THE IMPACT OF MATERNALLY TRANSMITTED MICROBES ON ANIMAL EVOLUTION

By

Lisa J. Funkhouser-Jones

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
in
Biological Sciences
May, 2016
Nashville, Tennessee

Approved:
Seth R. Bordenstein, Ph.D.
Katherine L. Friedman, Ph.D.
Julian F. Hillyer, Ph.D.
Antonis Rokas, Ph.D.
Timothy L. Cover, M.D.

To my families, old and new:

Mom, Dad, Erik, and Keith

&

Jeff, Watson, Cricket and Mendel

ACKNOWLEDGEMENTS

The question I am asked most often by recruits is why I decided to come to Vanderbilt. My answer is always the same: Vanderbilt truly cares about graduate education. I have to thank Dr. James Patton, Dr. Roger Chalkley and the entire BRET office for all the time and effort they put into running the IGP program, as well as Angela Titus and (formerly) Leslie Maxwell, who have had the unenviable task of keeping all of the graduate students in the Department of Biological Sciences on track. I would also like to thank the members of my thesis committee, past and present, for all of their support and helpful advice, both professionally and personally: Dr. Kathy Friedman (chair), Dr. Julián Hillyer, Dr. Antonis Rokas, Dr. Tim Cover, the late Dr. Dave McCauley and Dr. Laurie Lee. Most importantly, I must thank my thesis advisor, Dr. Seth Bordenstein, for his unflagging encouragement and unwavering support, and for always pushing me out of my comfort zone: I am a much better scientist today because of it.

I have been extremely lucky to have worked with an amazing group of people over the past several years. Thank you to the entire Bordenstein lab, especially to Sarah Bordenstein, who is the glue that holds the lab together, and to my fellow graduate students, who have become some of my best friends: Dr. Robert Brucker, Dr. Jason Metcalf, Daniel LePage, Teddy van Opstal, Andy Brooks, Jessie Perlmutter and Dylan Shropshire. A special thanks to Rob for all the pep talks and to Jason for always laughing at my cat videos. I would also like to acknowledge my rotation mentor, Dr. Bethany Kent, and the rest of the postdocs and lab assistants who have provided invaluable advice and helped the lab run smoothly: Dr. Kristin Jernigan, Dr. Kevin Kohl, Megan Chafee, Victor Schmidt, Joey Simmons, Andrew Williams, Shefali Setia, Rini Pauly and Didi Bojanova. Thank you also to my undergraduate mentees, Stephanie Sehnert, Caitlyn Le and Ananya Sharma, and to my collaborators from the Universidad de Madrid, José Bella, Paloma Martínez-Rodríguez, Raquel Toribio-Fernández, and Miguel Pita, for all of their hard work on the projects presented in this thesis. Finally, thank you to my IGP clique, Cassidy Cobbs, Dr. Sarah Lawson, Dr. Elise Pfaltzgraff, and Gage Matthews, for your friendship and for providing much needed stress relief.

Finally, I would like to thank my family for all their love and support. To my mom and dad, thank you for being the best parents and the best role models that I could ever ask for (and

for giving me your genes). To my brothers, your intelligence and creativity inspire me every day, and I am proud to call myself your big sister. To my in-laws, Caroline and Jerry, thank you for welcoming me into your family with open arms and open hearts and for always treating me like one of your own. Lastly, to my husband, Jeff: thank you for all of your encouragement, support, and patience, especially over the last few months. You are one in a million, and I could not have done this without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMP	Antimicrobial peptide
BLAST	Basic Local Alignment Search Tool
cDNA	complementary DNA
CI	Cytoplasmic incompatibility
ColA	Coleoptericin-A
Cpe	*Chorthippus parallelus erythropus*
Cpp	*Chorthippus parallelus parallelus*
DNA	Deoxyribonucleic acid
dsRNA	double-stranded RNA
FISH	Fluorescent in situ hybridization
FDR	False discovery rate
Gb	Gigabases
GFP	Green fluorescent protein
GSCN	Germ-line stem cell niche
HGT	Horizontal gene transfer
HMM	Hidden Markov Models
IIID	Insect Innate Immunity Database
kb	kilobases
Mb	Megabases
MYA	Million years ago
NCBI	National Center for Biotechnology Information
ns	Not significant
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with Triton X-100
PCR	Polymerase chain reaction
PGRP	Peptidoglycan recognition receptor protein
RNA	Ribonucleic acid
RNAi	RNA interference
RT-qPCR	Reverse transcription quantitative PCR
SSCN	Somatic stem cell niche
TEM	Transmission electron microscopy
QC	Quality control
qPCR	Quantitative PCR
QTL	Quantitative trait loci

CHAPTER I. INTRODUCTION

Symbiont-driven genome evolution

At no point in time have animals existed without the influence of microbes shaping their evolutionary trajectory. Despite the antiquity of host-microbe interactions, only recently have we begun to appreciate the importance of microbial participation in almost every aspect of animal physiology and behavior (Eisthen and Theis, 2016; McFall-Ngai et al., 2013). In fact, many argue that animals and plants should be viewed as "holobionts" consisting of the animal host plus its associated symbiotic microbes instead of individual entities (Bordenstein and Theis, 2015; Margulis, 1993; Rosenberg et al., 2007). For humans, this would include the estimated 39 trillion bacteria in the gut (Sender et al., 2016), plus any organisms living inside the mouth, on the skin, or in the sinuses. Since these microbial communities are acquired from the environment and shaped by personal experiences (like antibiotic treatment or a vegan diet) (Costello et al., 2012), every human has a unique microbiota with very few, if any, "core" microbial species shared among all humans (Human Microbiome Project, 2012; Huse et al., 2012; Li et al., 2013). However, despite the immense inter-individual variation in microbial inhabitants at the species level, the metabolic pathways present in the microbiome are strikingly similar across individuals (Human Microbiome Project, 2012), indicating that the functions our symbiotic partners perform, such as the degradation of complex sugars in the intestines (Qin et al., 2010), are more important than which specific microbial species are performing these functions.

In contrast, many animals, especially invertebrates, require specific microbial species for their survival. These symbionts often play a crucial role in the provisioning of essential amino acids or vitamins to hosts with highly specialized diets. For example, sap-sucking insects like aphids and whiteflies harbor obligate, mutualistic bacteria that produce essential amino acids lacking in plant phloem such as tryptophan, arginine, and threonine (Baumann, 2005; International Aphid Genomics, 2010), while blood-feeding insects like tsetse flies and bed bugs use nutritional symbionts to synthesize the B vitamins thiamine, biotin and riboflavin (Akman et al., 2002; Nikoh et al., 2014). To ensure that that these vital symbionts are maintained within the host population every generation, symbionts are vertically transmitted from mother to offspring instead of acquired from the surrounding environment. While many methods of maternal

microbial transmission have evolved in the animal kingdom (CHAPTER II), all result in a close association between host and symbiont over evolutionary time, allowing host-symbiont interactions to drive genetic change in the genomes of both partners. In this thesis, I will focus on how the presence of a maternally-inherited intracellular symbiont, *Wolbachia*, can influence host genome evolution both (1) directly through horizontal gene transfer (HGT) of symbiont DNA to the host genome and (2) indirectly through selection for host genetic variants that regulate *Wolbachia's* titer and transmission (Figure I-1).

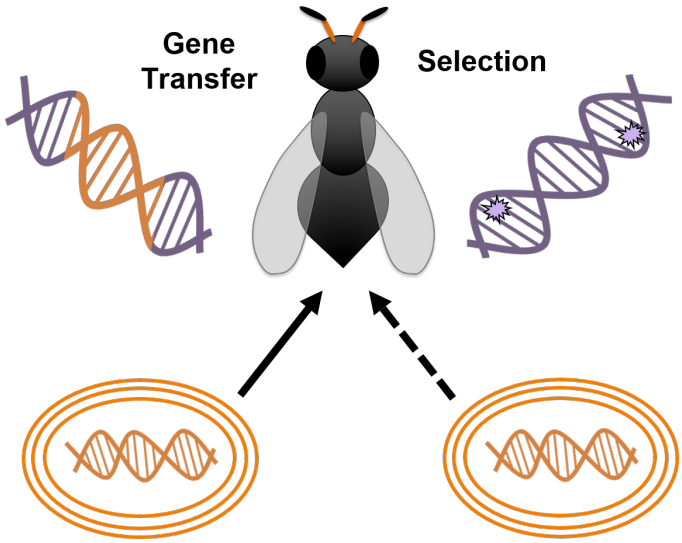

Figure I-1. Two mechanisms for symbiont-driven host genome evolution
Bacterial symbionts can affect host genome evolution directly through bacteria-to-eukaryote horizontal gene transfer or indirectly by placing pressure on the host genome to maintain genetic variants that help the host recognize and control its bacterial inhabitants.

Wolbachia as a model for maternally-transmitted endosymbionts

The obligate intracellular bacteria *Wolbachia* (Order: Rickettsiales) infects an estimated 40-52% of all arthropod species (Weinert et al., 2015; Zug and Hammerstein, 2012) and 47% of the Onchocercidae family of filarial nematodes (Ferri et al., 2011), making it arguably the most prevalent bacterial infection in the world and an excellent model for maternally-transmitted symbionts. The *Wolbachia* genus is divided into lettered supergroups instead of species, and individual *Wolbachia* strains are typically named based on their host species, such as *w*Mel from *Drosophila melanogaster*. Most *Wolbachia* found in arthropods belong to the A and B supergroups, while nematodes are predominantly infected with C and D *Wolbachia* (Casiraghi et

al., 2001). Interestingly, only one supergroup to date (the F supergroup) has been found in both arthropods and nematodes (Casiraghi et al., 2005).

In filarial nematodes, *Wolbachia* are obligate mutualists required for worm reproduction and survival (Fenn and Blaxter, 2004; Hoerauf et al., 1999). *Wolbachia* also serve as nutritional mutualists in blood-sucking bed bugs (Hosokawa et al., 2010; Moriyama et al., 2015; Nikoh et al., 2014) and are absolutely required for oogenesis in rice water weevils (Chen et al., 2012) and the wasp *Asobara tabida* (Dedeine et al., 2001; Pannebakker et al., 2007). However, in most arthropods, *Wolbachia* function as reproductive parasites that distort host sex ratio through a variety of mechanisms including male-killing, feminization (genotypic males function as fertile females) and parthenogenesis (virgin females produce only female offspring) (Serbus et al., 2008; Werren et al., 2008). The most common form of reproductive distortion, termed cytoplasmic incompatibility (CI), inhibits infected males from producing viable offspring with females that are uninfected or infected with a different *Wolbachia* strain. Cytologically, this phenomenon is caused by asynchronous development of the male and female pronuclei in embryos after fertilization with subsequent loss of paternal chromosomes and, in most cases, embryonic lethality (Tram and Sullivan, 2002). Despite extensive research, the underlying molecular modifications that *Wolbachia* use to induce CI and other reproductive phenotypes remain elusive.

Whether mutualistic or parasitic, all *Wolbachia* strains infect the ovaries of their hosts, where they generally target the somatic stem cell niche (SSCN) and, in some species of *Drosophila*, the germ-line stem cell niche (GSCN) (Fast et al., 2011; Frydman et al., 2006; Hosokawa et al., 2010; Toomey et al., 2013). Targeting the stem cell niches provides *Wolbachia* with at least three mechanisms for gaining access to a developing oocyte: (1) by directly infecting the germ-line stem cell, (2) by infecting the germ cell as it moves past the somatic stem cell niche, and (3) by moving into the germ cell after infecting the surrounding somatic follicle cells (Toomey et al., 2013). Once in the oocyte, many *Wolbachia* strains localize to the posterior pole where the reproductive organs will eventually develop (Chafee et al., 2011; Veneti et al., 2004), placing themselves in the perfect position to repeat the infection process.

Animal genome evolution through horizontal gene transfer

Intracellular bacteria like *Wolbachia* that infect the germ-line stem cells of their hosts are perfectly poised for bacteria-to-eukaryote horizontal gene transfer, especially if the bacteria harbor mobile genetic elements like bacteriophages. Historically, genomes of obligate, intracellular bacteria were assumed to lack bacteriophages and other mobile genetic elements because their restrictive intracellular environment would limit their exposure to foreign DNA and other bacteria (Bordenstein and Reznikoff, 2005). Of examples studied to date, this remains true for obligate, mutualistic bacteria that are strictly vertically-transmitted, presumably because lytic phage activity would endanger the symbionts that the host relies upon for survival (Tamas et al., 2002). Evidence for this theory can be seen in nematode *Wolbachia* genomes, which have remnants of phage genes but no intact phage (Foster et al., 2005; Kent and Bordenstein, 2010; Koutsovoulos et al., 2014). Even the *Wolbachia* strain in bed bugs, which has transitioned to mutualism more recently than nematode *Wolbachia*, lacks intact phage (Nikoh et al., 2014). On the other hand, host-switching, facultative intracellular symbionts like most arthropod *Wolbachia* often harbor bacteriophages. Approximately 89% of *Wolbachia* from the A and B supergroups have at least one prophage in their genomes from a temperate bacteriophage called WO (Bordenstein and Wernegreen, 2004; Gavotte et al., 2004). Some WO prophages retain the ability to become lytic and have been shown to transfer between different *Wolbachia* infections in the same host (Bordenstein and Wernegreen, 2004; Chafee et al., 2010; Gavotte et al., 2007; Kent et al., 2011; Masui et al., 2000). All of these observations led to the proposal of the Intracellular Arena Hypothesis, which posits that obligate, intracellular bacteria exchange bacteriophages and other mobile genetic elements when multiple endosymbionts co-infect the same host (Bordenstein and Wernegreen, 2004).

If we extend the concept of an intracellular arena for DNA exchange to include not just bacteria in the same host cytoplasm but all genomes within a single cell, then we would expect to see gene exchange between bacteria and the host nuclear genome. If a bacteria-to-eukaryote gene transfer occurred in a germ cell, then the symbiont DNA would be passed on by the host to its offspring. Since *Wolbachia* infect the germ cells of its hosts, almost every animal that is or once was infected with *Wolbachia* has *Wolbachia* DNA in its genome (Dunning Hotopp et al., 2007; Robinson et al., 2013). These inserts range in size from a couple hundred base pairs in *Nasonia* to the length of an entire *Wolbachia* genome in *Drosophila ananassae* (Dunning Hotopp et al.,

2007; Klasson et al., 2014). In CHAPTER III, I present the discovery of extensive *Wolbachia*-to-host HGT from two divergent *Wolbachia* supergroups to the meadow grasshopper *Chorthippus parallelus* and discuss how inherited bacteria like *Wolbachia* could directly impact the ever-expanding, gigantic genomes of Orthopterans in general.

Endosymbiont density regulation

While horizontal gene transfer is the most direct route for symbionts to affect host genome evolution, microorganisms can also act as a selective pressure to indirectly influence host genomic change. The most obvious example of microbes driving adaptive changes in animals is the evolution of immunity genes that recognize and control infectious agents. Indeed, studies in fruit flies and mosquitoes have found that recognition receptors and effector molecules like antimicrobial peptides are under positive selection and evolving more rapidly than other immune genes, such as those involved in signaling cascades (Sackton et al., 2007; Waterhouse et al., 2007). For maternally-inherited bacteria, host regulation of symbiont densities is critical. If symbiont titers become too low, not all host offspring will acquire the bacteria, and the bacteria will not fully express the adaptations that help them maintain their niche within the host (Anbutsu and Fukatsu, 2003; Jaenike, 2009; Kageyama et al., 2007). Conversely, if symbiont titers become excessive or aberrantly distributed in host tissues, the bacterial infection can turn virulent (Hughes et al., 2011; McMeniman et al., 2009; Min and Benzer, 1997). Thus, selection over time drives adaptations in the host genome to control these symbionts, and studies in tsetse flies (Rio et al., 2006), mosquitos (Berticat et al., 2002), fruit flies (Boyle et al., 1993; Dyer et al., 2005; Veneti et al., 2004), parasitoid wasps (Chafee et al., 2011; Mouton et al., 2003), adzuki bean beetles (Ijichi et al., 2002; Kondo et al., 2005), and weevils (Anselme et al., 2006) have all shown that host genotype influences the infection densities of endosymbionts.

For *Wolbachia* in particular, the efficient transmission of *Wolbachia* to the next generation as well as its ability to manipulate host reproduction, is dependent upon sufficient levels of *Wolbachia* within its host (Beeuwer and Werren, 1993; Bordenstein et al., 2006; Dyer et al., 2005; Jaenike, 2009; Unckless et al., 2009; Werren, 1999). Infection levels that are too high, however, can prove harmful to the host. In one extreme case, *Wolbachia* strain *w*MelPop over-proliferates in both reproductive and somatic tissues, including the brain, and cuts the lifespan of its *D. melongaster* host in half (Min and Benzer, 1997). Natural populations of insects

do not display these deleteriously high *Wolbachia* levels (Dobson et al., 1999; McGraw et al., 2001, 2002), presumably because co-evolution of insects and their resident *Wolbachia* strains has selected for regulatory mechanisms to control *Wolbachia* densities. For example, one study found that within pairs of insect species with the same *Wolbachia* strain, total *Wolbachia* densities were lower in the species that had harbored the *Wolbachia* strain for a longer period of time (Dobson et al., 1999). Furthermore, transinfection and introgression experiments have shown that the same *Wolbachia* strain will establish itself at different densities in different insect host species (Boyle et al., 1993; Chafee et al., 2011; Dobson et al., 1999; Kondo et al., 2005; McGraw et al., 2002; Veneti et al., 2004). One study even found that *Leptopilina heterotoma* parasitoid wasps naturally infected with three different strains of *Wolbachia* maintain each strain at a specific density that remains unchanged in wasps infected with only one or two of the strains (Mouton et al., 2003). Thus, host genetic factors appear pivotal in regulating the endosymbiont *Wolbachia*, although the extent to which *Wolbachia* influences its own densities, as well as the specific host-microbe interactions involved in density regulation, is not known.

Nasonia as a model organism for studying host regulation of endosymbiont titers

The *Nasonia* genus of parasitoid wasps serves as an excellent model system for studies of the genetic changes driving evolution in interspecific traits like wing size (Loehlin and Werren, 2012), cuticular hydrocarbon profiles (Niehuis et al., 2013), and memory retention (Hoedjes et al., 2014). *Nasonia* have many of the same advantages as the well-developed *Drosophila* model system including short-generation times, fully-sequenced genomes (Werren et al., 2010), detailed genetic maps (Desjardins et al., 2013b), and RNAi (Lynch and Desplan, 2006; Werren et al., 2009). However, *Nasonia* also have several unique advantages over other insect models, including a haplodiploid sex determination system, where fertilized eggs become diploid females and unfertilized eggs develop into haploid males. With only one copy of the genome, haploid males serve as powerful genetic tools for studying recessive genes. Furthermore, RNAi is systemic in *Nasonia*, which means that injection of dsRNA into the abdomen effectively knocks down gene expression throughout the body and in early embryos of injected parents (Lynch and Desplan, 2006; Werren et al., 2009)) Most importantly though for evolutionary genetic studies is the fact that the *Nasonia* genus is composed of four closely-related species that have all diverged within the last 1 MYA (Campbell et al., 1993; Raychoudhury et al., 2010). Once cured of their

incompatible *Wolbachia* infections, individuals from any *Nasonia* species can interbreed to produce fertile hybrid offspring, allowing genomic material to easily be moved from one species to another through a process termed introgression.

The two species used in this thesis, *N. vitripennis* and *N. giraulti*, last shared a common ancestor approximately 1 MYA (Campbell et al., 1993). Since diverging, each species has acquired different *Wolbachia* strains from both the A and B *Wolbachia* supergroups through independent horizontal transfer events (Bordenstein and Werren, 2007; Raychoudhury et al., 2009). Thus, neither species has been naturally exposed to the *Wolbachia* strains that the other species harbors. Within their resident host, *Wolbachia* densities are relatively low (less than 1 *Wolbachia* genome per host genome) regardless of the host or bacterial strain (Chafee et al., 2011) (Figure I-2). However, when *Wolbachia* strain *w*VitA from *N. vitripennis* is transferred by cytoplasmic introgression into an *N. giraulti* genetic background (IntG12.1), *w*VitA densities increase at least 80-fold (Chafee et al., 2011) (Figure I-2). When the same strain is transferred back to *N. vitripennis* from the high density *N. giraulti* IntG12.1 line, *w*VitA levels return to normal (Chafee et al., 2011) (Figure I-2). Since the *N. vitripennis* cytoplasm is introgressed along with the *Wolbachia*, only the nuclear genomes differ between the high density infection in *N. giraulti* and the naturally low density infection in *N. vitripennis*. Thus, there must be interspecific genetic differences between the two *Nasonia* species that influence titer levels of *w*VitA. Interestingly, when the experiment is repeated with the *N. vitripennis Wolbachia* strain *w*VitB, there is no change in *w*VitB titers in an *N. giraulti* background (Chafee et al., 2011) (Figure I-2). This suggests that host factors may interact specifically with a bacterial protein from one supergroup of *Wolbachia* but not another, which is plausible given that the A and B *Wolbachia* supergroups diverged approximately 60 million years ago (Werren et al., 1995).

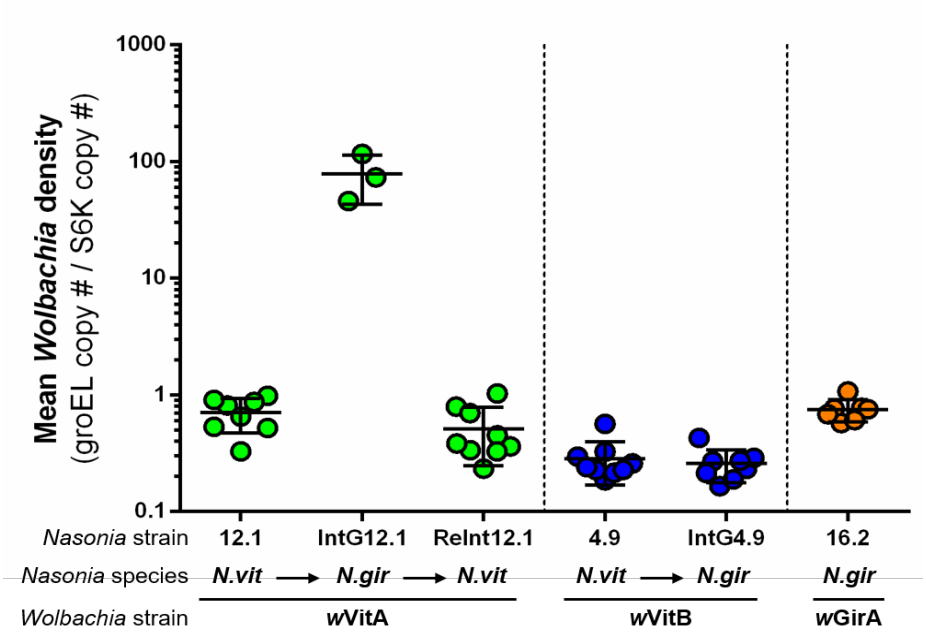

Figure I-2. Strain-specific proliferation of *Wolbachia* when transferred to a novel host
Wolbachia strains *w*VitA, *w*VitB and *w*GirA infect their native hosts at low titers. Transfer of *w*VitA from *N. vitripennis* (12.1) to *N. giraulti* increases mean *Wolbachia* density in adult females (IntG12.1) but returns to normal when transferred back into an *N. vitripennis* genomic background (ReInt12.1). *w*VitB titers remain the same when transferred from *N. vitripennis* to *N. giraulti*. . Original data was published in (Chafee et al., 2011).

In addition to higher titers, superinfections of *w*VitA in *N. giraulti* display an expanded tissue tropism relative to their normal localization in *N. vitripennis* (Chafee et al., 2011). In its resident host, *w*VitA localizes almost exclusively to the ovaries, but in the new host it infects nearly every somatic tissue (Chafee et al., 2011). This difference in *w*VitA distribution in adult *Nasonia* is likely established at the embryonic stage of *Nasonia* development since *w*VitA are concentrated exclusively at the posterior pole in *N. vitripennis* embryos but can be observed throughout *N. giraulti* embryos (Figure I-3). The expanded distribution of *w*VitA in the *N. giraulti* embryo places *Wolbachia* near cells that will develop into somatic tissues in the adult insect rather than restricting them to the reproductive tissues like in *N. vitripennis*. However, the majority of *w*VitA cells still localize to the posterior pole in *N. giraulti* and at higher titers than in *N. vitripennis* (Figure I-3). Thus, the 80- to 100-fold higher *w*VitA levels observed in adult *N. giraulti* may be due to higher initial titers in the embryo plus an expanded tissue tropism in the adult.

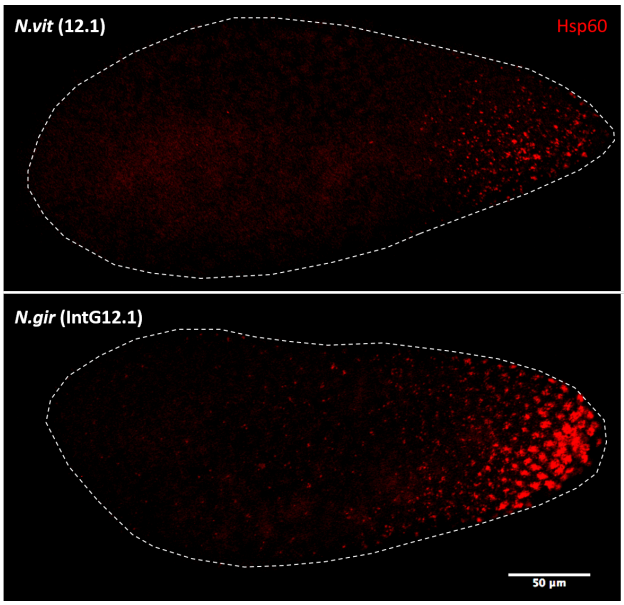

Figure I-3. Immunofluorescent staining of *w*VitA in *Nasonia* embryos
N. vitripennis (top) and *N. giraulti* IntG (bottom) embryos were stained for *Wolbachia* using an anti-Hsp60 antibody (red). Embryos are positioned from anterior (left) to posterior (right). Scale bar represents 50 µm.

In CHAPTER IV, I use the *w*VitA density difference between *N. vitripennis* and *N. giraulti* as a tractable phenotype to investigate the genetic basis of endosymbiont regulation in *Nasonia*. Reciprocal crosses of low-density *N. vitripennis* and high-density *N. giraulti* revealed that maternal *N. vitripennis* genes dominantly suppress *w*VitA titers in the subsequent generation. Using selective introgressions combined with genotyping microarrays and quantitative trait loci (QTL) analyses, I identify two regions on Chromosomes 2 and 3 that are associated with low *w*VitA densities, and confirm their role in *Wolbachia* density regulation with segmental introgression lines.

Conclusions and future directions

The research presented in this thesis makes significant progress in understanding host genome evolution in light of symbiosis, but many questions remain, including the specific genes in *Nasonia* that control *w*VitA titer and transmission. CHAPTER V will discuss future experiments that could help answer some of these questions.

CHAPTER II. MOM KNOWS BEST: THE UNVERSALITY OF MATERNAL MICROBIAL TRANSMISSION*

Summary

The sterile womb paradigm is an enduring premise in biology that human infants are born sterile. Recent studies suggest that infants incorporate an initial microbiome before birth and receive copious supplementation of maternal microbes through birth and breastfeeding. Moreover, evidence for microbial maternal transmission is increasingly widespread across animals. This collective knowledge compels a paradigm shift—one in which maternal transmission of microbes advances from a taxonomically specialized phenomenon to a universal one in animals. It also engenders fresh views on the assembly of the microbiome, its role in animal evolution, and applications to human health and disease.

Introduction

While the human microbiota comprises only 1–3% of an individual's total body mass, this small percentage represents over 100 trillion microbial cells, outnumbering human cells 10 to 1 and adding over 8 million genes to our set of 22,000 (Gill et al., 2006; Whitman et al., 1998). This complexity establishes a network of interactions between the host genome and microbiome spanning gut development (Murgas Torrazza and Neu, 2011), digestion (Ley et al., 2006; Turnbaugh et al., 2006), immune cell development (Ivanov et al., 2009; Ivanov et al., 2008; Round et al., 2011), dental health (Colombo et al., 2006; Ling et al., 2010), and resistance to pathogens (Candela et al., 2008; Fukuda et al., 2011). Recent studies have also provided a greater understanding of how the composition of an individual's microbiota changes throughout development, especially during the first year of life (Murgas Torrazza and Neu, 2011; Palmer et al., 2007). While the general dogma is that the placental barrier keeps infants sterile throughout pregnancy, increasing evidence suggests that an infant's initial inoculum can be provided by its

* This chapter is published in *PLOS Biology* (2013) 11(8):e1001631 with Seth R. Bordenstein as a co-author.

mother before birth (Bearfield et al., 2002; DiGiulio, 2012; DiGiulio et al., 2008; Jimenez et al., 2005; Jimenez et al., 2008) and is supplemented by maternal microbes through the birthing (Dominguez-Bello et al., 2010) and breastfeeding (Gronlund et al., 2007; Martin et al., 2012) processes.

While maternal transmission of microbes in humans has attracted considerable attention in the last few years, nearly a century's worth of research is available for vertical transmission of symbionts in invertebrates (Buchner, 1965). Similar to gut bacteria in humans that assist nutrient intake, many insect-associated bacteria function as nutritional symbionts that supplement the nutrient-poor diet of their host with essential vitamins or amino acids (Douglas, 1998; Feldhaar and Gross, 2009). Since these indispensable symbionts cannot live outside of host cells, they cannot be acquired from the environment and are faithfully transferred from mother to offspring (Buchner, 1965; Douglas, 1989). Maternal transmission in invertebrates has been reviewed elsewhere (Baumann, 2005; Bright and Bulgheresi, 2010; Buchner, 1965), but here we highlight several examples of heritable symbioses across invertebrate phyla.

Maternal Transmission in Insects

Insects that thrive on unbalanced diets such as plant sap, blood, or wood depend upon microbial symbionts for the provision of essential amino acids or vitamins lacking in their food source. In turn, hosts provide a wide range of metabolites to their symbionts as well as protection from environmental stressors. This codependence requires faithful transfer of symbionts to all offspring, usually through transovarial transmission (Douglas, 1998; Feldhaar and Gross, 2009). Reproductive parasites, such as the obligate, intracellular bacteria *Wolbachia*, are also widespread in insects and hijack maternal transmission routes to ensure their spread within an insect population (LePage and Bordenstein, 2013; Saridaki and Bourtzis, 2010).

*Pea Aphid (*Acrythosiphon pisum*)*

The pea aphid *Acrythosiphon pisum* (Figure II-2A) and its nutritional endosymbiont *Buchnera aphidicola* are a preeminent example of obligate mutualism in insects. The ancestral *Buchnera* gammaproteobacteria was acquired by aphids between 160 and 280 million years ago (Moran et al., 1993) and has since diverged in parallel with its aphid hosts through strict vertical transmission (Baumann, 2005; Moran et al., 1993). *Buchnera* are housed within the cytoplasm of

bacteriocytes arranged into dual bacteriome structures located in the aphid body cavity adjacent to the ovaries (Baumann et al., 1995), allowing efficient transfer of *Buchnera* symbionts to developing oocytes or embryos during the sexual and asexual phases of aphid reproduction, respectively. At the cellular level, symbiont transfer occurs when maternal bacteriocytes release *Buchnera* symbionts through exocytosis into the extracellular space between the bacteriocyte and oocyte or embryo, which then actively endocytoses the extracellular *Buchnera* symbionts (Koga et al., 2012).

Cockroaches (Order Blattodea)

Just as insects are morphologically diverse, the mechanisms by which insects transport symbionts to oocytes are highly varied. In cockroaches, *Blattabacterium*-filled bacteriocyte cells migrate from the abdominal fat body to the distantly located ovarioles where they adhere to the oocyte membrane (Sacchi et al., 1985; Sacchi et al., 1988). Interestingly, the bacteriocytes remain associated with the oocyte for eight to nine days before finally expelling their symbionts through exocytosis. The *Blattabacterium* cells then squeeze between the follicle cells surrounding the oocyte and are engulfed into the oocyte cytoplasm via endocytosis just prior to ovulation (Sacchi et al., 1988).

Whiteflies (Family Aleyrodidae)

The whitefly circumvents exocytosis of its intracellular nutritional symbiont, *Portiera aleyrodidarum*, by depositing entire bacteriocytes into its eggs. These maternal bacteriocytes remain intact yet separate from the developing embryo until the embryonic bacteriomes form, at which point the maternal bacteriocytes deteriorate (Buchner, 1965).

Tsetse Flies, Bat Flies, and Louse Flies (Superfamily Hippoboscoidea)

Members of the Hippoboscoidea superfamily (Order Diptera) are obligate blood feeders that have developed a unique reproductive strategy termed adenotrophic viviparity that offers a different solution to internal maternal transfer of symbionts. Females of this superfamily develop a single fertilized embryo at a time within their uterus (modified vaginal canal) until it is deposited as a mature third instar larva immediately preceding pupation. During their internal development, the larvae are nourished with milk produced by modified accessory glands that

empty into the uterus (Tobe, 1978). The milk primarily consists of protein and lipids (Cmelik et al., 1969), but it also serves as a reservoir for maternally transmitted microbial symbionts (Attardo et al., 2008). For example, the obligate mutualistic symbiont of tsetse flies, *Wigglesworthia glossinidia,* is absent from the female germ line and surrounding reproductive tissues but is found extracellularly in the female milk glands and is first detected in tsetse offspring once milk consumption begins during the first larval stage (Attardo et al., 2008).

Stinkbugs (Superfamily Pentatomoidea)

One of the most common mechanisms of external maternal transmission in insects is that of "egg smearing," which occurs when a female contaminates the surface of her eggs with symbiont-laden feces during oviposition. Upon hatching, offspring probe or consume the discarded egg shells to acquire the maternal bacteria. This mode of transmission is commonly found in plant-sucking stinkbugs, including the Pentatomidae and Acanthosomatidae families (Prado and Zucchi, 2012). In the Cynidae family of stinkbugs, along with the Coreidae family of leaf-footed bugs, gut symbionts are transferred maternally via coprophagy, in which offspring consume maternal feces, sometimes directly from the mother's anus (Buchner, 1965; Prado and Zucchi, 2012). Stinkbugs of the Plataspidae family, on the other hand, have developed a unique mode of transmission via a maternally provided "symbiont capsule" deposited on the underside of the egg mass (Fukatsu and Hosokawa, 2002). These capsules are comprised of bacterial cells dispersed throughout a resin-like matrix surrounded by a brown, cuticle-like envelope that protects the symbionts from environmental stressors such as UV irradiation or dissection (Hosokawa et al., 2005). After hatching, plataspid nymphs immediately probe the capsules to ingest the symbionts (Fukatsu and Hosokawa, 2002; Hosokawa et al., 2008).

*European Beewolf (*Philanthus triangulum*)*

While nutritional symbionts appear to be the most common type of bacteria transmitted via external maternal transmission in insects, the European beewolf (*Philanthus triangulum*) instead cultivates a symbiotic bacteria that protects offspring against microbial infection during development. Beewolves are solitary digger wasps that deposit their offspring in moist, underground nests, making them susceptible to fungal and bacterial infections (Strohm and Linsenmair, 1995). To combat these pathogens, female beewolves cultivate *Streptomyces*

philanthi bacteria in specialized glands in their antennae, which they copiously spread on the ceiling of the brood cell before oviposition (Goettler et al., 2007; Kaltenpoth et al., 2006; Kaltenpoth et al., 2005). After hatching, the larvae take up the bacterial cells and incorporate them into their cocoon that they build before pupation. When adult beewolves emerge from their cocoon in the summer, female beewolves acquire the maternally provided *Streptomyces* symbiont and house them in the female-specific gland reservoirs along each antenna (Goettler et al., 2007; Kaltenpoth et al., 2010).

Maternal Transmission in Marine Invertebrates

Marine Sponges (Phylum Porifera)

Sponges are ancient metazoans that evolved over 600 million years ago as one of the first multicellular animals (Li et al., 1998). In marine sponges, a remarkably large consortium of extracellular microbial symbionts thrives within the sponge's mesohyl, a gelatinous connective tissue located between the external and internal cell layers. Many of these bacterial residents are found in diverse species of sponges with nonoverlapping distributions but not in the surrounding seawater (Fieseler et al., 2004; Hentschel et al., 2002; Taylor et al., 2007). These "sponge-specific" microbes are hypothesized to have originated from ancient colonization events before the diversification of marine sponges and are maintained as symbionts through vertical transmission (Wilkinson, 1984). Independent studies have estimated that up to 33 phylogenetically distinct microbial clusters spanning ten bacterial phyla and one archaeal phylum are vertically transmitted in sponges (Hentschel et al., 2002; Schmitt et al., 2008; Taylor et al., 2007; Webster et al., 2010). Both transmission electron microscopy (TEM) and fluorescent in situ hybridization (FISH) studies have confirmed the presence of microorganisms of different shapes and sizes in the oocytes of oviparous sponges (Schmitt et al., 2008) and in the embryos of viviparous sponges (Ereskovsky et al., 2005; Schmitt et al., 2007; Sharp et al., 2007).

Vesicomyid Clams (Phylum Mollusca)

Deep-sea hydrothermal vent communities rely upon chemosynthetic bacteria to harness chemical energy stored in reduced sulfur compounds extruding from the vents. Metazoans that live in this extreme environment harbor chemosynthetic endosymbionts in their tissues that provide most, if not all, of the host's nutrition (Cavanaugh et al., 2006). Somewhat surprisingly,

most invertebrates that live near hydrothermal events acquire their endosymbionts anew from the environment each generation (Di Meo et al., 2000; Laue and Nelson, 1997), even though chemosynthetic bacteria are crucial for survival in such a harsh habitat. A major exception to this trend is found in the Vesicomyidae family of clams (Goffredi and Barry, 2002). Vesicomyid clams retain a rudimentary gut and rely primarily on sulfur-oxidizing bacteria sequestered intracellularly within specialized host cells called bacteriocytes in the clam's large, fleshy gills (Cavanaugh, 1983). Vertical transmission via transovarial transmission appears to be the dominant mechanism for maintenance of these thioautotrophic bacterial symbionts given that follicle cells surrounding an oocyte and the oocyte itself are heavily infected with the chemosynthetic bacteria (Cary and Giovannoni, 1993; Endow and Ohta, 1990).

By integrating previous studies in invertebrates with recent evidence for maternal microbial transmission in humans and other vertebrates, we contend that maternal provisioning of microbes is a universal phenomenon in the animal kingdom. As a result, a considerable new phase of study in heritable symbiont transmission is underway. Thus, this essay presents current evidence for maternal microbial transmission and provides new insights into its impact on microbiome assembly and evolution, with applications to human health and disease.

Internal Maternal Transmission

At the turn of the twentieth century, French pediatrician Henry Tissier asserted that human infants develop within a sterile environment and acquire their initial bacterial inoculum while traveling through the maternal birth canal (Tissier, 1900). More than a century later, the sterile womb hypothesis remains dogma, as any bacterial presence in the uterus is assumed to be dangerous for the infant. Indeed, studies of preterm deliveries have found a strong correlation between intrauterine infections and preterm labor, especially when birth occurs less than 30 weeks into the pregnancy (Goldenberg et al., 2000; Goncalves et al., 2002). Since preterm birth is the leading cause of infant mortality worldwide (Lawn et al., 2005), much attention has focused on identifying the bacterial culprits responsible for spontaneous preterm labor. Surprisingly, most of the bacteria detected in intrauterine infections are commonly found in the female vaginal tract (Goldenberg et al., 2000), and risk of preterm birth is markedly increased in women diagnosed with bacterial vaginosis during pregnancy (Fiscella, 1996). Interestingly, the

vaginal microbial community varies significantly among American women of different ethnicities (Caucasian, African-American, Asian, or Hispanic), with African-American and Hispanic women more likely to have a microbiota traditionally associated with bacterial vaginosis (predominance of anaerobic bacteria over *Lactobacillus* species) (Ravel et al., 2011) and a higher rate of spontaneous preterm deliveries (reviewed in (Menon et al., 2011)).

While intrauterine infection and inflammation is important in understanding the etiology of preterm birth, relatively few studies have examined the uterine microbiome of healthy, term pregnancies owing to the sterile womb paradigm. Investigations into the potential for bacterial transmission through the placental barrier have detected bacteria in umbilical cord blood (Jimenez et al., 2005), amniotic fluid (Bearfield et al., 2002; Rautava et al., 2012), and fetal membranes (Rautava et al., 2012; Steel et al., 2005) from babies without any indication of inflammation (Figure II-1). Furthermore, an infant's first postpartum bowel movement of ingested amniotic fluid (meconium) is not sterile as previously assumed, but instead harbors a complex community of microbes, albeit less diverse than that of adults (Gosalbes et al., 2013; Jimenez et al., 2008). Interestingly, many of the bacterial genera found in the meconium, including *Enterococcus* and *Escherichia*, are common inhabitants of the gastrointestinal tract (Gosalbes et al., 2013; Jimenez et al., 2008). To test whether maternal gut bacteria can be provisioned to fetuses *in utero*, Jiménez et al., 2008 fed pregnant mice milk inoculated with genetically-labeled *Enterococcus faecium* and then examined the meconium microbes of term offspring after sterile C-section. Remarkably, *E. faecium* with the genetic label was cultured from the meconium of pups from inoculated mothers, but not from pups of control mice fed noninoculated milk. Meconium from the treatment group also had a higher abundance of bacteria than that of the control group. Importantly, the study controlled for potential bacterial contamination from contact between skin and the meconium by sampling an internal portion of the meconium (Jimenez et al., 2008). Thus, this study provides foundational evidence for maternal microbial transmission in mammals.

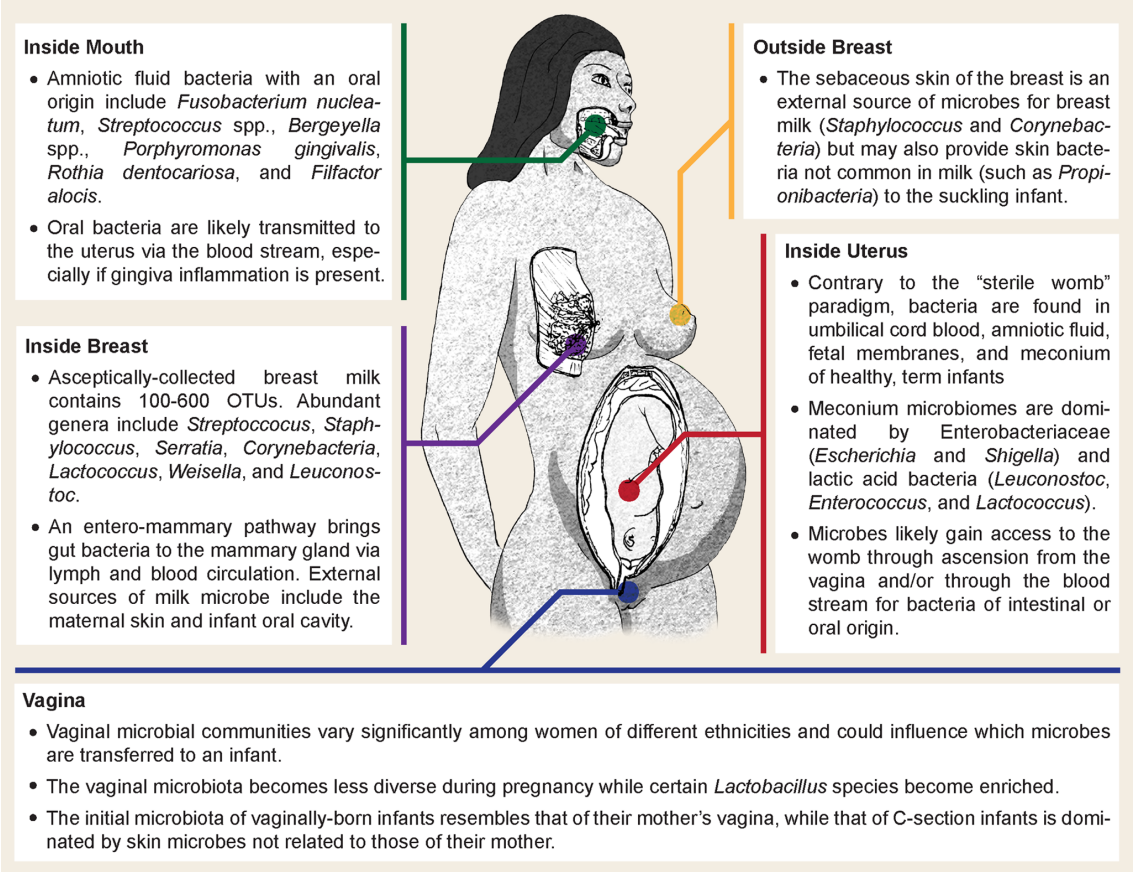

Figure II-1. Sources of microbial transmission in humans from mother to child.
Cut-away diagram highlighting the various internal and external sources of maternal microbial transmission as well as the species that are commonly associated with transfer from those regions.

Other than ascension of vaginal microbes associated with preterm births, the mechanisms by which gut bacteria gain access to the uterine environment are not well understood. One possibility is that bacteria travel to the placenta via the bloodstream after translocation of the gut epithelium. While the intestinal epithelial barrier generally prevents microbial entry into the circulatory system, dendritic cells can actively penetrate the gut epithelium, take up bacteria from the intestinal lumen, and transport the live bacteria throughout the body as they migrate to lymphoid organs (Rescigno et al., 2001; Vazquez-Torres et al., 1999). Interestingly, microbial translocation may even increase during pregnancy, as one study showed that pregnant mice were 60% more likely to harbor bacteria in their mesenteric lymph node (presumably brought there by dendritic cells) than nonpregnant mice (Perez et al., 2007). Bacterial species normally found in the human oral cavity have also been isolated from amniotic fluid and likely enter the

bloodstream during periodontal infections, facilitated by gingiva inflammation (Bearfield et al., 2002; DiGiulio, 2012) (Figure II-1).

Overall, the study of internal maternal transmission of microbes in mammals is in its infancy due to the enduring influence of the sterile womb paradigm and to the ethical and technical difficulties of collecting samples from healthy pregnancies before birth. Thus, we still know very little about the number and identity of innocuous microbes that traverse the placenta, whether they persist in the infant, or whether their presence has long-term health consequences for the child. Similarly, we know almost nothing about nonpathogenic viruses or archaea that may be transferred from mother to child alongside their bacterial counterparts. Fortunately, the advent of culture-independent, high-throughput sequencing will serve as a tremendous resource for this field and will hopefully lead to a characterization of the "fetal microbiome" *in utero*.

Maternal provisioning of microbes to developing offspring is widespread in animals, with evidence of internal microbial transmission in animal phyla as diverse as Porifera (Enticknap et al., 2006; Ereskovsky et al., 2005; Schmitt et al., 2008; Schmitt et al., 2007; Sharp et al., 2007), Mollusca (Cary and Giovannoni, 1993; Peek et al., 1998; Stewart and Cavanaugh, 2006; Stewart et al., 2008), Arthropoda (Balmand et al., 2013; Koga et al., 2012; Moran et al., 2008) (Figure II-2), and Chordata (Carlier et al., 2012; Dominguez-Bello et al., 2010; Inoue and Ushida, 2003) (Figure II-2). The presence of maternal transmission at the base of the Animalia kingdom and the surprising plasticity by which microbes gain access to germ cells or embryos in these systems signifies that maternal symbiont transmission is an ancient and evolutionarily advantageous mechanism inherent in animals, including humans. Therefore, we can no longer ignore the fact that exposure to microbes in the womb is likely and may even be a universal part of human pregnancy, serving as the first inoculation of beneficial microbes before birth.

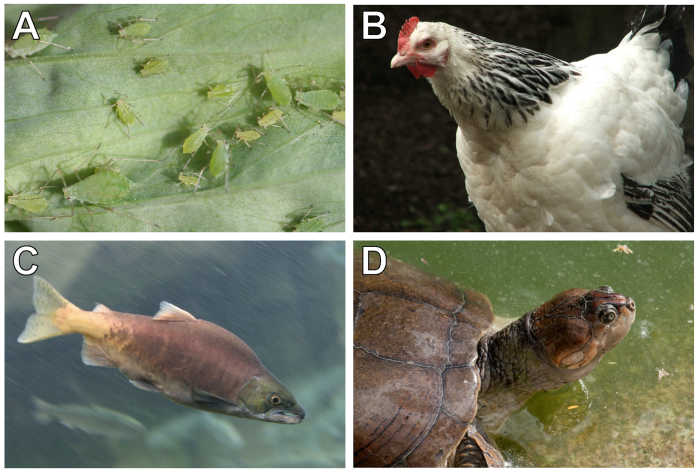

Figure II-2. Examples of animals that exhibit microbial maternal transmission.
(A) Pea aphid (Acyrthosiphon pisum), photo credit: Whitney Cranshaw, Colorado State University/©Bugwood.org/CC-BY-3.0-US; (B) Domesticated chicken hen (Gallus gallus domesticus), photo credit: Ben Scicluna; (C) Sockeye salmon (Oncorhynchus nerka), photo credit: Cacophony; (D) South American river turtle (Podocnemis expansa), photo credit: Wilfredor. All photos were obtained from Wikimedia Commons (www.commons.wikimedia.org).

Maternal Transmission in Vertebrates

Aside from studies in human and mouse models, very little is known about maternal transmission of microbial communities in vertebrates, especially outside Class Mammalia. Furthermore, research on vertical transmission in nonmammalians has largely focused on maternally transmitted pathogens, especially in animals of agricultural importance like chickens and fish.

Domesticated chickens (Gallus gallus domesticus)

Zoonotic *Salmonella* infections acquired from contaminated chicken eggs is estimated to cause more than 100,000 illnesses each year in the United States (Schroeder et al., 2005). In addition to horizontal transmission of *Salmonella* on eggs through surface contamination, direct transovarial transmission also occurs when *Salmonella* colonizes the reproductive tissues of hens (Figure II-2B). Depending on the infection location within the female reproductive tract, the bacteria are deposited into the yolk, albumen, eggshell membrane, and/or eggshell of the developing egg before oviposition (Gantois et al., 2009). Other poultry pathogens, such as *Mycoplasma synoviae* in chickens (MacOwan et al., 1984) and *M. gallisepticum*, *M. cloacale*,

and *M. anatis* in ducks (Bencina et al., 1988), have also been cultured from the yolk of embryonated eggs, though whether commensal flora are incorporated into the egg is not known.

Ray-finned fish (Class Actinopterygii)

Several bacterial pathogens of economically important fish are transmitted transovarially in the egg yolk including *Renibacterium salmoninarum*, the agent of bacterial kidney disease in salmonids (Figure II-2C), and *Flavobacterium psychrophilum*, which causes bacterial cold water disease in salmonids and rainbow trout fry disease in trout (reviewed in (Brock and Bullis, 2001)). *F. psychrophilum* has also been found in ovarian fluid and on the surface of eggs of steelhead trout (Brown et al., 1997). Additionally, an obligate, intracellular eukaryotic parasite, *Pseudoloma neurophilia*, is a common pathogen found in zebrafish (*Danio rerio*) facilities and has been observed in spores of the ovarian stroma and within developing follicle cells of spawning females, suggesting that it can be vertically transmitted, though it is primarily spread from fish to fish in contaminated water (Sanders et al., 2012).

Turtles (Order Cheloni)

The formation of egg components in the uterine tube and uterus of turtles takes approximately two weeks, providing ample opportunity for maternal transmission of intestinal or reproductive microbes to the egg (Alkindi et al., 2006). One study of unhatched (dead) eggs from loggerhead sea turtle (*Caretta caretta*) nests found several potential pathogens, including *Pseudomonas aeruginosa* and *Serratia marcesans*, in fluid from the interior of the eggs, though environmental contamination of the eggs cannot be ruled out (Craven et al., 2007). A similar study of eggs from two species of South American river turtles, *Podocnemis expansa* (Figure II-2D) and *P. unifilis*, identified several Enterobacteriaceae species, including *Escherichia coli*, *Shigella flexneri*, and *Salmonella cholerasuis*, in the eggs but not in the environmental samples taken from the turtle nests (Benevides de Morais et al., 2010), suggesting that they may have a maternal origin. In support of this hypothesis, a separate study in green turtles (*Chelonia mydas*) that collected eggs directly from the maternal cloacal opening during egg laying isolated *Pseudomonas*, *Salmonella*, *Enterobacter*, and *Citrobacter* from the eggshell, albumen, and yolk. In fact, the yolk was the egg component most heavily infected with bacteria (Al-Bahry et al., 2009). Altogether, many potentially pathogenic species have been isolated from turtle eggs, but

whether these bacteria actually cause disease in turtles or are part of their natural flora remains to be determined.

External Maternal Transmission

External maternal transmission encompasses any transfer of maternal symbionts to offspring during or after birth. In invertebrates, it is often accomplished by "egg smearing," in which females coat eggs with microbes as they are deposited (Kaltenpoth et al., 2009), or through the provision of a microbe-rich maternal fecal pellet that is consumed by larval offspring upon hatching (Fukatsu and Hosokawa, 2002; Hosokawa et al., 2007; Hosokawa et al., 2005; Hosokawa et al., 2008). Similarly, human infants are "smeared" with maternal vaginal and fecal microbes as they exit the birth canal (Bager et al., 2008; Huh et al., 2012; Thavagnanam et al., 2008) (Figure II-1). Several studies have shown that the human neonatal microbiota across all body habitats (skin, oral, nasopharyngeal, and gut) is influenced by their mode of delivery (Biasucci et al., 2008; Dominguez-Bello et al., 2010; Li et al., 2005; Penders et al., 2006), with infants born vaginally acquiring microbes common in the female vagina while C-section infants display a microbiota more similar to that of human skin (Dominguez-Bello et al., 2010). Furthermore, while the microbiota of a vaginally delivered infant clusters with the vaginal bacteria of its mother, the microbiota of C-section babies is no more related to the skin flora of its mother than that of a stranger, indicating that most microbes are transmitted to the neonate from those handling the infant (Dominguez-Bello et al., 2010). Importantly, epidemiological data suggest that a Cesarean delivery can have long-term consequences on the health of a child, especially concerning immune-mediated diseases. For example, children born via C-section are significantly more likely to develop allergic rhinitis (Renz-Polster et al., 2005), asthma (Renz-Polster et al., 2005), celiac disease (Decker et al., 2010), type 1 diabetes (Cardwell et al., 2008), and inflammatory bowel disease (Bager et al., 2012). These statistics are alarming given that 32.8% of all births in the United States in 2010 were delivered via C-section with similar rates on the rise in most developed countries (Gibbons et al., 2010).

The higher rate of immune-mediated diseases in C-section children may indicate that maternally transferred vaginal or fecal microbes are unique in their ability to elicit immune maturation in the neonate. Development of the intestinal mucosa and secondary lymphoid tissues in the gut is contingent upon recognition of microbial components by pattern-recognition

receptors on intestinal epithelial cells (Maynard et al., 2012; McElroy and Weitkamp, 2011). It is possible that these receptors cannot properly interact with the community of microbes acquired during Cesarean deliveries, leading to disrupted immune development and an increased risk for immune-mediated disorders in C-section children. Conversely, transmission for thousands of years of vaginal and fecal microbes at birth has likely produced specific human-microbe interactions important for neonatal gut development. In fact, a recent study found that the vaginal microbial community changes during pregnancy, becoming less diverse as the pregnancy progresses (Aagaard et al., 2012); yet, in spite of the general decrease in richness, certain *Lactobacillus* bacterial species are enriched in the vaginal community during pregnancy and are hypothesized to be important for establishing the neonatal upper GI microbiota after vaginal delivery (Aagaard et al., 2012).

Breastfeeding provides a secondary route of maternal microbial transmission as shown in humans (Fernandez et al., 2013) (Figure II-1) and nonhuman primates such as rhesus monkeys (Jin et al., 2011). In humans, maternal milk microbes are implicated in infant immune system development (Diaz-Ropero et al., 2007), resistance against infection (Maldonado et al., 2012), and protection against the development of allergies and asthma later in childhood (Fernandez et al., 2013). High-throughput sequencing of breast milk from 16 healthy women identified 100–600 species of bacteria in each sample with nine genera present in every sample: *Staphylococcus, Streptococcus, Serratia, Pseudomonas, Corynebacterium, Ralstonia, Propionibacterium, Sphingomonas,* and *Bradyrhizobiaceae* (Hunt et al., 2011). This "core" milk microbiome represented approximately 50% of all bacteria in each sample, with the other half representing individual variation in microbial composition (Hunt et al., 2011). A similar study found that the bacterial composition in breast milk changes over time: milk produced immediately after labor harbored more lactic acid bacteria along with *Staphylococcus, Streptococcus,* and *Lactococcus,* while breast milk after six months of lactation had a significant increase in typical inhabitants of the oral cavity, such as *Veillonella, Leptotrichia,* and *Prevotella* (Cabrera-Rubio et al., 2012), perhaps to prime the infant for the switch to solid food. However, as with any DNA-based, culture-independent study that does not discriminate between live and dead bacteria, the number and identity of bacteria detected in these studies should be interpreted with some caution.

Given that milk is only produced temporarily in a woman's life, the origin of milk microbes is still somewhat of a mystery. Breast milk was traditionally thought to be sterile;

however, colostrum (the first milk produced after delivery) collected aseptically already harbors hundreds of bacterial species (Cabrera-Rubio et al., 2012). Breast milk does share many taxa with the microbiota found on sebaceous skin tissue around the nipple (Grice et al., 2009; Hunt et al., 2011), and high levels of *Streptococcus* in breast milk may be a result of retrograde flow from an infant's oral cavity back to the milk ducts during suckling (Ramsay et al., 2004) since *Streptococcus* is the dominant phylotype in infant saliva (Cephas et al., 2011). However, the presence of anaerobic gut bacteria in human milk suggests that an entero-mammary route of transfer also exists that may utilize phagocytic dendritic cells to traffic gut microbes to the mammary glands, similar to microbial transfer to amniotic fluid as discussed earlier. To support this hypothesis, Perez *et al.* (Perez et al., 2007) found identical strains of bacteria in milk cells, blood cells, and fecal samples from lactating women, but more work is needed to directly connect bacterial translocation in the gut to incorporation in breast milk.

Overall, maternal transmission of beneficial microbes in humans has widespread relevance for human health. Evolution with these microbes has resulted in our dependence on them for the proper maturation and development of the immune system and gastrointestinal tract. Somewhat paradoxically, modern medicine designed to prevent infant mortality (such as emergency Cesarean sections and formula feeding) has likely contributed to the rise in immune-mediated diseases in developed countries due to the inherent lack of exposure to maternal microbes associated with these practices. Fortunately, biomedicine is also making strides in finding effective probiotic supplements to promote immune development and ameliorate some of the risks that C-section or formula-fed infants face as children and adults. Hopefully, as we gain understanding of the diversity and function of maternally transmitted microbes in humans, more complete and effective probiotic blends will recapitulate the microbial communities found in vaginally delivered, breast-fed infants and restore the microbe-host interactions that humans depend upon for proper development.

Conclusions

Since the early twentieth century, the study of maternal microbial transmission has focused heavily on animal systems in which maternal transmission maintains sophisticated partnerships with one or two microbial species. However, with the development of high-throughput sequencing technologies, it is now possible to identify entire microbiomes that are

transferred from mother to offspring in systems not traditionally considered to exhibit maternal transmission, such as humans. By expanding the definition of maternal transmission to include all internal and external microbial transfers from mother to offspring, we contend that maternal transmission is universal in the animal kingdom and is used to provision offspring with important microbes at birth, rather than leave their acquisition to chance.

Finally, with microbes contributing 99% of all unique genetic information present in the human body, maternal microbial transmission should be viewed as an additional and important mechanism of genetic and functional change in human evolution. Similar to deleterious mutations in our genetic code, disruption of maternal microbial acquisition during infancy could "mutate" the composition of the microbial community, leading to improper and detrimental host-microbe interactions during development. Maternal transmission is also a key factor in shaping the structure of the microbiome in animal species over evolutionary time, since microbes that promote host fitness, especially in females, will simultaneously increase their odds of being transferred to the next generation. Thus, whether internal or external, the universality and implications of maternal microbial transmission are nothing short of a paradigm shift for the basic and biomedical life sciences.

CHAPTER III. WOLBACHIA CO-INFECTION IN A HYBRID ZONE: DISCOVERY OF HORIZONTAL GENE TRANSFERS FROM TWO WOLBACHIA SUPERGROUPS INTO AN ANIMAL GENOME[†]

Abstract

Hybrid zones and the consequences of hybridization have contributed greatly to our understanding of evolutionary processes. Hybrid zones also provide valuable insight into the dynamics of symbiosis since each subspecies or species brings its unique microbial symbionts, including germline bacteria such as *Wolbachia*, to the hybrid zone. Here, we investigate a natural hybrid zone of two subspecies of the meadow grasshopper *Chorthippus parallelus* in the Pyrenees Mountains. We set out to test whether co-infections of B and F *Wolbachia* in hybrid grasshoppers enabled horizontal transfer of phage WO, similar to the numerous examples of phage WO transfer between A and B *Wolbachia* co-infections. While we found no evidence for transfer between the divergent co-infections, we discovered horizontal transfer of at least three phage WO haplotypes to the grasshopper genome. Subsequent genome sequencing of uninfected grasshoppers uncovered the first evidence for two discrete *Wolbachia* supergroups (B and F) contributing at least 448 kb and 144 kb of DNA, respectively, into the host nuclear genome. Fluorescent *in situ* hybridization verified the presence of *Wolbachia* DNA in *C. parallelus* chromosomes and revealed that some inserts are subspecies-specific while others are present in both subspecies. We discuss our findings in light of symbiont dynamics in an animal hybrid zone**.**

[†] This chapter was published in *PeerJ* (2015) 3:e1479 with Stephanie R. Sehnert, Paloma Martínez-Rodríguez, Raquel Toribio-Fernández, Miguel Pita, José L. Bella and Seth R. Bordenstein as co-authors. The research was performed in collaboration with José L. Bella's lab at the Universidad de Madrid.

Introduction

Microbial communities of many arthropod species are dominated numerically by heritable bacterial symbionts whose phenotypic effects range from mutualism to parasitism (Douglas, 2011). In some cases, millennia of co-evolution have produced obligate, mutualistic relationships in which microbial symbionts make essential amino acids and/or vitamins to complement the nutritionally incomplete diet of their hosts (Pais et al., 2008; Tamas et al., 2002; van Ham et al., 2003). In other cases, maternally-transmitted bacteria directly impact arthropod host reproduction by manipulating sex determination, fecundity, and the ratio of infected females (the transmitting-sex) within a population (LePage and Bordenstein, 2013). The alphaproteobacterium *Wolbachia* is the most widespread of these reproductive manipulators, infecting an estimated 40-52% of all terrestrial arthropod species (Weinert et al., 2015; Zug and Hammerstein, 2012). It uses a variety of mechanisms to increase the number of host females in a population including feminization of genetic males, male-killing, parthenogenesis, and cytoplasmic incompatibility (CI), which typically results in embryonic death of offspring produced by an uninfected female mated with an infected male (Serbus et al., 2008).

Hybrid zones are excellent model systems for studying the impact of interactions between heritable endosymbionts on animal evolution. For example, *Drosophila recens* and *D. subquinaria* meet in secondary contact in a hybrid zone spanning central Canada where *D. recens* is infected by a *Wolbachia* strain that causes strong CI (~90% reduction in progeny) when males mate with naturally uninfected *D. subquinaria* females (Jaenike et al., 2006; Shoemaker et al., 1999). In contrast, weak levels of CI in a hybrid zone could promote *Wolbachia* exchange between animal species. Two closely related species of field crickets, *Gryllus firmus* and *G. pennsylvanicus*, hybridize in a north-south zone along the eastern front of the Appalachian Mountains in the United States (Harrison and Arnold, 1982). Though each cricket species is predominantly infected with different *Wolbachia* strains, *Wolbachia* is not a primary source of hybrid incompatibility in this system (Mandel et al., 2001). This may partly explain why a significant portion of *G. pennsylvanicus* are infected with both *Wolbachia* strains (Mandel et al., 2001). *Wolbachia* co-infection of the same host can readily facilitate gene exchange and transfer of mobile elements between intracellular bacteria according to the intracellular arena concept (Bordenstein and Reznikoff, 2005; Bordenstein and Wernegreen, 2004; Newton and Bordenstein, 2011). Indeed, we previously showed that the co-infecting *Wolbachia* strains in *G.*

pennsylvanicus crickets harbor a nearly identical infection of *Wolbachia's* temperate bacteriophage WO (Chafee et al., 2010). Thus, hybrid zones that permit mixing of *Wolbachia* symbionts may in turn enable horizontal gene transfer between the coinfections.

Here, we investigate horizontal gene transfer of bacteriophage WO in a natural hybrid zone of the meadow grasshopper *Chorthippus parallelus*. During the last Ice Age, *C. parallelus* populations on the Iberian Peninsula were geographically isolated from those in continental Europe, resulting in the divergence of Iberian *C. parallelus erythropus* (Cpe) subspecies from the contemporary continental subspecies, *C. parallelus parallelus* (Cpp) (Shuker et al., 2005a). Now in secondary contact, hybrids of the two subspecies have interbred for an estimated 9,000 generations along a hybrid zone in the Pyrenees Mountains between France and Spain (Hewitt, 1993; Shuker et al., 2005a) (Figure III-1). Due to low dispersal rates, all grasshoppers collected from populations in the hybrid zone (i.e., Portalet) are hybrids of the two subspecies, while pure Cpp and Cpe populations reside on the edges of the hybrid zone (Gabas for Cpp and Escarrilla for Cpe) (Bella et al., 2007; Hewitt, 1993; Shuker et al., 2005a) (Figure III-1). F1 hybrids produced in laboratory crosses between the subspecies follow Haldane's rule and produce sterile F1 hybrid males, but both hybrid males and females in the field are fertile, possibly due to selection against deleterious allelic combinations that result in hybrid sterility (Bella et al., 1990; Shuker et al., 2005b).

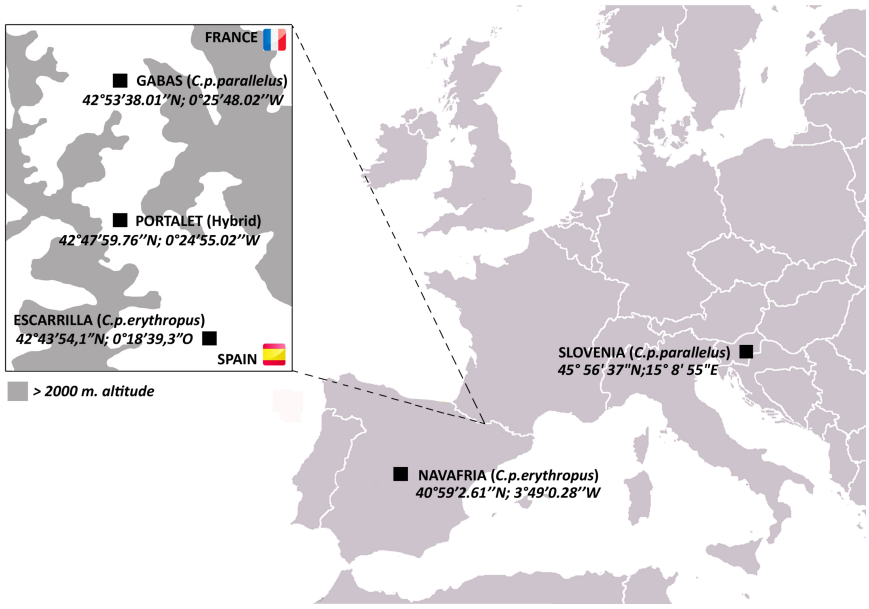

Figure III-1. Map of *C. parallelus* collection sites with their geographical coordinates.
Boxed inset shows the hybrid zone of *C. p. parallelus* and *C. p. erythropus* subspecies in the D'Ossau and Tena valleys of the Pyrenees Mountains between France and Spain.

C. parallelus subspecies are infected with *Wolbachia* strains from two divergent supergroups: Cpp are primarily infected with B *Wolbachia* while Cpe mostly harbor F *Wolbachia* (Zabal-Aguirre et al., 2010). In natural hybrid populations, the B and F *Wolbachia* each cause a significant amount of unidirectional CI, reducing embryo viability by approximately 33% and 23%, respectively, in incompatible crosses (Zabal-Aguirre et al., 2014). Bidirectional CI is weaker, with a 15% reduction in viable embryos in crosses between F-infected and B-infected grasshoppers (Zabal-Aguirre et al., 2014). With these incomplete CI rates permitting the mixture of *Wolbachia* strains, the incidence of *Wolbachia* infection is highly variable in the hybrid zone, and individuals collected from a single population are either uninfected, singly-infected with B or F *Wolbachia*, or co-infected by both (Zabal-Aguirre et al., 2010).

As the temperate bacteriophage WO is well known to transfer between A and B supergroup co-infections in arthropods (Bordenstein and Wernegreen, 2004; Chafee et al., 2010; Gavotte et al., 2007; Kent et al., 2011; Masui et al., 2000; Metcalf and Bordenstein, 2012), we used the *C. parallelus* hybrid zone to investigate whether phage WO can also transfer between co-infections of B and F *Wolbachia*. Here, we present the first screen for phage WO in the *C. parallelus* hybrid zone. While we do not find evidence for WO transfer between B and F *Wolbachia*, we identify three main WO haplotypes in the grasshopper genome. We also report, for the first time to our knowledge, the transfer of large amounts of DNA from two divergent *Wolbachia* supergroups into the host nuclear genome.

Materials and Methods

Sample collection, DNA extraction, and Wolbachia strain typing

The Spanish Comunidad de Madrid, the Gobierno de Aragón and the French Parc National des Pyrénées gave permission (permit numbers 10/103410.9/15; INAGA 500201/24/2012/12140; and Autorisation 2015-9, respectively) to collect *Chorthippus parallelus* individuals from five European and Iberian populations (Figure III-1). Gonads (or the whole body) were dissected and fixed in 100% ethanol. DNA was extracted as described elsewhere (Martinez-Rodriguez et al., 2013). *Wolbachia* was detected by PCR amplification of the *Wolbachia* 16S rRNA gene using *Wolbachia*-specific primers (Zabal-Aguirre et al., 2010), followed by nested PCR amplifications using B and F supergroup-specific primers (Martinez-

Rodriguez et al., 2013). 10 µl of each amplification product were electrophoretically separated on 1% agarose gels, which were stained with 0.5 mg/ml ethidium bromide and visualized under UV light (UVIdoc, Uvitec Cambridge).

Phage PCR amplification, cloning and sequencing

All PCR amplifications for phage and *Wolbachia* gene analyses were performed using 7.5 µl 2X GoTaq Green Master Mix (Promega), 3.6 µl sterile water, 1.2 µl of each primer (5 µM) and 1.5 µl template DNA for a 15 µl total reaction volume (scaled up as necessary) on a Veriti Thermal Cycler (Applied Biosystems) with the following primers: phgWOF (5'-CCCACATGAGCCAATGACGTCTG-3') and phgWOR (5'-CGTTCGCTCTGCAAGTAACTCCATTAAAAC-3') for the WO minor capsid gene (Masui et al., 2001); WolbF (5'-GAAGATAATGACGGTACTCAC-3') and WolbR3 (5'-GTCACTGATCCCACTTTAAATAAC-3') for the *Wolbachia* 16S ribosomal RNA gene (Casiraghi et al., 2001); ftsZunif (5'-GGYAARGGTGCRGCAGAAGA-3') and ftsZunir (5'-ATCRATRCCAGTTGCAAG-3') for *Wolbachia* ftsZ (Lo et al., 2002). The following primers were designed as part of this study to amplify specific WO alleles: forward primer WOPar1_F1 (5'-AATCTAAAAAGCGAAGTGAATCGTT-3') paired with phgWOR to amplify Cpar-WO1 alleles; reverse primer WOPar3_R1 (5'-CGACAGTTCTCGTAGCCTTCCTCA-3') paired with phgWOF to amplify Cpar-WO3 alleles.

To clone and sequence the *orf7* gene, PCR products were run on a 1% TBE agarose gel, then excised and purified using the Wizard PCR and Gel Clean-up Kit (Promega). 4 µl of each purified PCR product was cloned into a pCR4-TOPO vector using the TOPO TA Cloning kit (Invitrogen). OneShot TOP10 *E. coli* cells (Life Technologies) were transformed with the recombinant plasmids through heat shock according to the manufacturer's protocol. Transformed *E. coli* were plated on LB + carbenicillin plates and incubated overnight at 37 ˚C. Fifteen to 26 colonies were picked per plate then sent to GENEWIZ, Inc. (South Plainfield, NJ) for plasmid purification and Sanger sequencing. Both forward and reverse directions were sequenced for each plasmid then assembled in Geneious v5.5.8. For Sanger sequencing with allele-specific primers, PCR products were excised and purified from agarose gels as described above then sent to GENEWIZ, Inc. for sequencing. Both forward and reverse directions were sequenced for each PCR product then assembled in Geneious v5.5.8.

Phylogenetic tree construction

All multiple sequence alignments and phylogenetic trees were constructed in Geneious v5.5.8. Minor capsid sequences obtained through cloning and/or Sanger sequencing were aligned with homologous sequences from other WO phages (Table A-1) using the Translation Align Tool with default parameters, and the *dnaA* and *fabG* contigs from high-throughput sequencing were aligned with their homologs in *Wolbachia* strains using the Geneious alignment tool with default parameters. *Wolbachia dnaA* and *fabG* genes were extracted from full genome sequences from NCBI (Genbank) as follows: *w*Ha [CP003884.1], *w*Mel [AE017196.1], *w*Ri [CP001391.1], *w*No [CP003883.1], *w*Pip strain Pel [AM999887.1], *w*Oo [HE660029.1], *w*Ov strain Cameroon [HG810405.1], *w*Bm strain TRS [AE017321.1], and *w*Cle [AP013028.1].

After indels were manually removed, the minor capsid gene alignment was 332 bp with 49 sequences, the *dnaA* alignment was 742 bp with 11 sequences, and the *fabG* alignment was 735 bp with 11 sequences. "N"s were added to the 5' or 3' ends of any sequences that were shorter than the total alignment length. jModelTest 0.1.1 was used to determine the best model of nucleotide evolution for each alignment based on the corrected Akaike information criterion (AICc). For each gene, PhyML (Guindon and Gascuel, 2003) and MrBayes (Huelsenbeck and Ronquist, 2001) were executed in Geneious with default parameters to construct a maximum likelihood tree with bootstrapping and a Bayesian tree with a burn-in of 100,000, respectively. For the minor capsid gene, the third best model of nucleotide evolution (HKY + G) was used to generate both the maximum likelihood and Bayesian trees since the first two best models were not available in PhyML or MrBayes. The Hasegawa-Kishino-Yano (HKY) model of nucleotide evolution allows variable base frequencies and separate rates for transitions and transversions (Hasegawa et al., 1985). For the *dnaA* gene, the 10th best model of HKY + G was used since the first 9 were not available in PhyML or MrBayes. For the *fabG* gene, the second best model of GTR + G was used. The general time reversible (GTR) model of nucleotide evolution allows variable base frequencies and assumes a symmetric substitution matrix (Lanave et al., 1984; Tavare, 1986). For both the HKY and GTR models, rate variation among sites was modeled as a gamma distribution (+G).

High throughput sequencing of Wolbachia genomic inserts

Pooled DNA from three uninfected grasshoppers (two gonadal and one whole-body extractions) from the Gabas population (pure Cpp) was sequenced as 100 bp, paired-end reads on a single lane of an Illumina HiSeq2000 at the Vanderbilt VANTAGE sequencing facility. All analysis of sequencing data was performed in CLC Genomics Workbench 8. Reads were trimmed based on a quality limit of 0.05 and minimum length of 50 bp. After trimming, the data consisted of 227,349,258 reads with an average length of 93.5 bp totaling 21,347,095,705 bp.

All reads were initially mapped to the B *Wolbachia* genome of *w*Pip strain Pel (Genbank AM999887) using the CLC mapping tool with the following parameters: 80% similarity over 80% read length, mismatch cost = 2, insertion cost = 3, deletion cost = 3, and random mapping of non-specific reads. To ensure that the mapped reads were indeed from *Wolbachia,* reads from core *Wolbachia* genes were searched against the NCBI nucleotide database using blastn (megablast). Since many of the reads were more similar to genomic sequences from the F *Wolbachia* *w*Cle than to *w*Pip or other B *Wolbachia* genomes, we re-mapped all reads to the *w*Cle (Genbank AP013028) and *w*Pip (Genbank AM999887) reference genomes simultaneously with more stringent parameters: 90% similarity over 90% read length, mismatch cost = 2, insertion cost = 3, deletion cost = 3, and random mapping of non-specific reads. Since read mapping to each genome was mutually exclusive, this generated a list of reads that preferentially mapped to one genome over the other. To ensure that this was the case, reads that mapped to *w*Pip were extracted and mapped to the *w*Cle genome and vice versa with the more stringent parameters (90% similarity over 90% read length) to generate a combined list of "non-specific reads". After excluding these non-specific reads, the remaining reads were mapped back to the genome that they preferentially mapped to in order to determine the final lengths of the B and F inserts.

To find genes shared between the inserts, we took the reads that preferentially mapped to either *w*Pip or *w*Cle (B and F reads, respectively) and mapped them with less stringent parameters (70% sequence similarity over 90% sequence length) to the reciprocal genome. Genes were considered shared between the two inserts if both B and F reads mapped to homologous genes on both the *w*Pip and *w*Cle genomes and total read length for both B and F reads on each gene exceeded 80 bp. B and F variants for each gene were manually verified by

using blastn (discontiguous megablast) to confirm that percent similarity of B variants to *w*Pip were higher than to *w*Cle and vice versa.

To determine whether reads preferentially mapped to *w*Pip and *w*Cle over *Wolbachia* strains from other supergroups, we mapped all reads simultaneously to *w*Pip, *w*Cle, *w*Mel, *w*Bm, and *w*Oo reference genomes with a cutoff of 90% sequence similarity over 90% read length or 65% similarity over 80% read length. All reads that ambiguously mapped to more than one location were discarded.

Visualization of read mapping coverage on the *w*Pip and *w*Cle circular genomes was generated using the BLAST Ring Image Generator v0.95 (Alikhan et al., 2011) with a maximum mapping coverage of 30.

FISH analysis

To perform the cytogenetic analyses, male adult specimens of Cpp and Cpe were collected from the Gabas (France) and Escarrilla (Spain) populations, respectively. Grasshopper gonads were extracted and fixed in fresh ethanol:acetic acid (3:1) and used to prepare slides. After identifying uninfected individuals with *Wolbachia*-specific primers, as mentioned above, we designed primers to amplify a *Wolbachia* contig (Cpar-Wb1) identified during genome sequencing: 177contigF (5'-ACAGGAATTACAGCCTCAGGT-3') and 177contigR (5'-AAAAGCGTGGCAACAAAGTT-3'). PCR amplifications used the following conditions: Buffer 1X, $MgCl_2$ 2 mM, dNTPs (Roche) 0.2 mM, 1.2 µM of each primer, BIOTAQ DNA polymerase 1.25 U (Biotools), and 100 ng of genomic DNA, adjusting the final volume to 25 µl. The PCR program started with a cycle of 3 min at 95 °C, followed by 35 cycles of denaturing (30 s at 95 °C), annealing (45 s at 56 °C), extension (3 min at 72 °C), and a final extension of 10 min at 72 °C. PCR products were run on a 0.7% TAE agarose gel and were purified using the Illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare).

The purified DNA from the PCR was used to generate FISH probes with the DecaLabel DNA Labeling kit (Thermo Scientific), which is based on the random-primed method (Feinberg and Vogelstein, 1983, 1984), including a digoxigenin-labeled nucleotide. The complete reaction consisted of: 10 µl of decanucleotide, 5X Reaction Buffer, 1 µg of cDNA, and nuclease-free H_2O till 42 µl, keeping this mix at 100°C for 10 min; afterwards, we added 1 mM dNTPs mix, 1.75 µl of Digoxigenin-11-dUTP (Roche), and 1 µl of Klenow enzyme then incubated at 30°C for 2

hours. Finally, the probes were purified again with the Illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare), and eluted in 50 µl of H_2O.

Chromosome slides were prepared from fixed gonads to observe hybridization to male meiotic chromosomes from Cpe and Cpp individuals. Gonads were adhered to slides by the conventional technique of squashing, and the coverslip was removed after immersing the slides in liquid nitrogen. The squashed biological material was then treated for 5 min with pepsin (50 µg/ml in 0.01 N HCl) at 37°C, followed by a 30 min incubation in 2% paraformaldehyde at room temperature. Endogenous peroxidases were inactivated by incubation for 30 min with 1% H_2O_2. Slides were then dehydrated in a series of ethanol washes (70%, 85%, and 100%) and dried out. Slides were denatured and hybridized in the presence of 50 µl of the hybridization mixture under a coverslip for 5 min at 70°C. Hybridization mixture was composed of 2 µl of labeled probe, 50% formamide, 2X SSC, 300 mM NaCl, 30 mM sodium citrate, pH 7.0. After denaturing, slides were left overnight in a wet chamber at 37°C. Posthybridization washing and visualization of FISH-TSA (tyramide signal amplification) probes were performed as described previously (Krylov et al., 2007; Krylov et al., 2008). Detection of probes with antidigoxigenin conjugated to horseradish peroxidase (Roche) was done at a concentration of 1:2000 in TNB (Tris-NaCl-blocking buffer). The tyramide solution (Perkin Elmer) was incubated onto the slides for 5 min at a concentration of 1:50. Chromosomes were counterstained with 50 ng/µl of DAPI (4',6-diamidino-2-phenylindole, Roche) diluted in Vectashield (Vector Laboratories). Results were observed in a digital image analysis platform based on Leica DMLB fluorescence microscope with independent green and blue filters. Images were captured as tiff files using a cooled CCD Leica DF35 monochrome camera (Leica Microsystem), and final images were processed employing Photoshop CS6 (Adobe).

Data Availability

All cloning and sequencing data were deposited in the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/) under accession numbers KR081342 – KR081347 and KT599860 – KT599861. High-throughput genomic raw sequence reads are available from the Sequence Read Archives (http://www.ncbi.nlm.nih.gov/sra) under BioSample accession number SAMN03469681

Results

Infected and uninfected grasshoppers across the hybrid zone harbor phage WO genes

To initially determine the prevalence of phage WO in the *C. parallelus* hybrid zone, we PCR-screened hybrid, Cpe, and Cpp grasshoppers of all infection types (co-infected, B-infected, F-infected and uninfected) for the minor capsid gene (*orf7*), a virion structural gene commonly used to identify WO haplotypes (Bordenstein and Wernegreen, 2004; Chafee et al., 2010; Gavotte et al., 2004; Masui et al., 2000). Surprisingly, *orf7* amplicons were detected in 42 out of 43 (98%) samples, including all uninfected grasshoppers (n = 8), which were determined to be *Wolbachia*-free using nested PCR for the *Wolbachia* 16S ribosomal RNA gene (Figure III-2). Blank controls were negative for the *orf7* amplicon. These results indicate that (i) phage WO is or once was ubiquitous in *C. parallelus* and (ii) at least part of phage WO has laterally transferred to the grasshopper genome.

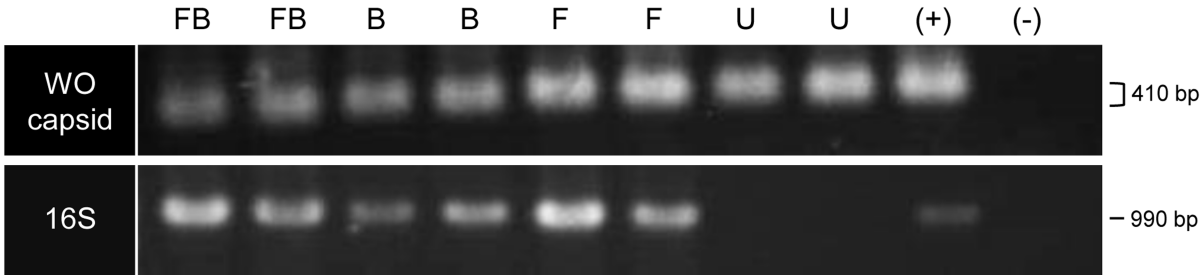

Figure III-2. PCR amplification of the WO minor capsid (*orf7*) gene and *Wolbachia* 16S ribosomal RNA gene

Two individuals of each infection type are shown: FB = co-infected, B = B infection only, F = F infection only, U = uninfected, (+) = positive DNA control, (-) = no template negative control. For the WO capsid gene, the gel ran askew, making some bands appear larger in size than others though all bands represent the same sized PCR amplicon (410 bp).

Diverse WO haplotypes are present in the grasshopper genome

To identify phage WO variation in a hybrid zone population, we cloned and sequenced an approximately 350 bp region of *orf7* from a co-infected (604FB), B-infected (603B), F-infected (607F) and uninfected (641U) hybrid grasshopper from the Portalet population. To confirm that these alleles were present in other individuals within the same population, we used allele-specific primers to amplify and sequence *orf7* from five additional individuals: three uninfected (167U, 169U and 186U), one F-infected (180F) and one co-infected (192FB). In total, we identified

eight unique *orf7* alleles spread throughout the phylogenetic tree of select WO minor capsid sequences (Figure III-3, Table A-1).

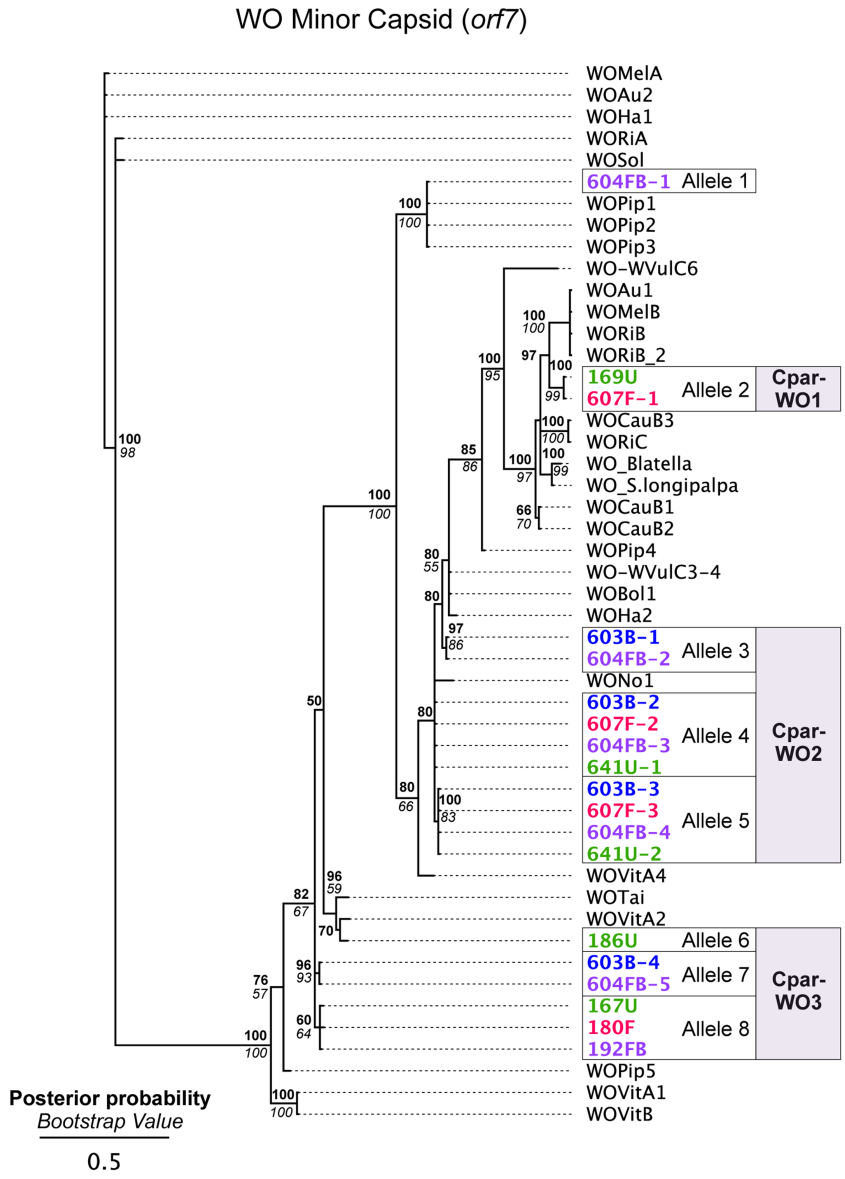

Figure III-3. Phylogeny of the WO minor capsid gene (*orf7*)

Bayesian phylogeny constructed using indel-free nucleotide alignment of the phage WO *orf7* gene. Sequences generated in this study are labeled with individual identification numbers and color-coded based on the grasshopper's infection status: FB = co-infected (purple), B = B-infection only (blue), F = F-infection only (red) and U = uninfected (green). Numbers after a hyphen designate different *orf7* sequences from the same individual. Posterior probability (Bayesian) and bootstrap (maximum likelihood) values over 50 are indicated in bold and italics, respectively. Accession numbers for sequences used in the tree, including the sequences from this study, are listed in Table A-1. The tree is arbitrarily rooted.

Seven of these alleles clustered into three haplotypes (Cpar-WO1, Cpar-WO2, and Cpar-WO3) based on a 96% identity cutoff (Figure III-3 and Figure III-4). Since all three haplotypes contain sequences obtained from uninfected individuals, we conclude that at least three phage WO insertions are present in the grasshopper nuclear genome. Two alleles without an identical sequence from an uninfected individual (alleles 3 and 7) may actually be present in a cytoplasmic *Wolbachia* strain rather than the host genome, but we have conservatively clustered them within the Cpar-WO2 and Cpar-WO3 haplotypes, respectively, since they are each 97.7% identical to an allele from an uninfected individual (alleles 4 and 8, respectively). An additional *orf7* allele (allele 1) was only found in a single co-infected individual, so we cannot conclude whether it was sequenced from a cytoplasmic *Wolbachia* infection or a nuclear insert.

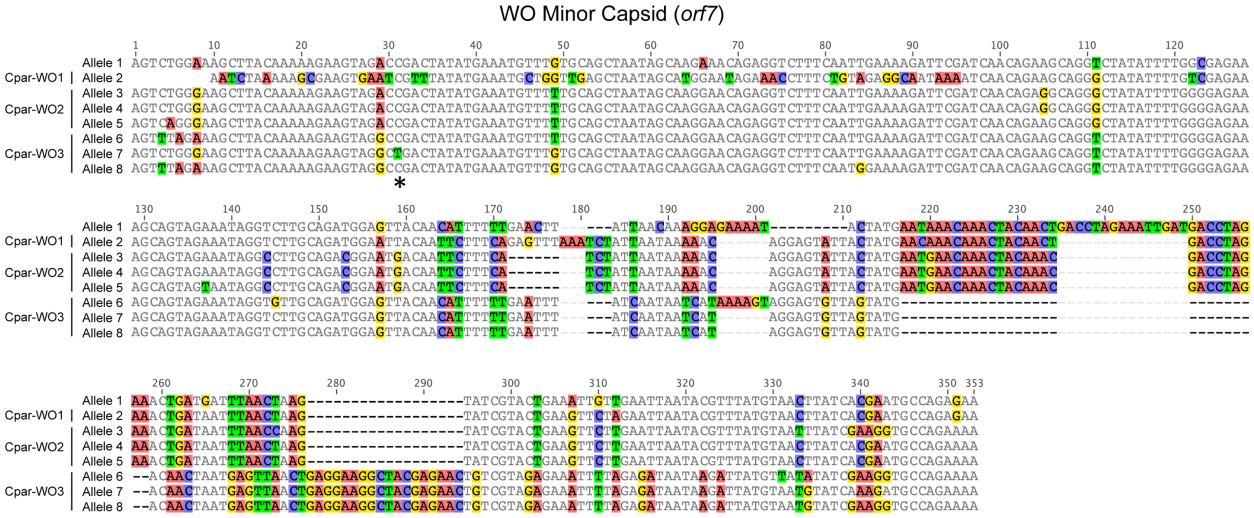

Figure III-4. Nucleotide alignment of WO minor capsid (*orf7*) alleles from hybrid grasshoppers. Asterisk indicates location of C to T substitution that introduces a premature stop codon in Cpar-WO3, allele 7. Nucleotides are counted from the start of the sequence alignment, not from the transcription start site of the gene

All alleles appear to be coding except for allele 7, which has a C to T substitution at nucleotide 31 that introduces a premature stop codon (Figure III-4). Since an identical allele was identified in another individual (604FB-5), it is unlikely that the SNP is a result of a PCR or sequencing error. Thus, at least one of the phage WO haplotypes may be undergoing pseudogenization, which is common for *Wolbachia* inserts in host genomes (Brelsfoard et al., 2014; Nikoh et al., 2008).

Genome sequencing reveals B and F Wolbachia DNA inserts in the grasshopper genome

The unexpected finding of intact phage WO genes in uninfected grasshoppers led us to characterize the genomic inserts in the *C. parallelus* genome. To do so, we pooled DNA from three uninfected grasshoppers from the Gabas population, which is a pure Cpp population in the northern tip of the hybrid zone (Figure III-1). Cpp grasshoppers were chosen for sequencing instead of hybrid individuals to limit the amount of genetic variation in the sequencing and because the Gabas population has a high prevalence of uninfected individuals (Zabal-Aguirre et al., 2010). We used Illumina high-throughput sequencing to generate 227,349,248 paired-end reads with an average length of 93.5 bp after trimming. To extract WO reads from grasshopper sequences, we first mapped all trimmed reads with a cutoff of 80% similarity over 80% read length to the reference genome of the B *Wolbachia* strain *w*Pip from *Culex quinquefasciatus* mosquitoes (Pel strain, Genbank AM999887), which has five WO prophages (Klasson et al., 2008). However, in addition to phage-related reads, we found that many of the 22,833 reads that mapped to *w*Pip fell outside of the WO prophage regions. Altogether, phage and non-phage *Wolbachia* reads covered a total of 655,940 bp (44%) of the *w*Pip reference genome when non-specific reads (i.e., reads with more than one match to the reference genome) were allowed to map randomly.

Manual observation of SNPs across the alignment revealed that many of the genes appeared to have multiple alleles, some of which were more closely related to homologs in the genome of F *Wolbachia* strain *w*Cle [Genbank AP013028] than to those in the *w*Pip B *Wolbachia* strain. Indeed, phylogenetic analyses of small contigs containing portions of the *dnaA* (Figure III-5a) or *fabG* (Figure III-5b) genes show one contig grouping with *w*Cle and the other contig grouping with its homologs from strains *w*Pip and *w*No (both B *Wolbachia* strains).

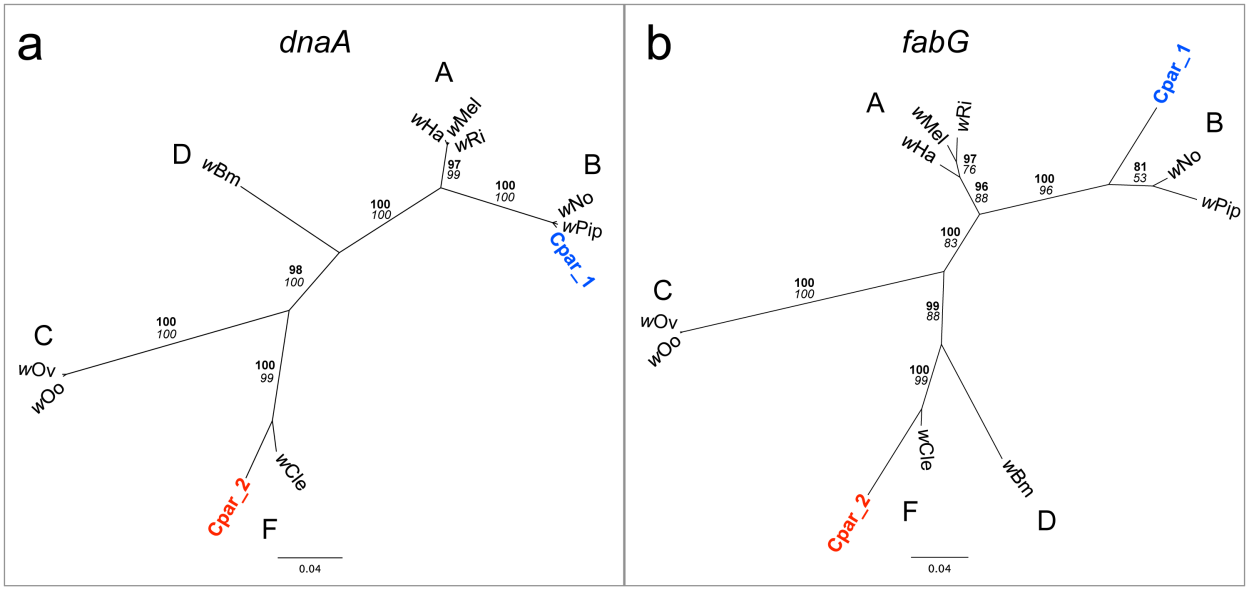

Figure III-5. Phylogenies of *Wolbachia dnaA* and *fabG* genes with *C. parallelus* inserts
Unrooted Bayesian phylogenies constructed using indel-free nucleotide alignments of Wolbachia (a) dnaA and (b) fabG genes with homologous contigs from C. parallelus genomic inserts (blue and red labels). Wolbachia supergroups (A-D, F) are indicated next to their respective clades. Posterior probability (bold) and bootstrap (italicized) values over 50 are indicated at each branch. Sequences for dnaA and fabG genes were extracted from the full genome sequences of their respective Wolbachia from NCBI (Genbank) as follows: wHa [CP003884.1], wMel [AE017196.1], wRi [CP001391.1], wNo [CP003883.1], wPip strain Pel [AM999887.1], wOo [HE660029.1], wOv strain Cameroon [HG810405.1], wBm strain TRS [AE017321.1], and wCle [AP013028.1].

To see if the sequencing reads preferentially map to *Wolbachia* from supergroups other than B or F, we simultaneously mapped all reads to the *w*Pip, *w*Cle, *w*Mel, *w*Bm, and *w*Oo reference genomes at a cutoff of 90% sequence similarity over 90% of read length. Reads were only allowed to map exclusively to one genome, and reads that mapped ambiguously to more than one genomic location were discarded. In total, 84.7% of all mapped reads (14,424 out of 17,031) mapped preferentially to *w*Pip and *w*Cle (Table III-1). A substantial number of reads (2,517) totaling 74,612 bp of the reference length also mapped to the genome of *w*Mel from the A supergroup. However, 63% of the *w*Mel reference covered by reads (47,054 out of 74,612 bp) are annotated as mobile genetic elements like phage WO, phage-associated regions adjacent to WO and transposases. Since phage WO and other mobile elements often transfer between *Wolbachia* strains (Bordenstein and Wernegreen, 2004; Chafee et al., 2010; Gavotte et al., 2007; Masui et al., 2000), these phage-related reads in the grasshopper genome do not necessarily originate from an A *Wolbachia* genome. Furthermore, the average contig length of those contigs mapping outside of the phage regions is only 113.4 bp (N50 = 100 bp), while contigs that map to

phage and mobile elements average 229.5 bp (N50 = 321 bp). With such short contigs in the non-phage regions, any mutational drift in the inserts due to relaxed selection could cause reads to incorrectly map to a supergroup that differed from that of the original donor.

Table III-1. Statistics for reads mapped to *Wolbachia* genomes from multiple supergroups

***Wolb.* genome**	**NCBI Reference #**	***Wolb.* supergroup**	**# of mapped reads**	**# of total contigs**	**Length of longest contig (bp)**	**Average length of contigs (bp)**	**Contig N50 (bp)**	**Total length of reference covered (bp)**
90% sequence similarity over 90% read length								
*w*Pip	AM999887	B	10,952	1,990	4,290	206	255	409,978
*w*Cle	AP013028	F	3,472	921	2,063	165.1	175	152,099
*w*Mel	AE017196	A	2,517	448	1,360	166.5	186	74,612
*w*Bm	AE017321	D	70	51	192	93.8	96	4,786
*w*Oo	HE660029	C	20	16	114	89.8	93	1,437
65% sequence similarity over 80% read length								
*w*Pip	AM999887	B	19,359	3,072	4,289	169.5	242	520,749
*w*Cle	AP013028	F	7,058	1,957	2,617	118.9	136	232,684
*w*Mel	AE017196	A	6,954	1,361	1,880	106.8	136	145,396
*w*Bm	AE017321	D	4,331	793	456	54.6	56	43,323
*w*Oo	HE660029	C	4,760	828	171	47.1	45	38,997

When mapping parameters were relaxed to 65% similarity over 80% read length, the number of reads mapping to all five genomes increased considerably, although the top two genomes with the most mapped reads and longest length of reference sequence covered were still *w*Pip and *w*Cle (Table III-1). Again, a substantial number of reads mapped to *w*Mel but those contigs in the non-phage regions only averaged 75 bp in length (N = 50), while those in phage regions were 228.6 bp long an average (N50 = 438). Likewise, contigs mapping to *w*Bm (D supergroup) and *w*Oo (C supergroup) only averaged 55 bp and 47 bp, respectively. With longest and thus most reliable contigs mapping to *w*Pip, *w*Cle, or phage regions, we conclude that most, if not all, *Wolbachia*-related reads in the grasshopper genome likely transferred from either a B or F *Wolbachia* strain.

When all trimmed reads were mapped simultaneously to only the *w*Pip and the *w*Cle genomes with cutoffs of 90% sequence similarity over 90% of read length and non-specific reads mapping randomly, 14,030 reads covering 493,855 bp and 3,768 reads covering 166,490 bp

mapped to the *w*Pip and *w*Cle genomes, respectively (Figure III-6). Together, both mappings covered a total of 660,345 bp, which is similar to the 655,940 bp covered when mapping to *w*Pip alone at an 80% sequence similarity over 80% read length cutoff, supporting the hypothesis that *Wolbachia* DNA in the grasshopper genome originated from both the B and F supergroups. To verify that reads mapped preferentially to one supergroup over the other, reads that mapped to either *w*Pip or *w*Cle were reciprocally mapped to the other genome with the same parameters as before (90% sequence similarity over 90% of read length). Only 12.5% of reads that mapped to *w*Pip also mapped to *w*Cle, while 18.6% of reads that mapped to *w*Cle also mapped to *w*Pip. This means that, in total, 89.1% of reads (15,332 out of 17,798) preferentially mapped to one supergroup over the other. After removing the non-specific reads, the reads that preferentially mapped to each genome covered approximately 448 kb of the *w*Pip and 144 kb of the *w*Cle reference genomes. We note appropriate caution that this analysis does not allow us to distinguish whether these are large, intact inserts or multiple smaller inserts spread throughout the genome.

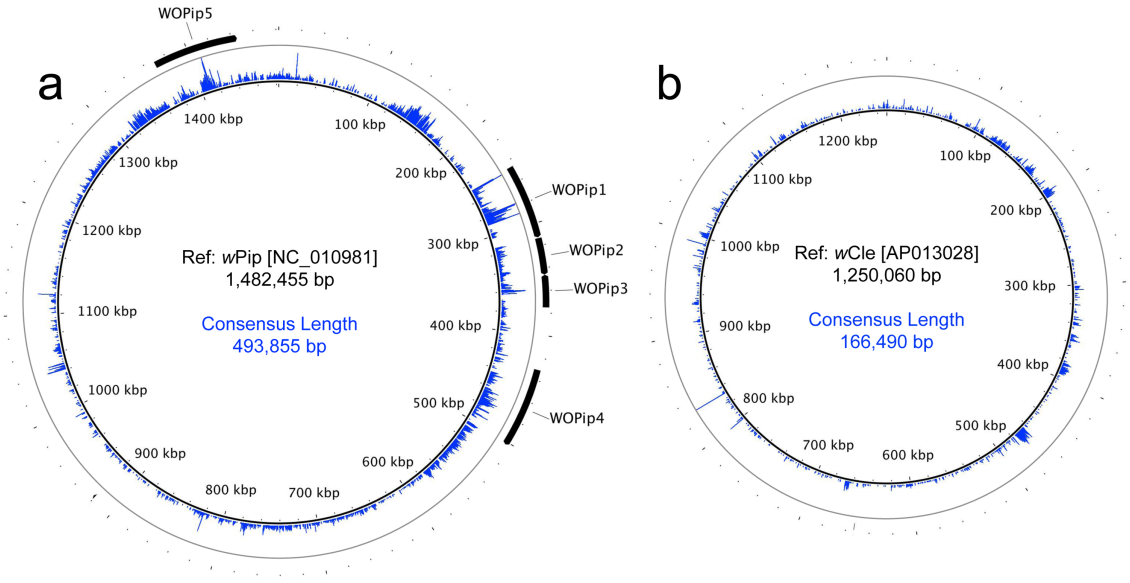

Figure III-6. Circular maps of sequencing coverage across the reference genomes of *w*Pip and *w*Cle
Mapping coverage at each base is represented in blue on the inner rings with the max coverage set at 30 (outer gray circles). WO phage regions are indicated with black arrows.

To further analyze the dual origin of the *Wolbachia* gene transfers, we computationally searched for evidence of B and F *Wolbachia* inserts that contain similar genetic repertories. In particular, we sought homologs in which the *w*Pip and the *w*Cle reference genes were both

covered by B- and F-specific reads of at least 80 bp. We then used blastn to verify that reads for each gene homolog from one insert had a greater percent sequence similarity to *w*Pip than to *w*Cle and vice versa. In total, we found 130 homologous genes that met these criteria, supporting a dual origin of the inserts.

Genome sequencing confirms multiple WO haplotypes in the grasshopper genome

Given the diversity of *orf7* alleles sequenced from uninfected hybrid grasshoppers, it is not surprising that when read coverage was mapped onto the *w*Pip (Figure III-6A) and *w*Cle (Figure III-6B) reference genomes, areas of higher coverage clustered mostly in the prophage regions (Figure III-6A). After extracting and assembling contigs from reads that mapped to the five WO minor capsid (*orf7*) genes in *w*Pip, we confirmed that there are at least three *orf7* alleles in the uninfected Cpp grasshopper genome (Figure III-7). One allele (WO2-contig) is 97.3% identical to allele 4 from the Cpar-WO2 haplotype (Figure III-7). The other two alleles are most similar to sequences from the Cpar-WO3 haplotype: WO3-contig1 is 97.5% identical to allele 6 and WO3-contig2 is 100% identical to allele 7 (Figure III-7). We did not find any *orf7* alleles from the Cpar-WO1 haplotype in the genomic contigs, which may be a consequence of low sequencing coverage. However, if Cpar-WO1 is absent from the Cpp genome, then it may be specific to the Cpe subspecies or could even be unique to hybrids if the horizontal transfer occurred after establishment of the hybrid zone.

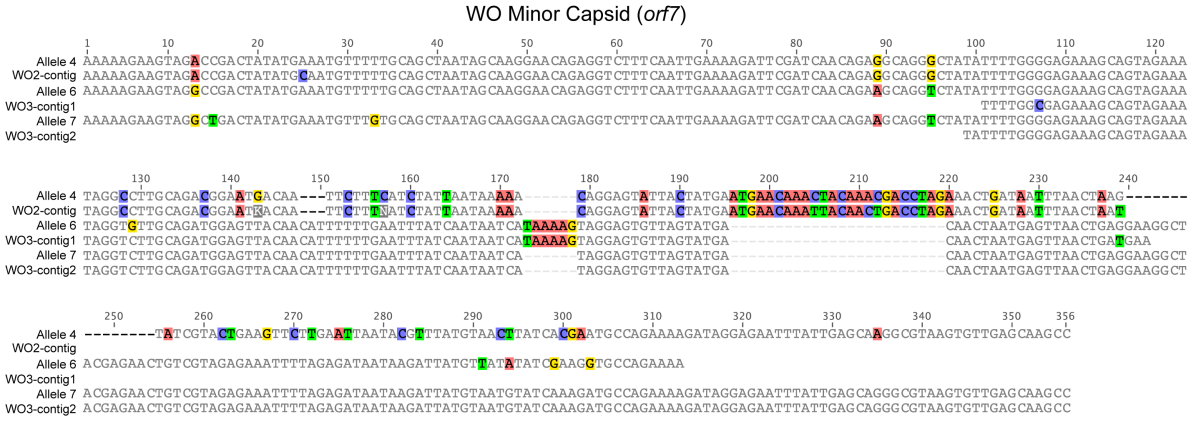

Figure III-7. Alignment of WO minor capsid sequences from cloning and Sanger sequencing with assembled contigs from Illumina sequencing

Contigs are grouped with their most similar WO allele identified through Sanger sequencing. Nucleotides are counted from the start of the sequence alignment, not from the transcription start site of the gene.

FISH localizes Wolbachia inserts in grasshopper chromosomes

Even though, on average, 70% of individual grasshoppers from the Gabas population are uninfected with *Wolbachia* (Zabal-Aguirre et al., 2010), it is possible that the "uninfected" grasshoppers from Gabas had a low-titer *Wolbachia* infection that accounts for the sequencing of copious *Wolbachia* genes. This explanation is highly unlikely because PCR for two essential bacterial genes, 16S rRNA and *ftsZ*, failed to detect a product in all three grasshoppers pooled for sequencing, while PCR of WO genes amplified a band in all individuals for the *orf7* gene (Figure III-8).

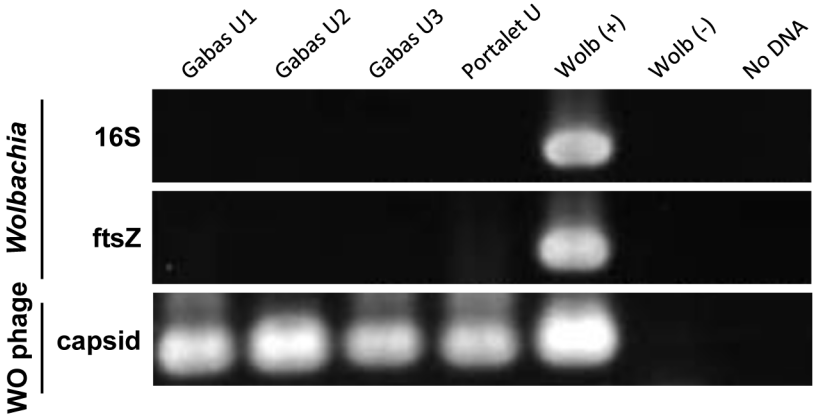

Figure III-8. PCR amplification of *Wolbachia* and WO genes
Wolbachia genes (16S rRNA and *ftsZ*) and the phage WO *orf7* gene were amplified from Gabas uninfected grasshoppers used for high-throughput sequencing (Gabas U1-U3) and a Portalet uninfected grasshopper (Portalet U). Wolb (+) and Wolb (-) controls are from *Wolbachia*-infected and tetracycline-cured lines of *Nasonia girault*i, respectively. No DNA control had no template added to the PCR reaction.

Moreover, to confirm *Wolbachia* insertions in the grasshopper genome, we used tyramide-coupled FISH to physically map *Wolbachia* genomic insertions in Cpe (Figure III-9a) and Cpp (Figure III-9b) chromosomes of uninfected male individuals. Hybridization of fluorescent DNA probes designed from a contig from the B *Wolbachia* insert revealed a discrete, repeatable distribution pattern along chromosomes in the karyotype (Figure III-9), particularly in telomeric constitutive heterochromatin and in some interstitial regions. When comparing the distribution of this contig on the chromosomes of Cpp and Cpe, some signals are present at homologous chromosomal locations in both genomes, such as on chromosome 4 (Figure III-9, white arrows), while other inserts, like that on chromosome 3 in Cpp (Figure III-9, red arrows), are subspecies-specific, suggesting that the former are ancestral to the last common ancestor of Cpp and Cpe, whilst the latter appeared after taxon divergence.

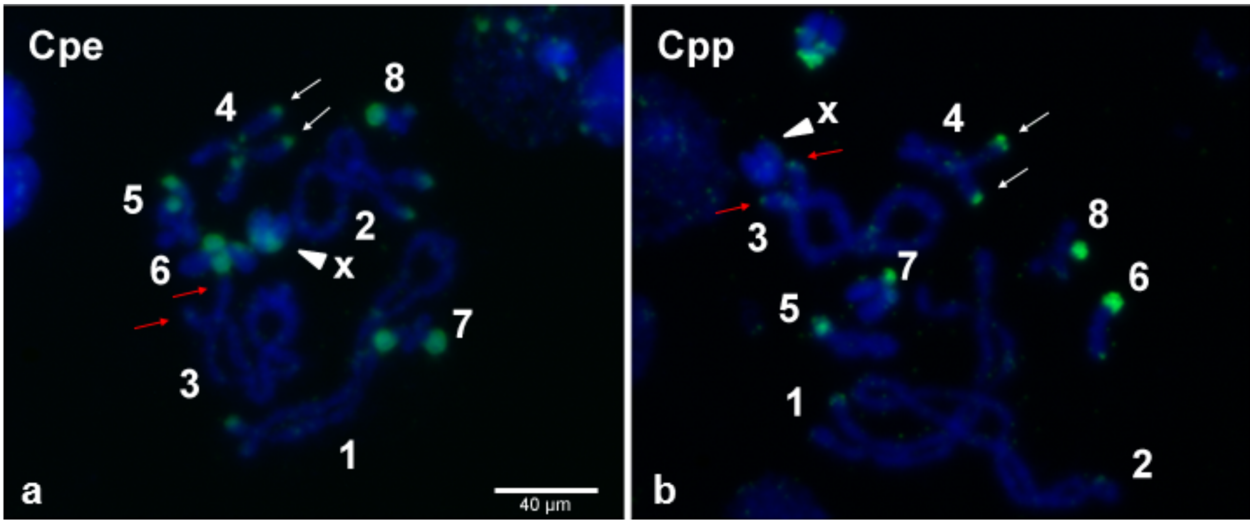

Figure III-9. *Wolbachia* inserts localized to *C. parallelus* chromosomes
Tyramide-coupled FISH Fluorescein signals using the Cpar-Wb1 probe reveal presence of *Wolbachia* genomic inserts (green fluorescence) in *C. parallelus erythropus*, Cpe (a) and *C. parallelus parallelus*, Cpp (b) meiotic chromosomes (blue fluorescence). Hybridization of *Wolbachia* insertions is abundant in telomeric regions of several chromosomes, certain interstitial regions and on chromosome X (arrowhead). White arrows mark a *Wolbachia* insert that coincides in homologous chromosomes of both Cpe and Cpp, while red arrows indicate a subspecies-specific insert present in Cpp but not Cpe. Numbers correspond to chromosome pairs (bivalents). Scale bar = 40 µm.

Discussion

The *Chorthippus parallelus* hybrid zone is an excellent model for symbiosis research since *Wolbachia* infection status is highly variable, with individuals collected at the same geographical location infected with either F or B *Wolbachia*, co-infected with both or naturally uninfected (Zabal-Aguirre et al., 2010). Though *Wolbachia* diversity has previously been investigated in this system, this work comprises the first screen for *Wolbachia's* temperate phage WO. We set out to characterize the types of phage WO present in the population and to determine whether co-infection with *Wolbachia* strains from divergent B and F supergroups facilitated transfer of phage WO between *Wolbachia.* Instead, we discovered an unexpected diversity of phage WO *orf7* alleles and multiple instances of horizontal transfer of the phage WO *orf7* gene to hybrid and non-hybrid grasshopper genomes. In total, we identified eight unique *orf7* alleles from nine different individuals collected from a single hybrid population. Genome sequencing of Cpp grasshoppers confirmed that three of these alleles (4, 6, and 7) predate

secondary hybridization of Cpp and Cpe subspecies, while the other alleles may have been introduced to the hybrid zone by Cpe or may be unique to hybrid populations.

Since many of the alleles are so similar to others (≥96% identical), they may represent allelic variation at the same locus in the diploid grasshopper genome instead of independent gene transfers. Thus, we conservatively classified similar alleles into three phage "haplotypes". Interestingly, we did not conclusively identify *orf7* alleles that were specific to *Wolbachia* cytoplasmic infections even though many of the grasshoppers were infected by B and/or F *Wolbachia.* It is likely that the cytoplasmic *Wolbachia* infections harbor phage WO with *orf7* sequences that are so similar to those in the host genome that we cannot distinguish between the two. For example, alleles 3 and 7 were only sequenced from B-infected individuals and may reside in the cytoplasmic B *Wolbachia* genome, but further genome sequencing of the cytoplasmic *Wolbachia* is needed to verify this observation.

After sequencing the genome of uninfected Cpp grasshopers, we discovered that not only phage genes had transferred to the host genome but also large regions of both B and F *Wolbachia*. Many animal hosts that harbor or once harbored *Wolbachia* have evidence of *Wolbachia* DNA in their genomes (Bordenstein, 2007; Dunning Hotopp, 2011; Dunning Hotopp et al., 2007), probably because *Wolbachia* are uniquely poised for symbiont-to-host gene exchange since they target the germ-line stem cell niche during host oogenesis (Fast et al., 2011; Robinson et al., 2013; Toomey et al., 2013). *Wolbachia* nuclear inserts can be quite large and cover a substantial portion of a *Wolbachia* genome. For example, approximately 30% of a *Wolbachia* genome is inserted in the X-chromosome of the bean beetle *Callosobruchus chinensis* (Kondo et al., 2002; Nikoh et al., 2008), while an estimated 180 kb of *Wolbachia* DNA is present in the genome of the longicorn beetle *Monochamus alternatus* (Aikawa et al., 2009). Multiple *Wolbachia* insertions in the same host genome have also been identified. Several *Drosophila ananassae* populations have multiple copies of an entire *Wolbachia* genome on one of their chromosomes (Dunning Hotopp et al., 2007; Klasson et al., 2014), while the tsetse fly *Glossina morsitans morsitans* genome has three *Wolbachia* chromosomal inserts with the two largest inserts each covering roughly half a *Wolbachia* genome at 527 kb and 484 kb (Brelsfoard et al., 2014). The large *Wolbachia* inserts in this case are highly similar to each other and also closely-related to the tsetse fly cytoplasmic *Wolbachia* strain, *w*Gmm, suggesting a single transfer from *w*Gmm to the tsetse fly genome followed by duplication of the insert, though independent

transfer events cannot be ruled out (Brelsfoard et al., 2014). Either way, both insertions came from the same *Wolbachia* supergroup and likely from the same *Wolbachia* strain.

In our study, phylogenetic analyses of variable contigs mapping to the same *Wolbachia* genes revealed that inserts in the *C. parallelus* genome likely originated from both B and F *Wolbachia.* To our knowledge, this is the first case of substantial *Wolbachia* DNA transfer from divergent supergroups into the same host genome. Similar techniques used to analyze the genomes of *Wolbachia*-free nematodes such as *Acanthocheilonema viteae, Onchocerca flexuosa*, *Loa loa,* and *Dictyocaulus viviparus* found ancient remnants of *Wolbachia* genes that appear to have originated from multiple supergroups when compared to present-day cytoplasmic *Wolbachia* genes (Desjardins et al., 2013a; Koutsovoulos et al., 2014; McNulty et al., 2010). However, the antiquity of these horizontal transfer events makes accurate phylogenetic inferences difficult, especially since the *Wolbachia* genes in the nematode host are no longer under the same selective pressures as cytoplasmic *Wolbachia* genes. For example, McNulty et al. 2010 estimates that "fossilized" evidence of *Wolbachia* sequences in the genomes of *A. viteae* and *O. flexuosa* must be several million years old based on their low percent identities (78% and 81%, respectively) to any contemporary *Wolbachia* sequences. In contrast, average percent identities of the B *Wolbachia* gene variants to *w*Pip and the F *Wolbachia* gene variants to *w*Cle for the 130 shared genes in the *C. parallelus* inserts are 94 ± 0.05% and 93 ± 0.04%, respectively.

The higher percent identity to a contemporary *Wolbachia* strain for the grasshopper inserts suggest that they have transferred more recently and/or are better preserved in the grasshopper genome due to the unique evolutionary dynamics of grasshopper genomes. Orthopterans like grasshoppers, locusts and crickets are known for their enormous genomes, and *C. parallelus* grasshoppers have one of the largest genomes in the order with estimates ranging from 12.3 to 14.7 Gb (Lechner et al., 2013). Genome gigantism in Orthoptera is thought to largely be due to frequent acquisition of new genetic material coupled with slow rates of DNA loss (Bensasson et al., 2001; Bensasson et al., 2000; Song et al., 2014). For example, Orthopteran genomes exhibit unusually high rates of DNA transfer from mitochondria to the nuclear genome (Bensasson et al., 2000; Song et al., 2014), and, with the slow rate of DNA loss, some of these inserts have remained intact for 150 million years (Song et al., 2014). Based on the rate of mitochondrial gene acquisition, grasshopper genomes may be more amenable to horizontal gene

transfer in general, especially from intracellular cytoplasmic entities like mitochondria or *Wolbachia.* It is not surprising, then, to presume that DNA from both B and F *Wolbachia* would eventually wind up in the *C. parallelus* host genome.

The dynamic nature of *Wolbachia* lateral gene transfer to the *C. parallelus* genome is evident when visualized with FISH. Some inserts are present at the same position on the chromosomes of both Cpp and Cpe while other inserts are subspecies-specific, indicating that insertion events likely occurred both before and after the divergence of the subspecies. Our sequencing of the WO *orf7* gene supports this hypothesis since the Cpar-WO2 and Cpar-WO3 haplotypes are present in the genomes of Cpp individuals from Gabas and in hybrids from Portalet, while the Cpar-WO1 haplotype was only detected in Portalet. Subspecies-specific sequences are likely relatively young since the two subspecies are estimated to have diverged between 0.2 and 2 MYA (Cooper and Hewitt, 1993; Lunt et al., 1998). If hybrid-specific inserts arose independently, they would be even younger since the transfer would have had to occur after the formation of the hybrid zone roughly 9,000 years ago (Hewitt, 1993; Shuker et al., 2005a). Thus, slow rates of DNA loss coupled with relatively recent transfer events allows standard phylogenetic analyses to easily identify and distinguish the inserts in the *C. parallelus* genome as originating from either a B or F *Wolbachia*, whereas *Wolbachia* inserts in nematode genomes may be too divergent to accurately predict the donor *Wolbachia's* supergroup.

Instead of independent transfers, B and F *Wolbachia* strains may have recombined to produce a *Wolbachia* strain with genes from both supergroups and part of this "hybrid" *Wolbachia* genome transferred as a single event into the *C. parallelus* genome. This scenario appears unlikely as we identified 130 *Wolbachia* genes with multiple alleles from both B and F *Wolbachia* in the genomic inserts. A recombinogenic genome with substantial genetic redundancy of essential genes is improbable given that endosymbiont genomes tend to be relatively streamlined (Newton and Bordenstein, 2011; Wernegreen, 2002). Furthermore, FISH analyses verified the presence of *Wolbachia* DNA in multiple locations on the *C. parallelus* chromosomes and further characterization of the inserts and their evolutionary history is in progress.

Conclusion

Alongside genetic introgression, animal hybrid zones offer an avenue for symbiont exchange, especially for heritable endosymbionts like *Wolbachia* (Mandel et al., 2001; Zabal-Aguirre et al., 2010). Resulting co-infections of multiple *Wolbachia* strains in a hybrid host provide opportunities for genetic exchange within the intracellular arena (Bordenstein and Reznikoff, 2005; Metcalf and Bordenstein, 2012; Newton and Bordenstein, 2011). Though exchange of bacteriophage WO occurs often between co-infections of A and B *Wolbachia* (Bordenstein and Wernegreen, 2004; Chafee et al., 2010; Kent et al., 2011; Masui et al., 2000), we found no evidence for phage WO transfer among B and F *Wolbachia* in hybrid *C. parallelus* grasshoppers. Instead, we found that horizontal gene transfer is clearly a dynamic process in *C. parallelus*, with two discrete *Wolbachia* supergroups (B and F) transferring approximately 448 kb and 144 kb of DNA, respectively, to the host genome. Since many insects are co-infected with *Wolbachia* from different supergroups, it is curious why there are not more insect genomes with *Wolbachia* inserts of dual origin. Part of the answer is likely that other genomes with inserts of dual origin have simply not been sequenced yet. However, grasshopper and other Orthopteran genomes, with their high rates of DNA acquisition and slow rates of DNA loss, may be uniquely poised for acquiring *Wolbachia* genes and maintaining them relatively intact for long periods of time, allowing phylogenetic analyses to accurately distinguish between different supergroups. Though the gigantic genomes of Orthopterans currently make them challenging to sequence and assemble, it will be interesting to see if more species of this undersampled insect order also have DNA from multiple endosymbionts in their genomes.

CHAPTER IV. THE GENETICS OF *WOLBACHIA* TITER REGULATION IN *NASONIA* PARASITOID WASPS

Abstract

Many animals maternally transmit microbial symbionts in the face of profound fitness consequences should symbiont titers go awry. However, little is known about the evolution of host genes involved in symbiont titer control or the molecular mechanisms by which this regulation is achieved. Here we use the first forward genetic analysis to dissect the host genetic architecture governing the regulation of a widespread bacterial symbiont, *Wolbachia,* in the *Nasonia* parasitoid wasp model. Interspecific transfer of *Wolbachia* strain *w*VitA from its resident host *N. vitripennis* to the closely-related species *N. giraulti* results in an 80-fold increase in infection titers. Using genetic tools including introgression, genotyping microarrays, quantitative trait loci analyses and RNA-seq, we identify the host genomic regions and potential mechanisms that underlie this symbiont regulation. We report three findings: (i) A maternal suppressor acts dominantly in *N. vitripennis* to establish the native low infection level in offspring. (ii) Two genomic regions, one each on chromosomes 2 and 3 (out of the five *Nasonia* chromosomes), underlie this maternal suppression trait and (iii) RNA-seq of *Nasonia* ovaries identified 33 differentially-expressed genes in the candidate regions, several of which function in pathways important for host control of intracellular bacteria including immunity, autophagy, and cell-to-cell trafficking. Taken together, this forward-genetic investigation highlights the significance of maternal regulation of inherited symbionts through a few key genomic regions and raises the prospects of identifying host symbiosis genes that control maternally-transmitted titers.

Introduction

All animals live in symbiosis with microbes, many of which play beneficial roles in host processes as diverse as nutritional uptake and metabolism (Ley et al., 2006; Turnbaugh et al., 2006), immune cell development (Ivanov et al., 2009; Ivanov et al., 2008; Round et al., 2011), and pathogen resistance (Candela et al., 2008; Fukuda et al., 2011). However, even innocuous microbes may become pathogenic when not properly regulated by the host (Calderone and Fonzi,

2001; Mitchell, 2011). In animals harboring vertically transmitted microbes, a delicate balance of symbiont regulation must be achieved where symbiont titers are high enough to ensure efficient transmission but not excessive enough to prove detrimental to host fitness. Co-evolution between insect hosts and their resident, maternally-transmitted symbionts may promote the development of unique host-symbiont interactions that maintain symbiont densities at specific levels within the host (Chafee et al., 2011; Kim et al., 2013; Kondo et al., 2005; Login and Heddi, 2013; Mouton et al., 2003; Rio et al., 2006). Some of these interactions are even strain-specific: each strain of a particular symbiont in a multiply-infected host is present at a unique but stable density, even when other, co-infecting strains are removed (Ijichi et al., 2002; Mouton et al., 2004; Mouton et al., 2003). When these unique host-symbiont interactions are disrupted through the transfer of a symbiont into a naïve host, control over the symbiont is often lost, leading to overproliferation and/or an expanded tissue tropism not witnessed in its original host species (Bian et al., 2013; Chafee et al., 2011; Le Clec'h et al., 2012; Le Clec'h et al., 2013).

The repeated evolution of maternal microbial transmission across diverse animal taxa (CHAPTER II) suggests that the evolutionary events required to balance different heritable host-microbe combinations may not be complex, but rather have a simple genetic basis. However, little is known about the molecular mechanisms by which vertically-transmitted symbionts gain access to the germ line or about how hosts regulate microbial titers within their reproductive organs. Reverse genetic studies in insects have intermittently discovered immune or developmental genes that affect endosymbiont densities, such as a peptidoglycan recognition protein (PGRP-LB) in tsetse flies (Wang et al., 2009), an antimicrobial peptide (ColA) in weevils (Login et al., 2011), and the embryonic axis determination gene *gurken* or actin-binding proteins profilin and villin in *Drosophila* (Newton et al., 2015; Serbus et al., 2011). However, these studies do not examine naturally-occurring species-specific genetic variation underlying endosymbiont densities. Thus, we are not aware of any forward genetic studies that have dissected the number and types of host genes that establish and maintain transmission and/or suppression of symbiont densities.

Here we utilize a major host interspecific difference in titers of the heritable endosymbiont *Wolbachia* to map *Nasonia* genes that control *Wolbachia* densities. The *Nasonia* genus (Order Hymenoptera) of parasitoid wasps is comprised of four closely-related species, with *N. vitripennis* last sharing a common ancestor with the other three species around 1 MYA

(Campbell et al., 1993; Raychoudhury et al., 2010). All four species are naturally infected with different *Wolbachia* strains, mostly acquired through horizontal transfer after species divergence (Raychoudhury et al., 2009). *Nasonia* have many advantages as a model genetic system, including haplodiploid sex determination, fully sequenced genomes (Werren et al., 2010), systemic RNAi (Lynch and Desplan, 2006; Werren et al., 2009), and the ability to produce viable, fertile hybrid offspring. The production of hybrid offspring means that genetic or cytoplasmic material (including intracellular bacteria) can easily be transferred between *Nasonia* species, making *Nasonia* an excellent model for studying the evolution of interspecific traits such as wing size (Gadau et al., 2002; Loehlin et al., 2010a; Loehlin et al., 2010b), head shape (Werren et al., 2015), sex pheromones (Niehuis et al., 2013; Niehuis et al., 2011), and memory retention (Hoedjes et al., 2014).

Wolbachia (Order Rickettsiales) are maternally-transmitted, obligate, intracellular bacteria that infect 40-52% of all arthropod species (Weinert et al., 2015; Zug and Hammerstein, 2012). In most insects, *Wolbachia* function as reproductive parasites that manipulate host reproduction through a variety of mechanisms to achieve a greater proportion of females in the host population (Serbus et al., 2008; Werren et al., 2008). Both efficient transovarial transmission of the parasite and their ability to manipulate host reproduction often depend upon sufficiently high within-host *Wolbachia* densities (Dyer et al., 2005; Perrot-Minnot and Werren, 1999), while overproliferation of *Wolbachia* has been shown to drastically reduce lifespan in *Drosophila* (McGraw et al., 2002; Min and Benzer, 1997), mosquitoes (McMeniman et al., 2009; Suh et al., 2009) and terrestrial isopods (Le Clec'h et al., 2012). Thus, over time, insects likely adapt to suppress the proliferation of their own resident *Wolbachia* strains.

We previously showed that transfer of a specific *Wolbachia* strain (*w*VitA) from one species of *Nasonia* (*N. vitripennis*) to a closely-related species (*N. giraulti*) through cytoplasmic introgression results in a remarkable 80-fold increase of the *Wolbachia* strain in its new host (Figure I-2) (Chafee et al., 2011). Additionally, tight localization of *Wolbachia* to the posterior pole of *N. giraulti* embryos breaks down (Figure I-3) and leads to an expanded tissue tropism beyond the reproductive organs in *w*VitA-infected *N. giraulti* adults (Figure IV-1A) (Chafee et al., 2011). The consequences of an increased bacterial load in *N. giraulti* include a reduction in fecundity, an increase in levels of cytoplasmic incompatibility, male-to-female transfer of the bacteria to uninfected females, and an increased acceptance of interspecific mates by densely-

infected females (Chafee et al., 2011). Importantly, *w*VitA densities return to normal when *w*VitA is introgressed back into a *N. vitripennis* genomic background from the high-density *N. girault*i line (IntG) (Figure IV-1B). Since both *Nasonia* lines have the same *N. vitripennis* cytoplasm, the interspecific *Wolbachia* density phenotype must be established by differences in the host nuclear genome (Chafee et al., 2011).

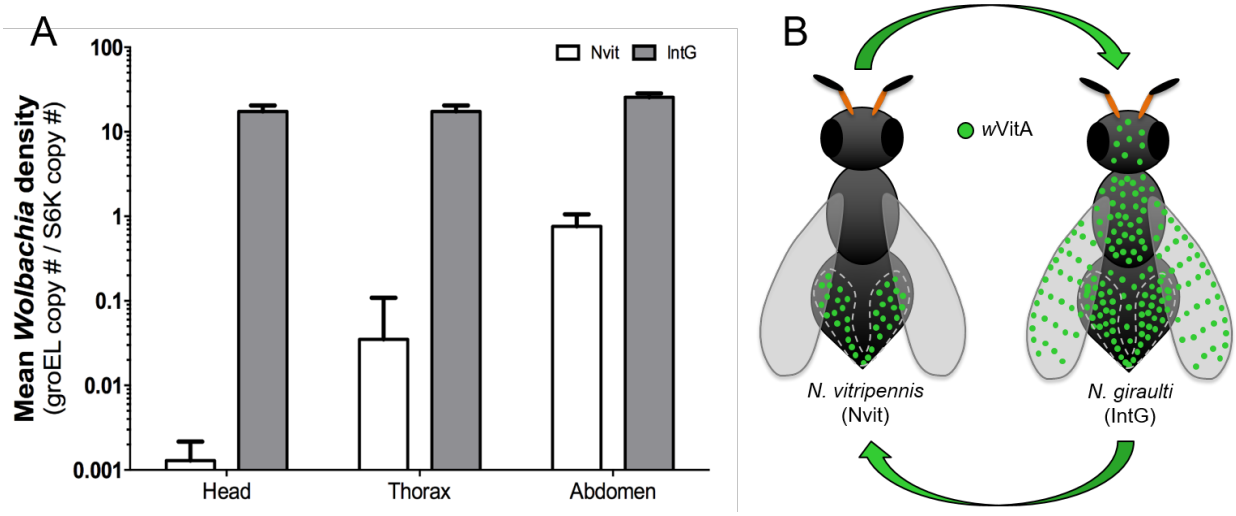

Figure IV-1. Expanded tissue tropism of *w*VitA in *N. girault*i
(A) Quantitative PCR of *w*VitA densities in the three body segments of *N. vitripennis* (white bars) and *N. girault*i IntG (gray bars). (B) Cartoon of *w*VitA (green dots) localization in its native host (*N. vitripennis*) versus a new host (*N. girault*i). Green arrows indicate that the density trait is reversible (*w*VitA transferred from IntG back to Nvit return to lower densities).

In this study, we used a comprehensive set of tools spanning selective introgressions, a *Nasonia* genotyping microarray (Desjardins et al., 2013b), quantitative trait loci analyses (QTL), marker-assisted introgression, and RNA-seq to uncover the underlying genetic architecture for *w*VitA density regulation in *Nasonia.*

Materials and Methods

Nasonia *strains and maintenance*

Experiments were performed with *N. vitripennis* 12.1, *N. girault*i IntG12.1 or hybrids of these two species. *N. vitripennis* 12.1 is singly-infected with native *Wolbachia* strain *w*VitA and was derived from *N. vitripennis* R511 (*w*VitA and *w*VitB-infected) after a prolonged period of diapause (Perrot-Minnot et al., 1996). *N. girault*i strain IntG12.1 was generated by backcrossing

N. vitripennis 12.1 females to uninfected *N. giraulti* Rv2x(u) males for nine generations (Chafee et al., 2011), producing hybrids with an *N. giraulti* genome and an *N. vitripennis* cytoplasm harboring *w*VitA. RNA-seq experiments included *N. giraulti* 16.2, which is singly-infected with native *Wolbachia* strain *w*GirA. The IntC3 introgression line contains an *N. vitripennis* chromosome 3 candidate region in an *N. giraulti* genomic background. All *Nasonia* were reared at 25°C in constant light on *Sarcophaga bullata* fly hosts.

Quantitative analysis of Wolbachia *densities*

Genomic DNA was extracted from pupae or adult *Nasonia* using the Gentra Puregene Tissue Kit (Qiagen) according to the manufacturer's protocol. Real-time quantitative PCR (qPCR) was performed on a CFX96 Real-Time system (Bio-Rad) using a total reaction volume of 25 µl: 12.5 µl of iQ SYBR Green Supermix (Bio-Rad), 8.5 µl of sterile water, 1.0 µl each of 5 µM forward and reverse primers, and 2 µl of target DNA in single wells of a 96-well plate (Bio-Rad). All qPCR reactions were performed in technical duplicates and included a melt curve analysis to check for primer dimers and nonspecific amplification. Selective amplification was performed using primers previously described for the *Wolbachia groEL* gene (Bordenstein et al., 2006) and *Nasonia NvS6K* gene (Bordenstein and Bordenstein, 2011). Standard curves for each gene were constructed as previously described (Bordenstein and Bordenstein, 2011) using a log10 dilution series of larger PCR products of known concentrations for each gene. *groEL* and *S6K* copy numbers for each sample were calculated based on the following standard curve equations: *groEL*: $y = -3.367x + 35.803$ and *S6K:* $y = -3.455x + 35.908$, where y = averaged Ct value between technical duplicates and x = log starting quantity of template DNA. *Wolbachia* density was calculated by dividing *groEL* copy number by *S6K* copy number for each sample.

Microsatellite marker genotyping

Primers used to amplify microsatellite markers that differ in size between *N. vitripennis* and *N. giraulti* are listed in Table B-1. Microsatellite markers not previously published were identified by aligning *N. vitripennis* and *N. giraulti* genomic sequences using the Geneious alignment tool in Geneious Pro v5.5.8 (Biomatters). The Geneious primer design tool was then used to generate primer sets spanning each microsatellite. All PCR reactions were run on a Veriti Thermal Cycler (Applied Biosystems) with a total reaction volume of 15 µl: 7.5 µl of GoTaq

Green Master Mix (Promega), 3.6 µl of sterile water, 1.2 µl of 5µM forward and reverse primers, and 1.5 µl of target DNA. PCR products were run on 4% agarose gels in TBE buffer (Sigma) at 90 volts for 2.5 to 6 hours, stained with GelRed (Biotium) according to manufacturer's protocol, and imaged on a Red Personal Gel Imager (Alpha Innotech).

Phenotype-based selection and introgression coupled with a genotyping microarray

N. vitripennis females (low *w*VitA density) were backcrossed with *N. giraulti* IntG males (high *w*VitA density) for nine generations. At each generation of backcrossing, five female pupal offspring were pooled from each hybrid female, and the pupal *Wolbachia* densities were measured using qPCR. Sisters of the pupae with the lowest *Wolbachia* densities were then used as mothers in the next round of backcrossing. Two independent selection lines were maintained simultaneously along with control lines of pure-breeding *N. vitripennis* and *N. giraulti*. After eight generations of selection, the three females from each introgression line that produced ninth-generation offspring with the lowest *Wolbachia* densities were pooled and their DNA extracted using the DNeasy Blood and Tissue Kit (Qiagen) with the protocol for purification of DNA from insects. To obtain enough DNA for microarray hybridization, we used the REPLI-g Mini Kit (Qiagen) with the protocol for 5 µl of DNA template to amplify genomic DNA overnight at 30 °C, then purified the DNA using ethanol precipitation. The final concentration for each sample was diluted to 1 µg/µl and a total of 10 µl was sent to The Center for Genomics and Bioinformatics at Indiana University to be processed on a *Nasonia* genotyping microarray (Roche NimbleGen) tiled with probes for 19,681 single nucleotide polymorphisms and indels that differ between *N. vitripennis* and *N. giraulti* (Desjardins et al., 2013b).

For each sample, the proportion of *N. vitripennis* alleles at each marker was determined based on the ratio of hybridization to the *N. vitripennis*-specific probe versus hybridization to the *N. giraulti*-specific probe, as previously described (Desjardins et al., 2013b). To verify species-specificity of these markers for our *Nasonia* strains, we also genotyped *N. vitripennis* 12.1 and *N. giraulti* IntG control females on the array, and markers that did not display the correct specificity within one standard deviation of the median were removed from subsequent analyses (5,301 markers total). The remaining markers were then manually mapped back to the most recent *Nasonia* linkage map (Desjardins et al., 2013b). Since all introgression females received one copy of their diploid genome from their *N. giraulti* father, the theoretical maximum proportion of

N. vitripennis alleles at each marker cluster for experimental samples is 0.5. The proportion of *N. vitripennis* alleles was averaged for every 22 consecutive markers across each chromosome, and heat maps were generated using the HeatMap function in MATLAB (MathWorks).

QTL Analysis

F2 hybrid females (N = 191) were generated by backcrossing F1 *N. vitripennis*/*N. giraulti* hybrids to *N. giraulti* IntG males. F2 females were then backcrossed again to *N. giraulti* IntG and allowed to lay offspring. Five female pupae from each F2 female were pooled and their *Wolbachia* densities measured using qPCR. Females that produced offspring with densities within the highest and lowest quartile of the density distribution (N = 42 for each quartile) were selectively genotyped with 47 microsatellite markers (Table B-1) spread across chromosomes 1, 2 and 3 with an average distance of 3 cM between markers. Phenotypic information for all 191 F2 females was included in the mapping analyses to prevent inflation of QTL effects due to the biased selection of extreme phenotypes (Lander and Botstein, 1989). QTL analyses were performed in R (version 3.0.2) with package R/qtl (Broman et al., 2003). Significance thresholds for our dataset were calculated by using a stratified permutation test with the scanone function (1000 permutations). To identify significant QTL and their interactions, we first conducted a one-dimensional, one-QTL scan and a two-dimensional, two-QTL scan using the EM algorithm with a step size of 1 cM and an assumed genotype error probability of 0.001. Two significant QTLs were identified, one each on chromosomes 2 and 3, which were predicted to act additively. The positions of identified QTL were then refined using multiple QTL modeling with the multiple imputation algorithm (200 imputations, step size = 1 cM) assuming a model with two additive QTLs. 95% Bayes credible intervals were calculated for each QTL after multiple QTL modeling using the bayesint function.

Marker-assisted segmental introgressions

Marker-assisted segmental introgression lines were generated by repeatedly backcrossing hybrid females to *N. giraulti* males for nine generations while selecting for *N. vitripennis* alleles at three microsatellite markers per QTL region (Chr2: MM2.17, MM2.26, and MM2.36; Chr3: MM3.17, NvC3-18, and MM3.37; Table B-1). After the ninth generation, families that maintained an *N. vitripennis* allele at one or more of these markers were selected, and siblings

were mated to each other to produce lines containing homozygous *N. vitripennis* regions at and around the markers. Unfortunately, due to hybrid incompatibilities that arose when some markers (MM2.26 and NvC3-18) located at or around the centromere were made homozygous, some lines were left heterozygous and allowed to mate randomly. For this reason, individual adult females from each segmental line were genotyped and phenotyped separately (N = 10 – 15 females per line). Females were hosted as virgins, five male pupal offspring per female were pooled, and pupal *Wolbachia* densities were measured using qPCR. Variation across plates for a single experiment was reduced by including a set of parental DNA controls on all plates. The parental fold-change was then calculated by dividing the average *N. giraulti* control density by the average *N. vitripennis* control density. To calculate the sample fold-change, the absolute density for each sample was divided by the average density of the *N. vitripennis* control. To determine how "effective" each segmental introgression line was at reducing densities, we calculated the percent effect on density suppression for each sample using the following equation:

$$\%\, effect\ \ on\ density\ suppresion = \left(1 - \frac{sample\ fold\ change}{parental\ fold\ change}\right) \times 100$$

Each female was genotyped with markers across the region of interest, all females with identical genotypes across all markers were grouped together into a single "haplotype", and their percent effects on density suppression were averaged.

RNA-seq of ovaries

One-day old females from *Nasonia* strains *N. vitripennis* 12.1, *N. giraulti* IntG and *N. giraulti* 16.2 were hosted as virgins on *S. bullata* pupae for 48 hours to stimulate feeding and oogenesis. Females were then dissected in RNase-free 1X PBS buffer, and their ovaries were immediately transferred to RNase-free Eppendorf tubes in liquid nitrogen. Fifty ovaries were pooled for each replicate and three biological replicates were collected per *Nasonia* strain. Ovaries were manually homogenized with RNase-free pestles, and their RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol for purification of total RNA from animal tissues. After RNA purification, samples were treated with RQ1 RNase-free DNase (Promega) for 1 hour at 37 °C, followed by an ethanol precipitation with 1/10th volume 3M sodium acetate and 3 volumes 100% ethanol incubated overnight at -20 °C. PCR of

samples with *Nasonia* primers NvS6KQTF4 and NVS6KQTR4 (Bordenstein and Bordenstein, 2011) revealed some residual DNA contamination, so DNase treatment and ethanol precipitation were repeated. After the second DNase treatment, PCR with the same primer set confirmed absence of contaminating DNA. Sample RNA concentrations were measured with a Qubit 2.0 Fluorometer (Life Technologies) using the RNA HS Assay kit (Life Technologies). Approximately 400 ng of each sample was converted to cDNA using the SuperScript VILO cDNA Synthesis Kit (Invitrogen), then shipped to the University of Rochester Genomics Research Center for sequencing. Library preparation was performed using the Illumina TruSeq Stranded mRNA Sample Preparation Kit, and all samples were run multiplexed on a single lane of the Illumina HiSeq2500 (single-end, 100 bp reads). Raw reads were trimmed and mapped to the *N. vitripennis* genome Nvit_2.1 (GCF_000002325.3) in CLC Genomics Workbench 8.5.1, allowing ten gene hits per read using a minimum length fraction of 0.9 and a minimum similarity fraction of 0.9. The number of reads generated for each sample and the percentage of reads that mapped to the *N. vitripennis* genic and intergenic regions are provided in Table IV-1. Significant differential gene expression was determined in CLC Genomics Workbench 8.5.1 at $\alpha = 0.05$ for unique gene reads using the Empirical analysis of DGE tool, which is based on the edgeR program commonly used for gene expression analyses (Robinson et al., 2010).

Table IV-1. Mapping statistics for RNA-seq of *Nasonia* ovaries

Sample	# of Reads after QC	# of Mapped Reads	% of Total Reads Mapped	# of Intergenic Gene Reads	# of Gene Reads	# of Unique Gene Reads
Nvit-1	12,622,234	11,663,493	92.40	11,166,442	497,051	464,442
Nvit-2	12,381,950	11,479,523	92.71	10,817,389	662,134	623,789
Nvit-3	10,524,703	9,758,516	92.72	9,137,620	620,896	587,458
IntG-1	11,207,434	10,327,329	92.15	8,688,965	525,186	495,981
IntG-2	9,830,279	9,107,587	92.65	8,577,889	529,698	501,814
IntG-3	10,306,862	9,550,099	92.66	9,045,192	504,907	477,929
Ngir-1	8,544,783	7,870,422	92.11	7,428,952	441,470	418,864
Ngir-2	12,457,440	11,482,452	92.17	10,785,001	697,451	661,703
Ngir-3	8,327,157	7,739,323	92.94	7,313,956	425,367	402,825

Nvit: *N. vitripennis* strain 12.1; IntG: *N. giraulti* strain IntG; Ngir: *N. giraulti* strain 16.2

RT-qPCR validation of RNA-seq results

One-day old females from *N. vitripennis* 12.1, *N. giraulti* IntG, and IntC3 were hosted with two *S. bullata* pupae and honey to encourage ovary development. After 48 hours, ovaries were removed in RNase-free PBS, flash-frozen in liquid nitrogen then stored at -80 ˚C. Five replicates of twenty ovaries per replicate were collected for each *Nasonia* strain. Total RNA was extracted from each sample using Trizol reagent (Invitrogen) with the Direct-zol RNA Miniprep kit (Zymo Research) then treated with the DNA-free DNA Removal kit (Ambion) for one hour at 37 ˚C. After ensuring with PCR that all DNA had been removed, RNA was converted to cDNA using the SuperScript VILO cDNA Synthesis kit (Invitrogen). All samples were diluted to a final cDNA concentration of 5 ng/µl in TE buffer.

RT-qPCR was performed on a CFX96 Real-Time system (Bio-Rad) using a total reaction volume of 10 µl: 5 µl of iTaq Universal SYBR Green Supermix (Bio-Rad), 2.5 µl of sterile water, 0.75 µl each of 5 µM forward and reverse primers, and 1 µl of target cDNA in single wells of a 96-well plate (Bio-Rad). All RT-qPCR reactions were performed in technical duplicates and included a melt curve analysis to check for nonspecific amplification. The 60S ribosomal protein L32 (also known as RP49) was used as an expression control. All primers for RT-qPCR are listed in Table B-2. Expression values for each candidate gene were calculated using the $\Delta\Delta C_t$ method of relative quantification (Livak and Schmittgen, 2001) with RP49 as the reference gene. Fold-change was determined by normalizing expression values to the mean expression value of *N. giraulti* IntG for each gene.

RNAi of candidate genes

To generate DNA template for dsRNA synthesis, gene-specific primers with a T7 promoter sequence on the 3’ end of each primer (Table B-3) were used to amplify a 500-700 bp region of the target gene by PCR using *N. vitripennis* whole-body cDNA as template. PCR amplicons were separated by electrophoresis on a 1% agarose gel, excised, and purified using the QIAquick Gel Extraction kit (Qiagen). The purified PCR products were used as template for a second PCR reaction with the same gene-specific T7 primers, then purified using the QIAquick PCR Purification kit (Qiagen). After quantification with the Qubit dsDNA Broad Range Assay kit (Thermo Fisher Scientific), approximately 500 – 800 ng of the purified PCR amplicon was used as template for dsRNA synthesis with the MEGAScript RNAi kit (Ambion). Each dsRNA

synthesis reaction was incubated for six hours at 37 ˚C, treated with RNase and DNase for one hour at 37 ˚C, then column-purified according to the manufacturer's protocol. To make dsRNA against GFP, the same protocol was followed except that the pGreen plasmid (Carolina) was used as template for PCR instead of *Nasonia* cDNA.

For injection, 4 ul of the dsRNA (or TE buffer as an injection control) was mixed with 1 ul of blue food coloring diluted 1:10,000 in TE buffer for a final concentration of approximately 1 ug/ul dsRNA. A Nanoject II (Drummond Scientific) was used to inject 13.8 nl of dsRNA (or buffer) into the ventral abdomen of female *Nasonia* at the yellow pupal stage. After emerging as adults, injected females were given honey and hosted individually on two *S. bullata* pupae for 48 hours. On the third day after emergence, they were transferred to new vials where they were presented with a single *S. bullata* host. After five hours, the hosts were opened and up to ten embryos were collected in a 1.5 ml Eppendorf tube for each female and stored at -80 ˚C. The females were given two hosts overnight, and then the same process was repeated again on the fourth day. On the fifth day after emergence, the abdomen of each female was removed with a razor blade, placed in an RNase-free 1.5 ml Eppendorf tube, flash-frozen in liquid nitrogen, and stored at -80 ˚C.

The number of *Wolbachia* cells per embryo from injected females three and four days post emergence was determined using qPCR with *Wolbachia groEL* primers as described above. *Wolbachia* titers were not normalized to *Nasonia* gene copy number because early embryos have varying numbers of genome copies depending on how many rounds of mitotic division they have undergone (Pultz et al., 2005). To determine the knock-down efficiency of each dsRNA injection, RNA extraction and RT-qPCR of female abdomens were performed as described above using the gene-specific QPCR primers in Table B-2.

Nuclear staining of Wolbachia *in* Nasonia *ovaries*

Female *Nasonia* were hosted on *Sarcophaga bullata* pupae for two to three days before dissection to encourage ovary development. Females were dissected in 1X phosphate-buffered saline (PBS) solution, where ovaries were removed with forceps and individual ovarioles were separated with fine needles. Ovaries were fixed in 4% formaldehyde in PBS with 0.2% Triton X-100 (PBST) for 20 minutes at room temperature then transferred to a 1.5 ml Eppendorf tube containing PBST. Ovaries were washed quickly three times with PBST then incubated in PBST

plus 1 mg/ml RNase A for three hours at room temperature then overnight at 4 ˚C. After removing the RNase A solution, ovaries were incubated at room temperature for 15 minutes in PBST with 1:300 SYTOX green nucleic acid stain (Thermo Fisher Scientific) before washing twice with PBST, 15 minutes each time. Ovaries were then transferred to a glass slide and mounted in ProLong Gold antifade solution (Thermo Fisher Scientific) and covered with a glass cover slip sealed with nail polish. Images were acquired on a Zeiss LSM 510 META inverted confocal microscope at the Vanderbilt Cell Imaging Shared Resource core.

Results

Inheritance of bacterial density trait: maternal versus zygotic effect and dominance

To determine the inheritance pattern of *w*VitA densities in *Nasonia* hybrids, we reciprocally crossed *N. vitripennis* (low-density) and *N. giraulti* IntG (high-density) individuals. Five female F1 hybrid pupae were pooled from each single-paired mating, and their *Wolbachia* densities were measured using quantitative PCR (qPCR) (Figure IV-2A). The average F1 pupal *Wolbachia* densities from pure-breeding *N. vitripennis* (N = 5) and *N. giraulti* control families (N = 5) were 0.057 ± 0.004 and 4.805 ± 1.071 (mean ± S.E.M.), respectively, which represents an 84-fold interspecific difference in *Wolbachia* titers and is consistent with previous studies (Chafee et al., 2011). Interestingly, even though F1 hybrid females from both crosses had identical genotypes (heterozygous at all loci), the average *Wolbachia* density in pupal F1 hybrid females from *N. vitripennis* mothers was 0.149 ± 0.029 (N = 10), while the average density in F1 pupae from *N. giraulti* mothers was significantly higher at 1.746 ± 0.187 (N = 10, p = 0.03, Kruskal-Wallis non-parametric test followed by a Dunn's test of multiple comparison, Figure IV-2A). This indicates that either the maternal genotype or maternal *Wolbachia* load is influencing *w*VitA titers in offspring since high-density *N. giraulti* mothers produce pupae with significantly higher densities than those observed in pupae from low-density *N. vitripennis* mothers.

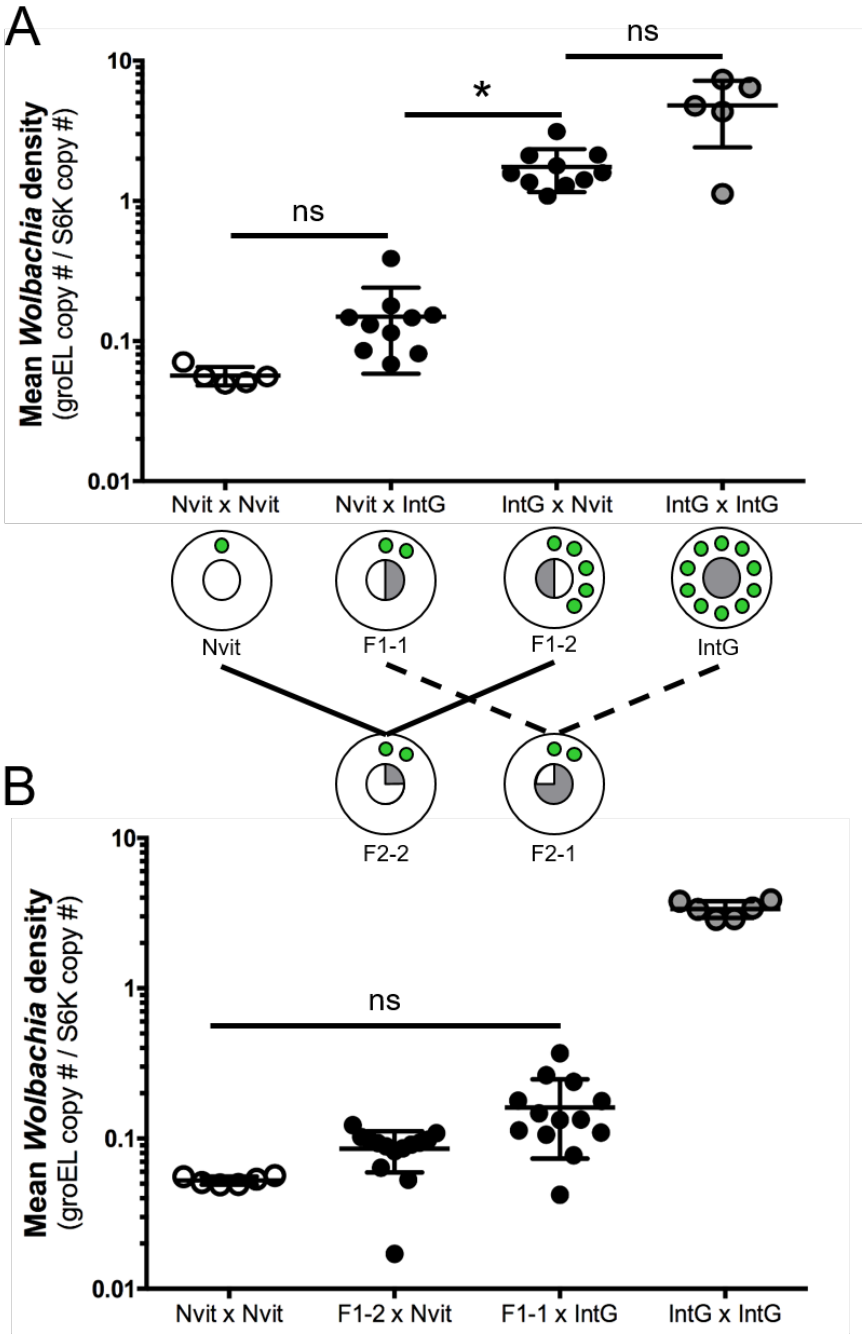

Figure IV-2. *w*VitA densities are controlled through a dominant *N. vitripennis* maternal effect
(A) *w*VitA titers in F1 pupae from crosses of *N. vitripennis* (Nvit) and *N. giraulti* IntG (female x male). (B) *w*VitA titers in F2 pupae from F1 females backcrossed to their paternal line. The fraction of white or gray in the inner circle of the diagrams in (A) and (B) indicates the average percentage of the genome that is of *N. vitripennis* (gray) or *N. giraulti* (white) origin in those pupae. The green circles (*w*VitA) in the outer circle is a representation of *w*VitA load (not drawn to scale). wVitA densities were measured using qPCR for a single-copy *Wolbachia* gene (*groEL*) normalized to a single-copy *Nasonia* gene (*NvS6K*).
*p = 0.03, Kruskal-Wallis test followed by a Dunn's test of multiple comparisons.

To test whether the difference in *Wolbachia* titers among genetically identical F1 hybrids is due to maternal *Wolbachia* load or to a partial genetic maternal effect, we backcrossed F1 females to their paternal line and pooled five female F2 pupae per F1 mother for qPCR (Figure IV-2B). If a genetic maternal effect is regulating *Wolbachia* densities, F2 pupae from both experimental lines would have similar *Wolbachia* levels since F1 hybrid mothers are genotypically identical. Indeed, the densities of F2 pupal offspring of both high- and low-density F1 mothers (F2-2 and F2-1, respectively, Figure IV-2B) were comparable, with only a 1.8-fold difference (0.161 ± 0.024, N = 13 and 0.086 ± 0.007, N = 14 respectively), supporting the hypothesis that host maternal genotype plays an important role in the regulation of *Wolbachia* densities. Furthermore, since the densities of both F2 hybrid groups were more similar to the *N. vitripennis* control (0.053 ± 0.001, N = 6) than to the *N. giraulti* control (3.364 ± 0.174, N = 6), the *N. vitripennis* low *Wolbachia* density phenotype is dominant (Figure IV-2B).

Phenotype-based selection and introgression to identify maternal suppressor genes

In an initial approach to determine the location and number of loci of major effect that regulate *w*VitA densities in *Nasonia*, we selected hybrid females that produced offspring with low *w*VitA titers and backcrossed them to high-density *N. giraulti* (IntG) males for nine generations, repeating the selection process each generation (Figure IV-3). Ideally, this introgression scheme would maintain *N. vitripennis* genomic regions that contribute to the low *Wolbachia* density trait in the genome while the rest of the genome is replaced with that of *N. giraulti*. Two independent selection lines were introgressed simultaneously to help discriminate between *N. vitripennis* regions maintained due to selection (present in both lines) and those randomly maintained through chance (present in only one line). Averages of the three lowest *Wolbachia* densities at the ninth generation for Line 1 (1.40 ± 0.07) and Line 2 (0.79 ± 0.04) were both significantly lower than the *N. giraulti* average (6.66 ± 0.71, N = 7, $p = 0.02$ for both, two-tailed Mann-Whitney U).

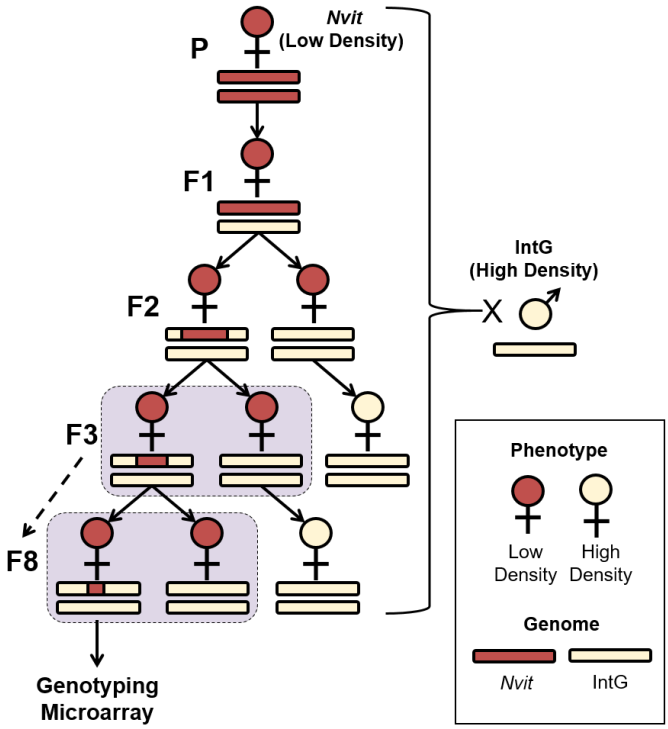

Figure IV-3. Introgression scheme using *Wolbachia* density as a selectable marker
Since *w*VitA densities are controlled through an *N. vitripennis*-dominant maternal effect, the phenotype of female hybrids (red female symbol = low, cream female symbiont = high) served as a proxy for the genotype of the mother (red bar = *N. vitripennis* origin, cream bar = *N. giraulti* origin). For the parental generation, *N. vitripennis* females (Nvit) were mated with *N. giraulti* IntG males to produce F1 hybrids. Female hybrids were then backcrossed with *N. giraulti* IntG males for nine generations. At each generation, pupal offspring were collected from each mated female, and their *w*VitA densities were measured by qPCR. The sisters of the pupae with the lowest *w*VitA densities were then chosen (purple boxes) as the mothers for the next round of mating and selection. Eighth generation females that produced ninth generation offspring with the lowest *Wolbachia* titers were genotyped on a *Nasonia* genotyping microarray.

For each independent line, DNA from the three females that produced ninth-generation offspring with the lowest *Wolbachia* densities were pooled and genotyped on a *Nasonia* genotyping microarray (Desjardins et al., 2013b) to identify the *N. vitripennis* regions that had been maintained through the entire introgression process. For each marker on the array, the proportion of *N. vitripennis* alleles was calculated based on hybridization intensity to the *N. vitripennis* probe versus the *N. giraulti* probe. A score of zero indicates that none of the females had an *N. vitripennis* allele at that marker, whereas the maximum score of 0.5 indicates that all three females were heterozygous at the locus. After nine generations of phenotype-based selection and introgression, both independent experimental lines displayed an enrichment of *N. vitripennis* alleles (proportion of *N. vitripennis* alleles $\geq$ 0.2) along the central portions of

chromosomes 2 and 3 (Figure IV-4). On the most recent *N. vitripennis* linkage map (Desjardins et al., 2013b), the area of enrichment on chromosome 2 for Line 1 occurs between 38 cM and 51.1 cM, while enrichment in Line 2 extends from 25.6 cM to 38 cM. Although overlap in *N. vitripennis* allele enrichment between Lines 1 and 2 on chromosome 2 occurs at 38 cM, the exact position and size of the overlap cannot be determined due to the fact that it falls within the poorly-assembled heterochromatic regions flanking the centromere (Desjardins et al., 2013b). For chromosome 3, the areas of enrichment for *N. vitripennis* alleles between Lines 1 and 2 mostly coincide starting at 35 cM and 34.3 cM, respectively, and ending at 47.5 cM for both lines.

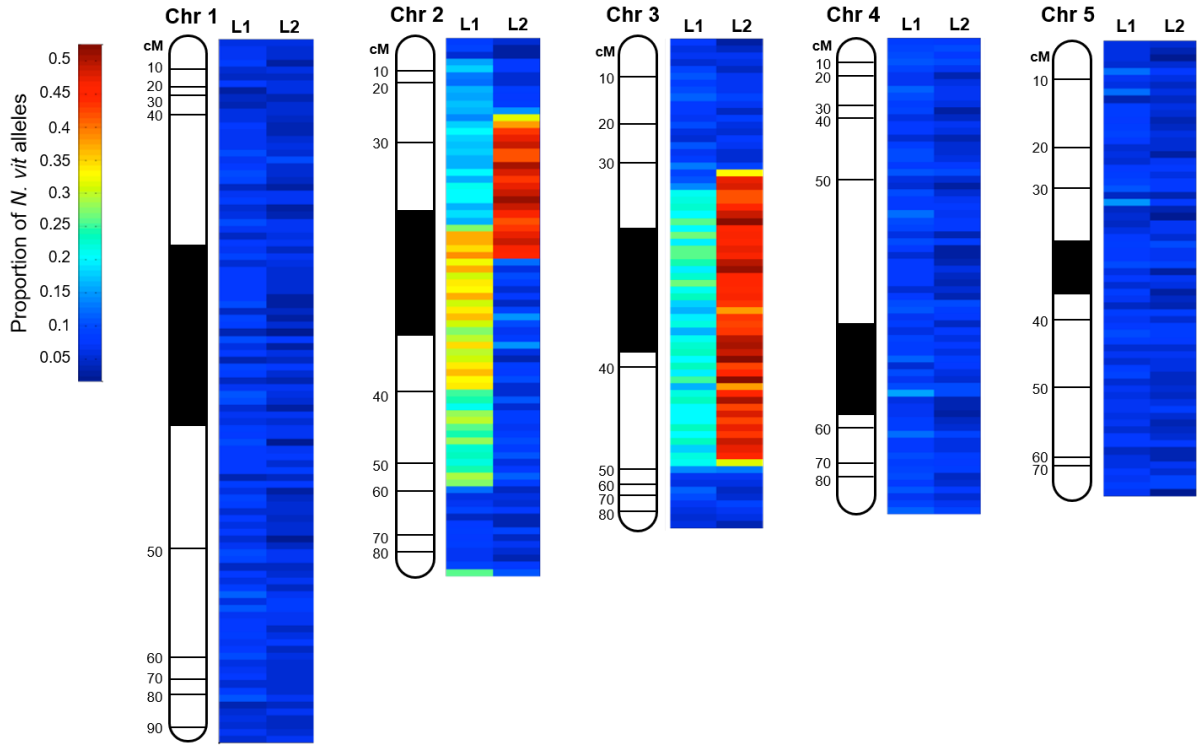

Figure IV-4. Regions of *N. vitripennis* allele enrichment on the five *Nasonia* chromosomes after selective introgression

Heatmap of the proportion of *N. vitripennis* alleles present in a pool of three females from each independent introgression line (L1 or L2) that produced offspring with the lowest *w*VitA densities after nine generations of selective backcrossing to *N. giraulti* IntG. The proportion of *N. vitripennis* alleles is based on a scale from 0 to 0.5, where 0 = no *N. vitripennis* alleles present in any of the three females and 0.5 = all females had one *N. vitripennis* allele at that marker. Each colored box represents an average of the proportion of *N. vitripennis* alleles present over 22 consecutive markers after markers were mapped back to the *Nasonia* genetic linkage map (Desjardins et al., 2013b). Black areas on the chromosome maps represent the centromeric regions.

Since chromosomes 1, 4 and 5 were not enriched for *N. vitripennis* alleles in our analysis, we conclude that chromosomes 2 and 3 are the most likely to harbor genes involved in *Wolbachia* density regulation, though chromosomes 1, 4 and 5 may harbor genes of minor effect not detected in our selection experiment.

QTL analysis and confirmation of maternal-effect suppressor regions

To confirm that our selection and introgression method accurately enriched for chromosomal regions affecting the *Wolbachia* density trait and to more precisely map those regions' chromosomal locations, we performed a quantitative trait loci (QTL) analysis in which F1 hybrid females were backcrossed to *N. giraulti* (IntG) males to obtain 191 F2 recombinant females. Assuming that F2 recombinant females with a dominant *N. vitripennis* allele at a gene important for *Wolbachia* regulation would produce offspring with low *Wolbachia* titers, each F2 female was "phenotyped" by measuring the *Wolbachia* titers in her F3 pupal offspring. Since the most informative individuals in QTL mapping are those with the most extreme phenotypes (Lander and Botstein, 1989), we selectively genotyped F2 females with the lowest (0.072 – 0.409, N = 42) and highest (2.958 – 10.674, N = 42) F3 pupal *Wolbachia* titers with a total of 47 microsatellite markers across chromosomes 1, 2 and 3 with an average distance between markers of 3 cM (Table B-1). Using genotype data for selected individuals and phenotype data for all F2 females, we identified two significant QTL peaks at a genome-wide significance level of $\alpha = 0.05$ (LOD > 2.29): one on chromosome 2 at 43 cM (LOD = 7.5, 95% Bayes credible interval of 38 cM – 50 cM) and one on chromosome 3 at 41.5 cM (LOD = 4.7, 95% Bayes credible interval of 35 cM – 61.5 cM) (Figure IV-5, Table IV-2). Interestingly, the 95% Bayes credible interval on chromosome 2 corresponds to the same region identified by the genotyping microarray as enriched for *N. vitripennis* alleles in introgression line 1 (38 cM – 51.1 cM). The 95% Bayes credible interval on chromosome 3 also contains the region on both introgression lines that was enriched for *N. vitripennis* alleles (35 cM – 47.5 cM) according to the genotyping microarray. Thus, the QTL analysis confirms that the regions identified in the selection experiment are important for *Wolbachia* regulation and predicts that the two significant QTLs act additively to explain approximately 23% of the phenotypic variance in *Wolbachia* densities between the two species.

As a negative control, we genotyped the same individuals with markers located on *Nasonia* chromosome 1, which was not enriched for *N. vitripennis* alleles after the selection introgression. The highest peak on chromosome 1 with a LOD score of 0.56 was not statistically significant, indicating that there is no correlation between *Wolbachia* density and a gene of major effect on chromosome 1, as expected.

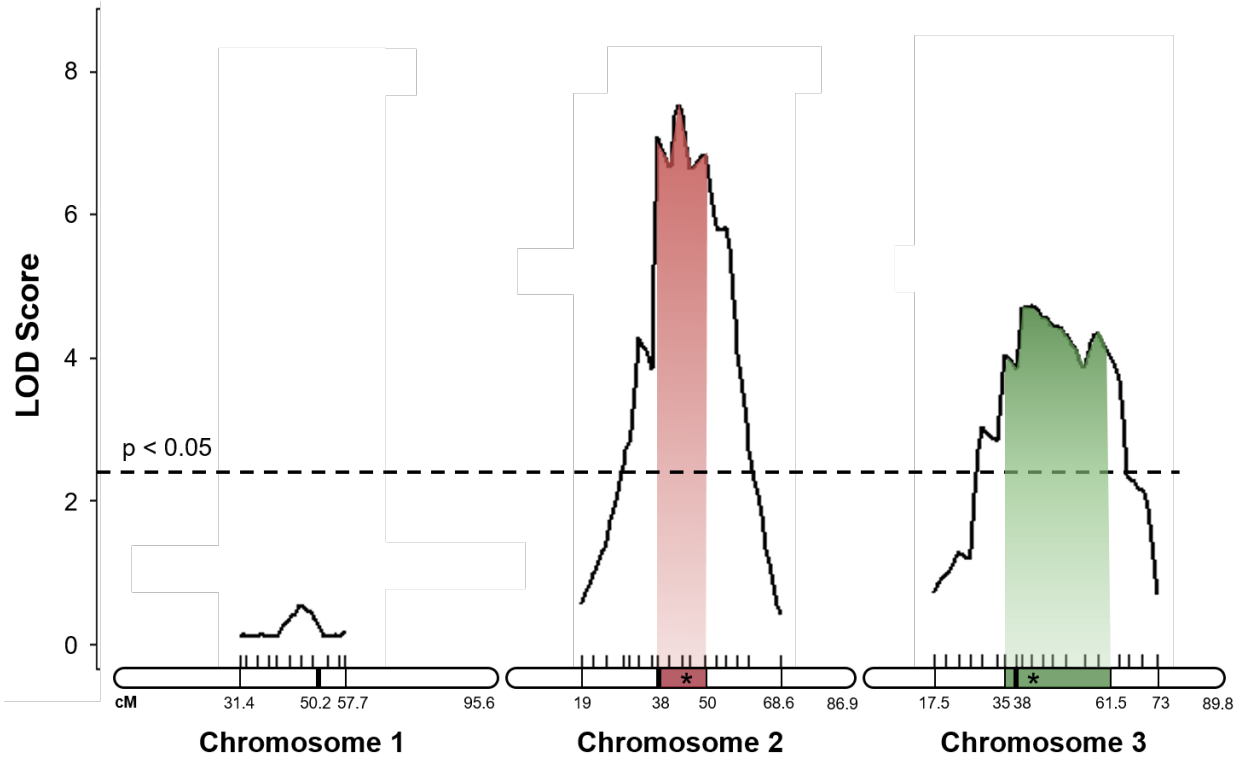

Figure IV-5. Significant QTL regions on Chromosomes 2 and 3 for the *w*VitA density trait
Plot of LOD score across all markers tested (black lines above the chromosome maps) on chromosomes 1, 2, and 3. Shaded regions represent the 95% Bayes credible interval for each significant QTL peak (star on the chromosome map). Dashed line represents genome-wide significance threshold at $\alpha = 0.05$. cM locations on chromosome maps are based on the *Nasonia* genetic linkage map (Desjardins et al., 2013b). The bolded tick marks and stars on the chromosomal maps indicate the locations of the centromeric regions and the predicted QTL peaks, respectively.

Table IV-2. Summary statistics for QTL analysis on *Wolbachia* density phenotype

QTL Location (Chr: cM)	LOD Score	p-value	95% Bayes CI (cM)	Approx. size of the CI (Mb)	Approx. # of genes in CI	Additive effect*	Variance explained (%)
2: 43	7.5	<0.001	38-50	30.1	889	1.4	14.5
3: 41.5	4.7	<0.001	35-61.5	27.3	1029	1.1	8.8

*Presence of *N. giraulti* allele increases *Wolbachia* density; LOD: Logarithm of odds; CI: credible interval

Marker-assisted introgression of maternal-effect suppressor QTLs

To verify the effect of each QTL on *Wolbachia* densities and to begin the process of fine-mapping, chromosomal regions surrounding the QTL peaks on chromosomes 2 and 3 were separately introgressed from *N. vitripennis* into an *N. giraulti* IntG background for at least nine generations using marker-assisted selection (Figure IV-6). After the ninth generation, sibling matings were performed in an attempt to produce segmental introgression lines that were homozygous *N. vitripennis* for the marker of interest. Unfortunately, hybrid incompatibilities that arose prevented us from generating *N. vitripennis* homozygous lines for some genomic regions, especially those near or around the centromere. Nevertheless, since *N. vitripennis* genes act dominantly to suppress *w*VitA titers, we tested females that are heterozygous at these regions instead since they do not exhibit hybrid incompatibilities. To determine the effect of an *N. vitripennis* allele on suppressing *w*VitA densities, we individually hosted virgin females from each line, measured the *Wolbachia* densities of their male pupae using qPCR, and calculated a percent effect on density suppression (see Materials and Methods). Females were then genotyped with microsatellite markers across the region of interest and all females with the same genotype, regardless of their original introgression line, were grouped into a common "haplotype" and their percent effects on density suppression were averaged.

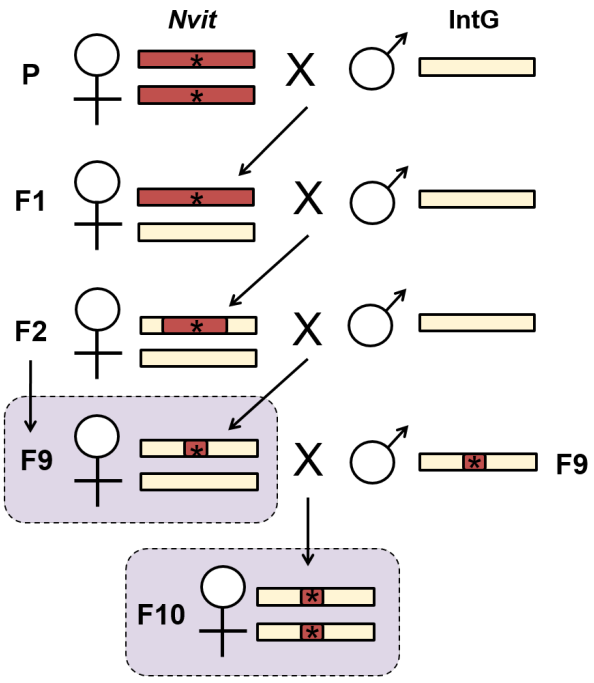

Figure IV-6. Marker-assisted introgression scheme
After an initial cross of *N. vitripennis* (Nvit) females with *N. giraulti* (IntG) males, hybrid females were backcrossed with IntG males for nine generations. At each generation, females were mated and allowed to lay offspring before being genotyped with a microsatellite marker (star) in the region being targeted for introgression (red bars). Only offspring of females that were heterozygous at the marker of interest were used in the next round of introgression. After nine generations, siblings were mated with each other to produce lines homozygous for the targeted *N. vitripennis* alleles. Since some lines could not be made homozygous, females with at least one copy of an *N. vitripennis* allele (purple boxes) were used to determine the effect of each region on *w*VitA densities.

For chromosome 2, 109 females from 22 introgression lines were genotyped with 23 markers (Figure IV-7, Table B-1) located between 32.1 cM and 49.6 cM, which includes, but is not limited to, the 95% Bayes credible interval (38 cM to 50 cM) for the chromosome 2 QTL. Females with at least one *N. vitripennis* allele at all 23 markers (haplotype C2-3, N = 4) suppressed *Wolbachia* titers in the offspring by 54.4 ± 8.8% (mean ± S.E.M., Figure IV-7) compared to *N. giraulti* (IntG) control females (N = 9), confirming the presence of a major QTL in the genotyped region. With the exception of haplotypes C2-6 and C2-7, the females with genotypes that suppressed *w*VitA titers by more than 40% (Figure IV-7, purple bars) all had at least one *N. vitripennis* allele at 38 cM near the centromere, between markers MM2.L5371 and MM2.L5543. The average percent suppression of *w*VitA densities for this region, which we will refer to as the "chromosome 2 candidate region," was 52.7 ± 2.5%. The chromosome 2 candidate region is approximately 1.8 Mb and contains approximately 137 genes. We could not determine

the exact size of the region because the markers span two scaffolds (5 and 15) in a genomic region that is poorly mapped (Desjardins et al., 2013b), so our calculations do not account for any scaffolds that may lie between scaffolds 5 and 15.

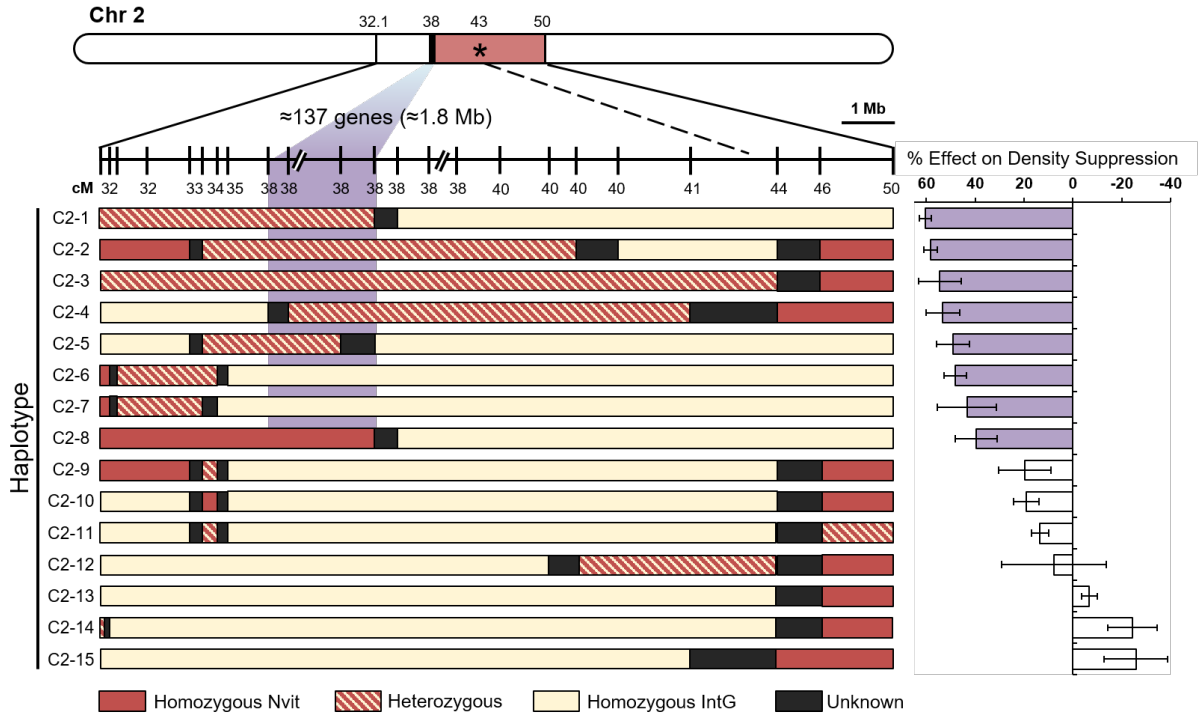

Figure IV-7. Segmental introgression haplotypes for chromosome 2 and their effects on *w*VitA density suppression

Diploid genotypes are depicted as haplotypes, where solid red bars represent *N. vitripennis* homozygous regions, dashed bars are heterozygous regions, solid cream bars are *N. giraulti* homozygous regions and black bars are recombination breakpoints between two markers. The star and red box on the chromosome map represent the QTL peak and 95% Bayes credible interval, respectively. The bar graph shows the mean percent effect on density suppression for all females with the same haplotype. Error bars are ± S.E.M. Purple area indicates the genomic region where the presence of an *N. vitripennis* allele generally corresponds with a mean % effect on density suppression over 40% (purple bars).

For chromosome 3, 86 females from 14 independent introgression lines were genotyped with 16 markers from 26.3 cM to 58.4 cM, which includes most of the 95% Bayes credible interval (35 cM to 61.5 cM) predicted for the chromosome 3 QTL (Figure IV-8, Table B-1). All haplotypes that suppressed *w*VitA titers by more than 40% (Figure IV-8, purple bars) had at least one *N. vitripennis* allele between 29.2 cM (marker MM3.22) and 37.2 cM (marker MM3.L8850). This "chromosome 3 candidate region" had an average percent suppression of 57.0 ± 2.5%, is 3.4 Mb in size, and contains 288 genes.

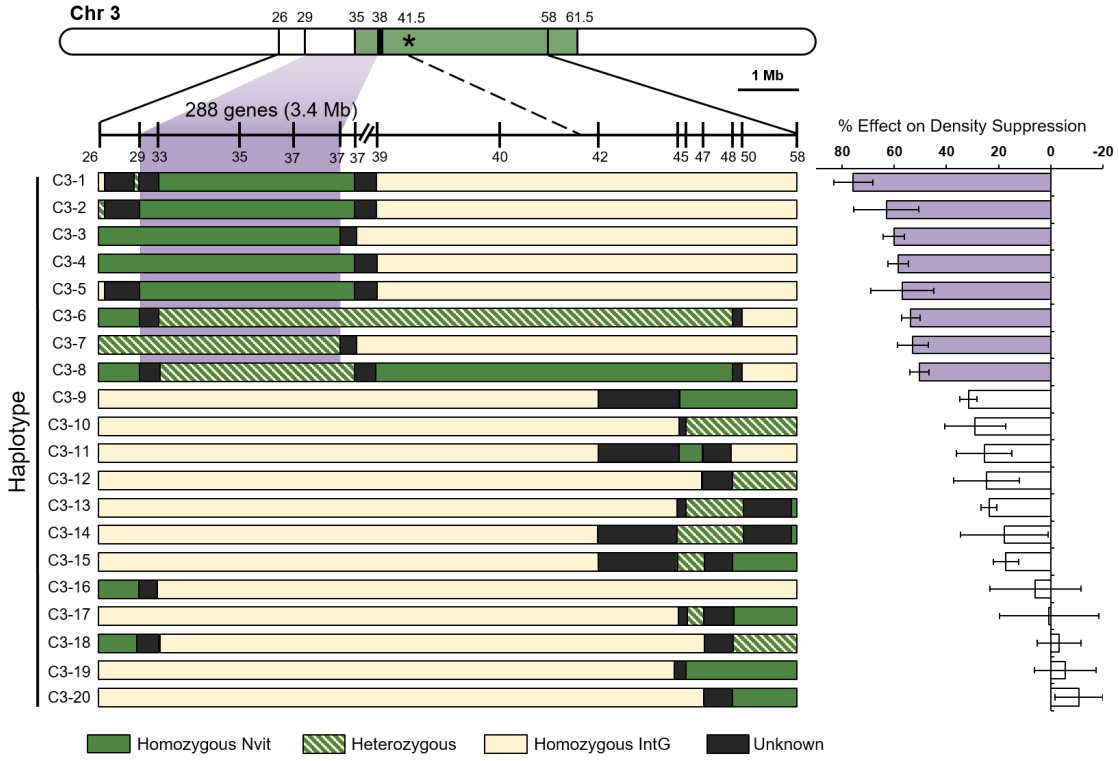

Figure IV-8. Segmental introgression haplotypes for chromosome 3 and their effects on *w*VitA density suppression

Diploid genotypes are depicted as haplotypes, where solid green bars represent *N. vitripennis* homozygous regions, dashed bars are heterozygous regions, solid cream bars are *N. giraulti* homozygous regions and black bars are recombination breakpoints between two markers. The star and green box on the chromosome map represent the QTL peak and 95% Bayes credible interval, respectively. The bar graph shows the mean percent effect on density suppression for all females with the same haplotype. Error bars are ± S.E.M. Purple area indicates the genomic region where the presence of an *N. vitripennis* allele corresponds with a mean % effect on density suppression over 40% (purple bars).

To see the interaction of the chromosome 2 and 3 candidate regions on *w*VitA density suppression, we crossed females with haplotype C2-2 with C3-3 males and compared the percent effect on density suppression in their offspring to females heterozygous for either the chromosome 2 or chromosome 3 candidate regions (Figure IV-9). Alone, the chromosome 2 and chromosome 3 candidate regions each suppressed densities by 58.2 ± 3.8% and 52.9 ± 6.2%, respectively, while females heterozygous at both candidate regions suppressed *w*VitA densities in their offspring by 83 ± 2.4%. Though we did not see a complete suppression of densities like we would expect if the genes acted strictly additively, both regions clearly had a combined effect on density suppression. This effect was significant between the chromosome 3 candidate region

alone and both regions in the same background (p = 0.004, Kruskal-Wallis followed by a Dunn's multiple comparisons test).

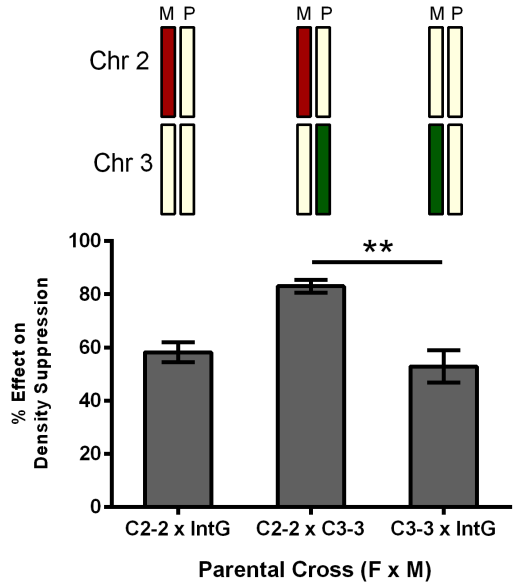

Figure IV-9. Combinatorial effect of candidate regions on *w*VitA density suppression
Individuals with haplotypes C2-2 and C3-3 were mated with either *N. giraulti* IntG or each other for the parental cross. The colored bars represent the genotype of F1 hybrid female offspring of each parental cross: red = *N. vitripennis* chromosome 2 allele, green = *N. vitripennis* chromosome 3 allele, and cream = *N. giraulti* allele; M = maternal allele and P = paternal allele. *w*VitA densities in male pupal offspring of the F1 hybrid females were measured with qPCR and a percent effect on density suppression was calculated compared to *N. vitripennis* and *N. giraulti* IntG controls. **p = 0.004, Kruskal-Wallis followed by a Dunn's multiple comparisons test.

RNA-seq of Nasonia ovaries

High-throughput RNA sequencing (RNA-seq) was performed on ovary samples from *N. vitripennis* 12.1 and *N. giraulti* IntG to identify genes within the candidate regions that are differentially expressed during oogenesis. We also sequenced the transcriptome of ovaries from *N. giraulti* 16.2 that is naturally infected with *Wolbachia* strain *w*GirA at low titers comparable to those of *w*VitA in *N. vitripennis* (Chafee et al., 2011) to analyze *w*VitA-specific effects on *N. giraulti* gene expression. Reads for all samples were mapped to the *N. vitripennis* genome Nvit_2.1 (GCF_000002325.3) with 14,321 annotated genes. A total of 9,786 genes in *N. vitripennis* and 9,764 genes in *N. giraulti* IntG had some level of expression (at least one uniquely-mapped read). Differential expression was analyzed using edgeR (Robinson et al., 2010), which identified 1,330 differentially-expressed genes with an FDR-corrected $p < 0.05$ between low-density *N. vitripennis* 12.1 and high-density *N. giraulti* IntG.

Table IV-3. Significantly differentially expressed genes

NCBI Gene ID	NCBI Gene Name	Mean Reads for Nvit	Mean Reads for IntG	EdgeR Fold Change (Nvit/IntG)	EdgeR p-value (FDR-corrected)
Chromosome 2 candidate region					
LOC100120281	Amyloid beta A4 precursor protein-binding	9.33	75.33	-9.37	9.6E-35
LOC100119653	Y + L amino acid transporter 2	86.67	159.33	-2.14	1.8E-12
LOC100120672	Protein TANC2	22.00	53.67	-2.84	9.9E-10
LOC100118571	Serine/threonine-protein phosphatase 2A	209.00	112.00	1.60	3.0E-06
LOC100120845	Band4.1-like protein 4	43.67	73.00	-1.95	1.9E-05
LOC100120971	Protein lethal(2)essential for life	0.67	7.33	-10.83	0.00072
LOC100121288	Voltage-dependent calcium channel	126.33	164.00	-1.53	0.00076
LOC100118094	Protein lethal(2)essential for life	252.67	147.00	1.47	0.0015
LOC100120822	Kin of IRRE-like protein 3	22.33	39.67	-2.08	0.0018
LOC100119051	uncharacterized	496.00	562.00	-1.34	0.0097
LOC100121425	Protein couch potato	3.67	9.67	-3.01	0.016
LOC100120755	DDRGK domain-containing protein 1	14.67	24.00	-1.91	0.021
LOC100118759	CTL-like protein 2	29.67	14.33	1.76	0.034
LOC100679688	Transcriptional repressor CTCFL-like	78.00	43.67	1.52	0.042
LOC100120500	Aldo-keto reductase-like	6.00	13.00	-2.52	0.045
Chromosome 3 candidate region					
LOC100122078	Uncharacterized	23.33	63.67	-3.17	7.6E-12
LOC100119248	Nephrin-like	157.67	431.00	-3.13	6.19E-11
LOC100121917	Trichohyalin-like	17.00	0.33	32.28	5.0E-10
LOC100121799	Uncharacterized	52.00	97.67	-2.20	5.4E-10
LOC100122001	Uncharacterized (possible Rho GTPase)	27.33	3.67	6.17	8.2E-08
LOC100117347	U4/U6.U5 small nuclear ribonucleoprotein	22.67	4.67	4.07	6.4E-06
LOC100679525	Uncharacterized	25.33	49.00	-2.25	6.8E-05
LOC100121657	Synapse-associated protein of 47 kDa	34.33	57.33	-1.96	0.00029
LOC100117496	Latrophilin Cirl	104.00	139.00	-1.57	0.00030
LOC100679322	Flocculation protein FLO11-like	49.00	72.33	-1.72	0.00049
LOC100121852	Contactin	10.33	21.67	-2.44	0.0015
LOC100121249	Tyrosine-protein phosphatase Lar	129.67	161.67	-1.46	0.0016
LOC100121400	Aryl hydrocarbon receptor translocator	110.67	138.33	-1.46	0.0024
LOC100679834	Myb-like protein 1	4.67	12.33	-3.04	0.0054
LOC100119601	Uncharacterized	3.67	10.67	-3.30	0.022
LOC100679276	Uncharacterized	103.67	59.67	1.48	0.031
LOC100119067	Rab11 family-interacting protein 4	239.00	162.67	1.25	0.039
LOC100118888	Protein lingerer	127.67	78.67	1.38	0.040

Of the 1,330 differentially-expressed genes, fifteen are located in the chromosome 2 candidate region (Table IV-3) and eighteen are located in the chromosome 3 candidate region (Table IV-3). Interestingly, only 21 genes total were significantly differentially expressed between *w*VitA-infected, high-density *N. giraulti* IntG and *w*GirA-infected, low-density *N. giraulti* 16.2 (Table C-1), indicating that *w*VitA does not induce a large change in the overall gene expression profile of *N. giraulti* ovaries. However, two genes in the candidate regions (Y+L amino acid transporter 2 and nephrin-like) are upregulated in *N. giraulti* (IntG) compared to both *N. vitripennis* (12.1) and *N. giraulti* (16.2), which could signify a specific interaction between *N. giraulti* and *w*VitA or a general *Nasonia* response to a high-density *Wolbachia* infection.

RT-qPCR was used to verify the expression differences for some of the RNA-seq candidate genes in a separate set of ovary-specific cDNA from *N. vitripennis* 12.1 and *N. giraulti* IntG. All genes tested in the chromosome 2 candidate region except for one (TANC2) were significantly differentially expressed by RT-qPCR (Figure IV-10A). Though none of the genes achieved the same magnitude of differential expression that was observed by RNA-seq (Figure IV-10A, gray bars), all but the calcium channel showed the same expression trend (either upregulated or downregulated in *N. vitripennis*) as the RNA-seq data. The same was not true for some of the genes tested from the chromosome 3 candidate region (Figure IV-10B). For example, the uncharacterized gene that is a possible Rho GTPase (LOC100122001) was 6.17X upregulated in the RNA-seq data but was found to be significantly downregulated in *N. vitripennis* by RT-qPCR ($p = 0.032$, Mann-Whitney U test). Three of the genes, latrophilin, FLO11 and lar phosphatase showed no significant differences in expression by RT-qPCR ($p = 0.31$, $p = 0.095$, and $p = 0.222$, respectively, individual Mann-Whitney U tests). Conversely, trichohyalin was 65.2X higher expressed in *N. vitripennis* than *N. giraulti* IntG as measured by RT-qPCR, which is double the 32.3X upregulation estimated by RNA-seq.

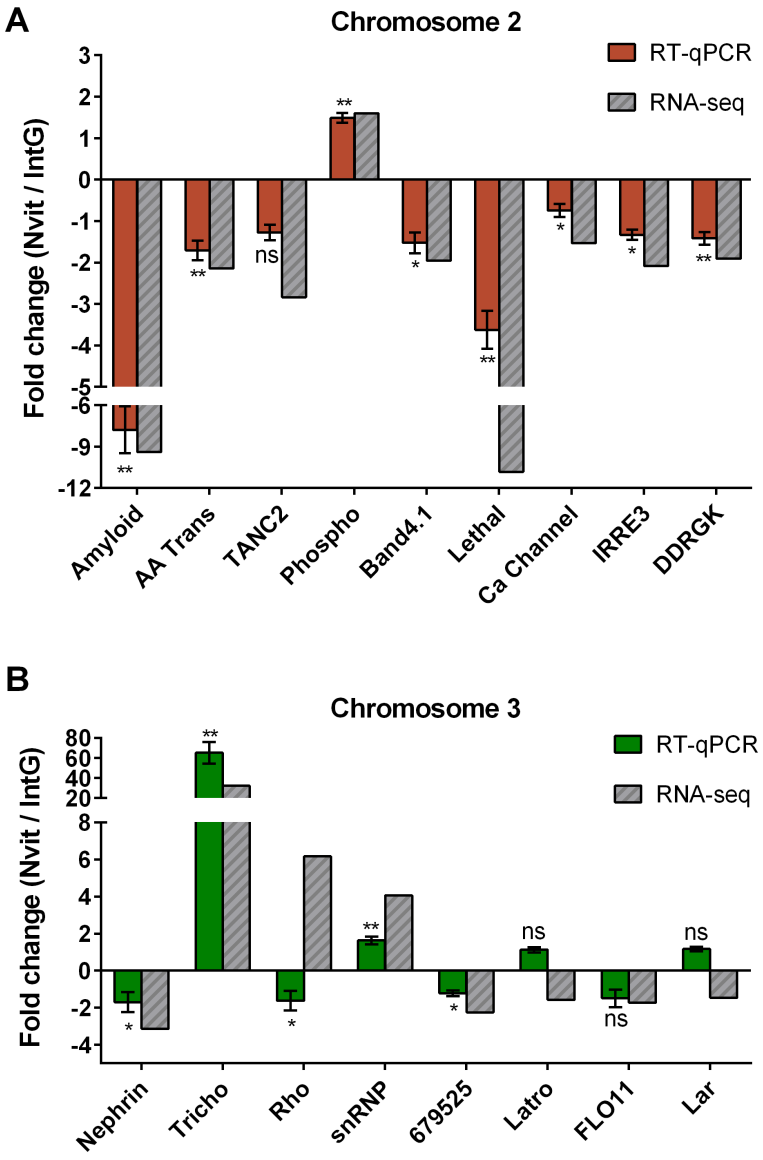

Figure IV-10. RT-qPCR validation of RNA-seq expression differences
Bars represent the average fold change of *N. vitripennis* ovarian gene expression compared to *N. giraulti* IntG expression (positive values = upregulated in Nvit, negative values = downregulated in Nvit) for (A) chromosome 2 candidate genes and (B) chromosome 3 candidate genes. Colored bars are values from RT-qPCR; dashed, gray bars are from RNA-seq. Errors bars are mean ± s.d. *p< 0.05, **p<0.01, Mann-Whitney U test between Nvit and IntG RT-qPCR expression values for each gene.

RNAi of trichohyalin

From the RNA-seq and RT-qPCR analyses, trichohyalin (LOC103317433) stood out as a strong candidate gene given its 65.2X higher expression levels in *N. vitripennis* than in *N. giraulti* IntG (Figure IV-10B). To see whether the *N. vitripennis* allele of trichohyalin remains overexpressed when present in an *N. giraulti* genomic background, we used RT-qPCR to

measure trichohyalin levels in the ovaries of a segmental introgression line (IntC3) that is homozygous *N. vitripennis* for the chromosome 3 candidate region (haplotype C3-1). We found that the mean fold change of trichohyalin in IntC3 ovaries is lower than for a pure *N. vitripennis* line, but that the gene is still 42.1X upregulated compared to expression in *N. giraulti* IntG ($p = 0.008$, Mann-Whitney U test, Figure IV-11A). In addition to gene expression differences, the *N. giraulti* allele of trichohyalin appears to be undergoing pseudogenization due to several large deletions (up to 177 bp) and a frameshift mutation that introduces a premature stop codon approximately halfway through the protein at amino acid 297 (out of 651 total) (Figure IV-11B). All mutations and indels were confirmed with PCR and Sanger sequencing. Thus, *N. giraulti* likely does not produce functional trichohyalin protein, which, by default, would allow the *N. vitripennis* allele to act dominantly to suppress titers in hybrids

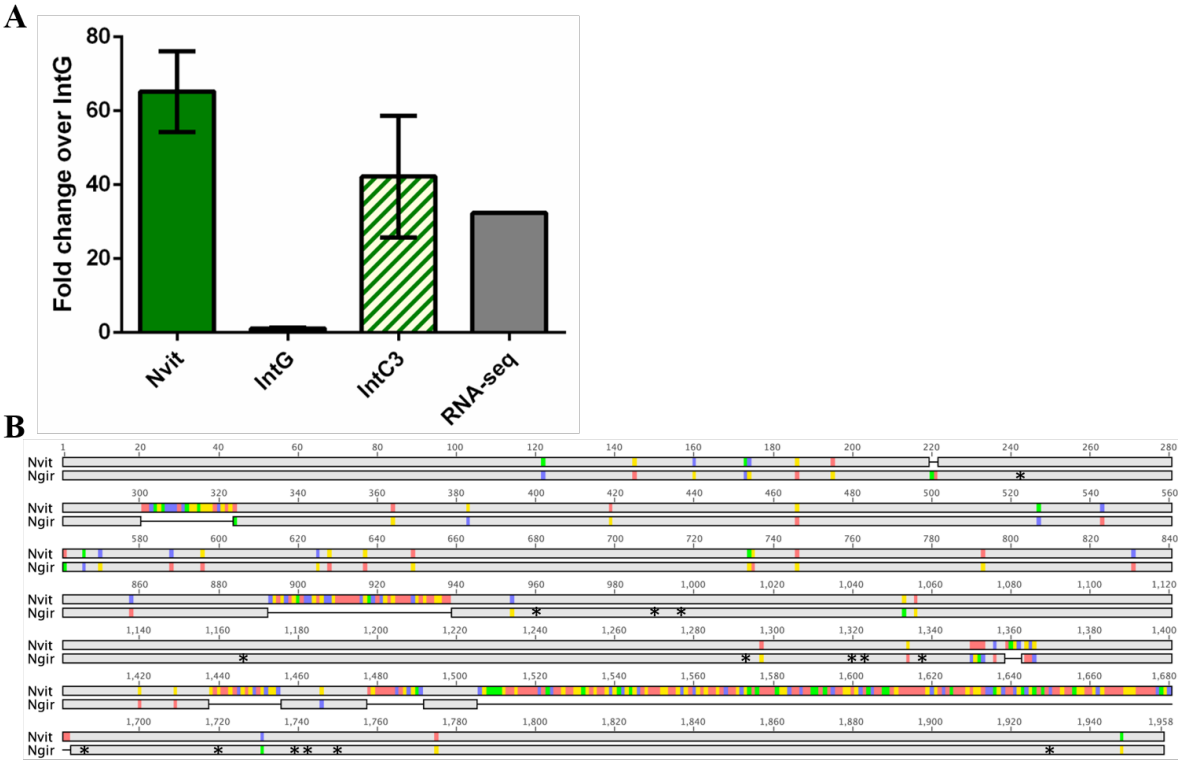

Figure IV-11. Interspecific differences in trichohyalin expression and coding sequence in *Nasonia*
(A) Mean fold-change expression of trichohyalin in ovaries of *N. vitripennis*, *N. giraulti* IntG and chromosome 3 segmental introgression line IntC3. Gray bar is the fold-change of *N. vitripennis* over IntG for RNA-seq. Error bars are mean ± s.d. (B) Alignment of the *N. vitripennis* (top) and *N. giraulti* (bottom) mRNA nucleotide sequences for trichohyalin. Colored regions indicate nucleotide changes, whereas gaps indicate nucleotide deletions. Black stars denote premature stop codons.

One of the benefits of the *Nasonia* model system is its amenability to gene knockdown using RNA interference (RNAi) (Lynch and Desplan, 2006; Werren et al., 2009). Since transgenic engineering of *Nasonia* is still in its infancy (Lynch, 2015), RNAi is currently the most accessible method to study gene function in this model system. Furthermore, several studies have successfully used parental RNAi in *Nasonia* to examine the effects of maternal genes on offspring development (Lynch et al., 2006a; Lynch and Desplan, 2010; Lynch et al., 2006b; Ozuak et al., 2014a, b; Verhulst et al., 2010). We used parental RNAi in the IntC3 hybrid line to test whether trichohyalin can regulate *w*VitA titers. The IntC3 line was chosen for this parental RNAi experiment so that we could target the dominant allele from *N. vitripennis* in a line with high enough titers to easily detect changes in *w*VitA levels by qPCR. If trichohyalin is responsible for regulating *Wolbachia* titers, then knocking down the *N. vitripennis* allele of trichohyalin in the IntC3 line should increase the number of *w*VitA in resulting offspring. In all, there were no significant differences in *w*VitA densities in embryos from Tricho-RNAi females than in embryos from GFP-RNAi or non-injected females (Figure IV-12A, $p = 0.2446$, Kruskal-Wallis test). However, embryos from Tricho-RNAi females did have the highest average number of *Wolbachia* per embryo (199.3 ± 133.5, mean ± s.d.) compared to either GFP-RNAi (129.6 ± 99.7) or non-injected females (168.2 ± 86.2). While we did see a significant reduction in trichohyalin expression in Tricho-RNAi females compared to GFP-RNAi females (Figure IV-12B, $p = 0.014$, 0.029, Mann-Whitney U test), the 55% knockdown efficiency achieved in this experiment may still be too low to produce a measurable effect on *w*VitA titer transmission.

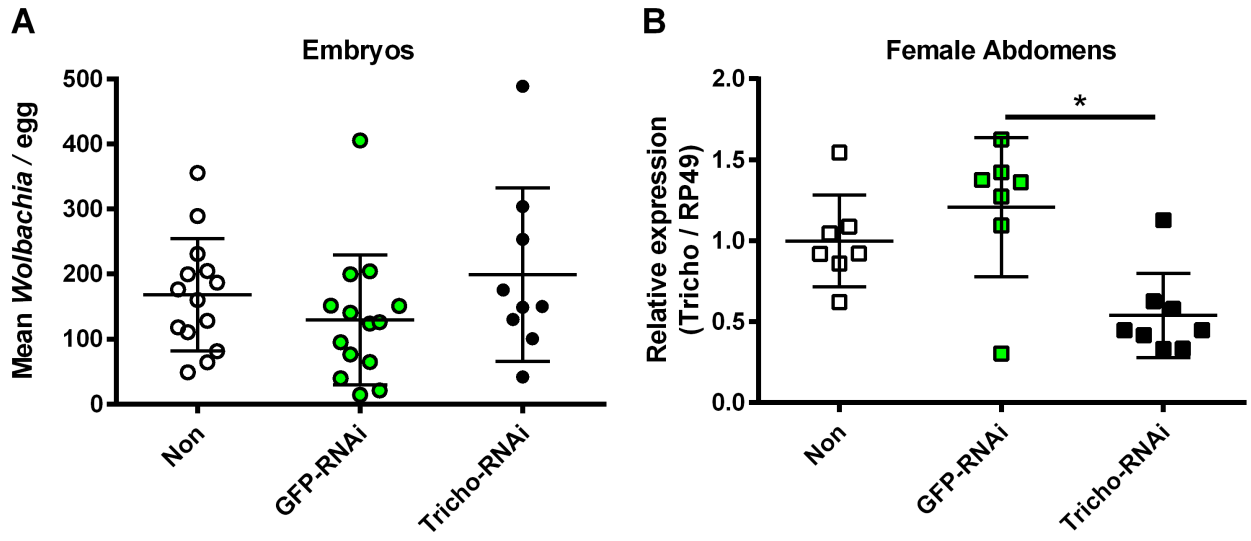

Figure IV-12. Effect of trichohyalin RNAi knockdown on *w*VitA densities and gene expression
(A) Mean number ± s.d. of *Wolbachia* per embryo or (B) relative expression of trichohyalin in abdomens from IntC3 females that were either uninjected (white), injected with dsRNA against GFP (green) or injected with dsRNA against trichohyalin (black). *p<0.05, Mann-Whitney U test

Distribution of wVitA during oogenesis

Since *w*VitA densities are controlled maternally (Figure IV-2) and disparities in *w*VitA titers are present even in early embryos (Figure I-3) (Chafee et al., 2011), maternal regulation of *w*VitA levels likely occurs sometime during the five stages of *Nasonia* oogenesis (King and Richards, 1969) (Figure IV-13A). In the first stage, germ-line stem cells produce daughter cells, which then undergo mitosis with incomplete cytokinesis to form an egg chamber with 16 interconnected cells. One of the cells becomes the oocyte and the other fifteen function as nurse cells that synthesize nutrients, proteins and maternal RNA for the oocyte. The entire egg chamber is covered with a layer of somatic follicle cells. By stage 2, the oocyte and nurse cells are distinguishable entities, with the oocyte surrounded by a visible ring of follicle cells (Figure IV-13A). At this stage, the oocyte is smaller than the group of nurse cells, but by stage 3 the oocyte has grown to the same size or larger than the nurse cells (Figure IV-13A). In stage 4, the nurse cells degenerate and dump their cytoplasmic contents into the oocyte, while the follicle cells secrete a vitelline membrane (Figure IV-13A). Finally, in stage 5 the oocyte is surrounded by the chorion and ready to be fertilized (Figure IV-13A) (King and Richards, 1969).

To see if there were any differences in *w*VitA localization during oogenesis between *N. vitripennis* and *N. giraulti* IntG, we fixed and stained ovarioles from each species with either

SYTO-11 or SYTOX green nucleic acid dye, both of which bind to host and *Wolbachia* DNA. In late stage 2 (Figure IV-13B,C) and early stage 3 (Figure IV-13D-F) egg chambers, *w*VitA is present in both the nurse cells and oocytes from *w*VitA-infected *N. vitripennis* (Figure IV-13B,E) but localized almost exclusively to the oocytes in *N. giraulti* IntG egg chambers at the same stages (Figure IV-13C,F). If this pattern remains consistent through the rest of oogenesis, then it could lead to *N. giraulti* embryos having a higher infection titer than *N. vitripennis* embryos. Unfortunately, fluorescent signals from host nuclei mask any signal emitted from the much smaller *Wolbachia* cells in the closely-packed stage 4 nurse cells, making it difficult to determine what percentage of *w*VitA cells remain in the nurse cells at the end of oogenesis. Future immunofluorescent staining of *w*VitA will hopefully resolve this issue by allowing us to visualize *w*VitA separately from host nuclei.

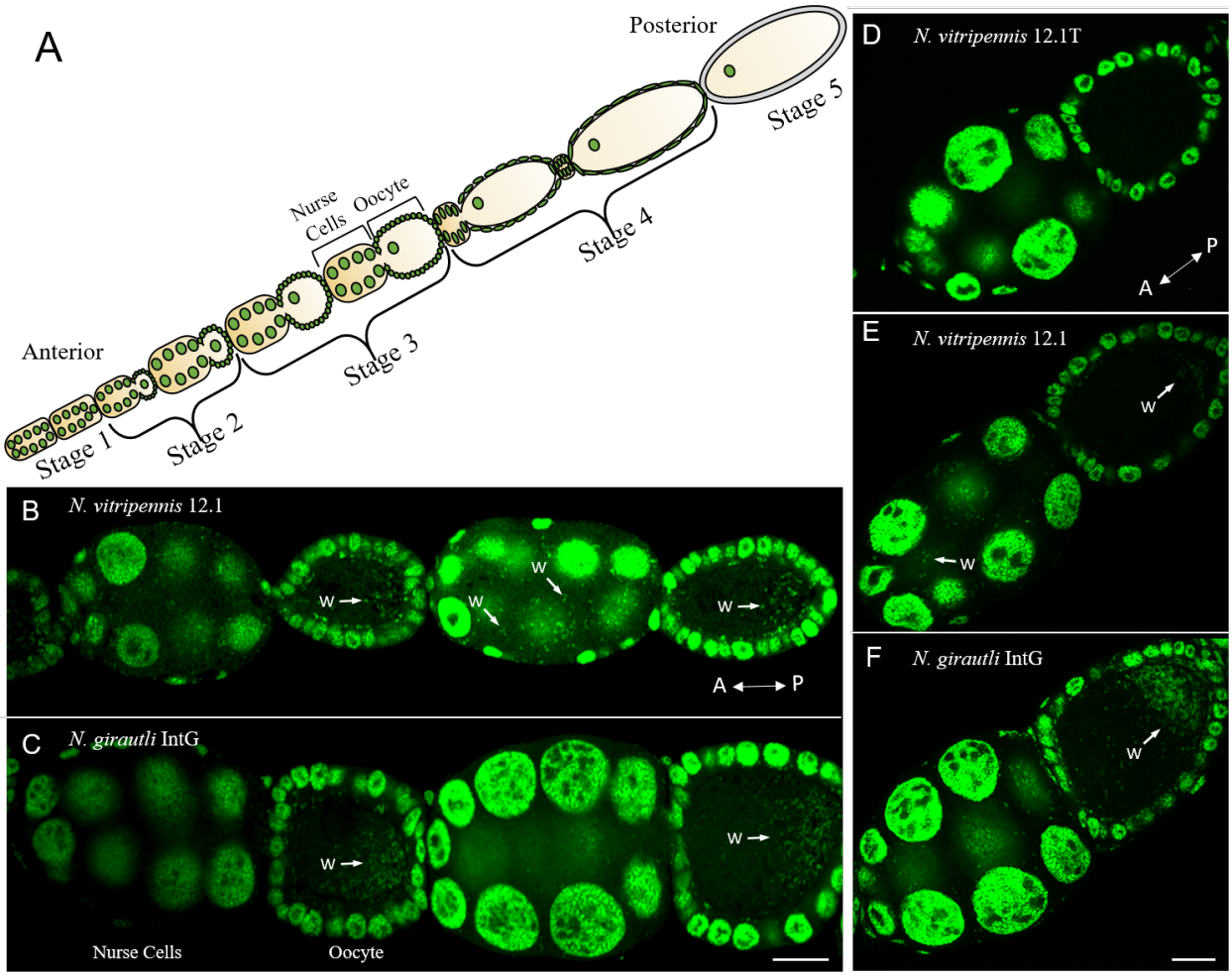

Figure IV-13. *w*VitA localization during *Nasonia* oogenesis

(A) Cartoon of a *Nasonia* ovariole depicting all five stages of oogenesis. Late stage 2 egg chambers of (B) *N. vitripennis* and (C) *N. girault*i IntG were stained with SYTO-11 nucleic acid stain. Early stage 3 egg chambers of (D) uninfected *N. vitripennis,* (E) *w*VitA-infected *N. vitripennis* and (F) *N. girault*i IntG were stained with SYTOX green nucleic acid stain. Arrows labeled with a "W" point to a population of *Wolbachia* cells. Double-headed arrows indicate the direction of the anterior-posterior axis. All scale bars = 15 μm.

RNAi knockdown of kinesin-A

The different localization patterns of *w*VitA in ovaries of *N. vitripennis* and *N. girault*i (Figure IV-13) could be established in part by differences in *w*VitA trafficking to the oocyte from the nurse cells. In *Drosophila*, the motor protein kinesin-1 is important for *Wolbachia* trafficking to the posterior pole of a developing oocyte (Serbus and Sullivan, 2007). *Nasonia* kinesin-A is located in the candidate region on chromosome 3 and, though it is not differentially expressed between *N. vitripennis* and *N. girault*i in the ovaries, the protein homologs differ at 26 amino acid sites (out of 1,375 total). If any of these amino acid substitutions result in a functional

change in the speed of kinesin-A or in its binding affinity to *w*VitA or to host cargo that *w*VitA interacts with, then this could result in the different localization patterns of *w*VitA in the egg chamber between *N. vitripennis* and *N. giraulti*. Thus, we used parental RNAi to knock down kinesin-A (LOC100115522) in the IntC3 introgression line. In the first experiment, there was a significant increase in the mean number of *Wolbachia* cells per egg in embryos from kinesin-RNAi females over those from non-injected females (Figure IV-14A, p = 0.025, Mann-Whitney U test) even when knock-down efficiency of dsRNA against kinesin-A was only 17.8% (Figure IV-14B). However, when this experiment was repeated with a larger sample size and a dsRNA control against a non-*Nasonia* gene (GFP), there were no significant differences between any of the injection groups (Figure IV-14A, p = 0.318, Kruskal-Wallis test). Similarly, in a third experiment that included an injection control (buffer-only) and a GFP-dsRNA control, there were no changes in *w*VitA levels in embryos from kinesin-RNAi females (Figure IV-14A, p = 0.3941, Kruskal-Wallis test). Although the knock-down efficiencies of the kinesin dsRNAs in experiments two and three have not yet been calculated, there is currently no definitive evidence that kinesin-A regulates *w*VitA titers in *Nasonia* embryos.

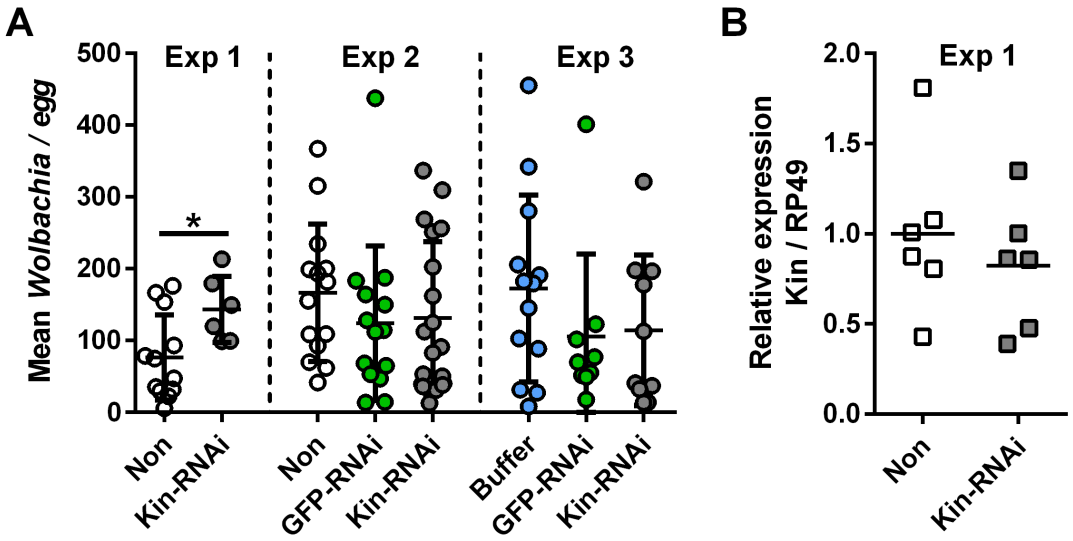

Figure IV-14. Effect of kinesin-A RNAi on *w*VitA densities and gene expression
(A) Mean number of *Wolbachia* per embryo or (B) relative expression of kinesin-A in abdomens from IntC3 females that were either uninjected (white), injected with buffer only (blue), injected with dsRNA against GFP (green) or injected with dsRNA against kinesin-A (gray). Data in (A) represents three independent experiments while data in (B) was generated from females from the first experiment only. *p<0.05, Mann-Whitney U test

Discussion

We have shown that host regulation of maternally-transmitted symbiont densities in *Nasonia* likely has a fairly simple genetic basis, with two major QTL regions explaining approximately 82% of the *Wolbachia* density disparity that arises when *w*VitA is transferred from its native *N. vitripennis* host to the closely-related but naïve host *N. giraulti*. Furthermore, this regulation occurs through a maternal effect, indicating that, at least in this symbiotic system, *Wolbachia* densities are primarily established during oogenesis or early embryogenesis before the maternal to zygotic transition. These data are corroborated by previous work that showed that *w*VitA-infected *N. giraulti* embryos already contain a higher *Wolbachia* titer than *N. vitripennis* embryos at the same stage (Figure I-3) (Chafee et al., 2011). Logically, this could be a result of a self-perpetuating cycle where female *Nasonia* with high *Wolbachia* densities in their ovaries incorporate more *Wolbachia* cells into their oocytes, producing offspring with high *Wolbachia* densities. However, our analyses show that F1 hybrid females from *N. giraulti* mothers (F1-2, Figure IV-2A) have significantly higher titers than those from *N. vitripennis* mothers (F1-1, Figure IV-2A), yet produce offspring with low titers (F2-1, Figure IV-2B). Thus, high maternal *Wolbachia* load is not automatically passed to the next generation but is a host-regulated process. This regulation is specific to *w*VitA since a second *Wolbachia* strain from *N. vitripennis*, *w*VitB, maintains low densities when transferred to *N. giraulti* (Figure I-2) (Chafee et al., 2011). *N. giraulti* is also naturally infected with its own strain of A *Wolbachia*, *w*GirA, at low titers similar to those of *w*VitA in *N. vitripennis* (Figure I-2) (Chafee et al., 2011). Interestingly, only 21 genes were significantly differentially regulated in the ovaries of high-titer, *w*VitA-infected *N. giraulti* compared to *N. giraulti* infected at low titers by *w*GirA (Table C-1), while 1330 genes were differentially regulated between *N. vitripennis* and *N. giraulti* infected with the same *Wolbachia* strain (*w*VitA). This suggests that, at least in the ovaries, gene expression differences are driven primarily by species-specific host genomic changes rather than as a direct response to *Wolbachia.*

While the resolution of our genetic mapping does not yet allow us to definitively identify the genes responsible for controlling *w*VitA levels, we propose several possible mechanisms for how the density difference may be established during oogenesis and discuss promising candidate genes for each scenario:

Differences in w*VitA trafficking could establish higher titers in* N. girault*i oocytes*

Although *Nasonia* and *Drosophila* are not closely-related, *Nasonia* oogenesis is remarkably similar to that of the well-characterized *Drosophila* system (King and Richards, 1969). The ovaries of both insects are comprised of individual ovarioles that continuously produce a series of egg chambers, budding new egg chambers at the anterior end of the ovariole while pushing older, more mature eggs toward the posterior (Figure IV-13A). In the most anterior part of the ovariole, germ-line stem cells within a germ-line stem cell niche (GSCN) produce a cytoblast that undergoes four rounds of mitosis with incomplete cytokinesis to produce a cyst with 16 interconnected cells (one oocyte and 15 nurse cells). As the cyst moves posteriorly, it passes by the somatic stem cell niche (SSCN) and is surrounded by somatic follicle cells to produce a complete egg chamber (Bastock and St Johnston, 2008; King and Richards, 1969).

Nuclear staining of stage 2 and 3 egg chambers revealed that *w*VitA appears to be efficiently shuttled to the oocyte in *N. giraulti* IntG but is present in both the nurse cells and the oocyte in *N. vitripennis* (Figure IV-13). Although all *Wolbachia* cells are dumped into the oocyte at the end of oogenesis in *Drosophila* (Ferree et al., 2005), unpublished work by Patrick Ferree (Claremont College) indicates that, in *Nasonia*, *Wolbachia* form a tight ring around the nurse cell nuclei during cytoplasmic dumping and are not transferred to the oocyte (P. Ferree, *personal communication*). Thus, though technical limitations of nuclear staining did not allow us to visualize *w*VitA distribution past stage 3, disparities in embryonic *w*VitA levels between *N. vitripennis* and *N. giraulti* IntG could be established during oogenesis if *N. vitripennis* is better at sequestering *w*VitA cells in the nurse cells, thereby limiting the number of *w*VitA cells that enter the oocyte. *N. giraulti* may lack this control mechanism, allowing *w*VitA free access to the developing oocyte.

In *Drosophila*, *Wolbachia* utilize cytoskeletal networks and motor proteins like dynein and kinesin to move around the egg chamber during oogenesis (Ferree et al., 2005; Serbus and Sullivan, 2007). In our study, knockdown of kinesin-A by parental RNAi in the IntC3 introgression line produced a significant increase in the average number of *w*VitA cells per embryo compared to non-injected controls. However, these results were not replicated in two subsequent experiments with additional control groups (Figure IV-14), so it is unclear whether kinesin plays a role in *w*VitA regulation in *Nasonia.* Another candidate gene located in the

chromosome 3 candidate region that could affect the microtubule network used by motor proteins like kinesin is the gene for tubulin-specific chaperone E (LOC100121708), which helps assemble α and β tubulin subunits into microtubules (Tian and Cowan, 2013). Tubulin-specific chaperone E is upregulated two-fold in *N. vitripennis* ovaries compared to *N. giraulti* IntG ovaries, which could be affecting the microtubule network in a way that prevents *w*VitA from moving efficiently between the nurse cells into the oocyte in *N. vitripennis*.

If *Wolbachia* hitchhikes on host RNA or protein cargo to move around the egg chamber instead of directly interacting with motor proteins, then regulation of *w*VitA localization could depend on its interaction with maternal RNAs and proteins that are also shuttled between the nurse cells and the oocyte. For example, disruption of kinesin-mediated trafficking of *oskar* maternal RNA to the pole plasm in *Drosophila* results in a loss of efficient *w*Mel posterior localization in the oocyte (Serbus and Sullivan, 2007). Reduction of another important maternal RNA, *gurken*, causes a microtubule-independent decrease in *w*Mel titers in both nurse cells and the oocyte (Serbus et al., 2011). Though the *oskar* and *gurken* genes are not located in our *Nasonia* candidate regions, the chromosome 3 candidate region does contain the gene for heat shock protein 83 (hsp83, LOC100117412). Maternal *hsp83* RNA in *Drosophila* is tightly regulated during oogenesis, where it is localized to the posterior pole of the oocyte along with other maternal RNAs like *oskar* (Ding et al., 1993). Mislocalization of *oskar* to the anterior pole leads to anterior localization of *hsp83*, suggesting that *hsp83* may be associated with *oskar* or other maternal factors in the pole plasm (Ding et al., 1993). If *w*VitA hitchhikes on *hsp83* maternal RNA or an associated protein complex to move between cells in the egg chamber, perhaps a localization difference of *hsp83* RNA in *N. vitripennis* versus *N. giraulti* ovaries dictates the distribution of *w*VitA during oogenesis.

Another candidate gene in the chromosome 3 region, the receptor-like tyrosine phosphatase Lar (LOC100121249), is required in somatic follicle cells around the oocyte to promote proper localization of *oskar* in *Drosophila* (Frydman and Spradling, 2001). Our study found that Lar was 1.5X upregulated in *N. giraulti* IntG ovaries (Table IV-3) by RNA-seq, though there was no significant difference in Lar expression by RT-qPCR (Figure IV-10B). Nevertheless, upregulation of Lar expression or signaling in the follicle cells of *N. giraulti* may help recruit more *w*VitA into the oocyte if *w*VitA is hitchhiking on maternal RNAs like *oskar*.

N. vitripennis *may prevent* w*VitA from moving into the oocyte from the follicle cells*

Certain strains of *Wolbachia* in *Drosophila* localize to the germ-line stem cell niche (GSCN), are incorporated directly into the germ-line stem cells, and are subsequently transmitted to all germ cells produced (Fast et al., 2011; Toomey et al., 2013). However, *Wolbachia* in general preferentially target the somatic stem cell niche (SSCN) (Frydman et al., 2006; Hosokawa et al., 2010; Toomey et al., 2013), including strains that also infect the GSCN (Toomey et al., 2013). In the SSCN, *Wolbachia* have direct access to germ cells as they pass by the SSCN, as well as indirect access when incorporated into the somatic follicle cells that surround the oocyte (Toomey et al., 2013). *Wolbachia* transport from follicle cells into the oocyte could function as a point of maternal regulation since intracellular *Wolbachia* likely utilize host pathways to cross cellular borders, though how *Wolbachia* travel between cells on a molecular level is not well understood. *Wolbachia* cells are surrounded by a eukaryotic membrane (Louis and Nigro, 1989), potentially of Golgi origin (Cho et al., 2011), that could facilitate transfer to the plasma membrane followed by exocytosis. Entry into the oocyte could then be accomplished through receptor-mediated endocytosis, which is used by many intracellular bacterial pathogens including *Rickettsia conorii* (Chan et al., 2009) and *Chlamydia trachomatis* (Hybiske and Stephens, 2007) to gain access to host cells. In fact, the vertically-transmitted endosymbiont *Spiroplasma poulsonii* invades *D. melanogaster* oocytes by interacting with yolk proteins that bind to the vitellogenin receptor (Yolkless in *Drosophila*) on the surface of the oocyte and are subsequently endocytosed (Herren et al., 2013). Other intracellular symbionts including *Buchnera* in aphids (Koga et al., 2012) and a yeast-like symbiont in brown planthoppers (Yukuhiro et al., 2014) are incorporated into oocytes through endocytosis.

If *Wolbachia* uses a similar mechanism to move from the follicle cells to the oocyte, then the interspecific difference in *w*VitA titers could be established if the rate of *w*VitA entry into the oocyte differed between *N. vitripennis* and *N. giraulti*. For example, *w*VitA may have a lower binding affinity to a host receptor in *N. vitripennis* than in *N. giraulti*, *N. vitripennis* may express less of the receptor on the surface of its cells, or endocytosis of the receptor in *N. vitripennis* may not be as efficient as in *N. giraulti*. An interesting membrane-associated protein located in the candidate region on chromosome 3 is contactin (LOC100121852), an extracellular GPI-anchored cell adhesion molecule that is 2.4-fold down-regulated in *N. vitripennis* than in *N. giraulti* (Table IV-3). A human homolog of contactin was shown to be important for *Streptococcus pneumoniae*

adhesion in a lung adenocarcinoma cell line (Muchnik et al., 2013). Another transmembrane protein on chromosome 3, nephrin-like (LOC100119248), is 3.1-fold down-regulated in *N. vitripennis* compared to *N. giraulti* IntG (Table IV-3). Nephrin is also one of only 21 genes that are differentially expressed between *w*VitA-infected and *w*GirA-infected *N. giraulti* lines. Since nephrin is 1.9-fold downregulated in low-density *w*GirA-infected *N. giraulti* compared to high-density *N. giraulti* IntG (Table C-1), its lower expression correlates with a decrease in *Wolbachia* titers in both *N. vitripennis* and *N. giraulti*. Thus, if either contactin or nephrin function as the receptor for *Wolbachia* binding on the surface of the germ cells, then their decreased expression could be limiting *w*VitA's access to the oocyte.

w*VitA may proliferate faster in* N. giraulti *oocytes*

As obligate, intracellular bacteria, *Wolbachia* depend on the host cell for many of its metabolic and nutritional needs. For example, *Wolbachia* likely obtain much of their energy from host amino acids (Caragata et al., 2013b; Wu et al., 2004) and have been shown to compete for host cholesterol (Caragata et al., 2013a; Caragata et al., 2013b). Thus, *N. giraulti* may naturally provide more of these nutrients to *w*VitA, resulting in increased proliferation and higher *w*VitA densities. For example, one of the chromosome 2 candidate genes is an amino acid transporter (LOC100119653) that is upregulated 2.1-fold in *N. giraulti* IntG compared to *N. vitripennis* (Table IV-3), and is 1.5X higher in *N. giraulti* IntG than in *w*GirA-infected *N. giraulti* 16.2 (Table C-1). Since *Wolbachia* are surrounded by host membranes, they may need host amino acid transporters on those membranes to access their food source (cytoplasmic amino acids). Upregulation of an amino acid transporter in *w*VitA-infected *N. giraulti* may result in a greater availability of host amino acids for *w*VitA, leading to faster proliferation of *w*VitA if the extra energy is allocated to replication.

Replication rates of *Wolbachia* during oogenesis may also depend on the location of *Wolbachia* in the egg chamber. In *D. melanogaster*, *Wolbachia w*Mel densities in the oocyte increase proportionally faster than those in the nurse cells during stages 3-7 of oogenesis (230% in the oocyte versus 160% in the nurse cells) (Ferree et al., 2005). If this observation holds true in *Nasonia*, then the high *w*VitA densities that we see in *N. giraulti* embryos could be a combination of more efficient shuttling of *w*VitA from the nurse cells into the oocyte (see above) and faster proliferation of *w*VitA once in the oocyte.

The immune system of N. vitripennis *may be better at detecting and destroying* w*VitA*

One of the most direct ways that *N. vitripennis* could suppress *w*VitA titers is through an active immune response to the bacteria that is lacking in some capacity in *N. giraulti*. For example, antimicrobial peptides (AMPs) are important effector molecules in the insect innate immune system. They are generally small (less than 100 amino acids) and directly kill bacteria by forming pores in the bacterial membrane or by inhibiting their metabolic processes (Brogden, 2005). AMPs are also one of the fastest evolving proteins in the insect innate immune system (Vilcinskas, 2013; Waterhouse et al., 2007), and *N. vitripennis* has developed a more complex repertoire of AMPs than other insects including *D. melanogaster* and *A. mellifera* (Tian et al., 2010a). Due to their rapid evolution, AMP sequences are likely to have diverged even among closely-related species groups like *Nasonia*, so an *N. giraulti* homolog of an *N. vitripennis* AMP that controls *w*VitA may not be able to recognize *w*VitA once in *N. giraulti*.

One candidate gene on chromosome 3 is annotated as "holotricin-3" (LOC100116930), and is 2.7X upregulated in *N. vitripennis* compared to *N. giraulti* IntG (Table IV-3). Holotricin-3 is an AMP first purified from the hemolymph of the *Holotrichia diomphalia* beetle (Lee et al., 1995). With 54 glycines out of 104 amino acids total, the *N. vitripennis* version of the protein is likely part of the family of glycine-rich AMPs, which function by directly inhibiting synthesis of bacterial outer membrane proteins to increase membrane permeability (Yi et al., 2014). The higher expression of holotricin-3 in *N. vitripennis* may help keep *w*VitA densities low by destroying *w*VitA cells once they reach a certain threshold. Furthermore, an alignment of the *N. vitripennis* holotricin-3 protein with its homolog in *N. giraulti* reveals that the AMP may not even be functional in *N. giraulti* since a premature stop codon occurs after the fourth amino acid, though this result needs to be confirmed by sequencing the *N. giraulti* mRNA for this gene.

Similarly, homologs of chromosome 3 candidate gene apolipophorin-III (gene LOC100117157) also directly targets bacterial cells by altering the structure and permeability of bacterial cell membranes and are important activators of the innate immune response in insects (Weers and Ryan, 2006; Zdybicka-Barabas et al., 2011). Interestingly, one study in *Manduca sexta* moths found that females from immune-challenged parents (injected with peptidoglycan) upregulated expression of apolipophorin-III in their ovaries compared to females from naïve parents (Trauer-Kizilelma and Hilker, 2015), indicating some sort of maternal contribution to

ovarian apolipophorin-III in response to bacteria. Although apolipophorin-III was only upregulated in *N. vitripennis* by 1.2-fold in our RNA-seq study, if it is a potent inhibitor against *w*VitA then even a small expression change could produce an effect on densities. Furthermore, three amino acids changes between the *N. vitripennis* and *N. giraulti* alleles in a relatively short protein (191 AA) may change the protein's ability to interact with *w*VitA.

Autophagy is another innate immune defense widely used for the destruction of intracellular pathogens. Autophagy sequesters bacteria present in the cytosol of a cell within a double membraned vesicle (the autophagosome), which eventually merges with a lysosome where the bacteria is degraded (Yano and Kurata, 2011). Induction of autophagy using the drug rapamycin significantly reduced *Wolbachia w*AlbB densities in *Ae. albopictus* cell lines and *w*MelPop densities in *D. melanogaster*, while suppression of autophagy increased *Wolbachia* titers in both systems (Voronin et al., 2012). In another study, overproliferation of *Wolbachia* strain *w*VulC from *Armadillidium vulgare* in the central nervous system of a naïve host, *Porcellio d. dilatatus*, induces a high density of autophagosomes, one of which is clearly shown a engulfing a *Wolbachia* cell in TEM images (Le Clec'h et al., 2012). While autophagy has not yet been associated with bacterial clearance in germ cells or in the ovaries, autophagy is essential for proper egg development in Dipterans due to its role in the breakdown of nurse cell constituents near the end of oogenesis (Nezis et al., 2006). Assuming this is true for Hymenopterans, *N. vitripennis* may have lower *w*VitA titers if it has higher rates of autophagy or is better at targeting *w*VitA for destruction via autophagy than *N. giraulti*.

One candidate gene that may be involved in autophagy in *Nasonia* is LOC100114497, which is upregulated 2.3-fold in *N. giraulti* IntG over *N. vitripennis.* Annotated as "girdin-like" in NCBI, BLASTp analysis reveals a high similarity of approximately 50% of the protein to the *Nasonia* eukaryotic translation initiation factor 4 gamma (eIF4G) gene. eIF4G is a subunit of the eIF4F complex, which is important for initiation of protein translation and is regulated through the same pathway that regulates autophagy (the mTOR signaling pathway) (Gingras et al., 1998). Interestingly, suppression of eIF4G, but not the other subunits of eIF4F, can activate autophagy through the mTOR pathway in human cell lines, indicating that it may be a negative regulator of autophagy (Ramirez-Valle et al., 2008). Thus, lower expression of eIF4G in *N. vitripennis* may activate higher levels of autophagy, decreasing the number of *w*VitA cells in the ovaries.

Testing genes with unknown function

One of the advantages of using a forward genetic approach to identify candidate genes is that it does not make assumptions about the types of genes involved in producing a phenotype. However, when fine-mapping is hindered by areas of low recombination like in this study, other unbiased methods may help identify the correct candidate gene, regardless of its annotation. Here, RNA-seq of *Nasonia* ovaries was employed to detect gene expression differences that may underlie the *w*VitA density phenotype. The gene in either candidate region with the largest expression difference between *N. vitripennis* and *N. giraulti* IntG ovaries was trichohyalin (LOC1003317433), which was 32.3-fold upregulated in *N. vitripennis* according to the RNA-seq data (Table IV-3) and 65.2 upregulated according to RT-qPCR (Figure IV-10). Trichohyalin protein is typically found in hair follicles and specialized epithelial tissues, where it functions as an intermediate-filament crosslinking protein and lends mechanical strength to tissues like hair, nails, skin and tongue (Steinert et al., 2003). However, although the *Nasonia* LOC1003317433 gene is annotated as "trichohyalin-like" in NCBI, its protein product has no close homologs in the non-redundant protein database, so it is unclear what function it actually serves in *Nasonia*.

RNAi of trichohyalin in the IntC3 introgression line resulted in a 55% decrease in trichohyalin gene expression but did not significantly alter the average number of *w*VitA in embryos of Tricho-RNAi females (199.3 ± 133.5) versus GFP-RNAi females (129.6 ± 99.7) or uninjected females (168.2 ±86.2, Figure IV-12). However, considering how highly upregulated trichohyalin is in *N. vitripennis* and the fact that the *N. giraulti* allele of this gene is likely non-functional (Figure IV-11), then knocking down the gene in *N. vitripennis* or IntC3 with RNAi may not affect *Wolbachia* densities unless gene expression knockdown is 100% effective. If that is the case, then a transgenic technique like CRISPR/Cas9 may be needed to knock out trichohyalin in *N. vitripennis* in order see an effect on *w*VitA levels in embryos.

Other uncharacterized proteins identified as significantly differentially-expressed in the RNA-seq experiment include LOC100119051, LOC100122078, LOC100121799, LOC100122001, LOC100679525, LOC100119601, and LOC100679276. Though some of these genes contain protein domains that could be important for their function, they would ultimately need to be experimentally characterized if they were to affect *w*VitA densities after RNAi knockdown.

Conclusion

Our study is the first, to our knowledge, that uses a forward genetic approach to dissect the genomic landscape underlying host control of inherited symbiont titers. Similar quantitative trait analyses in mice have identified host genomic regions that shape the bacterial composition of the murine gut (Benson et al., 2010; McKnite et al., 2012) and skin (Srinivas et al., 2013) microbiota, but these are transient communities that are not directly transmitted to offspring. Conversely, reverse genetic screens in *Drosophila* have identified mutations in genes like *gurken* and the actin-binding proteins profilin and villin that produced significant changes in *Wolbachia* densities in the oocyte (Newton et al., 2015; Serbus et al., 2011). Our *Wolbachia* density phenotype, on the other hand, results from natural genetic variation between two closely-related species, with the dominant suppression of *w*VitA titers in its resident host *N. vitripennis* supporting the hypothesis that the candidate regions we have identified are biologically relevant to this symbiosis. Furthermore, reverse genetic mutant analyses are limited in the fact that they target a single gene or pathway, while forward genetics can estimate the total number of genes involved in generating a phenotype. Despite the complicated interactions that must occur between host and intracellular symbionts, our study indicates that the control of inherited symbionts can have a relatively simple genetic basis, with only two genomic regions acting additively to explain almost all of the 80-fold *w*VitA density difference in *Nasonia* wasps. Lastly, though this system reflects a particular host-microbe interaction, it has the potential to elucidate general molecular mechanisms, if any, by which maternal transmission is accomplished throughout the animal kingdom.

CHAPTER V. CONCLUSIONS AND FUTURE DIRECTIONS

Maternally-transmitted bacteria are ubiquitous throughout the animal kingdom, yet their impact on animal genome evolution is an important and underexplored area of research. The data presented in this thesis sheds light on how the pervasive intracellular symbiont *Wolbachia* can influence host genome evolution directly through horizontal gene transfer and indirectly through host factors that regulate *Wolbachia* titers. However, these studies leave many unanswered questions and active areas of future research, some of which are discussed herein.

Sequencing cytoplasmic and chromosomal *Wolbachia* in *C. parallelus*

Our discovery that the *Chorthippus parallelus* genome contains large amounts of *Wolbachia* DNA from multiple supergroups opens many new avenues of research into the number of total inserts; the size, chromosomal location and gene content of each insert; whether any chromosomal *Wolbachia* genes are expressed; and the history behind each insert, such as when the transfer occurred and the identity of the donor *Wolbachia*. However, to answer any of these questions, we will need much higher sequencing coverage of the inserts than what was achieved in CHAPTER III. Furthermore, sequencing the cytoplasmic *Wolbachia* genomes in *C. parallelus* would allow comparisons between the inserts and their potential donor strains and would also facilitate annotation of phage WO in this system (one of our original goals for the project).

Theoretically, sequencing of cytoplasmic and chromosomal *Wolbachia* could be accomplished by sequencing B-infected, F-infected and uninfected grasshoppers from the same population. With high enough sequence coverage and long enough read lengths, each cytoplasmic *Wolbachia* genome and genomic insert could be assembled computationally based on SNP variation. A successful example of this approach is the assembly of a cytoplasmic *Wolbachia* genome (*w*Gmm) and three separate chromosomal insertions of *w*Gmm in the tsetse fly genome with sequences generated from *Wolbachia*-infected and tetracycline-treated tsetse flies (*Glossina morsitans morsitans*) (Brelsfoard et al., 2014). However, the genome of *C. parallelus* is so large that any whole genome sequencing project would be very challenging from a technical, computational and economic perspective. For comparison, the only orthopteran

genome sequenced to date is that of the *Locusta migratoria*, which required 721 Gb of data to achieve 114X coverage for its 6.3 Gb genome (Wang et al., 2014). Even with high sequence coverage, the *L. migratoria* genome is still broken into over 550,000 scaffolds due to gaps in sequence assembly (Wang et al., 2014). At approximately 14 Gb in size, the *C. parallelus* genome is more than double the size of the *L. migratoria* genome (Lechner et al., 2013) and would require twice as much data to reach the same sequence coverage. Furthermore, with such a large genome, *Wolbachia* reads will only constitute a small fraction of any *C. parallelus* genome sequencing project. In our own study, *Wolbachia* reads only represented 0.01% of the 227M reads generated.

For assembly of *Wolbachia*-related sequences in *C. parallelus,* an alternative to whole genome sequencing would be to use a targeted sequence capture array tiled with *Wolbachia* probes. The array would capture *Wolbachia* DNA, both cytoplasmic and genomic, in the sample while excluding the rest of the enormous *C. parallelus* genome before high-throughput sequencing. This method has been used successfully in our lab to sequence phage WO from the *w*VitB infection of *N. vitripennis* (Kent et al., 2011), and it would drastically reduce the number of sequence reads needed to properly assemble cytoplasmic B and F *Wolbachia* genomes and the *Wolbachia* inserts in the grasshopper genome.

The hunt for phage WO in F Wolbachia

One benefit for assembling the genome of the cytoplasmic *Wolbachia* infections in *C. parallelus* would be the opportunity to characterize phage WO in this system. The original goal of the project in CHAPTER III was to determine whether phage WO had jumped between co-infecting B and F *Wolbachia* infections. However, we discovered so many phage WO minor capsid alleles in uninfected grasshoppers (Figure III-3) that we were unable to distinguish any phage WO alleles in cytoplasmic *Wolbachia* infections from laterally-transferred WO in the grasshopper genome. Yet, the abundance of phage WO in the grasshopper genome indicates that the cytoplasmic *Wolbachia* infections likely contain their own WO phages. Thus, a major question that remains unanswered is if the F *Wolbachia* in *C. parallelus* harbor phage WO, and, if so, whether the phage is unique to F *Wolbachia* or was transferred from a co-infecting B *Wolbachia* strain.

The ecology of the F *Wolbachia* supergroup is unique among *Wolbachia* because it has been identified in both arthropods (Baldo et al., 2007; Covacin and Barker, 2007; Panaram and Marshall, 2007; Salunke et al., 2010; Zabal-Aguirre et al., 2010) and nematodes (Bordenstein et al., 2009; Ferri et al., 2011; Lefoulon et al., 2012). *Wolbachia* that infect nematodes function as obligate mutualists and are required for host development, fertility and viability (Hoerauf et al., 1999; Taylor and Hoerauf, 1999). As such, nematode *Wolbachia* have reduced genomes with remnants of past phage infections but no intact WO phage (Darby et al., 2012; Foster et al., 2005; Kent et al., 2011). Since nematode F *Wolbachia* are unlikely to harbor WO due to their mutualistic lifestyle, it remains unclear whether the genomes of arthropod F *Wolbachia* also lack phage. The only arthropod F *Wolbachia* genome sequenced to date is that of *w*Cle from *Cimex lectularis* (bed bugs), which contains some isolated WO phage genes but no intact prophage regions (Nikoh et al., 2014). However, *w*Cle is not a good model for most arthropod F *Wolbachia* infections because it has independently transitioned from parasitism to mutualism due to its ability to produce biotin for its blood-sucking bed bug host (Nikoh et al., 2014). To the best of our knowledge, the only WO minor capsid genes identified in F *Wolbachia* were isolated from two sympatric cockroach species in Pune, India (Vaishampayan et al., 2007). Given that the cockroach specimens were collected from the same geographical area, the prevalence of WO in other arthropod F *Wolbachia* strains remains unknown. Thus, if we discovered phage WO in our sequencing of the grasshopper F *Wolbachia*, it would be one of the first characterizations of phage WO in any F *Wolbachia* strain.

Testing the effects of candidate genes on *Wolbachia* densities

Since knockdowns of kinesin and trichohyalin using parental RNAi did not conclusively affect embryonic *w*VitA densities, the identities of *Wolbachia* regulation genes in *Nasonia* remain elusive. Larger RNAi screens will be necessary to pinpoint the major genes of effect for this trait, and Table V-1 lists several genes in the chromosome 3 candidate region that are strong candidates based on their expression profiles and putative functions. Immediate RNAi screening endeavors will focus exclusively on chromosome 3 candidates using the homozygous IntC3 introgression line, but future efforts could test chromosome 2 candidate genes in pure *N. vitripennis* or *N. giraulti* IntG lines or in a hybrid line after the generation of a stable, *N. vitripennis* homozygous IntC2 introgression line.

Table V-1. Interesting candidate genes in the chromosome 3 candidate region

Gene	Protein Name	Fold Change (Nvit/IntG)	Reasons for knockdown
LOC100122001	Uncharacterized	6.17	Significantly overexpressed in Nvit Has a Rho GTPase domain that could regulate microtubules and actin
LOC100116930	Holotricin-3	2.69	Significantly overexpressed in Nvit Antimicrobial peptide May have a premature stop codon in IntG
LOC100121708	Tubulin-specific chaperone E	1.99	Overexpressed in Nvit Regulates microtubule assembly dynamics
LOC100117157	Apolipophorin-III	1.23	Involved in immune activation and pathogen recognition in insects
LOC100117412	Heat shock protein 83	1.28	Maternal RNA is localized to the posterior pole of oocytes
LOC100121249	Tyrosine-protein phosphatase Lar	-1.46	Significantly underexpressed in Nvit Required for proper localization of maternal factors like *oskar*
LOC100114497	Girdin	-2.28	Underexpressed in Nvit May be a negative regulator of autophagy
LOC100121852	Contactin	-2.44	Significantly underexpressed in Nvit GPI-anchored cell adhesion molecule
LOC100119248	Nephrin-like	-3.13	Underexpressed in Nvit and low-density Ngir Transmembrane protein that regulates actin dynamics

In cases like trichohyalin where the *N. giraulti* allele appears to be non-functional, complete ablation of *N. vitripennis* gene expression may be necessary to produce an effect on *w*VitA titers. If this is the case, then an alternative to RNAi would be to create transgenic *Nasonia* using genomic engineering technology like zinc-finger nucleases, TALENS or CRISPR/Cas9 (Gaj et al., 2013). The parasitoid lifestyle of *Nasonia* makes genomic engineering more challenging than in other insects because embryos develop inside another organism (fly pupae). However, a recent study was able to inject morpholinos into *Nasonia* embryos without killing them (Rosenberg et al., 2014), and the creation of a *Nasonia* rearing media in our lab could eliminate the need for a fly host during *Nasonia* development (Brucker and Bordenstein,

2012a). Several labs are currently working on using TALENs and CRISPR/Cas9 to manipulate the *Nasonia* genome with preliminary success (Lynch, 2015). Once genome editing technology is available in *Nasonia*, we can delete or disrupt candidate genes in our *Nasonia* lines and test their effects on *w*VitA density suppression.

CRISPR/Cas9 can also be used in combination with a homology repair donor to remove a gene and replace it with another sequence (Zheng et al., 2014). Instead of introgressing large regions of *N. vitripennis* into an *N. giraulti* background, this technology would allow us to replace a single *N. giraulti* gene with its homologous *N. vitripennis* sequence (and vice versa) to test the effects of genetic variation at a single locus in multiple *Nasonia* species without the confounding influence of other introgressed genes.

Investigating a parent-of-origin effect in IntC3

In CHAPTER IV, reciprocal crosses of *N. vitripennis* and *N. giraulti* IntG revealed that *w*VitA levels in *Nasonia* pupae reflected the maternal genotype: titers in genetically identical F1 females differed depending on the maternal species (Figure IV-2A), and heterogeneous F2 populations from identical F1 mothers had similar titers (Figure IV-2B). Thus, densities appeared to be controlled through a genetic maternal effect, which occurs when the phenotype of an organism is determined by the genotype of the mother. However, while analyzing crosses with the IntC3 segmental introgression line, we discovered that females heterozygous at the IntC3 region produced offspring with different *w*VitA densities based on which parent provided her the *N. vitripennis* IntC3 allele (Figure V-1). Heterozygous females that inherited a maternal copy of the *N. vitripennis* IntC3 region suppressed *w*VitA densities in offspring by 52.9% on average, which was not significantly different than an average suppression of 61.8% in offspring from females homozygous for the *N. vitripennis* allele. In contrast, heterozygous females that inherited a paternal copy of the *N. vitripennis* IntC3 region only suppressed *w*VitA titers by 27.3%. However, this study was confounded by the fact that the females that received a paternal copy of the *N. vitripennis* IntC3 region were also heterozygous downstream of the IntC3 region (Figure V-1).

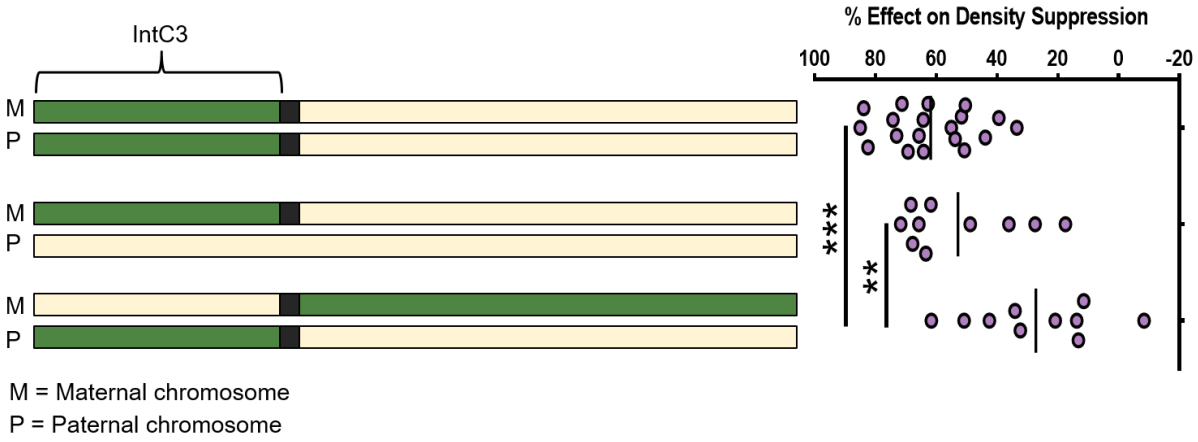

Figure V-1. Chromosome 3 parent-of-origin effect on density suppression
Sets of chromosomes represent the genotype of females at the chromosome 3 candidate region (IntC3) or downstream of the centromeric region (black box). Green represents an *N. vitripennis* region while cream represents an *N. giraulti* region. For each genotype, the maternal chromosome (M) is displayed above the paternal chromosome (P). The percent effect of each genotype on *w*VitA density suppression in offspring of each genotype is graphed on the right. **p<0.01, ***p<0.001, One-way ANOVA followed by a Tukey-Kramer test of multiple comparisons

To rule out any epistatic interactions between the two heterozygous regions that may have caused a decrease in suppression (and an increase in titers), we set up reciprocal crosses of a homozygous *N. vitripennis* IntC3 segmental introgression line with *N. giraulti* IntG to generate F1 hybrid females that were identical except for which parent provided the *N. vitripennis* allele of the IntC3 region (Figure V-2A). If the density suppression gene on chromosome 3 was acting through a pure maternal effect, then *w*VitA levels in F1 hybrids should have reflected that of their mother, not their own genotype. While *w*VitA densities of F1 pupae from the IntC3 (F) x IntG (M) cross were identical to homozygous IntC3 pupae, F1 pupae from the IntG (F) x IntC3 (M) cross had approximately 60% lower *w*VitA densities than pure IntG pupae (Figure V-2B). This indicates that zygotic expression of the *N. vitripennis* IntC3 allele could be driving densities in F1 hybrids. If this is the case, then an *N. vitripennis* IntC3 maternal allele is better than a paternal allele at suppressing densities since *w*VitA titers are significantly lower in heterozygous females with the maternal allele (Figure V-2B, $p = 0.009$, One-way ANOVA with a Tukey-Kramer's multiple comparisons test).

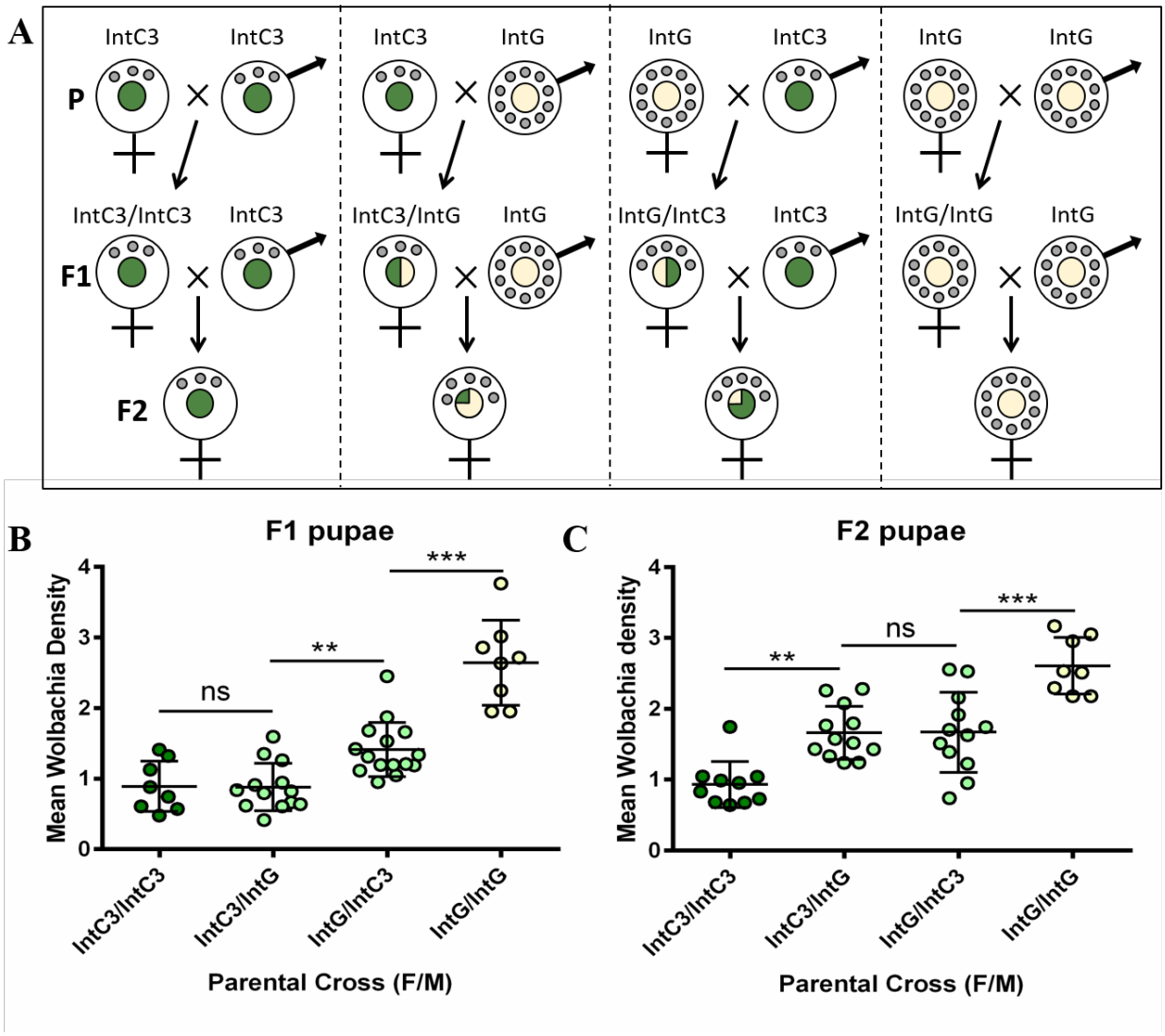

Figure V-2. IntC3 parent-of-origin effect on F1 and F2 pupae

(A) Crossing scheme for testing the IntC3 parent-of-origin effect on *w*VitA densities. Reciprocal crosses of IntC3 (green nuclei) and *N. giraulti* IntG (cream nuclei) produced identical F1 hybrid females, which were then crossed with their paternal line to produce heterogenous F2 pupae. Pure-breeding IntC3 and IntG lines were used as controls. Small, gray circles around the nuclei in the gender symbol represent the pupal *w*VitA densities for each genotype. Pupal *w*VitA densities were measured by qPCR for the (B) F1 and (C) F2 generations, where dark green is the IntC3 control cross, light green are the hybrid crosses, and cream is the pure IntG cross. F = female, M = male **p<0.01, ***p<0.001, One-way ANOVA followed by a Tukey-Kramer test of multiple comparisons

F1 heterozygous females were then backcrossed to their paternal line, and *w*VitA densities were measured for a pool of five F2 female pupae per cross (Figure V-2A). In this case, *w*VitA densities in heterogeneous F2 pupae were similar for both experimental groups, consistent with a genetic maternal effect (Figure V-2C). Thus, it is still unclear whether a

maternal effect or a zygotic parent-of-origin effect (or a combination of the two) is more important in establishing *w*VitA densities. To tease apart these possibilities, the experiment will be repeated again, but will measure *w*VitA levels in individual F2 female pupae instead of pools (Figure V-3). This way, if a parent-of-origin effect is important, than F2 pupae from the same mother will have different densities (Figure V-3A) depending on which IntC3 alleles they inherit. If the maternal genotype is more important in regulating offspring titers, then we expect to see relatively equal levels of *w*VitA across all F2 pupae regardless of zygotic genotype (Figure V-3B). If we see something in between these two possibilities, then there is likely a combinatorial effect where both maternal and zygotic expression of IntC3 influence *w*VitA titers.

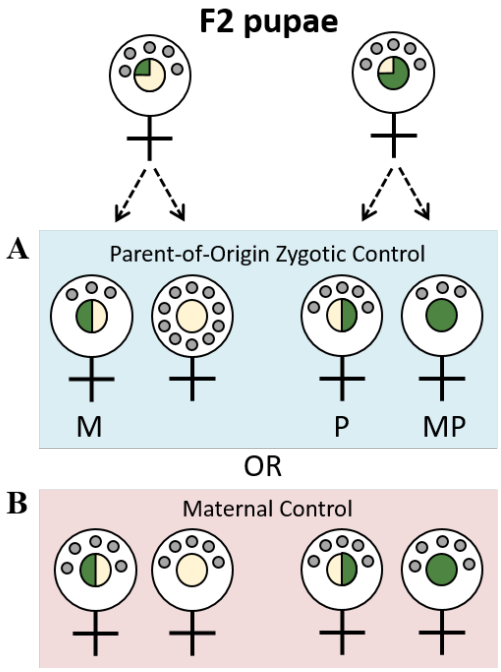

Figure V-3. Separating parent-of-origin effect from maternal effect
Ongoing experiment where *w*VitA densities (gray circles) are tested in individual F2 females from reciprocal crosses of IntC3 (green) and *N. giraulti* IntG (cream) instead of a heterogeneous pool of F2 pupae. (A) If a zygotic parent-of-origin effect controls *w*VitA densities, then females that receive an *N. vitripennis* IntC3 maternal allele (M) will have lower densities than those that receive an *N. vitripennis* IntC3 paternal allele (P). Females homozygous *N. giraulti* for IntC3 will have the same densities as pure *N. giraulti* IntG. (B) If *w*VitA densities are controlled through a maternal effect, all pupae will have similar *Wolbachia* densities, regardless of their genotype.

Parent-of-origin effects are often a result of allele-specific gene expression. For example, in mammals, genomic imprinting through DNA methylation of regulatory regions can

specifically silence either the maternal or paternal copy of a gene, allowing the transcriptionally-active gene to dictate the phenotype (Lawson et al., 2013). However, many insects, including *Drosophila melanogaster*, have almost undetectable levels of genome methylation (Raddatz et al., 2013; Yan et al., 2015). Although *Nasonia vitripennis* has one of the highest overall methylation levels in Hymenoptera at 1.6% of methylated CpG sites (Wang et al., 2013), its methylation profile is still extremely sparse compared to the 70-80% of CpGs methylated in the human genome (Bird, 2002). Despite the low levels of methylation, one study found that approximately 80% of genes in the *N. vitripennis* genome have a significant sex-biased gene expression, though only 7.6% were classified as "sex-specific" or "extremely-biased" in one sex over the other (Wang et al., 2015). Furthermore, they found little evidence of sex differences in DNA methylation, indicating that sex-biased expression in *Nasonia* is likely regulated through a mechanism other than DNA methylation (Wang et al., 2015).

Predicting new candidate genes

If the gene responsible for suppressing *w*VitA densities in the chromosome 3 candidate region does follow a parent-of-origin effect, then we could potentially narrow down our list of candidate genes considerably by identifying *N. vitripennis* alleles that are differentially expressed in reciprocal F1 hybrid females. Currently, we are extracting total RNA from the abdomens and/or ovaries of reciprocal F1 hybrid females to identify gene expression differences in the chromosome 3 candidate region with RNA-seq and RT-qPCR. Since the number of reads that mapped to coding genes from our last RNA sequencing run was so low (less than 1M reads per sample) due to high levels of ribosomal RNA contamination, we will also be resequencing the ovary transcriptomes of *N. vitripennis* 12.1 and *N. giraulti* IntG. This time, we will be using a TagSeq approach, which only sequences the 3' end of mRNA molecules with poly-A tails (Meyer et al., 2011). This approach differs from normal poly-A selection before RNA-sequencing because, with TagSeq, only a portion of the entire RNA molecule is sequenced, allowing accurate read counts while lowering the number of reads needed to detect rare transcripts (Meyer et al., 2011). On the same lane, we will also be sequencing the ovary transcriptome of the IntC3 segmental introgression line and uninfected strains of *N. vitripennis* (12.1T) and *N. giraulti* (IntG12.1T). Considering that some proteins in the ovaries may be synthesized elsewhere, like AMPs from the fat body (Meister et al., 1997), and not reflected in

an ovary-specific transcriptome, we also plan on using mass spectrometry to generate species-specific proteomic profiles of *Nasonia* ovaries.

Determining the molecular mechanisms behind host regulation of *Wolbachia* titers

Even without knowing the identity of the host genes involved in *w*VitA regulation, we can still gain insight into how *Wolbachia* titers are controlled mechanistically. For example, staining *N. vitripennis* and *N. giraulti* ovarioles with a nucleic acid dye revealed different localization patterns of *w*VitA during oogenesis (Figure IV-13). These observations can be taken a step further by staining with an antibody that recognizes *Wolbachia*, such as anit-hsp60 (Ferree et al., 2005; Serbus and Sullivan, 2007), so that *w*VitA can be distinguished from host nuclei in the early and late stages of oogenesis. Doing so would help determine whether *w*VitA cells stay within the nurse cells at the end oogenesis in *N. vitripennis* instead of being transferred to the oocyte. To investigate the role for the cytoskeleton in regulating *w*VitA densities, microtubule or actin depolymerizing drugs like colchicine and cytochalasin-D, respectively, could be used to see how *w*VitA distribution or density changes in the ovariole after disruption of cytoskeletal networks (Ferree et al., 2005; Serbus and Sullivan, 2007). However, results from these drug assays would need to be interpreted with caution since disrupting microtubules or actin will impact many other processes, such as transport of maternal RNAs into the oocyte, that may have a confounding influence on *Wolbachia* densities.

Once a candidate gene has been shown to affect *w*VitA densities after RNAi knock-down, it may be useful to visualize the RNA or protein product of the gene in conjunction with wVitA. For example, if *w*VitA is suspected of hitchhiking on a maternal RNA to move into the oocyte, then fluorescent in situ hybridization against the maternal RNA could confirm that the RNA and *w*VitA co-localize together in the cell. If a protein is thought to interact with *w*VitA, an antibody generated against the protein could be also be used for co-localization studies or in immunoprecipitation assays to determine its *Wolbachia* protein binding partners. Similarly, if a host uncharacterized protein with no known function affects *w*VitA densities, then immunoprecipitation with an antibody against the protein could help identify which host processes the protein participates in based on its binding partners.

Concluding Remarks

This body of work contributes only a small piece to the much larger puzzle that is microbe-mediated genome evolution. Nevertheless, the research presented here has broadened our understanding of maternal microbial transmission to include all animals, even humans; redefined the limits of *Wolbachia* horizontal gene transfer to animal genomes; and uncovered a potential mechanism for *Wolbachia* density control at the maternal-zygotic interface. Future studies will hopefully provide greater insight into the genetic and molecular mechanisms underlying the observations presented here, which could directly impact fields as diverse as evolutionary biology, microbiology and health care.

REFERENCES

Aagaard, K., Riehle, K., Ma, J., Segata, N., Mistretta, T.A., Coarfa, C., Raza, S., Rosenbaum, S., Van den Veyver, I., Milosavljevic, A.*, et al.* (2012). A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. PLoS One *7*, e36466.

Aikawa, T., Anbutsu, H., Nikoh, N., Kikuchi, T., Shibata, F., and Fukatsu, T. (2009). Longicorn beetle that vectors pinewood nematode carries many Wolbachia genes on an autosome. Proc Biol Sci *276*, 3791-3798.

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., and Aksoy, S. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies, Wigglesworthia glossinidia. Nat Genet *32*, 402-407.

Al-Bahry, S., Mahmoud, I., Elshafie, A., Al-Harthy, A., Al-Ghafri, S., Al-Amri, I., and Alkindi, A. (2009). Bacterial flora and antibiotic resistance from eggs of green turtles Chelonia mydas: an indication of polluted effluents. Marine pollution bulletin *58*, 720-725.

Alikhan, N.F., Petty, N.K., Ben Zakour, N.L., and Beatson, S.A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics *12*, 402.

Alkindi, A.Y., Mahmoud, I.Y., Woller, M.J., and Plude, J.L. (2006). Oviductal morphology in relation to hormonal levels in the snapping turtle, Chelydra serpentina. Tissue Cell *38*, 19-33.

Anbutsu, H., and Fukatsu, T. (2003). Population dynamics of male-killing and non-male-killing spiroplasmas in Drosophila melanogaster. Appl Environ Microbiol *69*, 1428-1434.

Anselme, C., Vallier, A., Balmand, S., Fauvarque, M.O., and Heddi, A. (2006). Host PGRP gene expression and bacterial release in endosymbiosis of the weevil Sitophilus zeamais. Appl Environ Microbiol *72*, 6766-6772.

Attardo, G.M., Lohs, C., Heddi, A., Alam, U.H., Yildirim, S., and Aksoy, S. (2008). Analysis of milk gland structure and function in Glossina morsitans: milk protein production, symbiont populations and fecundity. J Insect Physiol *54*, 1236-1242.

Bager, P., Simonsen, J., Nielsen, N.M., and Frisch, M. (2012). Cesarean section and offspring's risk of inflammatory bowel disease: a national cohort study. Inflamm Bowel Dis *18*, 857-862.

Bager, P., Wohlfahrt, J., and Westergaard, T. (2008). Caesarean delivery and risk of atopy and allergic disease: meta-analyses. Clin Exp Allergy *38*, 634-642.

Baldo, L., Prendini, L., Corthals, A., and Werren, J.H. (2007). Wolbachia are present in southern african scorpions and cluster with supergroup F. Curr Microbiol *55*, 367-373.

Balmand, S., Lohs, C., Aksoy, S., and Heddi, A. (2013). Tissue distribution and transmission routes for the tsetse fly endosymbionts. J Invertebr Pathol *112*, S116-S122.

Bastock, R., and St Johnston, D. (2008). Drosophila oogenesis. Current biology : CB *18*, R1082-1087.

Baumann, P. (2005). Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annu Rev Microbiol *59*, 155-189.

Baumann, P., Baumann, L., Lai, C.Y., Rouhbakhsh, D., Moran, N.A., and Clark, M.A. (1995). Genetics, physiology, and evolutionary relationships of the genus Buchnera: intracellular symbionts of aphids. Annu Rev Microbiol *49*, 55-94.

Bearfield, C., Davenport, E.S., Sivapathasundaram, V., and Allaker, R.P. (2002). Possible association between amniotic fluid micro-organism infection and microflora in the mouth. Bjog *109*, 527-533.

Beckage, N.E. (2008). Insect Immunology (Oxford: Academic Press).

Beeuwer, J., and Werren, J.H. (1993). Cytoplasmic incompatibility and bacterial density in *Nasonia vitripennis*. Genetics *135*, 565-574.

Bella, J.L., Hewitt, G.M., and Gosalvez, J. (1990). Meiotic imbalance in laboratory produced hybrid males of *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus*. Genetics Research *56*, 43-48.

Bella, J.L., Serrano, L., Orellana, J., and Mason, P.L. (2007). The origin of the Chorthippus parallelus hybrid zone: chromosomal evidence of multiple refugia for Iberian populations. J Evol Biol *20*, 568-576.

Bencina, D., Tadina, T., and Dorrer, D. (1988). Natural infection of ducks with Mycoplasma synoviae and Mycoplasma gallisepticum and Mycoplasma egg transmission. Avian pathology : journal of the WVPA *17*, 441-449.

Benevides de Morais, P., Wessel de Oliveira, K., Malvasio, A., Gomes de Ataide, A., and Pimenta, R.S. (2010). Enterobacteriaceae associated with eggs of *Podocnemis expansa* and *Podocnemis unifilis* (Testudines: Chelonia) in nonpolluted sites of National Park of Araguaia Plains, Brazil. J Zoo Wildl Med *41*, 656-661.

Bensasson, D., Petrov, D.A., Zhang, D.X., Hartl, D.L., and Hewitt, G.M. (2001). Genomic gigantism: DNA loss is slow in mountain grasshoppers. Mol Biol Evol *18*, 246-253.

Bensasson, D., Zhang, D.X., and Hewitt, G.M. (2000). Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. Mol Biol Evol *17*, 406-415.

Benson, A.K., Kelly, S.A., Legge, R., Ma, F., Low, S.J., Kim, J., Zhang, M., Oh, P.L., Nehrenberg, D., Hua, K.*, et al.* (2010). Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci U S A *107*, 18933-18938.

Berticat, C., Rousset, F., Raymond, M., Berthomieu, A., and Weill, M. (2002). High Wolbachia density in insecticide-resistant mosquitoes. Proc Biol Sci *269*, 1413-1416.

Beukeboom, L.W., Niehuis, O., Pannebakker, B.A., Koevoets, T., Gibson, J.D., Shuker, D.M., van de Zande, L., and Gadau, J. (2010). A comparison of recombination frequencies in intraspecific versus interspecific mapping populations of Nasonia. Heredity (Edinb) *104*, 302-309.

Bian, G., Zhou, G., Lu, P., and Xi, Z. (2013). Replacing a native Wolbachia with a novel strain results in an increase in endosymbiont load and resistance to dengue virus in a mosquito vector. PLoS Negl Trop Dis *7*, e2250.

Biasucci, G., Benenati, B., Morelli, L., Bessi, E., and Boehm, G. (2008). Cesarean delivery may affect the early biodiversity of intestinal bacteria. J Nutr *138*, 1796S-1800S.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev *16*, 6-21.

Blaul, B., and Ruther, J. (2011). How parasitoid females produce sexy sons: a causal link between oviposition preference, dietary lipids and mate choice in *Nasonia*. Proc R Soc B *278*, 3286-3293.

Bordenstein, S.R. (2007). Evolutionary genomics: transdomain gene transfers. Current Biology *17*, R935-936.

Bordenstein, S.R., and Bordenstein, S.R. (2011). Temperature affects the tripartite interactions between bacteriophage WO, Wolbachia, and cytoplasmic incompatibility. PLoS One *6*, e29106.

Bordenstein, S.R., Marshall, M.L., Fry, A.J., Kim, U., and Wernegreen, J.J. (2006). The tripartite associations between bacteriophage, Wolbachia, and arthropods. PLoS Pathog *2*, e43.

Bordenstein, S.R., Paraskevopoulos, C., Dunning Hotopp, J.C., Sapountzis, P., Lo, N., Bandi, C., Tettelin, H., Werren, J.H., and Bourtzis, K. (2009). Parasitism and mutualism in Wolbachia: what the phylogenomic trees can and cannot say. Mol Biol Evol *26*, 231-241.

Bordenstein, S.R., and Reznikoff, W.S. (2005). Mobile DNA in obligate intracellular bacteria. Nat Rev Microbiol *3*, 688-699.

Bordenstein, S.R., and Theis, K.R. (2015). Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. PLoS Biol *13*, e1002226.

Bordenstein, S.R., and Wernegreen, J.J. (2004). Bacteriophage flux in endosymbionts (Wolbachia): infection frequency, lateral transfer, and recombination rates. Mol Biol Evol *21*, 1981-1991.

Bordenstein, S.R., and Werren, J.H. (2007). Bidirectional incompatibility among divergent Wolbachia and incompatibility level differences among closely related Wolbachia in Nasonia. Heredity (Edinb) *99*, 278-287.

Boyle, L., O'Neill, S.L., Robertson, H.M., and Karr, T.L. (1993). Interspecific and intraspecific horizontal transfer of Wolbachia in Drosophila. Science *260*, 1796-1799.

Brelsfoard, C., Tsiamis, G., Falchetto, M., Gomulski, L.M., Telleria, E., Alam, U., Doudoumis, V., Scolari, F., Benoit, J.B., Swain, M.*, et al.* (2014). Presence of extensive Wolbachia symbiont insertions discovered in the genome of its host Glossina morsitans morsitans. PLoS Negl Trop Dis *8*, e2728.

Bright, M., and Bulgheresi, S. (2010). A complex journey: transmission of microbial symbionts. Nat Rev Microbiol *8*, 218-230.

Brock, J.A., and Bullis, R. (2001). Disease prevention and control for gametes and embryos of fish and marine shrimp. Aquaculture *197*, 137-159.

Brogden, K.A. (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat Rev Microbiol *3*, 238-250.

Broman, K.W., Wu, H., Sen, S., and Churchill, G.A. (2003). R/qtl: QTL mapping in experimental crosses. Bioinformatics *19*, 889-890.

Brown, L.L., Cox, W.T., and Levine, R.P. (1997). Evidence that the causal agent of bacterial cold-water disease *Flavobacterium psychrophilum* is transmitted within salmonid eggs. Dis Aquat Org *29*, 213-218.

Brucker, R.M., and Bordenstein, S.R. (2012a). In vitro cultivation of the hymenoptera genetic model, Nasonia. PLoS One *7*, e51269.

Brucker, R.M., and Bordenstein, S.R. (2012b). The roles of host evolutionary relationships (genus: *Nasonia*) and development in structuring microbial communities. Evolution *66*, 349-362.

Buchner, P. (1965). Endosymbiosis of Animals with Plant Microorganisms (New York: Interscience Publishers).

Cabrera-Rubio, R., Collado, M.C., Laitinen, K., Salminen, S., Isolauri, E., and Mira, A. (2012). The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr *96*, 544-551.

Calderone, R.A., and Fonzi, W.A. (2001). Virulence factors of Candida albicans. Trends Microbiol *9*, 327-335.

Campbell, B.C., Steffen-Campbell, J.D., and Werren, J.H. (1993). Phylogeny of the Nasonia species complex (Hymenoptera: Pteromalidae) inferred from an internal transcribed spacer (ITS2) and 28S rDNA sequences. Insect Mol Biol *2*, 225-237.

Candela, M., Perna, F., Carnevali, P., Vitali, B., Ciati, R., Gionchetti, P., Rizzello, F., Campieri, M., and Brigidi, P. (2008). Interaction of probiotic Lactobacillus and Bifidobacterium strains with human intestinal epithelial cells: adhesion properties, competition against enteropathogens and modulation of IL-8 production. Int J Food Microbiol *125*, 286-292.

Caragata, E.P., Rances, E., Hedges, L.M., Gofton, A.W., Johnson, K.N., O'Neill, S.L., and McGraw, E.A. (2013a). Dietary cholesterol modulates pathogen blocking by Wolbachia. PLoS Pathog *9*, e1003459.

Caragata, E.P., Rances, E., O'Neill, S.L., and McGraw, E.A. (2013b). Competition for Amino Acids Between Wolbachia and the Mosquito Host, Aedes aegypti. Microbial ecology *67*, 205-218.

Cardwell, C.R., Stene, L.C., Joner, G., Cinek, O., Svensson, J., Goldacre, M.J., Parslow, R.C., Pozzilli, P., Brigis, G., Stoyanov, D*., et al.* (2008). Caesarean section is associated with an increased risk of childhood-onset type 1 diabetes mellitus: a meta-analysis of observational studies. Diabetologia *51*, 726-735.

Carlier, Y., Truyens, C., Deloron, P., and Peyron, F. (2012). Congenital parasitic infections: a review. Acta Trop *121*, 55-70.

Cary, S.C., and Giovannoni, S.J. (1993). Transovarial inheritance of endosymbiotic bacteria in clams inhabiting deep-sea hydrothermal vents and cold seeps. Proc Natl Acad Sci U S A *90*, 5695-5699.

Casiraghi, M., Anderson, T.J., Bandi, C., Bazzocchi, C., and Genchi, C. (2001). A phylogenetic analysis of filarial nematodes: comparison with the phylogeny of Wolbachia endosymbionts. Parasitology *122 Pt 1*, 93-103.

Casiraghi, M., Bordenstein, S.R., Baldo, L., Lo, N., Beninati, T., Wernegreen, J.J., Werren, J.H., and Bandi, C. (2005). Phylogeny of Wolbachia pipientis based on gltA, groEL and ftsZ gene sequences: clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the Wolbachia tree. Microbiology *151*, 4015-4022.

Cavanaugh, C.M. (1983). Symbiotic chemoautotrophic bacteria in marine invertebrates from sulfide-rich habitats. Nature *302*, 58-61.

Cavanaugh, C.M., P., M.Z., Newton, I.L.G., and Stewart, F.J. (2006). Marine chemosynthetic symbioses. In The prokaryotes, M. Dworkin, S. Falkow, E. Rosenberg, K.H. Schleifer, and S. E., eds. (New York: Springer).

Cephas, K.D., Kim, J., Mathai, R.A., Barry, K.A., Dowd, S.E., Meline, B.S., and Swanson, K.S. (2011). Comparative analysis of salivary bacterial microbiome diversity in edentulous infants and their mothers or primary care givers using pyrosequencing. PLoS One *6*, e23503.

Chafee, M.E., Funk, D.J., Harrison, R.G., and Bordenstein, S.R. (2010). Lateral phage transfer in obligate intracellular bacteria (wolbachia): verification from natural populations. Mol Biol Evol *27*, 501-505.

Chafee, M.E., Zecher, C.N., Gourley, M.L., Schmidt, V.T., Chen, J.H., Bordenstein, S.R., and Clark, M.E. (2011). Decoupling of host-symbiont-phage coadaptations following transfer between insect species. Genetics *187*, 203-215.

Chan, Y.G., Cardwell, M.M., Hermanas, T.M., Uchiyama, T., and Martinez, J.J. (2009). Rickettsial outer-membrane protein B (rOmpB) mediates bacterial invasion through Ku70 in an actin, c-Cbl, clathrin and caveolin 2-dependent manner. Cell Microbiol *11*, 629-644.

Chasan, R., and Anderson, K.V. (1989). The role of easter, an apparent serine protease, in organizing the dorsal-ventral pattern of the Drosophila embryo. Cell *56*, 391-400.

Chen, S.J., Lu, F., Cheng, J.A., Jiang, M.X., and Way, M.O. (2012). Identification and biological role of the endosymbionts Wolbachia in rice water weevil (Coleoptera: Curculionidae). Environ Entomol *41*, 469-477.

Cho, K.O., Kim, G.W., and Lee, O.K. (2011). Wolbachia bacteria reside in host Golgi-related vesicles whose position is regulated by polarity proteins. PLoS One *6*, e22703.

Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G*., et al.* (2002). Immunity-related genes and gene families in *Anopheles gambiae*. Science *298*, 159-165.

Clark, M.E., O'Hara, F.P., Chawla, A., and Werren, J.H. (2010). Behavioral and spermatogenic hybrid male breakdown in *Nasonia*. Heredity (Edinb) *104*, 289-301.

Cmelik, S.H.W., Bursell, E., and Slack, E. (1969). Composition of gut contents of third-instar tsetse larvae (*Glossina morsitans* Westwood). Comp Biochem Physiol B Biochem Mol Biol *29*, 447.

Colombo, A.V., Silva, C.M., Haffajee, A., and Colombo, A.P. (2006). Identification of oral bacteria associated with crevicular epithelial cells from chronic periodontitis lesions. J Med Microbiol *55*, 609-615.

Cooper, M.D., and Alder, M.N. (2006). The evolution of adaptive immune systems. Cell *124*, 815-822.

Cooper, S.J., and Hewitt, G.M. (1993). Nuclear DNA sequence divergence between parapatric subspecies of the grasshopper Chorthippus parallelus. Insect Mol Biol *2*, 185-194.

Costello, E.K., Stagaman, K., Dethlefsen, L., Bohannan, B.J., and Relman, D.A. (2012). The application of ecological theory toward an understanding of the human microbiome. Science *336*, 1255-1262.

Covacin, C., and Barker, S.C. (2007). Supergroup F Wolbachia bacteria parasitise lice (Insecta: Phthiraptera). Parasitol Res *100*, 479-485.

Craven, K.S., Awong-Taylor, J., Griffiths, L., Bass, C., and Muscarella, M. (2007). Identification of bacterial isolates from unhatched loggerhead (*Caretta caretta*) sea turtle eggs in Georgia, USA. Marine Turtle Newsletter *115*, 9-11.

Darby, A.C., Armstrong, S.D., Bah, G.S., Kaur, G., Hughes, M.A., Kay, S.M., Koldkjaer, P., Rainbow, L., Radford, A.D., Blaxter, M.L*., et al.* (2012). Analysis of gene expression from the

Wolbachia genome of a filarial nematode supports both metabolic and defensive roles within the symbiosis. Genome Res *22*, 2467-2477.

De Gregorio, E., Spellman, P.T., Rubin, G.M., and Lemaitre, B. (2001). Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. Proc Natl Acad Sci U S A *98*, 12590-12595.

Decker, E., Engelmann, G., Findeisen, A., Gerner, P., Laass, M., Ney, D., Posovszky, C., Hoy, L., and Hornef, M.W. (2010). Cesarean delivery is associated with celiac disease but not inflammatory bowel disease in children. Pediatrics *125*, e1433-1440.

Dedeine, F., Vavre, F., Fleury, F., Loppin, B., Hochberg, M.E., and Bouletreau, M. (2001). Removing symbiotic Wolbachia bacteria specifically inhibits oogenesis in a parasitic wasp. Proc Natl Acad Sci U S A *98*, 6247-6252.

Desjardins, C.A., Cerqueira, G.C., Goldberg, J.M., Dunning Hotopp, J.C., Haas, B.J., Zucker, J., Ribeiro, J.M., Saif, S., Levin, J.Z., Fan, L.*, et al.* (2013a). Genomics of Loa loa, a Wolbachia-free filarial parasite of humans. Nature Genetics *45*, 495-500.

Desjardins, C.A., Gadau, J., Lopez, J.A., Niehuis, O., Avery, A.R., Loehlin, D.W., Richards, S., Colbourne, J.K., and Werren, J.H. (2013b). Fine-scale mapping of the Nasonia genome to chromosomes using a high-density genotyping microarray. G3 *3*, 205-215.

Desjardins, C.A., Perfectti, F., Bartos, J.D., Enders, L.S., and Werren, J.H. (2010). The genetic basis of interspecies host preference differences in the model parasitoid *Nasonia*. Heredity (Edinb) *104*, 270-277.

Di Meo, C.A., Wilbur, A.E., Holben, W.E., Feldman, R.A., Vrijenhoek, R.C., and Cary, S.C. (2000). Genetic variation among endosymbionts of widely distributed vestimentiferan tubeworms. Appl Environ Microbiol *66*, 651-658.

Diaz-Ropero, M.P., Martin, R., Sierra, S., Lara-Villoslada, F., Rodriguez, J.M., Xaus, J., and Olivares, M. (2007). Two Lactobacillus strains, isolated from breast milk, differently modulate the immune response. Journal of applied microbiology *102*, 337-343.

DiGiulio, D.B. (2012). Diversity of microbes in amniotic fluid. Semin Fetal Neonatal Med *17*, 2-11.

DiGiulio, D.B., Romero, R., Amogan, H.P., Kusanovic, J.P., Bik, E.M., Gotsch, F., Kim, C.J., Erez, O., Edwin, S., and Relman, D.A. (2008). Microbial prevalence, diversity and abundance in amniotic fluid during preterm labor: a molecular and culture-based investigation. PLoS One *3*, e3056.

Ding, D., Parkhurst, S.M., Halsell, S.R., and Lipshitz, H.D. (1993). Dynamic Hsp83 RNA localization during Drosophila oogenesis and embryogenesis. Molecular and cellular biology *13*, 3773-3781.

Dobson, S.L., Bourtzis, K., Braig, H.R., Jones, B.F., Zhou, W., Rousset, F., and O'Neill, S.L. (1999). Wolbachia infections are distributed throughout insect somatic and germ line tissues. Insect Biochem Mol Biol *29*, 153-160.

Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci U S A *107*, 11971-11975.

Douglas, A.E. (1989). Mycetocyte symbiosis in insects. Biological Reviews *64*, 409-434.

Douglas, A.E. (1998). Nutritional interactions in insect-microbial symbioses: Aphids and their symbiotic bacteria *Buchnera*. Annual review of entomology *43*, 17-37.

Douglas, A.E. (2011). Lessons from studying insect symbioses. Cell Host Microbe *10*, 359-367.

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

Dunning Hotopp, J.C., Clark, M.E., Oliveira, D.C., Foster, J.M., Fischer, P., Munoz 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*, 1753-1756.

Dyer, K.A., Minhas, M.S., and Jaenike, J. (2005). Expression and modulation of embryonic male-killing in Drosophila innubila: opportunities for multilevel selection. Evolution *59*, 838-848.

Eisthen, H.L., and Theis, K.R. (2016). Animal-microbe interactions and the evolution of nervous systems. Philos Trans R Soc Lond B Biol Sci *371*.

Endow, K., and Ohta, S. (1990). Occurrence of bacteria in the primary oocytes of vesicomyid clam *Calyptogena soyoae*. Marine Ecology Progress Series *64*, 309-311.

Enticknap, J.J., Kelly, M., Peraud, O., and Hill, R.T. (2006). Characterization of a culturable alphaproteobacterial symbiont common to many marine sponges and evidence for vertical transmission via sponge larvae. Appl Environ Microbiol *72*, 3724-3732.

Ereskovsky, A.V., Gonobobleva, E., and Vishnyakov, A. (2005). Morphological evidence for vertical transmission of symbiotic bacteria in the viviparous sponge *Halisarca dujardini* Johnston (Porifera, Demospongiae, Halisarcida). Marine Biology *146*, 869-875.

Evans, J.D., Aronstein, K., Chen, Y.P., Hetru, C., Imler, J.L., Jiang, H., Kanost, M., Thompson, G.J., Zou, Z., and Hultmark, D. (2006). Immune pathways and defence mechanisms in honey bees *Apis mellifera*. Insect Mol Biol *15*, 645-656.

Fast, E.M., Toomey, M.E., Panaram, K., Desjardins, D., Kolaczyk, E.D., and Frydman, H.M. (2011). Wolbachia enhance Drosophila stem cell proliferation and target the germline stem cell niche. Science *334*, 990-992.

Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Analytical biochemistry *132*, 6-13.

Feinberg, A.P., and Vogelstein, B. (1984). "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". Addendum. Analytical biochemistry *137*, 266-267.

Feldhaar, H., and Gross, R. (2009). Insects as hosts for mutualistic bacteria. Int J Med Microbiol *299*, 1-8.

Fenn, K., and Blaxter, M. (2004). Are filarial nematode Wolbachia obligate mutualist symbionts? Trends Ecol Evol *19*, 163-166.

Fernandez, L., Langa, S., Martin, V., Maldonado, A., Jimenez, E., Martin, R., and Rodriguez, J.M. (2013). The human milk microbiota: Origin and potential roles in health and disease. Pharmacol Res *69*, 1-10.

Ferree, P.M., Frydman, H.M., Li, J.M., Cao, J., Wieschaus, E., and Sullivan, W. (2005). Wolbachia utilizes host microtubules and Dynein for anterior localization in the Drosophila oocyte. PLoS Pathog *1*, e14.

Ferri, E., Bain, O., Barbuto, M., Martin, C., Lo, N., Uni, S., Landmann, F., Baccei, S.G., Guerrero, R., de Souza Lima, S.*, et al.* (2011). New insights into the evolution of Wolbachia infections in filarial nematodes inferred from a large range of screened species. PLoS One *6*, e20843.

Fieseler, L., Horn, M., Wagner, M., and Hentschel, U. (2004). Discovery of the novel candidate phylum "Poribacteria" in marine sponges. Appl Environ Microbiol *70*, 3724-3732.

Fiscella, K. (1996). Racial disparities in preterm births. The role of urogenital infections. Public Health Rep *111*, 104-113.

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*, e121.

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

Frydman, H.M., and Spradling, A.C. (2001). The receptor-like tyrosine phosphatase lar is required for epithelial planar polarity and for axis determination within drosophila ovarian follicles. Development *128*, 3209-3220.

Fukatsu, T., and Hosokawa, T. (2002). Capsule-transmitted gut symbiotic bacterium of the Japanese common plataspid stinkbug, Megacopta punctatissima. Appl Environ Microbiol *68*, 389-396.

Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., Tobe, T., Clarke, J.M., Topping, D.L., Suzuki, T.*, et al.* (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature *469*, 543-547.

Gadau, J., Page, R.E., and Werren, J.H. (2002). The genetic basis of the interspecific differences in wing size in Nasonia (Hymenoptera; Pteromalidae): major quantitative trait loci and epistasis. Genetics *161*, 673-684.

Gaj, T., Gersbach, C.A., and Barbas, C.F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol *31*, 397-405.

Gantois, I., Ducatelle, R., Pasmans, F., Haesebrouck, F., Gast, R., Humphrey, T.J., and Van Immerseel, F. (2009). Mechanisms of egg contamination by Salmonella Enteritidis. FEMS Microbiol Rev *33*, 718-738.

Gavotte, L., Henri, H., Stouthamer, R., Charif, D., Charlat, S., Bouletreau, M., and Vavre, F. (2007). A Survey of the bacteriophage WO in the endosymbiotic bacteria Wolbachia. Mol Biol Evol *24*, 427-435.

Gavotte, L., Vavre, F., Henri, H., Ravallec, M., Stouthamer, R., and Bouletreau, M. (2004). Diversity, distribution and specificity of WO phage infection in Wolbachia of four insect species. Insect Mol Biol *13*, 147-153.

Gerardo, N.M., Altincicek, B., Anselme, C., Atamian, H., Barribeau, S.M., de Vos, M., Duncan, E.J., Evans, J.D., Gabaldon, T., Ghanim, M.*, et al.* (2010). Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome Biol *11*, R21.

Gibbons, L., Belizan, J.M., Lauer, J.A., Betran, A.P., Merialdi, M., and F., A. (2010). The global numbers and costs of additionally needed and unnecessary Cesarean sections performed by year: overuse as a barrier to universal coverage. World Health Report (2010) Background Paper *30*.

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M., and Nelson, K.E. (2006). Metagenomic analysis of the human distal gut microbiome. Science *312*, 1355-1359.

Gingras, A.C., Kennedy, S.G., O'Leary, M.A., Sonenberg, N., and Hay, N. (1998). 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. Genes Dev *12*, 502-513.

Goettler, W., Kaltenpoth, M., Herzner, G., and Strohm, E. (2007). Morphology and ultrastructure of a bacteria cultivation organ: the antennal glands of female European beewolves, Philanthus triangulum (Hymenoptera, Crabronidae). Arthropod Struct Dev *36*, 1-9.

Goffredi, S.K., and Barry, J.P. (2002). Species-specific variation in sulfide physiology between closely-related Vesicomyid clams. Mar Ecol Prog Ser *225*, 227-238.

Goldenberg, R.L., Hauth, J.C., and Andrews, W.W. (2000). Intrauterine infection and preterm delivery. N Engl J Med *342*, 1500-1507.

Goncalves, L.F., Chaiworapongsa, T., and Romero, R. (2002). Intrauterine infection and prematurity. Ment Retard Dev Disabil Res Rev *8*, 3-13.

Gosalbes, M.J., Llop, S., Valles, Y., Moya, A., Ballester, F., and Francino, M.P. (2013). Meconium microbiota types dominated by lactic acid or enteric bacteria are differentially associated with maternal eczema and respiratory problems in infants. Clin Exp Allergy *43*, 198-211.

Grice, E.A., Kong, H.H., Conlan, S., Deming, C.B., Davis, J., Young, A.C., Program, N.C.S., Bouffard, G.G., Blakesley, R.W., Murray, P.R.*, et al.* (2009). Topographical and temporal diversity of the human skin microbiome. Science *324*, 1190-1192.

Gronlund, M.M., Gueimonde, M., Laitinen, K., Kociubinski, G., Gronroos, T., Salminen, S., and Isolauri, E. (2007). Maternal breast-milk and intestinal bifidobacteria guide the compositional development of the Bifidobacterium microbiota in infants at risk of allergic disease. Clin Exp Allergy *37*, 1764-1772.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology *52*, 696-704.

Harrison, R.G., and Arnold, J. (1982). A narrow hybrid zone between closely related cricket species. Evolution *36*, 535-552.

Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol *22*, 160-174.

Hentschel, U., Hopke, J., Horn, M., Friedrich, A.B., Wagner, M., Hacker, J., and Moore, B.S. (2002). Molecular evidence for a uniform microbial community in sponges from different oceans. Appl Environ Microbiol *68*, 4431-4440.

Herren, J.K., Paredes, J.C., Schupfer, F., and Lemaitre, B. (2013). Vertical transmission of a Drosophila endosymbiont via cooption of the yolk transport and internalization machinery. MBio *4: doi:10.1128/mBio.00532-12*.

Hewitt, G.M. (1993). After the ice: parallelus meets erythropus in the Pyrenees. In Hybrid Zones and the Evolutionary Process, R.G. Harrison, ed. (Oxford: Oxford University Press), pp. 140-164.

Hoedjes, K.M., Smid, H.M., Vet, L.E., and Werren, J.H. (2014). Introgression study reveals two quantitative trait loci involved in interspecific variation in memory retention among Nasonia wasp species. Heredity (Edinb) *113*, 542-550.

Hoerauf, A., Nissen-Pahle, K., Schmetz, C., Henkle-Duhrsen, K., Blaxter, M.L., Buttner, D.W., Gallin, M.Y., Al-Qaoud, K.M., Lucius, R., and Fleischer, B. (1999). Tetracycline therapy targets intracellular bacteria in the filarial nematode Litomosoides sigmodontis and results in filarial infertility. The Journal of clinical investigation *103*, 11-18.

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

Hosokawa, T., Kikuchi, Y., Meng, X.Y., and Fukatsu, T. (2005). The making of symbiont capsule in the plataspid stinkbug Megacopta punctatissima. FEMS Microbiol Ecol *54*, 471-477.

Hosokawa, T., Kikuchi, Y., Shimada, M., and Fukatsu, T. (2008). Symbiont acquisition alters behaviour of stinkbug nymphs. Biology letters *4*, 45-48.

Hosokawa, T., Koga, R., Kikuchi, Y., Meng, X.Y., and Fukatsu, T. (2010). Wolbachia as a bacteriocyte-associated nutritional mutualist. Proc Natl Acad Sci U S A *107*, 769-774.

Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754-755.

Hughes, G.L., Koga, R., Xue, P., Fukatsu, T., and Rasgon, J.L. (2011). Wolbachia infections are virulent and inhibit the human malaria parasite Plasmodium falciparum in anopheles gambiae. PLoS Pathog *7*, e1002043.

Huh, S.Y., Rifas-Shiman, S.L., Zera, C.A., Edwards, J.W., Oken, E., Weiss, S.T., and Gillman, M.W. (2012). Delivery by caesarean section and risk of obesity in preschool age children: a prospective cohort study. Archives of disease in childhood *97*, 610-616.

Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207-214.

Hunt, K.M., Foster, J.A., Forney, L.J., Schutte, U.M., Beck, D.L., Abdo, Z., Fox, L.K., Williams, J.E., McGuire, M.K., and McGuire, M.A. (2011). Characterization of the diversity and temporal stability of bacterial communities in human milk. PLoS One *6*, e21313.

Huse, S.M., Ye, Y., Zhou, Y., and Fodor, A.A. (2012). A core human microbiome as viewed through 16S rRNA sequence clusters. PLoS One *7*, e34242.

Hybiske, K., and Stephens, R.S. (2007). Mechanisms of Chlamydia trachomatis entry into nonphagocytic cells. Infect Immun *75*, 3925-3934.

Ijichi, N., Kondo, N., Matsumoto, R., Shimada, M., Ishikawa, H., and Fukatsu, T. (2002). Internal spatiotemporal population dynamics of infection with three Wolbachia strains in the adzuki bean beetle, Callosobruchus chinensis (Coleoptera: Bruchidae). Appl Environ Microbiol *68*, 4074-4080.

Inoue, R., and Ushida, K. (2003). Vertical and horizontal transmission of intestinal commensal bacteria in the rat model. FEMS Microbiol Ecol *46*, 213-219.

International Aphid Genomics, C. (2010). Genome sequence of the pea aphid Acyrthosiphon pisum. PLoS Biol *8*, e1000313.

Ivanov, II, Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V., *et al.* (2009). Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell *139*, 485-498.

Ivanov, II, Frutos Rde, L., Manel, N., Yoshinaga, K., Rifkin, D.B., Sartor, R.B., Finlay, B.B., and Littman, D.R. (2008). Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. Cell Host Microbe *4*, 337-349.

Jaenike, J. (2009). Coupled population dynamics of endosymbionts within and between hosts. Oikos *118*, 353-362.

Jaenike, J., Dyer, K.A., Cornish, C., and Minhas, M.S. (2006). Asymmetrical reinforcement and Wolbachia infection in Drosophila. PLoS Biol *4*, e325.

Jang, I.H., Chosa, N., Kim, S.H., Nam, H.J., Lemaitre, B., Ochiai, M., Kambris, Z., Brun, S., Hashimoto, C., Ashida, M., *et al.* (2006). A Spatzle-processing enzyme required for toll signaling activation in *Drosophila* innate immunity. Dev Cell *10*, 45-55.

Jimenez, E., Fernandez, L., Marin, M.L., Martin, R., Odriozola, J.M., Nueno-Palop, C., Narbad, A., Olivares, M., Xaus, J., and Rodriguez, J.M. (2005). Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. Curr Microbiol *51*, 270-274.

Jimenez, E., Marin, M.L., Martin, R., Odriozola, J.M., Olivares, M., Xaus, J., Fernandez, L., and Rodriguez, J.M. (2008). Is meconium from healthy newborns actually sterile? Res Microbiol *159*, 187-193.

Jin, L., Hinde, K., and Tao, L. (2011). Species diversity and relative abundance of lactic acid bacteria in the milk of rhesus monkeys (Macaca mulatta). J Med Primatol *40*, 52-58.

Kageyama, D., Anbutsu, H., Shimada, M., and Fukatsu, T. (2007). Spiroplasma infection causes either early or late male killing in Drosophila, depending on maternal host age. Naturwissenschaften *94*, 333-337.

Kaltenpoth, M., Goettler, W., Dale, C., Stubblefield, J.W., Herzner, G., Roeser-Mueller, K., and Strohm, E. (2006). 'Candidatus Streptomyces philanthi', an endosymbiotic streptomycete in the antennae of Philanthus digger wasps. Int J Syst Evol Microbiol *56*, 1403-1411.

Kaltenpoth, M., Goettler, W., Koehler, S., and Strohm, E. (2010). Life cycle and population dynamics of a protective insect symbiont reveal severe bottlenecks during vertical transmission. Evolutionary Ecology *24*, 463-477.

Kaltenpoth, M., Gottler, W., Herzner, G., and Strohm, E. (2005). Symbiotic bacteria protect wasp larvae from fungal infestation. Current biology : CB *15*, 475-479.

Kaltenpoth, M., Winter, S.A., and Kleinhammer, A. (2009). Localization and transmission route of Coriobacterium glomerans, the endosymbiont of pyrrhocorid bugs. FEMS Microbiol Ecol *69*, 373-383.

Katsumi, Y., Kihara, H., Ochiai, M., and Ashida, M. (1995). A serine protease zymogen in insect plasma. Purification and activation by microbial cell wall components. Eur J Biochem *228*, 870-877.

Keller, R.G., Desplan, C., and Rosenberg, M.I. (2010). Identification and characterization of *Nasonia* Pax genes. Insect Mol Biol *19 Suppl 1*, 109-120.

Kent, B.N., and Bordenstein, S.R. (2010). Phage WO of Wolbachia: lambda of the endosymbiont world. Trends Microbiol *18*, 173-181.

Kent, B.N., Salichos, L., Gibbons, J.G., Rokas, A., Newton, I.L., Clark, M.E., and Bordenstein, S.R. (2011). Complete bacteriophage transfer in a bacterial endosymbiont (Wolbachia) determined by targeted genome capture. Genome Biol Evol *3*, 209-218.

Kim, J.K., Kim, N.H., Jang, H.A., Kikuchi, Y., Kim, C.H., Fukatsu, T., and Lee, B.L. (2013). Specific midgut region controlling the symbiont population in an insect-microbe gut symbiotic association. Appl Environ Microbiol *79*, 7229-7233.

King, P.E., and Richards, J.G. (1969). Oogenesis in *Nasonia vitripennis* (Walker) (Hymenoptera: Pteromalidae). Proc Entomol Soc Lond, Ser A, Gen Entomol *44*, 143-157.

Klasson, L., Kumar, N., Bromley, R., Sieber, K., Flowers, M., Ott, S.H., Tallon, L.J., Andersson, S.G., and Dunning Hotopp, J.C. (2014). Extensive duplication of the Wolbachia DNA in chromosome four of Drosophila ananassae. BMC Genomics *15*, 1097.

Klasson, L., Walker, T., Sebaihia, M., Sanders, M.J., Quail, M.A., Lord, A., Sanders, S., Earl, J., O'Neill, S.L., Thomson, N.*, et al.* (2008). Genome evolution of Wolbachia strain wPip from the Culex pipiens group. Mol Biol Evol *25*, 1877-1887.

Koga, R., Meng, X.Y., Tsuchida, T., and Fukatsu, T. (2012). Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte-embryo interface. Proc Natl Acad Sci U S A *109*, E1230-1237.

Kondo, N., Nikoh, N., Ijichi, N., Shimada, M., and Fukatsu, T. (2002). Genome fragment of Wolbachia endosymbiont transferred to X chromosome of host insect. Proc Natl Acad Sci U S A *99*, 14280-14285.

Kondo, N., Shimada, M., and Fukatsu, T. (2005). Infection density of Wolbachia endosymbiont affected by co-infection and host genotype. Biology letters *1*, 488-491.

Koutsovoulos, G., Makepeace, B., Tanya, V.N., and Blaxter, M. (2014). Palaeosymbiosis revealed by genomic fossils of Wolbachia in a strongyloidean nematode. PLoS Genet *10*, e1004397.

Krylov, V., Tlapakova, T., and Macha, J. (2007). Localization of the single copy gene Mdh2 on Xenopus tropicalis chromosomes by FISH-TSA. Cytogenetic and genome research *116*, 110-112.

Krylov, V., Tlapakova, T., Macha, J., Curlej, J., Ryban, L., and Chrenek, P. (2008). Localization of human coagulation factor VIII (hFVIII) in transgenic rabbit by FISH-TSA: identification of transgene copy number and transmission to the next generation. Folia biologica *54*, 121-124.

Lanave, C., Preparata, G., Saccone, C., and Serio, G. (1984). A new method for calculating evolutionary substitution rates. J Mol Evol *20*, 86-93.

Lander, E.S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics *121*, 185-199.

Laue, B.E., and Nelson, D.C. (1997). Sulfur-oxidizing symbionts have not co-evolved with their hydrothermal vent tube worm hosts: an RFLP analysis. Mol Mar Biol Biotechnol *6*, 180-188.

Lawn, J.E., Cousens, S., and Zupan, J. (2005). 4 million neonatal deaths: when? Where? Why? Lancet *365*, 891-900.

Lawson, H.A., Cheverud, J.M., and Wolf, J.B. (2013). Genomic imprinting and parent-of-origin effects on complex traits. Nat Rev Genet *14*, 609-617.

Lazzaro, B.P., and Little, T.J. (2009). Immunity in a variable world. Philos Trans R Soc B-Biol Sci *364*, 15-26.

Le Clec'h, W., Braquart-Varnier, C., Raimond, M., Ferdy, J.B., Bouchon, D., and Sicard, M. (2012). High virulence of Wolbachia after host switching: when autophagy hurts. PLoS Pathog *8*, e1002844.

Le Clec'h, W., Raimond, M., Guillot, S., Bouchon, D., and Sicard, M. (2013). Horizontal transfers of feminizing versus non-feminizing Wolbachia strains: from harmless passengers to pathogens. Environ Microbiol.

Lechner, M., Marz, M., Ihling, C., Sinz, A., Stadler, P.F., and Krauss, V. (2013). The correlation of genome size and DNA methylation rate in metazoans. Theory Biosci *132*, 47-60.

Leclerc, V., Pelte, N., El Chamy, L., Martinelli, C., Ligoxygakis, P., Hoffmann, J.A., and Reichhart, J.M. (2006). Prophenoloxidase activation is not required for survival to microbial infections in *Drosophila*. EMBO Rep *7*, 231-235.

Lee, S.Y., Moon, H.J., Kurata, S., Natori, S., and Lee, B.L. (1995). Purification and cDNA cloning of an antifungal protein from the hemolymph of Holotrichia diomphalia larvae. Biol Pharm Bull *18*, 1049-1052.

Lefoulon, E., Gavotte, L., Junker, K., Barbuto, M., Uni, S., Landmann, F., Laaksonen, S., Saari, S., Nikander, S., de Souza Lima, S*., et al.* (2012). A new type F Wolbachia from Splendidofilariinae (Onchocercidae) supports the recent emergence of this supergroup. Int J Parasitol *42*, 1025-1036.

LePage, D., and Bordenstein, S.R. (2013). Wolbachia: Can we save lives with a great pandemic? Trends in Parasitology *29*, 385-393.

Ley, R.E., Turnbaugh, P.J., Klein, S., and Gordon, J.I. (2006). Microbial ecology: human gut microbes associated with obesity. Nature *444*, 1022-1023.

Li, C.W., Chen, J.Y., and Hua, T.E. (1998). Precambrian sponges with cellular structures. Science *279*, 879-882.

Li, K., Bihan, M., and Methe, B.A. (2013). Analyses of the stability and core taxonomic memberships of the human microbiome. PLoS One *8*, e63139.

Li, Y., Caufield, P.W., Dasanayake, A.P., Wiener, H.W., and Vermund, S.H. (2005). Mode of delivery and other maternal factors influence the acquisition of Streptococcus mutans in infants. J Dent Res *84*, 806-811.

Ligoxygakis, P., Pelte, N., Hoffmann, J.A., and Reichhart, J.M. (2002). Activation of *Drosophila* Toll during fungal infection by a blood serine protease. Science *297*, 114-116.

Ling, Z., Kong, J., Jia, P., Wei, C., Wang, Y., Pan, Z., Huang, W., Li, L., Chen, H., and Xiang, C. (2010). Analysis of oral microbiota in children with dental caries by PCR-DGGE and barcoded pyrosequencing. Microbial ecology *60*, 677-690.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Lo, N., Casiraghi, M., Salati, E., Bazzocchi, C., and Bandi, C. (2002). How many wolbachia supergroups exist? Mol Biol Evol *19*, 341-346.

Loehlin, D.W., Enders, L.S., and Werren, J.H. (2010a). Evolution of sex-specific wing shape at the widerwing locus in four species of Nasonia. Heredity (Edinb) *104*, 260-269.

Loehlin, D.W., Oliveira, D.C., Edwards, R., Giebel, J.D., Clark, M.E., Cattani, M.V., van de Zande, L., Verhulst, E.C., Beukeboom, L.W., Munoz-Torres, M.*, et al.* (2010b). Non-coding changes cause sex-specific wing size differences between closely related species of Nasonia. PLoS Genet *6*, e1000821.

Loehlin, D.W., and Werren, J.H. (2012). Evolution of shape by multiple regulatory changes to a growth gene. Science *335*, 943-947.

Login, F.H., Balmand, S., Vallier, A., Vincent-Monegat, C., Vigneron, A., Weiss-Gayet, M., Rochat, D., and Heddi, A. (2011). Antimicrobial peptides keep insect endosymbionts under control. Science *334*, 362-365.

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

Loker, E.S., Adema, C.M., Zhang, S.M., and Kepler, T.B. (2004). Invertebrate immune systems--not homogeneous, not simple, not well understood. Immunological reviews *198*, 10-24.

Louis, C., and Nigro, L. (1989). Ultrastructural evidence of *Wolbachia* Rickettsiales in *Drosophila simulans* and their relationships with unidirectional cross-compatibility. J Invertebr Pathol *54*, 39-44.

Lunt, D.H., Ibrahim, K.M., and Hewitt, G.M. (1998). mtDNA phylogeography and postglacial patterns of subdivision in the meadow grasshopper Chorthippus parallelus. Heredity (Edinb) *80 (Pt 5)*, 633-641.

Lynch, J.A. (2015). The expanding genetic toolbox of the wasp Nasonia vitripennis and its relatives. Genetics *199*, 897-904.

Lynch, J.A., Brent, A.E., Leaf, D.S., Pultz, M.A., and Desplan, C. (2006a). Localized maternal orthodenticle patterns anterior and posterior in the long germ wasp Nasonia. Nature *439*, 728-732.

Lynch, J.A., and Desplan, C. (2006). A method for parental RNA interference in the wasp Nasonia vitripennis. Nature protocols *1*, 486-494.

Lynch, J.A., and Desplan, C. (2010). Novel modes of localization and function of nanos in the wasp Nasonia. Development *137*, 3813-3821.

Lynch, J.A., El-Sherif, E., and Brown, S.J. (2012). Comparisons of the embryonic development of *Drosophila*, *Nasonia*, and *Tribolium*. Wiley Interdisciplinary Reviews: Developmental Biology *1*, 16-39.

Lynch, J.A., Olesnicky, E.C., and Desplan, C. (2006b). Regulation and function of tailless in the long germ wasp Nasonia vitripennis. Dev Genes Evol *216*, 493-498.

MacOwan, K.J., Atkinson, M.J., Bell, M.A., Brand, T.F., and Randall, C.J. (1984). Egg transmission of a respiratory isolate of Mycoplasma synoviae and infection of the chicken embryo. Avian pathology : journal of the WVPA *13*, 51-58.

Maldonado, J., Canabate, F., Sempere, L., Vela, F., Sanchez, A.R., Narbona, E., Lopez-Huertas, E., Geerlings, A., Valero, A.D., Olivares, M.*, et al.* (2012). Human milk probiotic Lactobacillus fermentum CECT5716 reduces the incidence of gastrointestinal and upper respiratory tract infections in infants. J Pediatr Gastroenterol Nutr *54*, 55-61.

Mandel, M.J., Ross, C.L., and Harrison, R.G. (2001). Do Wolbachia infections play a role in unidirectional incompatibilities in a field cricket hybrid zone? Mol Ecol *10*, 703-709.

Margulis, L. (1993). Symbiosis in Cell Evolution (New York: W.H. Freeman).

Martin, V., Maldonado-Barragan, A., Moles, L., Rodriguez-Banos, M., Campo, R.D., Fernandez, L., Rodriguez, J.M., and Jimenez, E. (2012). Sharing of bacterial strains between breast milk and infant feces. J Hum Lact *28*, 36-44.

Martinez-Rodriguez, P., Hernandez-Perez, M., and Bella, J.L. (2013). Detection of Spiroplasma and Wolbachia in the bacterial gonad community of Chorthippus parallelus. Microbial ecology *66*, 211-223.

Masui, S., Kamoda, S., Sasaki, T., and Ishikawa, H. (2000). Distribution and evolution of bacteriophage WO in Wolbachia, the endosymbiont causing sexual alterations in arthropods. J Mol Evol *51*, 491-497.

Masui, S., Kuroiwa, H., Sasaki, T., Inui, M., Kuroiwa, T., and Ishikawa, H. (2001). Bacteriophage WO and virus-like particles in Wolbachia, an endosymbiont of arthropods. Biochem Biophys Res Commun *283*, 1099-1104.

Maynard, C.L., Elson, C.O., Hatton, R.D., and Weaver, C.T. (2012). Reciprocal interactions of the intestinal microbiota and immune system. Nature *489*, 231-241.

McElroy, S.J., and Weitkamp, J.H. (2011). Innate Immunity in the Small Intestine of the Preterm Infant. Neoreviews *12*, e517-e526.

McFall-Ngai, M., Hadfield, M.G., Bosch, T.C., Carey, H.V., Domazet-Loso, T., Douglas, A.E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F.*, et al.* (2013). Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci U S A *110*, 3229-3236.

McGraw, E.A., Merritt, D.J., Droller, J.N., and O'Neill, S.L. (2001). Wolbachia-mediated sperm modification is dependent on the host genotype in Drosophila. Proc Biol Sci *268*, 2565-2570.

McGraw, E.A., Merritt, D.J., Droller, J.N., and O'Neill, S.L. (2002). Wolbachia density and virulence attenuation after transfer into a novel host. Proc Natl Acad Sci U S A *99*, 2918-2923.

McKnite, A.M., Perez-Munoz, M.E., Lu, L., Williams, E.G., Brewer, S., Andreux, P.A., Bastiaansen, J.W., Wang, X., Kachman, S.D., Auwerx, J.*, et al.* (2012). Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits. PLoS One *7*, e39191.

McMeniman, C.J., Lane, R.V., Cass, B.N., Fong, A.W., Sidhu, M., Wang, Y.F., and O'Neill, S.L. (2009). Stable introduction of a life-shortening Wolbachia infection into the mosquito Aedes aegypti. Science *323*, 141-144.

McNulty, S.N., Foster, J.M., Mitreva, M., Dunning Hotopp, J.C., Martin, J., Fischer, K., Wu, B., Davis, P.J., Kumar, S., Brattig, N.W.*, et al.* (2010). Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. PLoS One *5*, e11029.

Meister, M., Lemaitre, B., and Hoffmann, J.A. (1997). Antimicrobial peptide defense in Drosophila. Bioessays *19*, 1019-1026.

Menon, R., Dunlop, A.L., Kramer, M.R., Fortunato, S.J., and Hogue, C.J. (2011). An overview of racial disparities in preterm birth rates: caused by infection or inflammatory response? Acta Obstet Gynecol Scand *90*, 1325-1331.

Metcalf, J.A., and Bordenstein, S.R. (2012). The complexity of virus systems: the case of endosymbionts. Current Opinion Microbiology *15*, 546-552.

Meyer, E., Aglyamova, G.V., and Matz, M.V. (2011). Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. Mol Ecol *20*, 3599-3616.

Min, K.T., and Benzer, S. (1997). Wolbachia, normally a symbiont of Drosophila, can be virulent, causing degeneration and early death. Proc Natl Acad Sci U S A *94*, 10792-10796.

Mitchell, J. (2011). Streptococcus mitis: walking the line between commensalism and pathogenesis. Mol Oral Microbiol *26*, 89-98.

Moran, N.A., McCutcheon, J.P., and Nakabachi, A. (2008). Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet *42*, 165-190.

Moran, N.A., Munson, M.A., Baumann, P., and Ishikawa, H. (1993). A molecular clock in endosymbiotic bacteria is calibrated using insect hosts. Proc R Soc Lond *253*, 167-171.

Moriyama, M., Nikoh, N., Hosokawa, T., and Fukatsu, T. (2015). Riboflavin Provisioning Underlies Wolbachia's Fitness Contribution to Its Insect Host. MBio *6*.

Moussian, B., and Roth, S. (2005). Dorsoventral axis formation in the *Drosophila* embryo--shaping and transducing a morphogen gradient. Current biology : CB *15*, R887-899.

Mouton, L., Dedeine, F., Henri, H., Bouletreau, 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.

Mouton, L., Henri, H., Bouletreau, M., and Vavre, F. (2003). Strain-specific regulation of intracellular Wolbachia density in multiply infected insects. Mol Ecol *12*, 3459-3465.

Muchnik, L., Adawi, A., Ohayon, A., Dotan, S., Malka, I., Azriel, S., Shagan, M., Portnoi, M., Kafka, D., Nahmani, H.*, et al.* (2013). NADH oxidase functions as an adhesin in Streptococcus pneumoniae and elicits a protective immune response in mice. PLoS One *8*, e61128.

Muhlia-Almazan, A., Sanchez-Paz, A., and Garcia-Carreno, F.L. (2008). Invertebrate trypsins: a review. J Comp Physiol B *178*, 655-672.

Murgas Torrazza, R., and Neu, J. (2011). The developing intestinal microbiome and its relationship to health and disease in the neonate. J Perinatol *31 Suppl 1*, S29-34.

Newton, I.L., and Bordenstein, S.R. (2011). Correlations between bacterial ecology and mobile DNA. Curr Microbiol *62*, 198-208.

Newton, I.L., Savytskyy, O., and Sheehan, K.B. (2015). Wolbachia utilize host actin for efficient maternal transmission in Drosophila melanogaster. PLoS Pathog *11*, e1004798.

Nezis, I.P., Stravopodis, D.J., Margaritis, L.H., and Papassideri, I.S. (2006). Autophagy is required for the degeneration of the ovarian follicular epithelium in higher Diptera. Autophagy *2*, 297-298.

Niehuis, O., Buellesbach, J., Gibson, J.D., Pothmann, D., Hanner, C., Mutti, N.S., Judson, A.K., Gadau, J., Ruther, J., and Schmitt, T. (2013). Behavioural and genetic analyses of Nasonia shed light on the evolution of sex pheromones. Nature *494*, 345-348.

Niehuis, O., Bullesbach, J., Judson, A.K., Schmitt, T., and Gadau, J. (2011). Genetics of cuticular hydrocarbon differences between males of the parasitoid wasps Nasonia giraulti and Nasonia vitripennis. Heredity (Edinb) *107*, 61-70.

Nikoh, N., Hosokawa, T., Moriyama, M., Oshima, K., Hattori, M., and Fukatsu, T. (2014). Evolutionary origin of insect-Wolbachia nutritional mutualism. Proc Natl Acad Sci U S A *111*, 10257-10262.

Nikoh, N., Tanaka, K., Shibata, F., Kondo, N., Hizume, M., Shimada, M., and Fukatsu, T. (2008). Wolbachia genome integrated in an insect chromosome: evolution and fate of laterally transferred endosymbiont genes. Genome Res *18*, 272-280.

Obbard, D.J., Welch, J.J., Kim, K.W., and Jiggins, F.M. (2009). Quantifying adaptive evolution in the *Drosophila* immune system. PLoS Genet *5*, e1000698.

Ozuak, O., Buchta, T., Roth, S., and Lynch, J.A. (2014a). Ancient and diverged TGF-beta signaling components in Nasonia vitripennis. Dev Genes Evol *224*, 223-233.

Ozuak, O., Buchta, T., Roth, S., and Lynch, J.A. (2014b). Dorsoventral polarity of the Nasonia embryo primarily relies on a BMP gradient formed without input from Toll. Curr Biol *24*, 2393-2398.

Pais, R., Lohs, C., Wu, Y., Wang, J., and Aksoy, S. (2008). The obligate mutualist Wigglesworthia glossinidia influences reproduction, digestion, and immunity processes of its host, the tsetse fly. Appl Environ Microbiol *74*, 5965-5974.

Palmer, C., Bik, E.M., DiGiulio, D.B., Relman, D.A., and Brown, P.O. (2007). Development of the human infant intestinal microbiota. PLoS Biol *5*, e177.

Panaram, K., and Marshall, J.L. (2007). F supergroup Wolbachia in bush crickets: what do patterns of sequence variation reveal about this supergroup and horizontal transfer between nematodes and arthropods? Genetica *130*, 53-60.

Pannebakker, B.A., Loppin, B., Elemans, C.P., Humblot, L., and Vavre, F. (2007). Parasitic inhibition of cell death facilitates symbiosis. Proc Natl Acad Sci U S A *104*, 213-215.

Parmakelis, A., Moustaka, M., Poulakakis, N., Louis, C., Slotman, M.A., Marshall, J.C., Awono-Ambene, P.H., Antonio-Nkondjio, C., Simard, F., Caccone, A.*, et al.* (2010). *Anopheles* immune genes and amino acid sites evolving under the effect of positive selection. PLoS One *5*, e8885.

Peek, A.S., Feldman, R.A., Lutz, R.A., and Vrijenhoek, R.C. (1998). Cospeciation of chemoautotrophic bacteria and deep sea clams. Proc Natl Acad Sci U S A *95*, 9962-9966.

Penders, J., Thijs, C., Vink, C., Stelma, F.F., Snijders, B., Kummeling, I., van den Brandt, P.A., and Stobberingh, E.E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. Pediatrics *118*, 511-521.

Perez, P.F., Dore, J., Leclerc, M., Levenez, F., Benyacoub, J., Serrant, P., Segura-Roggero, I., Schiffrin, E.J., and Donnet-Hughes, A. (2007). Bacterial imprinting of the neonatal immune system: lessons from maternal cells? Pediatrics *119*, e724-732.

Perrot-Minnot, M.J., Guo, L.R., and Werren, J.H. (1996). Single and double infections with Wolbachia in the parasitic wasp Nasonia vitripennis: effects on compatibility. Genetics *143*, 961-972.

Perrot-Minnot, M.J., and Werren, J.H. (1999). *Wolbachia* infection and incompatibility dynamics in experimental selection lines. J Evol Biol *12*, 272-282.

Prado, S.S., and Zucchi, T.D. (2012). Host-symbiont interactions for potentially managing Heteropteran pests. Psyche *2012*, 1-9.

Pultz, M.A., Westendorf, L., Gale, S.D., Hawkins, K., Lynch, J., Pitt, J.N., Reeves, N.L., Yao, J.C., Small, S., Desplan, C.*, et al.* (2005). A major role for zygotic hunchback in patterning the Nasonia embryo. Development *132*, 3705-3715.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T.*, et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59-65.

Raddatz, G., Guzzardo, P.M., Olova, N., Fantappie, M.R., Rampp, M., Schaefer, M., Reik, W., Hannon, G.J., and Lyko, F. (2013). Dnmt2-dependent methylomes lack defined DNA methylation patterns. Proc Natl Acad Sci U S A *110*, 8627-8631.

Ramirez-Valle, F., Braunstein, S., Zavadil, J., Formenti, S.C., and Schneider, R.J. (2008). eIF4GI links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy. J Cell Biol *181*, 293-307.

Ramsay, D.T., Kent, J.C., Owens, R.A., and Hartmann, P.E. (2004). Ultrasound imaging of milk ejection in the breast of lactating women. Pediatrics *113*, 361-367.

Rautava, S., Collado, M.C., Salminen, S., and Isolauri, E. (2012). Probiotics modulate host-microbe interaction in the placenta and fetal gut: a randomized, double-blind, placebo-controlled trial. Neonatology *102*, 178-184.

Ravel, J., Gajer, P., Abdo, Z., Schneider, G.M., Koenig, S.S., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O.*, et al.* (2011). Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A *108 Suppl 1*, 4680-4687.

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. Evolution *63*, 165-183.

Raychoudhury, R., Desjardins, C.A., Buellesbach, J., Loehlin, D.W., Grillenberger, B.K., Beukeboom, L., Schmitt, T., and Werren, J.H. (2010). Behavioral and genetic characteristics of a new species of Nasonia. Heredity (Edinb) *104*, 278-288.

Renz-Polster, H., David, M.R., Buist, A.S., Vollmer, W.M., O'Connor, E.A., Frazier, E.A., and Wall, M.A. (2005). Caesarean section delivery and the risk of allergic disorders in childhood. Clin Exp Allergy *35*, 1466-1472.

Rescigno, M., Urbano, M., Valzasina, B., Francolini, M., Rotta, G., Bonasio, R., Granucci, F., Kraehenbuhl, J.P., and Ricciardi-Castagnoli, P. (2001). Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. Nat Immunol *2*, 361-367.

Rio, R.V., Wu, Y.N., Filardo, G., and Aksoy, S. (2006). Dynamics of multiple symbiont density regulation during host development: tsetse fly and its microbial flora. Proc Biol Sci *273*, 805-814.

Robinson, G.E., Hackett, K.J., Purcell-Miramontes, M., Brown, S.J., Evans, J.D., Goldsmith, M.R., Lawson, D., Okamuro, J., Robertson, H.M., and Schneider, D.J. (2011). Creating a buzz about insect genomes. Science *331*, 1386.

Robinson, K.M., Sieber, K.B., and Dunning Hotopp, J.C. (2013). A review of bacteria-animal lateral gene transfer may inform our understanding of diseases like cancer. PLoS Genet *9*, e1003877.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. (2007). The role of microorganisms in coral health, disease and evolution. Nat Rev Microbiol *5*, 355-362.

Rosenberg, M.I., Brent, A.E., Payre, F., and Desplan, C. (2014). Dual mode of embryonic development is highlighted by expression and function of Nasonia pair-rule genes. Elife *3*, e01440.

Round, J.L., Lee, S.M., Li, J., Tran, G., Jabri, B., Chatila, T.A., and Mazmanian, S.K. (2011). The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. Science *332*, 974-977.

Rutten, K.B., Pietsch, C., Olek, K., Neusser, M., Beukeboom, L.W., and Gadau, J. (2004). Chromosomal anchoring of linkage groups and identification of wing size QTL using markers and FISH probes derived from microdissected chromosomes in Nasonia (Pteromalidae: Hymenoptera). Cytogenetic and genome research *105*, 126-133.

Sacchi, L., Grigolo, A., Laudani, U., Ricevuti, G., and Dealessi, F. (1985). Bavohior of symbionts during oogenesis and early stages of development in the German cockroach, *Blatella germanica* (Blattodea). Journal of invertebrate pathology *46*, 139-152.

Sacchi, L., Grigolo, A., Mazzini, M., Bigliardi, E., Baccetti, B., and Laudani, U. (1988). Symbionts in the oocytes of *Blattella germanica* (L.) (Dictyoptera: Blattellidae): their mode of transmission. Journal of insect morphology and embryology *17*, 437-446.

Sackton, T.B., Lazzaro, B.P., Schlenke, T.A., Evans, J.D., Hultmark, D., and Clark, A.G. (2007). Dynamic evolution of the innate immune system in Drosophila. Nat Genet *39*, 1461-1468.

Salunke, B.K., Salunkhe, R.C., Dhotre, D.P., Khandagale, A.B., Walujkar, S.A., Kirwale, G.S., Ghate, H.V., Patole, M.S., and Shouche, Y.S. (2010). Diversity of Wolbachia in Odontotermes spp. (Termitidae) and Coptotermes heimi (Rhinotermitidae) using the multigene approach. FEMS Microbiol Lett *307*, 55-64.

Sanders, J.L., Watral, V., and Kent, M.L. (2012). Microsporidiosis in zebrafish research facilities. ILAR journal / National Research Council, Institute of Laboratory Animal Resources *53*, 106-113.

Saridaki, A., and Bourtzis, K. (2010). Wolbachia: more than just a bug in insects genitals. Curr Opin Microbiol *13*, 67-72.

Schmitt, S., Angermeier, H., Schiller, R., Lindquist, N., and Hentschel, U. (2008). Molecular microbial diversity survey of sponge reproductive stages and mechanistic insights into vertical transmission of microbial symbionts. Appl Environ Microbiol *74*, 7694-7708.

Schmitt, S., Weisz, J.B., Lindquist, N., and Hentschel, U. (2007). Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge Ircinia felix. Appl Environ Microbiol *73*, 2067-2078.

Schneider, D.S., Jin, Y., Morisato, D., and Anderson, K.V. (1994). A processed form of the Spatzle protein defines dorsal-ventral polarity in the *Drosophila* embryo. Development *120*, 1243-1250.

Schroeder, C.M., Naugle, A.L., Schlosser, W.D., Hogue, A.T., Angulo, F.J., Rose, J.S., Ebel, E.D., Disney, W.T., Holt, K.G., and Goldman, D.P. (2005). Estimate of illnesses from *Salmonella enteritdis* in eggs, United States, 2000. Emerg Infect Dis *11*, 113-115.

Sender, R., Fuchs, S., and Milo, R. (2016). Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. Cell *164*, 337-340.

Serbus, L.R., Casper-Lindley, C., Landmann, F., and Sullivan, W. (2008). The genetics and cell biology of Wolbachia-host interactions. Annual review of genetics *42*, 683-707.

Serbus, L.R., Ferreccio, A., Zhukova, M., McMorris, C.L., Kiseleva, E., and Sullivan, W. (2011). A feedback loop between Wolbachia and the Drosophila gurken mRNP complex influences Wolbachia titer. J Cell Sci *124*, 4299-4308.

Serbus, L.R., and Sullivan, W. (2007). A cellular basis for Wolbachia recruitment to the host germline. PLoS Pathog *3*, e190.

Sharp, K.H., Eam, B., Faulkner, D.J., and Haygood, M.G. (2007). Vertical transmission of diverse microbes in the tropical sponge Corticium sp. Appl Environ Microbiol *73*, 622-629.

Shoemaker, D.D., Katju, V., and Jaenike, J. (1999). *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. Evolution *53*, 1157-1164.

Shuker, D.M., King, T.M., Bella, J.L., and Butlin, R.K. (2005a). The genetic basis of speciation in a grasshopper hybrid zone. In Insect Evolutionary Biology, M. Fellowes, G. Holloway, and J. Roff, eds. (Wallingford, Oxon, UK: CABI Publishing, Oxford University Press), pp. 427-454.

Shuker, D.M., Underwood, K., King, T.M., and Butlin, R.K. (2005b). Patterns of male sterility in a grasshopper hybrid zone imply accumulation of hybrid incompatibilities without selection. Proc Biol Sci *272*, 2491-2497.

Song, H., Moulton, M.J., and Whiting, M.F. (2014). Rampant nuclear insertion of mtDNA across diverse lineages within Orthoptera (Insecta). PLoS One *9*, e110508.

Srinivas, G., Moller, S., Wang, J., Kunzel, S., Zillikens, D., Baines, J.F., and Ibrahim, S.M. (2013). Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering. Nat Commun *4*, 2462.

Steel, J.H., Malatos, S., Kennea, N., Edwards, A.D., Miles, L., Duggan, P., Reynolds, P.R., Feldman, R.G., and Sullivan, M.H. (2005). Bacteria and inflammatory cells in fetal membranes do not always cause preterm labor. Pediatr Res *57*, 404-411.

Steinert, P.M., Parry, D.A., and Marekov, L.N. (2003). Trichohyalin mechanically strengthens the hair follicle: multiple cross-bridging roles in the inner root shealth. J Biol Chem *278*, 41409-41419.

Stewart, F.J., and Cavanaugh, C.M. (2006). Bacterial endosymbioses in Solemya (Mollusca: Bivalvia)--model systems for studies of symbiont-host adaptation. Antonie Van Leeuwenhoek *90*, 343-360.

Stewart, F.J., Young, C.R., and Cavanaugh, C.M. (2008). Lateral symbiont acquisition in a maternally transmitted chemosynthetic clam endosymbiosis. Mol Biol Evol *25*, 673-687.

Strohm, E., and Linsenmair, K.E. (1995). Leaving the cradle: how beewolves (*Philanthus triangulum* F.) obtain the necessary spatial information for emergence. Zoology *98*, 137-146.

Suh, E., Mercer, D.R., Fu, Y., and Dobson, S.L. (2009). Pathogenicity of life-shortening Wolbachia in Aedes albopictus after transfer from Drosophila melanogaster. Appl Environ Microbiol *75*, 7783-7788.

Tamas, I., Klasson, L., Canback, B., Naslund, A.K., Eriksson, A.S., Wernegreen, J.J., Sandstrom, J.P., Moran, N.A., and Andersson, S.G. (2002). 50 million years of genomic stasis in endosymbiotic bacteria. Science *296*, 2376-2379.

Tang, H., Kambris, Z., Lemaitre, B., and Hashimoto, C. (2006). Two proteases defining a melanization cascade in the immune system of *Drosophila*. J Biol Chem *281*, 28097-28104.

Tavare, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. Lect Math Life Sci *17*, 57-86.

Taylor, M.J., and Hoerauf, A. (1999). Wolbachia bacteria of filarial nematodes. Parasitology today *15*, 437-442.

Taylor, M.W., Radax, R., Steger, D., and Wagner, M. (2007). Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. Microbiol Mol Biol Rev *71*, 295-347.

Thavagnanam, S., Fleming, J., Bromley, A., Shields, M.D., and Cardwell, C.R. (2008). A meta-analysis of the association between Caesarean section and childhood asthma. Clin Exp Allergy *38*, 629-633.

Tian, C., Gao, B., Fang, Q., Ye, G., and Zhu, S. (2010a). Antimicrobial peptide-like genes in *Nasonia vitripennis*: a genomic perspective. BMC Genomics *11*, 187.

Tian, C., Wang, L., Ye, G., and Zhu, S. (2010b). Inhibition of melanization by a *Nasonia* defensin-like peptide: implications for host immune suppression. J Insect Physiol *56*, 1857-1862.

Tian, G., and Cowan, N.J. (2013). Tubulin-specific chaperones: components of a molecular machine that assembles the alpha/beta heterodimer. Methods Cell Biol *115*, 155-171.

Tissier, H. (1900). Recherches sur la flore intestinale des nourrissons (état normal et pathologique) (Paris: G. Carre and C. Naud).

Tobe, S.S. (1978). Reproductive physiology of Glossina. Annu Rev Entomol *23*, 283-307.

Toomey, M.E., Panaram, K., Fast, E.M., Beatty, C., and Frydman, H.M. (2013). Evolutionarily conserved Wolbachia-encoded factors control pattern of stem-cell niche tropism in Drosophila ovaries and favor infection. Proc Natl Acad Sci U S A *110*, 10788-10793.

Tram, U., and Sullivan, W. (2002). Role of delayed nuclear envelope breakdown and mitosis in Wolbachia-induced cytoplasmic incompatibility. Science *296*, 1124-1126.

Trauer-Kizilelma, U., and Hilker, M. (2015). Insect parents improve the anti-parasitic and anti-bacterial defence of their offspring by priming the expression of immune-relevant genes. Insect Biochem Mol Biol *64*, 91-99.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. Nature *444*, 1027-1031.

Unckless, R.L., Boelio, L.M., Herren, J.K., and Jaenike, J. (2009). Wolbachia as populations within individual insects: causes and consequences of density variation in natural populations. Proc Biol Sci *276*, 2805-2811.

Vaishampayan, P.A., Dhotre, D.P., Gupta, R.P., Lalwani, P., Ghate, H., Patole, M.S., and Shouche, Y.S. (2007). Molecular evidence and phylogenetic affiliations of Wolbachia in cockroaches. Mol Phylogenet Evol *44*, 1346-1351.

van Ham, R.C., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., Fernandez, J.M., Jimenez, L., Postigo, M., Silva, F.J.*, et al.* (2003). Reductive genome evolution in Buchnera aphidicola. Proc Natl Acad Sci U S A *100*, 581-586.

Vazquez-Torres, A., Jones-Carson, J., Baumler, A.J., Falkow, S., Valdivia, R., Brown, W., Le, M., Berggren, R., Parks, W.T., and Fang, F.C. (1999). Extraintestinal dissemination of Salmonella by CD18-expressing phagocytes. Nature *401*, 804-808.

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.

Verhulst, E.C., Beukeboom, L.W., and van de Zande, L. (2010). Maternal control of haplodiploid sex determination in the wasp Nasonia. Science *328*, 620-623.

Vilcinskas, A. (2013). Evolutionary plasticity of insect immunity. J Insect Physiol *59*, 123-129.

Voronin, D., Cook, D.A., Steven, A., and Taylor, M.J. (2012). Autophagy regulates Wolbachia populations across diverse symbiotic associations. Proc Natl Acad Sci U S A *109*, E1638-1646.

Wang, J., Wu, Y., Yang, G., and Aksoy, S. (2009). Interactions between mutualist Wigglesworthia and tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission. Proc Natl Acad Sci U S A *106*, 12133-12138.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., Li, B., Cui, F., Wei, J., Ma, C.*, et al.* (2014). The locust genome provides insight into swarm formation and long-distance flight. Nat Commun *5*, 2957.

Wang, X., Werren, J.H., and Clark, A.G. (2015). Genetic and epigenetic architecture of sex-biased expression in the jewel wasps Nasonia vitripennis and giraulti. Proc Natl Acad Sci U S A *112*, E3545-3554.

Wang, X., Wheeler, D., Avery, A., Rago, A., Choi, J.H., Colbourne, J.K., Clark, A.G., and Werren, J.H. (2013). Function and evolution of DNA methylation in Nasonia vitripennis. PLoS Genet *9*, e1003872.

Waterhouse, R.M., Kriventseva, E.V., Meister, S., Xi, Z., Alvarez, K.S., Bartholomay, L.C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B.M., *et al.* (2007). Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science *316*, 1738-1743.

Webster, N.S., Taylor, M.W., Behnam, F., Lucker, S., Rattei, T., Whalan, S., Horn, M., and Wagner, M. (2010). Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. Environmental microbiology *12*, 2070-2082.

Weers, P.M., and Ryan, R.O. (2006). Apolipophorin III: role model apolipoprotein. Insect Biochem Mol Biol *36*, 231-240.

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.

Wernegreen, J.J. (2002). Genome evolution in bacterial endosymbionts of insects. Nature Reviews Genetics *3*, 850-861.

Werren, J.H. (1999). *Wolbachia* infection and incompatibility dynamics in experimental selection lines. J Evol Biol *12*, 272-282.

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

Werren, J.H., Cohen, L.B., Gadau, J., Ponce, R., Baudry, E., and Lynch, J.A. (2015). Dissection of the complex genetic basis of craniofacial anomalies using haploid genetics and interspecies hybrids in Nasonia wasps. Developmental biology.

Werren, J.H., Loehlin, D.W., and Giebel, J.D. (2009). Larval RNAi in Nasonia (parasitoid wasp). Cold Spring Harbor protocols *2009*, pdb prot5311.

Werren, J.H., Richards, S., Desjardins, C.A., Niehuis, O., Gadau, J., Colbourne, J.K., Beukeboom, L.W., Desplan, C., Elsik, C.G., Grimmelikhuijzen, C.J., *et al.* (2010). Functional and evolutionary insights from the genomes of three parasitoid Nasonia species. Science *327*, 343-348.

Werren, J.H., Zhang, W., and Guo, L.R. (1995). Evolution and phylogeny of Wolbachia: reproductive parasites of arthropods. Proc Biol Sci *261*, 55-63.

Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998). Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A *95*, 6578-6583.

Wilkinson, C.R. (1984). Immunological evidence for the Precambrian origin of bacterial symbioses in marine sponges. Proc R Soc Lond *220*, 509-517.

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*, E69.

Yan, H., Bonasio, R., Simola, D.F., Liebig, J., Berger, S.L., and Reinberg, D. (2015). DNA methylation in social insects: how epigenetics can control behavior and longevity. Annu Rev Entomol *60*, 435-452.

Yano, T., and Kurata, S. (2011). Intracellular recognition of pathogens and autophagy as an innate immune host defence. Journal of biochemistry *150*, 143-149.

Ye, J., Zhao, H., Wang, H., Bian, J., and Zheng, R. (2010). A defensin antimicrobial peptide from the venoms of *Nasonia vitripennis*. Toxicon : official journal of the International Society on Toxinology *56*, 101-106.

Yi, H.Y., Chowdhury, M., Huang, Y.D., and Yu, X.Q. (2014). Insect antimicrobial peptides and their applications. Appl Microbiol Biotechnol *98*, 5807-5822.

Yukuhiro, F., Miyoshi, T., and Noda, H. (2014). Actin-mediated transovarial transmission of a yeastlike symbiont in the brown planthopper. J Insect Physiol *60*, 111-117.

Zabal-Aguirre, M., Arroyo, F., and Bella, J.L. (2010). Distribution of Wolbachia infection in Chorthippus parallelus populations within and beyond a Pyrenean hybrid zone. Heredity (Edinb) *104*, 174-184.

Zabal-Aguirre, M., Arroyo, F., Garcia-Hurtado, J., de la Torre, J., Hewitt, G.M., and Bella, J.L. (2014). Wolbachia effects in natural populations of Chorthippus parallelus from the Pyrenean hybrid zone. J Evol Biol *27*, 1136-1148.

Zdybicka-Barabas, A., Januszanis, B., Mak, P., and Cytrynska, M. (2011). An atomic force microscopy study of Galleria mellonella apolipophorin III effect on bacteria. Biochim Biophys Acta *1808*, 1896-1906.

Zheng, Q., Cai, X., Tan, M.H., Schaffert, S., Arnold, C.P., Gong, X., Chen, C.Z., and Huang, S. (2014). Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. Biotechniques *57*, 115-124.

Zou, Z., Shin, S.W., Alvarez, K.S., Kokoza, V., and Raikhel, A.S. (2010). Distinct melanization pathways in the mosquito *Aedes aegypti*. Immunity *32*, 41-53.

Zug, R., and Hammerstein, P. (2012). Still a host of hosts for Wolbachia: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. PLoS One *7*, e38544.

APPENDIX A. SEQUENCE ACCESSION NUMBERS

Table A-1. Locus tags for WO minor capsid variants used in the *orf7* phylogeny

WO Haplotype	Minor capsid locus tag	*Wolbachia* strain	NCBI Accession #
WORiA	WRi_012630	*w*Ri	CP001391
WORiB-1	WRi_005560	*w*Ri	CP001391
WORiB-2	WRi_010220	*w*Ri	CP001391
WORiC	WRi_007170	*w*Ri	CP001391
WOSol	So0014	*w*Sol	KC955252
WOMelA	WD0271	*w*Mel	AE017196
WOMelB	WD0602	*w*Mel	AE017196
WOAu1	WPWAU_0301	*w*Au	LK055284
WOAu2	WPWAU_0654	*w*Au	LK055284
WOHa1	wHa_02460	*w*Ha	CP003884
WOHa2	wHa_03530	*w*Ha	CP003884
WOPip1	WP0252	*w*Pip (Pel)	AM999887
WOPip2	WP0311	*w*Pip (Pel)	AM999887
WOPip3	WP0326	*w*Pip (Pel)	AM999887
WOPip4	WP0426	*w*Pip (Pel)	AM999887
WOPip5	WP1303	*w*Pip (Pel)	AM999887
WOCauB1	WOCauB1_gp3	*w*CauB	AB161975
WOCauB2	WOCauB2_B2gp17	*w*CauB	AB478515
WOCauB3	WOCauB3_B3gp18	*w*CauB	AB478516
WOVitA1	ADW80142	*w*VitA	HQ906662
WOVitA2	No annotation	*w*VitA	HQ906663
WOVitA4	No annotation	*w*VitA	HQ906664
WOVitB	ADW80201	*w*VitB	HQ906665
WO-WVulC3-4	HM452368	*w*Vul	N/A
WO-WVulC6	HM452370	*w*Vul	N/A
WOTai	wTai_orf7	*w*Tai	AB036665
WONo1	wNo_01210	*w*No	CP003883
WOBol1	wBol1_1361	*w*Bol	**CAOH00000000**
Orf7 (allele 1)			KR081343
Cpar-WO1 (allele 2)			KR081342
Cpar-WO2 (allele 3)			KR081345
Cpar-WO2 (allele 4)			KR081346
Cpar-WO2 (allele 5)			KR081347
Cpar-WO3 (allele 6)			KT599860
Cpar-WO3 (allele 7)			KR081344
Cpar-WO3 (allele 8)			KT599861

APPENDIX B. PRIMER INFORMATION

Table B-1. *Nasonia* microsatellite markers

Primer name	Chr	cM*	Primer Set (5' to 3')	Size *Nvit* (bp)	Size *Ngir* (bp)	Used For
MM1.12	1	31.4	F: GCGGTCCTGCTCCATTAACCGC R: CCAGACTCGCGCGGGTGTATTT	284	242	QTL
MM1.13	1	32.9	F: AGCTCCGAGAGCGCGAGTGA R: TCCCGTGCCGACGCATACAC	224	167	QTL
MM1.14	1	35.8	F: GCCGTCGAGAGACGAGCGAG R: GCGCGGCTGGAGGATGCTTT	219	266	QTL
MM1.L521	1	38.7	F: ACACGTCCCGATCCTTCTTTGAC R: GCGCCTCACTTGTTGTGCAT	118	160	QTL
MM1.16	1	40.9	F: ACGCGACTCCTTTCTCCGCA R: GCGGAAATCGAATGCGCGGC	233	199	QTL
MM1.17	1	43.8	F: TGCCTCGCGAGAGCGCAAAA R: ACTGCTCTCGTCAAGGCCGC	177	217	QTL
NvC1-21	1	46.7	F: GTAACAGTGAGATAAATGTG R: TAGCAACGATAGTCCACG	148	N/A	QTL
MM1.057	1	49.6	F: CTACCACATCTTTCGCCAGTTT R: TCGAGTGATTAGAGATCGACGTT	180	206	QTL
MM1.L3567	1	53.3	F: CGCTCTGTCTACCTGTCCCT R: CGGCCACAAAGCAAATAGGC	154	184	QTL
MM1.31	1	56.2	F: CGCATCATCAACCCCCGACCA R: TCCGCGGCATAACCACTTGCT	266	297	QTL
MM1.32	1	57.7	F: ACCGGGACGACTTGAGCGTA R: ACAATGGGCGAATTTTTCTGCCG	183	220	QTL
MM2.13	2	19	F: AAGACGAGAGCCGACGTTGC R: GGCCTGCACGAGTGTGTATAGGG	240	206	QTL
MM2.15	2	21.9	F: TGGCAGATGACTCACGGAAATTAACAG R: CAGTTTTAGATGAGTTTATGAACTGTGTC	87	154	QTL
MM2.17	2	24.8	F: CGCCGACGTCGTTGCTGCTT R: AGCTCCACAACGGCGGCATC	143	99	QTL
MM2.20	2	29.2	F: TCTCCGTTAATTTCCAGCGCGT R: TCTTCCAATCCACGGGAAAACTGGT	207	168	QTL
Nv-20	2	30.7	F: TGACGAAGTATCCGAGAAG R: TCGAAAAACGATATTGCTCG	105	87	QTL
MM.Nasonins	2	32.1	F: GATGCGAAAGAAGGCGCACC R: ACAGGACTTTGCACGAGCGC	145	174	FM
MM2.L5217	2	32.1	F: GCGAGAGGCTATGCAAACAAG R: GCCAACGAAACATAAACACGCG	165	133	FM
MM2.L5223	2	32.1	F: AGTACATCCATCGTCGCATCG R: GCGAGTGAACGACTTCTTTGTGG	140	182	FM
MM2.L5251	2	32.1	F: AAACTGGAGGCATGAACGCG R: AACACGTCTCTACGCCGCTC	74	120	FM

MM2.L5304	2	32.9	F2: TCGCGCCTCCATTCTTTCGA R2: GACGCTCGCTACTGCACTGT	178	140	FM
MM2.26	2	32.9	F: GCATCGCGTATGCTAATCTGCCG R: GGCGGAGTGAGAGAGCGTTTCA	220	172	QTL, FM
MM2.L5327	2	33.6	F: ACGTGAAAGGCACAATAAAGCCG R: TTGCTGCGGAGAGAGGTTCG	154	123	FM
MM2.L5331	2	33.6	F: GAATCACAAGCAGATCGCGC R: TTATCCCACACCACGGCTGC	141	107	FM
MM2.L5335	2	36.5	F: CGCACGCGGTAATTGGCTTT R: TGTCCACGGCTGCGATTTGT	202	168	QTL
MM2.L5371	2	38	F: AGGCTAATTGAACGGCGGCG R: GCGCTTCCGAGGAGAATGCT	125	94	FM
MM2.L5414	2	38	F: CGCCGTACACGTCCCAATAA R: GGAGCTGCGTAGTTTCGGAG	232	194	FM
MM2.L5476	2	38	F: AGCATCACCGCACGATAAGGG R: TGACCGACGACCCATATCGC	104	137	FM
MM2.L5543	2	38	F: TTTCGTACCTCCGCCGATGC R: GCACATTTCTCGCCACAACGA	144	98	FM
MM2.L5572	2	38	F: CGCGAGTCTACAAGCGCAAC R: GGGAGGGAAATGCGAGAGCT	179	124	FM
MM2.28	2	38	F: ACGCTTACACGCTGGTGAATGAA R: ACACCGTAATGCAATTTCCCGCT	256	287	QTL, FM
MM2.L6283	2	38	F: GAGTCATTCCCCAGCAGAATCTT R: CTCATCCGCGTGAAACGAGT	183	225	FM
MM2.L6354	2	39.4	F: CAGTCGGAAGAAAGAGCGCG R: CCGAGAGCTGCCGTAAGAGA	159	127	FM
MM2.L6428	2	39.4	F: GGGTACAAGTTTGAGCGATTCTCG R: TTTCGCACCGGACGAGATTA	129	161	FM
MM2.L6480	2	39.4	F: TCCAACTGCTGAATGCAAACA R: TTGTAGTTGTTGCGCTGGGA	99	139	FM
MM2.L6542	2	39.4	F: CGGCGGGTGCAAAGTGAAA R: AAGTGTGCGTGCTTGTATCG	201	166	FM
MM2.30	2	40.9	F: TGGATGCGAGCGCGGGTTAT R: CCCATCGCTGATCCACGTTCTT	135	172	QTL, FM
MM2.L6870	2	43.8	F: GCTCTACACGGCGAAGGTCA R: CGCGCTTCTCTTTATGCCCG	140	191	QTL, FM
MM2.33	2	46	F: ACGAAACTCTGTACTGTATACTCCGGT R: CGGCGAGTCCTCGAGAGCAG	204	250	QTL, FM
MM2.36	2	49.6	F: GCCGTTGGAGAAATGTGCGGGA R: TCGCGTATATTTTCCGTAGTCACGC	178	139	QTL, FM
MM2.39	2	52.6	F: ACCGTTACAAAGCGAGCGAGAAT R: GCCGCCGCATAGCTCGATGA	161	207	QTL
MM2.40	2	54.8	F: TCCGTTTATCGCGCTTCGGACG R: CATCGGGCTGACCTTGGCCG	179	211	QTL
MM2.L7336	2	57.7	F: CATTCATCGCTCGTGTGCGC R: ACACATCTCTCCGAACGGCG	118	85	QTL
MM2.44	2	60.6	F: TCGACGGAAGCGAGGACGAG R: CTGGGCCGCAACGGTAAGCA	203	172	QTL

MM2.49	2	68.6	F: ACTGTTGCAGATGATGATGGTAATTT R: TCTGAAACATGCAACAATCAGGT	146	92	QTL
MM3.14	3	17.5	F: CTCTCGAAGCCGCGCGTGAA R: AGCCAGCTTTGCTTTCGACCG	231	206	QTL
MM3.15	3	20.4	F: ACACACGTTGTGCGGGGGTG R: GGTCGAAAATTTCTGCGCAGCCT	106	152	QTL
MM3.17	3	23.4	F: TGCGCGATGGCTGCTGTGAT R: TCGAGCGCAATAAACGCCGC	126	170	QTL
MM3.19	3	26.3	F: GCGGAAATTCTCGCCCCTGC R: TCCCATCATCAAAACGAAAAAGTCGC	177	220	QTL, FM
MM3.22	3	29.2	F: TCTCCTCCTGCTTCGGCCCC R: TCGTTCATCGTTCGTCATCGCA	116	146	QTL, FM
MM3.23	3	32.9	F: TTGAAGGGCTCATGGTCGCA R: CGCGAAACAGCGCACACG	183	219	QTL, FM
MM3.L8678	3	35	F: GCAGCCAGGGAGTGATATGCT R: AAAGGCCGACGACGAGAGAC	186	138	QTL, FM
MM3.L8756	3	36.5	F: CGCGTGTCGTGTGGACGTAA R: TCAAACATCCGCGAGAGTCGA	115	157	FM
MM3.L8813	3	37.2	F: CCGAGTGTGGGAGGTTTGACA R: TGTCAGCCGAGAATAGGCCG	177	148	FM
MM3.L8850	3	37.2	F: TGGTTGAGAGATCCACGCGA R: TCCGCGTTTACAACCAACATGG	159	206	FM
NvC3-18	3	38	F: GCCCAAATCATGCTTTCG R: GTTGTTCTTAAATGTGTATTCC	104	N/A	QTL
MM3.29	3	39.4	F: GGCCGATTTTCTCGACAGACC R: GCGAGGGAGAGCGAACGTC	241	285	QTL, FM
MM3.L10131	3	40.2	F: TGATGCGTTCTCGCCTTTCC R: CGACCGCAGAGCAACGATCA	155	204	FM
MM3.L10212	3	41.6	F: CCTCCCAAATCACTTCCGCGT R: TCAGCGCAATCGTTACCCTT	108	135	QTL, FM
Nv184	3	44.5	F: GCGTCATCGATGCATTTCTT R: TCTCGGGAGAGATTCAGTACG	209	141	QTL
MM3.L10340	3	45.3	F: CGAAACACCATTCGCAACGAGT R: TGTCGCATCGAGAACTGCA	194	167	FM
MM3.29.7M	3	45.3	F: CCAGTTGGATAATTCTTGAGGTCTTTC R: ACTTTGCTTGGCCCGACGAT	148	118	FM
MM3.35	3	46.7	F: GTACGTGAACCGGAAGTGTTT R: GACGGCTGCTACCGGCTATA	111	161	QTL, FM
MM3.36	3	48.2	F: ATTCGCGCCGCGGCTAATGG R: TTCCATACGTGTGGCAGGCG	150	197	FM
MM3.37	3	50.4	F: ACAAGCTTCGCACACACCGCA R: CGGTCGAAGAAGCGTCGCACA	185	157	QTL, FM
MM3.L10502	3	54.8	F: GCGCGAAACGACGAGGAATT R: CGAGCGTCGTGTGCTCTTCT	63	94	QTL
MM3.41	3	58.4	F: ACCGTGGGTCCGTGCAAC R: GGTTTGTACTTCATCGTGAGGCAATCG	186	142	QTL, FM
MM3.L10553	3	63.5	F: GCGCTTAATTGCGTCGTGTT R: CCGGTGCGGTTTCTTCTCCT	196	234	QTL

MM3.43	3	65.7	F: CGGCTGTTTATATTCCTCACCTGACGC R: GCAGCGACGAATCAGGAAATGCG	138	158	QTL
MM3.45	3	69.4	F: CGATTATGCAAACGACGCGA R: TTCCGATCACGATTCTCTCCTT	222	168	QTL
MM3.L10661	3	73	F: CCCTCCGATTATAGATGCAAGTGTCA R: GGCAGTAGTGGCTCTCTTTGCT	159	181	QTL

* cM location based on genetic map from (Desjardins et al., 2013b)

N/A: Sequence is absent in *N. giraulti* so no PCR product is generated.

Primers were used for quantitative trait loci (QTL) mapping, fine mapping with segmental introgression lines (FM) or both

All primers were designed as part of this project with the following exceptions:

NvC1-21, Nv-20, and NvC3-18 are from (Rutten et al., 2004); Nv184 is from (Beukeboom et al., 2010).

Table B-2. RT-qPCR Primers

NCBI Gene ID	Gene Name	Name of Primer Set	Primer sequences (5' to 3')	Product Size (bp)
LOC100115795	60S Ribosomal protein L32	RP49	F2: CAAGCGTAACTGGAGGAAGC R2: CTGCTAACTCCATGGGCAAT	221
LOC100120281	Amyloid beta A4 precursor protein-binding	Amyloid QPCR	F: TGAACGAGATGACGGGCAGC R: GACGCCCATGCTACCGATGT	80
LOC100119653	Y + L amino acid transporter 2	AATrans QPCR	F: CCACGAGGGTACAGGATGCT R: ATTACCGAGGCACAGCCAGT	92
LOC100120672	Protein TANC2	TANC QPCR	F: CTGCGAGATGCCCTTCGACA R: TTATCGTGGAGTCGGCTGCG	141
LOC100118571	Serine/threonine-protein phosphatase 2A	L18571 QPCR	F: CCTGCCTCTCACTGCACTTGT R: ACCTCCTGCAAACGATCTAATGC	111
LOC100120845	Band4.1-like protein 4	Band4.1 QPCR	F: CGTCACTCCTACGGCTACGT R: AACTCGACTATCCGGGCCTC	80
LOC100120971	Protein lethal(2)essential for life	Lethal QPCR	F: AGCGAGACGAACATGGCTGG R: ATGGTGAGCACTCCGTCCGA	115
LOC100121288	Voltage-dependent calcium channel	Cac QPCR	F: ATGCCACGACAAGACCACCC R: ATAGGAGCCGCGATGGAGGA	127
LOC100120822	Kin of IRRE-like protein 3	IRRE3 QPCR	F: TGGTTCAAGGACAGCTCGCC R: GATCGTGTTGTTGGGCAGCG	87
LOC100120755	DDRGK domain-containing protein 1	DDRGK QPCR	F: CGACAGACAAAGCGTTCGCG R: TTCTCCTCTTGCGCGGCCTT	75
LOC100119248	Nephrin-like	Nephrin QPCR	F: CGAGGCCGTGAACGTGACCC R: GATGCGAGGTGGGCAGACCG	80
LOC100121917	Trichohyalin-like	Tricho QPCR	F: CGTCATGCAATCCACCGAATACGC R: GATGGCTCAGACGGCCACGG	79
LOC100122001	Uncharacterized (possible Rho GTPase)	Rho QPCR	F: CAACCCTACGACCCCCAAGC R: CGAGTGCGGCTTCTCCTTGT	85
LOC100117347	U4/U6.U5 small nuclear RNP	snRNP QPCR	F: TCCGAGACGACGCGACCGAT R: TCTCTGTGTCTGTCAATATCCCTATCGC	138
LOC100679525	Uncharacterized	L79525 QPCR	F: ACGGACTCGATAGACGGCGA R: AGTTCCGACAACAGCGACGG	79
LOC100117496	Latrophilin Cirl	Latro QPCR	F: CACCTGATCCGCGCCAACTA R: TGACGCTCCAGTCGGTGTTG	76
LOC100679322	Flocculation protein FLO11-like	LAMP QPCR	F: AGCCGTGGAAAGTGAAAGTCCT R: TATCTTCGGCGCTTCTCGGG	74
LOC100121249	Tyrosine-protein phosphatase Lar	Lar QPCR	F: AGAACGCCAAGGACGACGAC R: AGCTTCCACCGTCACCGAGA	99

Table B-3. RNAi Primers

NCBI Gene ID	Gene Name	Name of Primer Set	Primer sequences (5' to 3')	Product Size (bp)*
LOC100121917	Trichohyalin-like	Tricho RNAi	F: TAATACGACTCACTATAGGGAGACC ACTCGCCATGTCAACTCGCGCC R: TAATACGACTCACTATAGGGAGACC ACTCGGCTCGGTACTGCTCGTCT	649
LOC100115522	Kinesin A	Kinesin RNAi	F: TAATACGACTCACTATAGGGAGACC ACACGTCCAACGATGAAATGGCT R: TAATACGACTCACTATAGGGAGACC ACTTCTTCTCTCACATTACACCTCGA	516

*Product size includes the T7 promoter sequence (27 bp) added to both ends of the PCR product

APPENDIX C. DIFFERENTIAL EXPRESSION ANALYSES FOR RNA-SEQ

Table C-1. Genes differentially expressed between *N. giraulti* IntG and *N. giraulti* 16.2

NCBI Gene ID	NCBI Gene Name	Mean Reads for IntG	Mean Reads for Ngir	EdgeR Fold Change (IntG/Ngir)	EdgeR p-value (FDR-corrected)
LOC100116940	Microtubule-associated protein futsch	132.67	53.33	2.46	1.3E-09
LOC100678553	Uncharacterized	192.67	88.67	2.19	3.7E-09
LOC100124063	Organic cation transporter protein	69.67	27.00	2.58	1.8E-07
LOC100118126	Sialin-like	140.33	231.33	-1.65	3.1E-05
LOC100120326	Protein dispatched homolog 1	4.67	21.67	-4.57	0.00017
LOC100122826	Activating transcription factor of chaperone	442.67	280.00	1.59	0.00017
LOC103317304	Uncharacterized	10.33	0.33	23.15	0.00017
LOC100116993	Alpha-glucosidase	139.00	85.33	1.64	0.0021
LOC100115498	Uncharacterized	46.33	22.00	2.11	0.0065
LOC100119225	Uncharacterized	40.67	18.00	2.25	0.0074
LOC100123191	Failed axon connections	1041.33	1394.67	-1.33	0.0074
LOC100680441	Glucose dehydrogenase [FAD, quinone]	49.00	23.33	2.09	0.0091
LOC100117466	Lipase 3-like	62.33	30.67	2.02	0.010
LOC100118694	Polycomb group protein Psc-like	46.67	23.00	2.02	0.011
LOC100118065	ATP-binding cassette sub-family D	18.67	6.00	3.07	0.025
LOC100122506	PDZ domain-containing protein 2-like	221.00	158.33	1.40	0.036
LOC100119653	Y+L amino acid transporter 2	159.33	107.33	1.48	0.041
SP142	Serine protease 142	55.33	31.00	1.78	0.041
LOC100117529	b(0,+)-type amino acid transporter 1	67.67	35.67	1.88	0.043
LOC100119248	Nephrin-like	431.00	227.00	1.86	0.043
LOC100119490	Uncharacterized transmembrane protein	11.00	1.33	7.96	0.043

APPENDIX D. INSECT INNATE IMMUNITY DATABASE (IIID): AN ANNOTATION TOOL FOR IDENTIFYING IMMUNE GENES IN INSECT GENOMES[‡]

Abstract

The innate immune system is an ancient component of host defense. Since innate immunity pathways are well conserved throughout many eukaryotes, immune genes in model animals can be used to putatively identify homologous genes in newly sequenced genomes of non-model organisms. With the initiation of the "i5k" project, which aims to sequence 5,000 insect genomes by 2016, many novel insect genomes will soon become publicly available, yet few annotation resources are currently available for insects. Thus, we developed an online tool called the Insect Innate Immunity Database (IIID) to provide an open access resource for insect immunity and comparative biology research (http://www.vanderbilt.edu/IIID). The database provides users with simple exploratory tools to search the immune repertoires of five insect models (including *Nasonia*), spanning three orders, for specific immunity genes or genes within a particular immunity pathway. As a proof of principle, we used an initial database with only four insect models to annotate potential immune genes in the parasitoid wasp genus *Nasonia*. Results specify 306 putative immune genes in the genomes of *N. vitripennis* and its two sister species *N. giraulti* and *N. longicornis*. Of these genes, 146 were not found in previous annotations of *Nasonia* immunity genes. Combining these newly identified immune genes with those in previous annotations, *Nasonia* possess 489 putative immunity genes, the largest immune repertoire found in insects to date. While these computational predictions need to be complemented with functional studies, the IIID database can help initiate and augment annotations of the immune system in the plethora of insect genomes that will soon become available.

[‡] This chapter was published in *PLOS ONE* (2012) 7(9): e45125 with Shefali Setia, Rini Pauly and Seth R. Bordenstein as co-authors. Robert M. Brucker was first-author.

Introduction

The innate immune system evolved early in the evolution of multicellular life, while the adaptive immune system evolved in the ancestor of the vertebrate lineage (Cooper and Alder, 2006). Thus, in insects and other invertebrates, the innate immune system not only combats foreign invaders, but it is also employed in wound healing, stress responses, and the management of microbial symbiont populations (Beckage, 2008). The versatility of the insect innate immune response is in part championed by the ability of insects to colonize diverse ecological niches across the planet while defending against pathogens that inhabit those niches (Loker et al., 2004). Indeed, immunity genes in general evolve at a faster rate than the genome as a whole (Lazzaro and Little, 2009), which is in part explained by the persistent selective pressures posed by a flux of new pathogens.

With the advent and growth of next-generation sequencing technology, rapid genome sequencing of non-model organisms is now feasible. The "i5k" initiative, launched in 2011, aims to sequence 5,000 insect genomes by 2016 (Robinson et al., 2011), generating vast amounts of data for comparative studies among insects. Annotation of immunity genes in these novel insect genomes will not only provide valuable insight into the diverse mechanisms insects employ for defense, but may also contribute to the development of new insecticides for the control of agricultural pests. To facilitate the annotation of immunity genes in insects, including our own model system of *Nasonia* parasitoid wasps, we have generated an open-access database called the Insect Innate Immunity Database (IIID, http://www.vanderbilt.edu/IIID) to serve as a starting point for researchers interested in using comparative biology to identify potential immune genes in insects. The database contains the immune repertoires of five insect models (including *Nasonia*) that span several orders, and each gene is categorized based on the pathway it participates in and the role it plays in that pathway. The intuitive web interface allows researchers to search for specific immunity genes by name, retrieve all immunity genes in the database for a particular species, pathway or class, and find putative homologs for a gene of interest using an internal BLAST tool.

The jewel wasp *Nasonia* is a genus of haplodiploid, parasitoid wasps composed of four closely related species (Order: Hymenoptera): *N. vitripennis, N. giraulti, N. longicornis*, and *N. oneida*. *Nasonia* is a model system to study the genetics of interspecific differences including host-microbe interactions (Bordenstein and Werren, 2007; Brucker and Bordenstein, 2012b;

Chafee et al., 2011), development (Keller et al., 2010; Loehlin and Werren, 2012; Lynch et al., 2012), and behavior (Blaul and Ruther, 2011; Clark et al., 2010; Desjardins et al., 2010; Niehuis et al., 2011). Recently, the genomes of the first three species mentioned above were sequenced (Werren et al., 2010). An initial characterization of immune genes in *N. vitripennis* was conducted as part of the *Nasonia* genome project (Werren et al., 2010) using two sets of Hidden Markov Models (HMMs). The first set of HMMs was generated based on alignments of select immune-related protein families from *Aedes aegypti, Anopheles gambiae* and *Drosophila melanogaster* (Waterhouse et al., 2007), and the second set was compiled using *A. aegypti* immune genes as seeds to find orthologous genes from five vertebrate and five insect species (Werren et al., 2010). Scanning the *N. vitripennis* gene set with these HMMs produced a total of 270 putative immunity genes (http://cegg.unige.ch/nasonia_genome). This number is likely an underestimate given that not all immune genes from the three Dipteran species above were used to generate the first set of HMMs. The second set of HMMs expanded the number of species incorporated in the models but only for those immune genes present in *A. aegypti*. Furthermore, only the *N. vitripennis* genome was examined; no study has attempted to identify immune genes in the sequenced sister species, *N. giraulti* and *N. longicornis*. Using the genes within the IIID to perform homology searches against the *Nasonia* genomes, we independently describe 306 putative immune genes in each of the *Nasonia* species, of which 146 genes were not found in previous annotations of *N. vitripennis* (Werren et al., 2010).

Materials and Methods

Initial construction of the IIID

To facilitate the annotation of innate immunity genes in insects, we initially created an Insect Immunity Database (IIID) composed of the published immune repertoires of four insect models spanning several different orders: *Drosophila melanogaster,* Diptera (De Gregorio et al., 2001; Obbard et al., 2009), *Anopheles gambiae*, Diptera (Parmakelis et al., 2010; Werren et al., 2010), *Apis mellifera*, Hymenoptera (Evans et al., 2006; Waterhouse et al., 2007), and *Acrythosiphon pisum*, Hemiptera (Gerardo et al., 2010). Our criteria for inclusion were that the species have a complete, publicly-available genome sequence, that the innate immune genes have been previously identified in computational or molecular studies, and that each species has an extensive review of its global immune pathways available as a resource. Sequence

information was obtained through NCBI for the 105 immunity genes described for *Acrythosiphon pisum* (Gerardo et al., 2010), 317 genes for *Anopheles gambiae* (Christophides et al., 2002; Parmakelis et al., 2010), 379 genes for *Drosophila melanogaster* (De Gregorio et al., 2001; Obbard et al., 2009), and 174 genes for *Apis mellifera* (Evans et al., 2006; Waterhouse et al., 2007). In total, 975 genes were included in the dataset used to analyze the *Nasonia* genomes. Each gene was categorized into its primary, secondary and tertiary pathways of putative function (i.e. Toll pathway, IMD pathway, humoral response, JAK/STAT, and cell cycle regulation) and into finite classes of function based upon its putative role in an immune response. Such classes include recognition (identifying potential pathogens and stressors), signaling (communicating between recognition and response), and response (molecules that interact with the pathogen or stressor).

Comparative analysis of N. vitripennis *immunity genes*

To validate the utility of this database, we used a sequence similarity BLASTx approach to mine for putative homologs of the 975 protein sequences in the IIID within the *N. vitripennis* transcriptome (OGS v1.2). A total of 18,941 unique transcripts were obtained from NasoniaBase (http://hymenopteragenome.org/nasonia/). For the BLASTx analyses, we used the BLOSUM62 matrix with a word size of 3 and a gap cost of 11, -1. The results were filtered to only contain hits with an *E*-value < 1e-10 and a bit score ≥ 30. A total of 1206 *N. vitripennis* transcripts were similar to entries in the IIID. To eliminate redundancies in the dataset, a reciprocal BLASTx analysis for each of the 1206 *Nasonia* transcripts was conducted against each of the four insect immunity gene datasets. This analysis resulted in 306 unique immune gene identifiers in *Nasonia vitripennis*.

Analysis of N. giraulti *and* N. longicornis *immunity genes*

Since the immune genes in the sister species *N. giraulti* or *N. longicornis* had not yet been evaluated, we conducted independent BLASTn analyses of the 489 *N. vitripennis* immunity genes (IIID predictions and previously annotated immune genes) against the *N. longicornis* (NCBI assembly name Nlon_1.0) and *N. giraulti* (NCBI assembly name Ngir_1.0) scaffolds (Werren et al., 2010). The parameters for the BLASTn search are as follows: *E*-value < 1e-10, word size 11, low complexity filter, and a gap cost 5, -2. For each species, best hits for the 489

genes were manually assessed as to the *E*-value and bit score, as previously described above, and nucleotide sequences were compiled for each gene in *N. giraulti* and *N. longicornis.*

Results and Discussion

The initial IIID was compiled using the immune repertoires of *D. melanogaster*, *A. gambiae*, *A. pisum*, and *A. mellifera* for a combined total of 975 genes. Using this dataset to perform homology searches against the *N. vitripennis* transcriptome, we identified 306 putative immune genes. 138 of these genes were previously reported as immune genes in the *Nasonia* genome (Nvit_1.2) paper, which identified a total of 270 putative immune genes using HMMs for protein domains common in immunity gene families (Werren et al., 2010). We also manually searched the *N. vitripennis* official gene set (v1.2) and the *Nasonia* literature (Tian et al., 2010a; Tian et al., 2010b; Ye et al., 2010) for genes with annotations similar to those of conserved immunity genes in other insect species. In total, we found 66 genes from our manual search that were not reported in Werren et al*.,* 2010. Importantly, 146 of the 306 genes identified using the IIID were not previously described in any of the *Nasonia* literature. Furthermore, using the IIID, we were able to assign names to 28 genes that were not previously annotated in the *N. vitripennis* gene set (Nvit_1.2). Conversely, a total of 183 immune genes identified previously in the *Nasonia* literature are absent from the IIID analyses of the *N. vitripennis* genome (see discussion).

Combining the immune genes identified using the IIID with the additional genes described in the literature, *N. vitripennis* possesses a total of 489 putative immunity genes. This is the largest predicted immune repertoire found in insects to date. None of the genes found in *N. vitripennis* were missing in either *N. giraulti* or *N. longicornis*.

Using the IIID, we increased the putative *Nasonia* immune repertoire by 58% in comparison to the number of immune genes originally published in the *Nasonia* genomes (Werren et al., 2010), while only finding 46% of the immune genes originally published. The missing genes are of interest. It is important to note that the *Nasonia* immune gene set in the genome sequence was identified using Hidden Markov Models (HMMs) that search for genes with protein domains common in immunity genes. One problem with this approach is that all members of a gene family with an immunity-related protein domain may not have a biological role in innate immunity if this domain can also function in other processes. Thus, using only

HMMs to find immune genes will increase the likelihood of false positives for any given protein family in which only a subset of its members are involved in immune pathways. For example, sixty-four of the innate immunity genes in the original *Nasonia* genome annotation are not found in our annotation using the IIID; these genes are classified as serine proteases. Several serine proteases play important roles in insect innate immune pathways, specifically the Toll pathway and the prophenoloxidase signaling cascade leading to melanization (Jang et al., 2006; Katsumi et al., 1995; Leclerc et al., 2006; Ligoxygakis et al., 2002; Tang et al., 2006; Waterhouse et al., 2007; Zou et al., 2010). However, the serine protease family is highly diverse, and most of its members function in other aspects of insect physiology (Chasan and Anderson, 1989; Moussian and Roth, 2005; Muhlia-Almazan et al., 2008; Schneider et al., 1994). A HMM that identifies conserved serine protease domains may simply find any serine protease, regardless of its biological function or relevance to insect immunity. Using the IIID for sequence similarity searches partially avoids this source of error because the search is performed using an entire gene, not just a protein domain, which has been identified as part of the innate immune system in another insect species. For example, the IIID predictions identified only 38 serine proteases while the HMMs found 97 serine proteases. Nevertheless, further experimental approaches are needed to determine whether the genes that we have identified actually function in the *Nasonia* immune system.

The other obvious limitation of using a sequence similarity based approach to find immune genes in a specific gene set is that the analysis misses any species-specific genes. For example, thirty-nine genes from our manual search of the literature (that were not detected by the BLASTx analysis) are antimicrobial peptides (AMPs) unique to the *Nasonia* genus, which were predicted computationally based on structural properties common to AMPs (Tian et al., 2010a; Tian et al., 2010b). Sequence similarity searches are also constrained by the reference species used to generate the database. Genes in the *Nasonia* immune repertoire present in an insect species not in the IIID would also be missed, although they are not unique to *Nasonia.*

In total, 489 unique genes have been described as potential immune genes in *N. vitripennis* when all previously published studies (Chasan and Anderson, 1989; Moussian and Roth, 2005; Tian et al., 2010a; Zou et al., 2010), manual annotations, and sequence similarity searches using the IIID are combined. To our knowledge, this list is the most complete set of insect immunity genes

currently available and the first to include those from *N. giraulti* and *N. longicornis*. While future studies are needed to confirm the functionality of these genes in the *Nasonia* immune response, the list will provide a stepping-stone for comparative analyses within the *Nasonia* genus and between *Nasonia* and other insect species. More importantly, the IIID will provide one more tool in the efforts to annotate complete immune gene repertoires in other insect genomes. Based on our investigation, we recommend the use of multiple annotation tools that will provide the most comprehensive set of predictions *in silico*, which can then be analyzed for their biological role *in vivo*.

APPENDIX E. PROTOCOLS

Modified Puregene DNA purification protocol

Cell Lysis

1. Grind *Nasonia* adults with pestle in liquid nitrogen. If working with larvae, pupae or embryos, do not use liquid nitrogen.

2. Add 100 ul Cell Lysis Solution and 0.5 ul Proteinase K to the tissue and incubate at 55C for 3 hours in the shaking water bath.

Protein Precipitation

3. Add 33 ul Protein Precipitation Solution and vortex vigorously for 20 s at high speed.

4. Centrifuge for 10 min at 16,000 *x g*.

5. Incubate on ice block for 5 min then repeat centrifugation (10 mins at 16,000 x *g*).

DNA Precipitation

6. Pipet 100 ul 100% isopropanol into a clean 1.5 ml tube. Add the supernatant from the previous step by pouring carefully. Be sure the protein pellet is not dislodged during pouring. Mix by inverting gently 50 times.

7. Centrifuge 1 min at 16,000 *x g*. Carefully discard the supernatant, and drain the tube by inverting on a clean piece of absorbent paper, taking care that the pellet remains in the tube.

8. Add 100 ul 70% ethanol and invert several times to wash the DNA pellet. Centrifuge for 1 min at 16,000 *x g*.

9. Carefully discard the supernatant. Drain the tube on a clean piece of absorbent paper, taking care that the pellet remains in the tube. Allow to air dry for 10 minutes.

DNA Hydration

10. Add 50 ul DNA Hydration Solution for a single individual, 100 ul for a pool of several *Nasonia*. Leave at room temperature overnight.

RNA extraction and cDNA synthesis

RNA extraction using the Direct-zol RNA Miniprep kit (Zymo Research)

1. If sample is in RNAlater, add approx. 2 volumes of RNase-free PBS and centrifuge at max speed for 1 min. Remove as much RNAlater as possible.

2. If extracting from soft samples (like tissues, larvae, pupae, etc.), directly crush with a new pestle in the chemical hood. If extracting from frozen, whole insects, crush with the vial in liquid nitrogen.

3. Add 200 ul Trizol reagent and continue crushing with the pestle. Remove pestle and let stand for 10 mins.

4. Centrifuge samples at max speed for 2 mins then transfer the supernatant to a fresh RNase-free Eppendorf tube.

5. Add 200 ul 100% ethanol. Mix well then transfer to a column in a collection tube. Centrifuge at max speed for 1 min. Transfer to a new collection tube.

6. Add 400 ul RNA PreWash buffer to each column. Centrifuge for 1 min and discard flow-through. Repeat.

7. Add 700 ul RNA Wash Buffer to each column. Centrifuge for 1 min and discard flow-through.

8. Centrifuge empty column for 2 mins.

9. Transfer column to an RNase-free Eppendorf tube. Add 30 ul of RNase-free water to the column and centrifuge for 1 min at max speed. Discard column.

DNase treatment of RNA with DNA-free kit (Thermo Fisher)

10. Add 3 ul of 10X DNase buffer and 1 ul DNase to the RNA sample. Incubate at 37˚C for 1 hour. **Important: Make sure to use the "DNA-free" kit, NOT the "TURBO DNA-free" kit.**

11. Add 3 ul DNase inactivation reagent and mix well. Incubate at room temperature for 5 mins, vortexing occasionally.

12. Centrifuge at 10,000 x g for 1.5 mins and carefully transfer the supernatant to a new RNase-free Eppendorf tube.

13. Run a PCR to make sure there is no DNA contamination left in the sample. If you do see bands, then repeat steps 10 - 13 until the PCR is clean (no bands) in the RNA samples.

cDNA synthesis using SuperScript VILO Mastermix (Invitrogen)

14. For each sample, combine the following components in a sterile PCR tube:

Component	Volume
SuperScript VILO Mastermix	4 ul
RNA (up to 2.5 ug)	x ul
RNase-free water	(16 – x) ul
Total	20 ul

15. Place PCR tubes in a thermocycler and run with the following PCR program settings:
    25˚C for 10 mins, 42˚C for 60 mins, 85˚C for 5 mins, 10˚C hold

RNAi with *Nasonia* pupae

Designing primers for dsRNA synthesis

1. For each gene of interest, design and order primers that will amplify a band between 500 and 1000 bp long for a cDNA template**.**
    - If you plan on using genomic DNA as your template, make sure your amplified product does not include any introns.
    - If you are working with more than one *Nasonia* species, make sure the primers you design will work for both species. This applies to the QPCR primers as well.

2. After extracting the primer sequences from Geneious, add the T7 promoter sequence to the **beginning (5' end)** of each primer. Do NOT reverse complement the T7 sequence for the reverse primer. These will be your RNAi primers.
    T7 promoter sequence: TAATACGACTCACTATAGGGAGACCAC

3. For each gene of interest, design and order primers that will amplify a band between 75 and 150 bp long with a cDNA template**.** These will be your QPCR primers to check efficiency of gene knockdown after dsRNA injections.

Making the PCR template for dsRNA synthesis

4. Resuspend lyophilized primer master stocks in Low TE Buffer to a concentration of 100 uM. Dilute working stock of each RNAi primer to 1 uM.

5. Run a 15 ul PCR reaction with your RNAi primers and 1.5 ul *Nasonia* cDNA as your template to make sure your primers work as expected. Run PCR products on a 1% SBA agarose gel at 300V for 20 minutes then stain in GelRed.
    - In general, a normal PCR program with an annealing Tm of 55˚C and extension time of 1 min for 40 cycles will work well.
    - If you see non-specific binding or a faint band with strong primer dimers, increase the annealing Tm.
    - If you are still seeing strong primer dimers, calculate the annealing temp for the gene specific portion of the primer and the whole RNAi primer separately. Run the first 5 cycles with the gene specific Tm and the last 35 cycles with the whole RNAi primer Tm.

6. Once you know that your primers work well, run the same PCR again except with 10 ul of cDNA as template in a total volume of 100 ul. Run on a 1% SBA agarose gel in 4 separate wells loaded with 25 ul each of the PCR product at 300V for 20 mins.
    - If you will be knocking-down the same gene in multiple species of *Nasonia*, you will need to run a separate reaction with cDNA from each species in order to make species-specific dsRNA. This does not matter as much if the genes are highly conserved with very little variation between the species.

7. After staining with GelRed, cut out two bands at a time with a razorblade and put in a 1.5 ml RNase-free Eppendorf tube. Each PCR product will thus have two Eppendorf tubes. From this point forward until after the dsRNA synthesis step, try to keep everything RNase-free.

8. Use the Qiagen QIAquick Gel Extraction kit to purify your PCR product. All centrifuge steps are performed at 17,900 x g:
    - Weigh each Eppendorf tube on the analytical scale after taring with an empty Eppendorf tube and write down weight in mg.
    - Add 3 volumes of Buffer QG to 1 volume gel weight (100 mg gel = 100 ul) for each sample and incubate at 50˚C for 10 mins.
    - Add 1 volume (of the original gel weight) of 100% isopropanol and mix.
    - Add up to 700 ul of the sample to a QIAquick spin column. Spin for 1 min then discard flow-through. Repeat this step until all the sample has run through the column.

- **IMPORTANT: Run both Eppendorf tubes of the same PCR product on the same column!**
- Add 500 ul of Buffer QG to the column and spin for 1 min. Discard flow-through.
- Add 750 ul Buffer PE to the column. Let column stand for 5 mins. Spin for 1 min and discard flow-through.
- Centrifuge again for 1 min to remove residual wash buffer.
- Place the column in a clean RNase-free 1.5 ul Eppendorf tube. Add 30 ul of Buffer EB and let column stand for 4 mins. Spin for 1 min then discard column.

9. Measure the concentration of your purified PCR product using the Nanodrop or the Qubit with the broad range dsDNA assay kit. If the concentration is greater than 50 ng/ul, then move on to the next step. If the concentration is less than 50 ng/ul, repeat steps 6 – 8 using the purified PCR product as your template instead of cDNA.

Synthesizing and purifying the dsRNA

10. Using the MEGAscript RNAi Kit, thaw all of the following reagents at RT except for the T7 Enzyme Mix, which should always be kept on ice. Set up the following dsRNA synthesis reaction in an RNase-free Eppendorf tube in the following order:

Component	Volume
PCR template	8 ul
10X T7 Reaction buffer	2 ul
ATP Solution	2 ul
CTP Solution	2 ul
GTP Solution	2 ul
UTP Solution	2 ul
T7 Enzyme Mix	2 ul
Total	20 ul

11. Incubate at 37˚C for at least 6 hours. Spin down every couple of hours to avoid evaporation.

12. Perform a nuclease digest using the RNAi kit to remove ssRNA and DNA from your newly synthesized dsRNA. Incubate at 37˚C for 1 hour. **Important: Do not perform this step or any further steps at the RNA bench to prevent contaminating the bench with RNase.**

Component	Volume
dsRNA from step 11	20 ul
Nuclease-free water	21 ul
10X Digestion Buffer	5 ul
DNase I	2 ul
RNase	2 ul
Total	50 ul

13. Purify the dsRNA using the RNAi kit by mixing the following components together:

Component	Volume
dsRNA from step 12	50 ul
10X Binding Buffer	50 ul
Nuclease-free water	150 ul
100% Ethanol	250 ul
Total	500

14. Pipet the 500 ul dsRNA binding mix unto a column provided by the kit. Centrifuge at max speed for 2 mins. Discard flow-through and replace the Collection Tube.

15. Add 500 ul of Wash Solution to the column. Spin at max speed for 2 mins. Discard flow-through. Repeat.
16. Centrifuge at max speed for 30 sec to remove residual wash solution.

17. Transfer the column to a fresh collection tube. Apply 50 ul of Elution Solution to the column, incubate at 65˚C for 2 mins, then spin at max speed for 2 mins. Discard the column.

18. Measure the concentration of your dsRNA using the Nanodrop. Make sure to blank with regular TE Buffer then measure as "DNA". Take the A_{260} number and multiply by 40 to get the concentration of the dsRNA in ng/ul.
    - I usually get a concentration around 1 ug/ul (approx. 50 ug dsRNA total).
    - If you need a higher concentration of dsRNA for injections, perform a DNA precipitation and resuspend the dsRNA in a smaller volume of elution buffer.
    - Tips for if you did not get enough dsRNA:
        - Incubate the dsRNA synthesis longer, up to 16 hours. Instead of doing an overnight incubation, freeze the sample at -20C overnight and then resume the synthesis reaction the next day.
        - Use a higher concentration of PCR template in the dsRNA synthesis reaction.

- Check that the kit is working properly using the positive control template DNA provided in the kit.

19. To check the quality of the dsRNA, make a 1:40 dilution of your dsRNA. Take 5 ul of the dilution and mix with 1 ul loading dye. Run on a 1% SBA agarose gel at 300V for 20 minutes. After staining, you should see a single band at the expected size. A lot of smearing of dsRNA smaller than the expected size indicates a low-quality preparation.

Making a GFP dsRNA control

To ensure that any effects on *Wolbachia* density are due to knocking down your candidate gene instead of the introduction of dsRNA in general, you must have a control dsRNA that targets a gene not found in the *Nasonia* genome. I use GFP from a bacterial plasmid.

20. Repeat steps 6 through 19 using the pGreen plasmid (Carolina #21-1449) as your PCR template with the following RNAi primers:

GFP RNAi F1:
TAATACGACTCACTATAGGGAGACCACGTGGAGAGGGTGAAGGTGATGC
GFP RNAi R1:
TAATACGACTCACTATAGGGAGACCACGGGCAGATTGTGTGGACAGGT

Injecting Nasonia pupae with dsRNA

21. Collect yellow female pupae, avoiding those that are white (too young) or have really dark eyes (too old).

22. Put double-sided sticky tape on a glass slide. With a paintbrush, place the *Nasonia* in a line facing the same direction and **gently** press down to adhere to tape. I usually put two rows of 10 on a single slide. Place the slides in a *Drosophila* vial (*Nasonia* stock vials are too small) for safe keeping.

23. Prepare your dsRNA for injection by mixing 4 ul of the dsRNA (gene of interest and GFP control) with 1 ul food coloring dye diluted 1:10,000 in 1X TE Buffer. Make an injection buffer control by mixing 4 ul of 1X TE Buffer with 1 ul of the dye solution.

24. Set up the Nanoject II:
    - Plug in the injector, the power supply and the foot pedal into the control box.
    - Fill a syringe with mineral oil then attach the needle provided with the Nanoject to the syringe.
    - Gently break the very end off of a glass capillary needle and fill it with mineral oil. Make sure there are no bubbles.

- Attach the glass needle to the injector by unscrewing the black plastic part, placing the plunger through the needle and pushing the needle down until it goes through the first black O-ring and sits on the white plastic spacer. Tighten the black plastic part.
- Press the "empty" button on the control box to push some of the oil out of the needle.

25. Pipet the 5 ul of Buffer control onto the top of a small, plastic petri dish. Stick the tip of the glass needle into the liquid and press "fill" on the control box to suck up the liquid. If you get any air bubbles, press "empty" until you push them out then repeat the process until you have enough liquid in the needle.

26. Press the foot pedal a couple times to make sure that a small amount of liquid is released each time. If no liquid is coming out, you may need to break more of the needle tip off.

27. Inject the *Nasonia* by sticking the needle into the side of their abdomen then pressing down on the foot pedal. You should see the abdomen acquire a slight blue tint.

28. Repeat until you've injected all the females for the buffer control. Remove the glass needle and repeat steps 24 through 27 with the dsRNA against the GFP control and dsRNA against the candidate gene.

29. After all injections are complete, place slides into the *Drosophila* tubes (one slide per tube). Take a cotton plug, moisten it with water, and then place on the side **opposite** the pupae. This will help prevent the pupae from drying out. Plug the vials with *Nasonia* plugs or cotton, then place in the incubator.

Collecting embryos from injected females

30. Once the injected pupae start turning black, carefully remove them from the double-sided sticky tape and transfer them to a regular glass vial. Do not let them emerge on the tape or their wings will get stuck.

31. The day that they emerge as adults is Day 1. Put each female into a separate vial and give her two hosts and a drop of honey. Wait 48 hours.

32. On the morning of Day 3, transfer each female to a new vial and give her one host in a foam plug with the head of the host facing out. Let her lay eggs for 5 hours.

33. After 5 hours, open the anterior part of the host and look for any eggs. Transfer eggs to an Eppendorf tube with a probe.

- Make sure the eggs make it inside the Eppendorf tube since static will often cause the eggs to end up anyone but the inside of the tube.
- Record the number of eggs inside each tube. You will need more than 5 eggs to get a decent QPCR signal. I aim for 10 eggs per tube since *N. giraulti* often does not lay more than that in a 5-hour window. If you are using *N. vitripennis*, you may be able to get 20 eggs per tube.
- Place the eggs in the -80C freezer.

34. Give each female 2 hosts and leave overnight. Use these hosts for pupae collections later if you want to test pupae in addition to embryos.

35. Repeat steps 32 – 34 on Day 4.

36. On Day 5, knock out each female with CO_2, cut off her abdomen with a razor blade, transfer the abdomen to an RNase-free 1.5 ml Eppendorf tube, and immediately submerge in liquid nitrogen. Repeat for all females then store the abdomens at -80C.

qPCR for Wolbachia titers of embryos from injected females

37. Extract DNA from embryos using the reagents from Qiagen's Puregene kit as follows. **Note: This protocol was modified from the DNA extraction protocol above in an attempt to get the sticky embryos off the side of the Eppendorf tube.**
    - Add 200 ul Cell Lysis Buffer then vortex at maximum speed for 5 seconds.
    - Centrifuge samples at maximum speed for 1 min.
    - Use a pestle to crush the eggs, making sure to get the sides of the tube where eggs may still be lingering.
    - Add 1 ul Proteinase K then incubate for 3 hours at 55˚C.
    - Add 66 ul Protein Precipitation Solution, vortex at max speed for 20 secs, then centrifuge at 16,000 x g for 10 mins.
    - Place on ice block for 5 mins
    - Centrifuge at 16,000 x g for 10 mins.
    - Transfer the supernatant to a new tube filled with 200 ul 100% isopropanol. Invert 50 times then centrifuge for 1 min at 16,000 x g.
    - Pour out the isopropanol, tap the tube on a paper towel, then add 100 ul 70% ethanol. Invert 10 times then centrifuge for 1 min at 16,000 x g.
    - Pour out the ethanol and invert the tube on a paper towel. Leave inverted for 10 mins.
    - Add 20 ul DNA hydration solution and leave overnight at room temperature.

38. Run a qPCR with *Wolbachia* primers **NvWGroQTF1** and **NvWGroQTR1.** Use 2 ul of DNA in a total reaction volume of 25 ul and run with the Standard QPCR Protocol. **Note: If you're working with a low-titer line like *N. vitripennis* 12.1, you may need to increase the DNA template to 4 ul unless you collected more than 10 eggs per tube.**

39. Calculate *Wolbachia* numbers using the Cq values and your standard curve equation, discarding any samples with Cq values higher than 30. Divide # of *Wolbachia* cells by number of eggs in each tube.

RT-qPCR of female abdomens to determine percent knockdown of GOI

40. Run a qPCR with your gene-specific QPCR primers and general cDNA template with a temperature gradient from 55˚C to 65˚C to see which annealing temperature produces the best amplification and sharpest melt curve.

41. For each abdomen sample, extract the RNA, DNase treat the RNA and convert into cDNA. Make sure the same amount of RNA (in ng) is added to the cDNA synthesis step for each sample.

42. Run a qPCR where each sample is tested with your QPCR primers against the gene that was knocked down as well as with primers for a housekeeping control gene: **RP49 F2** and **RP49 R2**. Use 2 ul of the cDNA in a total reaction volume of 25 ul with the annealing temperature that is optimal for your QPCR primers. **Note: If the annealing temp is several degrees different than 59C (the optimum for RP49), the gene of interest and the control gene may need to be run on different plates.**

43. To calculate the percent knockdown, follow the excel table outlined below. Basically, for each sample, subtract its RP49 Cq from its GOI Cq to get the ΔCq then calculate its 2^(-ΔCq) value. Average the 2^(-ΔCq) replicate values for each of the treatment groups, then normalize the experimental average to the control average to find the percent knockdown.

1	A	B	C	D	E	F	G
2	Sample	GOI Cq	RP49 Cq	ΔCq	2^-ΔCq	Avg.	% knockdown
3	Buffer 1			=B3-C3	=2^-D3	=average(E3:E4)	=(1-F3/F3)*100
4	Buffer 2			=B4-C4	=2^-D4		
5	GFP 1			=B5-C5	=2^-D5	=average(E5:E6)	=(1-F5/F3)*100
6	GFP 2			=B6-C6	=2^-D6		
7	GOI 1			=B7-C7	=2^-D7	=average(E7:E8)	=(1-F7/F3)*100
8	GOI 2			=B8-C8	=2^-D8		

Antibody staining of *Wolbachia* in *Nasonia* embryos

Collecting and fixing embryos

1. Remove embryos from host with probe and transfer to a glass vial with 5 ml heptane. Shake for 2 mins. Add 5 ml methanol to glass vial. Shake for another 2 mins.

2. Transfer any embryos that fall to the bottom of the vial to a 1.5 ml eppendorf tube. Store embryos in methanol at 4C if not staining immediately.

Staining embryos

3. Rehydrate embryos stepwise by washing in:
    - 90% methanol / 10% PBST for 1 min
    - 75% methanol / 25% PBST for 1 min
    - 50% methanol / 50% PBST for 1 min
    - 25% methanol / 75% PBST for 1 min
    - 100% PBST for 5 mins (twice)

4. Block in PBST-BSA (PBST + 0.2% BSA) for 30 minutes then in PBANG (PBST-BSA + 5% normal goat serum) for 1 hour.

5. Incubate with 1 mg/ml RNase diluted in PBANG for 2 hours at RT.

6. Add primary antibody (1:250 Hsp60) and incubate overnight at 4C.

7. Remove primary antibody. Wash 2X quickly with PBST then wash 4X for 15 min each with PBST-BSA.

8. Incubate with secondary antibody (1:500 Alexa 594) diluted in PBANG for 2 hours at RT.

9. Remove secondary antibody. Wash 2X quickly with PBST then wash 4X for 15 min each with PBST-BSA.

10. Incubate 15 min with DNA stain (1:300 SYTOX Green) diluted in PBST.

11. Wash 2X with PBST, 15 mins each time.

12. Remove PBST. Add Prolong Diamond Antifade reagent then transfer embryos to a glass slide. Place a cover slip on top and seal with nail polish.

Nuclear staining of *Nasonia* ovaries

Collecting and fixing ovaries

1. Dissect ovaries in ice cold PBS with forceps. Transfer with forceps to a clean well filled with PBS. Separate ovarioles using microdissecting needles with insect pins.

2. Carefully remove PBS and add 4% formaldehyde in PBST (PBS + 0.2% Triton X-100). Fix for 20 mins at RT.

3. Transfer to a 1.5 ml Eppendorf tube containing PBST. Wash quickly 3 times with PBST.

4. Incubate in PBST with 1 mg/ml RNase for 3 hours at RT, then incubate overnight at 4C.

Staining ovaries

5. Incubate ovaries 15 min with DNA stain (1:300 SYTOX) diluted in PBST.

6. Wash 2 times with PBST, 15 mins each time.

7. Transfer ovaries to a slide. Remove extra PBST. Add a drop of ProLong Diamond Antifade reagent, place a cover slip on top and seal with nail polish.

APPENDIX F. LIST OF PUBLICATIONS

Funkhouser-Jones, L.J., Sehnert, S.R., Martínez-Rodríguez, P., Toribio-Fernández, R., Pita, M., Bella, J.L., and Bordenstein, S.R. (2015) *Wolbachia* co-infection in a hybrid zone: discovery of horizontal gene transfers from two *Wolbachia* supergroups into an animal genome. PeerJ 3: e1479.

Metcalf, J.A., **Funkhouser-Jones, L.J.**, Brileya, K., Reysenbach, A., and Bordenstein, S.R. (2014) Antibacterial gene transfer across the tree of life. eLife 3: e04266.

Funkhouser, L.J., and Bordenstein, S.R. (2013) Mom knows best: the universality of maternal microbial transmission. PLoS Biol 11(8): e1001631.

Brucker, R.M., **Funkhouser, L.J.**, Setia, S., Pauly, R. and Bordenstein, S.R. (2012) Insect Innate Immunity Database (IIID): an annotation tool for identifying immune genes in insect genomes. PLoS ONE 7(9): e45125.

Hillyer, J.F., Estévez-Lao, T.Y., **Funkhouser, L.J.,** and Aluoch, V.A. (2012) *Anopheles gambiae* corazonin: gene structure, expression and effect on mosquito heart physiology. Insect Mol Biol 21(3): 343-355.

Kent, B.N., **Funkhouser, L.J.**, Setia, S., and Bordenstein, S.R. (2011) Evolutionary genomics of a temperate bacteriophage in an obligate intracellular bacteria (*Wolbachia*). PLoS ONE 6(9): e24984.