Identifying and Characterizing Phage Genes Involved in *Wolbachia*-Induced Cytoplasmic Incompatibility

By

J. Dylan Shropshire

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
August 31, 2020
Nashville, Tennessee

Approved:

Seth R. Bordenstein, Ph.D.
Antonis Rokas, Ph.D.
Ann Tate, Ph.D.
Jared Nordman, Ph.D.
Borden Lacy, Ph.D.

To my girlfriend Laci Baker for showing me endless support throughout my education and for making sure that I maintain some semblance of a life outside of academia.

and

To my parents for never pressuring me to be anything other than what I wanted.

and

To my dog Tyrion for never letting me forget that his head pats are more important than working an extra hour each night at home.

Acknowledgements

My thesis work would not have been possible without a set of incredible people in my life. At Soddy Daisy High School, Mr. David Potter and Mr. Tony Boydston taught me to think critically and deeply about subjects they knew I was interested in, and regularly challenged my viewpoints with alternative perspectives and opinions. I am eternally grateful for their encouragement and support. Ms. Nancy Zuber believed in me and gave me a second chance to recover my GPA by enrolling me in a program to retake classes I did poorly in during a time in my life I missed a lot of school. Without her intervention I may not have been able to go to college.

At East Tennessee State University, Dr. Karl Joplin, Dr. Darrel Moore, and Dr. Edith Seier taught me what it means to be a scientist. I joined their research group through a program called Collaborative Research on the Arthropod Way of Life (CRAWL) and they guided me through the scientific process from the conception of ideas to the publication of my works. I am grateful for the opportunities they gave me, the constant support, and the pressure they put on me to continue growing. Their mentorship taught me what it means to be an educator, mentor, and scientist.

At Vanderbilt University, I have many to thank. For starters, Dr. Seth Bordenstein consistently pushed me to excel and regularly provided affirmation of my potential and ability to succeed. His guidance has led me to be a more critical, diligent, and thorough scientist. I have also been fortunate to have an amazing group of colleagues that regularly provided scientific support: Sarah Bordenstein, Dr. Daniel LePage, Dr. Kevin Kohl, Dr. Aram Mikaelyan, Dr. Brittany Leigh, Dr. Edward van Opstal, Danny On, Dr. Andrew Brooks, Dr. Karissa Cross, Caitlin Sprowls, and Dr. Rupinder Kaur. I must also extend a special thanks to Aram, Brittany, Caitlin, and Sarah for regularly providing personal support, guidance, and perspective. These people have shaped my views on science as a career and encouraged me to continue when times were hard. Finally, I thank the undergraduate mentees that I had the pleasure of supervising: Katie Carbonell, Melissa Halstead, Emily Layton, Helen Zhou, Mahip Kalra, and Rachel Rosenberg. Their excitement in science often reignited my own. I would like to particularly thank Emily Layton for being my first long-term student and working with me as I develop skills as a mentor.

My work has been generously supported by Vanderbilt University, the Gisela Mosig Travel Fund, the National Science Foundation Graduate Research Fellowship Program, and the National Institute for Health.

Table of Contents

List of Tables

Table Page

List of Figures

Figure Page

Chapter I.

Introduction to the history, implications, and mechanisms underlying *Wolbachia*-induced cytoplasmic incompatibility

Introduction

Wolbachia are among the world's most common animal-associated infections (Hilgenboecker et al., 2008; Zug and Hammerstein, 2012) and often hijack their host's reproduction via cytoplasmic incompatibility (CI) (LePage and Bordenstein, 2013; Serbus et al., 2008; Taylor et al., 2018; Werren et al., 2008). CI is a powerful selfish-drive mechanism that has a considerable impact on *Wolbachia*'s spread to high frequencies, and is characterized by embryonic death when infected males fertilize uninfected eggs (Figure I-1a) or eggs infected with incompatible *Wolbachia* (Figure I-1b, c). Despite *Wolbachia*'s discovery in 1924 (Hertig and Wolbach, 1924), and its linkage to CI in 1973 (Yen and Barr, 1973), little is known about how CI actually works. The contents of this dissertation identify and functionally interogate genes that cause and rescue CI (Chapters II, III, & IV), describe how genetic variation relates to phenotypic variation (Chapter V), and determine regions of the proteins necessary for function (Chapter VI). These works reveal that CI is not strictly a *Wolbachia* trait and is instead encoded in *Wolbachia*'s prophage WO. As such, these bacteriophage genes hold the power to control arthropod reproduction. In this Introduction, we discuss the history of CI research, why CI is useful to *Wolbachia*, CI's implications in speciation and vector control, and what is known about CI's mechanistic basis. These works motivated our studies and provided a foundation for our work determining CI's genetic basis.

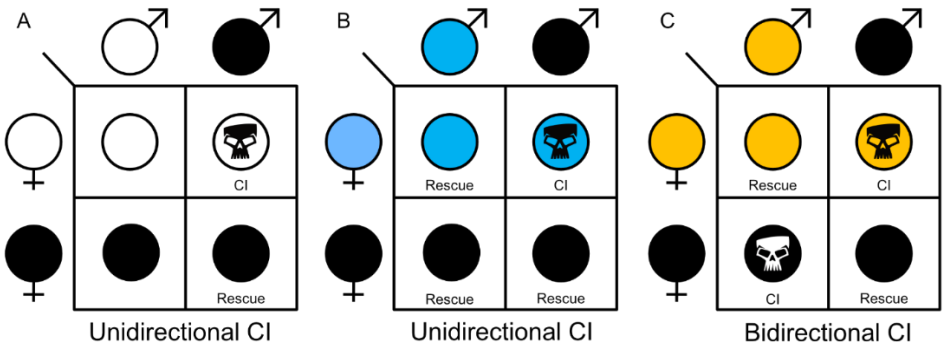

Figure I-1. Cytoplasmic incompatibility occurs in three forms.

(A) Unidirectional CI results in embryonic lethality when infected males are crossed with uninfected females. Rescue of this embryonic lethality occurs if the female carries a compatible infection. (B) In some cases, unidirectional CI can emerge when one strain can rescue another strain, but the other strain cannot reciprocate the rescue. (C) Bidirectional CI happens when numerous incompatible strains are present in a population. Rescue occurs if the female is likewise infected. Filled sex symbols represent infections. Different colors represent different infection types. Filled symbols of the inner circles represent the offspring's resulting infection state.

The enigma of insect biology

As early as 1938 and continuing into the 50s and 60s, incompatibility was reported between geographically isolated strains of *Culex pipiens* mosquitos (Laven, 1951; Marshall, 1938), *Aedes scutellaris* mosquitos (Smith-White and Woodhill, 1955), and *Nasonia* parasitoid wasps (Ryan and Saul, 1968). In *Culex* and *Aedes*, these discoveries led to debate that these incompatible strains should be designated as different species based on the Biological Species Concept (Laven, 1951), which is founded on the premise that species are biological entities incapable of successful interbreeding (Dobzhansky, 1937; Mayr, 1963). However, crossing experiments in both *Culex* and *Nasonia* revealed that this incompatibility was maternally inherited and the contributing factor was cytoplasmic (Laven, 1951; Ryan and Saul, 1968). Importantly, some cross types were incompatible in one direction (unidirectional; Figure I-1a,b) and others were incompatible regardless of the direction (bidirectional; Figure I-1c).

In 1924 Hertig and Wolbach discovered a peculiar Rickettsial bacteria residing in female reproductive cells of *Cu. pipiens* (Hertig and Wolbach, 1924). This bacteria would later be named *Wolbachia pipientis*: *Wolbachia* for Burt Wolbach and *pipientis* for the mosquito it was first found in (Hertig, 1936). In 1971 Yen and Barr hypothesized a connection between CI and this long-overlooked bacteria (Yen and Barr, 1971), and later tested this hypothesis using crossing experiments with antibiotic treated mosquitos to confirm that *Wolbachia* is the etiological agent of CI (Yen and Barr, 1973).

Since then, CI-inducing *Wolbachia* have been found in many Diptera (Baton et al., 2013; Bian et al., 2013; Riegler and Stauffer, 2002), Hymenoptera (Betelman et al., 2017; Dittmer et al., 2016), Coleoptera (Kajtoch and Kotásková, 2018), Hemiptera (Ju et al., 2017; Ramírez-Puebla et al., 2016), Orthoptera (Martinez-Rodriguez and Bella, 2018), Lepidoptera (Arai et al., 2018; Hornett et al., 2008), Thysanoptera (Nguyen et al., 2017), Acari (Gotoh et al., 2007, 2003; Vala et al., 2002), Isopoda (Cordaux et al., 2012; Sicard et al., 2014), and Arachnids (Curry et al., 2015). *Wolbachia* are highly diverse and phylogenetically divided into 17 "supergroups" (denoted A-R, excluding G), and CI-inducing *Wolbachia* are restricted to supergroups A and B (Lo et al., 2007a,

2007b; G. H. Wang et al., 2016a). However, despite the considerable diversity in the *Wolbachia* strains explored in published works (Table I-1), the most studied models are the *Wolbachia* of *Culex* (*w*Pip), *Drosophila* (*w*Ri and *w*Mel), *Nasonia* (*w*VitA), and the planthopper *Laodelphax striatellus* (*w*Str).

Table I-1. List of *Wolbachia* strains directly referenced in this dissertation and their native host species.
Some strains have been transinfected across species. Transinfected host species will be mentioned in the text when appropriate.

Host species	**Strain**	**Supergroup**	**Phenotype**
Aedes albopictus	*w*AlbA	A	CI
Ae. albopictus	*w*AlbB	B	CI
Brugia malayi	*w*Bm	D	Mutualist
Culex pipiens	*w*Pip	B	CI
Drosophila mauritiana	*w*Mau	B	Rescue only
D. melanogaster	*w*Mel	A	CI
D. melanogaster	*w*MelPop	A	CI
D. pandora	*w*PanCI	A	CI
D. pandora	*w*PanMK	A	MK
D. recens	*w*Rec	A	CI
D. santomea	*w*San	A	CI
D. simulans	*w*Au	A	None
D. simulans	*w*Ha	A	CI
D. simulans	*w*No	B	CI
D. simulans	*w*Ri	A	CI
D. suzukii	*w*Suz	A	None
D. teissieri	*w*Tei	A	CI
D. yakuba	*w*Yak	A	CI
Hypolimnus bolina	*w*Bol	B	CI, MK
Laodelphax striatellus	*w*Stri	B	CI
Nasonia vitripennis	*w*VitA	A	CI
N. vitripennis	*w*VitB	B	CI
Rhagoletis cerasi	*w*Cer	A	CI

Invasion and spread of *Wolbachia*

Understanding why CI-inducing *Wolbachia* are so common requires an appreciation for the models that explain *Wolbachia*'s spread with and without CI. *Wolbachia* commonly transmit

horizontally across evolutionary timescales (Boyle et al., 1993; Conner et al., 2017; Gerth et al., 2013; Huigens et al., 2004; Tolley et al., 2019). For example, *w*Ri has jumped between *Drosophila* hosts 10-50 million years diverged from each other but has done so in only the past 5-27 thousand years (Turelli et al., 2018a). The first boundary to a horizontally transmitted infection is population invasion from low frequencies. Theoretical and empirical studies agree that *Wolbachia* invasion is dependent on the relative fitness of infected and uninfected females (Bakovic et al., 2018; Carrington et al., 2011; Hoffmann et al., 2011; Turelli, 1994; Turelli and Hoffmann, 1991). *Wolbachia* confer a plethora of fitness advantages to their hosts including increased longevity and fecundity (Fry et al., 2004; Maistrenko et al., 2016; Olsen et al., 2001; Versace et al., 2014), reduced pathogen susceptibility (Ant et al., 2018; Blagrove et al., 2012; Caragata et al., 2016), and nutrient provisioning (Hosokawa et al., 2010; Ju et al., 2019; Newton and Rice, 2019). If these mutualistic attributes outweigh the fitness costs of carrying *Wolbachia* then *Wolbachia* will spread regardless of CI expression (Kriesner et al., 2013; Turelli, 1994). For example, male-killing *w*Bol *Wolbachia* of the butterfly *Hypolimnus bolina* and the non-parasitic *w*Au *Wolbachia* of *D. simulans* have reached very high frequencies in their respective populations (Duplouy et al., 2010; Hornett et al., 2009; Kriesner et al., 2013) presumably due to conferred relative fitness advantages *Wolbachia* confers by increasing fecundity or protecting its host from viral infection (Alexandrov et al., 2007; Teixeira et al., 2008). The ceiling to infection frequency is then determined by the frequency of maternal transmission (Turelli, 1994). If maternal transmission is perfect and yields 100% infected offspring, then *Wolbachia* can be expected to reach fixation (Narita et al., 2009; Turelli, 1994). Otherwise, the maximum infection frequency is dependent on the frequency of imperfect maternal transmission.

CI does not help with *Wolbachia*'s initial invasion, nor does it influence the infection ceiling. Instead, CI increases *Wolbachia*'s rate of spread through a population in a manner dependent on its infection frequency. For instance, when *Wolbachia* are common, infected females are relatively more fit than uninfected females since they are compatible with infected CI-inducing males. However, these fitness gains are only achieved after *Wolbachia* has reached a frequency exceeding the fitness cost of infection (Turelli, 1994). If infected females are 20% less fecund than uninfected females, then *Wolbachia's* frequency must exceed 20% of the population before CI will help with *Wolbachia's* spread (Hoffmann et al., 2011). Exceeding this fitness threshold is possible via drift if fitness cost is minimal, but if *Wolbachia* yield fitness gains then CI will immediately

help spread *Wolbachia* into the population. Higher infection frequencies will be required for spread if maternal transmission is not perfect since a proportion of their offspring will be uninfected (Adekunle et al., 2019).

After CI-inducing *Wolbachia* exceed the invasion threshold, they spread more quickly when CI is strong (high % embryonic death), the host organism disperses far, and infected females are more fit than uninfected females. The most well-studied example is the *w*Ri *Wolbachia* of *D. simulans* which spread through California at a rate of 100 km per year in the late 1980s, reaching a stable frequency above 95% in as little as 3 years (Turelli and Hoffmann, 1991). *w*Ri-infected females were as fecund as uninfected females, thus allowing *Wolbachia* to spread from very low frequencies. Additionally, *w*Ri caused an average of 45% CI (45% embryonic death) and approximately 96% of offspring were infected when their mom was infected (Hoffmann et al., 1990; Turelli et al., 1992; Turelli and Hoffmann, 1995). Since maternal transmission was not perfect, *Wolbachia*-infection never completely reached fixation but reached a high frequency that was stable with maternal transmission efficiency (Turelli and Hoffmann, 1995). Similar invasion dynamics have been observed in the cherry fruit fly *Rhagoletis cerasi* where *w*Cer2 *Wolbachia* have been spreading across central Europe at a rate of 1-1.9 km per year (Bakovic et al., 2018; Riegler and Stauffer, 2002; Schebeck et al., 2019).

A contrasting example is in *Ae. aegypti* mosquitos that were transinfected with *w*Mel *Wolbachia* for the purposes of vector control (discussed more below). These mosquitos were released into uninfected populations in Cairns, Australia to replace the population with *Wolbachia*-infected mosquitos. Unlike *w*Ri which spread from low frequencies, *w*Mel transinfected *Aedes* were ~20% less fecund than wild uninfected females. As such, they were released to frequencies between 20-30% before *Wolbachia* began to spread without intervention (Hoffmann et al., 2011). Moreover, the rate of spatial spread was significantly slower than *w*Ri's spread through California, at a rate of 100-200 m per year (Hoffmann et al., 2011). Since both strains are known to cause strong CI (Blagrove et al., 2012; Dutra et al., 2015), factors such as host dispersal rate and generation times may influence rate of spread. However, these *Wolbachia* have remained stable at high infection frequencies since soon after release, supporting CI's ability to not only spread *Wolbachia* but to also reinforce high frequencies (O'Neill et al., 2018).

A number of additional factors within a population can influence *Wolbachia*'s spread dynamics via CI. These include variable age structures due to mortality and life history strategies

(Engelstadter and Telschow, 2009; Farkas and Hinow, 2010; Turelli, 2010), co-existence with multiple CI-inducing *Wolbachia* (Yoshida et al., 2019), heat-stress and environmental fluctuations that impact *Wolbachia* density and CI strength (Foo et al., 2019; Hu et al., 2019; Ross and Hoffmann, 2018), assortative mate choice that encourages uninfected females to avoid infected males (Arbuthnott et al., 2016; Jaenike et al., 2006; Shropshire and Bordenstein, 2016), the range of host dispersal, and other factors that can influence CI strength, maternal transmission, or population structure. However, models that simply include CI strength, maternal transmission efficiency, and fitness have been sufficient to predict the spread of *Wolbachia* in populations monitored to date (Bakovic et al., 2018; Hoffmann et al., 2011; Turelli and Hoffmann, 1991).

CI contributes to reproductive isolation

Charles Darwin proposed our modern framework for understanding organismal evolution in the *Origin of Species* (Darwin, 1869). Dobzhansky, Mahr, and others then refined our definition of species into the Biological Species Concept which describes groups of individuals as different species if they cannot interbreed (Dobzhansky, 1937; Mayr, 1963). Under this paradigm, species emerge through divergence caused by reproductive isolation between two populations. Reproductive isolation can be subdivided into pre-mating barriers including geographic isolation and mate discrimination and post-mating barriers including zygotic mortality and hybrid sterility (Coyne, 2001). The concept that symbiosis can drive speciation was first proposed by the botanist Konstantin Mereschkowski in the early 1900's and later advanced by Ivan Wallin and Lynn Margulis (Margulis, 1967; Mereschkowsky, 1910; Wallin, 1927). Since then, scientific consensus has converged on an endosymbiotic theory of evolution where symbiotic interactions between two entities became the basis for Eukaryotic life (Imachi et al., 2020; Zachar and Boza, 2020). Moreover, it is increasingly appreciated that microbes are important in host nutrition, physiology, development, behavior, and reproduction (Gilbert et al., 2012). Many such symbiotic interactions are predicted to help drive speciation (Bordenstein, 2003; Brucker and Bordenstein, 2012a; Shropshire and Bordenstein, 2016). *Wolbachia*-induced CI is of particular interest as a mechanism for symbiotic speciation since it can reduce nuclear gene flow between host individuals in the absence of host genetic divergence or geographic isolation (Brucker and Bordenstein, 2012a; Coyne, 2001; Hurst and Schilthuizen, 1998; Laven, 1967a; Shropshire and Bordenstein, 2016; Thompson, 1987). Since *Wolbachia* and CI are common among arthropods (Weinert et al., 2015;

Zug and Hammerstein, 2012), CI has been hypothesized to be a major contributor to the considerable species richness of arthropods. Here, we highlight a few studies that suggest CI can contribute to reproductive isolation.

Bidirectional CI and unidirectional CI are predicted to have different consequences on reproductive isolation. Since bidirectional CI restricts geneflow in both directions, it is likely to establish reproductive isolation between populations with different infection states. Indeed, the *Nasonia* species group has bidirectionally incompatible *Wolbachia* in a younger species pair which diverged ~0.25 million years ago (*N. giraulti* and *N. longicornis*) and older species pair which diverged ~0.8 million years ago (*N. giraulti* and *N. vitripennis*) (Bordenstein et al., 2001; Bordenstein and Werren, 2007; Campbell et al., 1994). Curing each species of their *Wolbachia* restored compatibility between the younger species pair, but only minimally improved compatibility between the older species pair which also suffered from behavioral, spermatogenic, and genetic incompatibilities (Bordenstein et al., 2001; Breeuwer and Werren, 1990; Clark et al., 2010). These data suggest that CI can be a form of reproductive isolation that has emerged early in the speciation of *Nasonia* (Bordenstein et al., 2001).

Although unidirectional CI does not appear to contribute to speciation in some host-*Wolbachia* symbioses (Cooper et al., 2017), it is likely a significant form of reproductive isolation in others. For example, unidirectional CI reduces gene flow in the hybrid zone of North American populations of *D. recens*, which are infected with *Wolbachia*, and *D. subquinaria*, which are not (Jaenike et al., 2006; Shoemaker et al., 1999). Whenever *D. recens* males hybridize with *D. subquinaria* females, their offspring are inviable due to CI, however the inverse crossing is compatible (Shoemaker et al., 1999). This incompatibility relationship yields an asymmetrical reduction in gene flow between these populations, but *D. subquinaria* females have strong mate discriminating behaviors that prevent them from mating with infected *D. recens* males (Jaenike et al., 2006). Together, these barriers prevent hybridization between these species (Jaenike et al., 2006). Additionally, CI putatively acts as a form of reproductive isolation in other arthropod species (Giordano et al., 1997; Maroja et al., 2008), and models suggest that low migration rates make CI more likely to influence reproductive isolation (Telschow et al., 2007). While these works suggest that unidirectional CI can be a form of reproductive isolation when coupled with other factors, more work is necessary to determine if CI is a common form of reproductive isolation.

For CI to commonly be a significant form of reproductive isolation, then it must be maintained long enough for host divergence to reinforce hybrid incompatibility. It remains unknown if CI can persist in a species over evolutionary timescales, but theoretical and empirical studies have investigated how long *Wolbachia* coevolves with particular host species. For instance, theory predicts that hosts will develop resistance to CI-inducing infections if *Wolbachia* maternal transmission is imperfect (Koehncke et al., 2009; Turelli, 1994), driving *Wolbachia* to low titers, ablating CI phenotypes, and otherwise reducing its ability to spread. However, in *D. melanogaster* populations that diverged 3263-13998 years ago, *Wolbachia* have been stably maintained with particular mitochondrial haplotypes, suggesting that *Wolbachia* has associated with this host for thousands of years (Ilinsky, 2013). Conversely, while numerous parasitic and non-parasitic endosymbionts have rapidly invaded populations (Carrington et al., 2011; Kriesner et al., 2013; Turelli and Hoffmann, 1991), rapid declines have been observed in as little as 10 years in *Rickettsia* infecting whiteflies in North America (Bockoven et al., 2019). These case studies suggest that while *Wolbachia* can quickly reach high prevalence in a population, its long-term association is not guaranteed. Moreover, CI strength (% embryonic death) will contribute to the amount of gene flow that is allowed through this form of reproductive isolation, and influence the amount of time necessary for CI to contribute to emergence of other types of reproductive isolation. Weak CI may still contribute to reproductive isolation, but it is likely to take longer for other forms of reproductive isolation to emerge. In nature, it is common that even so-called "strong" CI inducers will only prevent hatching of half of CI-affected offspring (Turelli and Hoffmann, 1991). More work is necessary to determine if *Wolbachia* can regularly be maintained over evolutionary timescales and if CI can be a persistent form of reproductive isolation.

CI is a tool in vector control

CI has achieved considerable scientific and public interest in recent years for its use in vector control programs to curb the spread of mosquito-borne diseases. These efforts leverage infertilities caused by CI-induction and/or pathogen blocking characteristics of some *Wolbachia* strains that prevent the replication of some RNA viruses in their host (Teixeira et al., 2008). The World Health Organization recommended the development and deployment of *Wolbachia*-based vector control strategies in response to Zika outbreaks in 2016 (Vector Control Advisory Group, 2016). Currently two distinct strategies are used in the field. First, the incompatible insect

technique (IIT) releases *Wolbachia*-infected males into vector populations that are otherwise uninfected. When the infected males mate with uninfected females, the population size will decrease due to CI. Second, the population replacement strategy releases both *Wolbachia*-infected males and females into uninfected populations. The combined ability of CI and rescue enables the replacement of the native mosquito population with the introduced population which is resistant to disease transmission (Teixeira et al., 2008). There are many reviews that discuss *Wolbachia* and vector control (Flores and O'Neill, 2018; Hoffmann et al., 2015; Jeffries and Walker, 2015, 2016; LePage and Bordenstein, 2013; Lindsey et al., 2018a; Mohanty et al., 2016; Mustafa et al., 2016; Niang et al., 2018; Nikolouli et al., 2018; O'Neill, 2018; Ritchie et al., 2018; Shaw and Catteruccia, 2019; Terradas and McGraw, 2017). Here, we discuss the historical perspective on CI's use in vector control and briefly review the recent literature regarding these two strategies.

The IIT was used in field trials long before *Wolbachia* was known to be responsible for CI (Laven, 1967b; Yen and Barr, 1973). The earliest study aimed to eradicate a *Cu. pipiens* population of mosquitos in the Okpo village north of Rangoon in Burma (Laven, 1967b). The population was estimated to range from 4-20 thousand mosquitos and carried multiple bidirectionally incompatible infections (Atyame et al., 2011a). When infected *Cu. pipiens* males, carrying *Wolbachia* bidirectionally incompatible with the Burma strain were released, the Burmese *Cu. pipiens* populations experienced a rapid decline. This decline was measured as the percentage of egg rafts that hatched per week for 12 weeks (Laven, 1967b). Hatching started above 90% and declined to below 10% after 8 weeks of releases. A monsoon prevented continued monitoring of mosquito population recovery (Laven, 1967b), but this study provided the first proof-of-concept that the IIT could be used as a means to quickly and significantly reduce arthropod population sizes.

Several decades passed between these initial studies and further investigation of CI's use with the IIT. More recent studies have investigated the IIT's efficacy in *Ae. polynesiensis, Ae. albopictus*, and *Ae. aegypti* revealing significant CI in controlled semi-field releases of *Wolbachia*-infected males (Caputo et al., 2019; Chambers et al., 2011; Mains et al., 2019, 2016; O'Connor et al., 2012; Puggioli et al., 2016). Additionally, multiple organizations are now working to leverage the IIT to reduce the spread of disease. These include MosquitoMate (Kentucky, USA), Verily (California, USA), Commonwealth Scientific and Industrial Research Organization (Australia), Singapore's National Environment Agency (Singapore), among others. Verily's Debug project is the largest project, releasing millions of *Wolbachia*-infected *Ae. aegypti* males in Fresno

California, Innisfail Australia, and the Tampines neighborhood in Singapore. These efforts have proven successful, seeing a 95% population reduction in California in 2018 (Debug Fresno, 2018a), more than 80% in Australia in 2018 (Debug Fresno, 2018b), and 90% in Singapore in 2019 (Debug Fresno, 2019). No reports have been published to suggest that these strategies are reducing disease burden in release areas, and the cost-effectiveness of these efforts also remain unknown. Despite this, work is being done to expand the use of the IIT to control programs for other vectors and pests including the fruit pest *Ceratitis capitate* (Kyritsis et al., 2019) and the protozoan vector *Cu. quinquefasciatus* (Ant et al., 2020). There are also efforts to combine the IIT with the sterile insect technique which employs mutagenesis to impose infertility on males (D. Zhang et al., 2015). These combined efforts would ensure that released males are completely sterile, increasing the rate of population decline.

Conversely, the population replacement strategy leverages two aspects of *Wolbachia* biology to curb disease transmission. The first is a pathogen suppression phenotype observed with some *Wolbachia* strains (Moreira et al., 2009; Teixeira et al., 2008), and the second is CI to spread the pathogen suppression phenotype into the target population. *Ae. aegypti* mosquitos, a vector of many human pathogens, do not naturally carry *Wolbachia* (Walker et al., 2011), making it a target for *Wolbachia*-mediated control. The *w*Mel *Wolbachia*, native to *D. melanogaster* was transinfected into *Ae. aegypti* and conferred upon its host resistance to viruses including dengue, Zika, Chikungunya, Yellow Fever, and others (Caragata et al., 2016; Moreira et al., 2009; Teixeira et al., 2008; van den Hurk et al., 2012). As such, if *w*Mel-infected mosquitos replace an otherwise uninfected population then they are predicted to reduce the vectoral competence of the host population, thus reducing disease burden in humans. The World Mosquito Program (previously the Eliminate Dengue Program) is the predominant entity using the population replacement strategy in vector control. Initial studies released mosquitos in Australia (Hoffmann et al., 2011; O'Neill, 2018), and the World Mosquito Program is now operating in Brazil, Colombia, Mexico, India, Indonesia, Sri Lanka, Vietnam, Fiji, Kiribati, New Caledonia, and Vanuatu (https://www.worldmosquitoprogram.org/en/global-progress). In Australia, field trials began in January of 2011 with the release of thousands of *Wolbachia*-infected mosquitos (Hoffmann et al., 2011). Releases continued for 3 months and *Wolbachia*-infection frequencies continued increasing after release until reaching a stable equilibrium near fixation that has persisted for several years (Hoffmann et al., 2011; O'Neill et al., 2018). Reports from another release site in Townsville,

Australia reveal that dengue transmission is nearly eliminated once *Wolbachia* reaches high frequencies (O'Neill et al., 2018). Taken together, these studies have shown that the population replacement strategy can work to drive *Wolbachia* to high frequencies to reduce the vectoral capacity of mosquitos.

The mechanistic basis of CI remains mostly unresolved

Decades of research on *Wolbachia*-induced CI (Yen and Barr, 1973), have resulted in a considerable body of evidence to describe and interrogate CI's specific impact on reproduction, factors that influence CI strength, and host proteins and RNAs that correlate with CI-induction. While these works do not solve the mechanistic basis of CI, they provide a significant foundation for future works. A key limitation of many of the works below is that they are often correlative studies that link *Wolbachia*-infection state with host outcomes. As such, it often remains unclear if observations are related to *Wolbachia*-infection or to CI-induction. Below, we review the literature regarding the cytological assessment of CI, CI strength variation, and host proteins and RNAs that correlate with *Wolbachia*-infection. We end with a review of the prior efforts to uncover CI's genetic basis and the findings derived from the following chapters in this dissertation.

Pre- and post-fertilization abnormalities during CI.

Since CI is the byproduct of a sperm-egg incompatibility, CI-associated defects are predicted to manifest during or prior to embryogenesis. Proper sperm maturation involves the replacement of histones with protamines for packaging and then replacement of those protamines post-fertilization with maternally derived histones. In CI-affected embryos, histone deposition is delayed after protamine removal (Landmann et al., 2009), presumably leading to retention of the DNA polymerase cofactor proliferating cell nuclear antigen and delayed Cdk1 activation and nuclear envelope breakdown in the male pronucleus (Landmann et al., 2009; Tram and Sullivan, 2002). These defects traditionally culminate in arrest of the first mitosis and chromatin bridging (Callaini et al., 1996; Lassy and Karr, 1996; Ryan and Saul, 1968). In *D. simulans, Ae. polynesiensis* and *Cu. pipiens*, embryonic arrests can occur after the first mitosis (Bonneau et al., 2018b; Callaini et al., 1997; Ryan and Saul, 1968). These later stage defects manifest as early mitotic failures where there are several successful rounds of division before embryonic arrest, or

regional mitotic failures where regions of the embryo fail to divide and others are successful. Chromatin bridging often accompanies these defects. The cause of these later stage defects remains unknown, but it has been hypothesized that the intensity of the earlier pronuclear delay may determine whether later stage defects are observed. More specifically, if the male pronucleus is considerably slowed, then it may result in complete exclusion of the male pronucleus from early development and the embryo would undergo haploid development (Callaini et al., 1997; Tram et al., 2006). In *N. vitripennis* where haploid individuals become males and diploid become females, CI can manifest in haploidization where even fertilized eggs develop as haploid (Bordenstein et al., 2003; Ryan and Saul, 1968; Vavre et al., 2001, 2000). However, haploids are not viable in *D. simulans* and other diploid species, resulting in arrest later in embryogenesis (Ferree and Sullivan, 2006). Understanding the underlying mechanisms that contribute to these distinct cytological outcomes will be of interest for future research.

To date, histone deposition defects are the earliest detected aberrations post-fertilization, but CI-associated abnormalities have also been observed pre-fertilization. For example, infected *D. simulans* and *Ephestia* moths produce fewer sperm, and stronger CI is induced when more sperm are transferred during copulation in *D. simulans* (Awrahman et al., 2014; Lewis et al., 2011; Snook et al., 2000). When females were mated to both infected and uninfected males, the sperm of uninfected males were more likely to fertilize eggs (Champion de Crespigny and Wedell, 2006), suggesting that *Wolbachia*-modified sperm are less competitive. *Wolbachia*-affected sperm cysts exhibit abnormal morphology with some sperm being fused together and other having randomly oriented axoneme-mitochondrial complexes which are responsible for sperm motility (Riparbelli et al., 2007), perhaps explaining fertility defects and variation in sperm competition. However, the underlying causes to sperm motility and morphology are unknown and it remains unclear if these observations are directly related to CI or are a byproduct of *Wolbachia* infection in the testes.

Understanding *Wolbachia* localization during spermatogenesis provides valuable insights into CI's mechanism. Spermatogenesis starts with the germline stem cell niche (GSCN) replicating into spermatogonia which undergo mitosis to form a spermatocyst with 16 spermatocytes. Each spermatocyte in the cyst then undergoes two rounds of meiosis to form four spermatids, for a total of 64 spermatids in each cyst. The spermatids then undergo elongation and individualization before becoming mature sperm and entering the seminal vesicle for storage until mating. *Wolbachia* are not symmetrically distributed within infected testes, with only some spermatocysts being infected

in the strong CI-inducing *w*Ri strain (Clark et al., 2003). Indeed, *w*Ri is almost exclusively localized to the GSCN , and some GSCN remain uninfected, suggesting that *Wolbachia*-derived products responsible for CI must either act early in spermatogenesis or are diffusible factors that can stably travel into later stages of spermatogenesis (Clark et al., 2003, 2002; Riparbelli et al., 2007). *Wolbachia* that persist to infect spermatid tails are stripped during the individuation process and are moved into waste bags where they are presumably degraded (Riparbelli et al., 2007). Not only does this indicate that *Wolbachia* create a diffusible factor that interacts with sperm to cause CI, but also helps to explain why paternal *Wolbachia* transmission has not been observed (Yeap et al., 2016).

Factors that influence CI strength variation.

The intensity of CI can be highly variable within and between arthropod species. Studying the factors that influence this variability can inform CI's mechanism. Even within *Drosophila* species, CI varies from 100% embryonic death to 10-15% reductions in hatching (Awrahman et al., 2014; Clark et al., 2003; Cooper et al., 2017; Hoffmann, 1988; Layton et al., 2019a; Reynolds and Hoffmann, 2002; Turelli et al., 2018b; Yamada et al., 2007; Zabalou et al., 2004). In fact, a number of *Wolbachia* including *w*Mel of *D. melanogaster* and *w*Yak of *D. yakuba* were initially thought not to be parasitic strains (Charlat et al., 2004; Holden et al., 1993; Zabalou et al., 2004). However, laboratory studies revealed both strains can cause CI when factors such as age are controlled (Reynolds and Hoffmann, 2002), suggesting that environmental or technical variation may contribute to CI strength. Additional support for this hypothesis came from *w*Ri which was known to cause relatively strong CI in nature (45% embryonic death) that enabled its rapid spread through California but caused even stronger CI in the lab (~90% embryonic death) (Carrington et al., 2011; Mouton et al., 2006; Turelli and Hoffmann, 1995).

Temperature is a significant co-correlate of CI strength variation and is likely to contribute to the dynamics that govern *Wolbachia*'s spread (Foo et al., 2019). High temperatures usually exceeding 27^{o}C can have a significant negative impact on CI strength in *Wolbachia*-infected *Ae. aegypti* (Ross et al., 2020, 2019), *T. urticae* (van Opijnen and Breeuwer, 1999), *D. simulans* (Hoffmann et al., 1986), *D. melanogaster* (Reynolds and Hoffmann, 2002), *Ae. scutellaris* (Trpis et al., 1981; Wright and Wang, 1980), and *Ae. albopictus* (Wiwatanaratanabutr and Kittayapong, 2009). There is considerable evidence that temperature impacts *Wolbachia* densities in insect

reproductive tissues. For instance, *Wolbachia* in *D. simulans* (Clancy and Hoffmann, 1998) and *Leptopilina heterotoma* wasps (Mouton et al., 2006) replicate more quickly at warmer temperatures, while *Wolbachia* decrease with rising temperatures in *Ae. albopictus* and *Ae. aegypti* (Foo et al., 2019; Ross et al., 2020, 2019) or even cure host infection (Jia et al., 2009). It is unknown what contributes to this variable impact of temperature on *Wolbachia* densities, but *Wolbachia* titers in *N. vitripennis* and *T. urticae* have been shown to decrease with temperature alongside an increase in lytic activity of *Wolbachia's* phage WO (Bordenstein and Bordenstein, 2011; Lu et al., 2012), suggesting that temperature may trigger phage lysis and reduce *Wolbachia* titers. In nature, *Wolbachia* densities vary with season in the butterfly *Zizeeria maha,* and climate change appears to be contributing to a decrease in infection frequencies in the tropics (Charlesworth et al., 2019; Sumi et al., 2017). Notably, temperatures exceeding 27^{o}C are common in the tropics, and can thus may impact CI variability in those regions. However, it is possible that some species can become at least partly resistant to temperature fluctuation by changing their behaviors. For example, infected *D. melanogaster* are attracted to colder temperatures than their uninfected counterparts (Truitt et al., 2018). This behavioral shift may mitigate the impact of temperature fluctuations on *Wolbachia* titers and CI strength, allowing for infected individuals in nature to seek out microenvironments that prevent loss of infection. Together, these data support a relationship between CI strength, *Wolbachia* titers, and temperature in some *Wolbachia*-host associations.

Other co-correlates of CI strength variation are related to male age (Awrahman et al., 2014; Reynolds and Hoffmann, 2002), male mating rate (Awrahman et al., 2014; De Crespigny et al., 2006), rearing density (Yamada et al., 2007), male developmental timing (Yamada et al., 2007), and host nutrition (Clancy and Hoffmann, 1998). All of these factors are significantly impacted by the structure of the population, resource availability, or behavior. For example, *D. simulans* males mate more frequently than uninfected females, and the increased mating rate yields weaker CI in later matings (Awrahman et al., 2014; De Crespigny et al., 2006). Infected males also transfer more sperm during copulation than uninfected males during the first mating encounter, and decreased sperm transfer in subsequent matings corresponds with weaker CI (Awrahman et al., 2014). As such, the increased mating frequency may be a behavioral adaptation employed by some hosts to restore reproductive compatibility between infected males and uninfected females (Awrahman et al., 2014). Older males, even when maintained as virgins for numerous days, cause weakened CI in some strains (Karr et al., 1998; Reynolds and Hoffmann, 2002; Turelli and

Hoffmann, 1995; Weeks et al., 2007), but age alone cannot explain decreased CI upon remating (Awrahman et al., 2014). Moreover, *Wolbachia* titers decrease with male age (Binnington and Hoffmann, 1989; Bressac and Rousset, 1993; Clark et al., 2002; Riparbelli et al., 2007; Veneti et al., 2003), and while it has been hypothesized that variation in *Wolbachia* titers decrease upon remating, this hypothesis has not been explicitly tested. Alternatively, it has also been hypothesized that the amount of time that sperm remains in contact with *Wolbachia* or CI-inducing products corresponds with how strong CI can be (Karr et al., 1998), thus remating may contribute to high sperm turnover that limits *Wolbachia*-sperm exposure.

Host genotype also co-correlates with CI strength variation. The relationship between *Wolbachia* phenotypes and host genotypes are frequently investigated through artificial transinfections were *Wolbachia* are purified from one strain or species and injected into adults, embryos, or cell culture of another strain or species (Hughes and Rasgon, 2014). If the infection becomes stable, then factors such as CI strength can be measured in the context of new genetic backgrounds. For example, *w*Mel *Wolbachia* of *D. melanogaster* traditionally cause weak CI (Holden et al., 1993), but induce consistently strong CI when transinfected into either *D. simulans* or *Ae. aegypti* (Poinsot et al., 1998; Walker et al., 2011). Similar results have also been observed when *w*Yak, *w*Tei, and *w*San, which induce weak CI in the *D. yakuba* complex (Charlat et al., 2004; Cooper et al., 2017; Zabalou et al., 2004), are transferred into *D. simulans* (Zabalou et al., 2008). Moreover, genetic variation in the wasp *N. longicornis* has been shown to correlate with whether its *Wolbachia* are unidirectionally or bidirectionally incompatible with other strains (Raychoudhury and Werren, 2012). These studies support models that predict hosts will be selected to develop resistance against CI-induction (Prout, 1994; Turelli, 1994), and raise many questions about how host genotypes control reproductive parasitism and how *Wolbachia* may enter an evolutionary arms race with the host. As such, the pathway(s) that CI act(s) on in the host must be conserved enough for CI to be transferable between species, but also malleable enough for the pathway(s) to become resistant to CI-induction.

These factors do not work on CI in isolation and instead seem to be mingled in a state of perpetual complexity. For instance, age has a variable impact on CI strength in different host backgrounds, suggesting that genotypic variation in either the host or *Wolbachia* strain may impact these relationships (Reynolds and Hoffmann, 2002). Moreover, the impact of temperature on CI strength in *D. melanogaster* is dependent on male age, where 1-day old males reared at 25^{o}C induce

stronger CI than those reared at 19^{o}C, but the inverse is true with 3 and 5-day old males (Reynolds and Hoffmann, 2002). Age does not impact CI intensity in *Cu. pipiens*, suggesting that host background or *Wolbachia* genotype significantly impact whether factors such as age can influence CI strength variation (O. Duron et al., 2007). These studies highlight the complexity of *Wolbachia*-host-environment interactions and should motivate additional studies to resolve the factors that underpin these variations and the genetic loci that influence how impactful each factor might be in that particular host.

Host RNAs and proteins linked to CI.

It is common that researchers leverage correlations between *Wolbachia* infection state and host expression phenotypes (RNA, protein, etc.) to understand how *Wolbachia* impact their host. When differential expression is correlated with CI phenotypes, these data can be valuable for generating hypotheses surrounding CI's mechanism. Significant correlations between *Wolbachia* infection state and host expression are measurable in *D. melanogaster* (Biwot et al., 2019; He et al., 2019; LePage et al., 2014; Liu et al., 2014; Ote et al., 2016; Xi et al., 2008; Yuan et al., 2015; Zheng et al., 2011; Y. Zheng et al., 2019), *D. simulans* (Brennan et al., 2012; Clark et al., 2006; Xi et al., 2008), *La. striatellus* (Huang et al., 2019; Ju et al., 2017; Liu et al., 2019), *T. urticae* (Bing et al., 2019; Y.-K. Zhang et al., 2015), *Cu. pipiens* (Pinto et al., 2013), and *Ae. albopictus* (Baldridge et al., 2017, 2014; Brennan et al., 2012, 2008). Problematically, as many as 1613 genes are differentially expressed between *Wolbachia* infection states (Bing et al., 2019).

However, the most promising candidates are those that can be experimentally over- or under-expressed to recapitulate CI-like hatch rates and cytological defects. For example, over-expression of the tumor suppressor gene *lethal giant larvae* [l(2)gl] and myosin II gene *zipper* in *Wolbachia*-uninfected *D. simulans* induces a considerable reduction in hatching that is accompanied with CI-associated cytological defects (Clark et al., 2006). However, CI is not just associated with hatch rate defects, but also the ability to rescue those defects. When l(2)gl and *zipper* overexpressing males were mated to infected females, no change in hatching was observed (Clark et al., 2006), suggesting that infertilities caused by overexpression of these genes could not be rescued and are thus unlikely to be directly associated with CI. That said, there have been numerous studies that have identified host factors that contribute to CI-like embryonic abnormalities and can be rescued by infected females: the aminotransferase *iLve* which mediated

branched-chain amino acid biosynthesis in *La. striatellus* (Ju et al., 2017), the sRNA nov-miR-12 which negatively regulated the DNA-binding protein *pipsqueak* (*psq*) in chromatin remodeling in *D. melanogaster* (Y. Zheng et al., 2019), cytosol amino-peptidase-like (CAL) which are in the sperm acrosome and involved in fertilization in *La. striatellus* (Huang et al., 2019), two seminal fluid proteins (CG9334 and CG2668) with unknown function in *D. melanogaster* (Yuan et al., 2015), the histone chaperone *Hira* in *D. melanogaster* and *D. simulans* (Zheng et al., 2011), a Juvenile Hormone protein (JHI-26) involved in development in *D. melanogaster* (Liu et al., 2014), and the immunity-related gene *kenny* (*key*) in *D. melanogaster* (Biwot et al., 2019). Since misexpression of these host products in uninfected males mimic CI-like embryonic defects in a way that can be rescued by infected females, there is robust support that these products or their pathways are involved in CI. However, how these factors relate to cause CI remains unknown.

Additionally, infected *D. melanogaster, D. simulans, Ae. albopictus, Ae. polynesiensis,* and *T. urticae* males often have higher reactive oxygen species (ROS) in their testes than uninfected males (Brennan et al., 2012, 2008; Zug and Hammerstein, 2015). It has been hypothesized that this variation in ROS expression patterns is due to an elevated host immune response to *Wolbachia* infection (Zug and Hammerstein, 2015). However, multiple lines of evidence link ROS expression with CI. For example, increased ROS levels are consistently observed among CI inducing strains (Zug and Hammerstein, 2015) and ROS leads to DNA damage in spermatocytes in *D. simulans* (Brennan et al., 2012). Additionally, lipid hydroperoxide markers of ROS-induced oxidative damage are higher in infected *D. melanogaster* (Driver et al., 2004), and PCNA retention is another marker for DNA damage and is observed during the first mitosis of CI-affected embryos (Landmann et al., 2009). Interestingly, overexpression of the *D. melanogaster* gene *key* increases ROS levels and DNA damage in males when mimicking rescuable CI-like hatching and embryonic defects (Biwot et al., 2019). Together, these data support a role for ROS in CI's mechanism, but direct connections remain unknown and it remains unclear if the DNA damage induced by *Wolbachia*-associated ROS can be rescued or may otherwise lead only to infertilities.

Determining the genetic basis of CI.

The genetic basis of CI has remained elusive for decades. The intangibility of the CI genes was due in no small part to the inability to genetically engineer *Wolbachia* (Iturbe-Ormaetxe et al., 2007; Thiem, 2014) which has prevented the use of standard assays, such as knock-out libraries,

to systematically identify genes involved in *Wolbachia* phenotypes (Cameron et al., 2008; Ito et al., 2005; Yajjala et al., 2016). Progress became possible with the sequencing of the *w*Mel *Wolbachia* genome in 2004 which revealed it to contain numerous mobile elements including phages (Wu et al., 2004). *w*Mel's genome is also enriched with ankyrins, proteins involved in protein-protein interactions and common among eukaryotes but relatively rare in bacteria (Al-Khodor et al., 2010). Additionally, the *w*Bm genome from *Brugia malayia*, a mutualistic strain found in nematodes, was sequenced in the following years and does not contain phages (Foster et al., 2005). These findings motivated the hypotheses that ankyrin genes and/or phage genes may be involved in CI.

The first study attempting to functionally dissect the genetic basis of CI generated a list of 12 gene candidates in the *w*Mel genome based on putative host interaction: nine ankyrin genes (WD0294, WD0385, WD0498, WD0514, WD0550, WD0633, WD0636, WD0754, and WD0776), two virulence-related genes (WD0579 and WD0580), and one phage-associated methylase gene (WD0594) (Yamada et al., 2011). Since *Wolbachia* are not genetically tractable, alternative ways of testing the genetic basis that did not require genetic manipulation of *Wolbachia* were required. Cleverly, Yamada et al. used the powerful genetic toolbox of *D. melanogaster* to test their gene candidates. More specifically, they used the GAL4-UAS system where a genetic construct containing a GAL4 transcription factor, native to yeast, is expressed under a tissue-specific promoter and then binds to an upstream activating site (UAS) engineered upstream of their candidate genes inside the *D. melanogaster* chromosome (Duffy, 2002). However, transgenic expression revealed that none of these genes could cause CI (Yamada et al., 2011). These results were supported by findings that variation in transcriptional regulation or sequence of *Wolbachia*'s ankyrin genes is not correlated with a strain's ability to induce CI (Olivier Duron et al., 2007; Papafotiou et al., 2011).

The studies presented in this dissertation aim to answer the genetic basis of CI (Chapters II-IV) and further characterize the genes involved (Chapters V and VI). First, Chapter II demonstrates that two genes from *Wolbachia*'s prophage WO that we call cytoplasmic incompatibility factors A and B (*cifA* and *cifB*) are responsible for causing CI. Chapter III then demonstrates that despite *cifA* contributing to CI-induction when expressed in males, it can rescue CI when expressed in females. Chapter IV further optimizes the expression system used to test these genes and reveals that complete CI and rescue phenotypes can be synthetically engineered

in the absence of *Wolbachia*-infection using the GAL4-UAS system in *D. melanogaster*. Chapter V then tests if CI gene homologs are capable of causing and rescuing CI and in the process reveals that variation in *cifB*, but not *cifA*, significantly influences the phenotypic output of the proteins when expressed in *D. melanogaster*. Chapter VI tests the impact of mutating conserved amino acids in predicted domains on phenotypic output and demonstrates that *cifA* has an overlapping function in CI and rescue and *cifB* is not amenable to changes in conserved sites. These works resolve the genetic basis of CI and make significant strides to understand how *cif* genetic variation can influence phenotypic output.

Chapter II.

Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility[*]

Introduction

The genus *Wolbachia* is an archetype of maternally inherited intracellular bacteria that infect the germline of numerous invertebrate species worldwide. They can selfishly alter arthropod sex ratios and reproductive strategies to increase the proportion of the infected matriline in the population. The most common reproductive manipulation is cytoplasmic incompatibility, which results in embryonic lethality in crosses between infected males and uninfected females. Females infected with the same *Wolbachia* strain rescue this lethality. Despite more than 40 years of research (Yen and Barr, 1971) and relevance to symbiont-induced speciation (Brucker and Bordenstein, 2012a; Shropshire and Bordenstein, 2016), as well as control of arbovirus vectors (Dutra et al., 2016; O'Connor et al., 2012; Walker et al., 2011) and agricultural pests (Zabalou et al., 2004), the bacterial genes underlying cytoplasmic incompatibility remain unknown. Here we use comparative and transgenic approaches to demonstrate that two differentially transcribed, co-diverging genes in the eukaryotic association module of prophage WO (Bordenstein and Bordenstein, 2016) from *Wolbachia* strain *w*Mel recapitulate and enhance cytoplasmic incompatibility. Dual expression in transgenic, uninfected males of *Drosophila melanogaster* crossed to uninfected females causes embryonic lethality. Each gene additively augments embryonic lethality in crosses between infected males and uninfected females. Lethality associates with embryonic defects that parallel those of wild-type cytoplasmic incompatibility and is notably rescued by *w*Mel-infected embryos in all cases. The discovery of cytoplasmic incompatibility factor genes *cifA* and *cifB* pioneers genetic studies of prophage WO induced reproductive manipulations and informs the continuing use of *Wolbachia* to control dengue and Zika virus transmission to humans.

[*] This chapter is published in 2017 in *Nature*, *543*(7644), 243-247 with Daniel LePage and Jason Metcalf as first authors. Jungmin On, Jessie Perlmutter, Dylan Shropshire, Emily Layton, Lisa Funkhouser-Jones, and John Beckmann were co-authors. Seth Bordenstein was senior author. I contributed Figures IIIb and IVg.

Results and discussion

We hypothesized that the genes responsible for cytoplasmic incompatibility (CI) (Figure A-1a) are present in all CI-inducing *Wolbachia* strains and absent or divergent in non-CI strains; we also predicted that these genes are expressed in the gonads of infected insects. To elucidate CI candidates, we determined the core genome shared by the CI-inducing *Wolbachia* strains *w*Mel (from *D. melanogaster*), *w*Ri (from *Drosophila simulans*), *w*Pip (Pel strain from *Culex pipiens*), and *w*Rec (from *Drosophila recens*), while excluding the pan-genome of the mutualistic strain *w*Bm (from *Brugia malayi*). This yielded 113 gene families representing 161 unique *w*Mel genes (Figure II-1a; Table A-1). We further narrowed this list by comparing it with (1) homologues of genes previously determined by comparative genomic hybridization to be absent or divergent in the strain *w*Au (Ishmael et al., 2009), a non-CI strain, (2) homologues to genes highly expressed at the RNA level in *w*VitA-infected *Nasonia vitripennis* ovaries, and (3) homologues detected at the protein level in *w*Pip (Buckeye)-infected *C. pipiens* ovaries. We included ovarian data with the reasoning that CI genes might be generally expressed in infected reproductive tissues, or that the CI induction and rescue genes might be the same. Remarkably, only two genes, *w*Mel locus tags WD0631 and WD0632, were shared among all four gene subsets (Figure II-1b; Table A-2-4). Notably, the homologue of WD0631 in the *Wolbachia* strain *w*Pip was found at the protein level in the fertilized spermathecae of infected mosquitoes, lending support to its role in reproductive manipulation (Beckmann and Fallon, 2013).

We analysed the evolution and predicted protein domains of these two genes and found that their homologues are always paired within the eukaryotic association module of prophage WO8, and they co-diverged into three distinct phylogenetic groups that we designate types I, II, and III (Figure II-1c, e; Table A-5). These relationships are not evident in the phylogeny of the *Wolbachia* cell division gene *ftsZ*, which exhibits the typical bifurcation of supergroup A and B *Wolbachia* (Figure A-1b), or in the phylogeny of prophage WO baseplate assembly gene *gpW* (Figure A-1c). This suggests that homologues of WD0631 and WD0632 evolve under different evolutionary pressures than genes in the core *Wolbachia* genome or in a structural module of phage WO.

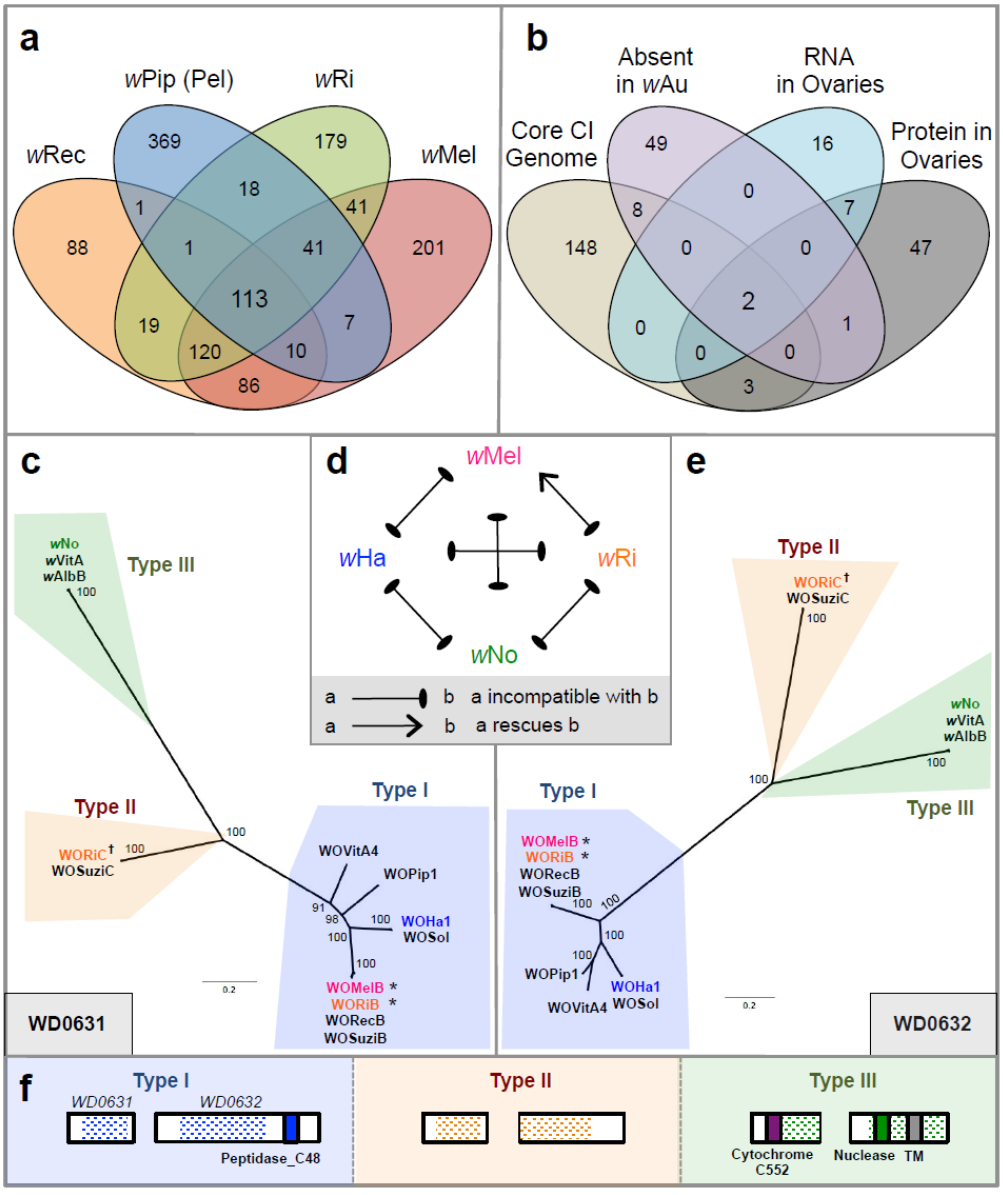

Figure II-1. Comparative analyses reveal WD0631 and WD0632 in the eukaryotic association module of prophage WO as candidate CI genes.

a, Venn diagram illustrating the number of unique and shared gene families from four CI-inducing *Wolbachia* strains. b, Venn diagram illustrating the number of unique and shared wMel genes matching each criteria combination. c, e, Bayesian phylogenies of (c) WD0631 and (e) WD0632 and their homologues, on the basis of a core 256-amino-acid (aa) alignment of WD0631 reciprocal BLASTp hits and a core 462-aa alignment of WD0632 reciprocal BLASTp hits. When multiple similar copies exist in the same strain, only one copy is shown. Consensus support values are shown at the nodes. Both trees are based on the JTT+ G model of evolution and are unrooted. d, CI patterns correlate with WD0631/ WD0632 sequence homology. wRi rescues wMel and both share a similar set of homologues (*). The inability of wMel to rescue wRi correlates with a type (†) that is present in wRi but absent in wMel. Likewise, bidirectional incompatibility of all other crosses correlates to divergent homologues. This diagram was adapted from ref. 30. f, Protein architecture of the WD0631 and WD0632 types is conserved for each clade and is classified according to the WD0632-like domain. TM, transmembrane. Dotted shading represents the region of shared homology used to construct phylogenetic trees. For c and e, the WO-prefix indicates a specific phage WO haplotype and the w-prefix refers to a 'WO-like island', a small subset of conserved phage genes, within that specific *Wolbachia* strain.

Type I variants are the most prevalent among ten sequenced *Wolbachia* strains, and are always associated with large prophage WO regions that often lack tail genes (Figure A-2); it is unclear whether these WO regions forge fully intact or defective interfering particles. The functions of type I WD0631 homologues are unknown, although type I WD0632 homologues contain weak homology to a putative Peptidase_C48 domain (*w*Mel, National Center for Biotechnology Information (NCBI) conserved domain $E = 6.69 \times 10{-}4$, Figure II-1f), a key feature of Ulp1 (ubiquitin-like-specific protease) proteases10. Type II variants are located within more complete phage haplotypes (Figure A-2), but the WD0632 homologues are truncated and lack recognized protein domains (Figure II-1f). Notably, all *Wolbachia* strains that contain type II variants contain at least one pair of intact type I variants. Type III variants possess WD0631 homologues with a weakly predicted cytochrome C552 domain involved in nitrate reduction (*w*No, NCBI conserved domain $E = 3.79 \times 10{-}3$), while type III WD0632 homologues contain weak homology to the PD-(D/E)XK nuclease superfamily (*w*No, NCBI conserved domain $E = 1.15 \times 10{-}3$) and to a transmembrane domain predicted by the transmembrane hidden Markov model (Krogh et al., 2001) (Figure II-1f). Finally, a putative type IV variant encoding a carboxy (C)-terminal PD-(D/E)XK nuclease superfamily (NCBI conserved domain $E = 3.69 \times 10{-}3$) was identified in *Wolbachia* strains *w*Pip and *w*AlbB, but not included in phylogenetic analyses because the WD0632 homologues are highly divergent (28% identity across 17% of the protein) and do not appear in reciprocal BLASTp analyses. The predicted functions of type III and IV protein domains are not well understood, but a homologue of the putative nuclease domain was previously found in a selfish genetic element that mediates embryonic lethality in *Tribolium* beetles (Lorenzen et al., 2008). Uncertain annotations and substantial unknown sequence across all of the phylogenetic types necessitate caution in extrapolating definitive gene functions. Importantly, the region of shared homology among the WD0632 homologues (Figure II-1f) is outside the putative C-terminal Peptidase_C48 domain, suggesting that the unannotated regions represent an ancestral CI sequence core that warrants closer inspection.

Consistent with a role in CI, the degree of relatedness and presence/ absence of homologues of WD0631 and WD0632 between *Wolbachia* strains correlates with known patterns of bidirectional incompatibility (Figure II-1d). Among the strains *w*Ri, *w*Ha, and *w*No, only *w*Ri rescues *w*Mel-induced CI in same-species crosses (Poinsot et al., 1998; Zabalou et al., 2008). We

postulate that this is due to *w*Ri and *w*Mel sharing highly related type I homologues (99% amino-acid identity), and thus probably sharing a rescue factor, while *w*Ri also has a type II homologue that may explain its ability to induce CI against *w*Mel. Meanwhile, bidirectionally incompatible pairs are highly divergent, with only 29–68% amino-acid identity (Figure A-3a). Additionally, variation in CI strength between strains appears to correlate with the number of copies of the WD0631/WD0632 pair (Figure A-3b). Strains with only one copy, such as *w*Mel, have a comparatively weak CI phenotype, while those with two or three copies, such as *w*Ri and *w*Ha, cause strong CI (Poinsot et al., 1998).

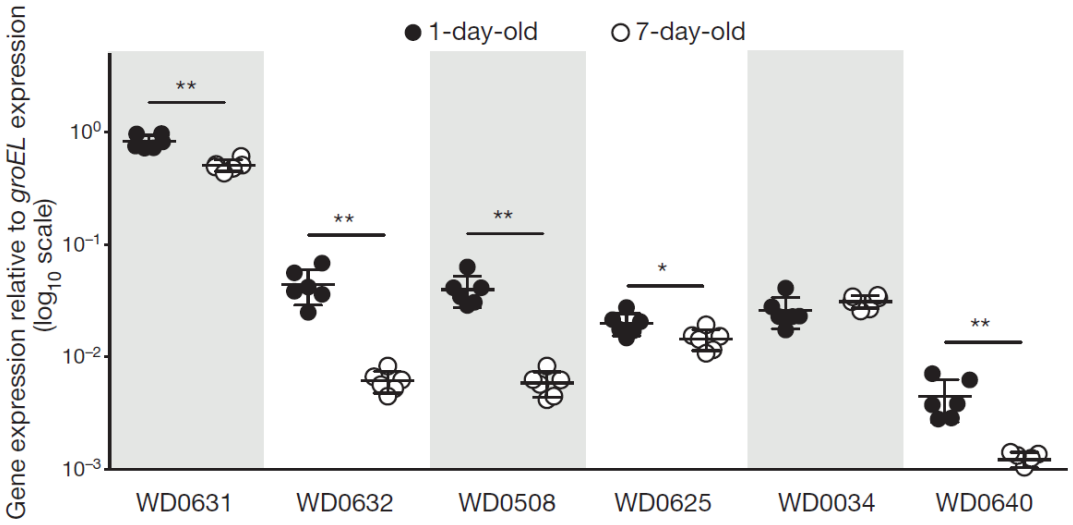

Figure II-2. Relative expression of CI candidate and prophage WO genes decreases as males age.
RNA expression in 1-day-old versus 7-day-old testes, normalized to expression of *groEL* in *w*Mel-infected *D. melanogaster* testes from the fastest-developing males. Values denote $2^{-\Delta Ct}$. $n = 6$ independent pools of 20 testes for each group. Bars, mean ± s.d. * $P < 0.05$, ** $P < 0.01$ by Mann–Whitney *U*-test. This experiment was performed once.

Given the various lines of evidence that associate these two genes with CI, we next examined the functional role of WD0631 and WD0632 in CI. For comparison, the following control genes were also used: WD0034, which has a predicted PAZ (Piwi, Argonaut, and Zwille) domain (NCBI conserved domain E = 1.85 × 10−18); WD0508, a prophage gene annotated as a putative transcriptional regulator with two helix–turn–helix domains (NCBI conserved domain E = 9.29 × 10−12) in the Octomom region; and WD0625, a prophage gene annotated as a DUF2466 with a JAB1/MPN/Mov34 metalloenzyme (JAMM) domain (NCBI conserved domain E = 1.60 × 10−41). We first examined the expression of these genes in the testes of *w*Mel-infected, 1-day-old

and 7-day-old *D. melanogaster* males. Since CI strength decreases significantly in aged males (Reynolds and Hoffmann, 2002), we predicted that a CI factor would be expressed at a lower level in 7-day-old males versus 1-day-old males that both emerged on day 1 of the cross. Indeed, WD0631 and WD0632 showed a significantly lower transcription level in aged males (Figure II-2). Moreover, WD0631 exhibited 18.6- and 83.0-fold higher expression than WD0632 for young and aged males, respectively (Figure II-2). Coupled with RNA-seq expression data (Gutzwiller et al., 2015) and operon predictor algorithms, evidence suggests that these genes are not generally acting as an operon in *w*Mel. Both prophage-associated control genes, WD0508 and WD0625, also exhibited this age-dependent expression pattern, but the non-prophage gene WD0034 did not (Figure II-2). WD0640, which encodes prophage WO structural protein GpW, was also reduced in older males, suggesting that prophage genes in general are relatively downregulated in 7-day-old testes (Figure II-2). The phenomenon of decreased CI in older males was not due to decreases in *Wolbachia* titre over time, as the copy number of *Wolbachia groEL* relative to *D. melanogaster rp49* increased as males aged, and there was no significant difference in absolute *Wolbachia* gene copies between 1-day-old and 7-day-old males (Figure A-4a, b). Since CI expression is also correlated with male development time, we examined gene expression in early emerging 'older brothers' (emerged on day 1) and later emerging 'younger brothers' (emerged on day 5). Expression was statistically equivalent for WD0631 (Figure A-4c), and slightly reduced in younger brothers for WD0632 (Figure A-4d). These results are consistent with a small younger brother effect (Yamada et al., 2007), although we did not observe a statistically significant effect on CI penetrance (Figure A-4e).

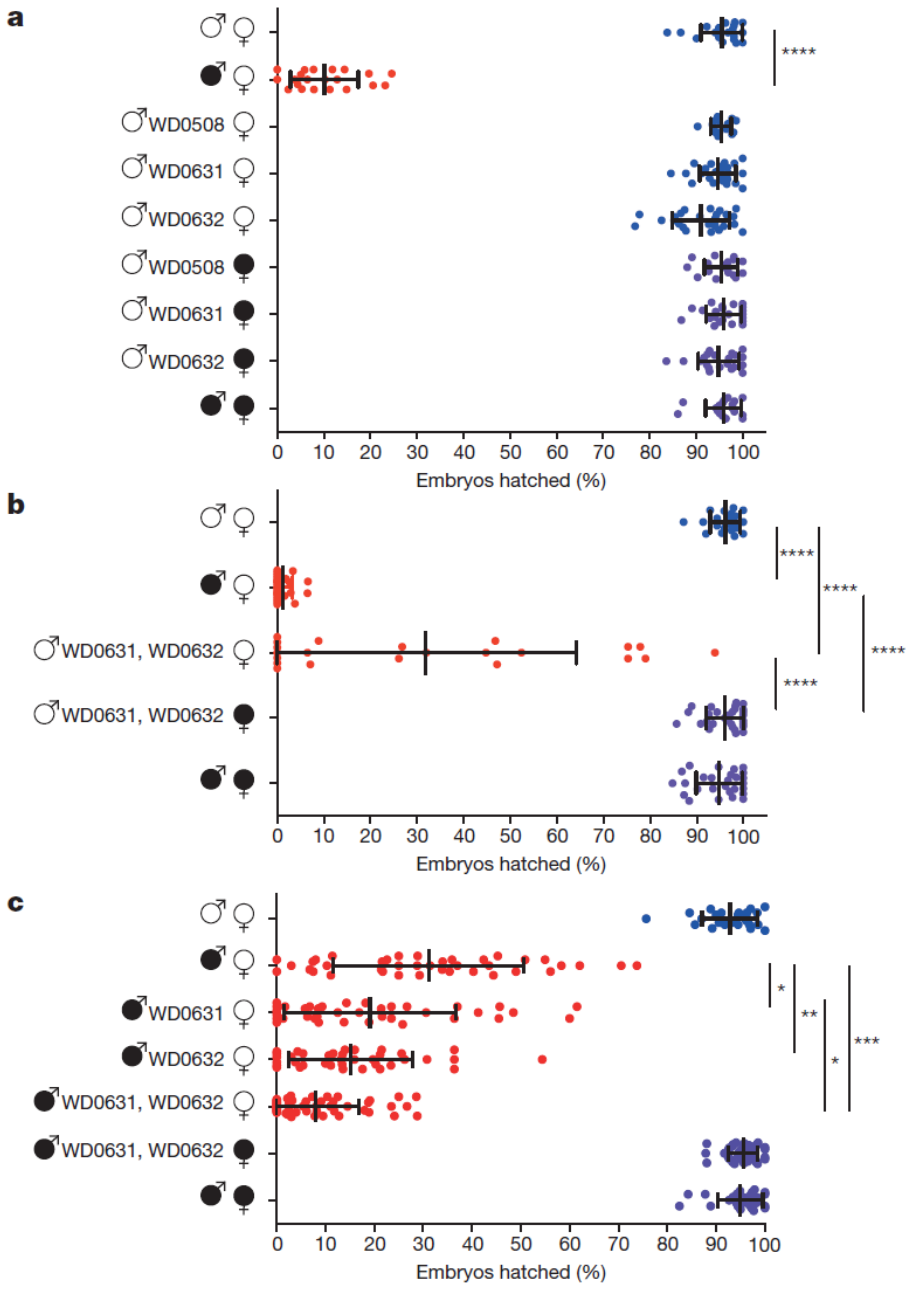

Figure II-3. Dual expression of WD0631 (*cifA*) and WD0632 (*cifB*) is necessary to induce CI-like defects.
a–c, Hatch rate assays used the fastest developing males that were aged either (a, b) 1 day or (c) 2–4 days in parental crosses; older males express incomplete CI. Parental infection status is designated with filled symbols for *w*Mel-infected parents or open symbols for uninfected parents. Transgenic flies are labelled with their transgene to the right of their male/female symbol. Unlabelled symbols represent WT flies. Data points are coloured according to the type of cross: blue, no CI; red, CI crosses; purple, rescue crosses with *w*Mel-infected females. n = 24–54 for each group. Bars, mean ± s.d. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ by analysis of variance (ANOVA) with Kruskal–Wallis test and Dunn's multiple test correction. Statistical comparisons are between all groups (a, b); or between CI crosses only (c). All experiments were performed at least twice, except for the increase of WT CI by WD0631, which was done once.

To directly test involvement of these genes in CI, we generated transgenic *D. melanogaster* expressing genes using an upstream activating sequence (UAS), since *Wolbachia* itself cannot be genetically transformed. We used a *nanos*-Gal4 driver line for tissue-specific expression in the male and female germline (Rørth, 1998; White-Cooper, 2012). We assessed CI by measuring the percentage of embryos that hatched into larvae. While wild-type (WT) CI between infected males and uninfected females led to significantly reduced hatch rates, expressing each of four candidate transgenes in uninfected (fastest-developing, 1 day old) males did not affect hatch rates when crossed to uninfected females (Figure II-3a; Figure A-5a). These transgenes also had no effect on sex ratios (Figure A-5b; Figure A-6a). There were no phenotypic effects despite confirmed expression of each transgene in the testes (Figure A-7a-d).

As WD0631 and WD0632 are adjacent, coevolving genes, we reasoned that dual expression of WD0631 and WD0632 might be required to induce CI. Indeed, expression of both transgenes in the same males significantly reduced hatch rates by 68% compared with uninfected WT crosses (Figure II-3b), with no effect on sex ratios (Figure A-6b). Roughly half of the crosses with transgenic males yielded hatch rates within the range observed in WT CI (3.8 ± 5.6% hatch rate). Interestingly, there was a strong positive correlation between hatch rate and clutch size when both transgenes were expressed (rs = 0.7; P = 0.0003), but not in WT CI, suggesting that dilution of transgene products across larger clutches may explain variation in transgene-induced CI. It is also possible that full transgene induction of CI requires other factors, or that transgenes are not expressed at the ideal time or place for complete CI, although transgene expression in adult testes was confirmed (Figure A-7c, d).

Importantly, transgene- induced lethality is fully rescued in embryos of *w*Mel-infected females (Figure II-3b), indicating that these genes produce probable CI factors rather than artefacts that reduce hatch rates through off target effects that would not be rescued. We therefore name and hereafter refer to these genes as cytoplasmic incompatibility factor A (*cifA*) and B (*cifB*) for WD0631 and WD0632, respectively. Type II, III, and IV homologues are designated *cif-like* until experimental evidence demonstrates that they recapitulate CI. To test whether *cifA* (WD0631) and *cifB* (WD0632) transgenes act additively with *Wolbachia* to enhance WT CI levels, *w*Mel-infected male flies expressing either transgene were aged 2–4 days to lower WT CI penetrance before crossing with uninfected females. In support of transgene-induced enhancement of CI, hatch rates in these aged males decreased significantly compared with WT CI crosses of the same age (Figure

II-3c), with no effect on sex ratios (Figure A-6c). In this context, wherein aged flies cause a weaker level of WT CI, the transgenes appear to add to the quantity of CI effectors in *w*Mel-infected tissues, causing stronger CI overall. This effect was not observed when control transgenes WD0508 or WD0625 were expressed individually in *w*Mel-infected males (Figure A-8a, b). Moreover, dual expression of *cifA* and *cifB* in *w*Mel-infected flies reduced hatch rates further than either gene alone, yet was still fully rescued in embryos of *w*Mel-infected females (Figure II-3c). Adding WD0625 to *cifB* in aged *w*Mel-infected males did not increase CI beyond *cifB* alone (Figure A-8b), and had no effect on embryonic hatch rates from crosses with 1-day-old uninfected males (Figure A-8c). Finally, none of these gene combinations affected offspring sex ratios (Figure A-9). Taken together, these findings support the central conclusion that *cifA* and *cifB* are both necessary to induce the CI phenotype, and they do not represent an artefact of the transgenic system.

To rule out the possibility that transgene-induced enhancement of CI in infected lines is due to increased *Wolbachia* titres, we quantitated amplicons of single-copy genes from *Wolbachia* and *D. melanogaster*. Although there were some differences in *Wolbachia* titres between infected transgenic lines (Figure A-10a-c), the variation did not correlate with induction or magnitude of CI, signifying that decreased offspring viability was due to the direct effect of transgene products rather than changes in *Wolbachia* density. Most notably, densities significantly increased in infected flies expressing the control Octomom transgene WD0508 (Figure A-10a) but did not enhance CI (Figure A-8).

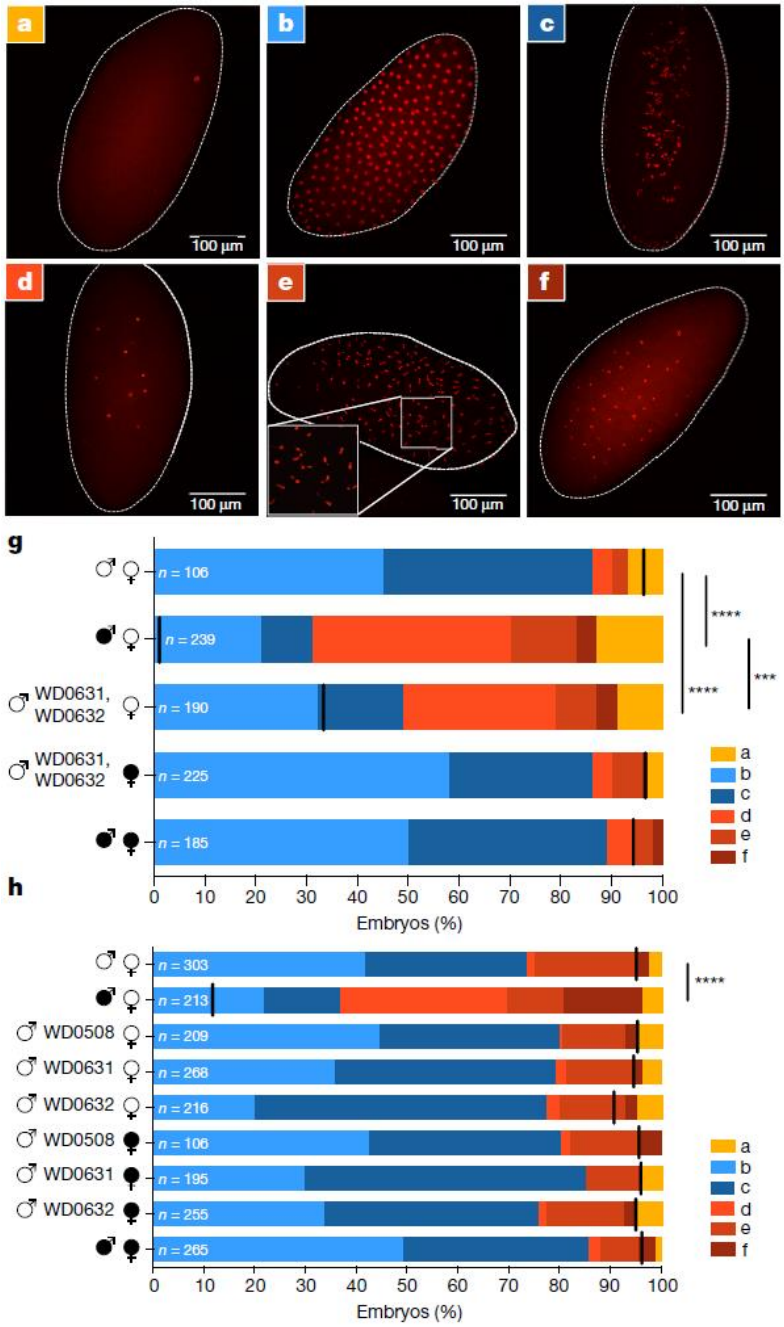

Figure II-4. Dual expression of WD0631 (*cifA*) and WD0632 (*cifB*) recapitulates CI-associated embryonic defects.

a–f, Representative embryo cytology is shown for (a) unfertilized embryos, (b) normal multi-nucleated embryos at 1 h of development, (c) normal embryos near 2 h of development in which nuclei begin to migrate to the periphery of the cytoplasm, and three different mitotic abnormalities: (d) failure of nuclear division after two to three mitoses, (e) chromatin bridging, and (f) regional mitotic failure. g, The number of embryos with each cytological phenotype resulting from the indicated crosses is shown. Infection status is designated with filled symbols for *w*Mel-infected parents or open symbols for uninfected parents. Transgenic flies are labelled with their transgene to the right of their male/female symbol. Unlabelled symbols represent WT flies. Black lines on each graph indicates mean hatch rate for the cross. *** $P < 0.001$, **** $P < 0.0001$ by two-tailed Fisher's exact test comparing normal (phenotypes b and c) with abnormal (phenotypes a, d–f) for each cross. h, Quantitation of cytological defects in control crosses. Cytology for g was performed twice and for h once.

Next, we tested whether transgene-induced CI associates with canonical cytological defects observed in *Wolbachia*-induced CI. Although CI is typically thought to cause failure of the first mitotic division (Landmann et al., 2009; Serbus et al., 2008), nearly half of the embryonic arrest in incompatible crosses occurs during advanced developmental stages in *D. simulans* (Callaini et al., 1997; Lassy and Karr, 1996), *Aedes polynesiensis* (Wright and Barr, 1981), and *C. pipiens* (Duron and Weill, 2006). We examined embryos from control and experimental crosses after 1–2 h of development and binned their cytology into one of six phenotypes. While a few embryos in each cross were unfertilized (Figure II-4a), most embryos in compatible crosses were either in normal late-stage preblastoderm (Figure II-4b) or syncytial blastoderm stages (Figure II-4c). In WT CI, significantly more embryos exhibited three defects: arrest of cellular division after two to three mitotic divisions (Figure II-4d), later stage arrest associated with chromatin bridging, as is classically associated with strong CI in *D. simulans* (Landmann et al., 2009) (Figure II-4e), or arrest associated with regional failure of division in one embryo region (Figure II-4f). After blindly scoring embryo cytology, we determined that aberrant phenotypes (a, d, e, and f) were significantly more common in the offspring of dual *cifA*/*cifB* transgenic males mated to uninfected females, but these abnormalities were rescued in embryos from *w*Mel-infected females (Figure II-4g). These effects were not seen with control transgene WD0508 or with singular expression of *cifA* or *cifB* (Figure II-4h). These data again validate that transgene-induced CI, as measured through cytological defects, recapitulates WT CI. Most of the embryos arrest after two to three mitotic divisions.

This study identifies, for the first time, two differentially transcribed and codiverging prophage WO genes that recapitulate and enhance CI. These rapidly evolving genes are not chromosomal *Wolbachia* genes per se, but rather occur widely in the eukaryotic association module of prophage WO (Bordenstein and Bordenstein, 2016). This module notably contains genes with amino-acid sequences homologous to eukaryotes or annotated to interact with animal cells, although *cifA* and *cifB* do not appear to have eukaryotic homology. CI can therefore be categorized as a prophage WO-induced phenotype rather than a *Wolbachia*-induced phenotype. We name the genes and close homologues cytoplasmic incompatibility factors A and B for WD0631 and WD0632, respectively. The *cif* name is conservatively grounded in phenotype and makes no assumptions about mechanism, which is notable because there are unannotated core

regions throughout the *cif* genes that may have as much bearing on mechanism as the annotated domains.

The discovery of *cifA* and *cifB* genes that functionally recapitulate and enhance CI is the first inroad in solving the genetic basis of reproductive parasitism, a phenomenon induced worldwide in potentially millions of arthropod species (Zug and Hammerstein, 2012). These prophage WO genes have implications for microbe-assisted speciation, because they can underlie CI-induced hybrid lethality observed between closely related species of *Nasonia* and *Drosophila* (Bordenstein et al., 2001; Jaenike et al., 2006). Finally, *cifA* and *cifB* are important for arthropod pest and vector control strategies, as they could be an alternative or adjunct to current *Wolbachia*-based efforts aimed at controlling agricultural pests or curbing arthropod-borne transmission of infectious diseases (Dutra et al., 2016; O'Connor et al., 2012; Walker et al., 2011; Zabalou et al., 2004).

Materials and methods

Comparative genomics and transcriptomics.

MicroScope (Vallenet et al., 2009) was used to select the set of genes comprising the core genomes of CI-inducing *Wolbachia* strains *w*Mel (NC_002978.6) (Wu et al., 2004), *w*Ri (NC_012416.1) (Klasson et al., 2009), *w*Pip (Pel) (NC_010981.1) (Klasson et al., 2008), and the recently sequenced *w*Rec (GCA_000742435.1) (Metcalf et al., 2014), while excluding the pan-genome of the mutualistic strain *w*Bm (NC_006833.1) (Foster et al., 2005), using cutoffs of 50% amino-acid identity and 80% alignment coverage. For the 'absent in *w*Au' criterion, *w*Au microarray data were obtained from the original authors (Ishmael et al., 2009) and genes that were present in CI-inducing strains *w*Ri and *w*Sim but absent or divergent in the non-CI strain *w*Au were selected.

For ovarian transcriptomics, 1-day-old females from *w*VitA-infected *N. vitripennis* 12.1 were hosted as virgins on *Sarcophaga bullata* pupae (Werren and Loehlin, 2009) for 48 h to stimulate feeding and oogenesis. Females were then dissected in RNase-free 1× PBS buffer, and their ovaries were immediately transferred to RNase-free microcentrifuge tubes in liquid nitrogen. Fifty ovaries were pooled for each of three biological replicates. 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), and ethanol precipitation was performed. PCR of RNA samples with *Nasonia* primers NvS6KQTF4 and NVS6KQTR4 (Bordenstein and Bordenstein, 2011) confirmed that all samples were free of DNA contamination. RNA concentrations were measured with a Qubit 2.0 Fluorometer (Life Technologies) using the RNA HS Assay kit (Life Technologies), and approximately 5 μ g of total RNA from each sample was used as input for the MICROB*Enrich* Kit (Ambion) to enrich for *Wolbachia* RNA in the samples. Bacterial-enriched RNA was then ethanol-precipitated, and rRNA was depleted from the samples using the Ribo-Zero Magnetic kit (Illumina) according to the manufacturer's protocol. Approximately 1.5 μ g of enriched, rRNA-depleted RNA for each replicate was shipped to the University of Rochester Genomics Research Center for sequencing. Library preparation was performed using the Illumina ScriptSeq version 2 RNA-Seq Library Preparation kit, and all samples were run multiplexed on a single lane of the Illumina HiSeq2500 (single-end, 100 base pair reads). Raw sequence reads were trimmed and mapped to the *w*VitA genome (PRJNA213627) in CLC Genomics Workbench 8.5.1 using a minimum length fraction of 0.9, a minimum similarity fraction of 0.8, and allowing one gene hit per read. With all three replicates combined, a total of 364,765 reads out of 41,894,651 (0.87%) mapped to the *w*VitA genome, with the remaining reads mapping to the *N. vitripennis* host genome (GCF_000002325.3). All *Wolbachia* genes with at least five RNA-seq reads, with the exception of the 16S and 23S RNA genes, were selected. For non-*w*Mel data sets, the closest homologues in *w*Mel were found using BLASTp in Geneious Pro version 5.5.6 (Kearse et al., 2012).

Protein extraction and mass spectrometry.

Protein was extracted from *C. pipiens* tissues as described previously (Beckmann and Fallon, 2013). Ovaries from 30 *w*Pip (Buckeye)- infected mosquitoes were dissected in 100% ethanol and collected in a 1.5 ml tube filled with 100% ethanol. Pooled tissues were sonicated at 40 mA for 10 s in a Kontes GE 70.1 ultrasonic processor, and trichloroacetic acid was added to a final concentration of 10% (v/v). After centrifugation at 13,000 r.p.m. in a microcentrifuge, pellets were washed with acetone:water (9:1), dried, and stored at − 20 °C. Samples were directly submitted to the University of Minnesota's Center for Mass Spectrometry and Proteomics for iTRAQ (isobaric tagging for relative and absolute quantification) analysis. Proteins were sorted according to their relative abundance as determined by the number of spectra from the single most

abundant peptide. Because proteins can often produce varying amounts of detectable tryptic peptides, depending upon protein size and lysine/arginine content, we counted only the single most abundant peptide for each protein. This quantification was justified by a previous report (Beckmann et al., 2013) showing that the two most abundant proteins are the *Wolbachia* surface protein (WSP; WP = 007302328.1) and another putative membrane protein (WP0576; WP = 012481859.1). Only proteins with at least three unique peptides (95% confidence) detected were reported; using this criterion the false discovery rate was zero.

Evolutionary analyses.

WD0631 (NCBI accession number AAS14330.1) and WD0632 (AAS14331.1) from *w*Mel were used as queries to perform a BLASTp search of NCBI's nonredundant (nr) protein sequence database with algorithm parameters based on a word size of six and BLOSUM62 scoring matrix (Johnson et al., 2008). Homologues were selected on the basis of the satisfaction of three criteria: (1) $E = \leq 10{-}20$, (2) query coverage greater than 60%, and (3) presence in fully sequenced *Wolbachia* and/or phage WO genomes. FtsZ and gpW proteins were identified for all representative *Wolbachia* and phage WO genomes, respectively. Protein alignments were performed using the MUSCLE plugin (Edgar, 2004) in Geneious Pro version 8.1.7 (Kearse et al., 2012); the best models of evolution, according to corrected Akaike information criteria (Hurvich and Tsai, 1993), were estimated using the ProtTest server (Abascal et al., 2005); and phylogenetic trees were built using the MrBayes plugin (Ronquist et al., 2012) in Geneious. Putative functional domains were identified using NCBI's BLASTp, Wellcome Trust Sanger Institute's PFAM database (Finn et al., 2016), a transmembrane hidden Markov model (Krogh et al., 2001), and EMBL's Simple Modular Architecture Research Tool (SMART) (Letunic et al., 2012). WD0631/WD0632 protein homology (percentage amino-acid identity) was based on a 1:1 BLASTp analysis for each pair. Prophage/WO-like island association for each pair of genes was based on prophage regions identified in a previous study8.

Gene expression assays and Wolbachia *titres.*

For the male age effect, native expression of CI candidates was tested with RT–qPCR on replicate pools of 20 pairs of testes from the fastest-developing virgin males that were aged 1 day or 7 days. RNA was extracted with a Qiagen RNeasy mini kit, DNase treated with TURBO DNase

(Life Technologies), and cDNA generated with Superscript III Reverse Transcriptase (Invitrogen). Primer sequences are listed in Table A-6. Quantitative PCR was performed on a Bio-Rad CFX-96 Real-Time System using iTaq Universal SYBR Green Supermix (Bio-Rad). Thirty cycles of PCR were performed against positive controls (extracted DNA), negative controls (water), RNA, and cDNA with the following conditions: 95 °C 2 min, 30× (95 °C 15 s, 56 °C 30 s, 72 °C 30 s), 72 °C 5 min. Values of $2^{-\Delta Ct}$ between the target gene and housekeeping gene *groEL* were used to determine relative gene expression. These experiments were performed once with multiple replicates for each condition.

For experiments on the younger brother effect, replicate pools of 20 pairs of testes were collected from the fastest-developing virgin males that emerged on the first day (older brothers) or fifth day (younger brothers). Male siblings for the younger brother effect analysis were also collected concurrently for hatch rates as described for hatch rate assays by crossing the *w*Mel-infected males to 3- to 5-day-old *w*Mel-infected or uninfected females. RNA was extracted using the Direct-zol RNA MiniPrep Kit (Zymo), DNase treated with DNA-free (Ambion, Life Technologies), cDNA was generated with SuperScript VILO (Invitrogen), and RT–qPCR was run using iTaq Universal SYBR Green (Bio-Rad). Primers, PCR conditions, and analysis were the same as for the male age effect above. These experiments were performed once with multiple replicates for each condition.

For gene expression in Figure A-7, six pools of six pairs of testes were dissected from parents used in hatch rate assays from a repeat of Figure II-3a and Figure A-5. In samples designated ‘High CI’ and ‘No CI’, the males correspond to crosses that had lower or normal hatch rates, respectively. For all other samples, the flies used were chosen at random. RNA was extracted using the same method as the younger brother experiment above. Thirty cycles of PCR were performed against positive controls (extracted DNA), negative controls (water), RNA, and cDNA with PCR conditions described above. Gel image size and brightness were adjusted in some cases for clarity. These experiments were performed once.

For the *Wolbachia* titres, pools of testes were dissected from 15 males in ice-cold PBS. For Figure A-10a-c, brothers of those used in the corresponding hatch rates were used. DNA was extracted using a Gentra Puregene Tissue kit (Qiagen). qPCR was done as described above. Absolute quantification was achieved by comparing all experimental samples with a standard curve generated on the same plate. Primers are listed in Table A-6. qPCR conditions were as

follows: 50 °C 10 min, 95 °C 5 min, 40× (95 °C 10 s, 55 °C 30 s), 95 °C 30 s. To obtain a more accurate *Wolbachia*:host cell ratio, it was assumed that each host cell had two copies of *rp49* and each *Wolbachia* cell had one copy of *groEL*. These experiments were performed once but with a sample size of eight for each condition.

Fly rearing.

D. melanogaster were reared on a standard cornmeal- and molasses- based media. Stocks were maintained at 25 °C while virgin flies were stored at room temperature. During virgin collections, stocks were kept at 18 °C overnight and 25 °C during the day. All flies were kept on a 12-h light/dark cycle. *Wolbachia*-uninfected lines were generated through tetracycline treatment for three generations. Briefly, tetracycline was dissolved in ethanol and then diluted in water to a final concentration of 1 mg/ml. One millilitre of this solution was added to 50 ml of media (final concentration of 20 μ g/ml). Freshly treated media was used for each generation. Infection status was confirmed with PCR using Wolb_F and Wolb_R3 primers (Casiraghi et al., 2005), and flies were reared on untreated media for at least three additional generations to allow for mitochondrial recovery before being used (Chatzispyrou et al., 2015).

Transgenic flies.

Each CI candidate gene was cloned into the pTIGER plasmid for transformation and expression in *D. melanogaster* (Ferguson et al., 2012). pTIGER, a pUASp-based vector designed for germline expression, exhibits targeted integration into the *D. melanogaster* genome using PhiC31 integrase (Groth et al., 2004) and tissue-specific, inducible expression through the Gal4–UAS system (Southall et al., 2008). Cloning was performed using standard molecular biology techniques and plasmids were purified and Sangersequenced for confirmation before injection. At least 200 *D. melanogaster* embryos were injected per gene by Best Gene (Chino Hills, California), and transformants were selected on the basis of *w*+ eye colour. All transgenic lines were made in the *yw D. melanogaster* background, and each was an isofemale line derived from the offspring of a single transformant. Homozygous lines were maintained when possible, or heterozygous flies were maintained when homozygous transgenics were inviable (WD0625/CyO). WD0508 and WD0631 insertion was performed with the y^1 M{vas-int.Dm}ZH-2A w* ; P{CaryP}attP40 line. WD0625 was inserted into BSC9723 with the genotype y^1 M{vas-int.Dm}ZH-2A w* ; PBac{y+

-attP-3B} VK00002. WD0632 insertion was done using BSC8622 with the genotype y^1 *w*67c23; P{CaryP}attP2.

Hatch rate and sex ratio assays.

Parental females were either infected or uninfected y^1w* flies (*w*Mel-infected or uninfected) and aged for 2–6 days before crossing. Uninfected y^1w* flies were generated as described for transgenic lines. Parental males were created by crossing *nanos*-Gal4 virgin females (*w*Mel-infected or uninfected) with either y^1w* or UAS-candidate gene-transgenic, uninfected males. Only the first males emerging between 0 and 30 h from these crosses were used in CI assays to control for the younger-brother effect associated with CI (Yamada et al., 2007). To test whether CI can be increased by transgenes, virgin, day 1 males were aged for 2–4 days before crossing to reduce the level of WT CI. Within experiments, care was taken to match the age of males between experimental and control crosses. Thirty-two to 64 individual crosses were set up for each crossing condition. The flies used were chosen at random from the desired group on the basis of age, sex, and genotype. These sample sizes were based on previous studies of CI in *D. melanogaster* that detected significant differences between treatment groups (LePage et al., 2014).

To perform the hatch rate assays, a male and female pair was placed in an 8-ounce, round bottom, polypropylene *Drosophila* stock bottle. A grape-juice–agar plate with a small amount of yeast mix smeared on top was placed in the bottle opening and affixed with tape. To create grape juice–agar plates, 12.5 g of agar was mixed in 350 ml of de-ionized water and autoclaved. In a separate flask, 10 ml of ethanol was used to dissolve 0.25 g tegosept (methyl 4-hyrdoxybenzoate). Welch's grape juice (150 ml) was added to the tegosept mix, combined with the agar, and poured into lids from 35 ×10-mm culture dishes (CytoOne).

Hatch rate bottles were placed in an incubator at 25 °C overnight (~16 h). After this initial incubation, the grape plates were discarded and replaced with freshly yeasted plates. After an additional 24 h, the adult flies were removed and frozen for expression analysis, and the embryos on each plate were counted. The counting was not blinded. These plates were then incubated at 25 °C for 36 h before the number of unhatched embryos was counted. Larvae from each pair of flies were moved from these plates using a probe and placed in vials of standard fly media with one vial being used for each individual grape plate to be assayed for sex ratios at adulthood. A total of 10–20 vials were used for each cross type. Any crosses with fewer than 25 embryos laid were discarded

from the hatching analysis while vials with fewer than ten adults emerging were discarded from the sex ratio analysis. Statistical analysis and outlier removal, using the ROUT method, were performed using Graphpad Prism version 6 software.

Embryo cytology.

Embryos were collected in a fashion similar to hatch rate assays except bottles contained 60–80 females and 15–20 males. All flies used were brothers and sisters of those used during corresponding hatch rates. Embryo collections and hatch rates were performed side-by-side. After initial mating overnight, fresh grape plates with yeast were provided and removed after 60 min. The embryo-covered plates were then placed in the incubator at 25 °C for a further 60 min to ensure each embryo was at least 1–2 h old. Embryos were then moved to a small mesh basket and dechorionated in 50% bleach for 1–3 min. These were then washed in embryo wash solution (0.7% NaCl, 0.05% Triton X-100) and moved to a small vial containing ~ 2 ml heptane. An equal amount of methanol was added to the vial and then vigorously shaken for 15 s. After the embryos settled, the upper heptane layer and as much methanol as possible were removed, and the embryos were moved into ~ 500 µ l fresh methanol in a 1.5 ml microcentrifuge tube. Embryos were stored overnight at 4 °C. The old methanol was then removed and replaced with 250 µ l of fresh methanol along with 750 µ l of PBTA (1× PBS, 1% BSA, 0.05% Triton X-100, 0.02% sodium azide). After inverting the tube several times, the solution was removed and replaced with 500 µ l PBTA. Embryos were then rehydrated for 15 min on a rotator at room temperature. After rehydrating, the PBTA was replaced with 100 µ l of a 10 mg/ml solution of RNase A (Clontech Labs) and incubated at 37 °C for 2 h. The RNase was then removed and embryos were washed several times with PBS followed by a final wash with PBS–azide (1× PBS, 0.02% sodium azide). After removing the PBS–azide, embryos were mounted on glass slides with ProLong Diamond Antifade (Life Technologies) spiked with propidium iodide (Sigma-Aldrich) to a final concentration of 1 µ g/ml. Imaging was performed at the Vanderbilt University Medical Center Cell Imaging Shared Resource using a Zeiss LSM 510 META inverted confocal microscope. All scores were performed blind (researcher was not aware of which slide represented which cross) and image analysis was done using ImageJ software (Schneider et al., 2012). Matched scoring, where embryos were derived from a side-by-side hatch rate, was performed once for conditions shown in Figure II-4h and twice for Figure II-4g.

Statistical analyses.

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment, except scoring of cytology (Figure II-4), which was done blindly. All statistical analyses used GraphPad Prism software (either Prism 6 or online tools). When comparing gene expression levels or *Wolbachia* titres between two sets of data, we used a two-tailed, non-parametric Mann–Whitney *U*-test since it does not require a normal distribution of the data. For comparisons between more than two data sets, we used a non-parametric Kruskal–Wallis one-way analysis of variance test that, if significant, was followed by a Dunn's test of multiple comparisons. This allowed robust testing between all data groups while avoiding multiple test bias. For the cytology studies, embryos were classified as either ' normal' or 'CI-like' in a 2 × 2 contingency table, and statistical differences between the groups were calculated using a Fisher's exact test.

Data availability.

*w*VitA transcriptome data have been deposited in the Sequence Read Archive with Bioproject PRJNA319204 and BioSample SAMN04881412. *w*Pip-infected ovarian proteome data have been deposited in the Proteome Xchange Consortium via the PRIDE (Vizcaíno et al., 2016) partner repository with the data set identifier PXD004047. All other source data are available as Supplementary Information with this publication.

Chapter III.

One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila* melanogaster[†]

Abstract

Wolbachia are maternally inherited, intracellular bacteria at the forefront of vector control efforts to curb arbovirus transmission. In international field trials, the cytoplasmic incompatibility (CI) drive system of *w*Mel *Wolbachia* is deployed to replace target vector populations, whereby a *Wolbachia*-induced modification of the sperm genome kills embryos. However, *Wolbachia* in the embryo rescue the sperm genome impairment, and therefore CI results in a strong fitness advantage for infected females that transmit the bacteria to offspring. The two genes responsible for the *w*Mel-induced sperm modification of CI, *cifA* and *cifB*, were recently identified in the eukaryotic association module of prophage WO, but the genetic basis of rescue is unresolved. Here we use transgenic and cytological approaches to demonstrate that maternal *cifA* expression independently rescues CI and nullifies embryonic death caused by *w*Mel *Wolbachia* in *Drosophila melanogaster*. Discovery of *cifA* as the rescue gene and previously one of two CI induction genes establishes a "Two-by-One" model that underpins the genetic basis of CI. Results highlight the central role of prophage WO in shaping *Wolbachia* phenotypes that are significant to arthropod evolution and vector control.

Significance

The World Health Organization recommended pilot deployment of *Wolbachia*-infected mosquitoes to curb viral transmission to humans. Releases of mosquitoes are underway worldwide because *Wolbachia* can block replication of these pathogenic viruses and deterministically spread by a drive system termed cytoplasmic incompatibility (CI). Despite extensive research, the underlying genetic basis of CI remains only half-solved. We recently reported that two prophage

[†] This chapter is published in 2018 in *Proceedings of the National Academy of Sciences, 115*(19), 4987-4991 with myself as first author. Jungmin On, Emily Layton, and Helen Zhou were co-authors. Seth Bordenstein was senior author.

WO genes recapitulate the modification component of CI in a released strain for vector control. Here we show that one of these genes underpins rescue of CI. Together, our results reveal the complete genetic basis of this selfish trait and pave the way for future studies exploring WO prophage genes as adjuncts or alternatives to current control efforts.

Introduction

The bacteria *Wolbachia* occur in an estimated 40–52% of arthropod species (Weinert et al., 2015; Zug and Hammerstein, 2012) and 47% of the Onchocercidae family of filarial nematodes (Ferri et al., 2011), making them the most widespread bacterial symbiont in the animal kingdom (Zug and Hammerstein, 2012). In arthropods, *Wolbachia* mainly reside in the cells of the reproductive tissues, transmit transovarially (Frydman et al., 2006), and often commandeer host fertility, sex ratios, and sex determination to enhance their maternal transmission via male killing, feminization, parthenogenesis, or cytoplasmic incompatibility (CI) (LePage and Bordenstein, 2013; Werren et al., 2008).

Discovered nearly half a century ago (Yen and Barr, 1973), *Wolbachia*-induced CI is the most common reproductive modification and results in embryonic lethality when an infected male mates with an uninfected female, but this lethality is rescued when the female is likewise infected (Serbus et al., 2008). As such, rescue can provide a strong fitness advantage to infected females, the transmitting sex of *Wolbachia* (Hancock et al., 2011; Turelli and Hoffmann, 1999). Alone, CI-induced lethality is deployed in vector control studies to crash the resident uninfected mosquito population through release of *Wolbachia*-infected males (Atyame et al., 2011b; Dobson et al., 2016; O'Connor et al., 2012; Ritchie et al., 2015; Zabalou et al., 2004; D. Zhang et al., 2015). Together, CI-induced lethality and rescue constitute a microbial drive system that is used in field studies worldwide to stably replace an uninfected mosquito population with an infected one via release of males and females harboring wMel *Wolbachia* (Hoffmann et al., 2014), which confer resistance against dengue and Zika viruses (Dutra et al., 2016; Walker et al., 2011). The efficacy of this drive system for spreading *Wolbachia* in target populations critically depends on *Wolbachia*'s ability to rescue its own lethal sperm modifications.

While CI is gaining momentum as a natural, sustainable, and inexpensive tool for vector control, the genes that underpin this microbial adaptation are not fully known. Our previous screen of *Wolbachia* genomes and transcriptomes from infected ovaries identified two adjacent genes, *cifA* and *cifB*, from the *w*Mel strain in *Drosophila melanogaster* as the only genes strictly

associated with CI (LePage et al., 2017). These two genes occur in the eukaryotic association module of prophage WO (Bordenstein and Bordenstein, 2016) and recapitulate CI when dually expressed in uninfected male flies (Beckmann et al., 2017; LePage et al., 2017). Each gene alone is incapable of inducing CI (LePage et al., 2017), and the rescue gene remains unknown. As *cifA* and *cifB* are the only two wMel genes strictly associated with CI, we previously hypothesized that the CI induction and rescue genes might be the same (LePage et al., 2017). Here we test the hypothesis that transgenic expression of *cifA* and/or *cifB* genes from wMel *Wolbachia* in ovaries can rescue CI and nullify the associated embryonic defects in *D. melanogaster*.

Results and discussion

Since *Wolbachia* cannot be genetically transformed, we tested the ability of *cifA* to transgenically rescue wild-type CI using a GAL4-UAS system for tissue-specific expression in uninfected *D. melanogaster* females. In transcriptomes of *w*Mel-infected *D. melanogaster*, *cifA* is a highly expressed prophage WO gene (Gutzwiller et al., 2015). As such, we conducted the transgenic experiments under the control of either *nos*-GAL4-*tubulin* in uninfected germline stem cells or the maternal triple driver, MTD-GAL4, to drive higher transgene expression throughout oogenesis. MTD-GAL4 utilizes two *nos*-GAL4 driver variants (including *nos*-GAL4-*tubulin*) and an ovarian tumor driver (Petrella et al., 2007). Control CI and rescue crosses with either driver yielded the expected hatching rates. Crosses between infected males and uninfected females expressing *cifA* under the control of MTD-GAL4 showed a markedly significant increase in embryonic hatching relative to *cifA* expression under *nos*-GAL4-*tubulin* and at levels similar to that in control rescue crosses (Figure III-1a). These results are consistent with complete rescue of CI by *cifA*, in association with increased expression throughout the developing egg chambers. Similar results with *nos*-GAL4-*tubulin* expression in uninfected ovaries resulted in a small increase in hatch rate that was inconsistently significant among replicates (Figure B-1). An analysis of *cifA* gene expression reveals MTD-GAL4 associates with a three-order-of-magnitude increase over *nos*- GAL4-*tubulin*, supporting strength of expression as a factor for rescue (Figure III-1b).

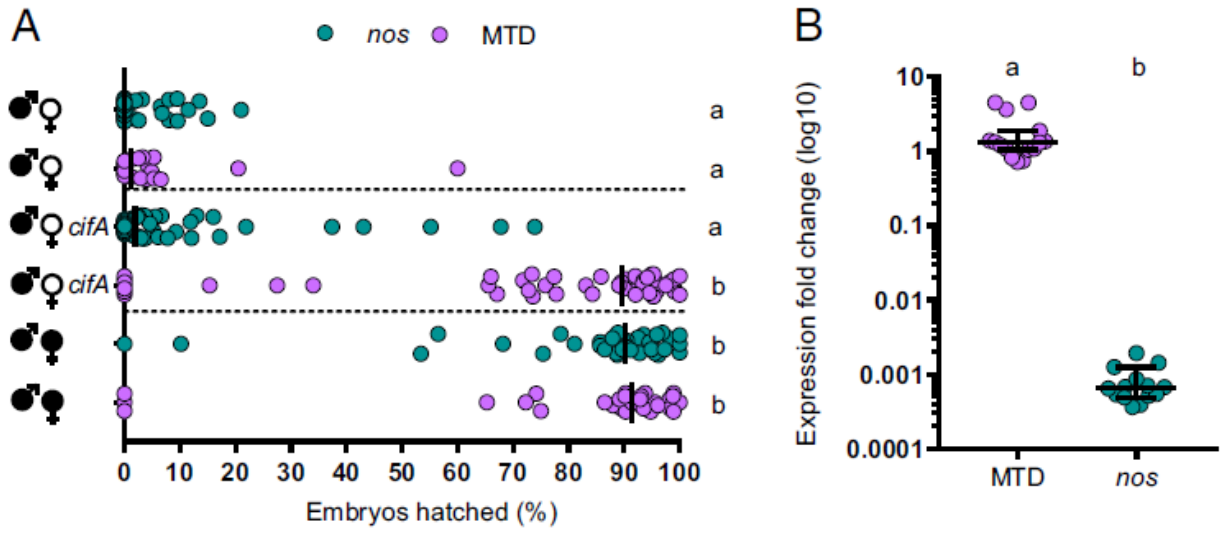

Figure III-1. *cifA* rescues cytoplasmic incompatibility when it is highly expressed throughout oogenesis.
(A) Hatch rate assays were conducted with transgenic expression of *cifA* under the control of *nos*-GAL4-tubulin or MTD-GAL4 drivers. Each dot represents a replicate. Rescue occurred only under MTD-GAL4 expression. Horizontal dotted lines from top to bottom separate cross-types with CI, *cifA* expression, and rescue. *Wolbachia* infections are represented by filled sex symbols, and expressed genes are noted to the right of the corresponding sex. n = 27–59 for each experimental cross across two experiments (both shown). Vertical bars represent medians, and letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by Kruskal–Wallis and Dunn's test for multiple comparisons. (B) Gene expression fold change of *cifA* relative to the *Drosophila* housekeeping gene *rp49* was determined on a subset of abdomens from females expressing *cifA* via MTD-GAL4 or *nos*-GAL4-*tubulin* with $2^{-\Delta\Delta Ct}$. Horizontal bars represent medians with 95% confidence intervals, and letters above indicate significance based on a Mann–Whitney U test. In both cases, statistical comparisons are between all groups. Exact P values are provided in Table B-2. Hatch rate experiments testing expression of *cifA* under MTD-GAL4 or *nos*-GAL4-*tubulin* have been repeated four and five times, respectively.

We expanded our evaluation of cif gene expression under the control of MTD-GAL4 in uninfected females to test if *cifB* alone or in combination with *cifA* impacts CI penetrance. As expected, infected males crossed to either uninfected females or females transgenically expressing *cifB* under MTD-GAL4 yielded similar CI penetrance (Figure III-2). These results suggest that *cifB* does not rescue CI when transgenically expressed in the ovaries, and its CI-related function is specific to testes. In contrast, MTD-GAL4 expression of *cifA*, by itself or in combination with *cifB*, significantly rescued CI to levels comparable to rescue by infected females (Figure III-2). These results are consistent with *cifA* independently functioning as the rescue factor and suggest that *cifB* does not inhibit *cifA*'s ability to rescue CI. As *Wolbachia* can induce phenotypes known to bias sex ratios, we collected the surviving offspring from the transgenic and control rescue crosses and sexed them to demonstrate normal sex ratios, indicating that rescue was not sex-specific (Figure B-2).

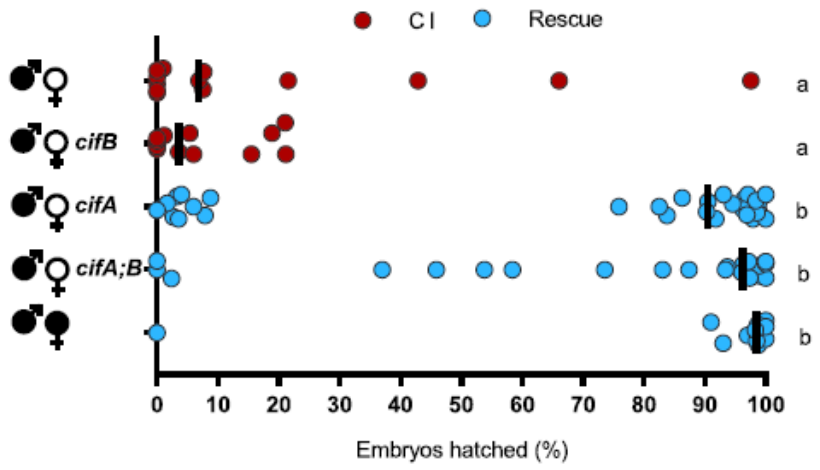

Figure III-2. Rescue of cytoplasmic incompatibility is specific to *cifA*.
Hatch rate assays were conducted with transgenic expression of *cifA*, *cifB*, and *cifA;B* using the MTD-GAL4 driver for expression throughout oogenesis. Each dot represents a replicate. *Wolbachia* infections are represented by filled sex symbols, and expressed genes are noted to the right of the corresponding sex. n = 11–29 for each experimental cross. Vertical bars represent medians, and letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by Kruskal–Wallis and Dunn's test for multiple comparisons. Statistical comparisons are between all groups. Exact P values are provided in Table B-2. Hatch rate experiments testing expression of *cifA* under MTD-GAL4 have been repeated four times.

Next, we tested if the canonical cytological defects observed in early CI embryos [early mitotic failure, chromatin bridging, and regional mitotic failure (Landmann et al., 2009)] were nullified under *cifA*-induced rescue. We examined embryos from control and transgenic crosses after 1–2 h of development and binned their cytology into one of five phenotypes as previously established for *D. melanogaster* CI (LePage et al., 2017). Nearly half of CI-induced lethality in embryos is the result of embryonic arrest during advanced developmental stages in Dipteran species (Callaini et al., 1996; Duron and Weill, 2006; Lassy and Karr, 1996; Wright and Barr, 1981). As expected, the control CI cross yielded high levels of all three CI-associated defects, and the embryos from the control rescue cross developed with significantly fewer abnormalities (Figure III-3). MTD-GAL4 transgenic expression of *cifA* in uninfected females, either alone or dually expressed with *cifB*, resulted in significantly fewer cytological defects (Figure III-3). These effects were not seen with transgenic *cifB* expression, again validating that *cifA* alone can recapitulate wildtype rescue by *Wolbachia*.

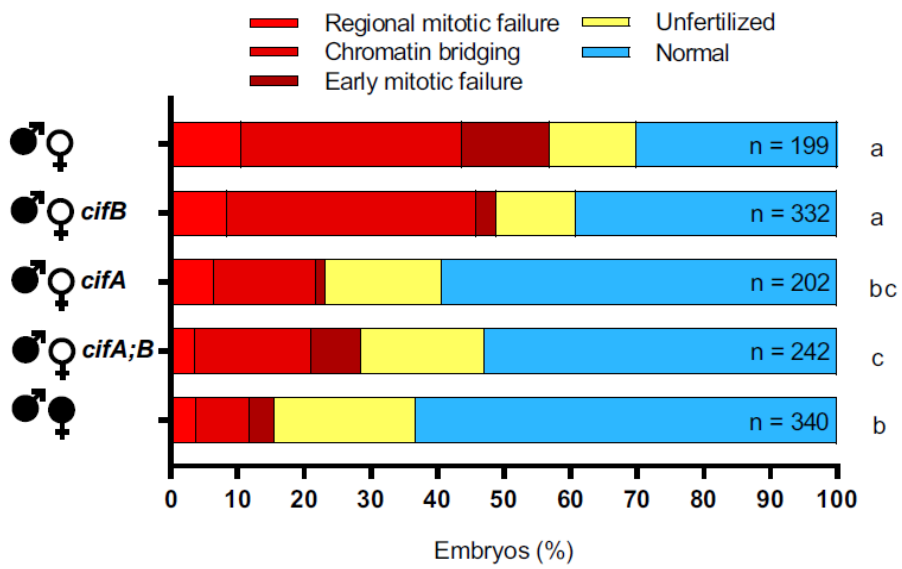

Figure III-3. *cifA* rescues embryonic defects caused by cytoplasmic incompatibility.
The percent of embryos with each cytological phenotype resulting from the indicated crosses are shown. All crosses were conducted in parallel and with sisters from the experiment in Figure III-2. *cifA*, *cifB*, and *cifA;B* transgene expression was under the control of MTD-GAL4. *Wolbachia* infections are represented by filled sex symbols and expressed genes are noted to the right of the corresponding sex. Letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by pairwise $\chi 2$ analyses comparing defects (all shades of red) against normal (blue) with Bonferroni adjusted P values. Exact P values are provided in Table B-2. This experiment has been conducted once.

These data are in contrast with previous work reporting the inability to transgenically rescue CI in *D. melanogaster* (Beckmann et al., 2017); however, there are three critical differences between the studies. First, *w*Pip's homologs from *Culex pipiens* were used in the prior work instead of *w*Mel's cif genes from *D. melanogaster* here. Thus, differences in host background interactions could explain the discrepancy. Second, a T2A sequence between the *w*Pip gene homologs was used to allow for bicistronic expression, but ribosome skipping results in a C-terminal sequence extension to the first protein and a proline addition to the second protein that generates sequence artifacts and could alter function (Donnelly et al., 2001b). Finally, different insertion sites are capable of different levels of expression due to their local chromatin environment (Akhtar et al., 2013), thus the chosen sites may produce insufficient product to cause rescue.

cifA encodes a putative catalase-rel domain, a sterile-like transcription factor (STE) domain, and a domain of unknown function (DUF3243) that shares homology with a putative Puf-family RNA binding domain in *cifA*-like homologs (Lindsey et al., 2018b), whereas *cifB* has nuclease and deubiquitilase domains (Beckmann et al., 2017; Lindsey et al., 2018b). Only the deubiquitilase annotation has been functionally tested and confirmed (Beckmann et al., 2017). Based on subcellular localization (PSORTb) and transmembrane helix predictors (TMbase), CifA

is a cytoplasmic protein without transmembrane helices (Figure B-3). Codon-based and Fisher's exact tests of neutrality demonstrate that closely related (76.2–99.8% pairwise nucleotide identity) type I *cifA* homologs (LePage et al., 2017) largely evolve by purifying selection (Figure B-4a,b), and sliding window analyses [sliding window analysis of Ka and Ks (SWAKK) and Java codon delimited alignment, JCoDA] reveal that purifying selection is strongest on the catalase-rel domain and the unannotated region at the N terminus, with considerably weaker purifying selection on the putative DUF3243 and STE domains (Figure III-4; Figure B-4c). This is supported by prior work reporting stronger amino acid conservation within the type I CifA N terminus relative to the C terminus (Lindsey et al., 2018b).

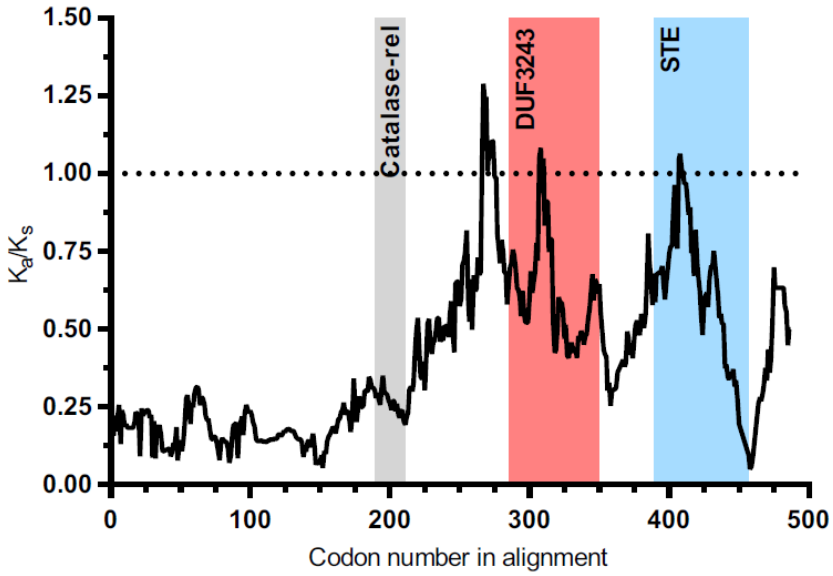

Figure III-4. Ka/Ks sliding window analysis identifies *cifA* regions evolving under negative selection.
A comparison between *cifA* homologs from *w*Mel and *w*Ha rejects the neutral expectation of Ka/Ks = 1 using a 25-amino-acid sliding window across most of *cifA*. Strong purifying selection is observed in several *cifA* regions including the sequence preceding the catalase-rel domain. Shaded regions denote previously described protein domain predictions and white regions are unannotated (Lindsey et al., 2018b).

These findings illustrate that the *Wolbachia* prophage WO gene *cifA* recapitulates rescue of wild-type CI. As *cifA* is one of two genes involved in induction of CI, results support the hypothesis that a gene involved in CI induction is also the rescue gene (LePage et al., 2017). In addition, transgenic expression of *cifA* in yeast inhibits a temperature-dependent growth defect caused by *cifB* expression (Beckmann et al., 2017). The discovery that CI is induced by *cifA* and *cifB* and rescued by *cifA* motivates a Two-by-One model of CI where two genes act as the CI modification factors (in the male), and one of these same genes acts as the rescue factor (in the female). This modification-rescue model posits that each strain of *Wolbachia* has its own set of

cifA- and *cifB*-associated CI modifications and one *cifA* rescue factor. The different roles of *cifA* in CI and rescue are intriguing. We predict that the function of *cifA* is dependent on differential localization and/or modification of gene products in testes/sperm (CI) relative to ovaries/embryos (rescue). Moreover, one could speculate that the putative antioxidant catalase-rel domain of the CifA protein acts as a functional switch in response to reactive oxygen species, known to be higher in *Wolbachia*-infected testes (Brennan et al., 2012), whereas the Puf-family RNA binding domain and STE are involved in RNA binding and transcriptional (mis)regulation of an unknown host factor.

It has been hypothesized that divergence in modification and rescue genes leads to bidirectional CI (Bonneau et al., 2018a; Charlat et al., 2001; LePage et al., 2017), which is a reciprocal incompatibility between males and females infected with different *Wolbachia* strains (Bordenstein et al., 2001; Bordenstein and Werren, 2007; O'Neill and Karr, 1990; Poinsot et al., 1998; Yen and Barr, 1973). Comparative genomic analyses of *cifA* and *cifB* genes reveal extremely high levels of amino acid divergence (LePage et al., 2017), strong codivergence (LePage et al., 2017; Lindsey et al., 2018b), and recombination (Bonneau et al., 2018a), consistent with the very rapid evolution of bidirectional CI across *Wolbachia* that can contribute to reproductive isolation and speciation (Bordenstein et al., 2001; Brucker and Bordenstein, 2012a). Indeed, divergence of the *cifA* and *cifB* genes into several phylogenetic types correlates with bidirectional CI patterns in *Drosophila* and *Culex* (Bonneau et al., 2018a; LePage et al., 2017). There are at least two explanations for how simple genetic changes in these genes can contribute to bidirectional CI. First, a single mutation in the *cifA* gene could produce variation in the modification and rescue components that render two *Wolbachia* strains incompatible. For instance, given an ancestral and derived allele of *cifA*, males and females with *Wolbachia* carrying the same *cifA* allele are compatible; however, males with *Wolbachia* carrying the ancestral *cifA* allele cause a sperm modification that is unable to be rescued by embryos with *Wolbachia* carrying the derived *cifA* allele, and vice versa. Thus, a single mutation in *cifA* alone can enable the switch from being compatible to incompatible *Wolbachia*. Second, mutations in both *cifA* and *cifB* could be required for the evolution of bidirectional CI. For example, CifA-CifB protein binding (Beckmann et al., 2017) and/or differential localization in the sperm and egg may underpin bidirectional CI between *Wolbachia* strains. In this model, amino acid divergence in the Cif proteins may contribute to weakened binding, which in turn yields *Wolbachia* strains incapable of CI but capable of rescuing

CI by the ancestral variant (Bourtzis et al., 1998; Zabalou et al., 2004). A compensatory substitution in the other Cif protein could in theory restore binding and yield bidirectional incompatibility with the ancestral Cif variants. Codivergence between amino acid sequences of these proteins is consistent with this model. Under both models, the presence of multiple WO prophages carrying *cifA* genes may also promote incompatibilities through the production of multiple CI product complexes simultaneously (LePage et al., 2017). In support of these hypotheses, complex diversification and duplication of *cifA* and *cifB* have been reported in *Drosophila* and *C. pipiens* that harbor a variety of incompatible *Wolbachia* strains (Bonneau et al., 2018a; LePage et al., 2017).

In conclusion, our findings reveal the connected genetic basis of CI and rescue and highlight the fundamental impact of prophage WO genes on the adaptive phenotypes of an obligate intracellular bacteria. In addition to genetically dissecting this widespread form of reproductive parasitism and microbial drive, we also establish a Two-by-One model to explain the modification and rescue components of CI. Finally, beneficial applications of CI and rescue genes as transgenic drive constructs may be possible as adjuncts or alternatives to pest control or vector control strategies currently deploying *Wolbachia*-infected mosquitoes (Hoffmann et al., 2014; Ritchie et al., 2015; Zabalou et al., 2004; D. Zhang et al., 2015).

Materials and methods

Fly rearing and strains.

D. melanogaster stocks y^1w^* (BDSC 1495), *nos*-GAL4- *tubulin* (BDSC 4442), MTD-GAL4 (containing *nos*-GAL4-*tubulin*, *nos*-GAL4-VP16, and otu-GAL4-VP16; BDSC 31777), and UAS transgenic lines homozygous for *cifA*, *cifB*, and *cifA*;B (LePage et al., 2017) were maintained at 12:12 light:dark at 25 °C and 70% relative humidity (RH) on 50 mL of a standard media. GAL4 lines were found to be infected with *w*Mel *Wolbachia*, and uninfected lines were produced through tetracycline treatment as previously described (LePage et al., 2017). Infection status was frequently confirmed via PCR using WolbF and WolbR3 primers (Casiraghi et al., 2005) (Table B-1). During virgin collections, flies were stored at 18 °C overnight to slow eclosion rate, and virgin flies were kept at room temperature.

Hatch rate and sex ratio assays.

Virgin MTD-GAL4 females were collected for the first 3 d of emergence and aged 9–11 d before crossing to nonvirgin homozygous UAS (*cifA*, *cifB*, or *cifA*;*B*) males. The start of collections for the maternal and paternal lineages was staggered by 7 d. Single pair matings occurred in 8-oz bottles, and a grape-juice agar plate was smeared with yeast and affixed to the opening of each bottle with tape. The flies and bottles were then stored at 25 °C and 70% RH for 24 h, at which time the plates were replaced with freshly smeared plates and again stored for 24 h. Plates were then removed and the number of embryos on each plate were counted and stored. After 30 h the remaining unhatched embryos were counted (Figure B-6). The hatch rate was calculated by dividing the number of hatched embryos by the initial embryo count and multiplying by 100. Hatch rate was plotted against clutch size for 3 MTD-GAL4 and 4 *nos*-GAL4-*tubulin* rescue crosses conducted in this study to reveal a significant correlation (Figure B-5), and a threshold clutch size for analysis was set equal to exclusion of 99% of plates with a hatch rate of 0 for each genotype (31 for *nos*-GAL4-*tubulin* and 48 for MTD-GAL4). Larvae were moved into vials of standard media and the offspring sex ratio determined after 15–18 d (Figure B-6). Hatch rates testing MTD-GAL4 or *nos*-GAL4-*tubulin* expression of *cifA* were conducted four and five times, respectively. Sex ratio experiments were conducted once.

Gene expression.

To compare the level of UAS-*cifA* expression between MTD-GAL4 and *nos*-GAL4-*tubulin* flies, mothers from hatch rate assays were collected after the allotted laying period, abdomens were immediately dissected, and samples were frozen in liquid nitrogen and stored at −80 °C until processing. RNA was extracted using the Direct-zol RNA MiniPrep Kit (Zymo), DNase treated with DNA-free (Ambion, Life Technologies), and cDNA was generated with SuperScript VILO (Invitrogen). Quantitative PCR was performed on a Bio-Rad CFX-96 Real-Time System using iTaq Universal SYBR Green Supermix (Bio-Rad). Forty cycles of PCR were performed against positive controls (extracted DNA), negative controls (water), no RT control (RNA), and cDNA with the following conditions: 50 °C 10 min, 95 °C 5 min, 40× (95 °C 10 s, 55 °C 30 s), 95 °C 30 s. Primers used were *cifA* opt and *rp49* forward and reverse (Table B-1). Fold expression of UAS-*cifA* relative to the *D. melanogaster* housekeeping gene *rp49* was determined with $2^{-\Delta\Delta Ct}$. This experiment and corresponding hatch rate were performed once.

Embryo cytology.

Flies were collected as described for the hatch rate assays, but with 60 females and 12males in each bottle with a grape-juice agar plate attached. All flies used were siblings of those from the hatch rate, grapejuice plates replaced as described above, and embryos collected in parallel to egg laying by hatch rate females. Embryos were collected, dechorionated, washed, methanol fixed, stained with propidium iodide, imaged, and categorized as previously described (LePage et al., 2017) (Figure B-6). This experiment was performed once. Putative CifA Localization. The PSORTb v3.0.2 web server (Yu et al., 2010) was used to predict subcellular localization of the *w*Mel CifA protein to either the cytoplasm, cytoplasmic membrane, periplasm, outer membrane, or extracellular space. A localization score is provided for each location, with scores of 7.5 or greater considered probable localizations. The TMpred web server (46) was used to predict transmembrane helices in *w*Mel CifA. TMpred scores were generated for transmembrane helices spanning from inside-to-outside (i-o) and outside-to-inside (o-i), and scores above 500 are considered significant.

cifA *selection analyses.*

Selection analyses were conducted using four independent tests of selection: codon-based Z test of neutrality, Fisher's exact test of neutrality, Sliding window analysis of Ka and Ks (SWAKK), and Java Codon Delimited Alignment (JCoDA) (Kumar et al., 2016; Liang et al., 2006; Steinway et al., 2010). The first two analyses were conducted using the MEGA7 desktop app with a MUSCLE translation alignment generated in Geneious v5.5.9. The SWAKK 2.1 web server and the JCoDA v1.4 desktop app were used to analyze divergence between wMel and wHa *cifA* with a sliding window of 25 codons and a jump size of 1 codon for SWAKK and 5 codons for JCoDA. Statistical Analyses. All statistical analyses were conducted in GraphPad Prism (Prism 7 or online tools). Hatch rate and sex ratio statistical comparisons were made using Kruskal–Wallis followed by a Dunn's multiple comparison test. Expression was compared using a Mann–Whitney U test. Correlations between hatch rate and clutch size were determined using Spearman rho. Pairwise $\chi 2$ analyses were used for cytology studies to compare defective and normal embryos followed by generation of Bonferroni adjusted P values. An unpaired t test was used for statistical comparison of RNA fold expression. All P values are reported in Table B-2.

Chapter IV.

Two-by-One model of cytoplasmic incompatibility: Synthetic recapitulation by transgenic expression of *cifA* and *cifB* in *Drosophila*[‡]

Abstract

Wolbachia are maternally inherited bacteria that infect arthropod species worldwide and are deployed in vector control to curb arboviral spread using cytoplasmic incompatibility (CI). CI kills embryos when an infected male mates with an uninfected female, but the lethality is rescued if the female and her embryos are likewise infected. Two phage WO genes, $cifA_{wMel}$ and $cifB_{wMel}$ from the *w*Mel *Wolbachia* deployed in vector control, transgenically recapitulate variably penetrant CI, and one of the same genes, $cifA_{wMel}$, rescues wild type CI. The proposed Two-by-One genetic model predicts that CI and rescue can be recapitulated by transgenic expression alone and that dual $cifA_{wMel}$ *and* $cifB_{wMel}$ expression can recapitulate strong CI. Here, we use hatch rate and gene expression analyses in transgenic *Drosophila melanogaster* to demonstrate that CI and rescue can be synthetically recapitulated in full, and strong, transgenic CI comparable to wild type CI is achievable. These data explicitly validate the Two-by-One model in *w*Mel-infected *D. melanogaster*, establish a robust system for transgenic studies of CI in a model system, and represent the first case of completely engineering male and female animal reproduction to depend upon bacteriophage gene products.

Author summary

Releases of *Wolbachia*-infected mosquitos are underway worldwide because *Wolbachia* block replication of Zika and Dengue viruses and spread themselves maternally through arthropod populations via cytoplasmic incompatibility (CI). The CI drive system depends on a *Wolbachia*-induced sperm modification that results in embryonic lethality when an infected male mates with an uninfected female, but this lethality is rescued when the female and her embryos are likewise

[‡] This chapter is published in 2019 in *PLOS Genetics.* 16(6), e1008221 with myself as first author. Seth Bordenstein was senior author.

infected. We recently reported that the phage WO genes, *cifA* and *cifB*, cause the sperm modification and *cifA* rescues the embryonic lethality caused by the *w*Mel *Wolbachia* strain deployed in vector control. These reports motivated proposal of the Two-by-One model of CI whereby two genes cause lethality and one gene rescues it. Here we provide unequivocal support for the model in the *Wolbachia* strain used in vector control via synthetic methods that recapitulate CI and rescue in the absence of a *Wolbachia* infections. Our results reveal the set of phage WOgenes responsible for this powerful genetic drive system, act as a proof-of-concept that these genes alone can induce gene drive like crossing patterns, and establish methodologies and hypotheses for future studies of CI in *Drosophila*. We discuss the implications of the Two-by-One model towards functional mechanisms of CI, the emergence of incompatibility between *Wolbachia* strains, vector control applications, and CI gene nomenclature.

Introduction

Wolbachia are the most widespread endosymbiotic bacteria on the planet and are estimated to infect half of all arthropod species (Weinert et al., 2015; Zug and Hammerstein, 2012) and half of the Onchocercidae family of filarial nematodes (Ferri et al., 2011). They specialize in infecting the cells of reproductive tissues, are primarily inherited maternally from ova to offspring, and often act in arthropods as reproductive parasites that enhance their maternal transmission by distorting host sex ratios and reproduction (LePage and Bordenstein, 2013; Taylor et al., 2018). The most common type of reproductive parasitism is cytoplasmic incompatibility (CI), which manifests as a sperm modification in infected males that causes embryonic lethality or haploidization in matings with uninfected females upon fertilization (Bordenstein et al., 2003; Serbus et al., 2008; Yen and Barr, 1973). This embryonic lethality is rescued if the female is infected with the same *Wolbachia* strain. As such, CI selfishly drives CI-inducing *Wolbachia* into host populations (Hancock et al., 2011; Hoffmann et al., 1990; Leftwich et al., 2018; Turelli, 1994; Turelli et al., 2018b), and the incompatibilities between host populations cause reproductive isolation between recently diverged or incipient species (Bordenstein et al., 2001; Brucker and Bordenstein, 2012a; Jaenike et al., 2006; Miller et al., 2010; Shropshire and Bordenstein, 2016).

In the last decade, *Wolbachia* and CI have garnered significant interest for their utility in combatting vector borne diseases worldwide. Two strategies are currently deployed: population

suppression and population replacement. The population suppression strategy markedly crashes vector population sizes through the release of only infected males that induce CI upon mating with wild uninfected females (Dobson et al., 2002; Lees et al., 2015; Nikolouli et al., 2018; O'Connor et al., 2012). In contrast, the population replacement strategy converts uninfected to infected populations through the release of both infected males and females that aid the spread *Wolbachia* via CI and rescue (Huang et al., 2018; O'Neill, 2018). Replacing a vector competent, uninfected population with infected individuals can notably reduce the spread of arthropod borne diseases such as Zika and dengue (Caragata et al., 2016; Hoffmann et al., 2011) because *Wolbachia* appear to inhibit various stages of viral replication within arthropods based on diverse manipulations of the host cellular environment (Bhattacharya et al., 2017; Brennan et al., 2012; Caragata et al., 2013; Geoghegan et al., 2017; Lindsey et al., 2018a; Molloy et al., 2016; Schultz et al., 2017). The combination of *Wolbachia*'s abilities to suppress arthropod populations, drive into host populations, and block the spread of viral pathogens have established *Wolbachia* in the vanguard of vector control efforts to curb arboviral transmission (Caragata et al., 2016; Huang et al., 2018; Hughes et al., 2011; O'Connor et al., 2012; O'Neill, 2018; Schmidt et al., 2017; Turelli and Barton, 2017).

An unbiased, multi-omic analysis of CI-inducing and CI-incapable *Wolbachia* strains revealed two adjacent genes, *cifA* and *cifB*, in the eukaryotic association module of prophage WO (Bordenstein and Bordenstein, 2016) that strictly associate with CI induction (LePage et al., 2017). Fragments of the CifA protein were found in the fertilized spermathecae of *w*Pip infected *Culex pipiens* mosquitoes (Beckmann and Fallon, 2013), and these genes are frequently missing or degraded in diverse CI-incapable strains (Lindsey et al., 2018b; Sutton et al., 2014). Dual transgenic expression of *cifA* and *cifB* from either of the CI-inducing strains *w*Mel or *w*Pip in uninfected male flies causes a decrease in embryonic hatching corresponding to an increase in CI associated cytological abnormalities including chromatin bridging and regional mitotic failures (Beckmann et al., 2017; LePage et al., 2017). Single transgenic expression of either $cifA_{wMel}$ or $cifB_{wMel}$ in an uninfected male was insufficient to recapitulate CI, but single transgenic expression of either gene in an infected male enhances *w*Mel-induced CI in a dose-dependent manner (LePage et al., 2017). Importantly, dual transgenic CI induced by $cifA_{wMel}$ and $cifB_{wMel}$ expressing males was rescued when they were mated with *w*Mel-infected females (LePage et al., 2017). Moreover, transgenic expression of $cifA_{wMel}$ alone in uninfected females rescues embryonic lethality and

nullifies cytological defects associated with wild type CI caused by a *w*Mel infection (Shropshire et al., 2018).

As such, we recently proposed the Two-by-One genetic model of CI wherein dual expression of $cifA_{wMel}$ and $cifB_{wMel}$ causes CI when expressed in males and expression of $cifA_{wMel}$ rescues CI when expressed in females (Shropshire et al., 2018). However, confirmation of the model's central prediction requires the complete synthetic replication of CI-induced lethality and rescue in the absence of any *Wolbachia* infections since it remains possible that other *Wolbachia* or phage WO genes besides *cifA* and *cifB* contribute to wild type CI and rescue by *w*Mel *Wolbachia*. Moreover, CI induced by dual $cifA_{wMel}$ and $cifB_{wMel}$ expression previously yielded variable offspring lethality with a median survival of 26.5% of embryos relative to survival of 0.0% of embryos from CI induced by a wild type infection under controlled conditions (LePage et al., 2017). The inability to recapitulate strong wild type CI suggests other CI genes are required, other environmental factors need to be controlled, or the transgenic system requires optimization.

Here, we utilize transgenic expression, hatch rates, and gene expression assays in *Drosophila melanogaster* to test if an optimized expression system can generate strong transgenic CI and whether bacteriophage genes $cifA_{wMel}$ and $cifB_{wMel}$ can fully control fly reproduction by inducing and rescuing CI in the complete absence of *Wolbachia* (Figure IV-1). We further assess if both cif_{wMel} genes are required for CI induction in the optimized system and whether $cifA_{wMel}$ in females can rescue transgenic CI. Results provide strong evidence for the Two-by-One model in *w*Mel-infected *D. melanogaster*, offer context for conceptualizing CI mechanisms and the volution of bidirectional incompatibilities between different *Wolbachia* strains, raise points for CI gene nomenclature, and motivate further research in developing these genes into a tool that combats vector borne diseases. To the best of our knowledge, they also represent the first case of completely engineering animal sexual reproduction to depend upon bacteriophage gene products.

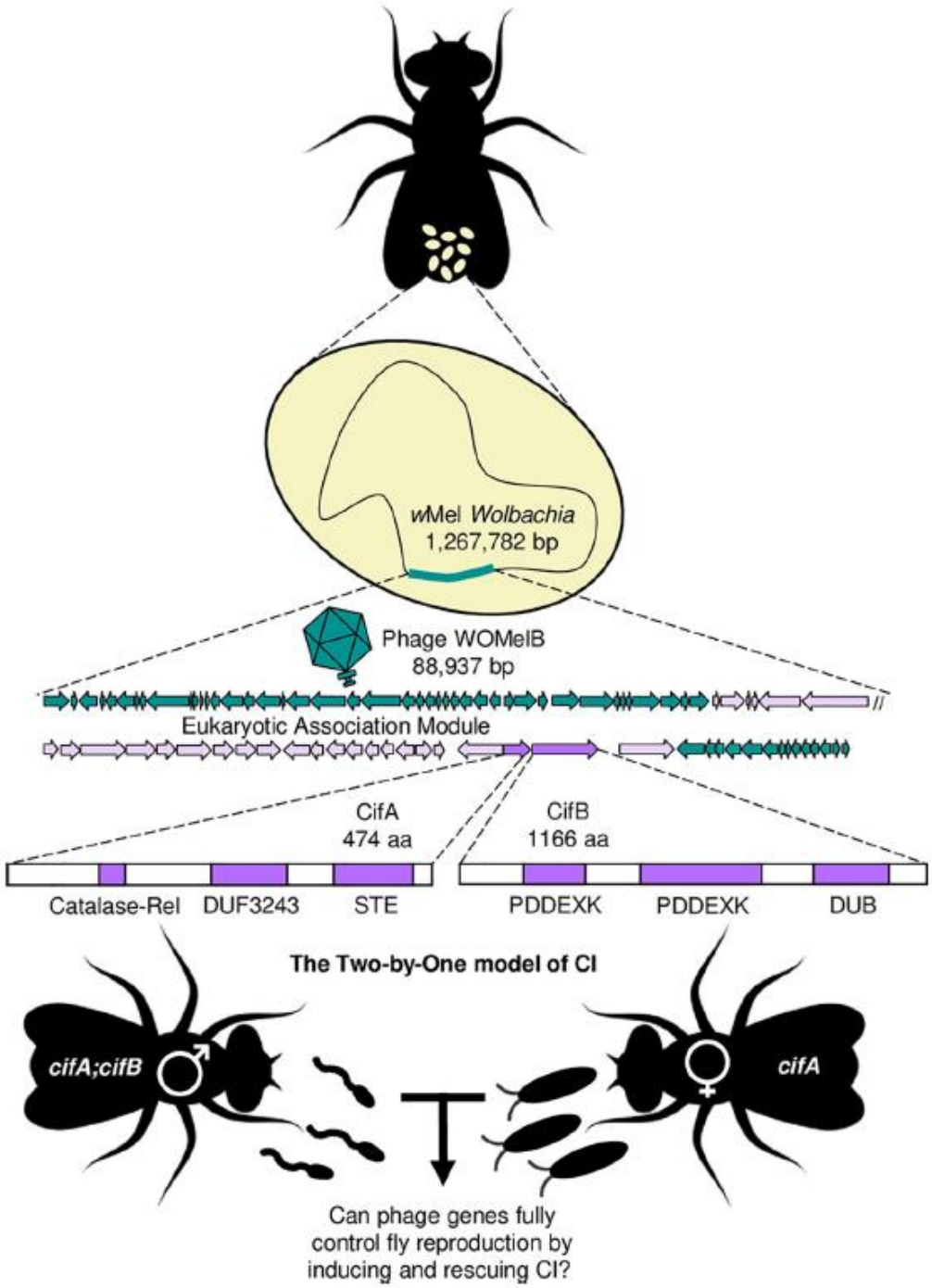

Figure IV-1. Two-by-One model of CI is governed by *cifA* and *cifB* genes in the eukaryotic association module of prophage WO in *Wolbachia*.

The Two-by-One model of CI predicts that *D. melanogaster* males and females can be engineered to recapitulate both CI and rescue phenotypes in the absence of *Wolbachia*, thus depending completely on phage genes for successful reproduction. Schematics are not to scale. Insect, sperm, and embryo art were obtained and modified using vecteezy.com. Phage gene schematics modified from (Beckmann and Fallon, 2013). CifA and CifB protein annotation from (Lindsey et al., 2018b). Purple indicates eukaryotic association module genes as indicated by (Bordenstein and Bordenstein, 2016).

Results

Optimizing transgenic CI.

Dual transgenic expression of *cifA*$_{wMel}$ and *cifB*$_{wMel}$ was previously reported to induce highly variable and incomplete CI relative to CI caused by an age-controlled *w*Mel infection [38], indicating either the presence of other genes necessary for strong CI, environmental factors uncontrolled in the study, or inefficiency of the transgenic system. Here, we test the latter hypothesis by dually expressing *cifA*$_{wMel}$ and *cifB*$_{wMel}$ in uninfected *D. melanogaster* males under two distinct GAL4 driver lines that express in reproductive tissues: *nos*-GAL4-*tubulin* and *nos*-GAL4:VP16 (Ni et al., 2011). Both driver lines contain a *nos* promoter region, but differ in that *nos*-GAL4-*tubulin* produces a transcription factor with both the DNA binding and transcriptional activating region of the GAL4 protein, and *nos*-GAL4:VP16 produces a fusion protein of the GAL4 DNA binding domain and the virion protein 16 (VP16) activating region (Doren et al., 1998; Tracey et al., 2000). The GAL4:VP16 transcription factor is a particularly potent transcriptional activator because of its binding efficiency to transcription factors (He et al., 1993; Sadowski et al., 1988). Additionally, the *nos*-GAL4-*tubulin* driver has a *tubulin* 3' UTR, and *nos*-GAL4:VP16 has a *nos* 3' UTR that may contribute to differences in localization within cells or between tissues (Doren et al., 1998; Ni et al., 2011; Tracey et al., 2000). As such, we predict that differences in the expression level or profile of these two driver lines will lead to differences in the penetrance of transgenic CI.

Since CI manifests as embryonic lethality, we measure hatching of *D. melanogaster* embryos into larvae to quantify the strength of CI. We confirm previous findings (LePage et al., 2017) that dual transgenic expression of *cifA*$_{wMel}$ and *cifB*$_{wMel}$ under *nos*-GAL4-*tubulin* in uninfected males yields low but variable embryonic hatching in crosses with uninfected females (Mdn = 26.3%, IQR = 10.4–38.1%) that can be rescued in crosses with *w*Mel-infected females (Mdn = 97.5%; IQR = 94.2–100%) (Figure IV-2a). However, dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression under *nos*-GAL4:VP16 in uninfected males yields significantly reduced embryonic hatching relative to *nos*-GAL4-*tubulin* ($p = 0.0002$) with less variability (Mdn = 0%; IQR = 0.0–0.75%) and can be comparably rescued (Mdn = 98.65%; IQR = 95.93–100%; $p > 0.99$) (Figure IV-2a). Together, these results support that dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression under *nos*-GAL4:VP16 induces the strongest CI and that the transgenic system, not the absence of necessary CI factors, contributed to the prior inability to recapitulate strong wild type CI.

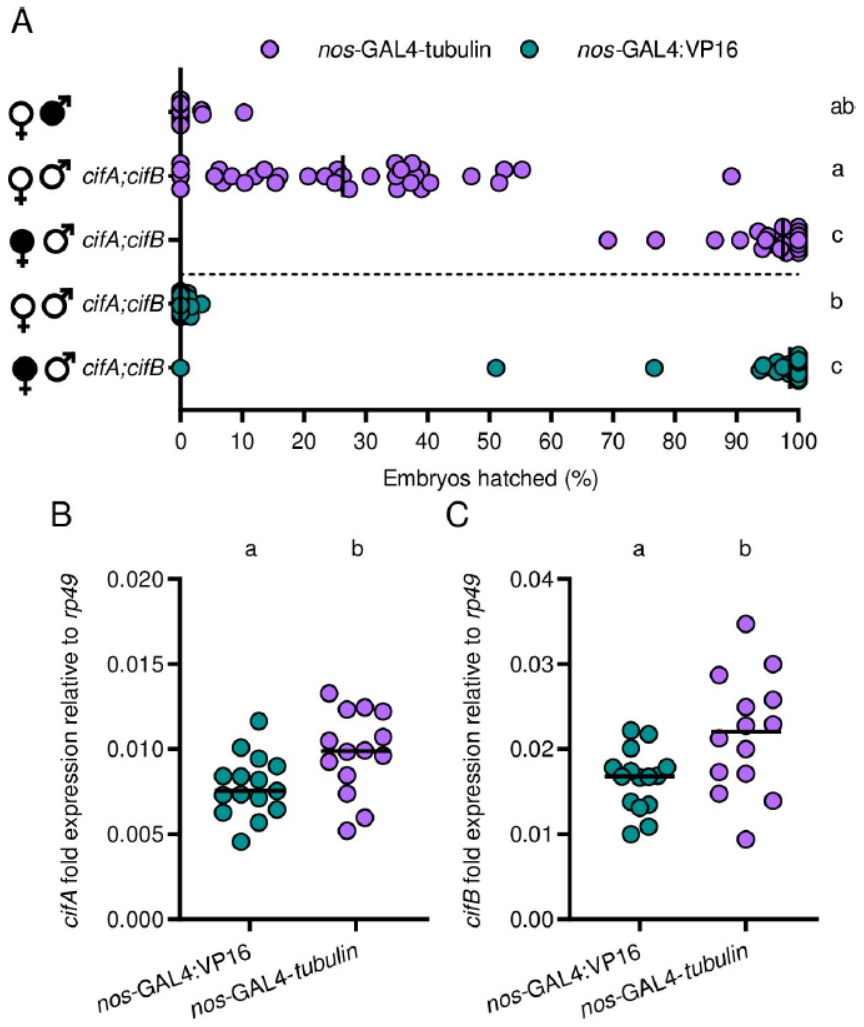

Figure IV-2. *cifA*$_{wMel}$ and *cifB*$_{wMel}$ induce strong CI when transgenically expressed in males under the *nos*-GAL4:VP16 driver.
(A) Two different driver lines, *nos*-GAL4-*tubulin* (purple; top) and *nos*-GAL4:VP16 (green; bottom) were tested for their ability to induce CI when transgenically expressed in uninfected, male *Drosophila*. Filled sex symbols represent infection with *w*Mel *Wolbachia*, and gene names to the right of a symbol represent expression of those genes in the male line. Vertical bars represent medians. Letters to the right indicate significant differences with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. (B,C) To test if *nos*-GAL4-*tubulin* and *nos*-GAL4:VP16 generate different levels of gene expression, (B) *cifA*$_{wMel}$ and (C) *cifB*$_{wMel}$ fold expression difference relative to the *Drosophila* housekeeping gene *rp49* in male abdomens under the two drivers was measured using qPCR. Males tested for gene expression were the same used in the hatch rate experiment in A. Letters above indicate significant differences with an $\alpha = 0.05$ calculated by a Mann-Whitney U test.

Next, we tested the hypothesis that differences in the penetrance of transgenic CI between the two drivers are due to differences in the strength of expression. To assess this, we used qPCR to measure the gene expression of *cifA*$_{wMel}$ and *cifB*$_{wMel}$ under the two drivers relative to a *Drosophila* housekeeping gene (*rp49*) in male abdomens (Figure IV-2b, c). Fold differences in RNA transcripts of *cifA*$_{wMel}$ relative to *rp49* reveal *nos*-GAL4-*tubulin* (Mdn = 0.0098; IQR =

0.0082–0.122) drives significantly stronger and more variable *cifA*$_{wMel}$ expression relative to *nos*-GAL4:VP16 (Mdn = 0.0075; IQR = 0.0064–0.0090) (p = 0.016, MWU, Figure IV-2b). The same is true for *cifB*$_{wMel}$ expression where *nos*-GAL4-*tubulin* (Mdn = 0.022; IQR = 0.0165– 0.0265) drives significantly stronger *cifB*$_{wMel}$ expression than *nos*-GAL4:VP16 (Mdn = 0.0168; IQR = 0.0135–0.0179) (p = 0.02, MWU, Figure IV-2c). Moreover, while *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression significantly correlate with each other under both *nos*-GAL4-*tubulin* (R2 = 0.85; p <0.0001) and *nos*-GAL4:VP16 (R2 = 0.75; p <0.0001; Figure C-1a), neither *cifA*$_{wMel}$ (R2 = 0.02; p = 0.62; Figure C-1b) nor *cifBwMel* (R2 = 0.04; p = 0.48; Figure C-1c) expression levels under the *nos*-GAL4-*tubulin* driver correlate with the strength of CI measured via hatch rates. Notably, *cifB*$_{wMel}$ is consistently more highly expressed than *cifA*$_{wMel}$ within the same line (Figure C-1a). We predict that expression differences are due to either differences in transgenic insertion sites or more rapid degradation of *cifA*$_{wMel}$ relative to *cifB*$_{wMel}$. Taken together, these results suggest that an increase in CI penetrance in these crosses is not positively associated with higher transgene transcript abundance from different drivers.

Optimizing transgenic rescue.

cifA$_{wMel}$ expression under the maternal triple driver (MTD) in uninfected females can rescue CI induced by a wild type infection (Shropshire et al., 2018). MTD is comprised of three drivers in the same line: *nos*-GAL4-*tubulin*, *nos*-GAL4:VP16, and *otu*-GAL4:VP16 (Ni et al., 2011). We previously reported that *cifA*$_{wMel}$ expression under the *nos*-GAL4-*tubulin* driver alone is rescue-incapable (Shropshire et al., 2018). Here, we test if *cifA*$_{wMel}$ expression under either of the other components of the MTD driver independently recapitulate rescue of *w*Mel CI. Hatch rate experiments indicate that CI is strong and expectedly not rescued when an infected male mates with a non-transgenic female whose genotype is otherwise *nos*-GAL4:VP16 (Mdn = 0.0%; IQR = 0.0–0.0%) or *otu*-GAL4:VP16 (Mdn = 0.0%; IQR = 0.0–0.0%) (Figure IV-3a). Transgenic expression of *cifA*$_{wMel}$ in uninfected females under either of the two drivers rescues CI induced by *w*Mel. However, rescue is significantly weaker under *cifA*$_{wMel}$ expression with the *otu*-GAL4:VP16 driver (Mdn = 70.4%; IQR = 0.0–90.45%) as compared to the *nos*-GAL4:VP16 driver (Mdn = 94.2%; IQR = 83.3–97.1%; p = 0.0491) which produced strong transgenic rescue (Figure IV-3a). Gene expression analysis of *cifA*$_{wMel}$ relative to *rp49* in the abdomens of uninfected females reveals that *nos*-GAL4: VP16 expresses *cifA*$_{wMel}$ significantly higher (Mdn = 1.08; p < 0.0001) than *otu*-

GAL4:VP16 (Mdn = 0.03) (Figure IV-3b), suggesting that high expression in females may underpin the ability to rescue. Alternatively, *nos*-GAL4:VP16 and *otu*-GAL4:VP16 are known to express GAL4 at different times in oogenesis, with the former in all egg chambers and the latter in late stage egg chambers (Ni et al., 2011).

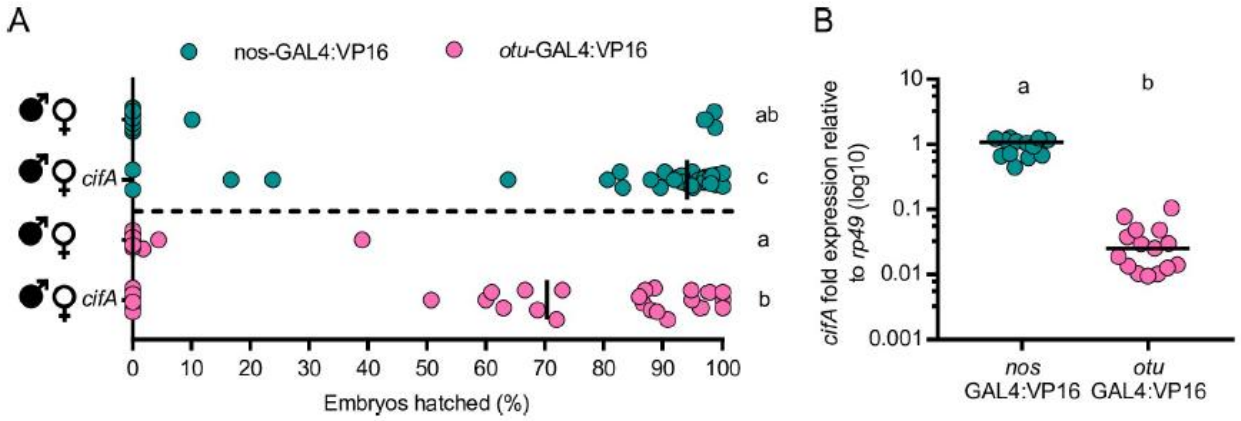

Figure IV-3. *cifA*$_{wMel}$ can induce strong rescue when expressed in uninfected females under the *nos*-GAL4:VP16 driver.
(A) Two different driver lines, *nos*-GAL4:VP16 (green; top) and *otu*-GAL4:VP16 (pink; bottom), were tested for their ability to rescue *w*Mel induced CI. Filled sex symbols represent infection with *w*Mel *Wolbachia*, and gene names to the right of a symbol represent expression of those genes in the corresponding sex of that cross. Vertical bars represent medians. Letters to the right indicate significant differences with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. (B) To test if nos-GAL4-tubulin and nos- GAL4:VP16 generate different levels of RNA expression, *cifA*$_{wMel}$ fold expression difference relative to the *Drosophila* housekeeping gene *rp49* in male abdomens under the two drivers was measured using qPCR. Females tested for gene expression were the same used in the hatch rate experiment in A. Letters above indicate significant differences with an $\alpha = 0.05$ calculated by a Mann-Whitney U test.

The Two-by-One model of CI.

With the transgenic expression system optimized for both transgenic CI and rescue, we then tested the hypothesis that the Two-by-One model can be synthetically recapitulated by dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression in uninfected males to cause CI and single *cifA*$_{wMel}$ expression in uninfected females to rescue that transgenic CI. Indeed, dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression in uninfected males causes hatch rates comparable to wild type CI (Mdn = 0.0%; IQR = 0.0%-2.55; $p > 0.99$) (Figure IV-4). Transgenic CI cannot be rescued by single *cifB*$_{wMel}$ expression in uninfected females (Mdn = 1.25%; IQR = 0.0–3.35%). Transgenic CI can be rescued by single *cifA*$_{wMel}$ expression (Mdn = 98.6%; IQR = 97.35–100%; $p = 0.41$) or dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression (Mdn = 96.7%; IQR = 88.3–98.2%; $p > 0.99$) to levels comparable to rescue from a wild type infection (Mdn = 95.6%; IQR = 92.5–97.4%). In addition, *cifA*$_{wMel}$ rescues a wild type infection at comparable levels to wild type rescue (Mdn = 96.6%; IQR = 93.5–98.85%; $p > 0.99$).

These data provide strong evidence for the Two-by-One model in *w*Mel infected *D. melanogaster*, namely that CI induced by transgenic dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression is sufficient to induce strong CI, and that *cifA*$_{wMel}$ alone is sufficient to rescue it.

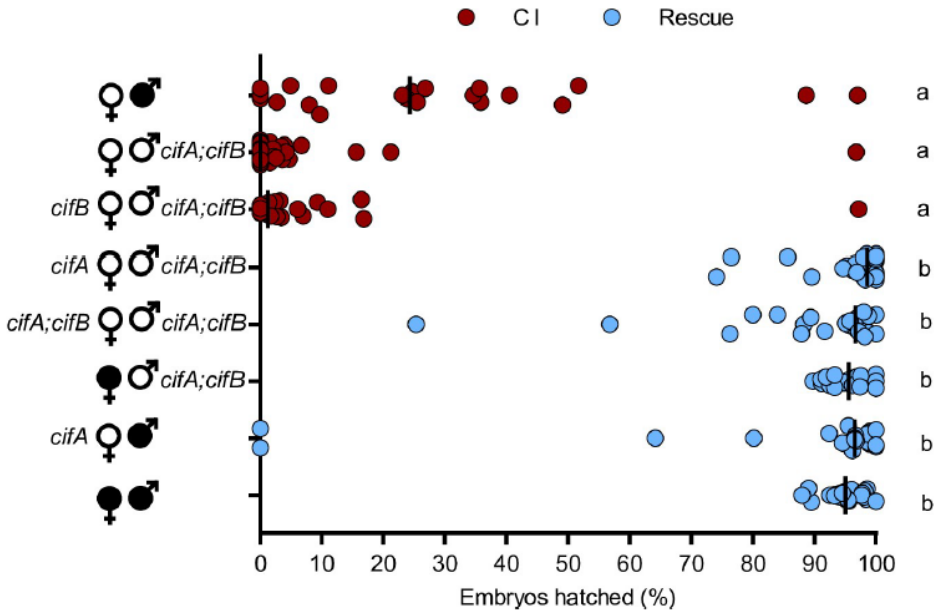

Figure IV-4. CI and rescue can be synthetically recapitulated under transgenic expression in the absence of *Wolbachia*.
Single *cifA*$_{wMel}$ and dual *cifA*$_{wMel}$ *and cifB*$_{wMel}$ expression under *nos*-GAL4:VP16 in uninfected females (open circles) were tested for their ability to rescue transgenic CI under the same driver in uninfected males. Filled sex symbols represent infection with *w*Mel *Wolbachia*, and gene names beside a symbol represent expression of those genes in the corresponding sex of that cross. Vertical bars represent medians. Letters to the right indicate significant differences with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test.

Next we reevaluated if single *cifA*$_{wMel}$ or *cifB*$_{wMel}$ expression under the more potent *nos*-GAL4:VP16 driver in uninfected males can recapitulate CI. Hatch rates indicate that dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression induces strong transgenic CI (Mdn = 0.0%; IQR = 0.0–1.15%) that can be rescued by a wild type infection (Mdn = 93.8%; IQR = 88.2–97.4%), whereas single expression of *cifA*$_{wMel}$ (Mdn = 96.1%; IQR = 97.78–98.55%; $p < 0.0001$) or *cifB*$_{wMel}$ (Mdn = 92.85%; IQR = 84.28–96.4%; $p < 0.0001$) failed once again to produce embryonic hatching comparable to expressing both genes together (Figure IV-5). In one replicate experiment, we note a statistically insignificant ($p = 0.182$) decrease in hatching under *cifB*$_{wMel}$ expression relative to wild type rescue cross. Thus, both *cifA*$_{wMel}$ and *cifB*$_{wMel}$ are required for strong CI. Together, these and earlier results validate the Two-by-One model of CI in *w*Mel whereby *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression are required and sufficient for strong CI, while *cifA*$_{wMel}$ expression is sufficient to rescue it.

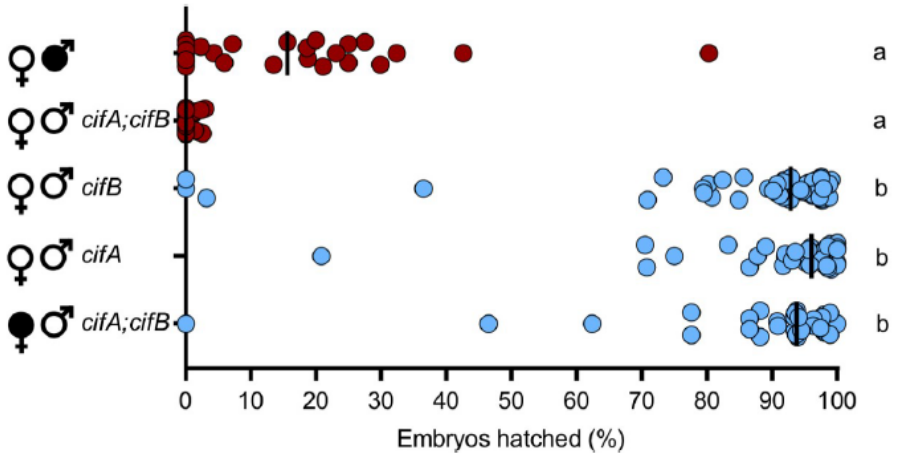

Figure IV-5. Neither *cifA*$_{wMel}$ or *cifB*$_{wMel}$ alone can induce CI when expressed under *nos*-GAL4:VP16.
cifA$_{wMel}$ and *cifB*$_{wMel}$ were tested for their ability to induce CI individually under *nos*-GAL4:VP16 expression in uninfected males (open circles). Filled sex symbols represent infection with *w*Mel *Wolbachia* and gene names to the right of a symbol represent expression of those genes in the corresponding sex of that cross. Vertical bars represent medians. Letters to the right indicate significant differences with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test.

Discussion

CI is the most common form of *Wolbachia*-induced reproductive parasitism and is currently at the forefront of vector control efforts to curb transmission of dengue, Zika, and other arthropod-borne human pathogens (Caragata et al., 2016; Huang et al., 2018; Hughes et al., 2011; O'Connor et al., 2012; O'Neill, 2018; Schmidt et al., 2017). Two prophage WO genes from *w*Mel *Wolbachia* cause CI (*cifA*$_{wMel}$ and *cifB*$_{wMel}$) and one rescues wild type CI (*cifA*$_{wMel}$) (LePage et al., 2017; Shropshire et al., 2018), supporting the proposal of a Two-by-One model for the genetic basis of CI (Shropshire et al., 2018). However, dual transgenic expression of *cifA*$_{wMel}$ and *cifB*$_{wMel}$ recapitulates only weak and highly variable CI as compared to CI induced by a wild type infection (LePage et al., 2017). In addition, the Two-by-One model predicts that both CI and rescue can be synthetically recapitulated by dual *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression in uninfected males and *cifA*$_{wMel}$ expression in uninfected females. Here we optimized the transgenic system for CI and rescue by these genes, further validated the necessity of expressing both *cifA*$_{wMel}$ and *cifB*$_{wMel}$ for CI, and synthetically recapitulated the Two-by-One model for CI with transgenics in the absence of *Wolbachia*.

CI induced by *w*Mel *Wolbachia* can be highly variable and correlates with numerous factors including *Wolbachia* density (Bourtzis et al., 1996), *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression levels

(LePage et al., 2017), host age (Awrahman et al., 2014; Reynolds et al., 2003; Reynolds and Hoffmann, 2002), mating rate (Awrahman et al., 2014), rearing density (Yamada et al., 2007), development time (Yamada et al., 2007), and host genetic factors (Cooper et al., 2017; Jaenike, 2007; Poinsot et al., 1998; Reynolds and Hoffmann, 2002). Some of these factors, such as age, are known to also correlate with the level of cif_{wMel} gene expression (LePage et al., 2017). As such, we hypothesized that prior reports of weakened transgenic CI could be explained by low levels of transgenic $cifA_{wMel}$ and $cifB_{wMel}$ expression in male testes (LePage et al., 2017).

Indeed, strong CI with a median of 0% embryonic hatching was induced when both $cifA_{wMel}$ and $cifB_{wMel}$ were expressed under the *nos*-GAL4:VP16 driver. However, contrary to our expectations, *nos*-GAL4:VP16 generates significantly weaker $cifA_{wMel}$ and $cifB_{wMel}$ expression than the *nos*-GAL4-*tubulin* driver previously used to recapitulate weak CI (LePage et al., 2017). Thus, the expression data conflict with previous reports in mammalian cells wherein the GAL4:VP16 fusion protein is a more potent transcriptional activator than GAL4 (Sadowski et al., 1988). Other differences between the two driver constructs may explain phenotypic differences, including the presence of different 3' UTRs that may contribute to differences in transcript localization (Ni et al., 2011). While it remains possible, though unlikely, that other *Wolbachia* or phage WO genes may contribute to CI, the induction of near complete embryonic lethality confirms that $cifA_{wMel}$ and $cifB_{wMel}$ are sufficient to transgenically induce strong CI and do not require other *Wolbachia* or phage WO genes to do so. Moreover, comparative multi-omics demonstrated that *cifA* and *cifB* are the only two genes strictly associated with CI capability (LePage et al., 2017).

We previously recapitulated transgenic rescue of *w*Mel-induced CI by expression of $cifA_{wMel}$ under the Maternal Triple Driver (MTD) (Shropshire et al., 2018), which is comprised of three independent drivers (Ni et al., 2011). Expression of $cifA_{wMel}$ using one of the MTD drivers in flies was previously shown to be rescue-incapable (Shropshire et al., 2018); the other drivers had not been evaluated. Here, we tested the hypothesis that expression of $cifA_{wMel}$ using either of the two remaining drivers is sufficient to rescue CI, and we found that $cifA_{wMel}$ expression under both driver lines recapitulates rescue, but at different strengths. Indeed, rescue is strongest when $cifA_{wMel}$ transgene expression is highest. These data are consistent with reports that $cifA_{wMel}$ is a highly expressed gene in transcriptomes of *w*Mel-infected females (Gutzwiller et al., 2015) and the hypothesis that rescue capability is largely determined by the strength of $cifA_{wMel}$ expression in

ovaries (Shropshire et al., 2018). These results combined with those for transgenic expression of CI now establish a robust set of methods for future studies of transgene-induced CI and rescue in the *D. melanogaster* model.

The central prediction of the Two-by-One model is that transgenic CI can be synthetically rescued in the absence of *Wolbachia* through dual *cifA* and *cifB* expression in uninfected males and *cifA* expression in uninfected females. Here, we explicitly validate the model that two genes are required in males to cause CI, and one in females is required to rescue it using *w*Mel *cif* gene variants. However, to confirm that the optimized expression system does not influence the ability of $cifA_{wMel}$ or *cifBwMel* alone to induce CI, we singly expressed them with the improved driver and found that embryonic hatching does not statistically differ from compatible crosses. Coupled with prior data in *w*Mel (LePage et al., 2017; Shropshire et al., 2018), these results strongly support the Two-by-One genetic model whereby dual $cifA_{wMel}$ and $cifB_{wMel}$ expression is required in the testes to cause a sperm modification that can then be rescued by $cifA_{wMel}$ expression in the ovaries (Figure IV-6).

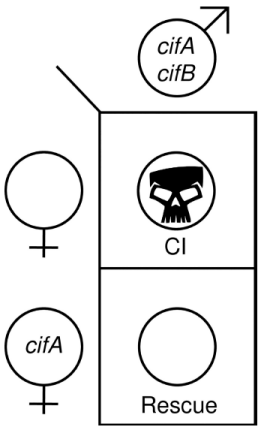

Figure IV-6. The Two-by-One model of CI.
The Two-by-One genetic model of CI explains that *cifA* and *cifB* dual expression in uninfected males is necessary for embryonic lethality (CI; skull) when crossed to uninfected and nonexpressing females. However, females expressing *cifA* can rescue CI in their offspring (rescue; open circle). Skull art is from vecteezy.com.

While the genetic basis of unidirectional CI appears resolved, it remains unclear how $cifA_{wMel}$ and $cifB_{wMel}$ functionally operate to generate these phenotypes. Numerous mechanistic models have been proposed over the last two decades (Beckmann et al., 2019a; Bossan et al., 2011; Ferree and Sullivan, 2006; Landmann et al., 2009; Poinsot et al., 2003; Shropshire et al., 2019; Tram and Sullivan, 2002). We can broadly summarize these models into either host-modification (HM) (Shropshire et al., 2019) or toxin-antidote (TA) (Beckmann et al., 2019a) models. HM

models suggest that CI-inducing factors modify host products in such a way that would be lethal unless they are later reversed by rescue factors (Bossan et al., 2011; Ferree and Sullivan, 2006; Landmann et al., 2009; Poinsot et al., 2003; Shropshire et al., 2019; Tram and Sullivan, 2002). Conversely, TA models state that the CI-inducing factor is toxic to the developing embryo unless it is crucially bound to a cognate antidote provided by the female (Beckmann et al., 2017, 2019a; Shropshire et al., 2019). There are numerous lines of evidence in support of both sets of hypotheses and while the Two-by-One genetic model does not explicitly support or favor one set of models over the other, it can be used to generate hypotheses related to the mechanism of CI.

HM models (Shropshire et al., 2019) predict that CI factors directly interact with host products in the testes, modify them, and are displaced. These modifications travel with the sperm, in the absence of pronuclei are delayed in the first mitosis during embryonic development in CI crosses (Callaini et al., 1997; Ferree and Sullivan, 2006; Tram et al., 2006). Since the first mitosis is initiated when the female pronucleus has developed, the delay of the male pronuclei leads to cytological defects (Tram and Sullivan, 2002). It is thus proposed that rescue occurs through resynchronization of the first mitosis by comparably delaying the female pronucleus (Ferree and Sullivan, 2006; Tram and Sullivan, 2002). The Goalkeeper model expands the mistiming model to propose that the strength of the delay is what drives incompatibility between different *Wolbachia* strains (Bossan et al., 2011). There are numerous hypotheses to explain the role of the Cif products in these kinds of models. One such hypothesis would be that CifA is responsible for pronuclear delay, thus capable of delaying both the male and female pronuclei, but it requires CifB to properly interact with testis-associated targets. This hypothesis may predict that CifB acts to either protect CifA from ubiquitin tagging and degradation, localize it to a host target, or bind CifA to elicit a conformational change required for interacting with male-specific targets. Alternatively, CI-affected embryos express defective paternal histone deposition, protamine development, delayed nuclear breakdown, and delays in replication machinery (Callaini et al., 1997; Ferree and Sullivan, 2006; Landmann et al., 2009; Lassy and Karr, 1996; Serbus et al., 2008; Tram et al., 2006; Tram and Sullivan, 2002). Any of these factors could be explained by modifications occurring from HM-type interactions between Cif and host products.

TA models (Beckmann et al., 2019a) contrast to HM models and require that the CI toxin transfers with or in the sperm and directly binds to a female-derived antidote in the embryo. If the antidote is absent, the CI toxin would induce cytological embryonic defects (Callaini et al., 1997;

Ferree and Sullivan, 2006; Landmann et al., 2009; Lassy and Karr, 1996; LePage et al., 2017; Serbus et al., 2008; Tram et al., 2006; Tram and Sullivan, 2002). There is mixed evidence in support of this model. First, mass spectometry and SDS-PAGE analyses in *Culex pipiens* reveal that $CifA_{wPip}$ peptides are present in female spermatheca after mating, suggesting $CifA_{wPip}$ is transferred with or in the sperm (Beckmann and Fallon, 2013). $CifB_{wPip}$ was not detected in these analyses, curiously suggesting that the CifB toxin was not transferred (Beckmann and Fallon, 2013). These results are inconsistent with the TA model, but the lack of transferred CifB may occur because *cifB* gene expression is up to nine-fold lower than that of *cifA* (Gutzwiller et al., 2015), and the concentration may have been too low to be observed via these methods. Second, CifA and CifB bind in vitro (Beckmann et al., 2017). However, it remains unclear if CifA-CifB binding enables rescue since this binding has no impact on known enzymatic activities of CifB (Beckmann et al., 2017). While the Two-by-One model does not explicitly support or reject the TA model, it does further inform it. Most intriguing is to understand how CifA acts as a contributor to CI when expressed in testes and as a rescue factor when expressed in ovaries. One hypothesis is that CifA and CifB bind to form a toxin complex that is later directly inhibited by female derived CifA (Shropshire et al., 2019, 2018). The difference in function between these two environments could be explained by post-translational modification and/or differential localization of CifA in testes and embryos (Shropshire et al., 2019, 2018). Alternatively, CifB may be the primary toxin, but is incapable of inducing CI unless a CifA antidote is present in both the testes and the ovaries (Beckmann et al., 2019a). This hypothesis predicts that male-derived CifA rapidly degrades, leaving CifB with or in the sperm. On its own, CifB would induce lethal cytological embryonic defects (Ferree and Sullivan, 2006; Landmann et al., 2009; Poinsot et al., 2003; Tram and Sullivan, 2002) unless provided with a fresh supply of CifA from the embryo.

It has been suggested that divergence in CI and rescue factors causes the incipient evolution of reciprocal incompatibility, or bidirectional CI, between different *Wolbachia* strains (Bonneau et al., 2018a; Charlat et al., 2001; LePage et al., 2017; Shropshire et al., 2018). Here, we review a non-exhaustive set of hypotheses that we previously proposed to explain the emergence of bidirectional CI and are consistent with the Two-by-One model (Shropshire et al., 2018). First, the simplest explanation for CifA's role in both CI and rescue is that it has similar functional effects in both testes/sperm and ovaries/embryos. Thus, instead of requiring a separate mutation for CI and another for rescue (Charlat et al., 2001), bidirectional CI may emerge from a single CifA

mutation that causes incompatibility against the ancestral strain while maintaining selfcompatibility. Second, CifA in testes and ovaries may also have different functions, localizations, or posttranslational modifications that contribute to CI and rescue. If this occurs, or if CifB is also an incompatibility factor, the evolution of bidirectional CI may require two or more mutations, and the strain may pass through an intermediate phenotype wherein it becomes unidirectionally incompatible with the ancestral variant or loses the capability to induce either CI or rescue before becoming bidirectionally incompatible with the ancestral variant. In fact, some *Wolbachia* strains are incapable of inducing CI but capable of rescuing CI induced by other strains (Bourtzis et al., 1998), and some can induce CI but cannot be rescued (Zabalou et al., 2008). Furthermore, sequence variation in both *cifA* and *cifB* from *Wolbachia* strains in *Drosophila* (LePage et al., 2017) and in small regions among strains of *w*Pip *Wolbachia* (Bonneau et al., 2018a) have been correlated to incompatibility, suggesting that variation in both genes influence incompatibility.

Additionally, it remains possible that significant divergence in *cifA, cifB,* or both may be necessary to generate new phenotypes. Indeed, comparative genomic analyses reveal high levels of amino acid divergence in CifA and CifB that correlates with incompatibility between strains (LePage et al., 2017; Lindsey et al., 2018b). Moreover, some *Wolbachia* strains harbor numerous phage WO variants, each with their own, often divergent, *cif* genes, and the presence of multiple variants likewise correlates with incompatibility (Bonneau et al., 2018a; LePage et al., 2017; Lindsey et al., 2018a). Thus, horizontal transfer of phage WO (Bordenstein and Bordenstein, 2016; Chafee et al., 2010; Kent et al., 2011; G. H. Wang et al., 2016b, 2016a; N. Wang et al., 2016) can in theory rapidly introduce new compatibility relationships, and duplication of phage WO regions, or specifically *cif* genes, in the same *Wolbachia* genome may relax the selective pressure on the *cif* genes and enable their divergence. Determining which of the aforementioned models best explains the evolution of incompatibilities between *Wolbachia* strains will be assisted by additional sequencing studies to identify incompatible strains with closely related *cif* variants.

The genetic bases of numerous gene drives have been elucidated in plants (Yang et al., 2012), fungi (Grognet et al., 2014; Hammond et al., 2012; Hu et al., 2017; Nuckolls et al., 2017), and nematodes (Ben-David et al., 2017; Seidel et al., 2011). Some gene drives have also been artificially replicated with transgenic constructs (Akbari et al., 2014, 2013; Chen et al., 2007). However, to our knowledge, the synthetic replication of the Two-by-One model of CI represents

the first instance that a gene drive has been constructed by engineering eukaryotic reproduction to depend on phage proteins. Additionally, vector control programs using *Wolbachia* rely on their ability to suppress pathogens such as Zika and dengue viruses, reduce the size of vector populations, and spread *Wolbachia* into a host population via CI and rescue. However, there are limitations to these approaches. Most critically, not all pathogens are inhibited by *Wolbachia* infection and some are enhanced, such as West Nile Virus in *Culex tarsalis* infected with *w*AlbB *Wolbachia* (Dodson et al., 2014). Additionally, it requires substantial effort to establish a *Wolbachia* transinfection in a target non-native species (Hughes and Rasgon, 2014) that could be obviated in genetically tractable vectors utilizing transgenic gene drives.

The complete synthetic replication of CI and rescue via the Two-by-One model represents a step towards transgenically using the *cif* genes in vector control efforts. The separation of CI mechanism from *Wolbachia* infection could theoretically expand CI's utility to spread 'payload' genes that reduce the vectoral capacity of their hosts (Champer et al., 2016) into a vector population by, for instance, expressing the CI genes and the payload gene polycistronically under the same promoter in the vector's nuclear or mitochondrial genomes. Moreover, these synthetic constructs have potential to increase the efficiency of *Wolbachia*-induced CI if they are transformed directly into *Wolbachia* genomes. For these efforts to be successful, considerable work is necessary to (i) generate a constitutively expressing *cif* gene drive that does not require GAL4 to operate, (ii) understand the spread dynamics of transgenic CI, (iii) characterize the impact of *cif* transgenic expression on insect fitness relative to wild vectors, (iv) generate and test effective payload genes in combination with *cif* drive, (v) explore and optimize the efficacy of *cif* drive in vector competent hosts such as mosquitoes, (vi) assess the impact of host factors on *cif* drive across age and development, (vii) compare the efficacy of a *cif* gene drive to other comparable technologies (CRISPR, homing drive, Medea, etc), in addition to numerous other lines of study. For example, while a substantial body of literature exists to describe the spread dynamics of CI (Hoffmann et al., 1990; Jansen et al., 2008; Rasgon, 2008; Turelli, 1994; Turelli et al., 2018b; Turelli and Barton, 2017), none yet describe how the Two-by-One model would translate into nuclear or mitochondrial spread dynamics in the absence of *Wolbachia*. As such, this study represents an early proof of concept that these genes alone are capable of biasing offspring survival in favor of flies expressing these genes under strictly controlled conditions, and should motivate additional study towards its application in vector control.

The generality of the Two-by-One model remains to be tested because it may be specific to certain strains of *Wolbachia* and/or phage haplotypes. For instance, transgenic expression of $cifB_{wPip}$ from *C. pipiens* in yeast yields temperature sensitive lethality that can be rescued by dual-expression of $cifA_{wPip}$ and $cifB_{wPip}$ (Beckmann et al., 2017). Moreover, attempts to generate a $cifB_{wPip}$ transgenic line failed, possibly due to generalized toxicity from leaky expression (Beckmann et al., 2017). Therefore, $cifB_{wPip}$ alone could in theory cause CI. However, this model has not been explicitly tested, it has not been explained how $cifA_{wPip}$ and $cifB_{wPip}$ dual-expression induces CI in transgenic *Drosophila* but prevents CI in yeast, and transgenic *w*Pip CI has not been rescued in an insect. As such, it remains possible that $cifB_{wPip}$ lethality could be explained by artefactual toxicity of overexpression or toxic expression in a heterologous system. Thus, confirmation of an alternative model for CI in *w*Pip is precluded by lack of evidence that $cifB_{wPip}$ alone can induce rescuable lethality in an insect. Since $cifB_{wPip}$ transgenic UAS constructs have not been generated due to toxicity from leaky expression, alternative PhiC31 landing sites or expression systems (i.e., the Q System) could prove valuable in addressing these questions.

Finally, these results further validate the importance of $cifA_{wMel}$ as an essential component of CI and underscore a community need to unify the nomenclature of the CI genes. When the CI genes were first reported, they were described as both CI factors (*cif*) and as CI deubiquitilases (*cid*), both of which are actively utilized in the literature. The *cif* nomenclature was proposed as a cautious naming strategy agnostic to the varied biochemical functions to be discovered, whereas the *cid* nomenclature was proposed based on the finding that the B protein is in part an *in vitro* deubiquitilase that, when ablated, inhibits CI-like induction (Beckmann et al., 2017; LePage et al., 2017). A recent nomenclature proposal suggested that the *cif* gene family name be used as an umbrella label to describe all CI-associated factors whereas *cidA* and *cidB* would be used to describe the specific genes (Beckmann et al., 2019a). However, we do not agree with this nomenclature revision despite the appeal of combining the two nomenclatures. CifA protein is not a putative deubiquitilase (Lindsey et al., 2018b), does not influence deubiquitilase activity of CifB (Beckmann et al., 2017), functions independently to rescue CI (Shropshire et al., 2018) and, as emphasized by the work in this study, is necessary for CI induction and rescue. The competing nomenclature presumes that it is appropriate to name the A protein *cid* because it could be expressed in an operon with the B protein. However, the evidence for the operon status of the genes is weak, and more work is needed to describe the regulatory control of these genes before

they can be categorized as an operon (Shropshire et al., 2019). Moreover, distant homologs that cluster into distinct phylogenetic groups are proposed to be named CI nucleases (*cin*) (Beckmann et al., 2017) yet the merger of these two groups into one name lacks phylogenetic rationality as the two lineages are as markedly divergent from each other as they are from *cid* (Shropshire et al., 2019). In addition, none of these distant homologs have been functionally characterized as CI genes (LePage et al., 2017; Lindsey et al., 2018b). As such, it is more appropriate to call these genes "*cif*-like" to reflect their homology and unknown phenotypes. Thus, the holistic and conservative *cif* nomenclature with Types (e.g., I-IV) used to delineate phylogenetic clades is appropriately warranted in utilizing and unifying CI gene naming.

In conclusion, the results presented here support that both $cifA_{wMel}$ and $cifB_{wMel}$ phage genes are necessary and sufficient to induce strong CI. In addition, $cifA_{wMel}$ is the only gene necessary for rescue of either transgenic or wild type *w*Mel CI. These results confirm the Two-by-One model of CI in *w*Mel *Wolbachia* and phage WO with implications for the mechanism of CI and for the diversity of incompatibility between strains, and they provide additional context for understanding CI currently deployed in vector control efforts. The synthetic replication of CI in the absence of *Wolbachia* marks an early step in developing CI as a tool for genetic and mechanistic studies in *D. melanogaster* and for vector control efforts that may drive payload genes into vector competent populations.

Materials and methods

Fly rearing and strains.

D. melanogaster stocks y^1w* (BDSC 1495), *nos*-GAL4-*tubulin* (BDSC 4442), *nos*-GAL4:VP16 (BDSC 4937), *otu*-GAL4:VP16 (BDSC 58424), and UAS transgenic lines homozygous for *cifA*, *cifB*, and *cifA;B* (LePage et al., 2017) were maintained at 12:12 light:dark at 25ºC and 70% relative humidity (RH) on 50 ml of a standard media. *cifA* insertion was performed with y^1M{vas-int.Dm}ZH-2A w*; P{CaryP}attP40 and *cifB* insertion was performed with y^1w67c23; P{CaryP}attP2, as previously described (LePage et al., 2017). UAS transgenic lines and *nos*-GAL4:VP16 were uninfected whereas *nos*-GAL4-*tubulin* and *otu*-GAL4:VP16 lines were infected with *w*Mel *Wolbachia*. Uninfected versions of infected lines were produced through tetracycline treatment as previously described (LePage et al., 2017). WolbF and WolbR3 primers

were regularly used to confirm infection status (LePage et al., 2017). Stocks for virgin collections were stored at 18^{o}C overnight to slow eclosion rate, and virgin flies were kept at room temperature.

Hatch rate assays.

To test for CI, hatch rate assays were used as previously described (LePage et al., 2017; Shropshire et al., 2018). Briefly, GAL4 adult females were aged 9–11 days post eclosion and mated with UAS males. Age controlled GAL4-UAS males and females were paired in 8 oz bottles affixed with a grape-juice agar plate smeared with yeast affixed to the opening with tape. 0–48 hour old males were used since CI strength rapidly declines with male age (Awrahman et al., 2014; Reynolds and Hoffmann, 2002). The flies and bottles were stored at 25^{o}C for 24 h at which time the plates were replaced with freshly smeared plates and again stored for 24 h. Plates were then removed and the number of embryos on each plate were counted and stored at 25^{o}C. After 30 h the remaining unhatched embryos were counted. The percent of embryos hatched into larvae was calculated by dividing the number of hatched embryos by the initial embryo count and multiplying by 100.

Expression analyses.

To assay transgenic RNA expression levels under the various gene drive systems, transgene expressing flies from hatch rates were immediately collected and frozen at -80˚C for downstream application as previously described (Shropshire et al., 2018). In brief, abdomens were dissected, RNA was extracted using the Direct-zol RNA MiniPrep Kit (Zymo), the DNA-free kit (Ambion, Life Technologies) was then used to remove DNA contamination, and cDNA was generated with SuperScript VILO (Invitrogen). Quantitative PCR was performed on a Bio-Rad CFX-96 Real-Time System in duplicate using iTaq Universal SYBR Green Supermix (Bio-Rad) using the *cifA*_opt and *rp49* forward and reverse primers as previously described (Shropshire et al., 2018). Samples with a standard deviation >0.3 between duplicates were excluded from analysis. Fold expression of *cifA* relative to *rp49* was determined with $2^{-\Delta\Delta Ct}$. Each expression study was conducted once.

Statistical analyses.

All statistical analyses were conducted in GraphPad Prism (Prism 8). Hatch rate statistical comparisons were made using Kruskal-Wallis followed by a Dunn's multiple comparison test. A Mann-Whitney-U was used for statistical comparison of RNA fold expression. A linear regression was used to assess correlations between hatch rate and expression. All p-values are reported in Table C-1.

Chapter V.

Cif genotypes and cytoplasmic incompatibility phenotypes: impacts on strain (in)compatibilities and penetrance

Abstract

Wolbachia are maternally-transmitted, intracellular bacteria that occur in roughly half of arthropod species. Within species, *Wolbachia* often rapidly spread through populations via a selfish drive system termed cytoplasmic incompatibility (CI). According to the Two-by-One genetic model of CI, offspring die as embryos when males dually expressing prophage WO genes *cifA* and *cifB* are crossed with uninfected females or females harboring an incompatible *Wolbachia* strain. However, females with a compatible strain expressing *cifA* rescue embryos from CI. Thus, CI mediated by *cif* genes confers a relative fitness advantage to females that transmit *Wolbachia* and phage WO. However, the genetic determinants of (in)compatibilities between *Wolbachia* strains and CI level variation remain unknown. *cifA* and *cifB* sequences diverged into at least 5 distinct phylogenetic clades referred to as Types 1 - 5. Here, we engineer *Drosophila melanogaster* to transgenically express cognate and non-cognate pairs of Type 1 and 2 *cif* variants and assess their CI and rescue potential. The combinatorial approach reveals that a resident cognate *cif* pair in *D. melanogaster* causes strong CI, but cognate *cif* pairs from Type 1 and Type 2 homologs from other *Drosophila* species cause weak transgenic CI. We take advantage of this variation in transgenic CI levels to explicitly link variation in the more rapidly evolving *cifB* sequence to weak transgenic CI, and determine that all Type 1 *cifA* sequences evaluated can contribute to strong transgenic CI and interchangeable rescue despite their evolutionary divergence. However, while we present the first evidence that Type 2 *cifA* and *cifB* can contribute to rescue and CI in cognate and non-cognate pairings with Type 1 genes, they cause only weak reductions in hatching. Finally, we find that Type 1 *cifA* can rescue Type 2 transgenic CI, but the inverse is not compatible, thus indicating a unidirectional CI between Type 1 and 2 *cifs*. Results add new support to the Two-by-One genetic model of CI and reveal previously unrecognized relationships between *cif* genotype, host genotype, and CI phenotype as it relates to transgenic CI levels and compatibility

relationships. We discuss the relevance of these findings to evolutionary and mechanistic models of the CI drive system.

Introduction

Wolbachia are the most common intracellular bacteria in animals, occurring in 40-65% of arthropod species (Charlesworth et al., 2019; Hilgenboecker et al., 2008; Weinert et al., 2015; Zug and Hammerstein, 2012). While often horizontally transmitted between species (Boyle et al., 1993; Frydman et al., 2006; Gerth et al., 2013; Huigens et al., 2004; Tolley et al., 2019), vertical transmission from mother to offspring predominates within species (Narita et al., 2009; Turelli and Hoffmann, 1991). *Wolbachia* can increase their rate of spread through the matriline by causing cytoplasmic incompatibility (CI), which results in embryonic death when *Wolbachia*-modified sperm fertilize uninfected embryos or embryos carrying an incompatible *Wolbachia* strain (Yen and Barr, 1973). Embryos with compatible *Wolbachia* are rescued from this lethality, yielding a relative fitness advantage to *Wolbachia*-infected females that transmit *Wolbachia* to their offspring (Hoffmann et al., 1990; Turelli, 1994; Turelli and Hoffmann, 1995). CI-inducing *Wolbachia* are at the forefront of vector control programs to reduce the spread of pathogenic RNA viruses, including dengue and Zika, since CI can be used to either suppress mosquito population sizes (Dobson et al., 2002; Lees et al., 2015; Nikolouli et al., 2018; O'Connor et al., 2012) or to replace uninfected mosquito populations with mosquitoes infected with *Wolbachia* that block arbovirus replication/transmission (Moretti et al., 2018; O'Neill, 2018). Moreover, CI can act as a mechanism of incipient speciation by reducing gene flow between populations of different infection states (Brucker and Bordenstein, 2012; Shropshire and Bordenstein, 2016), such as in *Nasonia* wasps (Bordenstein et al., 2001) and *Drosophila* flies (Jaenike et al., 2006)

The genetic basis of CI is captured by the Two-by-One genetic model (Figure V-1A) (Shropshire and Bordenstein, 2019). Two adjacent genes in the eukaryotic association module of *Wolbachia's* prophage WO cause CI when expressed in males (*cifA* and *cifB*) (Beckmann et al., 2017; Bordenstein and Bordenstein, 2016; Chen et al., 2019; LePage et al., 2017; Shropshire and Bordenstein, 2019) and one gene expressed in females (*cifA*) rescues CI (Chen et al., 2019; Shropshire et al., 2018; Shropshire and Bordenstein, 2019). Singly expressing *cifA* or *cifB* does not cause CI (Beckmann et al., 2017; Chen et al., 2019; LePage et al., 2017; Shropshire and

Bordenstein, 2019). The Two-by-One model has been validated using *cif* transgenes from the *w*Mel *Wolbachia* of *D. melanogaster* and is supported by transgenic expression of *cifA* and *cifB* genes from *w*Pip *Wolbachia* of *Culex pipiens* mosquitoes (Beckmann et al., 2017; Chen et al., 2019; Shropshire and Bordenstein, 2019). Phylogenetic analysis of the *cif* genes reveals at least five distinct clades designated Types 1 – 5 (Bing et al., 2020; LePage et al., 2017; Lindsey et al., 2018). The only *cif* genes functionally validated to induce and rescue CI are those in Type 1 (Beckmann et al., 2017; LePage et al., 2017; Shropshire et al., 2018; Shropshire and Bordenstein, 2019) and Type 4 clades (Chen et al., 2019), leaving a considerable amount of phylogenetic diversity untested.

CI frequently manifests between arthropods infected with different *Wolbachia* strains. Strains may be reciprocally incompatible (bidirectional CI), or only one of the two strains can rescue the other's sperm modification (unidirectional CI). The genetic basis of incompatibilities between *Wolbachia* strains remains unknown but is hypothesized to be caused by divergence in the gene(s) underpinning CI and rescue (Bonneau et al., 2018; Charlat et al., 2001; LePage et al., 2017; Shropshire et al., 2018). Phylogenetic and sequence analyses of *cif* genes from incompatible *Wolbachia* strains in *Drosophila* or *Culex* reveal that incompatible strains differ in genetic relationship and/or copy number (Bonneau et al., 2019, 2018; LePage et al., 2017). Moreover, since *cifA* is involved in both CI-induction and rescue, a simple, single-step evolutionary model for bidirectional CI can be hypothesized where potentially a single mutation in *cifA* leads to incompatibility between the ancestral and derived variants while retaining compatibility with the emergent variant (Shropshire et al., 2018). These studies and models suggest a correlation between *cif* gene sequence variation (and perhaps copy number variation) with (in)compatibility relationships between strains. However, these hypotheses have not been empirically tested.

Here, we test *cif* homologs for their ability to induce and rescue CI when transgenically expressed in uninfected *D. melanogaster*. We focus on three strains of CI-inducing *Wolbachia*: *w*Mel from *D. melanogaster*, *w*Rec from *D. recens,* and *w*Ri from *D. simulans*. *w*Rec and *w*Ri are strong CI inducers that cause high degrees of embryonic death (Shoemaker et al., 1999; Turelli and Hoffmann, 1991; Werren and Jaenike, 1995) and both have phylogenetic Type 1 *cif* genes similar to *w*Mel (LePage et al., 2017) (Figure V-1B). *w*Ri also harbors phylogenetic Type 2 *cif* genes highly diverged from *w*Mel (LePage et al., 2017) (Figure V-1B). The *w*Ri and *w*Mel *Wolbachia* are unidirectionally incompatible: *w*Ri can rescue *w*Mel-induced CI, but *w*Mel cannot

rescue *w*Ri-induced CI (Poinsot et al., 1998) (Figure V-1C). We hypothesize that both *w*Ri's Type 1 and 2 genes are functional, *w*Ri can rescue *w*Mel-induced CI because it has Type 1 genes comparable to *w*Mel, and *w*Mel cannot rescue *w*Ri because it does not have genes capable of rescuing *w*Ri's Type 2 genes (LePage et al., 2017) (Figure V-1D). We also predict that *w*Rec Type 1 genes cause CI that is rescuable by the closely-related *w*Mel genes (Figure V-1C, D). We discuss our results in the context of evolutionary (co)divergence between cognate *cifA* and *cifB* genes, mechanistic models of CI, the emergence of incompatibility phenotypes, and the role of host background in CI expression.

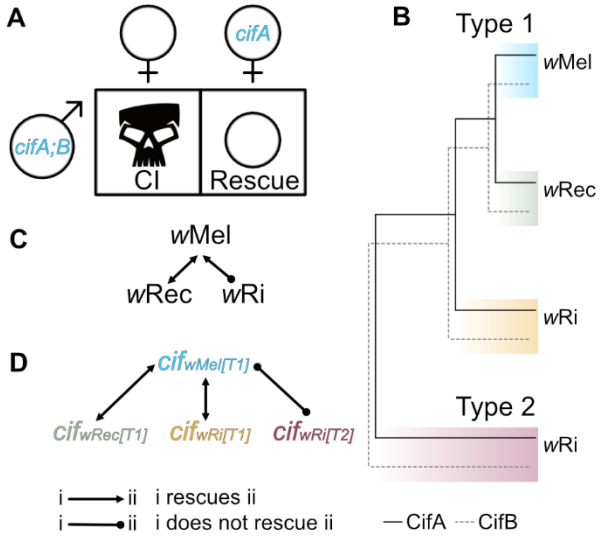

Figure V-1. Two-by-One model, Cif phylogeny, and predicted relationships between wMel, wRec, and wRi strains and cif gene variants.
(A) The Two-by-One genetic model of CI states that males expressing *cifA;B* cause CI that can be rescued by females expressing *cifA*. This model is based on experimental evidence from transgenic experiments using the CI genes of *w*Mel (Shropshire and Bordenstein, 2019). (B) Schematic of the evolutionary relationships between CifA and CifB proteins from *w*Mel, *w*Rec, and *w*Ri (LePage et al., 2017). (C) Putative (in)compatibilities between *w*Mel, *w*Rec, and *w*Ri *Wolbachia* strains. Unidirectional CI between *w*Mel and *w*Ri is expected based on crossing experiments after the transinfection of *w*Mel into *D. simulans* (Poinsot et al., 1998). Compatibility between *w*Mel and *w*Rec is based on the prediction that strains with closely-related *cif* gene sequences would be compatible. (D) Predicted (in)compatibility relationships between *cif* homologs from each of the three strains, based on phylogenetic Type and previously published CI data (LePage et al., 2017; Shropshire et al., 2018). *cif* is purposely used here instead of *cifA* and *cifB* to maintain an agnostic view that *cifA* in females may rescue potential *cifA*- *and*/or *cifB*-induced modifications expressed in males. Lines between strains/genes indicate compatibility relationships. If the line ends in an arrowhead then the strain/gene(s) at the beginning of the arrow can rescue CI caused by the strain/gene(s) the arrow points towards. If the line ends in a circle then rescue is not expected. Skull art is modified from vecteezy.com with permissions.

Results

To distinguish between different *cifA* and *cifB* genetic variants, we use a gene nomenclature system that identifies the *Wolbachia* strain in subscript and the *cif* phylogenetic Type associated with the variant in brackets (Shropshire et al., 2019). For instance, *cif* genes of the *w*Mel strain belong to the Type 1 clade and are referred to as $cifA_{wMel[T1]}$ and $cifB_{wMel[T1]}$. We test *cif* genetic variants using the GAL4-UAS system (Duffy, 2002) to drive the germline expression of *cif* transgenes in *D. melanogaster* males and females and evaluate their role in CI and rescue, respectively. Since CI manifests as embryonic death, we measure CI as the percentage of embryos that hatch into larvae. All statistical comparisons within hatch rate experiments are relative to a compatible control where CI caused by $cifA;B_{wMel[T1]}$ males is rescued by $cifA_{wMel[T1]}$ females, as expected from prior transgenic studies (LePage et al., 2017; Shropshire et al., 2018; Shropshire and Bordenstein, 2019). This cross is included in all experiments and will hereafter be referred to as the "compatible control". All transgenes are expressed in uninfected flies using the *nos*-GAL4:VP16 driver previously shown to allow for strong CI and rescue when driving the expression of $cif_{wMel[T1]}$ transgenes (Shropshire and Bordenstein, 2019).

Do phylogenetic Type 1 cif *genes from* w*Rec transgenically induce and rescue CI?*

*w*Rec naturally occurs in *D. recens* and is a strong CI inducer that causes near-complete embryonic death (Shoemaker et al., 1999; Werren and Jaenike, 1995). Genomic sequencing of the *w*Rec *Wolbachia* strain revealed it harbors a highly reduced prophage WO genome with retainment of approximately a quarter of its genes (Metcalf et al., 2014). This reduced phage WO genome helped narrow the list of candidate CI genes in an unbiased, multi-omics analysis of genes in *w*Mel *Wolbachia* (LePage et al., 2017). Relative to $Cif_{wMel[T1]}$ genes, $CifA_{wRec[T1]}$ has two amino acid substitutions in unannotated regions: one prior to CifA's putative DUF3243 and another after the annotated STE domain (Figure V-2A). $CifB_{wRec[T1]}$ has 13 amino acid changes that include a seven amino acid extension on the N-terminus, four substitutions in the N-terminal unannotated region, a single substitution in the first putative PD-(D/E)XK-like nuclease domain (hereafter PDDEXK), and a stop codon that truncates 1032-1173 amino acids on the C-terminus of the protein (Figure V-2A). All annotations are taken from a prior structural homology search using HHpred (Lindsey et al., 2018). Since *w*Rec causes strong CI in *D. recens* (Shoemaker et al., 1999; Werren and Jaenike, 1995), the *w*Rec genome lacks other *cif* genes (Metcalf et al., 2014), and these variants

are highly similar to *cif*$_{wMel[T1]}$ genes (Figure V-2A), we predicted that transgenic *cifA;B*$_{wRec[T1]}$ expression in uninfected males causes CI, transgenic *cifA*$_{wRec[T1]}$ expression in uninfected females rescues that CI, and that CI induced by *cif*$_{wRec[T1]}$ transgenes are compatible with *cif*$_{wMel[T1]}$ transgenes.

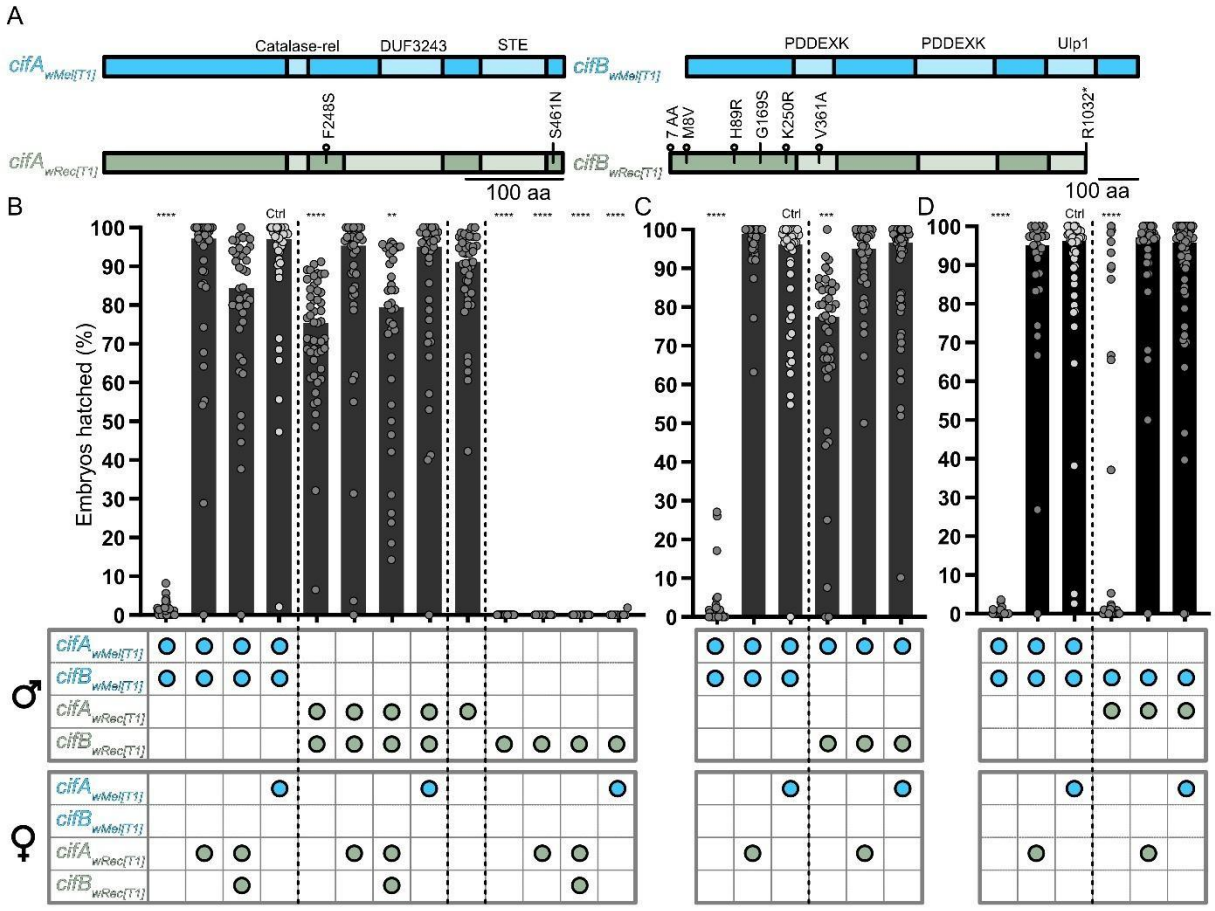

Figure V-2. Schematics and functions of Cif$_{wRec[T1]}$ proteins transgenically expressed in *D. melanogaster*.
(A) Protein architecture of Cif$_{wMel[T1]}$ and Cif$_{wRec[T1]}$ (Lindsey et al., 2018). Substitutions inside schematics represent sequence identity relative to the Cif$_{wMel[T1]}$ reference protein based on a pairwise MUSCLE alignment. Substitutions marked with a circle above the protein schema are shared between Cif$_{wRec[T1]}$ and Cif$_{wRi[T1]}$ proteins. Hatch rate analyses testing (B) *cifA*$_{wRec[T1]}$, *cifB*$_{wRec[T1]}$, and *cifA;B*$_{wRec[T1]}$ for CI and rescue (N = 12-51 where each dot represents a clutch of embryos from a single mating pair), (C) *cifA*$_{wMel[T1]}$*;cifB*$_{wRec[T1]}$ for CI (N = 36-55), and (D) *cifA*$_{wRec[T1]}$*;cifB*$_{wMel[T1]}$ for CI (N = 27-58). Horizontal bars represent median embryonic hatching from single pair matings. Genotypes for each cross are illustrated below the bars where the genes expressed in each sex are represented by colored circles. Blue circles represent *cif*$_{wMel[T1]}$ genes and green circles represent *cif*$_{wRec[T1]}$ genes. All flies were uninfected with *Wolbachia*. Each hatch rate contains the combined data of two replicate experiments, each containing all crosses shown. Asterisks above bars represent significant differences relative to a control transgenic rescue cross (denoted Ctrl) with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Replicate data was statistically comparable in all cases.

Consistent with prior reports in *D. melanogaster*, *cifA;B*$_{wMel[T1]}$ expression in uninfected males induces strong CI and *cifA*$_{wMel[T1]}$ uninfected females rescue that CI (Figure V-2B) (Shropshire and Bordenstein, 2019). Dual cognate expression of *cifA;B*$_{wRec[T1]}$ in uninfected males also reveals a small but statistically significant reduction in hatching (Mdn = 75.4% hatching; $p < 0.0001$; Figure V-2B), and *cifA*$_{wRec[T1]}$ females rescue this hatch rate reduction in a manner similar to the compatible control cross ($p > 0.99$), suggesting that *cifA*$_{wRec[T1]}$ is a rescue gene, and weak *cifA;B*$_{wRec[T1]}$-induced CI is rescuable (Figure V-2B). Rescue did not occur when dual *cifA;B*$_{wRec[T1]}$ males were crossed with dual *cifA;B*$_{wRec[T1]}$ females (Mdn = 79.6% hatching; $p = 0.0054$), signifying that *cifB*$_{wRec[T1]}$ may reduce *cifA*$_{wRec[T1]}$ rescue capacity in embryos, perhaps owing to expression in the foreign *D. melanogaster* background. These data support the Two-by-One genetic model of CI for *w*Rec CI, namely that *cifA;B*$_{wRec[T1]}$ induces weak CI, and *cifA*$_{wRec[T1]}$ alone rescues CI in *D. melanogaster*.

We also tested if singly expressing *cifA*$_{wRec[T1]}$ or *cifB*$_{wRec[T1]}$ causes CI. Since our previous work demonstrated that *cifA*$_{wMel[T1]}$ and *cifB*$_{wMel[T1]}$ do not induce CI alone, we predicted that neither *cifA*$_{wRec[T1]}$ nor *cifB*$_{wRec[T1]}$ would reduce hatching. Indeed, *cifA*$_{wRec[T1]}$ males did not statistically reduce hatching ($p > 0.99$). However, *cifB*$_{wRec[T1]}$ males caused near-complete lethality of embryos (Mdn = 0% hatching; $p < 0.0001$), but it was not rescuable by crossing to *cifA*$_{wRec[T1]}$ (Mdn = 0% hatching), *cifA;B*$_{wRec[T1]}$ (Mdn = 0% hatching), *cifA*$_{wMel[T1]}$ (Mdn = 0% hatching), or *w*Mel-infected (Mdn = 0% hatching) females (Figure V-2B; Figure D-1). We interpret the *cifB*$_{wRec[T1]}$-induced embryonic death to be a likely sterility artifact and not bona fide CI due to the lack of rescue (Figure V-2B). Next, we aimed to determine if the *cifB*$_{wRec[T1]}$ artifact was associated with cytological defects in early embryos (LePage et al., 2017) using a propidium iodide nucleotide stain on RNAse treated 1-2hr old embryos. As anticipated, control *cifA;B*$_{wMel[T1]}$ males induce high levels of embryonic defects when mated to uninfected females but few defects when mated to *cifA*$_{wMel[T1]}$ or *cifA;B*$_{wRec[T1]}$ females owing to rescue (Figure D-2). Consistent with a sterility artifact, single *cifB*$_{wRec[T1]}$ males had a high percentage of early mitotic failures and single puncta indicative of unfertilized embryos or embryos undergoing mitotic failure in the first division that are not rescuable (Figure D-2).

Next, we tested the hypothesis that transgenic *cifA;B*$_{wMel[T1]}$ CI is rescuable by *cifA*$_{wRec[T1]}$ and *vice versa*. When *cifA;B*$_{wMel[T1]}$ males mate *cifA*$_{wRec[T1]}$ ($p > 0.99$) or *cifA;B*$_{wRec[T1]}$ ($p = 0.10$) females, hatching was statistically similar to the compatible control (Fig. 2B). These data indicate

that ovarian expression of *cifA*$_{wRec[T1]}$ can rescue *cifA;B*$_{wMel[T1]}$ CI, and thus the two amino acid changes in *cifA*$_{wRec[T1]}$ are not sufficient to alter rescue between these phylogenetic Type 1 variants. Similarly, hatching levels were relatively high when *cifA;B*$_{wMel[T1]}$ males were crossed to *cifA;B*$_{wRec[T1]}$ females (Mdn = 84.4% hatching), suggesting that *cifA;B*$_{wRec[T1]}$ females may partially rescue *cifA;B*$_{wMel[T1]}$ CI relative to *cifA* females (Figure V-2B). However and curiously, *cifA;B*$_{wRec[T1]}$ females appear unable to rescue CI induced by *cifA;B*$_{wRec[T1]}$ males relative to *cifA*$_{wRec[T1]}$ females (Figure V-2B). Thus, a firm conclusion cannot be made on whether or not *cifA;B*$_{wRec[T1]}$ females can rescue *cifA;B*$_{wMel[T1]}$ CI. Notably, full rescue occurs when *cifA;B*$_{wRec[T1]}$ males are mated to both *cifA*$_{wRec[T1]}$ and *cifA*$_{wMel[T1]}$ females ($p > 0.99$) relative to the compatible control, suggesting that *cifB*$_{wRec[T1]}$ hinders the ability of *cifA*$_{wRec[T1]}$ to rescue. Together, these data indicate that *cifA*$_{wMel[T1]}$ and *cifA*$_{wRec[T1]}$ rescue the other strain's transgenic CI, but females expressing *cifB*$_{wRec[T1]}$ along with *cifA*$_{wRec[T1]}$ hamper *cifA*-mediated rescue and/or artifactually alters the fertility of expressing mothers or viability of resulting embryos, thus causing a reduction in hatching.

Finally, since *cifA;B*$_{wRec[T1]}$ males induce weak CI relative to *cifA;B*$_{wMel[T1]}$ males, we hypothesized that sequence variation in either *cifA*$_{wRec[T1]}$ or *cifB*$_{wRec[T1]}$ underpins that variability. We tested this hypothesis by engineering and dual expressing combinations of *cif*$_{wRec[T1]}$ genes with *cif*$_{wMel[T1]}$ genes. When *cifA*$_{wMel[T1]}$;*cifB*$_{wRec[T1]}$ males mate with uninfected females, we observe a weak but statistically significant reduction in hatching relative to the compatible control (Mdn = 77.6% hatching; $p = 0.0008$; Figure V-2C). Notably, this hatch rate reduction was comparable to that of cognate *cifA;B*$_{wRec[T1]}$ (Mdn = 75.4% hatching; Figure V-2B), and it was likewise rescued when expressing males were crossed with *cifA*$_{wMel[T1]}$ ($p > 0.99$) or *cifA*$_{wRec[T1]}$ ($p > 0.99$) females (Figure V-2C). However, *cifA*$_{wRec[T1]}$;*cifB*$_{wMel[T1]}$ males caused strong transgenic CI (Mdn = 0% hatching; $p < 0.0001$) that was rescued by *cifA*$_{wRec[T1]}$ (Mdn = 97.1% hatching; $p > 0.99$) or *cifA*$_{wMel[T1]}$ (Mdn = 95.9% hatching; $p > 0.99$) females (Figure V-2D). Together, these data demonstrate that non-cognate pairings between closely-related *cif*$_{wMel[T1]}$ and *cif*$_{wRec[T1]}$ are interchangeable. They cause rescuable CI, and sequence variation in *cifB* determines CI level variability when *cif*$_{wRec[T1]}$ transgenes are expressed in *D. melanogaster*.

Do phylogenetic Type 1 cif *genes from* w*Ri transgenically induce and rescue CI?*

*w*Ri naturally infects *D. simulans* and is a strong CI inducer that can cause near complete embryonic death in the lab and ~45% CI in the field (Carrington et al., 2011; Mouton et al., 2006; Turelli and Hoffmann, 1995). On multiple occasions, *w*Ri-like *Wolbachia* have been observed to make rapid sweeps through *D. simulans* populations (Turelli et al., 2018; Turelli and Hoffmann, 1991). Genomic and phylogenetic analyses of *w*Ri reveal two distinct *cif* gene pairs in the Type 1 and Type 2 clades (LePage et al., 2017; Lindsey et al., 2018), and crossing experiments revealed unidirectional CI is caused between *w*Mel- and *w*Ri-infected *D. simulans* flies (Poinsot et al., 1998). Here, we focus our attention on testing the $cif_{wRi[T1]}$ genes for CI and rescue and (in)compatibility between the $cif_{wMel[T1]}$ and $cif_{wRi[T1]}$ gene variants.

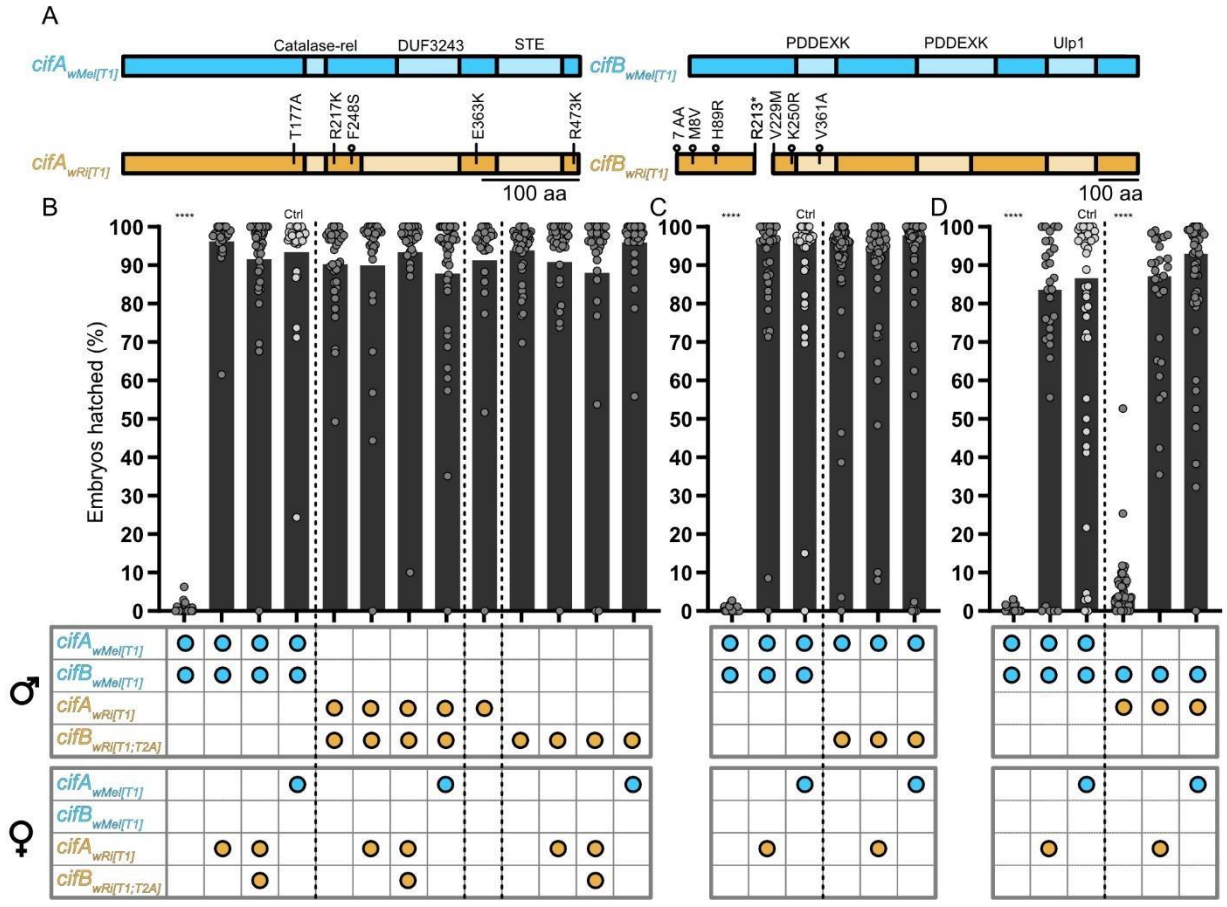

Figure V-3. Schematics and functions of $cif_{wRi[T1]}$ proteins transgenically expressed in *D. melanogaster*.
(A) Protein architecture of $Cif_{wMel[T1]}$ and $Cif_{wRi[T1]}$ (Lindsey et al., 2018). Substitutions inside schematics represent sequence identity relative to the $Cif_{wMel[T1]}$ reference protein based on a pairwise MUSCLE alignment. Substitutions marked with a circle above the protein schema are shared between $Cif_{wRec[T1]}$ and $Cif_{wRi[T1]}$ proteins. Hatch rate analyses testing (B) $cifA_{wRi[T1]}$, $cifB_{wRi[T1;T2A]}$, and $cifA;B_{wRi[T1;T2A]}$ for CI and rescue (N = 26-44 where each dot represents a

clutch of embryos from a single mating pair), (C) *cifA*$_{wMel[T1]}$*;cifB*$_{wRi[T1;T2A]}$ for CI (N = 32-56), and (D) *cifA*$_{wRi[T1]}$*;cifB*$_{wMel[T1]}$ for CI (N = 27-47). Horizontal bars represent median embryonic hatching from single pair matings. Genotypes for each cross are illustrated below the bars where the genes expressed in each sex are represented by colored circles. Blue circles represent *cif*$_{wMel[T1]}$ genes and orange circles represent *cif*$_{wRi[T1]}$ genes. All flies were uninfected with *Wolbachia*. Each hatch rate contains the combined data of two replicate experiments, each containing all crosses shown. Asterisks above bars represent significant differences relative to a control transgenic rescue cross (denoted Ctrl) with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Replicate data was statistically comparable in all cases.

Relative to CifA$_{w\text{Mel}[T1]}$, CifA$_{w\text{Ri}[T1]}$ protein sequence length is identical and has five amino acid substitutions in unannotated regions (Figure V-3A). One of these CifA amino acid substitutions is also present in CifA$_{w\text{Rec}[T1]}$. Relative to CifB$_{w\text{Mel}[T1]}$, CifB$_{w\text{Ri}[T1]}$ has an in-frame stop codon introduced at residue 213 in the 1173 amino acid long protein (Figure V-3A). Glimmer 3 predicts another protein begins with a valine start codon upstream of this stop codon at residue 229. Thus, we predicted that *cifB*$_{wRi[T1]}$ may yield two protein products: an N-terminal 212 amino acid protein and a C-terminal 945 amino acid protein. We refer to the gene sequence yielding the N-terminal peptide as *cifB*$_{wRi[T1;N]}$ and the gene sequence yielding the C-terminal peptide as *cifB*$_{wRi[T1;C]}$. CifB$_{w\text{Ri}[T1;N]}$ has two amino acid substitutions, a seven amino acid N-terminal extension, and an early stop codon relative to CifB$_{w\text{Mel}[T1]}$. In this region, CifB$_{w\text{Rec}[T1]}$ has the same sequence variations, excluding the early stop codon in addition to one extra substitution. CifB$_{w\text{Ri}[T1;C]}$ has three substitutions relative to CifB$_{w\text{Mel}[T1]}$, one of which is in the first PDDEXK domain (Figure V-3A). In this C-terminal region, CifB$_{w\text{Rec}[T1]}$ shares two of these substitutions. There are 15 amino acids in the gap between the N-terminal stop codon and the C-terminal start codon. We predicted that *cifA;B*$_{wRi[T1]}$ expressing *D. melanogaster* males would cause CI and that *cifA*$_{wRi[T1]}$ females would rescue that CI. In addition to testing transgene-induced CI of *cifB*$_{wRi[T1;N]}$ and *cifB*$_{wRi[T1;C]}$ singly or dually with *cifA* genes, we also generated a polycistronic *cifB*$_{wRi[T1]}$ transgene that expressed both the N-terminal and C-terminal peptides from a single transcript using a T2A sequence between the two proteins (Donnelly et al., 2001b, 2001a). We refer to this polycistronic transgenic construct as *cifB*$_{wRi[T1;T2A]}$.

Again, control *cifA;B*$_{wMel[T1]}$ males induce strong CI that is rescued by *cifA*$_{wMel[T1]}$ females (Figure V-3B). Since CI manifests under cognate male expression of both *cif* genes, we first tested *cifA;B*$_{wRi[T1;T2A]}$ males for their ability to induce CI. However, *cifA;B*$_{wRi[T1;T2A]}$ males did not significantly reduce hatching ($p = 0.55$) (Figure V-3B). Males dually expressing *cifA*$_{wRi[T1]}$ with either *cifB*$_{wRi[T1;N]}$ ($p = 0.55$; Figure D-3A) or *cifB*$_{wRi[T1;C]}$ ($p = 0.32$; Figure D-4A) also failed to reduce hatching, suggesting that dual expression of *cif*$_{wRi[T1]}$ transgenes cannot recapitulate CI-like

hatching. Additionally, we tested if singly expressing *cifA*$_{wRi[T1]}$ ($p > 0.99$) or *cifB*$_{wRi[T1;T2A]}$ ($p > 0.99$) males cause CI when crossed to uninfected females (Figure V-3B). Neither transgene alone induced significant reductions in embryonic hatching (Figure V-3B), consistent with results for singly expressing *cif*$_{wMel}$ genes (LePage et al., 2017; Shropshire and Bordenstein, 2019). Next, to test if *cif*$_{wRi[T1]}$ genes can rescue *cif*$_{wMel[T1]}$ transgenic CI, we crossed *cifA;B*$_{wMel[T1]}$ males with *cifA*$_{wRi[T1]}$ ($p > 0.99$) and *cifA;B*$_{wRi[T1;T2A]}$ ($p > 0.99$) females, both of which yielded hatching levels comparable to *cifA*$_{wMel[T1]}$ rescue (Figure V-3B). Taken together, these results indicate that *cifA*$_{wRi[T1]}$ is a rescue gene and *cif*$_{wRi[T1]}$ transgenes do not cause CI when singly or dually expressed in *D. melanogaster*. We discuss the implications of these results below with regards to host background effects on CI, the *cif* genotype – CI phenotypic relationship, and technical artifacts of polycistronic transgene expression.

To further evaluate if *cif*$_{wRi[T1]}$ transgenes are capable of CI-induction and whether variation in *cifA* or *cifB* may underpin the lack of CI above, we engineered and dually expressed non-cognate pairs of *cif*$_{wRi[T1]}$ genes with *cif*$_{wMel[T1]}$ genes. *cifA*$_{wMel[T1]}$*;cifB*$_{wRi[T1;T2A]}$ males did not yield a reduction in hatching compared to the compatible cross ($p > 0.99$; Figure V-3C). Similarly, dual males with *cifA*$_{wMel[T1]}$ and either *cifB*$_{wRi[T1;N]}$ ($p > 0.99$; Figure D-3B) or *cifB*$_{wRi[T1;C]}$ ($p > 0.99$; Figure D-4B) do not reduce hatching. However, *cifA*$_{wRi[T1];}$*cifB*$_{wMel[T1]}$ males caused near-complete embryonic death (Mdn = 0% hatching; $p < 0.0001$) that could be rescued by *cifA*$_{wRi[T1]}$ and *cifA*$_{wMel[T1]}$ females; (Figure V-3D). Taken together with the rescue results above, the findings support the conclusion that *cifA*$_{wRi[T1]}$ is functional and contributes to both rescue and CI-induction in *D. melanogaster*, but *cifB*$_{wRi[T1]}$ transgenes also fail to cause CI-induction in *D. melanogaster*. The non-cognate expression experiments here mirror the results with *cif*$_{wRec[T1]}$. Namely, *cifA* homologs with two-to-five amino acid changes are interchangeable, whereas sequence variation in *cifB* encompassing six amino acid changes across the protein, an N-terminal seven amino acid extension, a C-terminal truncation, and an in-frame stop codon in *cifB* inhibits CI inducibility or causes artifacts in *D. melanogaster*. We discuss the centrality of *cifA* to strong CI and rescue below.

Do the phylogenetic Type 2 cif *genes from* w*Ri transgenically induce and rescue CI?*

As noted, *w*Ri has two *cif* genes in the Type 1 and Type 2 *cif* clades. Pairwise MUSCLE alignments of Cif$_{wMel[T1]}$ and Cif$_{wRi[T2]}$ proteins (488 and 1239 amino acids for CifA and CifB respectively) reveal significant divergence between the proteins. First, CifA$_{wMel[T1]}$ and CifA$_{wRi[T2]}$

differ by 267 sites (45.3% identical sites), with 221 amino acid substitutions and 46 gap sites in the alignment (Figure V-4A; Figure D-5). CifA$_{wRi[T2]}$ also has substitutions in all six of the sites that vary in CifA$_{wRec[T1]}$ and CifA$_{wRi[T1]}$, and two of the CifA$_{wRi[T2]}$ substitutions are shared with both proteins and a third is shared with CifA$_{wRi[T1]}$. Second, CifB$_{wMel[T1]}$ and CifB$_{wRi[T2]}$ differ by 991 sites (20% identical sites), with 433 substitutions and 558 gap sites in the alignment (Figure V-4A; Figure D-5). Additionally, *cifB*$_{wRi[T2]}$ has substitutions in four of the six sites that vary in *cifB*$_{wRi[T1]}$ and *cifB*$_{wRec[T1]}$, but the specific amino acids are unique to *cifB*$_{wRi[T2]}$ (Figure D-5). Moreover, while the sequence lengths of the two CifA variants are comparable, CifB$_{wRi[T2]}$ does not have the C-terminal Ulp1 domain that for other distant Type 1 Cif variants acts *in vitro* as a deubiquitinase (Beckmann et al., 2017). It also has an eight amino acid N-terminal extension (Figure D-5), of which four amino acids are shared in the N-terminal extensions of CifB$_{wRec[T1]}$ and CifB$_{wRi[T1]}$. Here, we test if *cif*$_{wRi[T2]}$ transgenes cause and rescue CI in *D. melanogaster* and if they are (in)compatible with flies expressing *cif*$_{wMel[T1]}$ transgenes.

As expected, control *cifA;B*$_{wMel[T1]}$ males induce strong CI that is rescued by *cifA*$_{wMel[T1]}$ in females (Figure V-4B). *cifA;B*$_{wRi[T2]}$ males caused a weak but statistically significant hatch rate reduction (Mdn = 84.4% hatching; p = 0.01; Figure V-4B) that was rescued upon crossing with *cifA*$_{wRi[T2]}$ females (p > 0.99; Figure V-4B), consistent with Two-by-One genetic model of CI. Similar to results with *cifA;B*$_{wRec[T1]}$ females above (Figure V-2B), crossing *cifA;B*$_{wRi[T2]}$ males with *cifA;B*$_{wRi[T2]}$ females only slightly improved hatching such that it was no longer statistically different from the compatible control (Mdn = 86.9% hatching; p = 0.15); however, the median hatch rate was comparable when *cifA;B*$_{wRi[T2]}$ males were mated to uninfected females (Mdn = 84.4% hatching; Figure V-4B). Thus, similar to *cif*$_{wRec[T1]}$, it cannot be concluded that *cifA;B*$_{wRi[T2]}$ females are rescue-capable, but *cifA*$_{wRi[T2]}$ females clearly do rescue *cifA;B*$_{wRi[T2]}$ CI. In parallel, we tested if either *cifA*$_{wRi[T2]}$ (p = 0.84) or *cifB*$_{wRi[T2]}$ (p = 0.13) males alone reduce hatching, and found that neither reduced hatch rates, as expected (Figure V-4B). These data suggest, for the first time, that phylogenetic Type 2 *cif* genes, and *cif*$_{wRi[T2]}$ in particular, can induce weak CI and rescue under a Two-by-One genetic model akin to *cif*$_{wMel[T1]}$ (Shropshire and Bordenstein, 2019) and *cif*$_{wRec[T1]}$ shown above.

Next, we aimed to determine if the divergence between *cifA*$_{wRi[T2]}$ and *cifA*$_{wMel[T1]}$, which are different phylogenetic Types, underpins incompatibility between the strains (Figure V-1D). Embryo death was observed when *cifA;B*$_{wMel[T1]}$ males mated with *cifA*$_{wRi[T2]}$ (Mdn = 0%; p <

0.0001) or *cifA;B*$_{wRi[T2]}$ (Mdn = 0%; p < 0.0001) females (Figure V-4B), suggesting incompatibility between the gene variants. Reciprocally, hatch rates of embryos increased to compatible levels when *cifA;B*$_{wRi[T2]}$ males mated with *cifA*$_{wMel[T1]}$ females (p > 0.99) (Figure V-4B). Together, these data indicate that *cifA;B*$_{wMel[T1]}$-induced CI cannot be rescued by any *w*Ri Type 2 *cif* variants, but that *cifA;B*$_{wRi[T2]}$-induced CI can be rescued by both *cifA*$_{wMel[T1]}$ and *cifA*$_{wRi[T2]}$. Notably, these results contrast with our initial predictions since previously published crossing experiments when *w*Mel *Wolbachia* were transinfected into *D. simulans* determined that *w*Mel cannot rescue *w*Ri-induced CI (Poinsot et al., 1998). We discuss our interpretations of these discrepancies below.

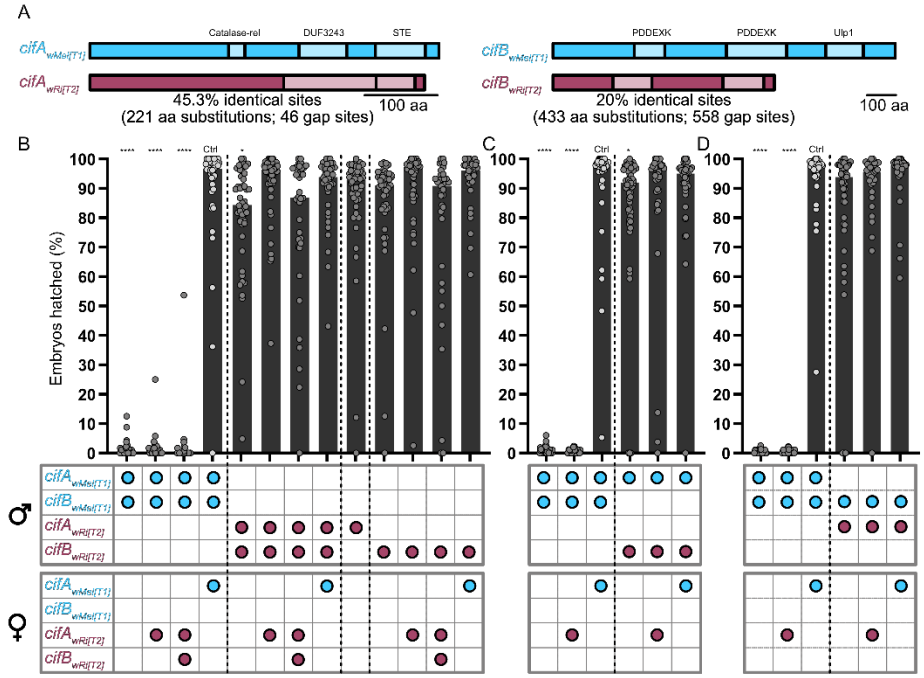

Figure V-4. Schematics and functions of *cif*$_{wRi[T2]}$ proteins transgenically expressed in *D. melanogaster*.
(A) Protein architecture of Cif$_{wMel[T1]}$ and Cif$_{wRi[T2]}$ (Lindsey et al., 2018). In a MUSCLE alignment of CifA$_{wMel[T1]}$ and CifA$_{wRi[T2]}$ (488 aa), there are 221 identical sites, 221 aa substitutions, and 46 gap sites. In an alignment of CifB$_{wMel[T1]}$ and CifB$_{wRi[T2]}$ (1239 aa), there are 248 identical sites, 433 aa substitutions, and 558 gap sites. Specific details on the kinds and locations of sequence variations are illustrated in Fig. S5. Hatch rate analyses testing (B) *cifA*$_{wRi[T2]}$, *cifB*$_{wRi[T2]}$, and *cifA;B*$_{wRi[T2]}$ for CI and rescue (N = 35-55 where each dot represents a clutch of embryos from a single mating pair), (C) *cifA*$_{wMel[T1]}$*;cifB*$_{wRi[T1;T2A]}$ for CI (N = 39-56), and (D) *cifA*$_{wRi[T2]}$*;cifB*$_{wMel[T1]}$ for CI (N = 31-45). Horizontal bars represent median embryonic hatching from single pair matings. Genotypes for each cross are illustrated below the bars where the genes expressed in each sex are represented by colored circles. Blue circles represent *cif*$_{wMel[T1]}$ genes and purple circles represent *cif*$_{wRi[T2]}$ genes. All flies were uninfected with *Wolbachia*. Each hatch rate contains the combined data of two replicate experiments, each containing all crosses shown. Asterisks above bars represent significant differences relative to a control transgenic rescue cross (denoted Ctrl) with an α = 0.05 calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Replicate data was statistically comparable in all cases.

Since dual expression of *cifA;B*$_{wRi[T2]}$ caused weak CI, we investigated if the non-cognate combinatory expression of *cif*$_{wRi[T2]}$ and *cif*$_{wMel[T1]}$ genes could cause CI and underpin variation in CI levels. First, *cifA*$_{wMel[T1]}$;*cifB*$_{wRi[T2]}$ males crossed to uninfected females yields a small but significant hatch rate reduction (Mdn = 92.0% hatching; p = 0.01), relative to the compatible control. Second, upon crossing these males to either *cifA*$_{wRi[T2]}$ (p > 0.99) or *cifA*$_{wMel[T1]}$ (p = 0.40) females, the weak hatch rate reduction was rescued (Figure V-4C), suggesting that *cifA*$_{wMel[T1]}$;*cifB*$_{wRi[T2]}$ induces a weak and rescuable CI phenotype. Third, switching the gene pairs and crossing of *cifA*$_{wRi[T2]}$;*cifB*$_{wMel[T1]}$ males to uninfected females had a similar, but slightly less significant, impact on hatching (p = 0.07) relative to *cifA*$_{wMel[T1]}$;*cifB*$_{wRi[T2]}$ males (Figure V-4D). Thus, dual expression of both non-cognate pairs yields a reduction in hatching. However, strong CI was not caused with either combination, unlike when *cifA*$_{wRec[T1]}$ was dually expressed with *cifB*$_{wMel[T1]}$, above. These data suggest that divergent *cif* variants can work together to cause a weak CI-like phenotype, but neither gene pair can contribute to strong CI.

Discussion

The Two-by-One genetic model indicates that *cifA;B* males cause CI that can be rescued by *cifA* females (Shropshire and Bordenstein, 2019). This model is well established for *w*Mel-induced CI in *D. melanogaster* (LePage et al., 2017; Shropshire et al., 2018; Shropshire and Bordenstein, 2019) and is in-line with current results for *w*Pip-induced CI in *C. pipiens* (Beckmann et al., 2017; Chen et al., 2019). However, the generality of the Two-by-One model across *cif* variants remains unknown. Additionally, the Two-by-One model explains CI between infected and uninfected insects, but the genetic basis of (in)compatibilities between *Wolbachia* strains and CI level variation is not known. Here, we use transgenic tools in *D. melanogaster* to expand upon the generality of the Two-by-One model, test the hypothesis that *cif* sequence variation relates to interstrain (in)compatibilities (Bonneau et al., 2019, 2018; Charlat et al., 2001; LePage et al., 2017; Shropshire et al., 2018), and assess the role of *cif* sequence variation in transgenic CI levels. We report four key findings (Figure V-5): (i) the Two-by-One genetic model applies to CI and rescue caused by the closely-related *cif*$_{wRec[T1]}$ and distantly related *cif*$_{wRi[T2]}$ (Figure V-5A); (ii) Type 1 *cifA* can rescue CI caused by other non-cognate Type 1 *cif* pairs (Figure V-5B); (iii) phylogenetic Type 2 *cifA* cannot rescue *cifA;B*$_{wMel[T1]}$-induced CI, but the inverse can rescue (Figure V-5C); and

(iv) genetic variation in *cifB* contributes to variation in CI levels among Type 1 *cifs*, but all Type 1 *cifA* homologs can contribute to strong CI when paired with $cifB_{wMel[T1]}$ (Figure V-5D). We also report two unanticipated anomalies: $cifB_{wRec[T1]}$ males express significant infertility artifacts, and $cifB_{wRi[T1]}$ does not induce transgenic CI alone or with any *cifA* variant. Below we interpret these findings in the context of genotype-phenotype relationships underpinning CI level variation, (in)compatibility relationships between *Wolbachia* strains, *cif* genotype by host genotype interactions, and CI mechanisms.

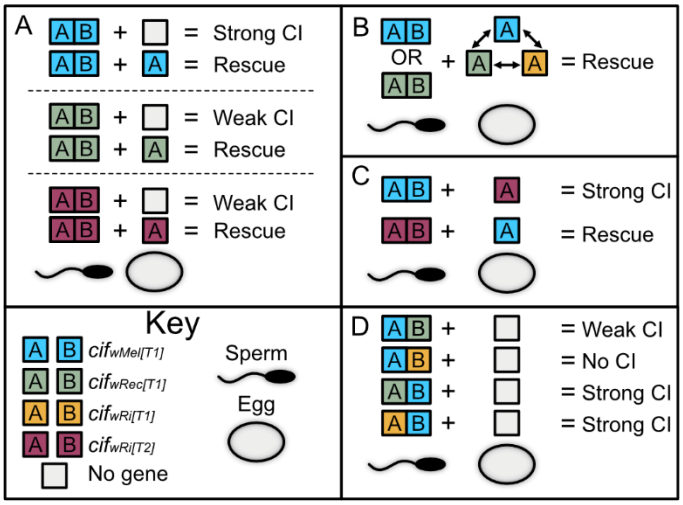

Figure V-5. Summary of key findings.
(A) $cif_{wRec[T1]}$ and $cif_{wRi[T2]}$ induce CI phenotypes in a manner consistent with the Two-by-One genetic model previously established with $cif_{wMel[T1]}$ genes (Shropshire and Bordenstein, 2019). (B) CI induced by Type 1 *cif* pairs can be interchangeably rescued by $cifA_{wMel[T1]}$, $cifA_{wRec[T1]}$, and $cifA_{wRi[T1]}$ transgene expressing females. (C) Unidirectional CI is caused between $cif_{wRi[T2]}$ and $cif_{wMel[T1]}$ transgenes such that $cifA_{wMel[T1]}$ can rescue Type 2 transgenic CI but $cifA_{wRi[T2]}$ fails to rescue Type 1 transgenic CI. (D) Dual non-cognate expression of Type 1 *cif* homologs and $cif_{wMel[T1]}$ transgenes reveal that *cifB* homologs cause weak or no CI while *cifA* homologs can contribute to strong transgenic CI. Non-cognate pairs that cause CI can be rescued by *cifA* expressing females.

CI level variation.

*w*Ri and *w*Rec induce strong CI in their native hosts (Shoemaker et al., 1999; Turelli and Hoffmann, 1991; Werren and Jaenike, 1995), leading to the prediction that their corresponding *cif* genes could yield strong transgenic CI in *D. melanogaster*. However, strong CI was not achieved under dual expression of cognate *cif* homologs. Instead, a small but significant and repeatable CI was observed when $cifA;B_{wRec[T1]}$ and $cifA;B_{wRi[T2]}$ were expressed in uninfected *D. melanogaster* males. CI was rescued when females expressed their cognate *cifA* variant or $cifA_{wMel[T1]}$. Thus, we

conclude these genes induce rescuable CI in a manner consistent with the Two-by-One model (Shropshire and Bordenstein, 2019). Notably, this is the first report of a CI-like phenotype caused by the phylogenetic Type 2 *cif* genes. Unlike *cifA;B*$_{wRec[T1]}$ and *cifA;B*$_{wRi[T2]}$, dual expression of *cifA*$_{wRi[T1]}$ and *cifB*$_{wRi[T1]}$ failed to cause CI. Interestingly, dual expression of the non-cognate *cifA*$_{wRec[T1]}$;*cifB*$_{wMel[T1]}$ in males yielded strong transgenic CI comparable to high levels of CI induced by cognate *cifA;B*$_{wMel[T1]}$, but the converse genotype *cifA*$_{wMel[T1]}$;*cifB*$_{wRec[T1]}$ yielded weak CI levels comparable to *cifA;B*$_{wRec[T1]}$ males. Comparable results were also observed when *cif*$_{wRi[T1]}$ transgenes were expressed with *cif*$_{wMel[T1]}$ transgenes. Thus, these data suggest that variation in *cifB*, not *cifA*, contributes to variation in CI level upon transgenic expression in *D. melanogaster*, and *cifA* from multiple Type 1 strains is capable of causing high CI levels. We propose two non-exclusive hypotheses for why the *cifA* and *cifB* transgenes variably impact CI penetrance in *D. melanogaster* and discuss our interpretations for why *cifA;B*$_{wRi[T1]}$ males fail to cause CI.

First, host genetic background can play a significant role in the level of CI (Bordenstein et al., 2003; Poinsot et al., 1998), and several transinfection and introgression studies lend support to an effect of host genotype on CI levels. For instance, *w*Yak, *w*Tei, and *w*San from the *D. yakuba* clade have been reported to induce no or weak CI in their native hosts but strong CI when transinfected into *D. simulans* (Zabalou et al., 2008), and *w*VitA from *N. vitripennis* wasps cause significantly stronger CI when introgressed into their sister species *N. giraulti* (Chafee et al., 2011). *D. melanogaster* are only known to harbor *w*Mel and *w*Mel-like *Wolbachia* that contain Type 1 *cif* genes (LePage et al., 2017; Lindsey et al., 2018). Thus, it is plausible that non-native CifB (CifB$_{wRec[T1]}$ and CifB$_{wRi[T2]}$) are not able to contribute to strong CI in *D. melanogaster* because divergent CifB may be unable to efficiently interact with *D. melanogaster* targets. Since a relatively small set of sequence changes are present between CifB$_{wMel[T1]}$ and our other Type 1 CifB homologs, and some variations are conserved in the Type 2 CifB, hypotheses can be built about what parts of CifB are related to CI strength variation.

For instance, relative to CifB$_{wMel[T1]}$, CifB$_{wRec[T1]}$ has five amino acid changes, a seven amino acid N-terminal extension, and a stop codon immediately after the Ulp1 domain which prevents translation of 141 C-terminal residues. Conversely, CifB$_{wRi[T2]}$ shares only 20% sequence identity to CifB$_{wMel[T1]}$, has an eight amino acid N-terminal extension, and does not encode a putative Ulp1 domain. The relative impact of each of these changes on CI levels remains unknown, but it is plausible that sequence variation shared between CifB$_{wRec[T1]}$ and CifB$_{wRi[T2]}$ may be

responsible for weak CI when transgenically expressed in *D. melanogaster*. For example, both CifB$_{wRi[T2]}$ and CifB$_{wRec[T1]}$ have N-terminal extensions, and four of the possible eight sites are shared between the two distantly related proteins. Additionally, three of the five sites with substitutions in CifB$_{wRec[T1]}$ relative to CifB$_{wMel[T1]}$, also vary in CifB$_{wRi[T2]}$ but with different amino acids. Alternatively, the weak CI induced by these different Cif Types may have different causes. For instance, structural homology searches of Type 1 and Type 2 CifB reveal that Type 2 proteins have two putative PDDEXK nuclease domains while Type 1 proteins have those two domains and a Ulp1 protease domain known to function as a deubiquitinase *in vitro* (Beckmann et al., 2017; LePage et al., 2017; Lindsey et al., 2018). Thus, functional differences between these proteins caused by the presence or absence of enzymatic domains or any of the hundreds of amino acid substitutions may contribute to host-specific interactions necessary to cause potent CI. These hypotheses can be tested through transgenic expression of *cif* genes in their native host background; however, transgenic CI has not previously been recapitulated in *D. recens* or *D. simulans*, and the expression system will likely require considerable optimization before testing is possible based on experiments in *D. melanogaster* (Shropshire et al., 2019). Additionally, mutagenesis assays will be necessary to determine the relative contributions of each of these sequence variations toward CI level variation in *D. melanogaster*.

Second, we previously posited that the full genetic basis of CI remains unclear until transgenic CI and rescue levels comparable to or higher than wild-type CI can be achieved (LePage et al., 2017; Shropshire and Bordenstein, 2019). In the context of *w*Mel-induced CI, a partial CI phenotype was caused by transgene expressing *cifA;B$_{wMel}$* males (LePage et al., 2017), and we could not conclude that the full microbial genetic basis of *w*Mel-induced CI was solved. Instead, this claim could only be made once the transgenic system was optimized to allow for strong CI and rescue with *w*Mel *cif* gene products alone (Shropshire and Bordenstein, 2019). Here, we tested *cif* homologs using the GAL4-UAS system optimized for transgenic expression of *w*Mel *cif* genes (Shropshire and Bordenstein, 2019). Thus, it is plausible that the level or location of expression optimal for *w*Mel-induced CI is not the same as for these other gene products, and optimization of the transgenic system may be necessary to yield complete phenotypes. Alternatively, other *Wolbachia* or prophage WO genes may be necessary to cause complete CI alongside *cif$_{wRec[T1]}$* and *cif$_{wRi[T2]}$* genes. Notably, this hypothesis is unlikely to apply to Type 1 *cifA* genes since non-cognate expression with *cifB$_{wMel[T1]}$* revealed that Type 1 *cifA* homologs can contribute to strong CI.

However, since *w*Ri contains both Type 1 and Type 2 *cif* genes (LePage et al., 2017), both gene sets may be required for strong CI and rescue. In fact, to date, all *Wolbachia* strains that are known to carry Type 2 *cifs* also harbor genes from other *cif* Types (LePage et al., 2017; Lindsey et al., 2018), suggesting the possibility that the Type 2 sequence variants only contribute to strong CI when coupled with other variants. Alternatively, other genes within the prophage WO's Eukaryotic Association Module are predicted to interact with eukaryotic processes (Bordenstein and Bordenstein, 2016), and *Wolbachia* has many polyvalent proteins including those containing ankyrins predicted to be involved in protein-protein interactions (Wu et al., 2004; Yamada et al., 2011). Any of these proteins could be candidates for modulation of phenotypic potency of CI. These hypotheses can be assessed through optimization of the transgenic expression system in *D. melanogaster* (Duffy, 2002), and co-expression of other genes with transgenic *cifs*.

Finally, $cifA;B_{wRi[T1]}$ males do not cause CI, unlike $cifA;B_{wRec[T1]}$ and $cifA;B_{wRi[T2]}$. Moreover, non-cognate expression of $cif_{wRi[T1]}$ genes with $cif_{wMel[T1]}$ genes revealed that $cifA_{wRi[T1]}$ can contribute to strong CI, but $cifB_{wRi[T1]}$ does not. While the two hypotheses regarding CI level variation above may also explain the lack of CI induced by $cifA;B_{wRi[T1]}$, there is an added layer of complexity. Since $cifB_{wRi[T1]}$ has an early in-frame stop codon relative to $cifB_{wMel[T1]}$, it was annotated in the *w*Ri genome as a pseudogene. We hypothesized CI may be caused when closely-related sequences, such as $cifA_{wRi[T1]}$ or $cifA_{wMel[T1]}$, were dually expressed with $cifB_{wRi[T1]}$'s N-terminal peptide prior to the stop codon, C-terminal peptide after the stop codon, or both. However, none of these combinations could cause CI. These data support the hypothesis that $cifB_{wRi[T1]}$ is a pseudogene, but there are alternative explanations. For instance, *w*Ri naturally harbors both Type 1 and 2 *cif* genes (LePage et al., 2017), and $cifB_{wRi[T1]}$ may contribute to CI when co-expressed with the Type 2 genes. Additionally, the early stop codon in $cifB_{wRi[T1]}$ may not prevent translation of the full-length protein since some stop codons slow translation instead of halting it (Wangen and Green, 2020). Thus, a full-length $CifB_{wRi[T1]}$ protein may be generated despite the introduced stop codon.

To co-express both the N-terminal and C-terminal $CifB_{wRi[T1]}$ proteins, we introduced a T2A sequence between the two proteins which causes translational slippage and multi-protein translation from a single transcript (Donnelly et al., 2001a, 2001b). Relative to the expression of two proteins from independent insertion sites in the *D. melanogaster* genome, this method of polycistronic expression yields an artifactual C-terminal sequence extension to the first protein

that may alter the function. Indeed, polycistronic transgenic expression has previously been used to attempt to cause transgenic CI but yields different results as compared to the expression system used here. For instance, while transgenic expression of *cifA* and *cifB* from a polycistronic T2A construct can cause significant reductions in hatching and embryonic, cytological defects, the induced lethality cannot be rescued when females expressed transgenic *cifA* or were infected with *w*Mel *Wolbachia* (Beckmann et al., 2017). Moreover, CI levels significantly vary between the dual expression of *cifA* and *cifB* from a polycistronic construct and from two separate transcripts (Chen et al., 2019). Thus, it is clear that there are phenotypic differences between these two means of expressing the CI genes, and it is likely driven by the unnatural sequence variations described above or from differences in relative transcript abundance between the two systems. It is plausible that our use of the T2A construct for dual expression of N- and C-terminal $CifB_{wRi[T1]}$ proteins resulted in similar transgenic artifacts that prevented CI phenotypes. Taken together, the evidence for the pseudogenization of $cifB_{wRi[T1]}$ remains equivocal and will require additional study to confirm. However, the results of this study confirm that, as annotated, neither $cifB_{wRi[T1]}$ protein produced by this gene can contribute to CI-induction in *D. melanogaster*.

(In)compatibility relationships.

Figure 1 summarizes the predicted and known (in)compatibility relationships between *w*Mel, *w*Rec, *w*Ri, and their *cif* variants (Fig. 1). *w*Mel and *w*Ri *Wolbachia* strains are unidirectionally incompatible when *w*Mel *Wolbachia* are transinfected into *D. simulans* (Poinsot et al., 1998) (Fig. 1C). Specifically, *w*Ri rescues *w*Mel CI, but *w*Mel cannot rescue *w*Ri CI. We hypothesized that sequence variation in *cif* genes controls these (in)compatibility relationships (LePage et al., 2017). Since *w*Ri has both Type 1 and 2 *cif* genes, we expected $cifA_{wRi[T1]}$ to rescue $cifA;B_{wMel[T1]}$-induced CI because the *cifA* variants are closely related, and $cifA_{wMel[T1]}$ would not rescue $cifA;B_{wRi[T2]}$-induced CI because $cifA_{wMel[T1]}$ is highly divergent from the Type 2 gene pair (LePage et al., 2017) (Fig. 1D). Additionally, *w*Rec and *w*Mel have only Type 1 genes with a few amino acid changes, leading to the prediction that they are compatible (Fig. 1C, D). We tested three key predictions of this *cif* genotype – CI phenotype hypothesis using transgenics in *D. melanogaster*: (i) Type 1 *cif* homologs rescue transgenic CI by $cif_{wMel[T1]}$, (ii) Type 2 *cif* genes cannot rescue transgenic CI by $cifA_{wMel[T1]}$, and (iii) $cifA_{wMel[T1]}$ cannot rescue Type 2 transgenic CI.

As predicted, *cifA*$_{wRec[T1]}$ and *cifA*$_{wRi[T1]}$ can rescue transgenic *cifA;B*$_{wMel[T1]}$ CI. Additionally, *cifA*$_{wRi[T2]}$ cannot rescue *cifA;B*$_{wMel[T1]}$ CI, despite being able to rescue *cifA;B*$_{wRi[T2]}$ CI. These data align with expectations that only closely-related *cif* homologs are compatible (Fig. 1D). However, we also hypothesized that *cifA*$_{wMel[T1]}$ does not rescue *cifA;B*$_{wRi[T2]}$ CI, but rescue occurred at the same levels for both *cifA*$_{wMel[T1]}$ or *cifA*$_{wRi[T2]}$ females, suggesting that both *cifA* variants were capable of rescuing transgenic *cifA;B*$_{wRi[T2]}$ CI. These results suggest a unidirectional incompatibility between Type 1 and Type 2 genes where Type 1 genes cannot be rescued by Type 2 genes, but the reciprocal cross is compatible. Not only are these results contrary to our expected results, but they also fail to sufficiently explain the unidirectional CI between *w*Mel and *w*Ri since *w*Mel's Type 1 genes were hypothesized to rescue both Type 1 and 2 genes. We propose two possible explanations for these results.

First, our initial hypothesis proposed that both Type 1 and Type 2 Cif protein pairs can cause and rescue CI (LePage et al., 2017), and incompatibilities were driven by one strain lacking the capacity to rescue CI induced by divergent variants. However, it remains possible that there are dynamic interactions between Cifs such that multiple phylogenetic Types interact with one-another to impact the phenotypic output. For instance, since *w*Ri naturally maintains both Type 1 and 2 *cif* genes (LePage et al., 2017; Lindsey et al., 2018), expression of both may be required to induce the published compatibility relationships between *w*Mel and *w*Ri (Poinsot et al., 1998). This hypothesis can be tested through the dual expression of both Type 1 and 2 gene pairs and crossing to *cif*$_{wMel}$ expressing flies.

Second, results in the prior section suggest an effect of host genotype on CI level variation when expressing *cif* homologs, specifically for *cifB*. Two studies have evaluated the CI relationships between *w*Mel and *w*Ri. The first determined that they are unidirectionally incompatible when *w*Mel is transinfected into a *D. simulans* background, such that wRi can rescue wMel-induced CI but wMel cannot rescue wRi-induced CI (Poinsot et al., 1998). The second found no evidence of CI when *w*Mel-infected *D. melanogaster* are bidirectionally crossed with *w*Ri-infected *D. simulans* (Gazla and Carracedo, 2009). Thus, in addition to impacting *cifB* mediated variation in CI level, host genotype may impact (in)compatibility relationships, possibly even switching the direction of compatibilities dependent on host background. It is unknown what kind of (in)compatibility relationships might occur if both *w*Mel and *w*Ri are in a *D. melanogaster* host background. However, our transgenic *cif* expression data indicate *w*Mel can rescue *w*Ri, but not

vice versa. Thus, we hypothesize that rescue, in particular, is impacted by host genotype such that *cifA* expressed natively (e.g., *w*Mel in *D. melanogaster* or *w*Ri in *D. simulans*) have expanded rescue capability as compared to introduced strains. Indeed, another example of a relationship between host genotype (in)compatibility relationships comes from an introgression study where two *Wolbachia* from the *N. longicornis* parasitoid wasp switched from being unidirectionally to bidirectionally incompatible when moved into the same genetic background (Raychoudhury and Werren, 2012). These results yield strong support for host control of *Wolbachia* reproductive parasitism and (in)compatibility relationships. This hypothesis can be tested through transinfection of *w*Ri into a *D. melanogaster* background or through transgenic expression of both $cif_{wMel[T1]}$, $cif_{wRi[T1]}$, and $cif_{wRi[T2]}$ in *D. simulans*.

CI mechanism.

How CifA and CifB mechanistically cause CI and/or rescue at the cellular and molecular levels remains an active area of investigation. To date, *in vitro* assays have determined that $CifB_{wMel[T1]}$ and $CifB_{wPip[T1]}$ act in part as deubiquitinases, $CifB_{wPip[T4]}$ acts in part as a nuclease, CifA and CifB can bind, and both CifA and CifB interact with host proteins when transgenically expressed in *D. melanogaster* (Beckmann et al., 2017, 2019c; Chen et al., 2019). There are two categories of mechanistic models for CI that are currently debated in the literature, termed Host Modification (HM) and Toxin Antidote (TA) (Beckmann et al., 2019a, 2019b; Hurst, 1991; Poinsot et al., 2003; Shropshire et al., 2019). HM models posit that CifA;B proteins cause CI by modifying host factors during spermatogenesis, and those modifications are transferred to the embryo. Rescue occurs when CifA in females reverses those sperm modifications in the embryo (Shropshire et al., 2019, 2018). Conversely, TA models suggest that CifB is transferred to the embryo via the sperm and kills the embryo unless its lethality is rescued through binding to CifA in the embryo (Beckmann et al., 2019a; Shropshire et al., 2019). Notably, there is no evidence of paternal transfer of Cif toxin(s), and there is not currently enough data to support one model over the other (Shropshire et al., 2019). Here, we place three findings from the experiments above into the context of CI's mechanistic basis: (i) CifB sequence variation impacts CI level, (ii) closely-related Type 1 CifA can be interchanged for both CI and rescue, and (iii) $CifB_{wRec[T1]}$ induces a non-CI sterility artifact when singly expressed.

First, a key finding of this study is that CifB sequence variation impacts CI level variation when transgenically expressed in *D. melanogaster*. CifB sequence variation impacts CI level variation under transgenic expression in a foreign host background (as discussed above). Additionally proteomic analyses reveal that CifB binds to at least 48 host proteins when singly expressed, CifA binds to at least 15 proteins when singly expressed, and CifA;B binds at least 60 host proteins (Beckmann et al., 2019c). Thus, it can be hypothesized that CifB homologs are less efficient or unable to bind CI-associated host proteins or to CifA when expressed in a novel host. Indeed, the sheer number of potential CifB binding partners may contribute to the large impact of CifB sequence variation on CI levels. Alternatively, CI levels have been correlated with the number of *Wolbachia*-infected spermatocytes and spermatids during spermatogenesis in wRi-infected *D. simulans* (Clark et al., 2003; Clark and Karr, 2002; Veneti et al., 2003), but even uninfected spermatocytes often result in modified sperm that can cause CI (Riparbelli et al., 2007), suggesting that CifA and/or CifB are diffusible between spermatocytes or during earlier stages of spermatogenesis. In the context of this study, CifB sequence variation may contribute to variation in its tissue localization, subcellular localization, or ability to diffuse between cellular components.

Second, while CifB sequence variation clearly impacts the level of transgenic CI in *D. melanogaster*, Type 1 CifA homologs were remarkably interchangeable and contribute to strong CI and rescue. These data suggest that while CifB sequence variation may be specifically attuned to a distinct host background, transgenic CifA is less subject to variation in host background. For instance, it is plausible that while CifB is interacting with rapidly-evolving host targets in an arms race, CifA interacts with a set of conserved targets. One prediction of this hypothesis would be that CifA would be under purifying selection to retain compatibility with conserved host targets. Indeed, comparative sequence analyses reveal not only that Type 1 CifAs are under strong purifying selection (Shropshire et al., 2018), but that CifA sequence length is highly conserved across the phylogenetic Types (LePage et al., 2017; Lindsey et al., 2018). Thus, a model could be proposed whereby CifB acts as an accessory to unlock CifA's access to conserved host processes not accessible in the absence of CifB.

Finally, $cifB_{wRec[T1]}$ males cause complete embryonic death, but this lethality is not rescuable. As such, $cifB_{wRec[T1]}$-induced embryonic death is not totally consistent with our expectations for CI-induction and may be a sterility artifact that disrupts cell biology in the testes. A prediction of the TA model is that CifB causes CI in the absence of CifA (Beckmann et al.,

2019a). Indeed, singly expressing *cifB*$_{wPip[T1]}$, *cifB*$_{wHa[T1]}$, *cifB*$_{wNo[T3]}$, *cifB*$_{wPip[T4]}$, and *cifB*$_{wStr[T5]}$ in yeast can result in temperature sensitive lethality (Beckmann et al., 2019, 2017; Chen et al., 2019). That lethality is significantly reduced when *cifB*$_{wPip[T1]}$, *cifB*$_{wHa[T1]}$, or *cifB*$_{wPip[T4]}$ are dually expressed with cognate *cifA* in yeast (Beckmann et al., 2019, 2017; Chen et al., 2019). However, when expressed in insects, singly expressing *cifB*$_{wMel[T1]}$, *cifB*$_{wRi[T1]}$, or *cifB*$_{wRi[T2]}$ males do not yield reduced hatching (LePage et al., 2017; Shropshire and Bordenstein, 2019). Aside from singly expressing *cifB*$_{wRec[T1]}$ in this study, only *cifB*$_{wPip[T4]}$ males cause a weak reduction in hatching, but there is no evidence that *cifB*$_{wPip[T4]}$-induced lethality in insects can be rescued (Chen et al., 2019). As such, the most conservative explanation is that CifB alone can cause some embryonic death, but that phenotype is an artifact of transgenic expression. As such, these data do not explicitly support the TA model over the HM model. However, it is also plausible that the inability to rescue *cifB*$_{wRec[T1]}$-induced embryonic death is due to a transgenic artifact akin to that preventing *cifA* or *w*Mel females from rescuing transgenic CI expressed under a polycistronic construct (Beckmann et al., 2017). Together, these works suggest an emerging trend wherein singly expressing some *cifB* variants yields reduced insect hatching, but we urge that the conclusion cannot be drawn that this is CI-associated lethality without evidence of rescue.

Conclusion

Here, we set out to investigate the hypothesis that *cif* sequence variation directly relates to CI phenotypic variation by evaluating cognate combinations of the *cif* genes and evaluating CI. Moreover, we engineered non-cognate gene sets to test CI capacity and links between *cif* sequence variation and variation in CI level. In summary, we determined for the first time that that Type 1 *cif* homologs from *w*Rec and Type 2 cif homologs from *w*Ri cause weak CI when transgenically expressed in *D. melanogaster*, that variation in *cifB* contributes to CI level variability, divergent *cifA* fail to rescue transgenic *cifA;B*$_{wMel[T1]}$ CI, and Type 1 *cifA* homologs are interchangeable for inducing both strong CI and rescue. We have discussed these results in the context of CI level variation, (in)compatibility relationships, and CI mechanism. The work expands upon our understanding of the genetics of CI and (in)compatibilities between *Wolbachia* strains, and they establish novel hypotheses regarding the CI mechanism, CI level variation, and the relationship between CI phenotypes and host genetics.

Materials and methods

Fly lines and maintenance.

The following UAS transgenic constructs were generated for this study: *cifA*$_{wRec[T1]}$, *cifB*$_{wRec[T1]}$, *cifA*$_{wRi[T1]}$, *cifB*$_{wRi[T1;N]}$, *cifB*$_{wRi[T1;C]}$,*cifB*$_{wRi[T1;T2A]}$, *cifA*$_{wRi[T2]}$, and *cifB*$_{wRi[T2]}$. Each gene was codon-optimized for expression in *D. melanogaster* and synthesized by GenScript (Hong Kong, China). Valine start codons were replaced with methionine. Each gene was cloned into the pTIGER plasmid at GenScript. pTIGER is a pUASp-based vector designed for germline expression and was previously used to generate *cifA*$_{wMel[T1]}$ and *cifB*$_{wMel[T1]}$ transgenes (LePage et al., 2017). pTIGER enables PhiC31 integration into the *D. melanogaster* genome, contains a UAS promoter region intended for GAL4/UAS expression, and has a red-eye marker for screening. *D. melanogaster* embryo injections were conducted by Best Gene (Chino Hills, California) using PhiC31 integrase to place *cifA* and *cifB* homologs into the Attp40 and Attp2 insert sites respectively. Transformants were screened via eye color and homozygous transgenic lines were generated for all lines. All lines were negative for *Wolbachia* based on PCR using Wolb_F and Wolb_R3 primers (Casiraghi et al., 2005). Dual expressing UAS transgenic lines were generated via standard genetic crossing schemes.

Additionally, the *D. melanogaster* stocks infected and uninfected y^1w^* (BDSC 1495), uninfected *nos*-GAL4:VP16 (BDSC 4937), and uninfected UAS transgenic lines homozygous for *cifA*$_{wMel[T1]}$, *cifB*$_{wMel[T1]}$, and *cifA;B*$_{wMel[T1]}$ (LePage et al., 2017) were used in this study. Genotypes and infection states were regularly confirmed for transgene expressing fly lines. *D. melanogaster* stocks were maintained at 12:12 light:dark at 25^o C on 50 ml of a standard media. Stocks for virgin collections were stored at 18^o C overnight to slow eclosion rate, and virgin flies were kept at room temperature.

Hatch rate assays.

To test for CI, hatch rate assays were used as previously described (LePage et al., 2017; Shropshire et al., 2018). Briefly, virgin *nos*-GAL4:VP16 adult females were aged 9-11 days post eclosion, to control for the paternal grandmother age effect (Layton et al., 2019), and mated with UAS transgenic males. GAL4-UAS males and females were paired in 8 oz bottles affixed with a

grape-juice agar plate smeared with yeast affixed to the opening with tape. Young early emerging males (0-48 h) were used to control the impact of male age and the younger brother effect on CI level (Reynolds and Hoffmann, 2002; Yamada et al., 2007), and 5-7 day old females were used since they are most fecund. The flies and bottles were stored at 25^{o} C for 24 h at which time the plates were replaced with freshly smeared plates and again stored for 24 h. Plates were then removed and the number of embryos on each plate were counted and stored at 25^{o} C. After 30 h the remaining unhatched embryos were counted. The percent of embryos that hatched into larvae was calculated by dividing the number of hatched embryos by the initial embryo count and multiplying by 100.

Embryonic cytology.

Flies were collected, aged, and crossed as described for hatch rate assays. However, 60 females and 12 males were included in each bottle with a grape-juice agar plate attached. Flies were siblings of those in hatch rate assays. Embryos laid in the first 24 h were discarded due to low egg-laying. During the second day, embryos were aged 1-2 hr and then dechorionated, washed, and fixed in methanol as previously described (LePage et al., 2017; Shropshire et al., 2018). Embryos were stained with propidium iodide and imaged (LePage et al., 2017; Shropshire et al., 2018). Scoring of cytological defects was conducted using previously defined characteristics (LePage et al., 2017).

Sequence analyses.

Sequence similarity between Cif proteins was determined using pairwise MUSCLE alignments of protein sequences using default settings. Glimmer 3 was used to identify open reading frames in $cifB_{wRi[T1]}$ after the early stop codon that truncates the gene. These analyses were conducted in Geneious Prime.

Statistical analyses.

All statistical analyses were conducted in GraphPad Prism 8. Hatch rate statistical comparisons were made using Kruskal-Wallis followed by a Dunn's multiple comparison test. Samples with fewer than 25 embryos laid were removed from hatch rate analyses as previously

described (LePage et al., 2017). Hatch rates in main text figures display the combination of two replicate experiments which were analyzed simultaneously, and those in the supplement display only single experiments (N = 8-58 per cross after exclusion). Cytological abnormalities were compared using a pairwise Fischer's exact test followed by a Bonferroni-Dunn correction test (N = 43-167 embryos per cross). Figure aesthetics were edited in Affinity Designer 1.7 (Serif Europe, Nottingham, UK).

Chapter VI.

Site-directed mutagenesis of *cif* genes reveal conserved sites essential for induction and rescue of cytoplasmic incompatibility

Abstract

Wolbachia are maternally-inherited symbionts that cause cytoplasmic incompatibility (CI) to increase the relative fitness of infected females in arthropod populations. CI results in embryonic death when sperm from infected males fertilize uninfected eggs, but embryonic lethality is rescued upon fertilization of infected eggs. Expression of two phage WO genes in males (*cifA* and *cifB*) causes CI, while expression of one of the same genes in females (*cifA*) rescues CI. Structural homology-based data mining predicts that CifA and CifB proteins may harbor three functional domains each, but their relative importance to CI and rescue is unclear. Here, in the absence of obvious catalytic motifs or binding residues, we use site-directed substitution mutagenesis to determine the functional importance of conserved amino acids across the gene, spanning each predicted domain and in N-terminal unannotated regions of each protein. We report that CifB amino acid mutations across the protein ablate CI without any particular significance to the domains or region of the protein. Interestingly, mutations on the 5' end of the CifA protein in the predicted catalase-rel domain and N-terminal unannotated region inhibit CI and rescue, whereas 3' sites in CifA's putative Puf-family RNA-binding domain ablate rescue. These results emphasize that multiple CifA and CifB regions and residues impact CI and/or rescue, and thus they contribute to resolving a more complete understanding of the genetics and potential mechanisms of CI.

Introduction

Wolbachia are maternally-inherited, intracellular α-Proteobacteria that infect 40-65% of all arthropod species (Charlesworth et al., 2019; Hilgenboecker et al., 2008; Weinert et al., 2015; Zug and Hammerstein, 2012). Residing mainly in the cells of reproductive tissues, *Wolbachia* commonly cause a sperm-egg incompatibility, termed cytoplasmic incompatibility (CI), to increase their frequency in host populations (Hoffmann et al., 1990; Turelli, 1994). CI manifests

as embryonic lethality when *Wolbachia*-infected males mate with uninfected females and is rescued when the female is infected with the same strain of *Wolbachia* (LePage and Bordenstein, 2013; Serbus et al., 2008; Taylor et al., 2018). As such, CI yields a relative advantage to females that transmit *Wolbachia* and assists *Wolbachia*'s spread through the matriline (Turelli, 1994; Turelli et al., 2018b; Turelli and Hoffmann, 1991). This drive system, in addition to *Wolbachia*'s ability to confer resistance to RNA viruses, has brought CI to the forefront of vector control efforts to reduce the spread of arboviral diseases including Zika and dengue (Aliota et al., 2016; Caragata et al., 2016; Kittayapong et al., 2018; Moreira et al., 2009; O'Connor et al., 2012; O'Neill, 2018; X. Zheng et al., 2019).

The Two-by-One genetic model of CI (Shropshire and Bordenstein, 2019) describes the microbial genetic basis of CI. Two genes, *cifA* and *cifB*, cause CI when dually expressed in testes (Beckmann et al., 2017; Chen et al., 2019; LePage et al., 2017; Shropshire and Bordenstein, 2019), and *cifA* rescues CI when singly expressed in ovaries (Chen et al., 2019; Shropshire et al., 2018; Shropshire and Bordenstein, 2019). However, the mechanism underlying CifA;CifB-induced CI and CifA-induced rescue remains mostly unknown. *In vitro* assays support CifB has a Ulp1 protease (hereafter Ulp1) domain that acts on poly-ubiquitin chains, and CifA and CifB bind to one another (Beckmann et al., 2017). These functions have not been validated *in vivo*. Otherwise, structural homology-based annotations of CifA and CifB protein architecture predict that both proteins harbor three functional domains (Lindsey et al., 2018b). We focus on the *w*Mel *Wolbachia* native to *Drosophila melanogaster*, which encodes *cif* genes capable of CI and rescue (LePage et al., 2017; Lindsey et al., 2018b; Shropshire et al., 2018; Shropshire and Bordenstein, 2019) and were transinfected into *Aedes* mosquitos for use in vector control by the World Mosquito Program (Hoffmann et al., 2011; Moreira et al., 2009; O'Neill, 2018).

$CifA_{wMel}$ is weakly predicted to encode a catalase-related domain (catalase-rel), domain of unknown function 3243 (DUF), and a sterile-like transcription factor (STE) (Lindsey et al., 2018b) (Figure VI-1a). Catalase-rel domains are predicted to catalyze the degradation of reactive oxygen species (ROS) (Guy et al., 2005; Loew, 1900). DUF has a distant homology to globin-like domains and Puf-family RNA-binding domains, which influence the stability of eukaryotic RNAs (Kumar and Subramaniam, 2018; Nishanth and Simon, 2020). Finally, STE domains mediate transcriptional induction in yeast (Wong Sak Hoi and Dumas, 2010). Importantly, all CifA annotations are weak predictions and there are no obvious catalytic motifs or binding sites

(Lindsey et al., 2018b). However, structural homology predictions identify the Puf-family RNA binding and STE domains across multiple phylogenetic Cif types (Bing et al., 2020; Lindsey et al., 2018b).

$CifB_{wMel}$ has two putative PD-(D/E)XK (hereafter PDDEXK) nuclease domains, and a Ulp1 domain shown to cleave poly-ubiquitin chains *in vitro* (Beckmann et al., 2017; LePage et al., 2017; Lindsey et al., 2018b) (Figure VI-2a). Among these three domains, the PDDEXK domains are conserved in multiple phylogenetic Cif types based on sequence or structural homology (Bing et al., 2020; LePage et al., 2017; Lindsey et al., 2018b), suggesting that these domains are important and central to CI. Indeed, a divergent variant of $CifB_{wPip}$ harbors two PDDEXK domains, and *in vitro* biochemical assays confirm that both domains nick DNA (Chen et al., 2019; Lindsey et al., 2018b). Moreover, Ulp1 domains are restricted to a single phylogenetic clade of CifB. The function of the PDDEXK dimer of CifB has not been assessed in the clade containing Ulp1, but neither PDDEXK domain contains canonical catalytic motifs and are instead only predicted as PDDEXK domains based on structural similarity (Knizewski et al., 2007; Steczkiewicz et al., 2012).

Here, we ask the question are conserved amino acid residues across the CifA and CifB proteins necessary for CI and/or rescue? To answer this question, we mutagenized a selection of conserved sites across the Cif proteins and transgenically expressed them in *D. melanogaster*. We report three key results: (i) conserved sites in CifA's N-terminal unannotated region and the catalase-rel domain are important in both rescue and CI, (ii) conserved sites in CifA's DUF are only involved in CI, and (iii) all tested conserved sites in CifB are required for CI. Taken together, this study identified sites in seven Cif mutants (both CifA and CifB) essential for CI, and characterized the bipartite nature of CifA in which the N-terminal end inclusive of the predicted catalase-rel is seemingly central to CI and rescue, while the middle region containing the DUF is specialized to CI. These results inform the mechanistic basis for CI and rescue and provide additional support for a Two-by-One genetic model where both CifA and CifB are critical for CI.

Results

CifA mutants impact CI and rescue.

We used a previous sequence analysis of conserved amino acid residues in an alignment of phylogenetically diverse CifA proteins (Lindsey et al., 2018b) to select highly conserved sites

across the protein for mutagenesis. CifA$_{*1}$, CifA$_{*2}$, CifA$_{*3}$, and CifA$_{*4}$ have alanine substitutions in the N-terminal unannotated region and putative catalase-rel, DUF, and STE domains, respectively (Figure VI-1a). Alanine mutagenesis is used to analyze the importance of specific amino acids in protein sequences without contributing significant structural variation to the protein (Cunningham and Wells, 1989). We tested mutant *cifA* transgenes for their ability to (i) induce CI when dually expressed with *cifB* in uninfected males and (ii) rescue when singly expressed in uninfected females. Since CI manifests as embryonic death, we measured the strength of CI induced under mutant transgenic expression by measuring the percentage of *D. melanogaster* embryos that hatch into larvae. Notably all mutants are expressed in the same insertion site within the *D. melanogaster* chromosome and with the *nos*-GAL4:VP16 driver. Since the level and location of transgene expression are determined by these two factors, we anticipate minimal impact of Cif expression variation on the phenotypic results.

We first tested if dual expression of *cifA*$_{\Delta}$*;cifB* in uninfected males could induce an appreciable reduction in hatching relative to *cifA;cifB* expression and as compared to compatible controls (Figure VI-1b). Consistent with prior studies (Shropshire and Bordenstein, 2019), dual *cifA;cifB* expression in males yielded nearly complete embryonic death when mated to uninfected females (Mdn = 0% hatching). Dual *cifA;cifB*-induced CI was statistically comparable to *w*Mel-induced CI ($p > 0.99$) and could be rescued when females were *w*Mel-infected (Mdn = 95.4% hatching). However, *cifA*$_{*1}$*;cifB* (Mdn = 93.7% hatching; $p > 0.99$) and *cifA*$_{*3}$*;cifB* (Mdn = 94.1% hatching; $p > 0.99$) males caused no significant hatch rate reduction relative to the rescue cross, suggesting that mutating conserved sites in CifA's unannotated region and putative DUF ablates CI. Conversely, transgenic expression of *cifA*$_{*4}$*;cifB* caused hatch rates statistically comparable to *cifA;cifB*-induced CI (Mdn = 0% hatching; $p > 0.99$), suggesting that mutation of conserved sites in the putative STE did not impact *cifA*'s ability to contribute to CI. Finally, transgenic expression of *cifA*$_{*2}$*;cifB* (Mdn=66.0%) yielded an intermediate phenotype where it was statistically different from both *cifA;cifB*-induced CI ($p = 0.0006$) and rescue of transgenic CI ($p = 0.0001$), indicating that the putative catalase-rel mutant induces a partial CI phenotype. Together, these results suggest the mutated sites in the unannotated region, catalase-rel, and DUF of CifA are important for CI-induction (Figure VI-1b). Intriguingly, mutations within the catalase-rel domain yielded only a partial loss in the phenotype, suggesting that function associated with this region has impacted CifA's efficiency.

Next, we tested if uninfected transgenic females expressing *cifA* mutants can rescue *cifA;cifB*-induced CI (Figure VI-1c). As above, dual *cifA;cifB* expressing males induced near-complete embryonic death consistent with strong CI (Mdn = 0% hatching), and this lethality could be rescued when the female expressed *cifA* (Mdn = 85.9% hatching). Transgenic expression of $cifA_{*1}$ (Mdn = 0.00%; $p < 0.0001$) and $cifA_{*2}$ (Mdn=27.6%; $p = 0.0390$) failed to rescue *cifA;cifB*-induced CI as compared to the standard transgenic rescue cross. Conversely, transgenic expression of $cifA_{*3}$ (Mdn=91.2%; $P > 0.9999$) and $cifA_{*4}$ (Mdn=97.6%; $P = 0.3039$) rescued *cifA;cifB*-induced CI at levels comparable to the standard transgenic rescue cross. These results suggest that the sites mutated in the unannotated and catalase-rel regions of CifA are important for rescue (Figure VI-1c).

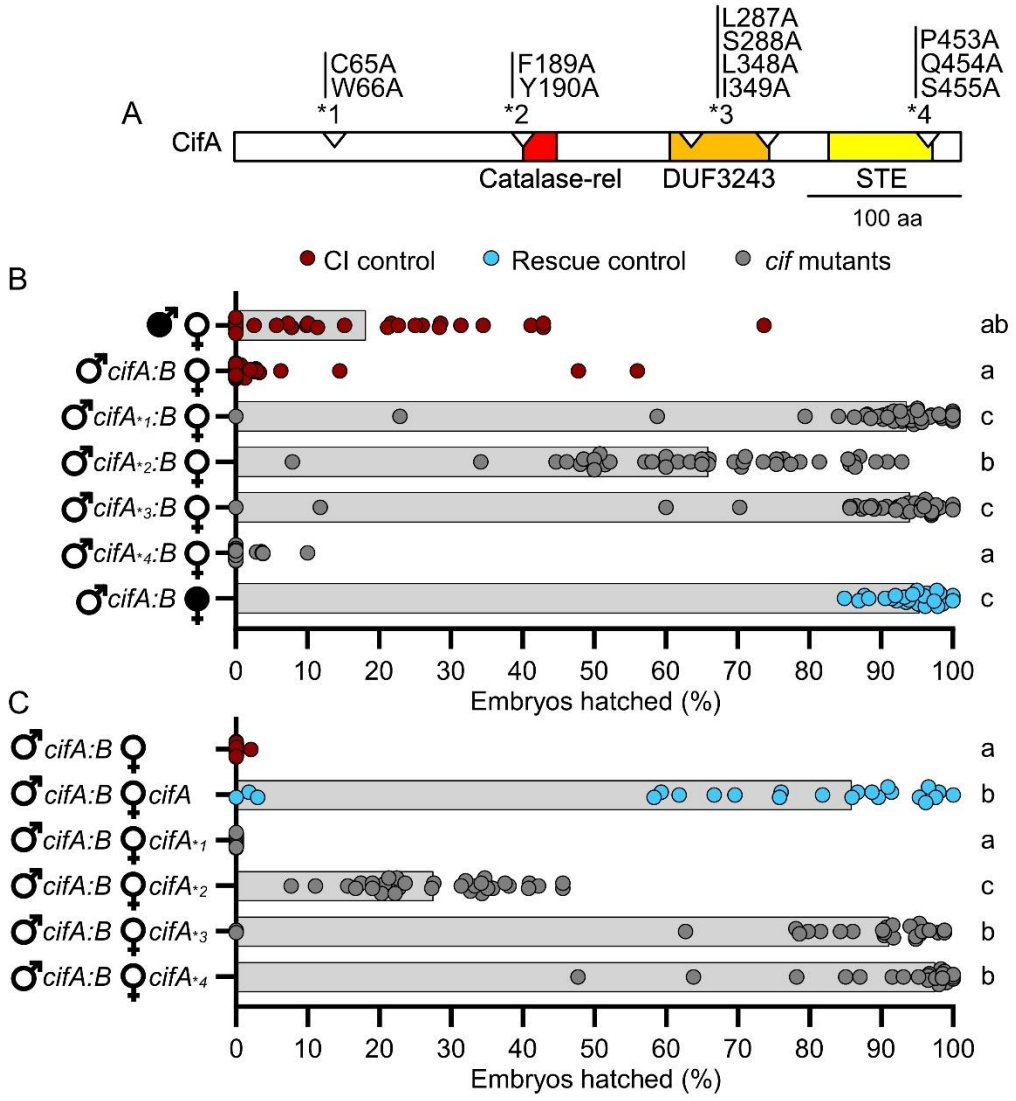

Figure VI-1. *cifA*$_{*1}$ and *cifA*$_{*2}$ mutants fail to cause or rescue CI, and *cifA*$_{*3}$ can rescue but fails to cause CI. (A) schematic showing the location of mutations in CifA relative to previously-predicted domains (Lindsey et al., 2018b). (B) Hatch rate experiment testing if cifA mutants can induce CI when dual expressed with cifB in uninfected males. (C) Hatch rate experiment testing if expressing cifA mutants can rescue transgenic CI when expressed in uninfected females. (B/C) Each dot represents the percent of embryos that hatched from a single male and female pair. Expressed genes are noted to the right of the corresponding sex. Gray bars represent median hatch rates for each cross and letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by Kruskal-Wallis and Dunn's test for multiple comparisons between all groups. All p-values are reported in Table E-1.

CifB mutants ablate CI.

Four CifB mutants were constructed based on the comparative sequence analysis of conserved residues (Lindsey et al., 2018b). All CifB mutations are similarly alanine substitutions, with the exception of one glycine mutation of a conserved alanine (Figure VI-2a). Glycine was chosen to replace alanine since it is comparably sized and would be less likely to impact protein structure than other amino acids. CifB$_{*1}$, CifB$_{*2}$, CifB$_{*3}$, and CifB$_{*4}$ have mutations in the N-terminal unannotated region, first PDDEXK, second PDDEXK, and Ulp1 respectively (Figure VI-2a). The Ulp1 mutation is the same used previously to test for the catalytic activity of the Ulp1 domain (Beckmann et al., 2017). As with *cifA* mutants above, we tested mutant *cifB* transgenes for their ability to induce CI when dual expressed with *cifA* in uninfected males.

As expected, dual *cifA;cifB* expression in uninfected males caused hatch rates statistically comparable to *w*Mel-induced CI ($p > 0.99$) and it could be rescued by *w*Mel-infected females (Mdn = 93.9% hatching). However, transgenic expression of *cifA;cifB*$_{*1}$ (Mdn = 96.3%; $p < 0.0001$), *cifA;cifB*$_{*2}$ (Mdn = 95.6%; $p < 0.0001$), *cifA;cifB*$_{*3}$ (Mdn = 94.3%; $p < 0.0001$), and *cifA;cifB*$_{*4}$ (Mdn = 93.0%; $p < 0.0001$) all failed to reduce hatch rates statistically comparable to *cifA;cifB*-induced CI (Mdn = 0.%). These results suggest that all mutated conserved sites are important for CifB in CI-induction, and validate prior reports that mutating the catalytic site of Ulp1 ablates CI-induction (Beckmann et al., 2017).

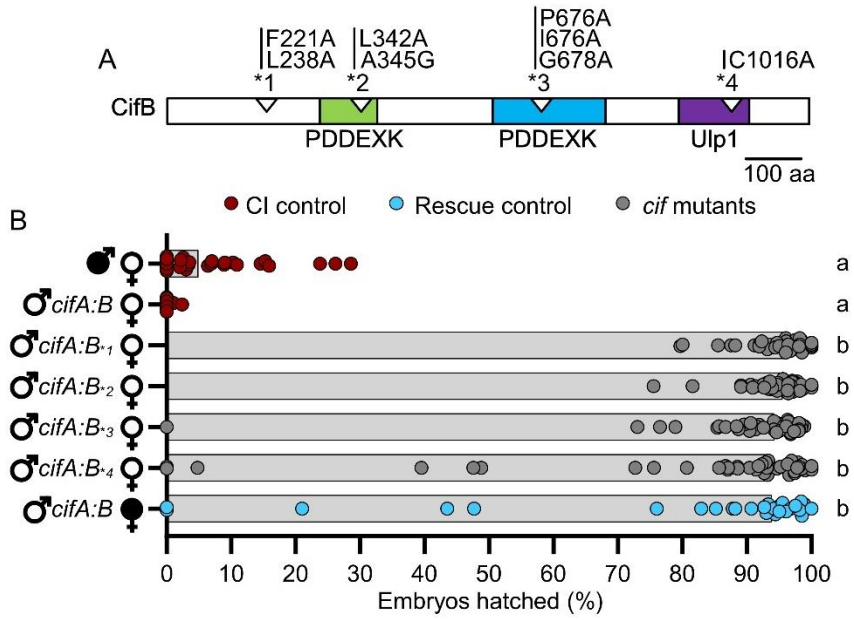

Figure VI-2. All *cifB* mutants fail to contribute to CI.
(A) schematic showing the location of mutations in CifB relative to previously predicted domains (LePage et al., 2017; Lindsey et al., 2018b). (B) Hatch rate experiment testing if *cifB* mutants can induce CI when dual expressed with *cifA* in uninfected males. Each dot represents the percent of embryos that hatched from a single male and female pair. Expressed genes are noted to the right of the corresponding sex. Gray bars represent median hatch rates for each cross and letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by Kruskal-Wallis and Dunn's test for multiple comparisons between all groups. All p-values are reported in Table E-1.

Cif structural predictions

There are numerous ways to interpret the impact of a mutation on a protein's function. These can include changes to catalytic motifs, ligand binding sites, or changes in local or global structure that ablate, enhance, or otherwise modify the phenotypic output of the protein. Aside from *cifB*$_{*4}$'s mutation of the catalytic site of the Ulp1 domain (Beckmann et al., 2017), no other catalytic motifs or binding sites are known. As such, we aimed to investigate the impact of these mutations on the structure of CifA and CifB proteins.

The Iterative Threading ASSembly Refinement (I-TASSER) webserver was used to generate a list of structural homologs from the protein databank (PDB) for each wildtype and mutant Cif protein and construct structural models based on these hits (Zhang, 2009). The shared and unique PDB hits for wild-type and mutant proteins are summarized in Figure VI-3a and detailed in Table E-2. The top 10 PDB hits for each protein were used to create structural models (Figure VI-3b). Each model is generated with confidence measures in the form of C-scores and TM-scores. C-scores range from -5 to 2 where 2 is the highest confidence, and TM-scores range

from 0-1 where 1 is the highest confidence (Zhang, 2009). The similarity between wildtype and mutant structures was then assessed using the Alignment plugin in PyMOL 2.3.2 which provides the root-mean-square deviation of atomic positions (RMSD) values. Higher RMSDs indicates a greater distance between the atoms of mutant proteins superimposed on the wildtype protein. Structural models were generated for CifA (C-score = -2.74;TM = 0.4±0.13), CifA$_{*1}$ (C-score = -1.42; TM = 0.54±0.15; RMSD = 19.7), CifA$_{*2}$ (C-score = -2.84;TM = 0.39±0.13; RMSD = 1.9), CifA$_{*3}$ (C-score = -1.39;TM = 0.54±0.15; RMSD = 20.2), and CifA$_{*4}$ (C-score = -1.43;TM = 0.54±0.15; RMSD = 19.8). These low C-scores and TM-scores indicate that the I-TASSER predictions for CifA are not robust, and that variation in structure between wildtype and mutant proteins could be the result of poor threading templates. However, these results suggest that CifA is structurally most comparable to CifA$_{*2}$ while the other structures are predicted to change to comparable degrees. Crucially, since the most divergent model (CifA$_{*3}$) remains rescue-capable and the second most divergent model CifA$_{*4}$ remains functional in both CI and rescue it is unlikely that global structural variation is responsible for phenotypic ablation in CifA. While these models suggest minimal changes in protein structure upon CifA mutation, they do not rule out the possibility.

I-TASSER was also used to identify PDB hits and create structures for wildtype and mutant CifB. The shared and unique PDB hits for wild-type and mutant proteins are summarized in Figure VI-3c and detailed in Table E-2. As above, I-TASSER protein structures were then created based on these threading templates and compared for RMSD. Structural models (Figure VI-3d) were generated for CifB (C-score = -1.02, TM-score = 0.59±14), CifB$_{*1}$ (C-score = -0.68, TM-score = 0.63±14; RMSD = 1.7), CifB$_{*2}$ (C-score = -1.03, TM-score = 0.58±14; RMSD = 10.6), CifB$_{*3}$ (C-score = -1.07, TM-score = 0.58±14; RMSD = 2.2), CifB$_{*4}$ (C-score = -0.65, TM-score = 0.63±14; RMSD = 2.2). Together, these results suggest that all mutant CifB variants are structurally comparable to the wildtype protein, but do not rule out the possibility. As with CifA, it remains unknown how small effects in protein structure may influence phenotype, but these results suggest that global structural changes are not responsible for phenotypic ablation.

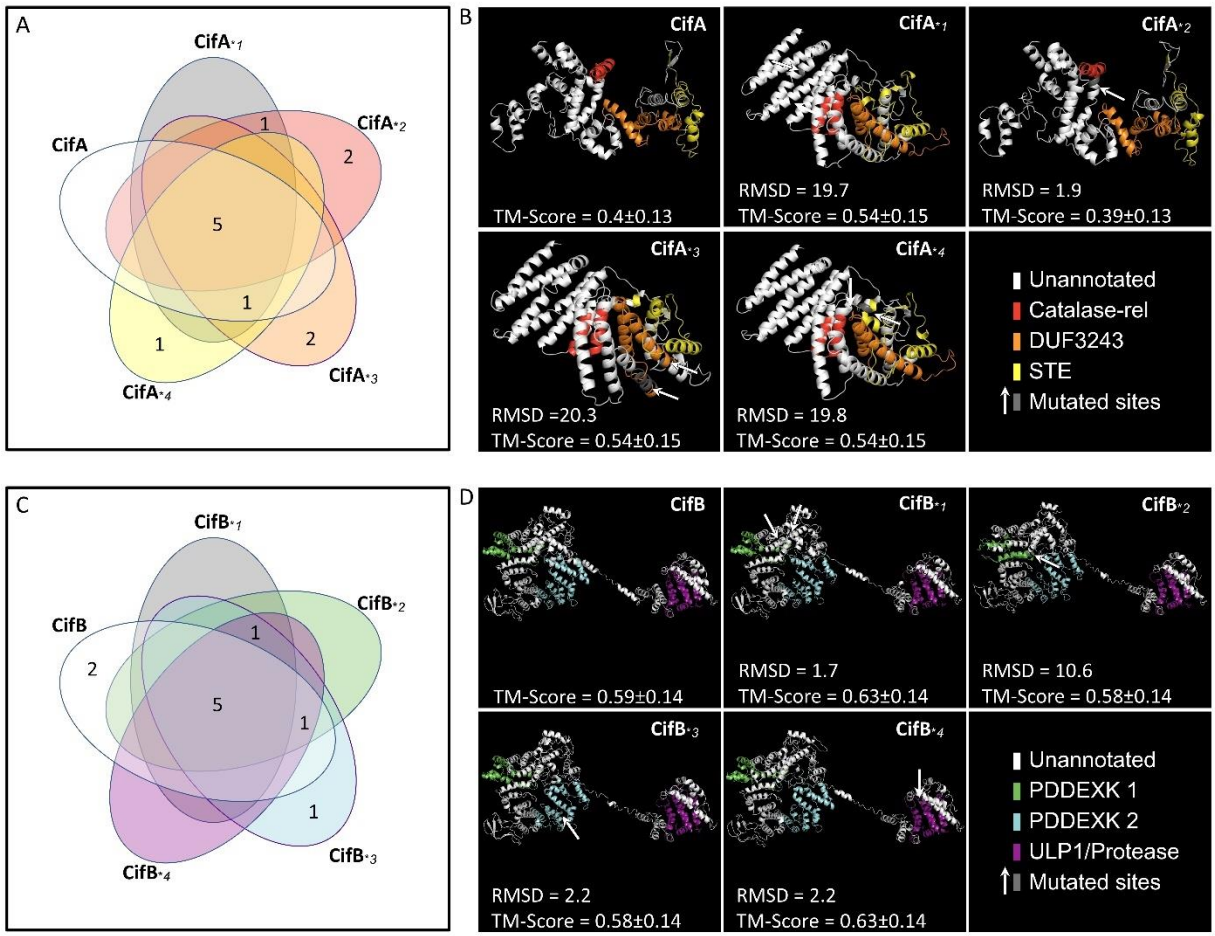

Figure VI-3. Summary of I-TASSER structural predictions for CifA, CifB, and their mutants.
(A, C) Venn-diagrams showing the number of PDB hits shared between wildtype and mutant (A) CifA and (C) CifB proteins. (B, D) I-TASSER uses the PDB hits to generate structural predictions for (C) CifA and (D) CifB. TM-scores range from 0-1 where 1 is the highest confidence. RMSD scores are from pairwise alignment of mutant proteins with the wildtype in PyMol. Higher RMSD scores represent more distance between the superimposed proteins. Mutated sites in the tertiary structure are indicated with a white arrow. Domain annotations were based on previous sequence analyses (Lindsey et al., 2018b).

Discussion

CI is the most prevalent form of *Wolbachia*-induced host manipulation and is explained by a Two-by-One genetic model where *cifA*;*cifB* expression in testes causes CI and *cifA* expression in ovaries rescues CI (Beckmann et al., 2017, 2019c; LePage et al., 2017; Shropshire et al., 2018; Shropshire and Bordenstein, 2019). However, the mechanistic basis of *cifA;cifB*-induced CI and *cifA*-induced rescue remains mostly unknown. Here, we test for the phenotypic importance of

conserved sites (Lindsey et al., 2018b) across the Cif_{wMel} proteins *in vivo* using site-directed mutagenesis and transgenic expression in *D. melanogaster*.

There are multiple explanations for phenotypic ablation caused by mutagenesis. First, the mutated site may represent a component of an unannotated catalytic motif. If so, this would mean that the conserved site, or nearby residues in 3D space, act in an enzymatic capacity to induce CI and/or rescue. Second, the conserved site may be critical for binding to host ligands or for CifA and CifB to bind to one another. In such a case, the mutation may prevent the formation of protein complexes, thus ablating phenotypes. Finally, amino acid substitutions may yield local or global changes to protein structure. Global folding abnormalities may block the enzymatic function of other domains or prevent binding elsewhere in the protein. Local changes to the structure are expected to impact only the region near the mutation in the tertiary structure. The hypotheses described above are applicable to all mutant variants; however, our analysis of protein structure predictions suggest that global changes are unlikely to explain phenotypic variation in mutants relative to wildtype proteins. Below, we discuss specific mutations and how their phenotypic results inform the mechanistic basis of CI and rescue.

Mutating conserved sites in both CifA and CifB's N-terminal unannotated region prevented all functionality in CI and rescue. For CifA, this was the only mutation that ablated both CI and rescue completely. Intriguingly, selection analyses using sliding windows have shown that the N-terminal region of CifA is under stronger purifying selection than the rest of the protein, suggesting that variation in this region is more likely to impact the phenotypic output of CifA (Shropshire et al., 2018). Given that the N-terminus of these proteins are unannotated, we cannot state with confidence why changes in these regions ablate CI and rescue phenotypes. However, the importance of the N-terminus in CifA to both CI and rescue helps to explain why it is under stronger purifying selection than the rest of the protein which had comparable impacts on CI and rescue. It is notable that regions within the N-terminal unannotated domain of *w*Pip homologs of both CifA and CifB have been predicted to encode ankyrin-interacting domains (Bonneau et al., 2018a). Thus, these mutations may have ablated the binding capacity of these genes to host ligands through sequence or structural modifications of the protein. Moreover, CifA and CifB bind *in vivo* (Beckmann et al., 2017), and these mutations may impact that binding affinity. The finding that the N-terminal unannotated regions of CifA and CifB are critical for phenotypic output motivates the continued study of their role in CI and rescue mechanisms.

Mutating conserved sites in the putative catalase-rel domain of CifA yielded a ~30% reduction in hatching when expressed with CifB in males and a ~30% reduction in CI intensity when expressed in females. Aside from the N-terminal unannotated domain, this was the only domain that when mutated impacted both CI and rescue equivalently. Catalases are enzymes that are involved in the decomposition of hydrogen peroxide and protect cells from reactive oxygen species (ROS) damage (Loew, 1900). Some catalase-like domains are involved in host immune pathways that use ROS to combat disease (Govind, 2008; Zug and Hammerstein, 2015), and high levels of ROS can cause male infertility in organisms as diverse as *Drosophila* and humans (Homa et al., 2015; Yu and Huang, 2015). While CifA is annotated with a catalase-rel domain (Lindsey et al., 2018b), the closest sequence homolog is from *Helicobacter pylori*, shares only ~22% sequence identity (Guy et al., 2005), and has no obvious active sites (Lindsey et al., 2018b). It is unclear if CifA's catalase-rel is capable of degrading ROS, but it may otherwise attract or interact with ROS while not successfully degrading them. For example, oxidative posttranslational modifications (PTM) can shift phenotypic output (Cai and Yan, 2013). CifA's catalase-rel may attract ROS, enabling PTM of itself, CifB, or other host-derived targets. Since oxidative PTMs can be reversible (Cai and Yan, 2013), rescue may in part occur through the removal of these PTMs in the embryo. Alternatively, CifA may help to localize ROS to host targets to induce oxidative damage or otherwise modify host targets.

Moreover, *Wolbachia* infection has been correlated with increases in ROS levels in *D. melanogaster, D. simulans, Ae. albopictus, Ae. polynesiensis*, and *T. urticae* (Brennan et al., 2012, 2008; Zug and Hammerstein, 2015). *Wolbachia*-induced increases in ROS levels correlate with DNA damage in *D. simulans* spermatocytes (Brennan et al., 2012) and an increase in lipid hydroperoxides in *D. melanogaster* which are markers for ROS-induced oxidative damage (Driver et al., 2004). While *Wolbachia*-induced ROS levels have been hypothesized as a host immune response to *Wolbachia* infection (Zug and Hammerstein, 2015), most *Wolbachia* known to increase ROS levels are CI-inducing strains. Neither the feminizing *Wolbachia* of *Armadillidium vulgare* nor the mutualist *Wolbachia* of *Asobara tabida* are evidenced to increase ROS levels (Chevalier et al., 2012; Kremer et al., 2012). Intriguingly, the immune-related gene *kenny* (*key*) is upregulated in *Wolbachia*-infected *D. melanogaster*, and experimental upregulation of *key* in uninfected male flies yielded increased ROS levels, DNA damage, and decreased hatching that could be rescued when mated to infected females (Biwot et al., 2019). Together, these data support

a role for ROS in CI's mechanism, but more work will be necessary to determine if CI works through interaction with ROS-associated host genes such as *key*, whether CifA directly influences ROS expression, and if ROS are directly responsible for CI/rescue-induction or are otherwise a symptom of other modifications in gametogenesis.

The DUF in CifA shares distant homology to Puf-family RNA-binding proteins and is the only putative domain shared in all CifA clades (Bing et al., 2020). Mutating conserved residues in CifA's DUF domain revealed those residues to be important for CI-induction but not for rescue. As such, this was the only domain in CifA that was differentially important between the two phenotypes. RNA-binding proteins are important in transcriptional regulation and can influence the stability, localization, and translation of bound RNA. Puf-family RNA-binding proteins typically influence the stability of mRNAs involved in cell maintenance, embryonic development, and other processes (Forbes and Lehmann, 1998; Macdonald, 1992; Parisi and Lin, 1999). For example, the *Drosophila* Puf-family RNA *Pumilio* (*pum*) is critically involved in the establishment of embryonic patterning by suppressing the translation of maternal hunchback RNA in the *Drosophila* embryo (Forbes and Lehmann, 1998; Weidmann and Goldstrohm, 2012). Moreover, *pum* in spermatogenesis negatively regulates the expression of the p53 pathway, increases apoptosis, and reduces sperm production and fertility (Chen et al., 2012). Intriguingly, p32 is a candidate suppressor of CI based on *in vitro* pull-down assays using *Drosophila* lysates and p32 is another regulator in p53 activation (Beckmann et al., 2019c; Ghate et al., 2019), suggesting that *pum* in spermatogenesis may influence similar pathways in CI. Additionally, *Wolbachia* infection has been shown to have a considerable impact on the fly transcriptome and even sRNA profiles (Baião et al., 2019; Pinto et al., 2013; Y. Zheng et al., 2019). On their own, these correlations do not sufficiently link transcription with CI. However, as described above, *key* is significantly upregulated and causes rescuable hatch rate defects when experimentally overexpressed (Biwot et al., 2019). Additionally, *Wolbachia* upregulate the sRNA nov-miR-12 which negatively regulates *pipsqueak* (*psq*), a DNA-binding protein that impacts chromatin structure (Horowitz and Berg, 1996; Siegmund and Lehmann, 2002), and knockdown of *psq* causes CI-like embryonic abnormalities and hatch rates in *D. melanogaster* (Y. Zheng et al., 2019). Thus, CifA's DUF may influence the expression of RNAs involved in the CI pathway, and by mutating the conserved site it may ablate the domain's ability to regulate these RNAs. It is also critical to note that since the DUF mutant only prevented CifA from contributing to CI, it supports prior hypotheses that CifA

has distinct mechanistic input to CI and rescue (Shropshire and Bordenstein, 2019). The phenotypic plasticity of CifA may be caused by distinct protein conformations in testes and ovaries or DUF-associated targets may only be present in testes. More work will be necessary to confirm that CifA can bind RNAs and what impact this binding has on downstream processes.

The final CifA domain shares homology to STE proteins which are found predominantly in fungi and encode a sequence-specific DNA-binding motif that influences yeast reproduction through pheromone-responsive elements (Wong Sak Hoi and Dumas, 2010). Mutation of conserved sites within the STE had no impact on either CI or rescue. This was surprising since the STE domain appeared structurally conserved across numerous CifA phylogenetic Types and the conserved nature of the residue suggested that it was critical for phenotypic expression (Bing et al., 2020; Lindsey et al., 2018b). However, since transgenes are expressed via host transcription and translation machinery, transgenic expression by-passes the need to export the proteins outside of *Wolbachia* and through the host-derived membranes that surround *Wolbachia* within the cell (Cho et al., 2011; Fattouh et al., 2019). As such, it is possible that the STE domain has a functional role in translation initiation within *Wolbachia* and/or in protein export. For example, CifA and CifB are adjacent proteins hypothesized to be regulated as an operon (Beckmann et al., 2019b, 2017; Bonneau et al., 2018a; Shropshire et al., 2019), but expression levels are considerably different between the two proteins (Lindsey et al., 2018b). It is common that proteins under an operon are differentially expressed (Güell et al., 2011; Murakawa et al., 1991), but regulators underpinning differential expression of *cif* genes remain unknown. If regulated as an operon, the STE-like domain of CifA may act as an auto-regulator of the transcriptional abundance of CifB. Additional research will be necessary to determine if any component of the STE domain is necessary for CI and whether this domain is essential when expressed inside *Wolbachia*.

Type I CifB have two domains with homologs in the PDDEXK nuclease family, but they do not encode canonical catalytic sites (Chen et al., 2019; Knizewski et al., 2007; Lindsey et al., 2018b). This nuclease family is heavily involved in DNA restriction, repair, recombination, and binding (Knizewski et al., 2007). Mutations in conserved sites in either domain ablated CI-phenotypes. This indicates that sites within both nuclease domains play a role in CI, but does not necessarily mean that they act to create DNA nicks. Interestingly, the Cif proteins are split into at least five phylogenetic "Types" and only the Type I CifB proteins encode Ulp1 domains; Type II, III, IV CifB only encode two nuclease domains (Lindsey et al., 2018b), and Type V CifB have two

nuclease domains and other domains unique to the clade (Bing et al., 2020). The Type IV CifB have functional nucleases with canonical catalytic motifs and can induce CI (Chen et al., 2019), but it remains unclear how these DNA nicks contribute to CI-induction and how this can be rescued by CifA expressing females. Moreover, while we show here that mutating conserved residues in either PDDEXK domain ablates CI-induction, this does not confirm its role as a nuclease. Instead, these domains may be essential for the localization of Cif proteins to host DNA or other host targets and primary catalytic activity that leads to CI may have switched from DNA damage to Ulp1 activity. More work will be necessary to determine CifB's nuclease domains are active and why mutating these conserved sites ablates CI.

CifB's final domain is a Ulp1 domain that contains the only known catalytic motif within the Cif proteins and is responsible for the deubiquitinase activity observed *in vitro* (Beckmann et al., 2017). Previous reports show that mutating the conserved cysteine active site ablates CI function in $CifB_{wPip}$ (Beckmann et al., 2017). Here, we confirm that mutating the cysteine active site ablated CI in the $CifB_{wMel}$ protein. However, it is premature to suggest that the Ulp1 domain is the "catalytic warhead" for CI (Beckmann et al., 2017) because several sites, when mutated, ablate the CI phenotype. Instead, it is evident that the Ulp1 domain plays a catalytically important role in CI-induction but is unlikely to be the only critically important part of the protein. More work will be necessary to dissect the relative importance of the unannotated region, the nuclease domains, and the Ulp1 in CI's mechanism.

In conclusion, we report conserved amino acids in CifA and CifB that are essential for CI and rescue phenotypes. For CifA, conserved sites in the unannotated region and catalase-rel domain were important for CifA-induced CI and rescue, while the mutated sites in the DUF was additionally important to CI. For CifB, mutating conserved sites in an unannotated region, both PDDEXK nuclease domains and the Ulp1 domain were important in CifB-induced CI. These works provide additional support for the necessity of expressing both CifA and CifB proteins to cause CI, insight into the phenotypic heterogeneity of these proteins, and avenues of research to link loss-of-function mutations with CI's molecular basis.

Materials and methods

Creating transgenic flies.

One mutant *cifA* and *cifB* gene variant was synthesized at GenScript and cloned into a pUC57 plasmid. Site-directed mutagenesis was then performed by GenScript to produce the remaining three mutant variants of these genes (Figure VI-1a, Figure VI-2a). UAS transgenic *cifA*$_*$ and *cifB*$_*$ mutant flies were then generated following previously described protocols (LePage et al., 2017). Briefly, each gene was subcloned into the pTIGER plasmid, which is a pUASp-based vector designed for germline expression, and was then integrated into a targeted region of the *D. melanogaster* genome using PhiC31 integrase via embryonic injections at BestGene (LePage et al., 2017).

Fly rearing and strains.

D. melanogaster stocks y^1w* (BDSC 1495), *nos*-GAL4:VP16 (BDSC 4937), and UAS transgenic lines homozygous for *cifA*, *cifA*$_*$, *cifB*, *cifB*$_*$, *cifA*;*cifB*, *cifA*;*cifB*$_*$, and *cifA*$_*$;*cifB* were maintained on a 12-hour light/dark cycle at 25^{o}C on 50mL of standard media. Dual transgenic lines were generated through standard genetic crossings and were all homozygous viable. Uninfected lines were produced by tetracycline treatment as previously described (LePage et al., 2017). Infection status for all lines was regularly confirmed by PCR using Wolb_F and Wolb_R3 primers (Casiraghi et al., 2005). Genotyping was confirmed by PCR and Sanger sequencing using the primers in Table E-3.

CI measurement assays.

CI was measured using hatch rate assays. To control for the paternal grandmother age effect on CI (Layton et al., 2019), virgin *nos*-GAL4:VP16 females were collected for the first 3 days of emergence and aged 9-11 days before crossing to nonvirgin UAS transgenic (*cifA*$_*$, *cifA*;*cifB*$_*$, or *cifA*$_*$;*cifB*) males. Collections for maternal and paternal lineages were separated by a 7-day period. Individual male and female mating occurred in 8-oz *Drosophila* stock bottles with a grape-juice agar plate smeared with yeast and secured to the opening of each bottle with tape. Only the first emerging and youngest males were used to control for the younger brother effect and age effects on CI (Reynolds and Hoffmann, 2002; Yamada et al., 2007). Grape-juice agar plates were produced as previously described (LePage et al., 2017). The flies and bottles were incubated at

25°C for 24 hours, at which time the grape plates were replaced with fresh plates and stored for an additional 24 hours. After this, the initial number of embryos on each plate were counted. The plates were incubated at 25°C and after 30 hours, the number of unhatched embryos were counted. The percentage of embryos that hatched was calculated by dividing the number of hatched embryos by the total number of embryos and multiplying by 100.

Predicting the impact of mutagenesis on protein structure.

The effect of mutations on protein structure was evaluated with the I-TASSER protein prediction tool (Zhang, 2009). I-TASSER generated protein tertiary structure predictions for Cif proteins and their mutants using the on-line server with default settings. Structures are build based on the top ten hits generated by querying the PDB. Hits were provided Z-scores that characterize the similarity to the query sequence. Higher Z-scores represent more confident matches. C-scores and TM-scores were generated for each tertiary structure. C-scores range from -5 to 2 where 2 is the highest confidence. TM-scores range from 0-1 where 1 is the highest confidence.

Statistical analysis.

All statistical analyses for hatch rates were conducted in GraphPad Prism 8. Hatch rate statistical comparisons were made using Kruskal-Wallis followed by a Dunn's multiple comparison test. All p-values from statistical comparisons are provided in Table E-1. Figure aesthetics were edited using Affinity Designer.

Chapter VII.

Recent genetic and biochemical advances in our understanding of *Wolbachia*-induced cytoplasmic incompatibility

CI genetics and phylogenetics

In the past decade, significant advances have been made toward our understanding of the genetics of *Wolbachia*-induced CI. The sequencing of the *w*Mel genome revealed a considerable list of candidates that may be related to CI-induction (Wu et al., 2004), including numerous ankyrin domains predicted to be involved in protein-protein interactions. Testing of these candidates proved not to be fruitful (Yamada et al., 2011). However, mass spectrometry and SDS page analyses of spermatheca extracts from *Cu. pipiens* females mated to *w*Pip infected males the prophage WO protein WPIP0282 was present in spermatheca (Beckmann and Fallon, 2013), thus identifying a new candidate for CI and/or rescue. Later, genomic comparisons of the *w*Mel genome against the genome of the non-parasitic *w*Au strain revealed 10 genes absent in the non-parasitic strain that were present in *w*Mel (Sutton et al., 2014). These genes included numerous genes from *Wolbachia*'s prophage WO such as WD0631 which is a *w*Mel homolog of *w*Pip's WPIP0282 (Sutton et al., 2014). Sequencing of the *w*Rec genome revealed a highly reduced phage (Metcalf et al., 2014) and formed the basis of an unbiased, comparative 'omic study assessing the genomes of CI-inducing *Wolbachia*, a genome of a non-parasitic strain (*w*Au of *D. simulans*), and a transcriptome and proteome of *Wolbachia*-infected ovaries revealing only WD0631 and the adgacent WD0632 as candidates for CI (LePage et al., 2017). These genes were named cytoplasmic incompatibility factors A and B (*cifA* and *cifB*) respectively, and experimental assays were conducted to test if *cifA* and *cifB* could cause CI. Two studies independently and simultaneously explored the role of the *cif* genes in the *w*Mel (LePage et al., 2017) and *w*Pip (Beckmann et al., 2017) strains using transgenic expression systems in *D. melanogaster*. With *w*Mel, singly expressing $cifA_{wMel}$ or $cifB_{wMel}$ in uninfected males failed to induce CI, but dual expressing the genes caused rescuable CI-like hatch rates and cytological defects (LePage et al., 2017). Similar results were reported under dual expression of $cifA;B_{wPip}$ genes when transgenically expressed in uninfected *D. melanogaster* males, but rescue was not achieved (Beckmann et al.,

2017). Later, similar transgenic experiments revealed that *cifA*$_{wMel}$ expression in uninfected *D. melanogaster* females can rescue CI (Shropshire et al., 2018), motivating a Two-by-One genetic model wherein *cifA* and *cifB* cause CI unless *cifA* is expressed in the ovaries or embryo. This model was experimentally validated through transgenic expression of *cifA*$_{wMel}$ and *cifB*$_{wMel}$ in males to induce transgenic CI, and through crossing them to *cifA*$_{wMel}$ expressing females, CI was rescued at levels comparable to an infected female (Shropshire and Bordenstein, 2019).

Comparative sequence analysis of Cif proteins reveals that CifA and CifB have concordant phylogenies with considerable divergence across five clades referred to as Types I-V (Bing et al., 2019; LePage et al., 2017; Lindsey et al., 2018b). The *w*Mel Cif proteins belong to the Type I clade and *w*Pip has both Type I and Type IV Cif proteins. The phylogenetic classification of a *cif* gene is indicated with a T# in brackets beside the gene name (i.e., *cifA*$_{wMel[T1]}$). While the *cif* genes are often associated with the Eukaryotic Association Module of prophage WO, Cif phylogeny is not concordant with phage WO or *Wolbachia* phylogeny, suggesting an independent evolutionary history (LePage et al., 2017). Some *cif* genes are flanked by ISWpi1 transposons which may mediate horizontal transfer between *Wolbachia* strains and/or phages, but it remains unclear if they alone are responsible for divergence between the phylogeny of Cifs, *Wolbachia*, and phage WO (Cooper et al., 2019; Madhav et al., 2020). It also remains unclear if all *cif* Types are capable of causing and/or rescuing CI. However, the *cif* loci of numerous CI-inducing *Wolbachia* strains have been tested to induce at least small and rescuable reductions in hatching, including *cif*$_{wMel[T1]}$, *cif*$_{wPip[T1]}$, *cif*$_{wRec[T1]}$, *cif*$_{wRi[T2]}$, and *cif*$_{wPip[T4]}$ (Chen et al., 2019) (Chapter V). The Type III and Type V loci remain untested. However, the CI-inducing strains *w*No of *D. simulans* and *w*Stri of *La. striatellus* only have Type III or Type V loci respectively (Bing et al., 2019; LePage et al., 2017), suggesting that these loci may cause CI and rescue. These studies further support that despite considerable divergence in sequence, proteins across the phylogenetic landscape of the Cif proteins remain capable of causing and rescuing CI.

These works provide strong evidence for the genetic basis of unidirectional CI between infected and uninfected insects. However, the genetic basis of bidirectional and unidirectional CI between insects with different *Wolbachia* remains less understood. Sequence divergence in CI-associated factors has long been thought to be a contributing factor to these incompatibilities (Charlat et al., 2001). Indeed, phylogenetic analyses of *cif*-alleles reveal that strains carrying similar alleles tend to be compatible, that more distantly related *cif*-alleles are not, and that a single

Wolbachia strain can have numerous *cif* copies (Bonneau et al., 2019, 2018a; LePage et al., 2017). For example, *w*Pip *Wolbachia* can be split into five distinct phylogenetic groups (wPipI-wPipV) and there are complex compatibility relationships between these strains (Atyame et al., 2014). Population genetic analyses of *cif* loci in *w*PipI and *w*PipIV reveal that while each strain carries multiple closely related *cif* variants that belong to Type I and Type IV *cif* clades, and a single genetic variant of $CifB_{wPip[T1]}$ correlates with *w*PipI's inability to rescue *w*PipIV-induced CI (Bonneau et al., 2019, 2018a). These data suggest that *cif* genetic variation alone can explain the diversity of incompatibility relationships. However, since *w*PipI and *w*PipIV *Wolbachia* are also from different *Cu. pipiens* populations (Atyame et al., 2011a), it still remains possible that host variation can contribute to these incompatibility relationships in a way that also correlates with *cif* genotypic diversity. It is also notable that the finding that *cifA* is involved in both CI and rescue enables a framework for the emergence of bidirectional CI wherein sequence variation in *cifA* alone can lead to divergence in CI relative to ancestral variants while remaining compatible with the emergent mutant (Shropshire et al., 2018). Functional studies will be necessary to confirm that *cif* sequence variation alone can explain the emergence of incompatibilities between *Wolbachia,* and which of the *cif* genes contributes most substantially to incompatibility relationships.

Biochemical basis of Cif-induced CI and rescue

Little is known about how the Cif proteins work to induce CI. To date, biochemical studies have been restricted to CifB. $CifB_{[T1]}$ from both *w*Mel and *w*Pip were initially characterized with a single putative Ulp1 Protease domain (Beckmann et al., 2017; LePage et al., 2017; Lindsey et al., 2018b). This domain has since been determined to cleave K6-, K11-, K27-, K29-, K33-, K48-, and K63-linked ubiquitin chains *in vitro*, but with a preference for deubiquitination of K63-ubiquitin chains (Beckmann et al., 2017). K63 chains are associated with numerous host activities including NF-κB signaling which regulates DNA transcription and are also found on autophagocytosed structures (Tan et al., 2008; Wertz and Dixit, 2010). As such, CifB's ability to cleave K63 chains may implicate it in these kind of processes. A single amino acid mutation in the catalytic site of the Ulp1 domain prevents the expression of transgenic CI and the breakdown of ubiquitin chains into ubiquitin (Beckmann et al., 2017). Expressing the Ulp1 catalytic mutant for

CifB in males alongside CifA did not induce CI, suggesting that deubiquitilase activity or structural changes are important for CI (Beckmann et al., 2017).

While CifB's Ulp1 domain is seemingly important for CI, only Type I CifB have the Ulp1 domain. However, all CifB proteins are annotated with dimers of PD-(D/E)XK (hereafter PDDEXK) nuclease domains (Bing et al., 2020; Lindsey et al., 2018b). Indeed, purified $CifB_{wPip[T4]}$ protein act as *in vitro* nucleases, and mutating PDDEXK catalytic sites prevents nuclease activity *in vitro* and CI-inducibility when expressed in *D. melanogaster* (Chen et al., 2019). $CifB_{[T1]}$ proteins do not have the canonical PDDEXK catalytic sites, thus lending doubt to the importance of these domains in Type I CifB (Beckmann et al., 2017). However, PDDEXK domains are often reported as functional without canonical catalytic motifs and some PDDEXK-like domains without catalytic sites are still involved in other DNA-associated processes (Knizewski et al., 2007). Indeed, mutagenesis of conserved residues in the PDDEXK domains of $CifB_{wMel[T1]}$ ablate CI under transgenic expression (Chapter IV), suggesting that these domains are important for CI expression. Moreover, despite *w*Pip having both Type I and IV genes, there are no notable differences in embryonic defects caused when both genes are expressed as compared to other strains that only have $CifB_{[T1]}$, suggesting that these genes have converged on similar cytological outcomes (Bonneau et al., 2018b). It remains unclear if these Cif variants have distinct mechanisms and more work understanding their enzymatic and biochemical outputs will be necessary.

Our understanding of CifA's function is purely based on predictions. For instance, structural homology-based annotations suggest that Type I CifA have three putative domains (Lindsey et al., 2018b). $CifA_{[T1]}$ is predicted to have a catalase-related domain involved in the degradation of reactive oxygen species, a domain of unknown function (DUF) 3243 with homology to a Puf-family RNA binding domain (RBD), and a sterile-like transcriptional regulator (STE) (Lindsey et al., 2018b). Importantly, for these annotations to be reported, the annotation needed only to exceed 20% probability, which is admittedly very low. While the catalase-rel domain is unique to the $CifA_{[T1]}$, the STE is maintained in Type I-IV genes (Lindsey et al., 2018b), and the Puf-family RBD exists in Type I-V loci (Bing et al., 2020). Sliding window analyses of selection for $CifA_{[T1]}$ suggest that while the entire protein is under purifying selection, the catalase-rel domain and the unannotated N-terminal region are under the strongest selection (Shropshire et al., 2018). Interestingly, mutagenesis of conserved sites across $CifA_{wMel[T1]}$ reveal that sites in the

N-terminal unannotated region and the catalase-rel are involved in both CI and rescue, sites in the Puf-family RBD contributes to CI, and the STE did not play a role in either phenotype (Chapter VI). These studies are suggestive that the N-terminal region of the protein has comparable function in CI and rescue, but uses residues in the Puf-family RBD specifically for CI (Chapter VI).

For many proteins to act, they need to bind to other proteins in the host. Emerging evidence has yielded a considerable list of potential binding factors for the Cif proteins. For example, CifA and CifB bind one another *in vitro* (Beckmann et al., 2017). It does, however, remain unknown if CifA binds CifB in the testes to cause CI or if maternal CifA binds to paternal CifB in the embryo to cause rescue (Beckmann et al., 2019a, 2019b; Shropshire et al., 2019). More work on the localization, co-localization, and binding profiles of these proteins will elucidate this question. Additionally, Cifs bind to a suite of host proteins that differ based on if CifA and CifB are expressed alone or together. Indeed, 67 host proteins were identified under co-expression of CifA and CifB and 45 different proteins with CifB expression alone (Beckmann et al., 2019c). Of these proteins, karyopherin-α (kap-α) was determined to be of particular importance since its overexpression in females could partially rescue CI. Kap-α is a nuclear import receptor and is also a regulator of p53 which has roles in the protamine-histone exchange process known to be involved in CI (Beckmann et al., 2019c; Emelyanov et al., 2014). Not only do these data suggest that Kap-α is directly related to CI-induction, but that it may be a host suppressor of CI intensity since overexpression of Kap-α in uninfected females yields partial rescue (Beckmann et al., 2019c). Intriguingly, Kap-α was only pulled down when CifB alone was expressed (Beckmann et al., 2019c). More work is necessary to determine CifA's role in CI-induction and how exactly CifB's binding to Kap-α contributes to CI-induction.

How CifA is involved in both CI and rescue remains largely a mystery. The simplest explanation is that CifA maintains the same function in both CI and rescue. Under this framework, CifA would act on a pathway that can be modified during spermatogenesis and in oogenesis to produce opposite affects. For example, the mistiming model for CI (discussed more below) suggests that a delay in male pronuclear development causes CI and that rescue occurs through a reciprocal delay in the female pronucleus. Under this scenario, delay of the female pronucleus alone is not anticipated to cause lethality sense the first embryonic mitosis begins when the female pronucleus is fully developed, not when the male pronucleus is (Ferree and Sullivan, 2006). If CifA were to drive this delay, then CifB's role in CI would seemingly be auxiliary and perhaps

only necessary for localization of CifA to particular targets or to protect CifA from degradation by ubiquitin pathways. Alternatively, CifA may be a multi-functional protein that employs one set of functions to cause CI and another to cause rescue. For instance, if CifA targets sex-specific host pathways, it can be expected that CifA can only affect its host in a particular way if that target is available. Additionally, CifA may be modified in some manner that differs between the two environments, unlocking unique biochemical functions by posttranslational modification, localization differences, or the expression of different protein isoforms.

Finally, some *Wolbachia* strains or combinations of strains have CI phenotypes that are uncharacteristic of standard CI-rescue relationships. For example, *w*Suz of *D. suzukii* carry both Type I and Type II *cif* loci that are highly similar to the strong CI-inducing strain of *w*Ri but do not themselves cause appreciable CI (Cattel et al., 2018; Conner et al., 2017; Hamm et al., 2014; Lindsey et al., 2018b). Between 2-4 amino acid substitutions are in each protein and it remains possible that this small variation in sequence could be responsible for *w*Suz being unable to cause CI. However, *w*Ri is a strong CI-inducer in *D. simulans* that fails to cause appreciable CI when transinfected into *D. suzukii* (Cattel et al., 2018), despite transinfections of strains such as *w*Ha or *w*Tei yielding strong CI. As such, these studies suggest that *D. suzukii* has host suppressors for CI-induction but that those suppressors are specific to strains that are closely related to *D. suzukii*'s native *w*Suz *Wolbachia*. Understanding the mechanism underlying these suppression mechanisms will not only inform our understanding of the evolutionary dynamics of *Wolbachia*-host interactions but also further unravel the diversity of mechanisms that contribute to CI. Similarly, two *Wolbachia* strains from *D. pandora*, *w*PanMK and *w*PanCI, induce male-killing and CI respectively. The male-killing strain has an early stop codon that putatively ablates function and allowing for the phenotypic switch from CI to male-killing (Asselin et al., 2018), but this remains untested. Moreover, some *Wolbachia* infection states yield CI-induction without being capable of self-rescue, as is the case with the triple-strain infection of *w*AlbA, *w*AlbB, and *w*Mel in *Ae. albopictus* (Ant and Sinkins, 2018). Since each of the individual *Wolbachia* strains in this triple-strain infection are self-compatible, genetic variation in the *cif* genes alone cannot explain the emergence of self-incompatibility under superinfection. It is likely that variation in host genotype corresponds to these descrepencies since theory predicts that hosts will suppress infection (Turelli, 1994).

Models for *Wolbachia*-induced CI and rescue

Numerous models have been developed to explain how CI and rescue work. These models leverage data involving phenotypic variation, embryonic and spermatogenic cytology, genetic advancements, and biochemical assays to create testable hypotheses for *Wolbachia*'s selfish spread mechanism. Below, we discuss the utility of the modification/rescue (mod/resc) model in a post-genomic world and mechanistic models currently used to explain CI. It is important to note, that despite considerable advances in our understanding of CI genetics and biochemistry, we have not yet reached a complete understanding of CI's mechanism and none of the below models can be completely ruled out given the current data. We discuss each model in the context of a Two-by-One genetic model, but it is critical to reemphasize that while both CifA and CifB are required to induce CI, the specific underlying contribution of each protein to the phenotype remains unknown.

The mod/resc model.

The mod/resc model defines a mod factor as a CI-inducing product produced in males and a resc factor as a rescue-inducing product produced in females (Werren, 1997). The mod/resc model does not make assumptions about the genetic, biochemical, enzymatic, or cytological basis of CI. Instead, the mod/resc model provides a framework for describing the CI-inducibility of different *Wolbachia* strains. For example, a standard CI-inducing strain that can self-rescue would be denoted as mod+/resc+. Less common phenotypes include so-called suicidal *Wolbachia* (mod+/resc-) and *Wolbachia* that do not cause CI but can rescue CI induced by other strains (mod-/resc+) (Ant and Sinkins, 2018; Meany et al., 2019; Zabalou et al., 2008). *Wolbachia* that do not cause CI or rescue are designated mod-/resc-.

The mod/resc model assumes that for bidirectional CI to occur, the mod and resc factors would differ in such a way that they remain functional but are incompatible with each other (Charlat et al., 2001; Werren, 1997). As such, a strain can carry multiple mod or resc factors that determine the compatibility relationships with other strains, and the mod/resc model can be used to estimate the number of mod and resc factors within a host (Zabalou et al., 2008). To do this, *Wolbachia* strains are transinfected into the same genetic background and then crossed to determine the incompatibility relationships between strains or against uninfected flies. A strain that causes CI against an uninfected female is considered to have at least one mod factor. If it can

rescue itself then it has at least one resc factor. If two CI-inducing and self-compatible strains are bidirectionally incompatible, then it is assumed that each carry at least one set of mod and resc factors but that they are not the same. Indeed, crossing experiments between various *Wolbachia* strains have revealed unidirectional and bidirectional incompatibilities which have led to agreement that *Wolbachia* frequently carry multiple mod and resc factors (Poinsot et al., 1998; Zabalou et al., 2008).

With the identification of the CI and rescue genes (Beckmann et al., 2017; Chen et al., 2019; LePage et al., 2017; Shropshire et al., 2018; Shropshire and Bordenstein, 2019), it is compelling to abandon the mod/resc model in favor of a purely genetic description of CI relationships. Indeed, with the ever-growing availability of genomic datasets, this may prove to be the simplest way to generate hypotheses about compatibilities between CI-inducing *Wolbachia* strains, where strains carrying similar profiles of *cif* genes would be predicted to be compatible. However, we urge that the mod/resc model still holds value since sequence information alone does not always correlate with expected phenotypic results. For example, *w*Suz does not induce CI while maintaining genes that at first glance may appear functional (Hamm et al., 2014). Thus, despite being classified as likely CI-inducers at the genetic level, the phenotypic mod/resc model would suggest *w*Suz maintains neither component. Moreover, the triple-strain infection of *w*AlbA, *w*AlbB, and *w*Mel in *Ae. albopictus* does not induce rescuable CI despite each individual strain being self-compatible (Ant and Sinkins, 2018). Thus, despite genetically appearing that this strain combination would be rescue-capable, this line would be marked as a mod+/resc- insect line. The areas where genetic and phenotypic information disagree will be exciting areas for future investigation.

Mechanistic models.

Despite considerable advances in the genetics and biochemistry of CI, there are still numerous conflicting mechanistic models used to describe CI and rescue (Beckmann et al., 2019a; Poinsot et al., 2003; Shropshire et al., 2019). Here we will describe each model, some of the supporting data, and how a Two-by-One genetic framework fits into the model. These models fall under two discrete categories: host-modification (HM) and toxin-antidote (TA) (Beckmann et al., 2019a; Shropshire et al., 2019). HM-based models assume that the CI-inducing factors act directly

on host products to modify them and that rescue occurs through either removal of these modifications or otherwise reversing the effects through a separate host modification in the female. TA-based models assume that the CI-inducing factors are transported into the embryo with the sperm and are toxic unless the rescue factor is present, binds to the CI-toxin, and inhibits its toxicity. Below, we discuss these two models in more detail and highlight variants of these models that make additional assumptions about CI's mechanism.

HM models require that host products are modified prior to fertilization by the Cif proteins and those modifications induce CI in the fertilized embryo unless the modification is replaced or otherwise negated by CifA in the embryo. Indeed, there are numerous pre-fertilization defects associated with infection with CI-inducing *Wolbachia* including changes in sperm morphology and competitive ability (Champion de Crespigny and Wedell, 2006; Riparbelli et al., 2007), supporting that the host is modified prior to fertilization. It is unknown whether these pre-fertilization defects are causatively related to CI or are just correlated with *Wolbachia* infection. Moreover, a key prediction of the HM model is that rescue does not work through direct binding of maternal CifA with male-derived Cif products. Instead, CifA may interact with host processes to reverse or otherwise stop the effects of CI caused by CifA;CifB expression in males. Since both CifA and CifB are required to cause CI (Beckmann et al., 2017; Chen et al., 2019; LePage et al., 2017; Shropshire and Bordenstein, 2019), the binding affinity of the two proteins can be explained by the proteins needing to interact in the testes to cause CI. As such, assessment of the location of CifA and CifB binding, the transfer of Cif products with the sperm, and the interactions that Cif have with the host will further inform this model. We discuss three additional HM-based models below: titration-restitution, mistiming, and goalkeeper.

The titration-restitution model posits that CI is induced by misregulation of host factors or pathways in the testes/sperm and rescue occurs when the same factors are misregulated in the opposite direction in the ovaries/embryo (Kose and Karr, 1995; Poinsot et al., 2003). Indeed, *Wolbachia* has a considerable impact on expression profiles, and some genes are differentially expressed in male and female reproductive tissues (Baldridge et al., 2017, 2014; Bing et al., 2019; Yuan et al., 2015). For example, qRT-PCR of full *D. melanogaster* bodies revealed the histone chaperone Hira is upregulated in infected females and downregulated in infected males, and RNAi of Hira in males leads to reduced hatching which can be rescued by infected females (Zheng et al., 2011). Since CI and rescue would occur through titration of the same host product, and CifA is

known to be crucial for both phenotypes (Shropshire and Bordenstein, 2019), CifA may be the primary agent in CI and rescue. As such, CifB may act as an accessory that enables CifA to get to a paternally derived product that it would otherwise not be able to target. The conserved nature of CifB's nuclease dimers, and relationship to nuclear import through binding to the Kap-α nuclear import receptor, suggests that the CifA and CifB complex might localize to host DNA, allowing for the addition or removal of host factors to paternal chromosomes or interactions with host mRNAs (Beckmann et al., 2019c; Lindsey et al., 2018b). Alternatively, CifA may on its own lead to up- or down-regulation of a host product but has the opposite impact on that product when CifB is present. As such, rescue would occur through CifA's lone action which counteracts the misregulation caused by CifA;CifB dual expression. Bidirectional CI can then be explained by variable impacts on multiple host expression pathways. Thus, rescue would not be possible from a second strain since it could be targeting the wrong host factor or pathway. Indeed, divergent CI genes may impact differential impact on host pathways. For example, only the Type I CifB maintain a functional Ulp1 domain while the other four Cif clades have a dimer of PDDEXK nucleases (Beckmann et al., 2017; Bing et al., 2020; Lindsey et al., 2018b). It is therefore feasible that CifB with different domains impact different host pathways. Alternatively, Cif proteins may have differential impacts on the level of misregulation instead of, or in addition to, impacting multiple host pathways which may influence incompatibility relationships. More work will be necessary to understand if Cif expression influences transcriptional and translational variation and how that variation corresponds to CI.

The mistiming model is based on the observation that the paternal pronucleus has slowed development relative to the female pronucleus in CI crosses, and the rescue cross has normal cell cycle timing (Ferree and Sullivan, 2006). This established the hypotheses that delayed male pronuclear development is responsible for emergent defects in early embryogenesis, and that resynchronization of the development may occur by comparably slowing down the development of the female pronucleus or slowing the cell cycle in rescue. Since the cell cycle timing of the female pronucleus is what establishes the timing for the first mitosis (Bossan et al., 2011), infected females do not induce CI since the male pronucleus will get into position prior to the female pronucleus. Though, the reciprocal cross would be incompatible because the female pronucleus will have finished development prior to the male. Thus, rescue would occur as long as the female pronucleus is delayed at least as long as the male pronucleus. Importantly, this model predicts that

CI crosses would be subject to haploidization of diploid offspring since the male pronucleus could be completely excluded from mitosis if it was significantly slowed. This is indeed the case in *N. vitripennis* where CI often manifests as only male offspring since haploid offspring are viable in this species but develop as males (Bordenstein et al., 2003).

The mistiming model proposes that CI and rescue have comparable impacts on the development of male and female gametes respectively. As such, a single gene could in theory be responsible for both CI and rescue (Poinsot et al., 2003). Under this paradigm, CifA may be enacting a slow-down in both tissues since it is involved in both phenotypes (Shropshire and Bordenstein, 2019). However, if this were the case, then what would be the purpose of CifB? It is possible that CifB is responsible for localizing CifA to a male specific target where it imposes the same outcomes on its host. Since this hypothetical male product would not be available in the embryo, CifB would not have a role in rescue. However, an alternative model for mistiming is that rescue may not occur through slowing down the female pronucleus but may instead work by removing the slow-down agents from the male pronucleus. More work will be necessary to understand if rescue occurs via slow-down of the female pronucleus or from speading-up the male pronucleus.

A major limitation of the mistiming model is that it cannot explain bidirectional CI. Since mistiming proposes that rescue happens through delaying the female pronucleus as much as or greater than the male pronucleus, a sufficiently strong delay should yield compatibility with any strain that has a weaker male delay. As such, only unidirectional CI should manifest between strains where the strain inducing the stronger delay is capable of rescue. The goalkeeper model was proposed as a way to address this limitation (Bossan et al., 2011). In addition to the expectations of the mistiming model, goalkeeper suggests that a secondary factor unassociated with this mistiming may also be involved in CI. The combined contribution of these two mod factors leads to CI. Under this paradigm, CifA and CifB may contribute to different types of defects during spermatogenesis, each contributing in somewhat independent ways to CI induction. Rescue must then negate the impact of both factors. Thus, for CifA to rescue CI it would not only need to contribute to a delay in the pronuclear development but also reverse the impacts of a secondary source of modification.

Since *Wolbachia* are not paternally inherited, Hurst proposed in 1991 that *Wolbachia* make a CI-inducing toxin that diffuses into the sperm cytoplasm and is transferred to the egg during

fertilization and causes death (Hurst, 1991). Rescue then occurs when *Wolbachia* in the egg produce an antidote that binds to toxin and prevents it from killing the embryo (Hurst, 1991). We now know this model as the TA model (Beckmann et al., 2019a; Shropshire et al., 2019) and that CifA and CifB are the CI-inducing proteins (Beckmann et al., 2017; Chen et al., 2019; LePage et al., 2017). The TA model makes two key predictions. First, the Cif-proteins are transferred to the embryo. Mass spectrometry of spermatheca from infected *Cu. pipiens* females mated with infected males reveal fragments of CifA (Beckmann and Fallon, 2013). These later data have been used to support this prediction, but since CifA is also the rescue protein (Chen et al., 2019; Shropshire et al., 2018) its presence in infected spermatheca is most easily explained by its relationship to rescue, not CI induction. It remains possible that Cif proteins are transferred, but the current evidence does not support this hypothesis (Beckmann and Fallon, 2013). Second, if the proteins are transferred, then maternal CifA must bind to the CI toxin to prevent function. *In vitro* biochemical assays reveal that CifA and CifB are capable of binding, but it remains unknown if they bind as a toxin complex to induce CI or if CifA binds to CifB in the embryo to rescue CI. Moreover, while CifB's Ulp1 domain is an *in vitro* deubiquitinase, CifA's binding to CifB does not inhibit deubiquitinase activity, suggesting that if binding is for the purpose of rescue it is not inhibiting CifB's biochemical function (Beckmann et al., 2017). As such, assays investigating if the Cif products are transferred to the embryo and where the proteins bind will inform the foundation of this hypothesis.

Moreover, The TA model traditionally states that the toxin and antidote are separate factors (Poinsot et al., 2003). However, our genetic understanding is that CifA is involved in both CI and rescue. There are two ways to sort-out this discrepancy while maintaining the key assumptions of the TA model. First, CifB may be the sole toxin but requires CifA even during spermatogenesis to prevent overly defective sperm (Beckmann et al., 2019a). For this to work, CifA is expected to degrade faster than CifB, leaving CifB alone to enter the egg as a toxin unless it binds to maternally derived CifA (Beckmann et al., 2019a). Alternatively, CifA and CifB could work together as a toxin complex that enters the embryo and is then rescued by maternally derived CifA.

The TA model as described above aims to explain unidirectional CI between *Wolbachia* infected and uninfected individuals. A modification of the TA model, called lock-and-key, expands the TA model to explain incompatibilities between *Wolbachia* strains. The lock-and-key model, like TA, proposes that a toxin is transferred from infected males to the embryo and will cause

embryonic death unless an antidote is supplied. Toxins in this case are called locks and antidotes are keys. The toxic lock is proposed to bind to or otherwise interfere with factors associated with proper embryonic development unless the antidote key is available to remove the lock. Bidirectional CI can then be explained by one strain carrying a set of locks and keys that are not compatible with the other strains' locks and keys because of differences in binding affinity. This model leveraged predictions of the mod/resc model that (i) strains can have multiple sets of locks and keys and that (ii) a key is more likely to bind to its associated lock than to a divergent lock. Indeed, *Wolbachia* exhibit considerable *cif* polymorphism (Bing et al., 2020; Bonneau et al., 2019; LePage et al., 2017; Lindsey et al., 2018b) and binding of CifA and CifB is strongest between cognate partners (Beckmann et al., 2017). However, the lingering questions with the TA model also apply with the lock-and-key model. Additionally, functional validation that divergent Cif proteins are functional, that they have differential impacts on the host, and contribute summatively to incompatibilities.

Finally, while HM-based and TA-based models are seemingly incompatible it is possible that both HM and TA type affects simultaneously occur in some *Wolbachia* strains. In fact, the goalkeeper model posits that the early and late embryonic defects associated with CI may be caused by different effects and that the ratio of these defects contributes to compatibility relationships between *Wolbachia* strains (Bossan et al., 2011). As such, one set of defects may be HM-based and the other TA-based. Moreover, since the titration-restitution model does not make predictions about the developmental timing of the male and female pronuclei, it is compatible with both mistiming and goalkeeper models. Indeed, titration-restitution can help explain mistiming through misregulation of host factors in a manner that leads to slowed development. Together, these models would help to explain the way that *Wolbachia* interacts with its host to cause CI and the downstream defects that ultimately culminate in embryonic lethality. Moreover, since these models are in essence HM-based models, they would be compatible with lock-and-key or TA-based models in the same way as described above. As such, evidence for one model does not necessarily exclude the possibility that both types of defects may be observable and that different *Wolbachia* strains may leverage HM- and TA- associated defects at different frequencies.

Conclusion

Wolbachia was first discovered in *Cu. pipiens* mosquitos in 1924 and was later linked to CI in 1973 (Hertig and Wolbach, 1924; Yen and Barr, 1973). Since then, CI has been reported in many insect orders where it can be used to rapidly spread itself to high frequencies in populations (Turelli, 1994; Weinert et al., 2015; Zug and Hammerstein, 2012). CI is associated with reproductive isolation (Bordenstein et al., 2001; Jaenike et al., 2006) and is being leveraged as a successful tool in the prevention of arboviral diseases that infect humans (O'Neill, 2018). Considerable effort has been made to untangle factors that influence CI strength (Layton et al., 2019; Reynolds and Hoffmann, 2002; Yamada et al., 2007), describe the embryonic and spermatogenic cytological defects CI causes (Ferree and Sullivan, 2006; Landmann et al., 2009), and link variation in host expression with CI phenotypes (Biwot et al., 2019; Liu et al., 2014; Zheng et al., 2011). Moreover, the last decade has seen a rapid expansion in our understanding of CI genetics (Beckmann et al., 2019c; Chen et al., 2019; LePage et al., 2017; Shropshire et al., 2018; Shropshire and Bordenstein, 2019), phylogenetics (Bing et al., 2020; LePage et al., 2017; Lindsey et al., 2018b), and mechanism (Beckmann et al., 2019c, 2017; Chen et al., 2019). Together, this significant body of literature has motivated models to explain how CI works (Beckmann et al., 2019a; Bossan et al., 2011; Poinsot et al., 2003; Shropshire et al., 2019). However, despite these advances, considerable work is still necessary to fully resolve how the CI genes interact with the host to cause CI, how CifA rescues CI, and how mechanisms may differ over phylogenetic landscapes.

Appendix A.

Chapter II supplementary information[§]

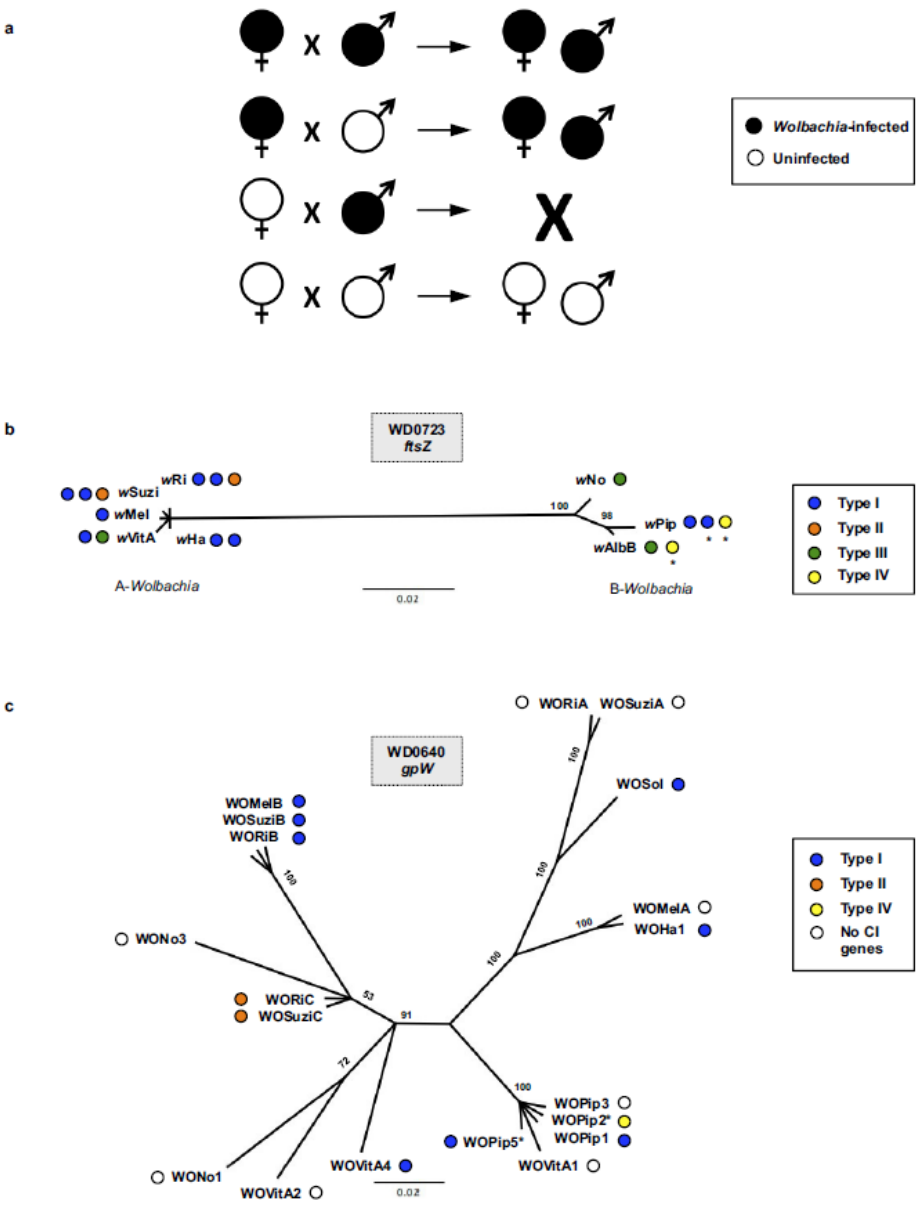

Figure A-1. CI and the evolution of *Wolbachia* and prophage WO genes.
a, The effect of parental *Wolbachia* infection on progeny viability and infection status. CI (embryonic inviability) occurs in crosses between *Wolbachia*-infected males and uninfected females. *Wolbachia*-infected females mated to infected males rescue the inviability. b, Bayesian phylogenies based on a 393-aa alignment of WD0723, the *w*Mel *ftsZ*

[§] This chapter is published in *Nature*, *543*(7644), 243-247 with Daniel LePage and Jason Metcalf as first authors. Jungmin On, Jessie Perlmutter, Dylan Shropshire, Emily Layton, Lisa Funkhouser-Jone, John Beckmann were co-authors. Seth Bordenstein was senior author.

gene, and its homologues, and (c) a 70-aa alignment of WD0640, the phage WO *gpW* gene, and its homologues. Trees are based on JTT+ G and CpRev+ I models of evolution, respectively, and are unrooted. Consensus support values are shown at the nodes. Asterisk indicates that the CI genes are not included in Figure II-1. The WOPip5 homologue is truncated while the WOPip2 and second *w*AlbB homologues are highly divergent from WD0632.

Figure A-2. WD0631/WD0632 homologues associatewith the eukaryotic association module in prophage WO regions.

CI gene homologues are labelled and coloured pink. Structural modules are labelled as baseplate, head, or tail. The WD0611–WD0621 label highlights a conserved gene cluster that is often associated with the CI genes. Only one phage haplotype is shown per *Wolbachia* strain when multiple copies of the same type are present.

a

WD0631-like % aa identity

	WORiB	WORiC	WOHa1	*w*No
WOMelB	99	46	67	31
WORiB		46	68	31
WORiC			44	33
WOHa1				31

WD0632-like % aa identity

	WORiB	WORiC	WOHa1	*w*No
WOMelB	99	30	62	29
WORiB		30	62	29
WORiC			31	36
WOHa1				30

b

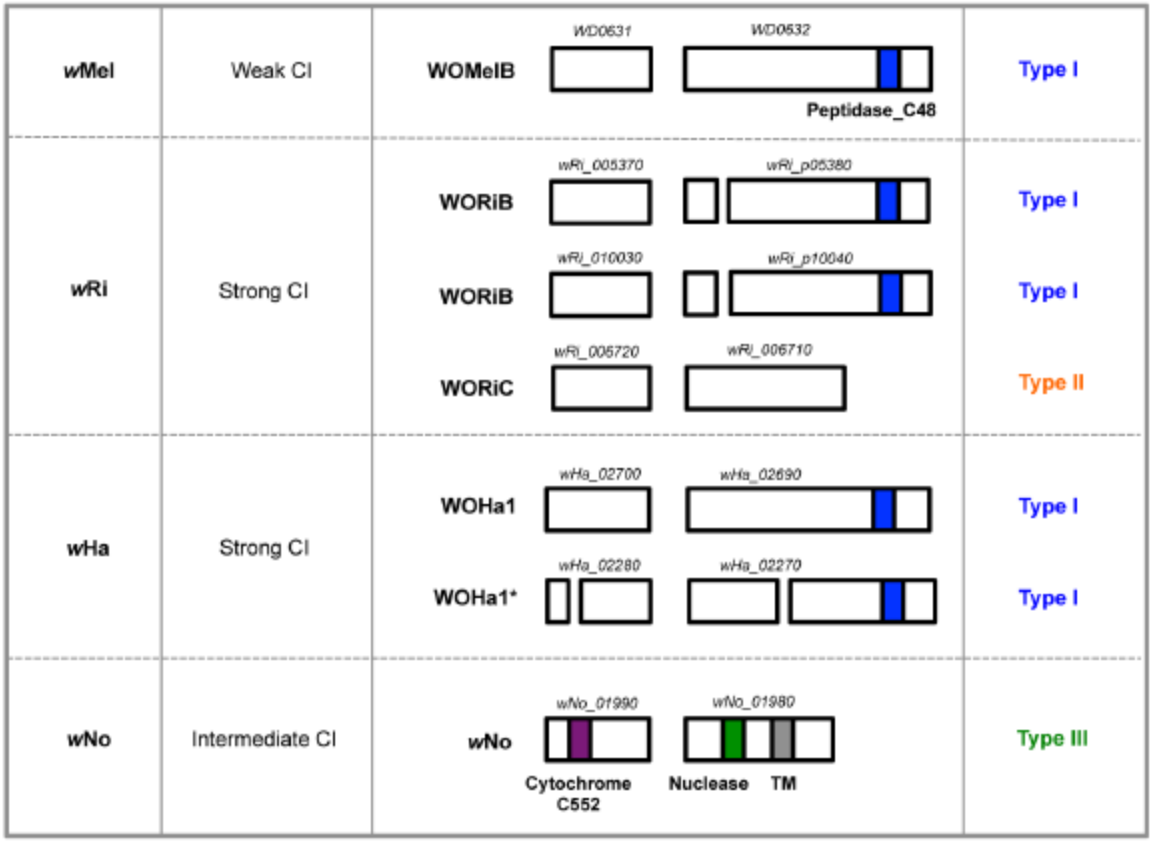

Figure A-3. *Wolbachia* CI patterns correlate with WD0631/WD0632 homologue similarity and copy number. a, The percentage aa identity between each WD0631/WD0632 homologue correlates with *Wolbachia* compatibility patterns. The only compatible cross, *w*Mel males × *w*Ri females, features close homology between WOMelB and WORiB. All other crosses are greater than 30% divergent and are bidirectionally incompatible. Each '% aa identity' value is based on the region of query coverage in a 1:1 BLASTp analysis. b, CI strength, protein architecture, and clade type are listed for each of the *Wolbachia* strains shown in Figure II-1d. Asterisk indicates the proteins are disrupted and not included in comparison analyses.

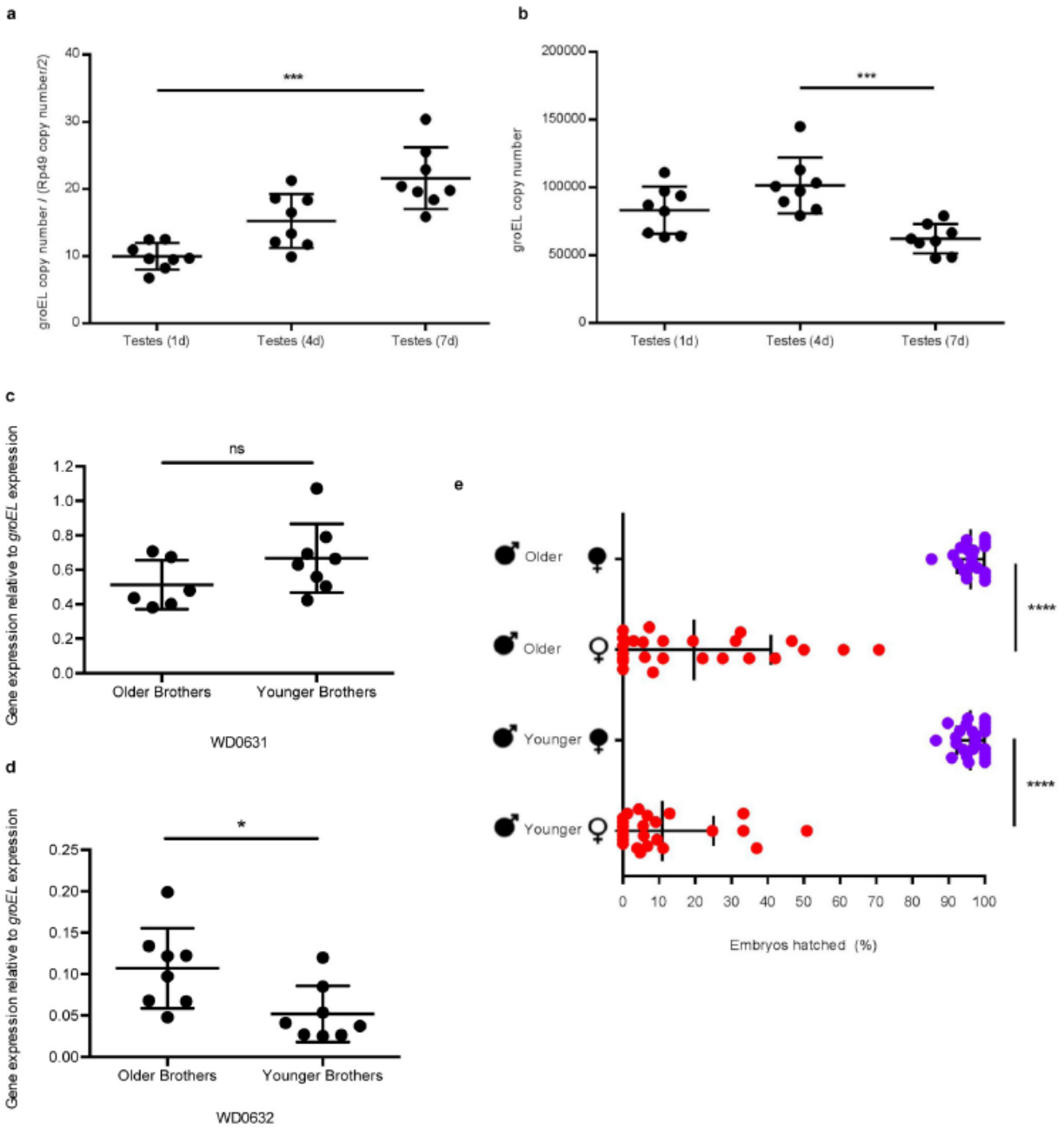

Figure A-4. *Wolbachia* titres, the male age effect, and the younger brother effect.
a, Relative *Wolbachia* titres in WT lines do not decrease with age. DNA copy number of the *w*Mel *groEL* gene is shown normalized to *D. melanogaster rp49* gene copy number in testes at the indicated ages. **b**, Absolute *Wolbachia* titres do not decrease from day 1 to day 7 males. **c**, **d**, In *w*Mel-infected males, WD0631 gene expression is equal between older (first day of emergence) and younger (fifth day of emergence) brothers while WD0632 gene expression is slightly higher in early emerging brothers. **e**, There is no statistical difference in CI penetrance between older and younger brothers. $n = 8$ for each group in **a**–**d**; $n = 19$–25 for each group in **e**. Bars, mean ± s.d. * $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$ by ANOVA with Kruskal–Wallis test and Dunn's multiple test correction for **a**, **b**, and **e**, and two-tailed Mann–Whitney *U*-test for **c** and **d**. Exact *P* values are provided in Supplementary able 7. These experiments were performed once.

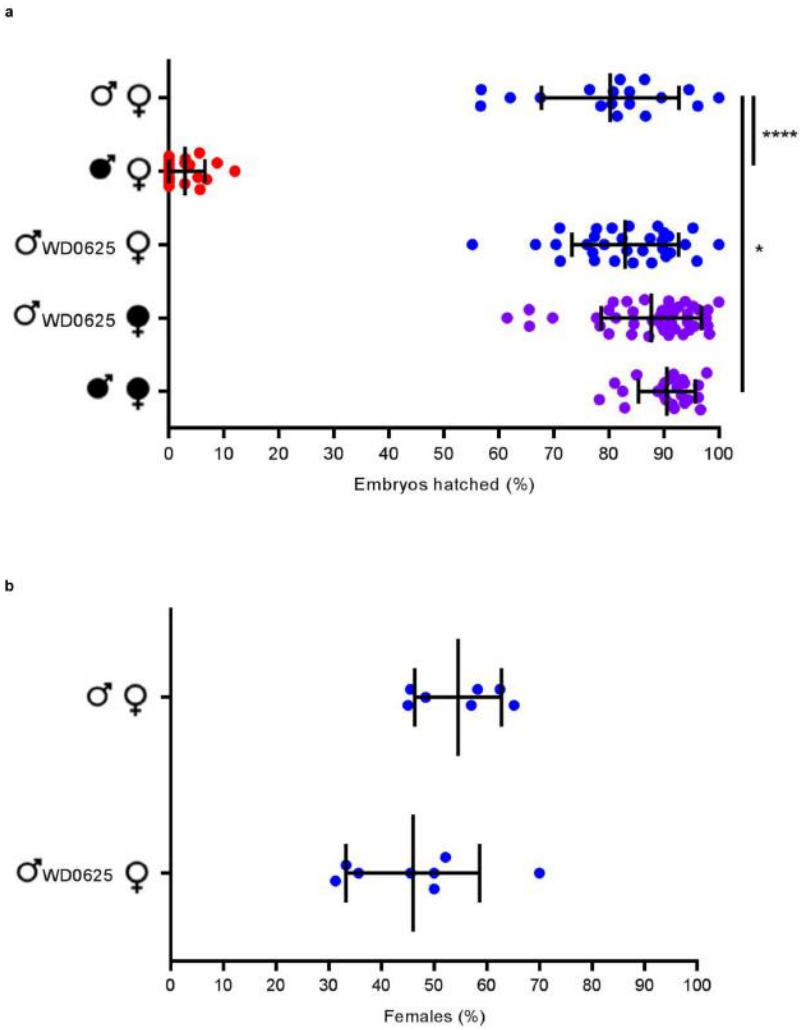

Figure A-5. WD0625 transgene expression does notinduce CI-like defects.
Expression of control gene WD0625 in 1-day-old uninfected males does not affect (**a**) embryo hatch rates or (**b**) sex ratios. Infection status is designated with filled symbols for a *w*Mel-infected parent or open symbols for an uninfected parent. Transgenic flies are labelled with their transgene to the right of their male/female symbol. Unlabelled symbols represent WT flies. Data points are coloured according to the type of cross: blue, no CI; red, a CI cross; purple, a rescue cross with *w*Mel-infected females. n = 18–47 for each group in **a**; n = 7–8 for **b**. Bars, mean ± s.d. * $P < 0.05$, *** $P < 0.001$ by ANOVA with Kruskal–Wallis test and Dunn's multiple test correction. Exact P values are provided in Table A-7. This experiment was replicated three times.

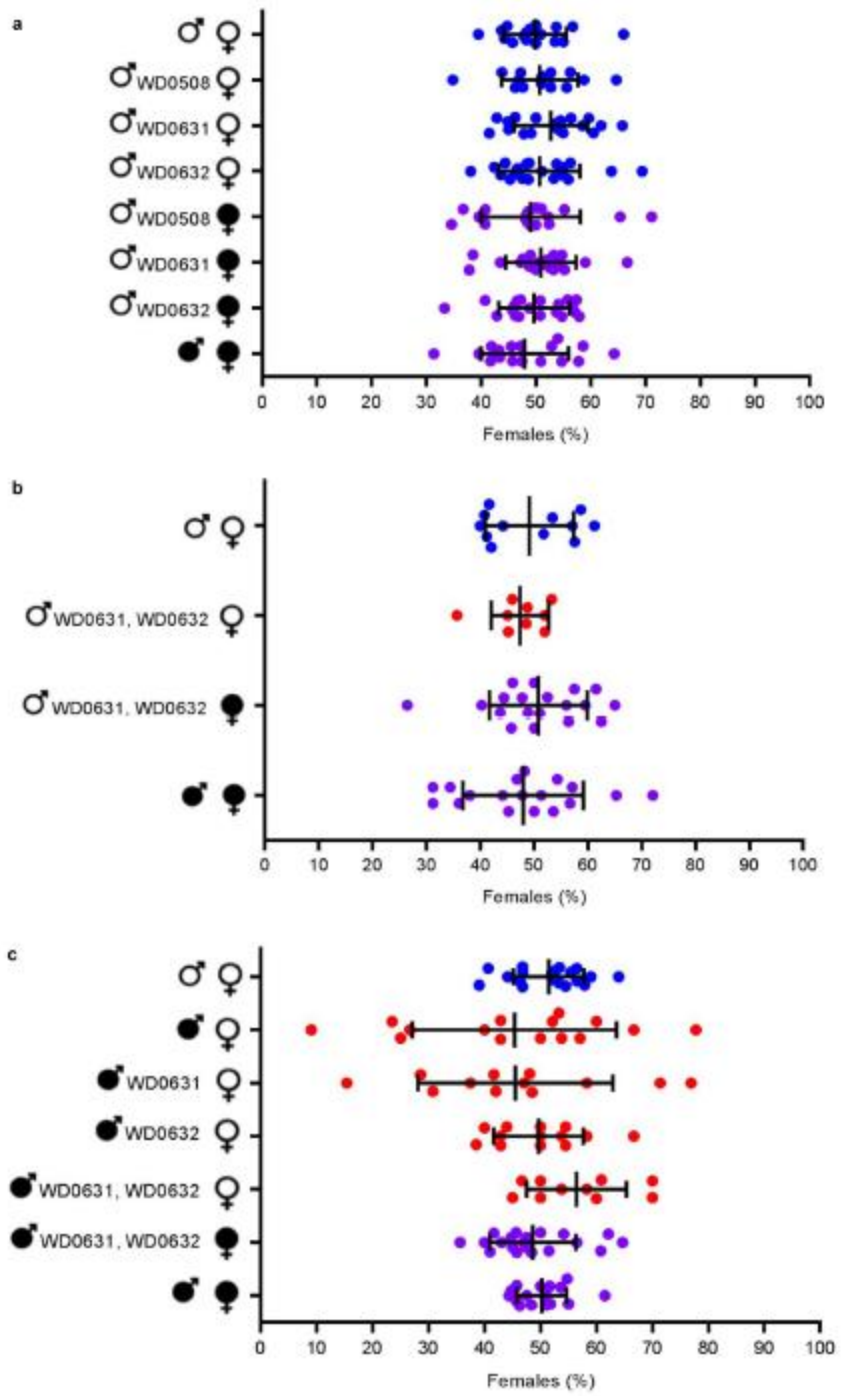

Figure A-6. Expression of transgenes does not alter sex ratios.
Graphs correspond to the same crosses as in Figure II-3. Infection status is designated with filled symbols for a *w*Mel-infected parent or open symbols for an uninfected parent. Transgenic flies are labelled with their transgene to the right of their gender symbol. Unlabelled gender symbols represent WT flies. Data points are coloured according to the type of cross: blue, no CI; red, a CI cross; purple, a rescue cross with *w*Mel-infected females. n = 10–36 for each group. Bars, mean ± s.d. Statistics include a Kruskal–Wallis tests and Dunn's multiple test corrections. The experiment in Figure A-6a, c was performed once, while that in Figure A-6b was performed twice.

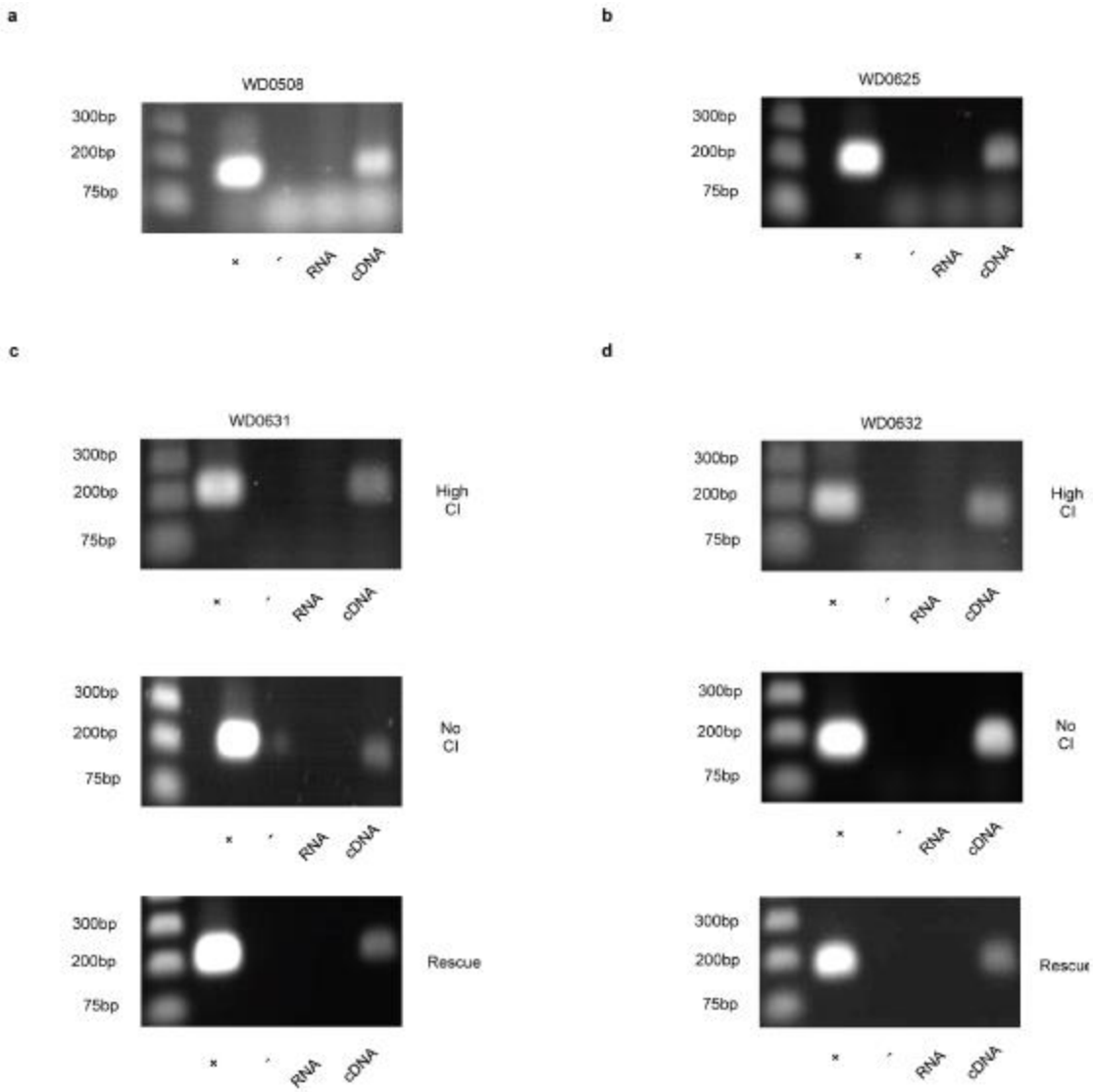

Figure A-7. Transgenes are expressed in testes.
a, b, WD0508 and WD0625 transgenes are expressed in testes as evident by PCR performed against cDNA generated from dissected males used in Figure II-3a and Figure A-5a, respectively. c, d, WD0631 and WD0632 transgenes are expressed in the testes from transgenic males specifically inducing high CI, no CI, or rescued CI. Testes were removed from males used in a replicate of Figure II-3b. n = six pools of six pairs of testes, with representative image shown. This experiment was performed once.

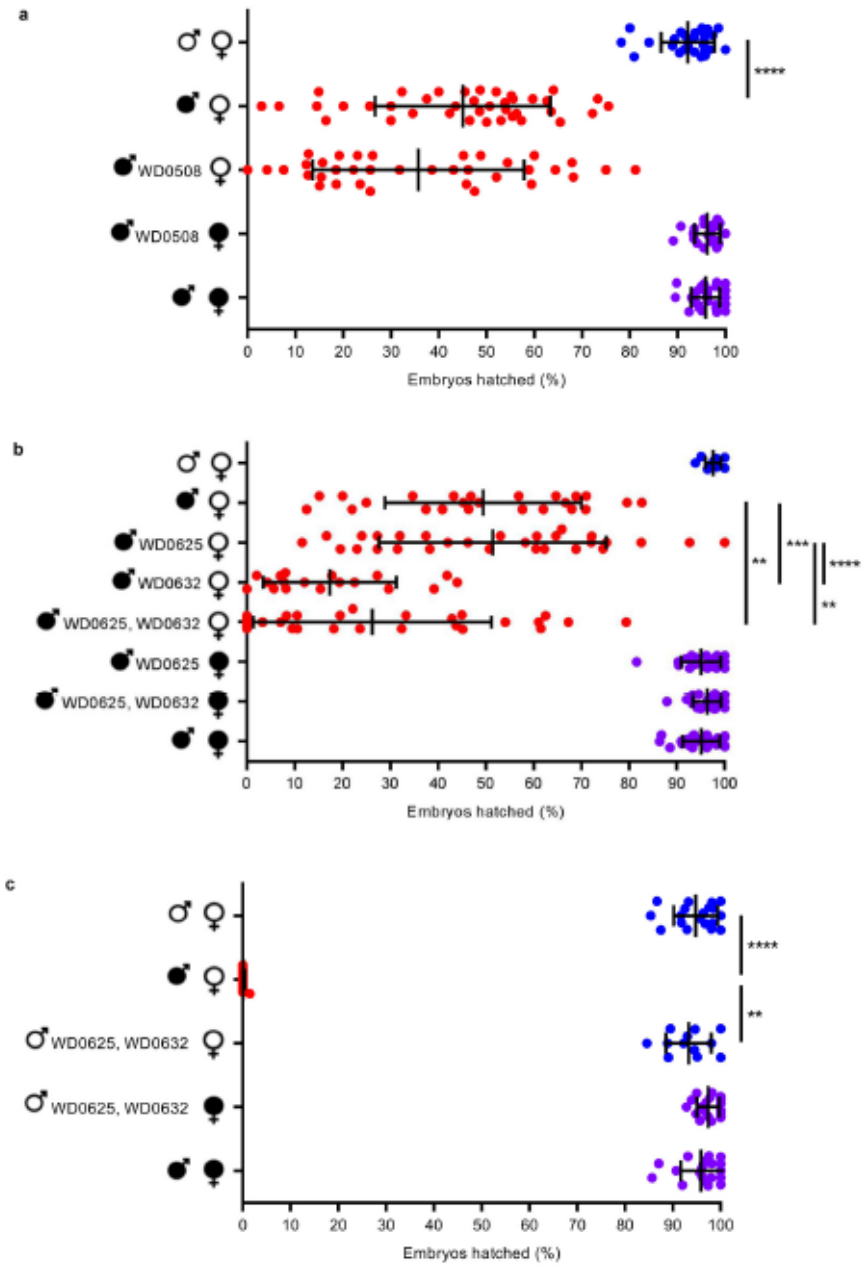

Figure A-8. Transgenic expression of WD0508, WD0625, and WD0625/WD0632 (*cifB*) does not enhance or induce CI.

a, The WD0508 transgene alone does not enhance CI in 2- to 4-day-old infected males. b, The WD0625 transgene alone does not enhance CI either; conversely, WD0632 does enhance CI as previously shown in Figure II-3c. The WD0625 transgene together with WD0632 does not enhance CI further than WD0632 alone. c, WD0625/WD0632 dual expression cannot induce CI in uninfected 1-day-old males. Infection status is designated with filled symbols for a *w*Mel-infected parent or open symbols for an uninfected parent. Transgenic flies are labelled with their transgene to the right of their male/female symbol. Unlabelled symbols represent WT flies. Data points are coloured according to the type of cross: blue, no CI; red, a CI cross; purple, a rescue cross with wMel-infected females. n = 12–44 for each group. Bars, mean ± s.d. ** P < 0.01, *** P < 0.001, **** P < 0.0001 by ANOVA with Kruskal–Wallis test and Dunn's multiple test correction. Exact P values are provided in Table A-7. These experiments were done twice (a, c), three times (b, WD0625, WD0632), or once (b, WD0625/WD0632).

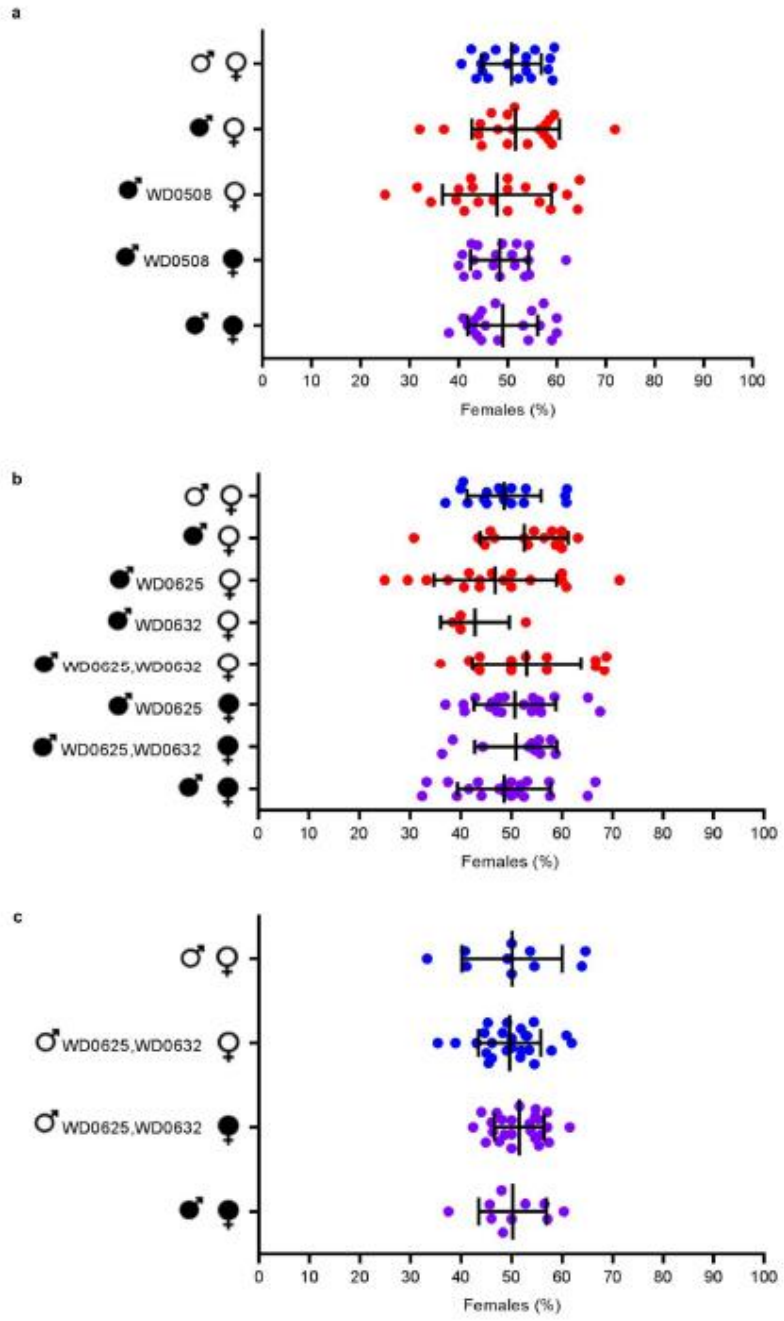

Figure A-9. Transgenic expression of control genes does not affect sex ratios.
All flies are from same crosses shown in Figure A-8, except for c, which comes from a replicate experiment. Infection status is designated with filled symbols for a *w*Mel-infected parent or open symbols for an uninfected parent. Transgenic flies are labelled with their transgene to the right of their male/female symbol. Unlabelled symbols represent WT flies. Data points are coloured according to the type of cross: blue, no CI; red, a CI cross; purple, a rescue cross with wMel-infected females. n = 4–27 for each group. Bars, mean ± s.d. Statistics performed by ANOVA with Kruskal–Wallis test and Dunn's multiple test correction. These experiments were done twice (b) or once (a, c).

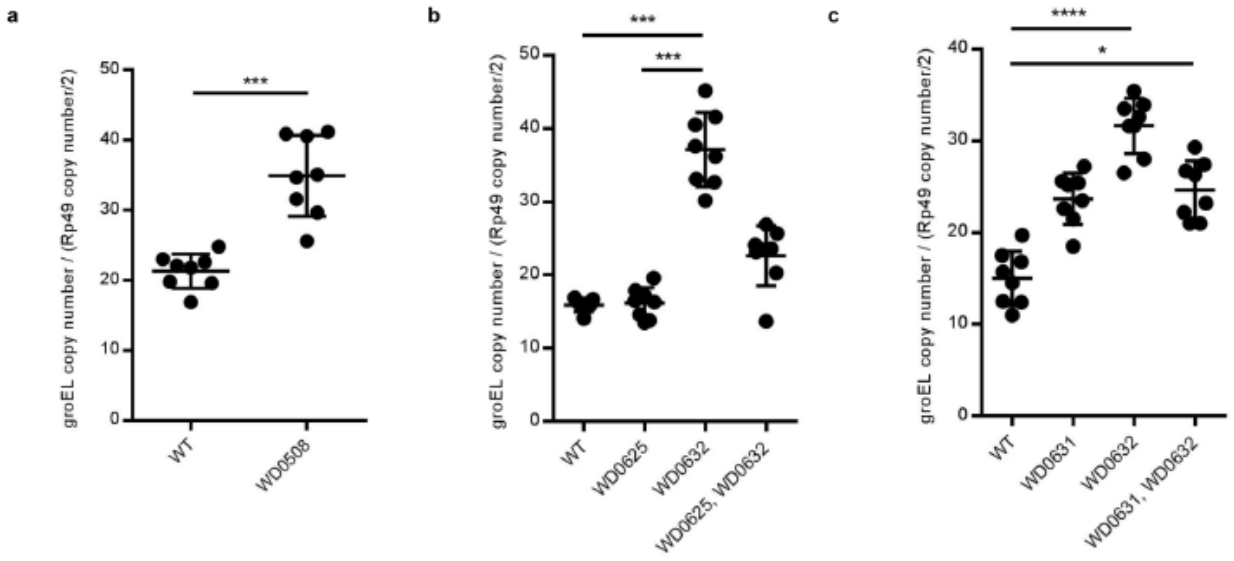

Figure A-10. There is variation in *Wolbachia* titres in transgenic lines.
a–c, Relative *Wolbachia* titres are higher in WD0508, WD0631, and WD0632 (*cifB*) transgenic lines than in WT lines. This does not occur in the WD0625 transgenic line, nor does there appear to be an additive effect. DNA copy number of the *w*Mel *groEL* gene is shown normalized to *D. melanogaster rp49* gene copy number in testes of the indicated strains. n = 8 independent pools of 15 pairs of testes for each group. Bars, mean ± s.d. * $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$ for two-tailed Mann–Whitney U-test (a) and Kruskal–Wallis test with Dunn's multiple test correction (b, c). Exact P values are provided in Table A-7. These experiments were done once.

Table A-1. Core CI Genome.
This table lists all unique *w*Mel genes from gene families present in CI-inducing strains, but absent from non-CI strains.

***w*Mel Locus Tag**	**Gene Description**
WD0035	ankyrin repeat-containing protein
WD0038	Protein tolB
WD0046	reverse transcriptase, interruption-N
WD0049	hypothetical protein
WD0056	major facilitator family transporter
WD0061	hypothetical protein
WD0064	Pyridoxine 5'-phosphate synthase
WD0069	hypothetical protein
WD0074	hypothetical protein
WD0077	hypothetical protein
WD0078	hypothetical protein
WD0079	hypothetical protein
WD0092	DNA processing chain A
WD0099	multidrug resistance protein
WD0100	sugE protein
WD0131	hypothetical protein
WD0139	TenA family transcription regulator
WD0140	TenA family transcription regulator
WD0168	major facilitator family transporter
WD0200	hypothetical protein
WD0208	hypothetical protein
WD0211	hypothetical protein
WD0214	hypothetical protein
WD0217	phage uncharacterized protein

WD0231	hypothetical protein
WD0234	hypothetical protein
WD0240	transposase, IS5 family, degenerate
WD0255	transcriptional regulator, putative
WD0257	DNA repair protein RadC, truncation
WD0258	Reverse transcriptase, frame shift
WD0274	hypothetical protein
WD0278	Prophage LambdaW1, minor tail protein Z
WD0279	hypothetical protein
WD0281	hypothetical protein
WD0282	prophage LambdaW1, baseplate assembly protein W, putative
WD0283	prophage LambdaW1, baseplate assembly protein J, putative
WD0284	hypothetical protein
WD0285	Prophage LambdaW1, ankyrin repeat protein
WD0286	ankyrin repeat-containing prophage LambdaW1
WD0288	prophage LambdaW1, site-specific recombinase resolvase family protein
WD0315	hypothetical protein
WD0324	hypothetical protein
WD0329	transposase, IS5 family, degenerate
WD0336	transposase, IS5 family, degenerate
WD0338	hypothetical protein
WD0345	RND family efflux transporter MFP subunit
WD0376	potassium uptake protein TrKH, frame shift
WD0382	hypothetical protein
WD0385	ankyrin repeat-containing protein
WD0396	reverse transcriptase, truncation
WD0407	Na+/H+ antiporter, putative
WD0426	hypothetical protein
WD0431	glycosyl transferase, group 2 family protein
WD0447	phage prohead protease
WD0458	HK97 family phage major capsid protein
WD0472	AAA family ATPase
WD0480	hypothetical protein
WD0481	hypothetical protein
WD0482	SPFH domain-containing protein/band 7 family protein
WD0483	M23/M37 peptidase domain-containing protein
WD0498	ankyrin repeat-containing protein
WD0501	surface antigen-related protein
WD0506	Reverse transcriptase, frame shift
WD0507	DNA repair protein RadC, truncation
WD0508	transcriptional regulator, putative
WD0515	reverse transcriptase, interruption-C
WD0518	reverse transcriptase, interruption-N
WD0538	reverse transcriptase, truncation
WD0604	hypothetical protein
WD0606	Reverse transcriptase, frame shift
WD0623	transcriptional regulator, putative
WD0624	conserved domain protein, frame shift
WD0625	DNA repair protein RadC, putative
WD0626	transcriptional regulator, putative
WD0628	hypothetical protein
WD0631	hypothetical protein
WD0632	hypothetical protein (Ulp1/Peptidase_C48)

WD0633	Prophage LambdaW5, ankyrin repeat domain protein
WD0634	prophage LambdaW5, site-specific recombinase resolvase family protein
WD0636	ankyrin repeat-containing prophage LambdaW1
WD0638	hypothetical protein
WD0639	prophage LambdaW5, baseplate assembly protein J, putative
WD0640	prophage LambdaW5, baseplate assembly protein W, putative
WD0641	hypothetical protein
WD0642	prophage LambdaW5, baseplate assembly protein V
WD0643	hypothetical protein
WD0644	Prophage LambdaW5, minor tail protein Z
WD0645	reverse transcriptase, truncation
WD0686	hypothetical protein
WD0693	reverse transcriptase, putative
WD0696	hypothetical protein
WD0702	hypothetical protein
WD0713	hypothetical protein
WD0718	conserved hypothetical protein, truncated
WD0721	Mg chelatase-related protein
WD0724	hypothetical protein
WD0730	phosphatidylglycerophosphatase A, putative
WD0733	hypothetical protein
WD0748	hypothetical protein
WD0749	transposase, IS5 family, degenerate
WD0750	PQQ repeat-containing protein
WD0764	hypothetical protein
WD0787	araM protein
WD0790	hypothetical protein
WD0818	hypothetical protein
WD0823	hypothetical protein
WD0826	hypothetical protein
WD0834	conserved hypothetical protein, degenerate
WD0835	hypothetical protein
WD0874	transposase, truncated
WD0875	IS5 family transposase
WD0880	coenzyme PQQ synthesis protein C, putative
WD0882	FolK
WD0883	dihydropteroate synthase, putative
WD0884	dihydrofolate reductase
WD0887	DNA repair protein RadA
WD0901	transposase, IS110 family, degenerate
WD0903	transposase, IS5 family, degenerate
WD0908	transposase, degenerate
WD0911	transposase, IS5 family, degenerate
WD0914	hypothetical protein
WD0932	IS5 family transposase
WD0935	transposase, IS5 family, interruption-C
WD0941	transposase, degenerate
WD0947	IS5 family transposase
WD0958	hypothetical protein
WD0964	hypothetical protein
WD0975	hypothetical protein
WD0995	reverse transcriptase
WD0999	hypothetical protein

WD1002	hypothetical protein
WD1012	HK97 family phage portal protein
WD1015	hypothetical protein
WD1016	phage uncharacterized protein
WD1041	surface protein-related protein
WD1047	sodium/alanine symporter family protein
WD1052	folylpolyglutamate synthase
WD1069	hypothetical protein
WD1073	N-acetylmuramoyl-L-alanine amidase
WD1091	tRNA (guanine-N(7)-)-methyltransferase
WD1114	LipB
WD1118	hypothetical protein
WD1126	hypothetical protein
WD1131	conserved hypothetical protein, degenerate
WD1132	phage uncharacterized protein
WD1138	reverse transcriptase, putative
WD1159	Pyridoxine/pyridoxamine 5'-phosphate oxidase
WD1160	ComEC/Rec2 family protein
WD1161	hypothetical protein
WD1162	ribosomal large subunit pseudouridine synthase D
WD1163	diacylglycerol kinase
WD1175	hypothetical protein
WD1179	hypothetical protein
WD1204	TPR domain-containing protein
WD1212	16S ribosomal RNA methyltransferase RsmE
WD1218	ParB family protein
WD1242	hypothetical protein
WD1272	hypothetical protein
WD1310	hypothetical protein
WD1320	multidrug resistance protein D
WD1321	hypothetical protein

Table A-2. Genes divergent in *w*Au.
*w*Mel genes that are absent or divergent in *w*Au as identified in Figure II-1b.

***w*Mel Locus Tag**	**Gene Description**
WD0019	transcription antitermination protein NusG, putative
WD0022	ribosomal protein L10
WD0034	PAZ Zwille/Arganaut/Piwi/ SiRNA binding domain
WD0072	hypothetical protein
WD0205	hypothetical protein
WD0244	hypothetical protein
WD0255	transcriptional regulator, putative
WD0256	hypothetical protein
WD0257	DNA repair protein RadC, truncation
WD0289	hypothetical protein
WD0297	hypothetical protein
WD0311	hypothetical protein
WD0320	trigger factor, putative
WD0349	hypothetical protein
WD0363	hypothetical protein
WD0366	hypothetical protein
WD0367	hypothetical protein
WD0369	hypothetical protein

WD0389	conserved hypothetical protein
WD0424	hypothetical protein
WD0449	hypothetical protein
WD0508	transcriptional regulator, putative
WD0512	ankyrin repeat domain protein
WD0553	hypothetical protein
WD0576	hypothetical protein
WD0577	hypothetical protein
WD0578	hypothetical protein
WD0579	hypothetical protein
WD0598	hypothetical protein
WD0607	hypothetical protein
WD0623	transcriptional regulator, putative
WD0624	conserved domain protein, authentic frameshift
WD0625	DNA repair protein RadC, putative
WD0626	transcriptional regulator, putative
WD0631	hypothetical protein
WD0632	hypothetical protein (Ulp1/Peptidase_C48)
WD0633	prophage LambdaWp5, ankyrin repeat domain protein
WD0704	hypothetical protein
WD0723	cell division protein FtsZ
WD0746	hypothetical protein
WD0747	hypothetical protein
WD0806	hypothetical protein
WD0808	hypothetical protein
WD0809	hypothetical protein
WD0836	hypothetical protein
WD0837	hypothetical protein
WD0840	hypothetical protein
WD0850	rpsU-divergently transcribed protein
WD0854	membrane protein, putative
WD0877	hypothetical protein
WD0940	hypothetical protein
WD0946	hypothetical protein
WD0971	hypothetical protein
WD1038	hypothetical protein
WD1151	citrate synthase
WD1260	hypothetical protein
WD1287	hypothetical protein
WD1291	hypothetical protein
WD1311	Glycoside hydrolase 24
WD1313	conserved domain protein

Table A-3. *w*VitA transcriptome.
*w*Mel homologs that are expressed (as measured by RNA) in the ovaries of *w*VitA-infected *Nasonia vitripennis* .

***w*VitA Feature ID**	***w*Mel Locus Tag**	**Gene Description**	**12.1 wVitA - Mean RPKM**	**12.1 average # of reads mapped to gene**	**12.1 wVitA - 12_1_2 M trimmed RNA-Seq - Total**	**12.1 wVitA - 12_1_3 M RNA-Seq - Total gene reads**	**12.1 wVitA - 12_1_1 M trimmed RNA-Seq - Total**

					gene reads		**gene reads**
gwv_835	EXCLUDED	16S rRNA	351983.2481	66816.0	148928	22578	28942
gwv_1180	EXCLUDED	23S rRNA	157771.9274	53996.0	118025	18597	25366
gwv_528	WD1063	outer surface protein	835.3599645	33.0	19	47	33
gwv_664	WD0838	hypothetical protein	216.8355074	26.3	14	35	30
gwv_400	WD0745	putative outer membrane protein	205.5588118	24.0	17	35	20
rpoC	WD0024	DNA-directed RNA polymerase	91.44973092	20.7	9	38	15
gwv_424	WD0722	hypothetical protein	723.0305001	19.0	11	25	21
gwv_788	WD0906	S1 RNA binding domain protein	122.5244426	14.3	4	27	12
groEL	WD0307	chaperonin GroEL	140.8718458	13.7	10	23	8
gwv_846	WD1236	DNA/RNA helicase	183.4533269	13.0	8	20	11
gwv_141	WD0632	hypothetical protein (Ulp1/Peptidase_C48)	45.47043156	11.3	4	16	14
gwv_138	WD0292	ankryin repeat protein	52.0888594	10.3	3	13	15
fusA	WD0016	translation elongation factor G	81.2935939	9.7	5	13	11
gwv_1314	WD0950	uncharacterised protein family UPF0005	207.2377169	8.3	4	13	8
gwv_1093	WD0147	tetratricopeptide repeat family protein	17.79714113	7.3	4	13	5
gwv_142	WD0631	hypothetical protein	71.02829671	7.0	7	8	6
gyrA	WD1202	DNA gyrase	46.24634563	6.7	3	13	4
dnaK	WD0928	chaperone protein DnaK	55.20689465	6.3	3	5	11
gwv_968	WD0039	metallo-beta-lactamase superfamily protein	62.20883237	6.3	5	8	6
gwv_219	WD0550	ankryin repeat protein	104.550262	6.0	4	6	8
gwv_874	WD1249	sodium/hydrogen exchanger family protein	63.25273067	5.7	2	11	4
gwv_971	WD0041	putative membrane protein	98.05217219	5.7	4	7	6
gwv_889	WD1278	hypothetical protein	40.99214546	5.7	1	11	5
gwv_294	WD1071	cytochrome b	68.14464273	5.3	5	8	3
gwv_726	WD1064	RNA polymerase sigma-32 factor	103.3470918	5.0	3	8	4

gwv_848	WD1238	fructose-bisphosphate aldolase class 1	97.14626625	5.0	3	8	4
gwv_127	WD0337	hypothetical protein	70.90369473	5.0	3	3	9
rpsD			140.1432593	4.7	2	8	4
gwv_193			51.53566326	4.7	2	7	5
lpdA (gwv.assembly.1 388700..390089)			47.34146102	4.7	5	4	5
ftsZ			63.02213521	4.3	2	5	6
clpX			61.55998273	4.3	2	7	4
gwv_592			49.87489713	4.3	0	7	6
gwv_1267			228.5442649	4.0	3	6	3
gwv_1351			98.71359598	4.0	1	6	5
gwv_134			87.53953909	4.0	1	8	3
rpoA			68.14353225	4.0	2	7	3
agcS (gwv.assembly.1 617423..618869)			45.85405828	4.0	2	3	7
typA			38.00061768	4.0	2	5	5
lon			28.79549911	4.0	2	5	5
gwv_27			65.63670247	3.7	2	4	5
gwv_865			60.59688572	3.7	2	6	3
iscS			51.15879624	3.7	3	8	0
sucC			51.03343902	3.7	2	2	7
gwv_660			47.49917482	3.7	1	7	3
dnaX			41.74839386	3.7	2	7	2
clpB			29.67965234	3.7	0	9	2
gwv_46			22.57157345	3.7	2	4	5
gyrB			21.39989203	3.7	4	3	4
gwv_603			184.7219284	3.3	0	5	5
gwv_275			75.59102504	3.3	3	4	3
nusG			70.68705809	3.3	1	3	6
gwv_734			67.32779895	3.3	0	7	3
rho			52.5016369	3.3	0	6	4
gwv_837			46.44631685	3.3	1	2	7
gwv_1163			36.66487753	3.3	3	4	3
hslU			32.09423701	3.3	5	4	1
rplL			144.584441	3.0	1	6	2
rpoZ			123.1224335	3.0	2	4	3
gwv_868			114.347465	3.0	1	4	4
rplQ			109.5519599	3.0	2	2	5
gwv_38			77.90440588	3.0	1	2	6
gwv_549			77.50677716	3.0	2	5	2
rpsB			66.31737504	3.0	0	2	7
gwv_361			64.9017479	3.0	2	4	3
gwv_407			44.42206307	3.0	1	5	3
gwv_631			44.21350879	3.0	1	5	3
lysS			37.00384036	3.0	1	5	3

atpA			34.54520243	3.0	1	3	5
ctaD			32.3881063	3.0	2	4	3
sdhA			26.91908739	3.0	2	3	4
gwv_333			26.0913448	3.0	0	6	3
gwv_1345			25.21412951	3.0	1	5	3
gwv_560			169.7933214	2.7	1	3	4
gwv_263			83.02918988	2.7	2	4	2
rplA			79.54918497	2.7	0	3	5
gwv_417			72.92232064	2.7	0	6	2
gwv_517			61.11517361	2.7	2	3	3
sucD			55.92513437	2.7	1	4	3
atpB			54.38477172	2.7	2	1	5
gwv_774			53.70391551	2.7	0	5	3
trpS			50.65915791	2.7	1	5	2
dnaN			47.36465385	2.7	1	2	5
dnaA			43.46751424	2.7	0	8	0
hflK			41.27488537	2.7	2	3	3
gwv_293			36.43743379	2.7	2	2	4
gwv_390			35.89644056	2.7	1	7	0
gwv_96			35.80675485	2.7	1	3	4
gwv_314			31.5950271	2.7	1	6	1
gwv_866			30.04184694	2.7	2	3	3
gwv_695			28.82698438	2.7	0	5	3
hscA			26.42417352	2.7	1	4	3
gwv_485			23.42918111	2.7	1	4	3
sucA			20.3494561	2.7	1	7	0
gwv_332			20.28271617	2.7	0	3	5
acnA			19.57018513	2.7	1	5	2
infB			19.06993006	2.7	2	4	2
pheT			14.27942126	2.7	4	3	1
gwv_1352			185.0654362	2.3	0	4	3
rplX			150.1474294	2.3	0	4	3
rpsL			93.88015955	2.3	2	2	3
rpsG			86.4082733	2.3	1	3	3
rplI			80.1994963	2.3	0	2	5
gwv_873			66.73614153	2.3	1	1	5
talC			62.08394103	2.3	1	2	4
gwv_989			36.52952064	2.3	1	5	1
gwv_415			34.33779635	2.3	3	4	0
gwv_227			32.30215824	2.3	1	3	3
tig			32.28018893	2.3	1	4	2
agcS (gwv.assembly.1 618891..620235)			32.06402695	2.3	1	4	2
purB			30.36431431	2.3	2	4	1
atpD			30.17790772	2.3	1	4	2
guaB			29.97354912	2.3	0	2	5
gwv_382			29.87270914	2.3	2	5	0
gwv_976			29.05271817	2.3	1	5	1
gwv_1043			28.97535789	2.3	2	2	3
thrS			26.80496563	2.3	0	6	1

gwv_484			23.90024428	2.3	2	2	3
gwv_404			21.71444884	2.3	3	2	2
gwv_1215			14.59109858	2.3	3	1	3
yaeT			13.76564423	2.3	3	3	1
gwv_1203			8.728108414	2.3	1	3	3
gwv_113			166.4780621	2.0	1	2	3
gwv_254			113.4160354	2.0	0	4	2
gwv_953			80.67937801	2.0	1	3	2
gwv_797			61.19193632	2.0	1	2	3
gwv_308			53.09546078	2.0	2	2	2
gwv_1030			49.53555013	2.0	2	4	0
gwv_1015			43.28419185	2.0	1	3	2
gwv_109			43.05287818	2.0	1	4	1
hemH			42.43927849	2.0	0	3	3
gwv_178			38.01239925	2.0	1	3	2
gwv_580			37.95175012	2.0	0	6	0
gwv_344			37.48356021	2.0	0	1	5
gwv_283			35.45854451	2.0	0	2	4
gwv_939			32.85282152	2.0	1	3	2
secY			29.20674777	2.0	1	5	0
recA			28.95765486	2.0	2	3	1
gwv_942			28.64511776	2.0	0	4	2
gwv_688			25.70804272	2.0	2	3	1
nusA			24.56383904	2.0	0	2	4
purD			24.31350267	2.0	2	3	1
proS			23.09589783	2.0	2	2	2
gidA			20.92170359	2.0	0	3	3
gwv_888			19.61903309	2.0	1	4	1
gatA			18.83715191	2.0	2	1	3
pyrG			14.35375888	2.0	3	1	2
gwv_582			10.92713149	2.0	1	2	3
rpsO			131.7001889	1.7	0	4	1
gwv_321			99.04724952	1.7	0	4	1
rplT			89.3934139	1.7	0	2	3
rpsJ			87.43040243	1.7	1	2	2
gwv_42			61.8400838	1.7	1	3	1
gwv_1007			51.14594978	1.7	1	4	0
gwv_651			51.01612357	1.7	3	1	1
gwv_274			48.26543126	1.7	1	3	1
gwv_1028			44.36184823	1.7	0	3	2
gwv_348			43.85916772	1.7	1	1	3
gwv_722			43.85194193	1.7	0	3	2
virB9 (gwv.assembly.1 894720..895515)			43.19002583	1.7	0	3	2
gwv_90			42.49899713	1.7	0	4	1
pdxJ			39.64005534	1.7	1	2	2
gwv_561			37.51358601	1.7	1	1	3
rplB			35.97968512	1.7	1	3	1
hflC			35.85489263	1.7	1	4	0
gwv_792			35.1693724	1.7	1	2	2

gwv_1022			30.8462831	1.7	2	2	1
trxB			29.88834843	1.7	1	2	2
gwv_414			28.26750691	1.7	2	1	2
gwv_844			26.93194874	1.7	0	4	1
gwv_630			26.92838641	1.7	0	2	3
gwv_842			26.79606019	1.7	0	2	3
gwv_288			25.32158594	1.7	0	3	2
gatB			25.28421348	1.7	0	4	1
gwv_83			25.05546538	1.7	2	3	0
purF			24.77349966	1.7	0	3	2
gwv_81			24.08659211	1.7	1	1	3
gwv_191			22.98188052	1.7	1	4	0
gwv_1220			21.55524675	1.7	0	4	1
recJ			20.6633055	1.7	0	4	1
gwv_392			20.52189061	1.7	2	1	2
argD (gwv.assembly.1 846447..847623)			19.90844292	1.7	2	2	1
ispG			19.38251139	1.7	1	0	4
gwv_320			19.38152429	1.7	2	0	3
gwv_686			18.02800146	1.7	1	1	3
gwv_575			17.60571781	1.7	1	3	1
gwv_909			17.34144976	1.7	0	3	2
pyrC			16.9654557	1.7	2	2	1
pheS			16.4189823	1.7	3	1	1
gwv_653			16.259985	1.7	1	0	4
gwv_285			15.48849844	1.7	1	2	2
gwv_698			12.68960914	1.7	2	2	1
gwv_475			11.36042875	1.7	1	1	3
tkt			10.51338535	1.7	2	1	2
alaS			9.426346656	1.7	1	0	4
gwv_1065			6.423485127	1.7	2	0	3
ppdK			6.265137982	1.7	3	1	1
rpmF			137.7242785	1.3	0	1	3
rpsU			133.4409155	1.3	0	2	2
rpmI			107.0956217	1.3	1	2	1
rpsS			99.78843877	1.3	0	3	1
gwv_594			97.17979713	1.3	0	2	2
gwv_727			72.68732793	1.3	0	2	2
gwv_262			66.5071607	1.3	1	1	2
gwv_679			62.36777423	1.3	0	3	1
hscB			61.23658449	1.3	0	2	2
rplP			60.8781231	1.3	0	1	3
bfr			59.62202317	1.3	0	3	1
gwv_1319			59.20888302	1.3	0	2	2
rplS			54.3669647	1.3	1	1	2
gwv_862			48.84407551	1.3	0	1	3
gwv_955			48.56173982	1.3	0	1	3
gwv_343			44.21674204	1.3	0	1	3
ssb			43.35593387	1.3	1	1	2
gwv_1075			43.19102095	1.3	0	2	2

gwv_714			42.54805933	1.3	1	1	2
coxB			37.17608503	1.3	0	3	1
gwv_1342			35.2412702	1.3	0	3	1
gwv_637			34.78809858	1.3	0	2	2
ribB			33.09293503	1.3	1	1	2
gwv_934			30.34075579	1.3	1	0	3
gwv_824			29.9137011	1.3	1	1	2
dnaQ			29.526886	1.3	1	1	2
tsf			29.27240762	1.3	0	1	3
gwv_477			27.0925495	1.3	0	2	2
rseP			25.41528601	1.3	0	3	1
nuoD			24.68724397	1.3	0	3	1
gwv_56			23.12171142	1.3	0	3	1
gltA			22.73357718	1.3	0	3	1
gwv_1194			22.39272091	1.3	1	2	1
gwv_1332			22.03265835	1.3	2	2	0
gwv_771			21.99285089	1.3	1	2	1
gwv_1020			21.6070114	1.3	1	2	1
gwv_281			19.93161697	1.3	3	0	1
guaA			19.23082923	1.3	0	4	0
fabF			18.70037322	1.3	1	3	0
gwv_860			17.51565738	1.3	3	1	0
xerD			16.10727695	1.3	2	1	1
gwv_372			15.36175487	1.3	0	2	2
gwv_959			13.88282816	1.3	0	2	2
gwv_926			11.78564258	1.3	1	2	1
secA			10.079528	1.3	0	2	2
gwv_40			9.886742753	1.3	2	1	1
gwv_525			9.753962639	1.3	2	0	2
murE			9.242589841	1.3	2	0	2
gwv_1096			9.145280562	1.3	1	3	0
glmS			8.730303321	1.3	2	1	1
uvrB			8.332223452	1.3	2	1	1
gwv_1178			7.999888686	1.3	2	0	2
uvrA			7.350040292	1.3	1	1	2
gwv_21			6.835890749	1.3	1	2	1
gwv_136			5.813095797	1.3	0	2	2
carB			4.710483658	1.3	2	1	1
glyS			4.545312084	1.3	3	0	1
gwv_933		acrB/AcrD/AcrF family protein	2.856484571	0.7	1	1	0

Table A-4. *w*Pip proteome.
*w*Mel homologs detected at the protein level in *w*Pip (Buckeye)-infected *C. pipiens* ovaries.

wMel Locus Tag	**Abundance**	**Name**	**Species**	**Accession #**	**TPS > (95%)**	**% Cov**	**most abundant peptide (95%)**	**function**	**Unique Peptides > (95%)**
WD1063	42	surface antigen Wsp	wPip	gi|190571332	128	53.5	LQYNGEVLPFK	Cell envelope biogenesis/Outer	31

								membrane	
absent	29	Putative membrane protein	wPip	gi|190570988	68	48.8	ASQIEEVNQGVLNACVK	Cell envelope biogenesis/Outer membrane	20
WD0928	15	chaperone protein dnak (hsp70)	wPip	gi|190570602	32	38	IINEPTAAALAYGLDKK	Protein modification/degradation/Chaperones	19
WD0307	12	chaperonin groEL, 60 kDa	wPip	gi|190570503	94	62.7	EMLEDIAALTGAK	Protein modification/degradation/Chaperones	50
WD0655	10	ATP synthase F1, alpha subunit	wPip	gi|190571573	75	16	VVDALGNAIDGKGEIK	Energy production/conversion/transfer	4
WD0308	9	chaperonin groES, 10 kDa	wPip	gi|190570502	24	72.9	ESDLLAVIK	Protein modification/degradation/Chaperones	8
WD0631	8	hypothetical protein (WP0282)	wPip	gi|190570728	68	61.9	VQSVEKDAPILDFCVNK	Function unknown	29
WD1255	8	peptidoglycan-associated lipoprotein, putative	wPip	gi|190571199	20	40.9	VTLTGHTDNR	Cell envelope biogenesis/Outer membrane	8
WD0745	7	Putative outer membrane protein	wPip	gi|190571111	59	64.1	FVPYAALHYFMTDEK	Cell envelope biogenesis/Outer membrane	29
WD0683	7	translation elongation factor tu	wPip	gi|190571544	25	38.2	TTLTAAITK	Ribosome stucture/biogenesis/Translation	10
(partial)	6	ankyrin repeat	wPip	gi|190570819	18	47.8	YLIEQGANPNATDHLGR	Function unknown	11

		domain protein							
WD0604	6	minor capsid protein E	wPip	gi|190570849	9	31.8	ALADVITDHLQLMR	Phage\Viral related proteins	4
WD0572	6	putative phage related protein	wPip	gi|190571703	12	54.1	VQEVLKDFFSPIIQKT	Phage\Viral related proteins	7
absent	5	Hypothetical protein WP0984	wPip	gi|190571376	37	65	EEVNHVNNMFGMDILNSFEGR	Function unknown	25
absent	5	Hypothetical protein WP0890	wPip	gi|190571287	23	58.8	IYNYITLAK	Function unknown	12
WD0674	5	ribosomal protein L16	wPip	gi|190571553	6	41.6	VLFEISSDVPMHLAR	Ribosome stucture/biogenesis/Translation	4
WD0016	4	translation elongation factor G	wPip	gi|190570976	14	28.1	FVPVLCGSAFK	Ribosome stucture/biogenesis/Translation	9
WD0906	3	polyribonucleotide nucleotidyltransferase	wPip	gi|190571231	16	39.8	APVAGIAMGLIK	Ribosome stucture/biogenesis/Translation	12
WD0023	3	ribosomal protein L7/L12	wPip	gi|190570969	15	93.2	EVNSTLNLK	Ribosome stucture/biogenesis/Translation	11
WD0590	3	putative phage related protein	wPip	gi|190571688	5	19.7	IVIFGPYGIGK	Phage\Viral related proteins	3
WD0879	3	thioredoxin	wPip	gi|190571104	5	60.2	AVNDQNFESEVANHK	Cellular defense mechanisms	5
WD1050	3	recA protein	wPip	gi|190571327	4	22.3	AEIEGDMGDQHMGLQAR	DNA replication/repair/packaging/Cell division	4

WD0024	2	DNA-directed RNA polymerase, beta/beta' subunits	wPip	gi|190570968	25	24	AIPGVNEENLYHLDDSGIVK	Transcription/Post-transcriptional modification	20
WD1085	2	surface antigen	wPip	gi|190571424	8	19.4	IRLDFGFPLVK	Cell envelope biogenesis/Outer membrane	7
WD0065	2	DNA-binding protein, HU family	wPip	gi|190571020	10	83.5	LKQDCVSQNIDITK	DNA replication/repair/packaging/Cell division	7
WD1271	2	enhancing lycopene biosynthesis protein 2, putative	wPip	gi|190571210	8	41.4	CFAPDINITQVMDHK	Secondary metabolite synthesis/catabolism	7
WD0658	2	DNA-directed RNA polymerase, alpha subunit	wPip	gi|190571569	7	34.9	ILQEQFQPFISSDMSYKK	Transcription/Post-transcriptional modification	7
WD0632	2	hypothetical protein (Ulp1/Peptidase_C48; WP0283)	wPip	gi|190570729	9	17	VISIDFGNPQSALDKIDGVSR	Function unknown	6
WD1090	2	ribosomal protein S1	wPip	gi|190571429	8	34.7	QIEYDPLEELIEK	Ribosome stucture/biogenesis/Translation	5
WD0631	2	hypothetical protein (WP0294)	wPip	gi|190570737	6	33.3	SAFEEDGSDDDLRR	Function unknown	5
WD0531	2	translation elongation factor Ts	wPip	gi|190571620	5	39.2	SIIEEQVK	Ribosome stucture/biogenesis/Translation	4

WD0253	2	transposase	wPip	gi|190571636	6	27.9	TTGLVDYKELETNILSSIR	DNA replication/repair/packaging/Cell division	4
WD0756	2	antioxidant, AhpC/Tsa family	wPip	gi|190570611	5	47.2	GKPAMQASDEGVADYLNSHSAEL	Cellular defense mechanisms	4
WD0790	2	Putative dnaj domain membrane protein	wPip	gi|190570961	4	18.5	DFDGLIAILK	Protein modification/degradation/Chaperones	3
WD0664	2	ribosomal protein S5	wPip	gi|190571563	4	31.8	SNDPHNIICAVFK	Ribosome stucture/biogenesis/Translation	3
WD0511	2	conserved hypothetical protein	wPip	gi|190570734	5	30	IMDEIAAFAQK	General function prediction only	3
WD0021	2	ribosomal protein L1	wPip	gi|190570971	4	37.8	FGTVTSNIAEATK	Ribosome stucture/biogenesis/Translation	3
WD0654	2	transcription elongation factor GreA	wPip	gi|190571574	4	36	DQGDLSENAEYHAAR	Ribosome stucture/biogenesis/Translation	3
WD1318	2	translation initiation factor IF-2	wPip	gi|190571749	2	13.6	ITFIDTPGHEAFTAMR	Ribosome stucture/biogenesis/Translation	1
WD0583*	2	putative phage related protein	wPip	gi|190571691	7	27.3	ILTPGGLLLLGGAPK	Phage\Viral related proteins	6
WD0585	2	putative phage related protein	wPip	gi|190571690	7	37.4	KINSIADLNGLEFTAK	Phage\Viral related proteins	5
WD0594	2	Phage related	wPip	gi|190571683	4	25.1	SDGTVVDGHLR	Phage\Viral	3

		DNA methylase						related proteins	
WD0061	2	ompA-like protein	wPip	gi|190571144	5	24.2	ILGAISYK	Cell envelope biogenesis/Outer membrane	4
WD0732	2	two component transcriptional regulator	wPip	gi|190570997	3	32.8	IGNMNINFDHR	Signal transduction	3
WD0675	2	ribosomal protein S3	wPip	gi|190571552	4	38.4	LHQDLFIR	Ribosome stucture/biogenesis/Translation	3
WD0738	2	superoxide dismutase, Fe	wPip	gi|190571001	3	20.8	LNELVENTDYQHMEIEELVTK	Cellular defense mechanisms	3
WD0589	1	putative phage related protein	wPip	gi|190571689	6	34.7	EYLNDQSSIPK	Phage\Viral related proteins	6
WD0833	1	protease DO	wPip	gi|190571439	5	21.8	INSDKDLPFVEFGNSDK	Protein modification/degradation/Chaperones	5
absent	1	Hypothetical protein WP1117	wPip	gi|190571499	6	46.3	AKTDTIPADLTAK	Function unknown	5
WD0227	1	membrane GTPase involved in stress response	wBm	gi|58584322	6	15.3	INIIDTPGHADFGGEVER	Signal transduction	4
WD0388	1	ribosomal protein S4	wPip	gi|190570680	4	40.2	IPILIEAEQKQER	Ribosome stucture/biogenesis/Translation	4
WD1319	1	N utilization substance protein A	wPip	gi|190571750	4	24.4	AITPAEVSK	Transcription/Post-transcriptional modification	4

WD0795	1	transcription termination factor Rho	wPip	gi|190570947	5	31.6	IFPAIDITK	Transcription/Post-transcriptional modification	4
WD1277	1	heat shock protein HtpG	wPip	gi|190571174	3	24.5	ELISNASDACDKLR	Protein modification/degradation/Chaperones	3
WD0678	1	ribosomal protein L2	wPip	gi|190571549	3	36.9	ATIGVVSNLDHK	Ribosome stucture/biogenesis/Translation	3
WD0320	1	trigger factor, putative	wPip	gi|190570981	3	26.6	LRFPEDYQVISLAGQEAAFSVR	Protein modification/degradation/Chaperones	3
WD0751	1	pyruvate dehydrogenase complex, E3 component, Dihydrolipoamide dehydrogenase	wPip	gi|190570560	4	29.6	DACIDAFFKK	Energy production/conversion/transfer	3
WD1210	1	succinyl-CoA synthase, beta subunit	wPip	gi|190571356	3	36.5	IVKFDIDPATGFTNLDNSK	Energy production/conversion/transfer	3
WD0762	1	peptidase, M16 family	wPip	gi|190570922	3	10.7	ELDTLLFK	General function prediction only	3
WD0174	1	ribosomal 5S rRNA E-loop binding protein Ctc/L25/TL5	wPip	gi|190571325	3	36	CSPEKIPQVIEIDLSGK	Ribosome stucture/biogenesis/Translation	3
WD0029	1	phosphoribosylamin	wPip	gi|190570964	4	17.3	ANGIAAGK	Nucleotide	4

		e--glycine ligase						metabolism	
WD0532	1	ribosomal protein S2	wPip	gi|190571619	3	45.8	ILNDEDSILTKK	Ribosome stucture/biogenesis/Translation	3
WD0832	1	hflC protein	wPip	gi|190571440	4	24.1	EIRAEGEQAGQEIR	Protein modification/degradation/Chaperones	4
WD0917*	1	hypothetical protein WP0593	wPip	gi|190571002	3	29	FSDANAEGVGSPSLSK	Function unknown	3
WD0897	1	iron compound ABC transporter, periplasmic iron compound-binding protein	wPip	gi|190571080	3	29.9	KEELVHSLFDDFTK	Transporters	3
WD0391	1	ribosomal protein L28	wPip	gi|190570684	3	39.4	TFLLNLHK	Ribosome stucture/biogenesis/Translation	3
WD0681	1	ribosomal protein L3	wPip	gi|190571546	3	37.9	IGLLMTNVGHTAMYFDNSR	Ribosome stucture/biogenesis/Translation	3
WD1111*	1	hypothetical protein WP0065	wPip	gi|190570536	3	32.4	IIDETKQEIAQHIENSDVESVQLR	Function unknown	3
WD1285	1	bacterioferritin comigratory protein	wPip	gi|190571297	3	20.7	TTFLIDKK	Cellular defense mechanisms	3

Table A-5. Accession numbers.

Accession numbers for WD0631/WD0632 homologs analyzed in Figure II-1c, e and Figure A-1-3.

	Genome	WD0631-like Accession #	WD0632-like Accession #	WD0632-like Domain	Prophage WO/ WO-like Island
Type I	WOMelB	WP_010962721	WP_010962722	Ulp1/Peptidase_C48	Prophage WO
	WOSuziB	WP_044471237	WP_044471243	Ulp1/Peptidase_C48	Prophage WO
	WORiB	WP_012673191	CP001391*	Ulp1/Peptidase_C48	Prophage WO

	WOHa1	WP_015588933	WP_015588932	Ulp1/Peptidase_C48	Prophage WO
	WOSol	AGK87106	AGK87078	Ulp1/Peptidase_C48	Prophage WO
	WORecB	WP_038198916	JQAM01000018*	Ulp1/Peptidase_C48	Prophage WO
	WOPip1	WP_012481787	WP_012481788	Ulp1/Peptidase_C48	Prophage WO
	WOVitA4	PRJNA213627*	PRJNA213627*	Ulp1/Peptidase_C48	Prophage WO
Type II	WORiC	WP_012673228	WP_012673227	None	Prophage WO
	WOSuziC	WP_044471252	WP_044471251	None	Prophage WO
Type III	wNo	WP_015587806	WP_015587805	Nuclease	WO-like island
	wAlbB	WP_006014162	WP_006014164	Nuclease	WO-like island
	wVitA	PRJNA213627*	PRJNA213627*	Nuclease	WO-like island

Table A-6. Primers.
Primers utilized in this study are listed. F = forward primer, R = reverse primer.

Primer	Sequence	Product Length
Rp49_F	CGGTTACGGATCGAACAAGC	154
Rp49_R	CTTGCGCTTCTTGGAGGAGA	
groELstd_F	GGTGAGCAGTTGCAAGAAGC	932
groELstd_R	AGATCTTCCATCTTGATTCC	
groEL_F	CTAAAGTGCTTAATGCTTCACCTTC	97
groEL_R	CAACCTTTACTTCCTATTCTTG	
WD0034_F	GGAAGAAACTTGCACACCACTTAC	151
WD0034_R	TGCTCTCCGACCATCTGGATATTT	
WD0508_F	TAGAGATCTAGCTTGCGGACAAGA	204
WD0508_R	TCCTTAACTAAACCCTTTGCCACC	
WD0625_F	GAGCCATCAGAAGAAGATCAAGCA	120
WD0625_R	TTCTCGAAAGCTGAAATAGCCTCC	
WD0631_F	TGTGGTAGGGAAGGAAAGAGGAAA	111
WD0631_R	ATTCCAAGGACCATCACCTACAGA	
WD0632_F	TGCGAGAGATTAGAGGGCAAAATC	197
WD0632_R	CCTAAGAAGGCTAATCTCAGACGC	
WD0640_F	CTACAACCTCATCGAAGCGAATCT	144
WD0640_R	CTGCAGAAGCTTTGGAAAAATGGG	
WD0508opt_F	GACGTGCTGATCAAGAGCCT	136
WD0508opt_R	TGCCCACTGTCTTCAGGATG	
WD0625opt_F	CGCGAGATGGATGACCTGAA	180
WD0625opt_R	CTCGCGCTCACTATGTCCAA	
WD0631opt_F	GGTGGATAGTCAGGGCAACC	191
WD0631opt_R	AAAAGTACTCCACGCCCTCG	
WD0632opt_F	CCTGCCCTACATTACACGCA	159
WD0632opt_R	GGCGACAGATCCAGGTCAAT	
Wolb_F	GAAGATAATGACGGTACTCAC	990
Wolb_R3	GTCACTGATCCCACTTTAAATAAC	

Table A-7. Exact p-values.
The exact p-values for all statistical calculations, along with method used, are listed.

Figure	**p-value**	**Comparison**	**Test**
Figure II-2	0.0022	WD0631 1 day old vs 7 days old	Mann-Whitney (two-tailed)
	0.0022	WD0632 1 day old vs 7 days old	Mann-Whitney (two-tailed)
	0.0022	WD0508 1 day old vs 7 days old	Mann-Whitney (two-tailed)
	0.0411	WD0625 1 day old vs 7 days old	Mann-Whitney (two-tailed)
	0.0022	WD0640 1 day old vs 7 days old	Mann-Whitney (two-tailed)
Figure II-3a	0.0022	(-) M x (-) F vs (+) M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)

Figure II-3b	<0.0001	(-) M x (-) F vs (+) M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	<0.0001	(-) M x (-) F vs (-)WD0631,WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	<0.0001	(+) M x (-) F vs (-)WD0631,WD0632 M x (+) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	<0.0001	(-)WD0631,WD0632 M x (-) F vs (-)WD0631,WD0632 M x (+) F	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure II-3c	0.0390	(+) M x (-) F vs (+)WD0631 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	0.0047	(+) M x (-) F vs (+)WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	0.0102	(+)WD0631 M x (-) F vs (+)WD0631,WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	<0.0001	(+) M x (-) F vs (+)WD0631,WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure II-4g	<0.0001	(-) M x (-) F vs (+) M x (-) F	Fisher's exact
	<0.0001	(-) M x (-) F vs (-)WD0631,WD0632 M x (-) F	Fisher's exact
	0.0002	(+) M x (-) F vs (-)WD0631,WD0632 M x (-) F	Fisher's exact
Figure II-4h	<0.0001	(-) M x (-) F vs (+) M x (-) F	Fisher's exact
Figure A-4a	0.0002	Testes (1 day old) vs Testes (7 days old)	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-4b	0.0007	Testes (4 days old) vs Testes (7 days old)	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-4d	0.0104	Older brothers vs younger brothers	Mann-Whitney (two-tailed)
Figure A-4e	<0.0001	(+) M x (+) F vs (+) M x (-) F (older)	ANOVA (Kruskal-Wallis with Dunn's correction)
	<0.0001	(+) M x (+) F vs (+) M x (-) F (younger)	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-5a	0.0004	(-) M x (-) F vs (+) M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	0.0220	(-) M x (-) F vs (+) M x (+) F	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-8a	<0.0001	(-) M x (-) F vs (+) M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-8b	0.0032	(+) M x (-) F vs (+)WD0625,WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	0.0002	(+) M x (-) F vs (+)WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	0.0011	(+)WD0625 M x (-) F vs (+)WD0625,WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	<0.0001	(+)WD0625 M x (-) F vs (+)WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-8c	<0.0001	(-) M x (-) F vs (+) M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
	0.0023	(+) M x (-) F vs (-)WD0625,WD0632 M x (-) F	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-10a	0.0002	WT vs WD0508	Mann-Whitney (two-tailed)
Figure A-10b	0.0003	WT vs WD0632	ANOVA (Kruskal-Wallis with Dunn's correction)

	0.0004	WD0625 vs WD0632	ANOVA (Kruskal-Wallis with Dunn's correction)
Figure A-10c	<0.0001	WT vs WD0632	ANOVA (Kruskal-Wallis with Dunn's correction)
	0.0334	WT vs WD0631,WD0632	ANOVA (Kruskal-Wallis with Dunn's correction)

Appendix B.

Chapter III supplementary information**

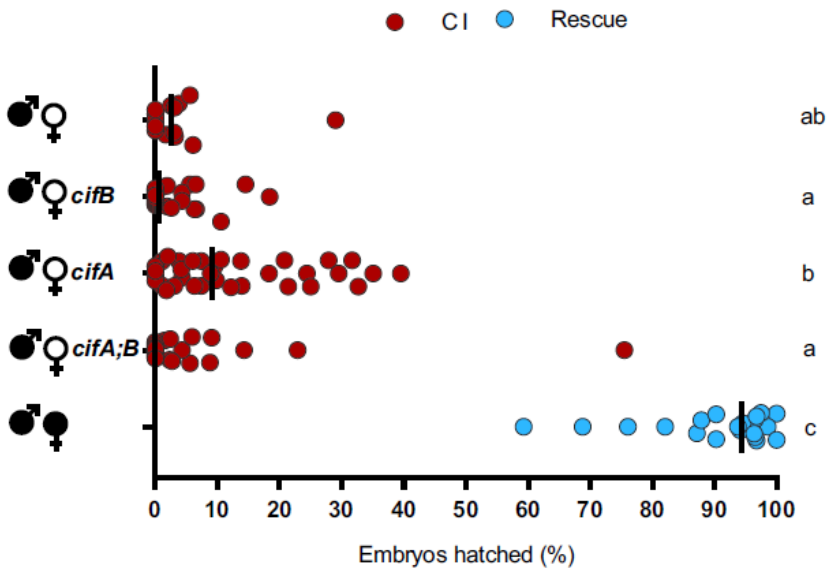

Figure B-1. Transgenic *cifA* expression in germline stem cells fails to elicit rescue.
Transgenic expression of *cifA*, *cifB*, and *cifA;B* using the *nos*-GAL4-*tubulin* driver does not lead to rescue of cytoplasmic incompatibility. Each dot represents a replicate. *Wolbachia* infections are represented by filled sex symbols, and expressed genes are noted to the right of the corresponding sex. n = 15–34 for each experimental cross. Vertical bars represent medians and letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by Kruskal–Wallis and Dunn's test for multiple comparisons. Statistical comparisons are between all groups. Exact P values are provided in Table B-2.

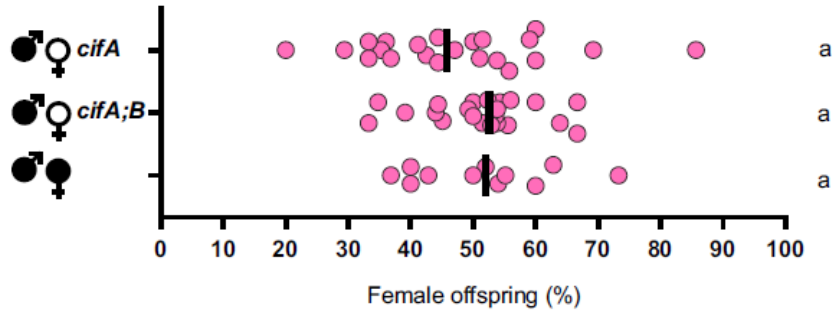

Figure B-2. *cifA* does not preferentially rescue one sex over the other.
Surviving offspring from the experiment displayed in Figure III-2 were collected for adult sex ratio counts. There was no significant difference between any of the crosses. A sex ratio count was not possible for CI crosses due to the low number of surviving offspring. *Wolbachia* infections are represented by filled sex symbols and expressed genes are noted to the right of the corresponding sex. n = 11–22 for each experimental cross. Vertical bars represent medians and letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by Kruskal–Wallis and Dunn's test

** This chapter is published in *Proceedings of the National Academy of Sciences*, *115*(19), 4987-4991 with myself as first author. Jungmin On, Emily Layton, and Helen Zhou were co-authors. Seth Bordenstein was senior author.

for multiple comparisons. Statistical comparisons are between all groups. Exact P values are provided in Table B-2. This experiment was conducted once.

A PSORTb subcellular protein localization

	Accession	Cytoplasmic	Cytoplasmic Membrane	Periplasmic	Outer Membrane	Extracellular
*w*Mel	WP_010962721.1	8.96	0.51	0.26	0.01	0.26
*w*Ri	WP_012673191.1	8.96	0.51	0.26	0.01	0.26
*w*Rec	WP_038198916.1	8.96	0.51	0.26	0.01	0.26
*w*Suzi	WP_044471237.1	8.96	0.51	0.26	0.01	0.26
*w*Ha	WP_015588933.1	8.96	0.51	0.26	0.01	0.26
*w*Sol	AGK87106.1	8.96	0.51	0.26	0.01	0.26
*w*Bol 1-b	WP_019236549.1	8.96	0.51	0.26	0.01	0.26
*w*Pip	WP_012481787.1	8.96	0.51	0.26	0.01	0.26

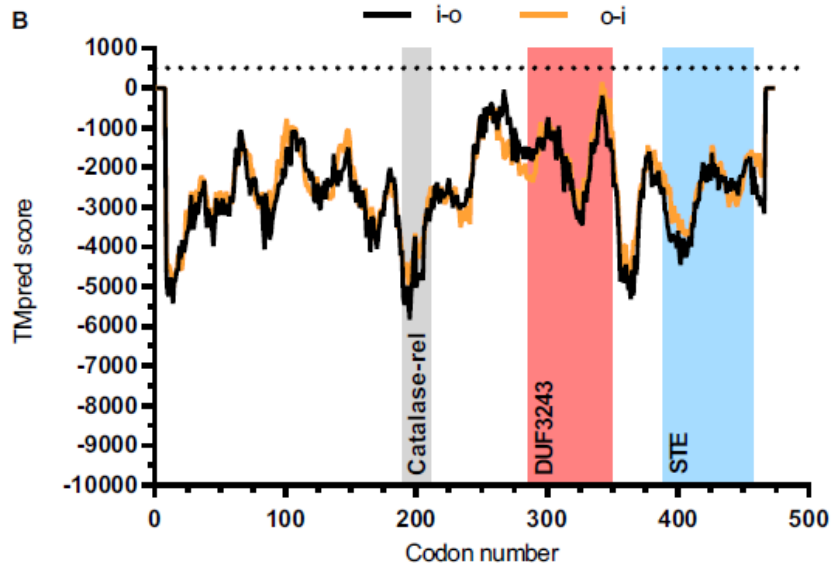

Figure B-3. CifA is a putative cytoplasmic protein.
(A) The PSORTb subcellular protein localization web server was used on type I CifA proteins to predict the protein's localization in the *Wolbachia* cell. Predictive scores above 7.5 are accepted to be sufficient to determine a single location of localization and suggest that CifA is a cytoplasmic protein. (B) The TMpred web server was used to predict transmembrane helices. TMpred scores exceeding 500 (denoted by horizontal dotted line) are considered significant. TMpred scores were generated for transmembrane helices spanning from inside-to-outside (i-o) and outside-to-inside (o-i). Shaded regions denote previously described protein domain predictions (Lindsey et al., 2018b).

A Codon-based test of neutrality

	Genome Accession	*w*Mel	*w*Ri	*w*Rec	*w*Suzi	*w*Ha	*w*Sol	*w*Bol 1-b	*w*Pip
***w*Mel**	NC_002978		2.241	-0.422	1.734	-7.316	-7.323	-6.388	-6.721
***w*Ri**	NC_012416	0.027		0.315	1.416	-7.355	-7.361	-6.426	-6.759
***w*Rec**	NZ_JQAM01000018	0.673	0.754		-0.161	-7.205	-7.211	-6.277	-6.609
***w*Suzi**	NZ_CAOU02000034	0.085	0.160	0.872		-7.412	-7.419	-6.484	-6.816
***w*Ha**	NC_021089	0.000	0.000	0.000	0.000		-1.206	-4.038	-3.929
***w*Sol**	KC955252	0.000	0.000	0.000	0.000	0.230		-4.229	-4.121
***w*Bol 1-b**	NZ_CAOH01000056	0.000	0.000	0.000	0.000	0.000	0.000		-1.603
***w*Pip**	NC_010981	0.000	0.000	0.000	0.000	0.000	0.000	0.112	

B Fisher's exact test of neutrality

	Genome Accession	*w*Mel	*w*Ri	*w*Rec	*w*Suzi	*w*Ha	*w*Sol	*w*Bol 1-b	*w*Pip
***w*Mel**	NC_002978								
***w*Ri**	NC_012416	0.297							
***w*Rec**	NZ_JQAM01000018	1.000	0.618						
***w*Suzi**	NZ_CAOU02000034	0.483	0.616	1.000					
***w*Ha**	NC_021089	1.000	1.000	1.000	1.000				
***w*Sol**	KC955252	1.000	1.000	1.000	1.000	1.000			
***w*Bol 1-b**	NZ_CAOH01000056	1.000	1.000	1.000	1.000	1.000	1.000		
***w*Pip**	NC_010981	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

C

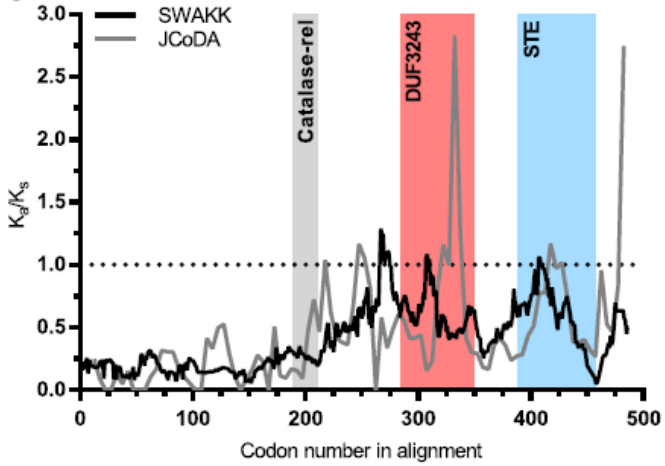

Figure B-4. CifA regions evolve under negative selection.
(A) Pairwise codon-based Z tests of selection suggest that regions of the *cifA* gene are not evolving under the neutral expectation of Ka = Ks. Values below the diagonal are P values for where there is a significant departure from neutrality or not. Values above the diagonal are the difference of Ka − Ks in which positive values suggest positive selection and negative values suggest purifying selection. (B) Pairwise Fisher's exact tests of neutrality suggest that *cifA* evolves under purifying selection. Values below the diagonal are P values. If the P value is less than 0.05, then the null hypothesis of strictly neutral or purifying selection is rejected. If the observed number of synonymous differences per synonymous site exceeds the number of nonsynonymous differences per nonsynonymous site then MEGA sets P = 1 to indicate purifying selection, rather than positive selection. (C) SWAKK and JCoDA were used for analysis between *cifA* homologs of *w*Mel and the bidirectionally incompatible *w*Ha. Both programs were performed with 25-aminoacid windows and yield Ka/Ks ratios evident of strong purifying selection in the N-terminal region preceding the catalase-rel domain and weaker purifying selection beyond it. Shaded regions denote previously described domain predictions (Lindsey et al., 2018b).

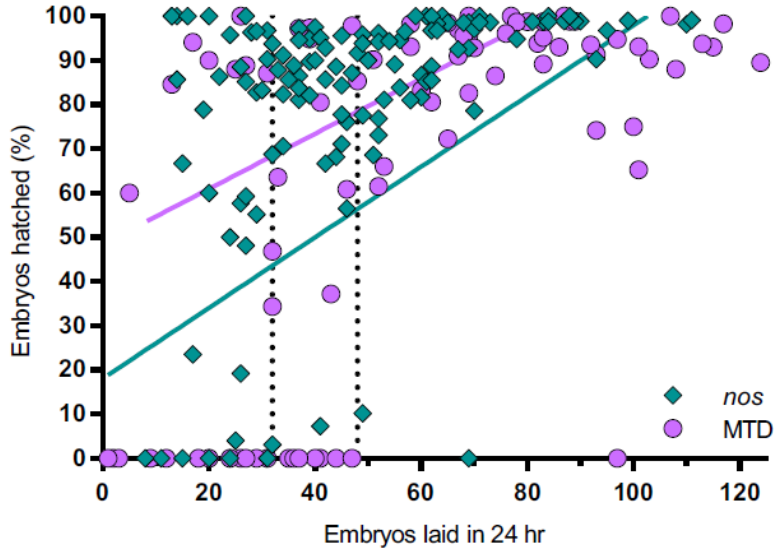

Figure B-5. Fertility is related to strain genotype.
A metaanalysis of control rescue crosses (infected male × infected female) without a transgene shows that clutch size and hatch rate are significantly correlated for both the MTD-GAL4 and *nos*-GAL4-*tubulin* genotypes (r = 0.59 and 0.50 for MTD-GAL4 and *nos*-GAL4-*tubulin*, respectively), but the two strains have different y intercepts (18.06 +/− 6.73 and 48.49 +/− 5.34 for MTD-GAL4 and *nos*-GAL4-*tubulin*, respectively). Each dot represents a replicate where circles and diamonds are MTD-GAL4 (n = 91) and nos-Gal4-tubulin (n = 134), respectively. Vertical dotted lines represent embryo counts where 99% of clutch sizes with 0% embryo hatch rate are to the left for *nos*-GAL4-*tubulin* (left line) and MTD-GAL4 (right line). Correlation was assessed with Spearman rho. A linear regression best-fit line is plotted for each genotype. Exact P values are provided in Table B-2.

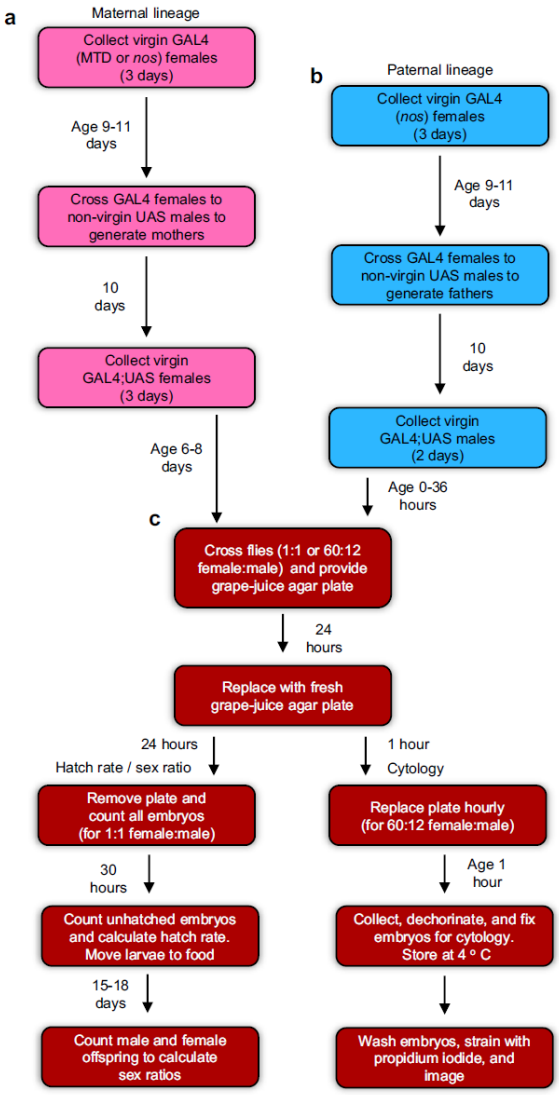

Figure B-6. Schematic of experimental methodology.
(A) All experimental setups begin with the generation of the maternal lineage (pink), derived from GAL4 driver lines and collected as virgins and aged for 6–8 d until the peak of their fecundity. (B) The paternal lineage (blue) is set up in a stagger such that the males used in the experiment emerge on the day of the experiment. (C) Flies are crossed in a fashion dependent on the ultimate intent, and grape-juice agar plates provided and replaced in a similar manner for all experiments. Sex ratio studies are derived from hatch rate assays.

Table B-1. Primers used in Chapter III for RT-qPCR (Figure III-1b) or for *Wolbachia* infection checks.

Primer	Sequence	Product Length (bp)
Rp49_F	CGGTTACGGATCGAACAAGC	154
Rp49_R	CTTGCGCTTCTTGGAGGAGA	
cifA opt_F	CCCGCTATTGCATCACAGGA	186
cifA opt_R	CGCGGTCGATCCAAAAATCG	
Wolb_F	GAAGATAATGACGGTACTCAC	990
Wolb_R3	GTCACTGATCCCACTTTAAATAAC	
F = forward primer, R = reverse primer.		

Table B-2. P-values associated with all statistical comparisons made in main and extended data figures.

Figure	Comparison	p-value	Test
Figure III-1a	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]MTD;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;+]MTD;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]nos;wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]nos;*cifA*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;+]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;+]MTD;wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;-]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;-]nos;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;+]nos;wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;+]MTD;wt vs. [M;+]nos;wt x [F;-]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;+]MTD;wt vs. [M;+]nos;wt x [F;-]nos;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;+]MTD;wt vs. [M;+]nos;wt x [F;+]nos;wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;wt vs. [M;+]nos;wt x [F;-]nos;*cifA*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;wt vs. [M;+]nos;wt x [F;+]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;*cifA* vs. [M;+]nos;wt x [F;+]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
Figure III-1b	nos-GAL4-tubulin vs. MTD-GAL4 *cifA* expression	**<0.0001**	Mann-Whitney test
Figure III-2	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]MTD;*cifA*	**0.0379**	Kruskal Wallis with Dunn's correction

	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]MTD;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]MTD;*cifA*;*cifB*	**0.0038**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;+]MTD;wt	**0.0006**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;-]MTD;*cifB*	**0.0058**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;-]MTD;*cifA*;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;+]MTD;wt	0.5436	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifB* vs. [M;+]nos;wt x [F;-]MTD;*cifA*;*cifB*	**0.0004**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifB* vs. [M;+]nos;wt x [F;+]MTD;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA*;*cifB* vs. [M;+]nos;wt x [F;+]MTD;wt	>0.9999	Kruskal Wallis with Dunn's correction
Figure III-3	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]MTD;*cifA*	**0.0010**	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]MTD;*cifB*	0.4680	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;-]MTD;*cifA*;*cifB*	**0.0010**	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;wt vs. [M;+]nos;wt x [F;+]MTD;wt	**0.0010**	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;-]MTD;*cifB*	**0.0010**	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;-]MTD;*cifA*;*cifB*	1.0000	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;+]MTD;wt	0.5740	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;*cifB* vs. [M;+]nos;wt x [F;-]MTD;*cifA*;*cifB*	**0.0010**	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;*cifB* vs. [M;+]nos;wt x [F;+]MTD;wt	**0.0010**	Chi-square with bonferroni adjusted p-value
	[M;+]nos;wt x [F;-]MTD;*cifA*;*cifB* vs. [M;+]nos;wt x [F;+]MTD;wt	**0.0030**	Chi-square with bonferroni adjusted p-value
Figure B-1	[M;+]nos;wt x [F;-]nos;wt vs. [M;+]nos;wt x [F;-]nos;*cifA*	0.1534	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;wt vs. [M;+]nos;wt x [F;-]nos;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;wt vs. [M;+]nos;wt x [F;-]nos;*cifA*;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;wt vs. [M;+]nos;wt x [F;+]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;*cifA* vs. [M;+]nos;wt x [F;-]nos;*cifB*	**0.0204**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;*cifA* vs. [M;+]nos;wt x [F;-]nos;*cifA*;*cifB*	**0.0306**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;*cifA* vs. [M;+]nos;wt x [F;+]nos;wt	**0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;*cifB* vs. [M;+]nos;wt x [F;-]nos;*cifA*;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction

	[M;+]nos;wt x [F;-]nos;*cifB* vs. [M;+]nos;wt x [F;+]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]nos;*cifA*;*cifB* vs. [M;+]nos;wt x [F;+]nos;wt	**<0.0001**	Kruskal Wallis with Dunn's correction
Figure B-2	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;-]MTD;*cifA*;*cifB*	0.5209	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA* vs. [M;+]nos;wt x [F;+]MTD;wt	0.8609	Kruskal Wallis with Dunn's correction
	[M;+]nos;wt x [F;-]MTD;*cifA*;*cifB* vs. [M;+]nos;wt x [F;+]MTD;wt	>0.9999	Kruskal Wallis with Dunn's correction
Figure B-5	Hatch rate vs clutch size (MTD-GAL4)	**<0.0001**	Spearman's Rho
	Hatch rate vs clutch size (nos-GAL4-tubulin)	**<0.0001**	Spearman's Rho
M = male, F = female, + = *Wolbachia* infected, - = *Wolbachia* uninfected, bold p-values = significant			

Appendix C.

Chapter IV supplementary information[††]

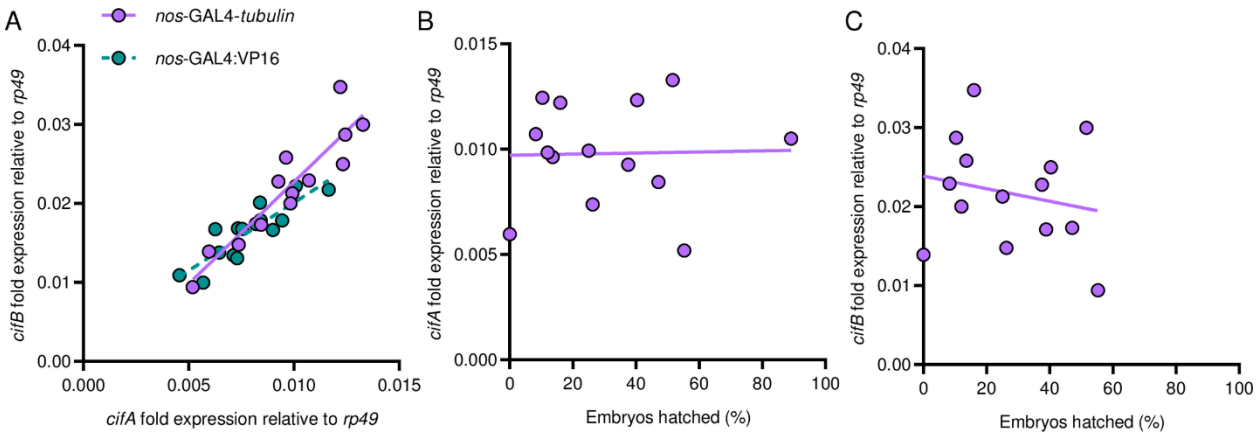

Figure C-1. Fold expression of transgenic *cifA*$_{wMel}$ correlates with *cifB*$_{wMel}$ in males relative to the *Drosophila* housekeeping gene *rp49* but neither correlate with hatch rate under the *nos*-GAL4-*tubulin* driver.
(A) A linear regression of *cifA*$_{wMel}$ and *cifB*$_{wMel}$ expression reveals a positive correlation for both *nos*-GAL4-*tubulin* and *nos*-GAL4VP16. (B,C) A linear regression of (B) *cifA*$_{wMel}$ and (C) *cifB*$_{wMel}$ expression and embryonic hatching reveals no correlation for *nos-GAL4-tubulin*. Removal of data points corresponding to 0% embryonic hatching did not change the significance of the correlation. The nos-GAL4:VP16 driver was not included in analysis A or B since the majority of data points corresponded with 0% hatching. This analysis uses hatch rate samples from the experiment in Figure IV-2a and expression data from Figure IV-2b and Figure IV-2c.

Table C-1. P-values associated with all statistical comparisons made in main and supporting information figures.

Figure	Comparison	p-value	Test
Figure IV-2a	[M;+]nos-gal4-tubulin;wt x [F;-]wt vs [M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;-]wt	0.1198	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]wt vs [M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]wt vs [M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;-]wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]wt vs [M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;-]wt vs [M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;+]wt	**0.0005**	Kruskal Wallis with Dunn's correction
	[M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;-]wt vs [M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;-]wt	**0.0002**	Kruskal Wallis with Dunn's correction
	[M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;-]wt vs [M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;+]wt	**0.0002**	Kruskal Wallis with Dunn's correction
	[M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;+]wt vs [M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;-]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-gal4-tubulin;*cifA*;*cifB* x [F;+]wt vs [M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;+]wt	>0.9999	Kruskal Wallis with Dunn's correction

[††] This chapter is published in *PLOS Genetics.* 16(6), e1008221 with myself as first author. Seth Bordenstein was senior author.

	[M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;-]wt vs [M;-]nos-gal4:VP16;*cifA*;*cifB* x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
Figure IV-2b	[M;-]nos-gal4-tubulin;*cifA*;*cifB cifA* expression vs [M;-]nos-gal4:VP16;*cifA*;*cifB cifA* expression	**0.0157**	Mann-Whitney U
Figure IV-2c	[M;-]nos-gal4-tubulin;*cifA*;*cifB cifB* expression vs [M;-]nos-gal4:VP16;*cifA*;*cifB cifB* expression	**0.0202**	Mann-Whitney U
Figure IV-3a	[M;+]nos-gal4-tubulin;wt x [F;-]otu-GAL4:VP16;wt vs [M;+]nos-gal4-tubulin;wt x [F;-]otu-GAL4:VP16;*cifA*	**0.0051**	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]otu-GAL4:VP16;wt vs [M;+]nos-gal4-tubulin;wt x [F;-]nos-GAL4:VP16;wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]otu-GAL4:VP16;wt vs [M;+]nos-gal4-tubulin;wt x [F;-]nos-GAL4:VP16;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]otu-GAL4:VP16;*cifA* vs [M;+]nos-gal4-tubulin;wt x [F;-]nos-GAL4:VP16;wt	0.0576	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]otu-GAL4:VP16;*cifA* vs [M;+]nos-gal4-tubulin;wt x [F;-]nos-GAL4:VP16;*cifA*	**0.0491**	Kruskal Wallis with Dunn's correction
	[M;+]nos-gal4-tubulin;wt x [F;-]nos-GAL4:VP16;wt vs [M;+]nos-gal4-tubulin;wt x [F;-]nos-GAL4:VP16;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
Figure IV-3b	[F;-]otu-GAL4:VP16;*cifA cifA* expression vs [F;-]nos-GAL4:VP16;*cifA cifA* expression	**<0.0001**	Mann-Whitney U
Figure IV-4	[M;+]wt x [F;-]nos-GAL4:VP16;wt vs[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*;*cifB*	**0.0024**	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	**0.0141**	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]nos-GAL4:VP16;wt vs [M;+]wt x [F;-]nos-GAL4:VP16;*cifA*	**0.004**	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]nos-GAL4:VP16;wt vs [M;+]wt x [F;+]wt	**0.01**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction

	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*;*cifB*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;wt vs [M;+]wt x [F;-]nos-GAL4:VP16;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;wt vs [M;+]wt x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA* vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifB*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA* vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*;*cifB*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA* vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA* vs [M;+]wt x [F;-]nos-GAL4:VP16;*cifA*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA* vs [M;+]wt x [F;+]wt	0.4176	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifB* vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*;*cifB*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifB* vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifB* vs [M;+]wt x [F;-]nos-GAL4:VP16;*cifA*	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifB* vs [M;+]wt x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*;*cifB* vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*;*cifB* vs [M;+]wt x [F;-]nos-GAL4:VP16;*cifA*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]nos-GAL4:VP16;*cifA*;*cifB* vs [M;+]wt x [F;+]wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt vs [M;+]wt x [F;-]nos-GAL4:VP16;*cifA*	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt vs [M;+]wt x [F;+]wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]nos-GAL4:VP16;*cifA* vs [M;+]wt x [F;+]wt	>0.9999	Kruskal Wallis with Dunn's correction

Figure IV-5	[M;+]wt x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA* x [F;-]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifB* x [F;-]wt	**0.0002**	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;+]wt x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA* x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifB* x [F;-]wt	0.1572	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA* x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA* x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	>0.9999	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifB* x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifB* x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	0.182	Kruskal Wallis with Dunn's correction
	[M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;-]wt vs [M;-]nos-GAL4:VP16;*cifA*;*cifB* x [F;+]wt	**<0.0001**	Kruskal Wallis with Dunn's correction
Figure C-1a	[M;-]nos-GAL4-tubulin;*cifA*;*cifB cifA* expression vs *cifB* expression	**<0.0001**	Linear regression
	[M;-]nos-GAL4:VP16;*cifA*;*cifB cifA* expression vs *cifB* expression	**<0.0001**	Linear regression
Figure C-1b	[M;-]nos-GAL4-tubulin;*cifA*;*cifB cifA* expression vs hatch rate	0.9307	Linear regression
	[M;-]nos-GAL4:VP16;*cifA*;*cifB cifA* expression vs hatch rate	0.6968	Linear regression
Figure C-1c	[M;-]nos-GAL4-tubulin;*cifA*;*cifB cifB* expression vs hatch rate	0.4816	Linear regression
	[M;-]nos-GAL4:VP16;*cifA*;*cifB cifB* expression vs hatch rate	0.6756	Linear regression
M = male, F = female, + = *Wolbachia* infected, - = *Wolbachia* uninfected, bold p-values = significant			

Appendix D.

Chapter V supplementary information

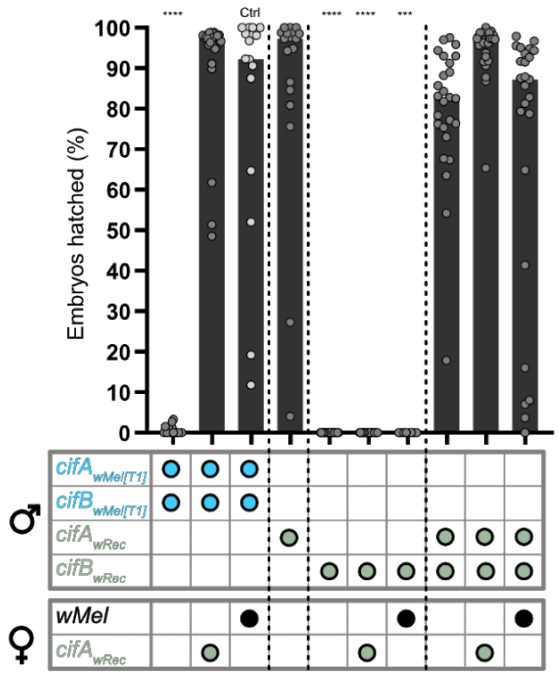

Figure D-1. *w*Mel-infected females fail to rescue embryonic death caused by *cifB*$_{wRec[T1]}$.
Hatch rate analyses testing if *w*Mel-infected females can rescue lethality induced by *cifA;B*$_{wMel[T1]}$, *cifB*$_{wRec}$, and *cifA;B*$_{wRec[T1]}$ (N = 8-26 where each dot represents a clutch of embryos from a single mating pair). Horizontal bars represent median embryonic hatching from single pair matings. Genotypes for each cross are illustrated below the bars where the genes expressed in each sex are represented by colored circles. Blue circles represent *cif*$_{wMel[T1]}$ genes, green circles represent *cif*$_{wRec[T1]}$ genes, and black circles represent *w*Mel-infected flies. All other flies were uninfected with *Wolbachia*. Each hatch rate was conducted once. Asterisks above bars represent significant differences relative to a control rescue cross (denoted Ctrl) with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

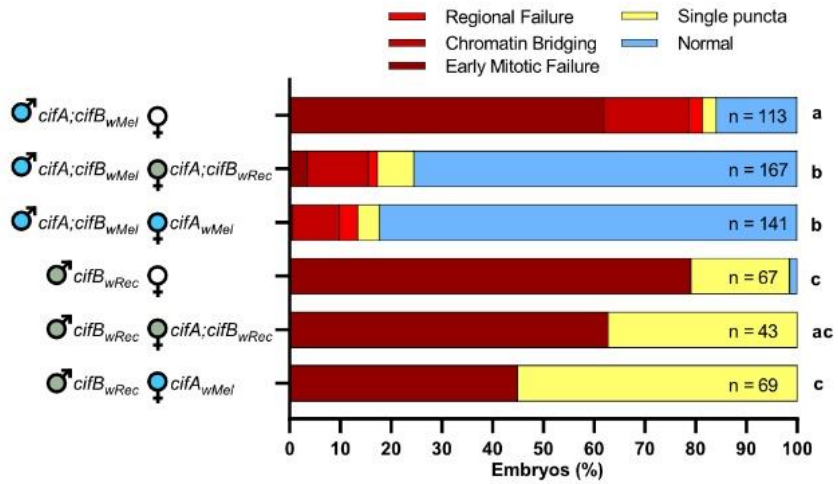

Figure D-2. *cifB*$_{wRec[T1]}$ expressing males cause non-CI, sterility artifacts in embryos.
The percentage of embryos with each cytological phenotype resulting from the indicated crosses are shown. All crosses were conducted in parallel and with sisters from the experiment in Fig. 2B. Filled sex symbols represent transgenic expression of *cif*$_{wMel[T1]}$ (blue) or *cif*$_{wRec[T1]}$ (green) genes. Expressed genes are noted to the right of the corresponding sex. All flies were uninfected with *Wolbachia*. Letters to the right indicate significant differences based on $\alpha = 0.05$ calculated by pairwise $\chi 2$ analyses comparing the sum of defects (red and yellow categories) against normal (blue) with Bonferroni adjusted P-values.

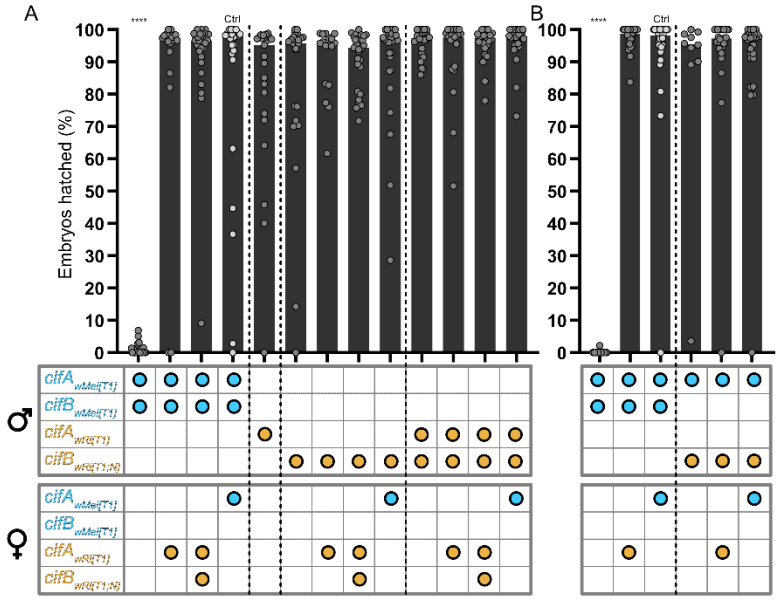

Figure D-3. *cifB*$_{wRi[T1;N]}$ does not contribute to CI.
(A,B) Hatch rate analyses testing *cif*$_{wRi[T1]}$ transgenes for rescue and CI. (A) Hatch rate analyses testing *cifA*$_{wRi[T1]}$, *cifB*$_{wRi[T1;N]}$, and *cifA;B*$_{wRi[T1;N]}$ for CI-induction (N = 17-30 where each dot represents a clutch of embryos from a single mating pair). (B) *cifB*$_{wRi[T1;N]}$ is tested for CI when dual expressed with *cifA*$_{wMel[T1]}$ (N = 10-32). Horizontal bars represent median embryonic hatching from single pair matings. Genotypes for each cross are illustrated below the bars where the genes expressed in each sex are represented by colored circles. Blue circles represent *cif*$_{wMel[T1]}$ genes and orange circles represent *cif*$_{wRi[T1]}$ genes. All flies were uninfected with *Wolbachia*. Each hatch rate was conducted once. Asterisks above bars represent significant differences relative to a control transgenic rescue cross (denoted Ctrl) with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

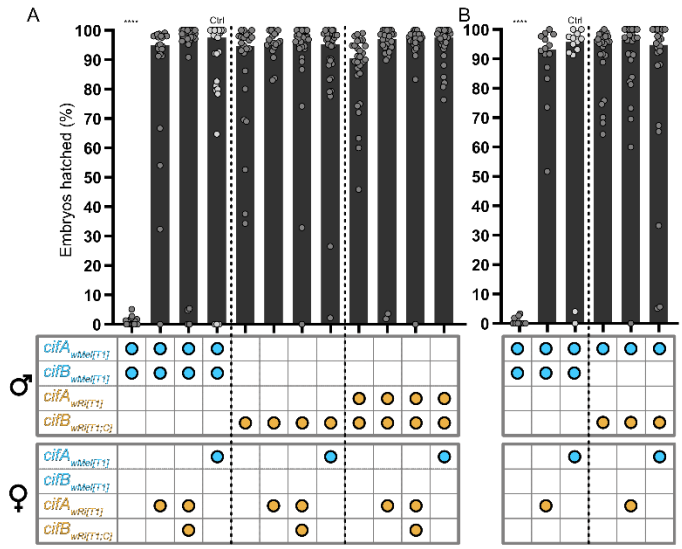

Figure D-4. *cifB*$_{wRi[T1;C]}$ does not contribute to CI.
(A,B) Hatch rate analyses testing *cif*$_{wRi[T1]}$ transgenes for rescue and CI. (A) Hatch rate analyses testing *cifA*$_{wRi[T1]}$, *cifB*$_{wRi[T1;C]}$, and *cifA;B*$_{wRi[T1;C]}$ for CI-induction (N = 19-31 where each dot represents a clutch of embryos from a single mating pair). (B) *cifB*$_{wRi[T1;C]}$ is tested for CI when dual expressed with *cifA*$_{wMel[T1]}$ (N = 13-25). Horizontal bars represent median embryonic hatching from single pair matings. Genotypes for each cross are illustrated below the bars where the genes expressed in each sex are represented by colored circles. Blue circles represent *cif*$_{wMel[T1]}$ genes and orange circles represent *cif*$_{wRi[T1]}$ genes. All flies were uninfected with *Wolbachia*. Each hatch rate was conducted once. Asterisks above bars represent significant differences relative to a control transgenic rescue cross (denoted Ctrl) with an $\alpha = 0.05$ calculated by a Kruskal-Wallis analysis followed by Dunn's multiple comparison test. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

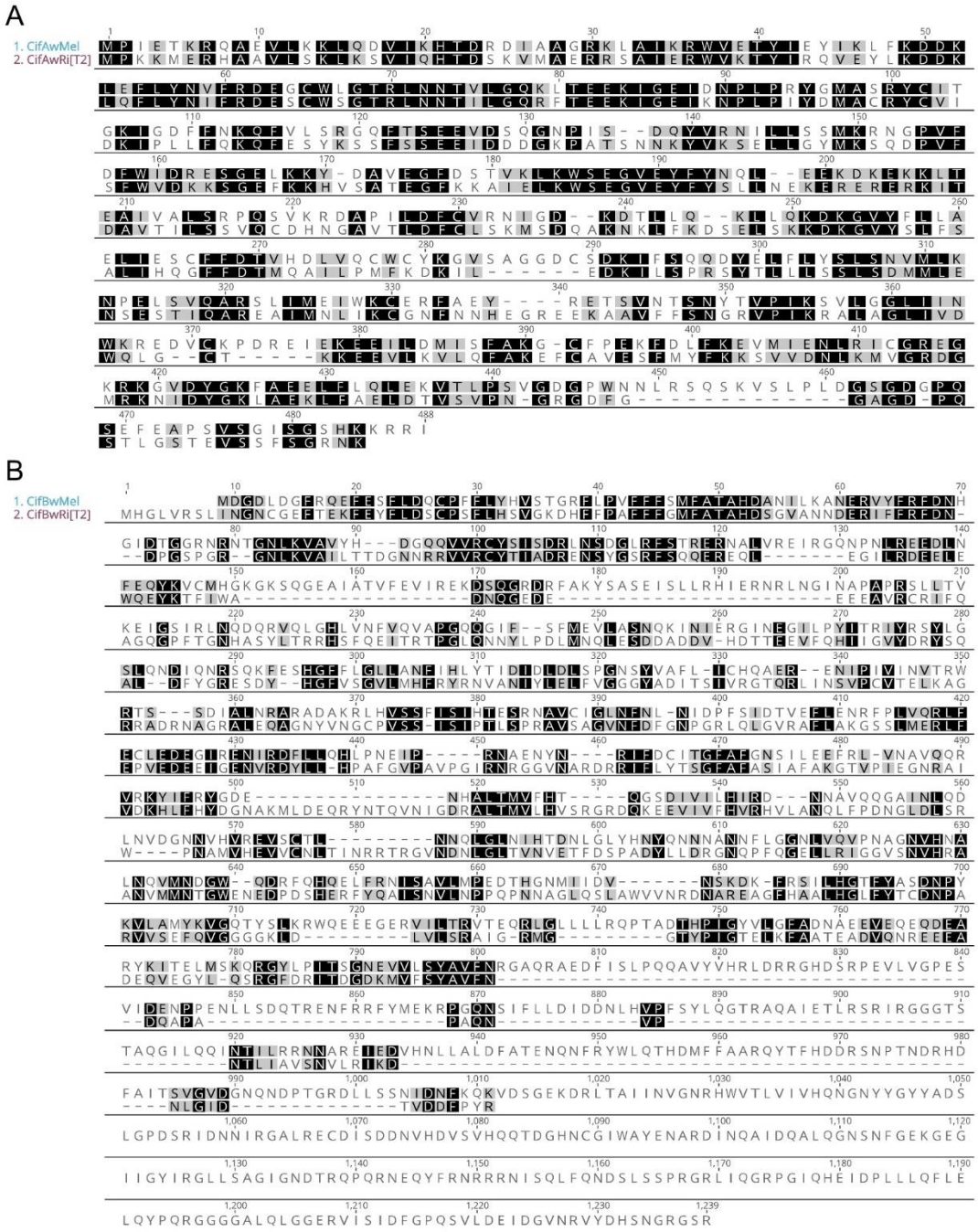

Figure D-5. MUSCLE alignment of Cif$_{wMel[T1]}$ and Cif$_{wRi[T2]}$.
(A) Alignment of CifA proteins. (B) Alignment of CifB proteins. Black shading represents conserved amino acids, grey is different but chemically comparable, and white is different. Hyphens represent gaps in the alignment. All alignments were pairwise and generated in Geneious using default settings.

Appendix E.

Chapter VI supplementary information

Table E-1. P-values associated with all statistical comparisons made in main and extended data hatch rate and cytology figures.
M=male, F=female ,+=Wolbachia infected, −=Wolbachia uninfected.

Figure	Comparison	Summary	P value	Test
Figure VI-1b	yw[+];yw[-] x yw[-] vs nos:VP16[-];cifA[-]; cifB[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vs nos:VP16[-];cifA*1[-]; cifB[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vs nos:VP16[-];cifA*2[-]; cifB[-] x yw[-]	ns	0.1977	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vs nos:VP16[-];cifA*3[-]; cifB[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vs nos:VP16[-];cifA*4[-]; cifB[-] x yw[-]	ns	0.7955	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vs nos:VP16[-];cifA[-]; cifB[-] x yw[+]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-] vs nos:VP16[-];cifA*1[-];cifB[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-] vs nos:VP16[-];cifA*2[-];cifB[-] x yw[-]	***	0.0006	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-] vs nos:VP16[-];cifA*3[-];cifB[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-] vs nos:VP16[-];cifA*4[-];cifB[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-] vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*1[-];cifB[-] x yw[-]vs nos:VP16[-];cifA*2[-];cifB[-] x yw[-]	**	0.0015	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*1[-];cifB[-] x yw[-]vs nos:VP16[-];cifA*3[-];cifB[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*1[-];cifB[-] x yw[-]vs nos:VP16[-];cifA*4[-];cifB[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*1[-];cifB[-] x yw[-]vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*2[-];cifB[-] x yw[-] vs nos:VP16[-];cifA*3[-];cifB[-] x yw[-]	**	0.0016	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*2[-];cifB[-] x yw[-] vs nos:VP16[-];cifA*4[-];cifB[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*2[-];cifB[-] x yw[-] vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	***	0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*3[-];cifB[-] x yw[-] vs nos:VP16[-];cifA*4[-];cifB[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*3[-];cifB[-] x yw[-] vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA*4[-];cifB[-] x yw[-]vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test

Figure VI-1c	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; yw[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; yw[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*1[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; yw[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*2[-]	*	0.0277	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; yw[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*3[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; yw[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*4[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*1[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*2[-]	*	0.039	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*3[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA[-] vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*4[-]	ns	0.3039	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA*1[-]vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*2[-]	*	0.0307	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA*1[-]vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*3[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA*1[-]vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*4[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA*2[-]vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*3[-]	**	0.0012	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA*2[-]vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*4[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-]; cifA*3[-]vs nos:VP16[-];cifA[-];cifB[-] x nos:VP16[-];cifA*4[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
Figure VI-2b	yw[+];yw[-] x yw[-] vsnos:VP16[-];cifA[-]; cifB[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vsnos:VP16[-];cifA[-]; cifB*1[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vsnos:VP16[-];cifA[-]; cifB*2[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vsnos:VP16[-];cifA[-]; cifB*3[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	yw[+];yw[-] x yw[-] vsnos:VP16[-];cifA[-]; cifB*4[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test

	yw[+];yw[-] x yw[-] vsnos:VP16[-];cifA[-]; cifB[-] x yw[+]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*1[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*2[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*3[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*4[-] x yw[-]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB[-] x yw[-]vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	****	<0.0001	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*1[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*2[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*1[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*3[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*1[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*4[-] x yw[-]	ns	0.4359	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*1[-] x yw[-]vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*2[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*3[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*2[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*4[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*2[-] x yw[-]vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*3[-] x yw[-]vs nos:VP16[-];cifA[-];cifB*4[-] x yw[-]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*3[-] x yw[-]vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test
	nos:VP16[-];cifA[-];cifB*4[-] x yw[-]vs nos:VP16[-];cifA[-];cifB[-] x yw[+]	ns	>0.9999	Kruskal-Wallis with Dunn's multiple correction test

Table E-2. Protein structural prediction software I-TASSER identifies homologous protein domains found in all of our Cif mutants, and those that differ.

Query	PDB Hits	Presence	Classification
CifA	1vt4	All	Apoptosis
	2pff	All	Transferase
	4um2A	All	Telomerase-binding protein
	4y21A	All	Exocytosis
	5lj3A	All but CifA$_{*2}$	Splicing
	5voxP	All	Hydrolase
	1ldjA	CifA$_{*1}$, CifA$_{*2}$	Ligase
	5a9qA	CifA$_{*2}$	Transport protein
	5ctqA	CifA$_{*2}$	Immune system nuclear protein RNA binding protein
	1vt4A	CifA$_{*3}$	Apoptosis
	4n5aA	CifA$_{*3}$	Protein binding

	4pjwA	CifA[*4]	Cell cycle
CifB	5ijoJ	All	Transport protein
	2oivA	CifB	Hydrolase
	5yz0A	CifB	Cell cycle
	1euv	All	Hydrolase
	5ham	All	Hydrolase
	2xphA	All but cifB[*1]	Hydrolase
	3j3iA	All	Virus
	5yfpB	All	Exocytosis
	6edoA	All but cifB	Motor protein
	3iayA	cifB[*3]	Transferase/DNA

Table E-3. Primers used for genotyping and sanger sequencing of *cif* transgenes.

Amplicon	Direction	Name	Sequence
1	Forward	pTIGMCS_F	GAGGAAAGGTTGTGTGCGGACGA
	Reverse	MelA1208_R	AGCCCTCCACGGCATCGTACTT
2	Forward	MelA1072_F	TAGCCAGGGCAACCCAATCTCG
	Reverse	MelA1763_R	TGCACACATCCTCGCGCTTC
3	Forward	MelA1667_F	TACCGCGAGACCAGCGTGAA
	Reverse	pTIGMCS_R	CGAATTGGTGCTATGTTTATGGCGCT

Appendix F.

Speciation by symbiosis: The microbiome and behavior[‡‡]

Abstract

Species are fundamental units of comparison in biology. The newly discovered importance and ubiquity of host-associated microorganisms is now stimulating work on the roles that microbes can play in animal speciation. We previously synthesized the literature and advanced concepts of speciation by symbiosis with notable attention to hybrid sterility and lethality. Here, we review recent studies and relevant data on microbes as players in host behavior and behavioral isolation, emphasizing the patterns seen in these analyses and highlighting areas worthy of additional exploration. We conclude that the role of microbial symbionts in behavior and speciation is gaining exciting traction, and the holobiont and hologenome concepts afford an evolving intellectual framework to promote research and intellectual exchange between disciplines such as behavior, microbiology, genetics, symbiosis and speciation. Given the increasing centrality of microbiology in macroscopic life, microbial symbiosis is arguably the most neglected aspect of animal and plant speciation, and studying it should yield a better understanding of the origin of species.

Introduction

In 1998, Carl Woese referred to the microbial world as the "sleeping giant" of biology (Woese, 1998). Almost two decades later, unprecedented attention to our microbial world has turned the fields of zoology (McFall-Ngai et al., 2013) and botany (Turner et al., 2013) inward - towards an increased awareness and understanding of individual animals and plants as holobionts (Bordenstein and Theis, 2015; Gilbert et al., 2012; Zilber-Rosenberg and Rosenberg, 2008). The term "holobiont" denotes a host plus all of its microbial symbionts, including inconstant and constant members that are either vertically or horizontally transmitted or environmentally acquired; it was first coined in 1991 by Lynn Margulis (reviewed in (Gilbert et al., 2012)). The ubiquity and importance of microbes in and on holobionts, including humans, is evident in studies of host development (McFall-Ngai et al., 2012), immunity (Lee and Mazmanian, 2010),

[‡‡] This chapter is published in 2016 in *mBio*, *7*(2), e01785-15 with myself as first author. Seth Bordenstein was senior author.

metabolism (Kamra, 2005; McCutcheon and von Dohlen, 2011; Nicholson et al., 2012; Taylor et al., 2007), behavior (Archie and Theis, 2011; Ezenwa et al., 2012), speciation (Brucker and Bordenstein, 2013a; Wang et al., 2015), and numerous other processes. Host-microbe interactions provide the holobiont with disadvantages (Morgan et al., 2005; Nougue et al., 2015; Polin et al., 2014) such as increasing the risk of cancer (Kodaman et al., 2014), and advantages (Chung et al., 2012; McFall-Ngai et al., 2012; Round and Mazmanian, 2009; Teixeira et al., 2008) such as driving the evolution of resistance to parasites and pathogens (Hornett et al., 2014; Oliver et al., 2008; Rigaud and Juchault, 1993), and among other things producing signal components (i.e., metabolites) used to recognize differences in potential mates (Leclaire et al., 2014; Venu et al., 2014).

The newfound importance of diverse microbial communities in and on animals and plants led to the development of the hologenome theory of evolution (Rosenberg and Zilber-Rosenberg, 2013; Zilber-Rosenberg and Rosenberg, 2008). The "hologenome" refers to all of the genomes of the host and its microbial symbionts, and the theory emphasizes that holobionts are a level of phenotypic selection in which many phenotypes are produced by the host and microbial members of the holobiont. This developing scientific framework distinguishes itself by placing importance not only on well-studied primary microbial symbionts and vertical microbial transmission, but also on the vast diversity of host-associated microbes and horizontal microbial transmission. The key reason for aligning these different transmission modes and levels of complexity into an eco-evolutionary framework is that the community-level parameters among host and symbionts in the holobiont (e.g., community heritability, selection and coinheritance) can be analyzed under a common set of concepts to the parameters that occur in the nuclear genome (Bordenstein and Theis, 2015; van Opstal and Bordenstein, 2015).

As natural selection operates on variation in phenotypes, the hologenome theory's most significant utility is that it reclassifies the target of "individual" selection for many animals and plants traits to the holobiont community. This claim is straightforward given the overwhelming influence of microbes on host traits (Berg et al., 2016; Gilbert et al., 2015; McFall-Ngai, 2015; Tsuchida et al., 2010). The question going forward is whether the response to this community-level selection is relevant to the biology of holobionts. In other words, can host-associated microbial communities be selected such that shifts in the microbial consortia over multiple generations are a response to selection on holobiont traits? Community selection at the holobiont

level is shaped by genetic variation in the host and microbial species and covariance between hosts and their microbial consortia, the latter of which can be driven by (i) inheritance of the microbial community from parents to offspring (Funkhouser and Bordenstein, 2013; Gilbert, 2014) and/or (ii) community heritability H^2_C (Shuster et al., 2006; van Opstal and Bordenstein, 2015). We recently summarized ten foundational principles of the holobiont and hologenome concepts, aligned them with pre-existing theories and frameworks in biology, and discussed critiques and questions to be answered by future research (Bordenstein and Theis, 2015).

In the context of the widely accepted Biological Species Concept (Dobzhansky, 1937; Mayr, 1942), the principles of holobionts and hologenomes offer an integrated paradigm for the study of the origin of species. The Biological Species Concept operationally defines species as populations no longer capable of interbreeding. Reproductive isolation mechanisms that prevent interbreeding between holobiont populations are either prezygotic (occurring before fertilization) or postzygotic (occurring after fertilization). In the absence of reproductive isolation and population structure, unrestricted interbreeding between holobiont populations will homogenize populations of their genetic and microbial differences (Bordenstein and Theis, 2015). While postzygotic isolation mechanisms include hybrid sterility or inviability, prezygotic isolation mechanisms can include biochemical mismatches between gametes and behavior mismatches between potential partners.

Symbionts can cause prezygotic reproductive isolation in two modes: broad-sense and narrow-sense (Brucker and Bordenstein, 2012a). Broad-sense symbiont-induced reproductive isolation refers to divergence in host genes that result in a reproductive barrier because of selection on the host to accommodate microorganisms. In this case, loss or alteration of the symbiont does not have an impact on the capacity to interbreed; rather host genetic divergence and reproductive isolation evolve in response to microbial symbiosis and cause isolation regardless of whether the hosts are germ-free or not. Conversely, narrow-sense symbiont-induced reproductive isolation occurs when host-microbe or microbe-microbe associations result in a reproductive barrier, namely one that can be ameliorated or removed via elimination of the microbes. Therefore, narrow-sense isolation can be experimentally validated if it is reversible under microbe-free rearing conditions and inducible with the reintroduction of microbes. Isolation barriers that require host and microbial component underpin hologenomic speciation (Bordenstein and Theis, 2015; Brucker and Bordenstein, 2013a).

We recently synthesized the literature and concepts of various speciation mechanisms related to symbiosis, with notable attention to postzygotic isolation (Brucker and Bordenstein, 2012a, 2012b, 2013a). While aspects of the microbiology of prezygotic isolation are less understood, seminal cases exist (Koukou et al., 2006; Miller et al., 2010; Sharon et al., 2010) and control of behavior by symbionts is an emerging area of widespread interest (Ezenwa et al., 2012; O'Mahony et al., 2017; Sampson and Mazmanian, 2015). Here we emphasize the patterns seen in these new and old analyses (Table 1) and highlight important and tractable questions about the microbiome, behavior, and speciation by symbiosis. For the purposes of this review, we refer to the microbiome as the community of microorganisms in and on a host.

Signaling & microbiome homogenization

Recognizing signals of species membership (Carlson et al., 1976), gender (Cator et al., 2009), relatedness (Lize et al., 2013), and colony or group membership (Matsuura, 2001) is relevant to choosing a mate. Visual (Carlson et al., 1976), auditory (Cator et al., 2009), and chemosensory signals (De Cock and Matthysen, 2005) can each be used to relay this information, with the latter being particularly influenced by the microbiome in either "microbe-specific" or "microbe-assisted" ways. Both mechanisms involve the expression of chemosensory cues, but microbe-specific processes involve bacterial-derived products such as metabolites while microbe-assisted mechanisms involve bacterial modulation of host-derived odorous products (Figure F-1).

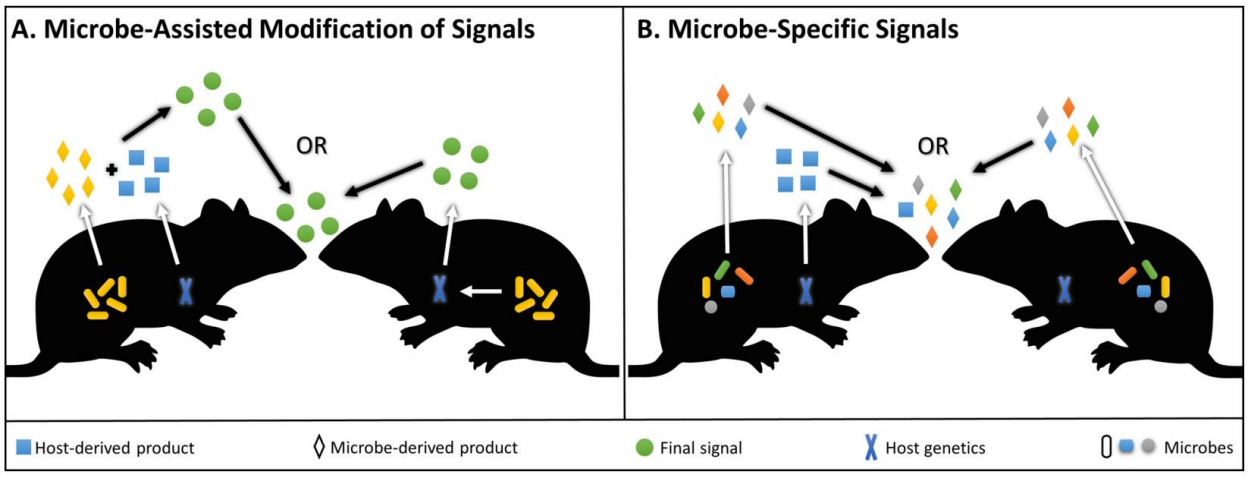

Figure F-1. Microbe-assisted and microbe-specific signaling.
(A) Microbe-assisted processes denote the production of a host signal with input from the microbiome. It occurs in two possible scenarios. On the left, the host and microbial symbionts produce products that interact or combine to form a signaling compound; on the right, microbial symbionts modify host signal expression, but they do not make a specific product directly involved in the signal itself. (B) Microbe-specific processes denote the production of a

microbial signal without input from the host. It occurs in two possible scenarios. On the left, the host and microbial symbionts produce products that are both required to elicit a response; on the right, microbial symbionts produce compounds used by the host for signaling. Mouse image source: Wikimedia Commons, Angelus (https://commons.wikimedia.org/wiki/File:Rat_2.svg).

The microbiome's capacity to provide identity information for mate recognition may rely on products being an honest signal of holobiont group membership, requiring that many or all members of the group (i.e., gender, population or species) contain appropriate microbial members that express equivalent signal profiles. Holobionts can be colonized by similar microbes via a number of different mechanisms, spanning behavioral similarities and contact with shared environmental sources (Lax et al., 2014; Tung et al., 2015), similar ecological niches and diets (Ley et al., 2008; Spor et al., 2011; Yatsunenko et al., 2012), and host genetic effects (Brucker and Bordenstein, 2013a; McKnite et al., 2012). Each of these mechanisms may explain a portion of the variation in the microbial communities of holobionts (Brucker and Bordenstein, 2013b, 2012a; Franzenburg et al., 2013; Ochman et al., 2010; Sanders et al., 2014).

In the context of group living, humans in the same household (Lax et al., 2014; Song et al., 2013) and chimpanzees (Degnan et al., 2012) or baboons (Tung et al., 2015) in the same social group have more similar microbial communities than non-group members. Among several mammalian species, microbial community composition covaries with odorous secretions, and similarities are shared based on host age, sex, and reproductive status allowing for potential signaling and recognition of these traits (Leclaire et al., 2014; Theis et al., 2013). In hyenas, there is less microbial community variation within species than between them, and clans have more comparable microbial communities due to the marking and remarking of collective territory to signal clan ownership (Theis et al., 2013). In baboons, there is less microbiome variation within social groups than between them, and baboons involved in communal grooming behaviors share even more similarities (Tung et al., 2015). Insect populations such as termites can stabilize their gut microbiomes by way of trophallaxis, a behavior in which nestmates supply nutrients and microbes (e.g., cellulolytic microbes) to other colony members through fluids they excrete from their hindgut (Klass et al., 2008). However, Tung *et al* appropriately note, "*one of the most important unanswered questions is whether social network-mediated microbiome sharing produces net fitness benefits or costs for hosts*" (Tung et al., 2015). From the perspective of the origin of species, it will be similarly important to determine if fitness impacts of the microbiome in turn affect the evolution of group living and reproductive isolation. On one hand, socially-shared

microbiomes could drive the evolution of population-specific mating signals and ensuing behavioral isolation. On the other hand, they could fuse incipient species in sympatry that socially share bacterial communities responsible for mating signals.

Similarities in diet can also influence microbiome homogenization, particularly in the digestive tract. For instance, *Drosophila melanogaster* reared on similar food sources carry comparable microbial communities (Sharon et al., 2010). Trophically similar ant species also share microbial species (Anderson et al., 2012). In humans, gut microbiome variation in taxonomy and functions correlates with dietary variation (Siddharth et al., 2013), and alterations in human diet can rapidly and reproducibly change the structure of the microbiome (Claesson et al., 2012; David et al., 2014). Seasonal variation in wild howler monkey diet is also correlated to shifts in the microbiome (Amato et al., 2015). Mediterranean fruit flies (Ben-Yosef et al., 2008) and olive flies (Ben-Yosef et al., 2010) acquire microbes from their food that increase clutch size and oviposition rate of females exposed to diets lacking essential amino acids (Ben-Yosef et al., 2008; Ben-Yosef et al., 2010). Intriguingly, male sexual competitiveness of Mediterranean fruit flies increases up to two-fold with diets enriched with *Klebsiella ozytoca* versus a conventional diet (Gavriel et al., 2011).

Host genetics also affects microbial community assembly. In mice, there are 18 candidate loci for modulation and homeostatic maintenance of Bacteroidetes, Firmicutes, Rikenellaceae, and Provetellaceae in the gut (Benson et al., 2010; McKnite et al., 2012). Moreover, the presence of many rare bacterial groups in the gills of the Pacific oyster are correlated to genetic relatedness (Wegner et al., 2013). Congruently, genetic variability in human immune-related pathways are associated with microbial profiles on several body sites including various locations along the digestive tract (Blekhman et al., 2015), and the largest twin cohort to date examined members of the gut microbiome and found that the bacterial family Christensenellaceae has the highest heritability ($h^2 = 0.39$), and associates closely with other heritable gut bacterial families (Goodrich et al., 2014). Human genetic background also influences the risk of developing gastric cancer caused by *Helicobacter pylori*, indicating that incompatibilities between hosts and symbionts can produce deleterious effects (Kodaman et al., 2014). Phylosymbiosis, characterized by microbial community relationships that reflect host phylogeny (van Opstal and Bordenstein, 2015), has also been reported in several cases. For instance, closely related *Nasonia* species that diverged roughly 400,000 years ago share more similar microbial communities than species pairs that diverged a

million years ago (Brucker and Bordenstein, 2013a, 2012a). Similar phylosymbiotic patterns are observed in hydra (Franzenburg et al., 2013), ants (Sanders et al., 2014) and primates (Ochman et al., 2010).

The overall complexity inherent in microbial community structures and processes may be problematic for animal holobionts seeking to interpret a vast array of signaling information. However, recognition and differentiation of these microbe-induced signals may be possible if a subset of the microbiome affects the production of the particular signal. Furthermore, it may also be challenging to disentangle social, environmental, and diet effects on microbial assemblages in natural populations (Tung et al., 2015). Nonetheless, the important theme among all of these cases is that microbial community variation often appears to be less within holobiont groups/species than between them. This pattern, if sustained in natural populations, could facilitate the evolution of microbe-specific and/or microbe-assisted mating signals that promote recognition within populations or species and discrimination between them. Once this critical point is passed, speciation has commenced. There are parallels here with inclusive fitness theory, which posits that individuals can influence their own reproductive success or the reproductive success of other individuals with which they share genes (Hughes et al., 2008; West et al., 2007). If one follows the continuity from genes to microbial symbionts, then the inclusive fitness framework may also apply to holobionts in which specific microbial symbionts may influence their reproductive success by increasing the reproductive success of their hosts through microbe-specific and/or microbe-assisted mating. A case-by-case analysis of the reliance of the symbiont on the host for transmission (e.g., maternal, social, environmental transmission) will augment the relevance of this framework.

Microbe-assisted modification of mating signals

A common, microbe-assisted modification involves manipulation of host signals (Figure F-1a). One seminal study found that *D. melanogaster* acquires more *Lactobacillus* when reared on starch than on a molasses-cornmeal-yeast mixture (Dodd, 1989; Sharon et al., 2010). The increased *Lactobacillus* colonization correlates with an upregulation of 7,11-heptacosadiene, a cuticular hydrocarbon sex pheromone in the female fly, resulting in an ability to distinguish fly holobionts raised in the starch environment from those reared on the molasses-cornmeal-yeast substrate (Ringo et al., 2011; Sharon et al., 2010). This microbe-assisted positive assortative mating is

reproducible, reversible, and maintained for several dozen generations after diet homogenization (Najarro et al., 2015; Sharon et al., 2010). Moreover, this diet-dependent homogamy appears to be directly mediated by different gut bacteria, as inoculation of germ-free flies with *Lactobacillus* causes a significant increase in mating between flies reared on the different diets (Sharon et al., 2010). Replication of these experiments found that inbred strains specifically followed this mating pattern (Najarro et al., 2015). Moreover, another *D. melanogaster* study involving male mate choice and antibiotics revealed that female attractiveness is mediated by commensal microbes (Arbuthnott et al., 2016). These laboratory studies provide a critical model for how microbe-assisted modifications in a signaling pathway, ensuing behavioral changes, and mating assortment can potentiate behavioral isolation and possibly speciation. Indeed, natural populations of *D. melanogaster* express positive assortative mating and differential signal production based on food sources (Stennett and Etges, 1997), and a bacterial role in these instances should be explored.

Microbe-assisted signaling also occurs in laboratory mice (*Mus musculus*), in which bacterial conversion of dietary choline into trimethylamine (TMA) leads to attraction of mice while also repelling rats (Li et al., 2013). Antibiotic treatment decreases TMA production, and genetic knockout of the mouse receptor for TMA leads to decreased attraction in mice (Li et al., 2013). Antibiotic treatment and subsequent depletion of TMA in mice could in turn result in a decrease in repellence of rats (Li et al., 2013), though this possibility has not yet been tested *in vivo*. Another study found that female mice are more attracted to males not infected with *Salmonella enterica* infected compared to those that are, yet females mated multiply and equally in mating choice tests with the two types of males (Zala et al., 2015).

Mate preference based on infection status fits well with the Hamilton-Zuk hypothesis of parasite-mediated sexual selection, which posits that traits related to infection status can influence mating success (Hamilton and Zuk, 1982). One seminal study showed that male jungle fowl infected with a parasitic roundworm produce less developed ornamentation and are less attractive to females (Hamilton and Zuk, 1982). In house finches, male plumage brightness indicates their quality of broodcare and is associated with resistance to the bacterial pathogen *Mycoplasma gallicepticum* (Hill and Farmer, 2005). The Hamilton-zuk hypothesis has been reviewed in detail (Balenger and Zuk, 2014).

Microbe-specific signals

Microbe-specific signals frequently involve the release of volatile microbial metabolites, often through excretions from specialized glands on the host's body (Figure F-1b). Microbial volatiles can transmit information utilized for social signaling (Archie and Theis, 2011; Ezenwa and Williams, 2014) and intra- or interspecies mate recognition (Leal, 1998; Li et al., 2013). For example, beetles (Leal, 1998), termites (Matsuura, 2001), nematodes (Meisel et al., 2014), hyenas (Theis et al., 2013), meerkats (Leclaire et al., 2014), and badgers (Sin et al., 2012) produce and recognize bacterial metabolites in communication that can modulate their behavior. In termites, fecal metabolites produced by intestinal bacteria (Matsuura, 2001) coat the termite body and hive walls to signal colony membership. Termite holobionts lacking colony-specific metabolite profiles are attacked and killed by the hive (Matsuura, 2001). In contrast, some beetles and mammal species excrete bacterial metabolites from colleterial and anal scent glands, respectively (Leal, 1998; Leclaire et al., 2014; Theis et al., 2013). For example, female grass grub beetles house bacteria within their colleterial glands peripheral to the vagina that are used to attract males to mate (Leal, 1998).

An exciting area of research regarding microbe-specific bacterial signaling involves mammalian fermentation. The mammalian fermentation hypothesis (Leclaire et al., 2014; Theis et al., 2013) states that fermentative bacteria within mammalian scent glands produce odorous metabolites involved in recognition. For example, hyena subcaudal scent pouches store bacteria that are mostly fermentative (Theis et al., 2013). When marking territory, hyenas deposit species-specific, bacterial-derived volatile fatty acids from this gland onto grass stalks (Theis et al., 2013) . Bacterial metabolite secretions are more variable in the social hyena species, presumably because the complexity of signals from social species improves intraspecies identification (Theis et al., 2013). Alternatively, social hyenas may permissively transmit more diverse bacteria leading to diverse metabolite profiles. Hyena microbiomes also covary with group membership, sex, and reproductive state (Theis et al., 2013). Similarly, bacterial communities in meerkat anal scent secretions vary with host sex, age, and group membership (Leclaire et al., 2014). In both cases, the signal diversity may allow animal holobionts to recognize diverse biotic characteristics.

Humans also carry bacteria related to odor production. Breath (Morita and Wang, 2001; Pianotti and Pitts, 1978), foot (Stevens et al., 2015), and underarm (James et al., 2013) odor covary with oral and skin microbiomes, respectively. Many diseases (e.g., smallpox, bacterial vagi*nos*is,

syphilis, etc.) are associated with distinct odors, and have historically been used by physicians in diag*nos*is (Penn and Potts, 1998). Clothing made from different materials even carry different odor profiles based on material-specific bacterial colonization (Callewaert et al., 2014; Tsuchiya et al., 2008). Male odor has been associated with women's interpretation of a male's attractiveness (Havlicek et al., 2005; Lubke and Pause, 2015; Saxton et al., 2008), possibly influencing their choice in a mate.

The salient theme among the aforementioned cases is that host-associated microbes frequently emit odors, and sometimes this microbe-specific chemosensory information can affect mate choice. Reciprocally, ample evidence shows that chemical signals mediate sexual isolation (Smadja and Butlin, 2009), and a full understanding of whether these signals are traceable to host-associated microbes is worthy of serious attention. Germ-free experiments and microbial inoculations should be a prerequisite for such studies; otherwise they risk missing the significance of microbes in chemosensory speciation (Smadja and Butlin, 2009). Additional behaviors involved in speciation, such as habitat choice and pollinator attraction, are also likely to be influenced by microbe-specific products. Indeed, classic model systems of speciation await further experimentation in this light. For example, food-specific odors on apples and hawthorn translate directly into premating isolation of incipient host races of fruit flies of the genus *Rhagoletis* (Linn et al., 2003). Furthermore, the fruit fly *Drosophila sechellia* exclusively reproduces on the ripe fruit of *Morinda citrifolia*, which is toxic to other phylogenetically-related *Drosophila* species, including *D. melanogaster* and *D. simulans*. Some of the volatile compounds involved in these interactions, such as isoamyl acetate, have been associated with fermentative bacteria like *Lactobacillus plantarum* (Lee et al., 2009), suggesting that food-based premating isolation may be related to bacterial associations with the food source, though this requires further study. In summary, new challenges necessitate the concerted effort of scientists of diverse backgrounds to explore questions at the boundaries of many biological disciplines and to develop the tools to untangle and interpret this intricate web of interactions. Critical topics to be explored in the future include determining the microbial role in animal mate choice, quantifying the extent to which microbe-induced mating assortment impacts the origin of species, and identifying the mechanisms involved in these interactions.

Endosymbionts and mate choice

Wolbachia, *Spiroplasma*, *Rickettsia*, *Cardinium*, and several other endosymbiotic bacteria can change animal sex ratios and sex determination mechanisms to increase their maternal transmission and thus frequency in the host population from one generation to the next. Notably, these reproductive alterations affect mate choice (Beltran-Bech and Richard, 2014), and here we highlight a few prominent examples and discuss how endosymbiotic bacteria can influence behavioral isolation and the origin of species.

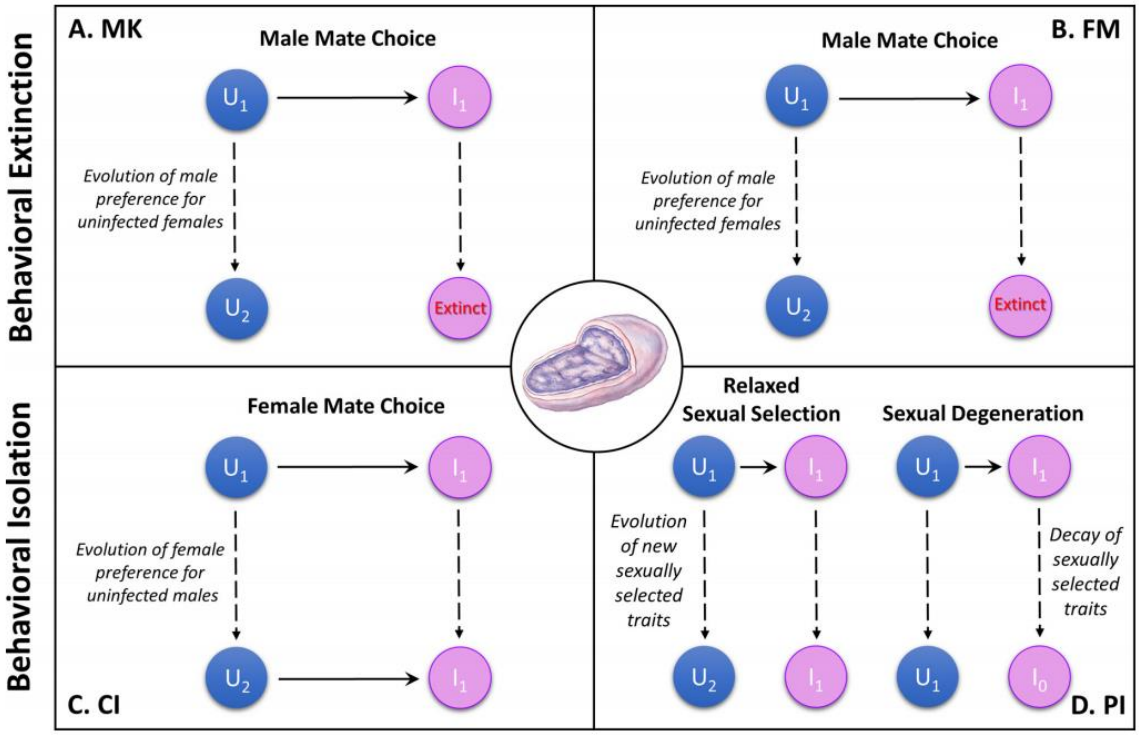

Figure F-2. Endosymbiont-induced behavioral isolation and extinction.
U (blue) and I (pink) represent the uninfected and infected populations, respectively. Horizontal solid arrows represent the direction of gene flow (from males to females), and vertical dashed arrows represent divergence time. Different subscript numbers for U and I represent evolutionary change in traits involved in behavioral extinction and behavioral isolation. (A and B) Behavioral changes induced by male killing (MK) (A) and feminization (FM) (B) evolve in response to selection on uninfected males to mate preferentially with uninfected females. If male preference is completely penetrant, then total loss of mating between the uninfected and infected population ensues, effectively leading the infected population to extinction, since infected females rely on (the now discriminating) uninfected males to reproduce. We term this model "behavioral extinction." (C and D) In contrast, behavioral changes induced by cytoplasmic incompatibility (CI) (C) and parthenogenesis induction (PI) (D) can result in reduced or no gene flow between the infected and uninfected populations. CI-assisted reproductive isolation can be enhanced by the evolution of mate discrimination and specifically uninfected female mate choice for uninfected males. While this model does not sever gene flow in reciprocal cross directions, asymmetric isolation barriers can act as an initial step in speciation. PI-assisted reproductive isolation is mediated by two possible mechanisms: (i) sexual degeneration which involves the degeneration of sexual traits in the infected population that ultimately lock the populations into uninfected sexual and infected parthenogenetic species, and (ii) relaxed sexual selection which involves the evolution of new sexual characteristics in the uninfected sexual population that prevent mating with the infected parthenogenetic population. *Wolbachia* image source: Tamara Clark, Encyclopedia of Life, *Wolbachia* page (http://eol.org/data_objects/466412).

Cytoplasmic incompatibility.

Wolbachia are the most well-studied reproductive distorters (Kageyama et al., 2012; Werren et al., 2008) and are estimated to infect approximately 40% of all arthropod species (Zug and Hammerstein, 2012). Across the major insect orders, *Wolbachia* cause cytoplasmic incompatibility (CI), a phenomenon in which *Wolbachia*-modified sperm from infected males leads to post-fertilization embryonic lethality in eggs from uninfected females or from females infected with a different strain of *Wolbachia*, but not in eggs from infected females (Serbus et al., 2008).

In this context, *Wolbachia*-induced CI can promote the evolution of mate discrimination between populations or species because females can be selected to avoid males that they are not compatible with (Figure F-2c). Among closely related species of mushroom-feeding flies, *Wolbachia*-infected *Drosophila recens* and uninfected *D. subquinaria* contact each other and interspecifically mate in their sympatric range in Eastern Canada. However, gene flow between them in either cross direction is severely reduced due to the complementary action of CI and behavioral isolation. *Wolbachia*-induced CI appears to be the agent for evolution of behavioral isolation as asymmetric mate discrimination occurs in flies from the zones of sympatry but not in flies from the allopatric ranges (Jaenike et al., 2006). A similar pattern of *Wolbachia*-induced mate discrimination occurs among strains of the two-spotted spider mite, *Tetranychus urticae* (Vala et al., 2004) and *D. melanogaster* cage populations (Koukou et al., 2006). Moreover, discrimination between particular semispecies of *D. paulistorum* is associated with their *Wolbachia* infections (Miller et al., 2010). In cases where host populations or species harbor different *Wolbachia* infections that are bidirectionally incompatible, for example in different *Nasonia* species that exist sympatrically (Bordenstein et al., 2001; Bordenstein and Werren, 2007), reciprocal mate discrimination has evolved (Bordenstein and Werren, 2007; Telschow et al., 2005). In contrast to these examples, interspecific mate discrimination in *Nasonia giraulti* is diminished when non-native transfections of *Wolbachia* spread throughout the whole body including to the brain, suggesting that *Wolbachia* can also inhibit pre-existing mate discrimination (Chafee et al., 2011).

These cases reveal, to varying degrees, that *Wolbachia* can be causal to the evolution of assortative mating within and between species. Indeed, population genetic theory demonstrates that mate choice alleles spread quicker in populations or species with CI than those with nuclear incompatibilities (Telschow et al., 2005). This is primarily due to the dominance of these

Wolbachia-induced incompatibilities since CI causes F1 inviability, while nuclear incompatibilities are typically expressed in the F2 hybrids due to the recessive nature of hybrid incompatibility alleles.

Male killing.

Male killing is the most common form of endosymbiont-induced sex-ratio manipulation and can occur during embryonic (Elnagdy et al., 2011; Hurst et al., 1993) or larval development (Hurst and Pomiankowski, 1991; Nakanishi et al., 2008). The effect of male killing is to increase the number of female hosts in a population, thereby increasing endosymbiont transmission rates. To prevent complete fixation of females and population extinction (Jiggins et al., 2002), selection can favor hosts to (i) suppress male killing via genes that reduce *Wolbachia* densities or functions (Charlat et al., 2007; Gilfillan et al., 2004; Hornett et al., 2014; Veneti et al., 2005) or (ii) electively choose mates whereby uninfected males preferentially mate with uninfected females (Randerson et al., 2000; Rigaud and Moreau, 2004). If mate choice evolves as a behavioral adaptation to avoid male killing, it could begin to splinter infected and uninfected populations and initiate the first steps of the speciation process (Figure F-2a). One significant caveat in this conceptual model is that the infected population will go extinct without uninfected males to mate with. Thus, if mate preference based on infection status was complete, it would cause speciation between the infected and uninfected populations, resulting in the immediate extinction of the infected population that requires uninfected males to reproduce. We term this phenomena "behavioral extinction" (Figure F-2).

Wolbachia-induced male killing can reach a state of equilibrium, as suggested by their long-term maintenance in natural populations of butterflies (Dyson and Hurst, 2004). Discriminatory males occasionally mate with infected females allowing for the infection to remain in the population (Randerson et al., 2000), and eventually an equilibrium is reached (Dyson and Hurst, 2004). However in some cases, the infection rate is high (>95%), and male preference for uninfected females has not been identified (Jiggins et al., 2002). It is not known what mechanisms are involved in preventing male killing from reaching fixation in these situations.

Feminization.

Feminization, or the conversion of genetic males to morphological and functional females, has similar evolutionary consequences to male killing (Figure F-2b). This process occurs in many different arthropod species including butterflies (Hiroki et al., 2004; Narita et al., 2007), leafhoppers (Negri et al., 2006), and woodlouse (Moreau et al., 2001). Resistance to these effects in the pillbug *Armadillidium vulgare* has evolved in the form of feminization suppressors and male preference towards uninfected females. Males that mate with infected females produce feminized males (Rigaud and Juchault, 1993, 1992). Ultimately, a female-biased sex-ratio in feminized woodlouse populations results in an increase in male mate choice, male mating multiplicity, and sperm depletion. In the context of sperm depletion, initial mating encounters are normal, but upon increased mating frequency, males provide less sperm to subsequent females. Moreover, infected females are curiously less fertile at lower sperm densities possibly because they are less efficient at utilizing small quantities of sperm (Rigaud and Moreau, 2004). Insufficient sperm utilization and slight differences in infected female courtship behaviors can result in male preference for uninfected females within the population (Moreau et al., 2001). Just as with male killing, assortative mating within infected and uninfected populations may initiate the early stages of speciation and lead to behavioral extinction (Figure F-2)

Parthenogenesis.

Microbial-induced parthenogenesis is common among haplodiploid arthropods such as wasps, mites, and thrips (Arakaki et al., 2001; Stouthamer, 1997; Stouthamer et al., 1999), wherein unfertilized eggs become females (Adachi-Hagimori et al., 2008; Dunn et al., 1995). As we previously discussed (Bordenstein, 2003), parthenogenesis-induced speciation by endosymbiotic bacteria falls neatly with the Biological Species Concept because parthenogenesis can sever gene flow and cause the evolution of reproductive isolation between sexual and asexual populations. Microbe-induced parthenogenesis does not necessarily exclude sexual capability of parthenogenetic females, but instead removes the necessity of sexual reproduction and can potentially drive divergence in sexual behaviors and mate choice (Stouthamer et al., 1990). Speciation therefore commences between sexual and asexual populations under two models: (i) Sexual Degeneration and (ii) Relaxed Sexual Selection (Bordenstein, 2003) (Figure F-2d).

The Sexual Degeneration model posits that the asexual population becomes incompetent to engage in sexual interactions due to mutational accumulation and thus trait degeneration while the sexual population remains otherwise the same (Bordenstein, 2003). In this case, parthenogenetic lineages accumulate mutations in genes involved in sexual reproduction. Traits subject to mutational meltdown may span secondary sexual characteristics, fertilization, mating behavior, signal production, among others (Gottlieb and Zchori-Fein, 2001; Kraaijeveld et al., 2009; Kremer et al., 2009). For instance, long-term *Wolbachia*-induced parthenogenesis in mealybugs and some parasitoid wasps prevents females from attracting mates or properly expressing sexual behaviors (Kremer et al., 2009; Pijls et al., 1996). Similarly in primarily asexual populations, male courtship behavior and sexual functionality is often impaired (Gottlieb and Zchori-Fein, 2001; Pannebakker et al., 2005; Zchori-Fein et al., 1995). The accrual of these mutations prevents sexual reproduction, thus causing the parthenogenetic population to become "locked in" to an asexual lifestyle. While this model is an attractive hypothesis for the onset of reproductive isolation between asexual and sexual populations, it is not always easily distinguishable from the alternative Relaxed Sexual Selection model (Bordenstein, 2003). In this model, the sexual population diverges by evolving new or altered mating factors (e.g., courtship sequence, signals, etc.) while the asexual population does not degrade, but rather stays the same and thus can no longer mate with individuals from the diverging sexual population (Bordenstein, 2003).

Conclusions

Over the past decade, biology has stood *vis-à-vis* with what Carl Woese referred to as the "sleeping giant" of biology - the microbial world (Woese, 1998). During this period of groundbreaking research, a new vision for the increasing importance of microbiology in many subdisciplines of the life sciences has emerged. As such, studies of animal and plant speciation that do not account for the microbial world are incomplete. We currently know that microbes are involved in a multitude of host processes spanning behavior, metabolite production, reproduction, and immunity. Each of these processes can in theory or in practice cause mating assortment and commence population divergence, the evolution of reproductive isolation, and thus speciation. Understanding the contributions of microbes to behavior and speciation will require concerted

efforts and exchanges among these biological disciplines, namely ones that embrace the recent "unified microbiome" proposal to merge disciplinary boundaries (Alivisatos et al., 2015).

Table F-1. Microbe-induced traits that associate with or cause changes in behavior and barriers to interbreeding.

Traits	Host Species	Common Name	Symbiont(s)	Behavior or reproductive outcome	References
Host signal modification	*Drosophila bifasciata*	Fruit fly	Unknown	Assortative mating based on familiarity	(Lewis et al., 2014)
	Drosophila subobscura	Fruit fly	Unknown	Assortative mating based on kinship	(Lewis et al., 2014)
	Drosophila melanogaster	Fruit fly	*Lactobacilli plantarum*	Assortative mating based on diet	(Ringo et al., 2011; Sharon et al., 2010)
	Mus musculus	House mouse	Unknown gut bacteria	Species recognition	(Li et al., 2013)
Bacterial metabolite production	*D. melanogaster*	Fruit fly	*L. brevis , L. plantarum*	Assortative mating based on diet	(Venu et al., 2014)
	Reticulitermes speratus	Termite	Unknown gut bacteria	Exclusion of non-colony members	(Matsuura, 2001)
	Costelytra zealandica	Grass grub	Unknown bacteria in colleterial glands	Mate attraction	(Leal, 1998)
	Crocuta crocuta	Spotted hyena	Unknown bacteria in anal scent glands	Clan, age, sex, and reproductive status recognition	(Theis et al., 2013)
	Hyaena hyaena	Striped hyena	Unknown bacteria in anal scent glands	Clan, age, sex, and reproductive status recognition	(Theis et al., 2013)
	Meles meles	European badger	Unknown bacteria in	Possible mate discrimination	(Sin et al., 2012)

			anal scent glands		
	Suricata suricatta	Meerkat	Unknown bacteria in anal scent glands	Group, age, and sex recognition	(Leclaire et al., 2014)
Odor production	*M. musculus*	House mouse	*Salmonella enterica*	Initial avoidance of infected males	(Zala et al., 2015)
	Homo sapiens	Humans	Unknown	Attractiveness	(Havlicek et al., 2005; Lubke and Pause, 2015; Saxton et al., 2008)
Cytoplasmic incompatibility	*Drosophila paulistorum*	Fruit fly	*Wolbachia*	Assortment within semispecies	(Miller et al., 2010)
	D. recens & *D. subquinaria*	Fruit fly	*Wolbachia* in *D. recens*	Asymmetric mating isolation	(Jaenike et al., 2006)
	D. melanogaster	Fruit fly	*Wolbachia*	Increased mate discrimination	(Koukou et al., 2006)
	Nasonia giraulti	Parasitoid wasp	*Wolbachia*	Decreased mate discrimination	(Chafee et al., 2011)
	Tetranychus urticae	Two-spotted spider mite	*Wolbachia*	Uninfected females prefer uninfected males	(Vala et al., 2004)
Male killing	*Armadallidium vulgare*	Pillbug	*Wolbachia*	Reduce sperm count and female fertility	(Rigaud and Moreau, 2004)
	D. melanogaster	Fruit fly	*Spiroplasma poulsonii*	Evolved suppressors to prevent male killing	(Veneti et al., 2005)
	Acraea encedon	Common Acraea butterfly	*Wolbachia*	Male mate-choice	(Randerson et al., 2000)
	A.encedon	Common Acraea butterfly	*Wolbachia*	Populations with high infection rates are not discriminatory	(Jiggins et al., 2002)

	Hypolimnas bolina	Great eggfly butterfly	*Wolbachia*	Reduced female fertility	(Charlat et al., 2007; Dyson and Hurst, 2004)
	H. boling	Great eggfly butterfly	*Wolbachia*	Evolved suppressor gene to prevent male killing	(Hornett et al., 2014)
Feminization	*A. vulgare*	Pillbug	*Wolbachia*	Males reproductively female but masculine males prefer true females	((Moreau et al., 2001)
	Eurema hecabe	Grass yellow butterfly	*Wolbachia*	Males reproductively female	(Hiroki et al., 2004; Narita et al., 2007)
	Zyginidia pullula	Leafhopper	*Wolbachia*	Males reproductively female	(Negri et al., 2006)
Parthenogenesis	*Apoanagyrus diversicornis*	Mealybug parasite	*Wolbachia*	Females less attractive to males	(Pijls et al., 1996)
	Asobara japonica	Parasitoid wasp	*Wolbachia*	Females less attractive to males	(Kremer et al., 2009)
	Leptopilina clavipes	Parasitoid wasp	*Wolbachia*	Reduction in male and female sexual traits and fertility	(Kraaijeveld et al., 2009; Pannebakker et al., 2005)
	Muscidifurax uniraptor	Parasitoid wasp	*Wolbachia*	Reduction in sexual traits	(Gottlieb and Zchori-Fein, 2001)
	Neochrysocharis Formosa	Parasitoid wasp	*Wolbachia*	Female biased sex ratio	(Adachi-Hagimori et al., 2008)
	Galeopsomyia fausta	Parasitoid wasp	*Unknown*	Females not receptive	(Argov et al., 2000)
	Franklinothrips vespiformis	Thrips	*Wolbachia*	Male sperm presumably do not fertilize female eggs	(Arakaki et al., 2001)

Appendix G.

Models and nomenclature for cytoplasmic incompatibility: Caution over premature conclusions – A reponse to Beckman et al.[§§]

Recent studies have identified two genes in bacteriophage WO, *cifA* and *cifB*, that contribute to the induction of cytoplasmic incompatibility (CI) (Beckmann et al., 2017; LePage et al., 2017), and one of these two genes, *cifA*, rescues it (Shropshire et al., 2018). These findings underpin a two-by-one genetic model (Figure G-1a) that reflects current understanding of CI genetics and embraces various functional models (Shropshire et al., 2018) (Figure G-1b). A recent article by Beckmann et al. (Beckmann et al., 2019a) provides interesting ideas about the mechanism and evolutionary history of the CI genes. Therein, they claim that it is 'clearer than ever that the CI induction and rescue stem from a toxin–antidote (TA) system', and that disputes regarding the operon status of the cif genes are semantic. They also propose a new nomenclature to describe the genes. It is important to test hypotheses and develop nomenclature carefully in the context of current data because misconceptions can sometimes become a narrative for those unfamiliar with the evidence. Here, we present and evaluate three points of criticism of the arguments related to the TA model, the operon hypothesis, and the proposed gene nomenclature. We recommend caution and nuance in interpreting current data (and lack thereof). As we will frequently note, more research will be necessary before a functional narrative should be prescribed for CI.

The TA model

The proposed TA model (Beckmann et al., 2019a) assumes that male-derived CifB (the presumed toxin) is transferred to the host embryo during fertilization and that its associated defects are rescued upon binding to embryoderived CifA (the presumed antidote). Although CifA and CifB bind to each other in vitro (Beckmann et al., 2017), there is no evidence for transfer of CifB to the embryo. In fact, there is evidence to the contrary that was mentioned by the authors. Mass

[§§] This chapter is published in 2019 in *Trends in genetics: TIG*. 35(6), 397-399 with myself as first author. Brittany Leigh, Sarah Bordenstein, Anne Duplouy, Mark Riegler, and Jeremy Brownlie were co-authors. Seth Bordenstein was senior author.

spectrometry and SDS-PAGE analyses indicate that CifA, but not CifB, is present in the spermatheca of females mated to infected males (Beckmann and Fallon, 2013). There are numerous technical explanations for why CifB is absent – for example, CifB protein expression levels may be below the threshold of detection – but the current evidence is consistent with a model wherein CifA, but not CifB, reaches the female reproductive tract. Therefore, it cannot be assumed that CifB from the male directly interacts with CifA in the embryo. For this reason the essential premise of the TA model is unfounded. One of several alternative hypotheses is that, instead of CifB transferring with the sperm, a host product modified by CifA and/or CifB leads to CI induction, and the host modification is then reversed by CifA in the embryo (Shropshire et al., 2018). It is notable that, if supported, this model and others (Figure G-1b) would contradict the proposed TA model.

The *cif* operon hypothesis

cifA and *cifB* transcriptional regulation is proposed to occur as an operon (Beckmann et al., 2019a), defined as a set of genes that are coregulated by a single promoter (Jacob et al., 1960). However, the number of *cif* promoters is crucially unknown. Moreover, there is a marked ninefold reduction in gene expression of *cifB* relative to *cifA* as well as a predicted hairpin termination element in the short intergenic region (Gutzwiller et al., 2015; Lindsey et al., 2018b). This evidence is consistent with either one or two promoters because the hairpin termination element could contribute to the transcriptional differences. The genes may have two separate promoters with two separate functional transcripts, but aberrant cotranscription could occur owing to an imperfect hairpin terminator. In such a Model, *cifA* and *cifB* would not form an operon. Alternatively, *cifA* and *cifB* may have a single promoter upstream of *cifA*, and the imperfect terminator would provide a mechanism to control the large transcriptional differences between *cifA* and *cifB*. Therefore, we do not see this as a 'semantic debate' but as a hypothesis that requires further testing, and conclusions can only be drawn once firm evidence of a single promoter is established.

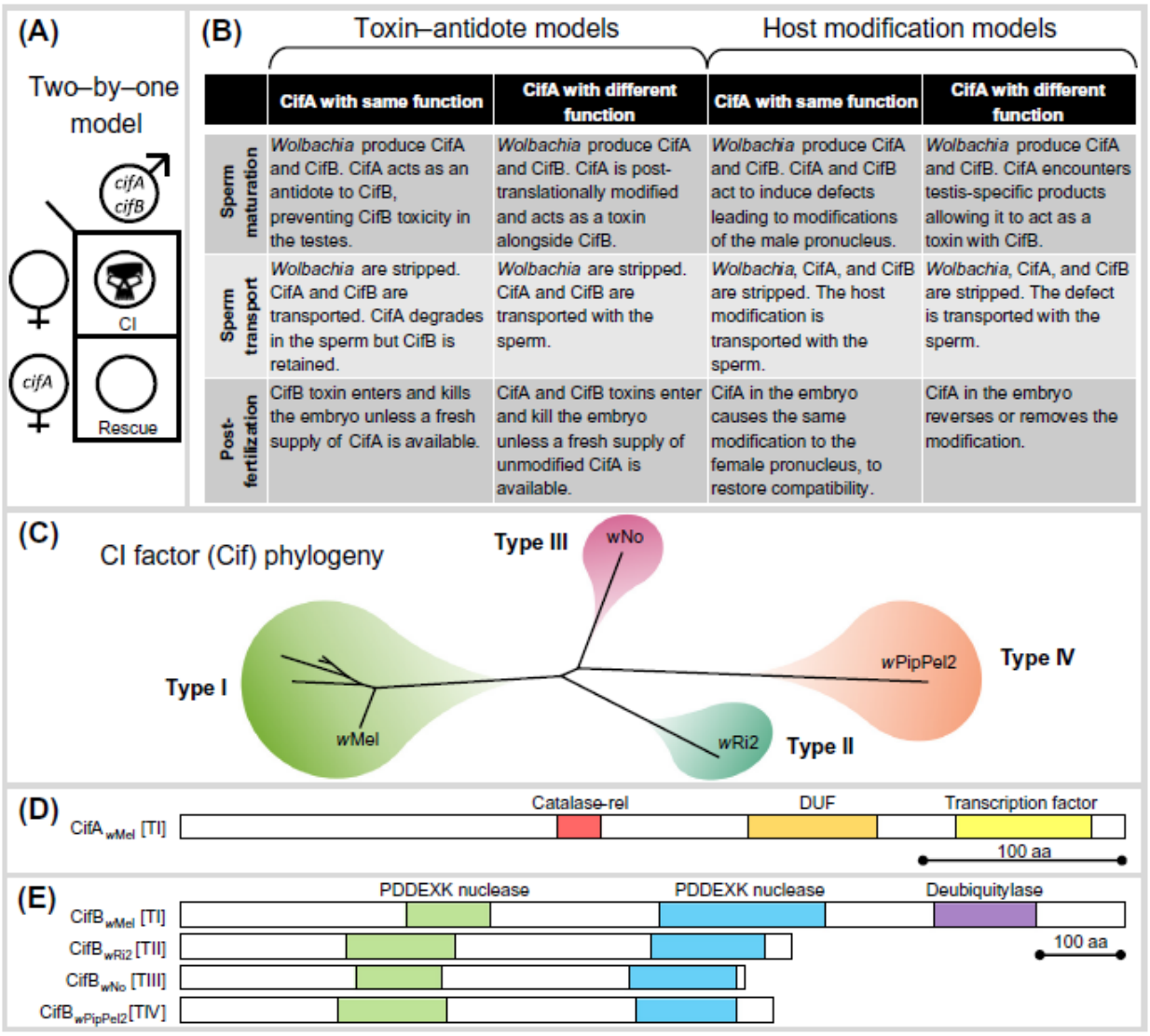

Figure G-1. Models for Cytoplasmic Incompatibility (CI), Phylogenetics, and Annotated Cif Protein Architecture.

(A) The two-by-one model posits that *cifA* and *cifB* expression in males causes CI that can be rescued if *cifA* is expressed in the female/embryo. (B) This genetic model remains conservatively agnostic towards the functional characterization of the CI genes, encompassing multiple toxin–antidote (TA) models and models wherein the host is modified but the Cif products are not transferred from the male to the embryo. We highlight here two examples of TA and host modification models wherein CifA may have the same function in both the testes and embryo, or different functions. The functional description of the models is separated into what takes place during sperm maturation (before individualization), during sperm transport, and after fertilization of the embryo. Importantly, these models do not represent a comprehensive set of possibilities. For example, the ability of CifA ability to act in both instances could be the result of differential localization, post-translational modifications, or comparable functions in both the testes and embryo [3]. (C) The phylogeny of Cif proteins (adapted from [8]) reveals at least four monophyletic clades (types I–IV). Representative alleles are labeled for each clade. (D) Architecture for the CifA protein of *w*Mel in the type I clade is shown with previously annotated domains. (E) Representative CifB proteins from types I–IV clades are shown. All clades have two putative PDDEXK nuclease domains based on structural homology, and type I also has a deubiquitylase domain. Amino acid scale bars are shown. White spaces in protein schematics are unannotated regions. Abbreviations: Catalase-rel, catalase-related; DUF, domain of unknown function.

Gene nomenclature

Useful gene nomenclature should accurately and conservatively describe the phenotype(s) of mutants of the corresponding gene (Demerec et al., 1966). The CI factor (*cif* and *cif*-like) nomenclature conservatively names two separate and codiverging genes involved in CI into at least four clades designated types I–IV (Figure G-1c). Only type I *cif* genes (CI deubiquitylases, *cid* in the competing nomenclature) have been shown to be involved in CI (Beckmann et al., 2017; LePage et al., 2017) and rescue (Shropshire et al., 2018). Homologous genes in distant clades (types II–IV) are therefore denoted *cif*-like to reflect the fact that CI function has not been established (LePage et al., 2017). By contrast, the CI nuclease (*cin*) and CI nuclease/ deubiquitylase (*cnd*) nomenclatures prematurely assigns CI function despite the absence of evidence for a causal role in CI. Moreover, the *cid*, *cin*, and *cnd* nomenclature inaccurately describes gene A (Figure G-1d) based on the purported function of gene B (Figure G-1e) and is not conservative to the unresolved functions of these putatively polyvalent proteins. It would also group types II–IV into one gene category, and type I in another, without any phylogenetic rationality with respect to divergence levels. Similar Gene Names for Similar Functions It would be inaccurate to describe one gene based on the phenotype of another gene unless the genes 'govern related functions' (https://jb.asm.org/sites/default/files/additional-assets/JB-ITA.pdf) or are in an operon. Because CifA has three predicted domains that are completely unrelated to deubiquitylase activity (Figure G-1) (Lindsey et al., 2018b), does not influence the deubiquitylase activity of CifB (Beckmann et al., 2017), and functions independently to rescue CI (Shropshire et al., 2018), it should not be designated as a Cid. Although the Cid nomenclature is based on the operon claim (Beckmann et al., 2019a), the evidence for the operon hypothesis, as explained above, is insufficient and should not be applied to the CI gene nomenclature. The same issue applies to CinA, which does not have nuclease annotations or confirmed phenotypes. Although further research is necessary, the *cif* nomenclature is based on current evidence that identifies both genes as being CI factors, and is therefore accurate because it makes no premature claims about rotein functions across the diversity of alleles in the phylogeny (Figure G-1c).

Concluding remarks

We conclude that, first, there is no evidence for transfer of CifB protein (the presumed toxin of the TA model) from males to the embryo. For this reason, the essential premise of the

proposed TA model is unfounded. Second, evidence for the operon hypothesis is equivocal, and consequently it remains unclear whether CI induction and rescue stem from an operon system. Third, the proposed cid, cin, and cnd gene nomenclature has several weaknesses that are solved with the cif nomenclature.

Appendix H.

Paternal grandmother age affects the strength of *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila melanogaster*[***]

Abstract

Wolbachia are obligate intracellular bacteria that are globally distributed in half of all arthropod species. As the most abundant maternally inherited microbe in animals, *Wolbachia* manipulate host reproduction via reproductive parasitism strategies, including cytoplasmic incompatibility (CI). CI manifests as embryonic death when *Wolbachia*-modified sperm fertilize uninfected eggs but not maternally infected eggs. Thus, CI can provide a relative fitness advantage to *Wolbachia*infected females and drive the infection through a population. In the genetic model *Drosophila melanogaster*, the *Wolbachia* strain *w*Mel induces variable CI, making mechanistic studies in *D. melanogaster* cumbersome. Here, we demonstrate that sons of older paternal *D. melanogaster* grandmothers induce stronger CI than sons of younger paternal grandmothers, and we term this relationship the "paternal grandmother age effect" (PGAE). Moreover, the embryos and adult sons of older *D. melanogaster* grandmothers have higher *Wolbachia* densities, correlating with their ability to induce stronger CI. In addition, we report that *Wolbachia* density positively correlates with female age and decreases after mating, suggesting that females transmit *Wolbachia* loads that are proportional to their own titers. These findings reveal a transgenerational impact of age on *w*Mel-induced CI, elucidate *Wolbachia* density dynamics in *D. melanogaster*, and provide a methodological advance to studies aimed at understanding *w*Mel-induced CI in the *D. melanogaster* model.

Importance

Unidirectional cytoplasmic incompatibility (CI) results in a postfertilization incompatibility between *Wolbachia*-infected males and uninfected females. CI contributes to reproductive isolation between closely related species and is used in worldwide vector control

[***] This chapter is published in 2019 in *mBio*. 10(6) with Emily Layton as first author. Jessie Perlmutter and Seth Bordenstein were co-authors. I was senior author.

programs to drastically lower arboviral vector population sizes or to replace populations that transmit arboviruses with those resistant to transmission. Despite decades of research on the factors that influence CI, penetrance is often variable under controlled laboratory conditions in various arthropods, suggesting that additional variables influence CI strength. Here, we demonstrate that paternal *D. melanogaster* grandmother age influences the strength of CI induced by their sons. Older *D. melanogaster* females have higher *Wolbachia* densities and produce offspring with higher *Wolbachia* densities that associate with stronger CI. This work reveals a multigenerational impact of age on CI and expands our understanding of host-*Wolbachia* interactions and the biology of CI induced by the *Wolbachia* strain infecting the most widely used arthropod model, *D. melanogaster*.

Introduction

Wolbachia are obligate intracellular bacteria that infect 40% to 65% of arthropod species (Hilgenboecker et al., 2008; Weinert et al., 2015; Zug and Hammerstein, 2012) and 37% of the members of the Onchocercidae family of filarial nematodes (Ferri et al., 2011). These bacteria are maternally transmitted from ova to offspring (Serbus et al., 2008) and often cause cytoplasmic incompatibility (CI) to selfishly increase their transmission through the matriline (Hancock et al., 2011; Hoffmann et al., 1990; Jansen et al., 2008; Turelli, 1994; Turelli et al., 2018b). CI manifests as embryonic death when *Wolbachia*modified sperm fertilize uninfected eggs but not when they fertilize infected eggs (LePage and Bordenstein, 2013; Taylor et al., 2018; Yen and Barr, 1973). Thus, infected transmitting females have a fitness advantage relative to their uninfected counterparts that leads to the spread of *Wolbachia* through host populations (Hancock et al., 2011; Hoffmann et al., 1990; Jansen et al., 2008; Turelli, 1994; Turelli et al., 2018b). Additionally, since CI reduces gene flow between *Wolbachia*-infected and uninfected populations or populations with different *Wolbachia* strains, it is associated with reproductive isolation and incipient speciation (Bordenstein et al., 2001; Brucker and Bordenstein, 2012a).

Global vector control efforts have successfully leveraged CI to either suppress native populations (Dobson et al., 2002; Laven, 1967b; Lees et al., 2015; O'Connor et al., 2012) or promote the spread of disease-resistant *Wolbachia* strains (Hoffmann et al., 2011; O'Neill, 2018; O'Neill et al., 2018) specifically through release of mosquitoes transinfected with the *w*Mel *Wolbachia* strain of *Drosophila melanogaster*. *w*Mel's success in these efforts is partially due to the strong CI that it induces in mosquito hosts (Blagrove et al., 2012; Dutra et al., 2015); however,

in the native host *D. melanogaster*, *w*Mel's CI strength can range from an average of nearly 0% (no CI) to 100% (complete CI) (Bourtzis et al., 1996; Hoffmann et al., 1998, 1994; Holden et al., 1993; LePage et al., 2017; Reynolds and Hoffmann, 2002; Solignac et al., 1994; Yamada et al., 2007). There are numerous factors reported to impact the penetrance of *w*Mel-induced CI: *Wolbachia* density in the testes (Bourtzis et al., 1996; Clark et al., 2003), expression level of the CI genes *cifA* and *cifB* (LePage et al., 2017; Shropshire et al., 2018), male age (Reynolds and Hoffmann, 2002), male mating rate (De Crespigny et al., 2006; Reynolds and Hoffmann, 2002), time of male emergence (Yamada et al., 2007), fly rearing density (Yamada et al., 2007), and temperature (Reynolds and Hoffmann, 2002). However, these factors are not independent, and they have likely hampered the researcher's ability to use the vast resources of *D. melanogaster* for the study of reproductive parasitism and endosymbiosis. For example, CI strength rapidly decreases with male age (Reynolds and Hoffmann, 2002), which also cocorrelates with *cifA* and *cifB* gene expression (LePage et al., 2017) and *Wolbachia* density in the testes (Clark et al., 2003).

Despite control of male age, time of emergence, rearing density, and temperature, we continued to see various levels of CI strength in our laboratory, suggesting that additional factors are involved. This variation in phenotype makes *w*Mel in *D. melanogaster* difficult to study despite the fly's extensive history as a powerful animal model. However, anecdotal observations in our laboratory suggested that stronger CI was induced in embryos when their infected paternal grandmothers were significantly aged before mating. Here, we used hatch rate analyses to formally test the hypothesis that paternal grandmother age influences the strength of CI induced by her sons. We also measured the effect of age and virginity on female *Wolbachia* titers and assessed whether females with higher *Wolbachia* titers deposited more *Wolbachia* into their progeny. Our results reveal a "paternal grandmother age effect" (PGAE) on CI strength, where older grandmothers produce males that induce stronger CI. We also characterize transgenerational *Wolbachia* density dynamics that correlate with CI penetrance. This work enhances our understanding of *Wolbachia*-host dynamics and provides methodological techniques of importance to studies of *w*Mel-induced CI in *D. melanogaster*.

Results

To test the hypothesis that *D. melanogaster* paternal grandmother age influences the strength of CI, we measured the percentage of surviving offspring produced by sons of

differentially aged, infected y^1w^* grandmothers. CI strength increased with grandmother age when uninfected females were mated to infected sons of 2-, 5-, 11-, 14-, and 18-day-old grandmothers (Figure H-1). Sons of 2-day-old grandmothers produced statistically weaker CI than those of either 14-day-old (P = 0.0031) or 18-day-old (P = 0.0005) grandmothers, and the same was true for sons of 5-day-old grandmothers compared to those of either 14-day-old (P = 0.0095) or 18-day-old (P = 0.0018) grandmothers. Importantly, sons of 11-day-old uninfected grandmothers produced high hatch rates (Figure H-1), suggesting that the reduction in hatch rate in the remaining crosses was not associated with further aging of the flies. Together, these data suggest that CI is strongest in sons of older grandmothers (Figure H-1).

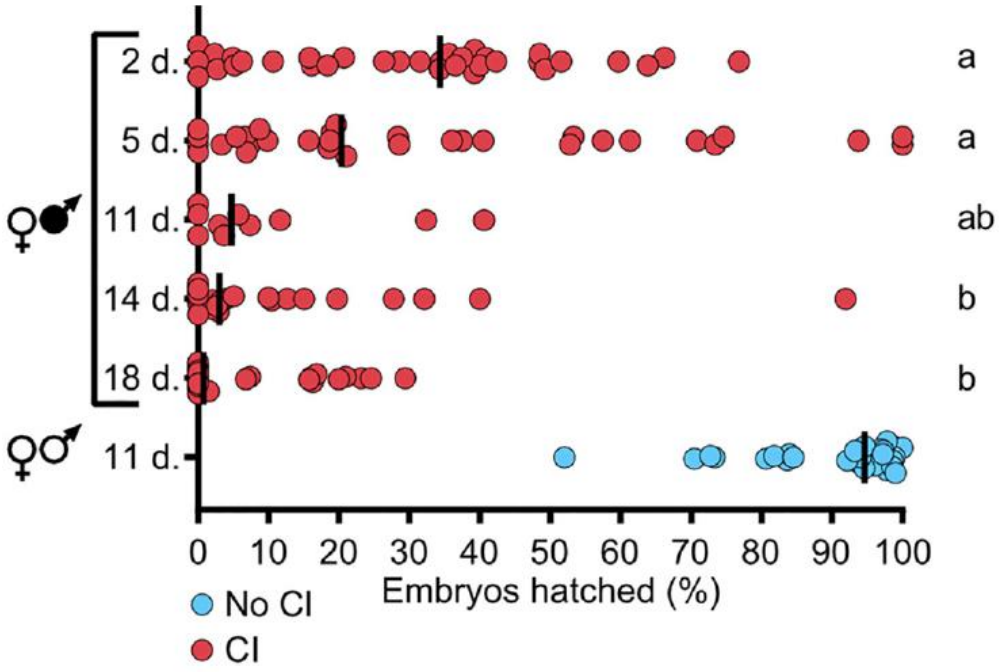

Figure H-1. Paternal grandmother age effect impacts CI strength.
Hatch rate assays were conducted with either uninfected y^1w^* males derived from uninfected females aged 11 days (d.) before mating or infected y^1w^* males derived from infected females aged 2, 5, 11, 14, or 18 days. *Wolbachia* infections are represented by filled sex symbols, and the age of the paternal grandmother is shown immediately to the left of the y axis. Each dot represents a replicate of offspring from single-pair matings. Vertical bars represent medians, and letters to the right indicate significant differences based on alpha = 0.05 calculated by a Kruskal-Wallis test followed by a Dunn's multiple-comparison test performed between all CI crosses. All statistical values are presented in Table S1 in the supplemental material.

Next, we tested whether the increase in embryonic death with *D. melanogaster* grandmother age indeed represented CI and not some other transgenerational embryonic defect. In accordance with prior results (Figure H-1), there was an overall trend indicating that older

grandmothers produced sons that induced stronger CI. Indeed, sons derived from 2-day-old infected grandmothers induced statistically weaker CI than sons of 11-day-old (P = 0.0008) and 14-day-old (P = 0.0110) grandmothers (data not shown). Sons of 11-day-old grandmothers produced a lower median hatch rate than sons of 14-day-old grandmothers; however, the differences were not statistically significant (P = 0.9999). As expected for CI rescue, high rates of embryonic hatching were observed when infected females were mated to sons of infected 2-, 5-, 11-, and 14-day-old grandmothers and the rates did not differ significantly between groups (P = 0.3705). Together, these results suggest that the PGAE is not attributable to other transgenerational, age-associated defects.

To test if the PGAE is specific to the y^1w^* strain, these experiments were repeated in a *nos*-GAL4-*tubulin* genetic background. The *nos*-GAL4-*tubulin* line was chosen because it was previously used to identify the *cifA* and *cifB* genes that underpin *w*Mel-induced CI (LePage et al., 2017). The 2-, 5-, and 11-day time points were selected because they had demonstrated the greatest differences in hatch rate in the previous experiments. As predicted, CI strength correlated with the age of paternal grandmothers when uninfected *nos*-GAL4-*tubulin* females were mated to infected sons of 2-, 5-, and 11-day-old *nos*-GAL4-*tubulin* grandmothers (data not shown). Sons of 11-day-old grandmothers induced significantly stronger CI than sons of 2-day-old grandmothers (P = 0.0033), suggesting that the PGAE is not specific to y^1w^* flies. When sons of uninfected grandmothers aged 2, 5, or 11 days were mated to uninfected females, there were no statistically significant differences in hatching rates across all three groups (P = 0.3907), indicating that the PGAE is CI associated in *nos*-GAL4-*tubulin* flies as seen with y^1w^* flies.

Since *Wolbachia* densities are positively associated with CI strength (Bourtzis et al., 1996; Boyle et al., 1993; Clark et al., 2002; Veneti et al., 2004), we then tested the hypothesis that infected sons derived from older *D. melanogaster* grandmothers have higher *Wolbachia* densities than infected sons from younger grandmothers. We did so by measuring the abundance of the single-copy *Wolbachia groEL* gene relative to that of the *Drosophila rp49* housekeeping gene. Abdomen samples were taken from virgin male siblings of those used in the hatch rate experiment represented in Figure H-1. As predicted, *Wolbachia* densities in male abdomens positively correlated with paternal grandmother age, and sons of 18-day-old grandmothers had significantly higher *Wolbachia* densities than sons of 2-day-old grandmothers (P = 0.0450) (Figure H-2a). However, no significant differences were observed between sons of 5-, 11-, or 14-day-old

grandmothers relative to any other group, presumably due to the variable penetrance of CI, low sample sizes, or biological reasons proposed in the Discussion. Taken together, these data suggest that older grandmothers produced sons with higher *Wolbachia* titers, which allowed the sons to induce stronger CI, though this density effect was weak relative to the effect that we see for CI.

Next, we tested the hypothesis that embryos from older *D. melanogaster* grandmothers had higher *Wolbachia* titers than those from younger grandmothers. *Wolbachia* densities were measured in 0-to-1-h-old embryos produced by both 2-day-old and 11-day-old grandmothers (Figure H-2b). The 2-day and 11-day time points were chosen because they exhibited the greatest differences in CI strength over the shortest time interval. Here, embryos produced by 11-day-old grandmothers had significantly higher *Wolbachia* densities than embryos from 2-day-old grandmothers (P = 0.0006) (Figure H-2b). Thus, these data indicate that older females produce embryos with higher *Wolbachia* titers.

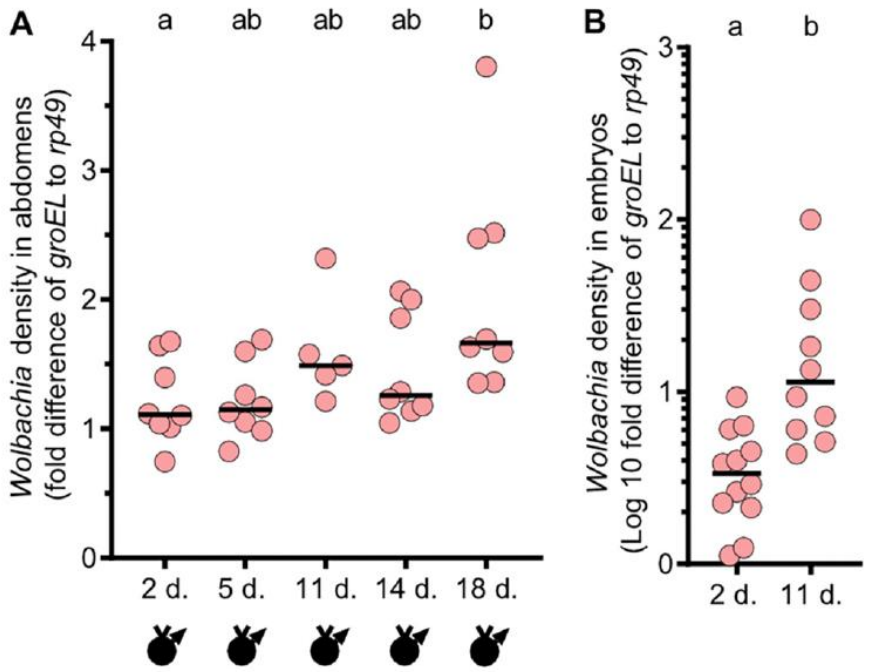

Figure H-2. *Wolbachia* densities are highest in sons and embryos of older *D. melanogaster* grandmothers.
(A) *Wolbachia* density assays were conducted with virgin females (indicated by a "v" above a sex symbol) and with infected y^1w^* males derived from grandmothers aged 2, 5, 11, 14, or 18 days (d.). *Wolbachia* infections are represented by filled sex symbols, and the age of the grandmother is shown immediately below the x axis. The samples analyzed were from abdomens of siblings of fathers corresponding to the hatch rate data in Figure H-1. (B) *Wolbachia* density assays were conducted with pools of 50 1-to-2-h-old embryos collected from 2-and-11-day-old grandmothers. The sex of the embryos was unknown since it cannot be determined visually. *Wolbachia* titers were lower in adults, requiring a standard linear scale (A), but higher in embryos, requiring a common logarithmic scale (B). Each dot represents the average of results from triplicate technical replicates for panel A and duplicates for panel B. Horizontal

bars indicate medians, and the letters above the bars indicate significant differences based on alpha = 0.05 calculated by Kruskal-Wallis test followed by a Dunn's multiple-comparison test performed between all groups (A) or by a Mann Whitney U test (B). All statistical values are presented in Table S1. Fold differences in *Wolbachia* densities (*groEL*) relative to *D. melanogaster* reference gene rp49 were determined with $2^{-\Delta\Delta Ct}$.

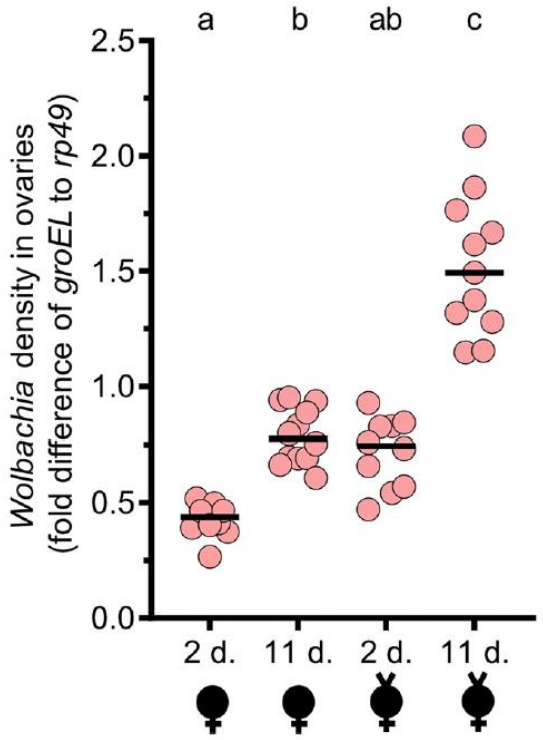

Figure H-3. *Wolbachia* densities increase with female age in ovaries and decrease after mating.
Wolbachia density assays were conducted with pools of 4 ovaries from virgin females (indicated by a "v" above a sex symbol) and nonvirgin females aged 2 or 11 days (d.). *Wolbachia* infections are represented by filled sex symbols, and the age of the sample is shown immediately below the *x* axis. Virgin and nonvirgin females were siblings. The nonvirgin females produced the embryos whose results are shown in Figure H-2b. Nonvirgin females were allowed to mate and lay for 48 h before ovary dissections. Nonvirgin and virgin females were incubated for that same period of time and dissected in parallel. Each dot represents the average of duplicate values. Horizontal bars indicate medians, and the letters above the bars indicate significant differences based on alpha = 0.05 calculated by a Kruskal-Wallis test followed by a Dunn's multiple-comparison test performed between all groups. All statistical values are presented in Table S1. Fold differences in *Wolbachia* density (*groEL*) relative to *D. melanogaster* reference gene *rp49* were determined with $2^{-\Delta\Delta Ct}$.

Finally, this led to the hypothesis that older *D. melanogaster* grandmothers have higher *Wolbachia* densities than younger grandmothers and that they transfer more *Wolbachia* to their offspring. Supporting this hypothesis, *Wolbachia* densities were significantly higher in the ovaries of 11-day-old virgin females than in those of 2-day old virgin females ($P = 0.0045$) (Figure H-3). Additionally, we predicted that *Wolbachia* densities would decrease in ovaries after egg-laying if grandmothers loaded *Wolbachia* into their offspring. As such, we measured *Wolbachia* densities in ovaries of mated grandmothers that laid eggs in the embryo density study described previously.

We found that ovaries from mated 11-day-old females had significantly less *Wolbachia* than virgin 11-day-old females (P = 0.0240) (Figure H-3). Likewise, mated 2-day-old females had lower *Wolbachia* titers than virgin 2-day-old females, though the differences were not significant (P = 0.0882) (Figure H-3). Despite the overall decrease in relative *Wolbachia* densities after mating, ovaries from 11-day-old mated grandmothers had significantly higher densities than ovaries from 2-day-old mated grandmothers (P = 0.0087). Importantly, threshold cycle (*CT*) values remained consistent across age and virginity states for the *Drosophila rp49* gene, suggesting that changes in the *Wolbachia groEL* gene, rather than in *rp49* copy number, were responsible for the density dynamics that we report here (data not shown).

Similar results can be observed in measuring *Wolbachia* densities in abdomens instead of in ovaries (data not shown). Measuring abdominal titers, 11-day-old virgin females had statistically higher *Wolbachia* densities than 2-day-old virgin females (P = 0.0001). There was a detectable trend indicating that the mated females had less *Wolbachia*, though neither mated 11-day-old females (P = 0.2291) nor mated 2-day-old females (P = 0.9999) had titers significantly different from those of their virgin counterparts. The titers in 11-day-old and 2-day-old mated females were not significantly different (P = 0.9999). Taken together, these data suggest that females accumulate *Wolbachia* as they age, that older females transfer more *Wolbachia* to their offspring, and that sons of older females induce stronger CI. Moreover, laying eggs appears to quickly reduce the amount of *Wolbachia* contained in the ovaries, suggesting that the PGAE is strongest soon after initial mating.

Discussion

D. melanogaster is a valued model system in studies of *Wolbachia*-host interactions due to its genetic tractability and the importance of its native *Wolbachia* strain, *w*Mel, in vector control (Flores and O'Neill, 2018). However, the study of *w*Mel-induced CI in *D. melanogaster* is inhibited by its variable penetrance, ranging from nearly complete embryonic death to none at all (Hoffmann et al., 1998, 1994; Holden et al., 1993; LePage et al., 2017; Reynolds and Hoffmann, 2002; Solignac et al., 1994; Yamada et al., 2007). Some phenotypic variation persists despite control of known variables of CI strength, leading to the hypothesis that as-yet-unknown factors contribute to CI variability. Anecdotal observations in our laboratory suggested that stronger CI

may be induced by offspring of older virgin females, leading to the formal hypothesis that variation in CI penetrance is partly controlled by a paternal grandmother age effect (PGAE).

Here, we report evidence in support of the PGAE, namely, that sons of older *D. melanogaster* grandmothers induce stronger CI than sons of younger grandmothers. Paternal grandmother age did not influence the ability of CI to be rescued, suggesting that no other age-associated transgenerational deficiencies contributed to the increased embryonic death. Additionally, we found that embryos of older grandmothers had higher *Wolbachia* densities than the offspring of younger grandmothers. Likewise, older virgin females had more *Wolbachia* than younger virgin females. As such, the data support a model whereby PGAE is caused by an accumulation of *Wolbachia* in a virgin as she ages, leading to an increase in levels of *Wolbachia* passed on to her sons, who induce stronger CI in their offspring than sons of younger grandmothers.

In this study, we measured *Wolbachia* densities by comparing the number of *Wolbachia groEL* gene copies to the number of *Drosophila rp49* gene copies. Note that we cannot make direct claims about the density of *Wolbachia* per host cell based on these analyses, since doing so would assume that the number of host cells and host ploidy remain constant. Recent work has highlighted that a protein-enriched diet can influence relative estimates of *Wolbachia* density analysis in *D. melanogaster* by increasing ovary size and *rp49* copy number (Christensen et al., 2019). While age and mating state may be hypothesized to influence *rp49* copy number, *rp49 CT* values remained constant across female age and mating states whereas *groEL CT* values changed (data not shown). These data suggest that despite possible fluctuations in *rp49* copy number across cell types within ovaries, the average *rp49* copy number remains consistent across the extracted tissue samples. As such, we conclude that changes in *Wolbachia groEL* copy number, not *rp49* copy number, underpin the results. However, future work will be necessary to describe how these density estimates explicitly relate to *Wolbachia* titers per host cell and across cell types in these tissues.

In addition to the PGAE, CI variation has previously been attributed to a "younger-brother effect" where the slowest-developing males, from a clutch of embryos within the 0-to-5-h age range, induced the weakest CI (Yamada et al., 2007). If embryo deposition order correlates with maturation rate, then the younger-brother effect is at least in part explained by our findings that (i) *Wolbachia* densities in ovaries quickly decrease after mating and egg laying, (ii) the *Wolbachia*

density in embryos correlates with ovary densities, and (iii) sons from eggs laid by mothers with lower *Wolbachia* densities induce weaker CI. As such, when a *D. melanogaster* female lays eggs, the amount of *Wolbachia* in her ovaries may be sequentially depleted after each embryo is produced.

Thus, younger brothers that take longer to develop may receive fewer *Wolbachia* and then induce a weaker CI than their older counterparts that originally had received more *Wolbachia*. Therefore, the dynamics of the interaction that we observed between CI induction and *Wolbachia* densities across generations may explain the younger-brother effect, although this remains to be precisely established in future research.

Additionally, this paper adds to a growing body of literature reporting an influence of female insect age on *Wolbachia* densities. Indeed, older females harbor higher *Wolbachia* titers in *w*AlbA- and *w*AlbB-infected *Aedes albopictus* (Calvitti et al., 2015; Tortosa et al., 2010), *w*VulC-infected *Armadillidium vulgare* (Genty et al., 2014), and *w*Stri-infected *Laodelphax striatellus* (Guo et al., 2018). The relationship between paternal grandmother age and the strength of *w*Mel-induced CI was explored once before; however, no relationship was found (Yamada et al., 2007). Crucially, the virginity status of the grandmothers differs between the cited study and the one presented here and may in part explain the discrepancy. Our study maintained the virginity of all grandmothers as they aged, and grandmothers were allowed only 24 h of mating prior to egg deposition for hatch rate analysis. In contrast, the grandmothers in the prior study remained virgin until 3 days old and were then allowed to continuously mate until they were 11 days old, and the CI levels from sons produced at each of the two time points were compared (Yamada et al., 2007). Our results suggest that mating has a detectable impact on *Wolbachia* densities and may explain why the PGAE was not observed in the earlier study. Additionally, we predict that the PGAE most strongly applies to aged virgins, since mating significantly reduced *Wolbachia* densities in our study.

The depletion of *Wolbachia* found in females following egg laying supports the hypothesis that the PGAE is caused by an effect of maternal loading of *Wolbachia* into her sons. However, the source of that loading is still unclear. In *D. melanogaster*, the following four sources of *Wolbachia* transfer to progeny are known: bacteriocyte-like cells (BLCs), germ line stem cells (GSCs), the somatic stem cell niche (SSCN), and late-stage oogenesis (Fast et al., 2011; Frydman et al., 2006; Sacchi et al., 2010; Serbus et al., 2008; Toomey et al., 2013; Veneti et al., 2004). BLCs

found at the tip of the ovarioles are densely packed with *Wolbachia* and are predicted to transfer *Wolbachia* to GSCs (Sacchi et al., 2010). When a GSC asymmetrically undergoes mitosis (Deng and Lin, 1997; Lin and Schagat, 1997), its population of *Wolbachia* is divided between two daughter cells, one of which is an identical GSC that remains in the ovaries and the other a differentiating cytoblast that develops into the egg (Serbus et al., 2008). Therefore, it is possible that the levels of *Wolbachia* allocated to the daughter cytoblast (and thus the offspring) are proportional to the densities in the parent GSC or the surrounding BLCs. Additionally, as the cytoblast develops into a germ line cyst, it comes into contact with the highly infected SSCN, acquiring additional *Wolbachia* (Fast et al., 2011; Frydman et al., 2006; Toomey et al., 2013). Finally, while *Wolbachia* replication in the oocyte occurs primarily at the beginning of oogenesis in *w*Mel-infected *D. melanogaster* and halts at the onset of vitellogenesis, it can resume at a lower rate before egg laying in late-stage oogenesis (Veneti et al., 2004). As such, prolonged retention of eggs in aged virgins may lead to an accumulation of *Wolbachia* in these developed oocytes. We hypothesize that *Wolbachia* replicate in the BLCs, GSCs, SSCNs, or late-stage oocytes as a mother ages, resulting in eggs with relatively high titers. Since eggs account for the greatest proportion of *Wolbachia* cells in the ovaries, this hypothesis could explain why titers are depleted after mating and egg laying.

Intriguingly, differences in CI strength more closely correlated with *Wolbachia* densities in embryos than with densities in adult males. CI is hypothesized to be caused by *cif* gene modifications of sperm-associated host products (Bossan et al., 2011; Ferree and Sullivan, 2006; Landmann et al., 2009; Poinsot et al., 2003; Presgraves, 2000; Shropshire et al., 2019; Shropshire and Bordenstein, 2019; Tram and Sullivan, 2002) or to be a consequence of loading of toxins into the sperm (Beckmann et al., 2019a, 2017; Shropshire et al., 2019; Shropshire and Bordenstein, 2019); however, *Wolbachia* are stripped from the sperm during individualization (Bressac and Rousset, 1993; Clark et al., 2002; Snook et al., 2000). Therefore, *Wolbachia* titers are likely more important during a specific stage of spermatogenesis than at the time of CI induction. In *D. melanogaster*, spermatogenesis is a continuous process lasting approximately 11 days (Lindsley, 1980). As such, there may be a lag of several days between the time that sperm are subjected to the actions of *cifA* and *cifB* gene products and the time of CI induction. Spermatogenesis begins during larval development (Lindsley, 1980) and continues throughout the adult life span, though the first batches of mature sperm are produced soon after adult hatching (Ruhmann et al., 2016).

Since the males in our study were mated shortly after adult hatching, the majority of their sperm would have started spermatogenesis at a time closer to embryonic deposition than adult hatching, which may explain why CI strength correlates better with *Wolbachia* densities in embryos than in adult males. Additionally, spermatogenesis may incorporate and eliminate *Wolbachia* faster than they can multiply, resulting in the reduction and equalization of titers in adults (Clark et al., 2002). This may explain why some studies, including studies analyzing the younger-brother effect, found that CI strength did not always correlate with *Wolbachia* densities in adults (Bourtzis et al., 1996; Karr et al., 1998; Yamada et al., 2007). As such, we predict that the PGAE is the result of the presence of high *Wolbachia* densities during a critical time point in spermatogenesis when CI-defining changes occur, which may become the subject of future research.

It remains unclear if the association between female age and *Wolbachia* densities would be the case in wild populations. Since wild *D. melanogaster* females are estimated to mate, on average, every 27 h (Giardina et al., 2017), it would seem unlikely that the *Wolbachia* accumulation reported here would occur in nature. However, infection status has been reported to influence mate choice behaviors in numerous animals, including *D. subquinaria, D. paulistorum, Nasonia vitripennis*, and *Tetranychus urticae* (Chafee et al., 2011; Jaenike et al., 2006; Miller et al., 2010; Vala et al., 2004). For example, male mating rate affects CI strength (Reynolds and Hoffmann, 2002), so *w*Mel-infected males mate more frequently to reduce the impact of CI strength and therefore improve their lifetime reproductive success (De Crespigny et al., 2006). Additionally, females infected with *Wolbachia* have a higher reproductive fitness when their daughters can sufficiently rescue CI and when their sons induce weak CI. Thus, it is plausible that the latency to copulation could be either lengthened in instances where a higher *Wolbachia* titer would be preferable (rescue efficiency) or, conversely, shortened in populations where a lower density is preferred (weakened CI). While it is unlikely that a fly in nature will remain virgin for as long as reported in this study, it is notable that CI strength increased substantially with every time point measured. As such, even small changes in mating latency may influence CI strength sufficiently to change the rate of spread through a population. Field studies measuring the latency toward copulation in sites with different infection rates would help determine if insects can modulate their mating latency, and thus *Wolbachia* titers,
to increase their fitness and the fitness of their offspring.

While this work reports a PGAE for *w*Mel in *D. melanogaster*, it is unknown if these dynamics occur for *w*Mel in mosquito hosts. In *w*Mel-infected *Aedes aegypti* mosquitoes, CI is consistently strong (Blagrove et al., 2012; Dutra et al., 2015). However, some factors such as *Wolbachia* densities and temperature were shown previously to correlate with CI penetrance (Ross et al., 2019). It is possible that other as-yet-unstudied factors in mosquitoes, such as the PGAE, can contribute to changes in CI strength. Since strong CI is crucial for rapid spread of *w*Mel-infected mosquitoes through populations for successful vector control applications (Ritchie et al., 2018), understanding the factors that contribute to variation in CI strength would further inform the efficacy of population replacement and rearing strategies. Moreover, comparative studies exploring *w*Mel-induced CI in *D. melanogaster* and *A. aegypti* could clarify the *Wolbachia*-host dynamics that govern the penetrance of CI.

Finally, there is a striking range of CI penetrance across *Wolbachia* and hosts, and more work is necessary to determine if the PGAE applies to other CI or reproductive parasite systems. For example, *w*Ri in *D. simulans* consistently induces strong CI (Hoffmann et al., 1990, 1986; Turelli et al., 2018b) and *w*Yak and *w*Tei in the *D. yakuba* clade cause weak and variable levels of CI similar to those seen with *w*Mel (Cooper et al., 2017). Intriguingly, *w*Mel and *w*Tei were initially thought not to cause CI until factors such as male age and host genotype were found to have a significant impact on CI strength (Charlat et al., 2004; Cooper et al., 2017; Martinez et al., 2014; Poinsot et al., 1998; Reynolds and Hoffmann, 2002; Zabalou et al., 2004). Since it is clear that some *Wolbachia* cause CI only under strictly limited conditions, it remains possible that other weak-CI inducing *Wolbachia* are mislabeled as non-CI strains because factors such as the PGAE had not been controlled for during initial testing. Indeed, while this work presents the first reported case of transgenerational *Wolbachia* titers influencing CI, it is not the first case of transgenerational *Wolbachia* titers influencing reproductive parasitism. In *D. innubila*, male-killing *Wolbachia* frequently kill all male offspring, but females with lower *Wolbachia* titers are known to produce some viable sons (Dyer et al., 2005). The surviving female offspring inherit lower-than-average *Wolbachia* titers, leading to a greater-than-average chance that those infected females would also produce sons (Dyer et al., 2005). Together, our results and those in *D. innubila* suggest that a transgenerational effect of titers may be common and consequential with respect to the expression of reproductive parasitism traits.

In conclusion, we characterize *Wolbachia* density dynamics in females in relation to age and mating, and we link a transgenerational influence of grandmother age to CI penetrance. This work highlights the importance of controlling grandparent age in future studies of *w*Mel-induced CI in *D. melanogaster* and has implications for laboratory experiments where precise control over levels of CI would be valuable for dissecting the genetic and functional basis of CI. Additionally, it expands our understanding of *Wolbachia*-host interactions in relation to CI penetrance and titer dynamics and should motivate additional studies exploring these interactions in *w*Mel-infected mosquitoes.

Materials and methods

Fly strains and maintenance.

The following *D. melanogaster* strains were used in this study: *wMel*-infected and uninfected variants of y^1w* (BDSC 1495) and *nos*-GAL4-*tubulin* (BDSC 4442). Uninfected lines were generated through three generations of tetracycline treatment as previously described (LePage et al., 2017). All stocks were reared on 50 ml of a standard medium containing cornmeal, molasses, and yeast and were maintained at 25°C with a 12-h/12-h light:dark cycle and at 70% relative humidity (RH). All virgin flies were collected using CO2 anesthetization per standard procedures. Briefly, virgin flies were collected in the morning based on the presence of a meconium, bottles were subsequently cleared of adult flies, and flies collected in the evening were assumed virgin due to the standard time of latency until mating. All virgin flies were kept at room temperature prior to experimentation.

Hatch rate assays.

Hatch rate assays were used to assess the impact of *D. melanogaster* paternal grandmother age on the strength of CI induced by their sons. We conducted 3 variant hatch rate assays to test (i) whether paternal grandmother age influences CI hatch rates, (ii) whether this effect is specific to the y^1w* genetic background, and (iii) whether the transgenerational impact of age on hatching is indeed caused by CI.

First, we assessed if *D. melanogaster* paternal grandmother age influences CI hatch rates in the y^1w* genetic background. Paternal y^1w* grandmothers were collected as virgins and allowed to reach 2, 5, 11, 14, or 18 days of age before mating in parallel with paternal grandfathers aged 0

to 2 days. Paternal grandparents from each age cohort were crossed in single-pair matings in standard vials of media. Since rearing density influences CI strength (Yamada et al., 2007), paternal grandparents were allowed 24 h to mate and to deposit eggs before the grandfathers were discarded and the grandmothers were flash frozen and stored at –80°C for *Wolbachia* titer analysis. To control for the younger-brother effect and the effect of male age on the strength of CI (Reynolds and Hoffmann, 2002; Yamada et al., 2007), the earliest eclosing fathers were collected as virgins and left to age 1 day at room temperature before being used in hatch rate assays.

Maternal y^1w* grandparents were crossed in standard medium bottles and allowed to mate for 4 days before flies were cleared, as described above for the paternal grandparents. Mothers were collected as virgins and allowed to reach 6 to 8 days of age at room temperature to maximize fertility (Miller et al., 2014).

Parental y^1w* mating pairs were placed in 8-oz *Drosophila* stock bottles (Genesee Scientific) with a grape juice agar plate covered in yeast affixed to the top to collect embryos for hatch rate analysis as previously described (LePage et al., 2017; Shropshire et al., 2018). Parents were allowed two back-to-back 24-h mating and laying periods, each with separate freshly yeasted grape juice agar plates. The plates from the first mating period were discarded due to the typically low levels of egg laying in the first 24 h. The embryos from the second mating period were immediately counted after 24 h of additional laying. Embryos were then incubated for 30 h at 25°C to allow time to hatch. The unhatched embryos were counted, and the percentage of embryonic hatching was determined by dividing the number of unhatched embryos by the total number of embryos laid during the second mating period.

To minimize the effect of female fecundity on embryo viability (Miller et al., 2014), any plate with fewer than 25 embryos was excluded. We measured the hatch rates of offspring produced by two sons of each paternal grandmother. If both sons from the same family produced 25 or more embryos, one was randomly selected and used in analysis.

Next, to assess if the PGAE was specific to the y^1w* genetic background, a separate hatch rate assay was conducted using *nos*-GAL4-*tubulin*-infected and uninfected flies. This experiment was conducted similarly to the hatch rate experiment described above, with the following adjustments: age and virginity of paternal grandfathers were not controlled. Paternal *nos*-GAL4-*tubulin*-infected and uninfected grandmothers were collected as virgins and allowed to reach 2, 5,

or 11 days of age before they were allowed to mate in standard medium bottles, and these bottles were cleared of flies after 4 days of laying to control rearing density (Yamada et al., 2007).

Finally, to determine if the PGAE was in fact due to *Wolbachia* and not to other forms of inviability induced by a transgenerational impact on age, we conducted compatible rescue crosses with males derived from 2-, 5-, 11-, or 14-day-old females. This experiment was conducted similarly to the hatch rate experiment described above, with the following adjustments: both infected and uninfected males were produced from virgin females aged 2, 4, 11, or 14 days; the uninfected males were mated to uninfected females; and infected males were mated to infected females. Paternal grandparents were paired in 8-oz *Drosophila* stock bottles (Genesee Scientific) with a grape juice agar plate (LePage et al., 2017) covered in yeast affixed to the top for a 24-h mating and laying period, and then grandparents were collected from the bottles. The plates were maintained for 24 h, and then 20 of the largest larvae were transferred from each plate to a standard medium vial to control rearing density (Yamada et al., 2007).

Wolbachia *titer assays.*

To assess the relationship between the PGAE and *Wolbachia* titers, the following tissues were collected: ovaries, female abdomens, embryos, and male abdomens. Since the low biomass of *Drosophila* testes requires them to be pooled, abdomens were used instead of testes so that samples could be taken directly from the males used in hatch rate assays. To test if virginity and age impact female *Wolbachia* titers, virgin and nonvirgin females 2 and 11 days of age were reared in parallel, ovaries were dissected in phosphate-buffered saline (PBS), and samples were frozen in liquid nitrogen followed by storage at –80°C. Samples consisted of 4 pairs of ovaries. Nonvirgin females were mated in cohorts of 60 females to 12 males, provided grape juice plates, and allowed 48 h to mate and lay eggs before dissection. Additionally, full bodies from 2-or-11-day-old paternal grandmothers from a hatch rate assay were collected alongside virgin paternal grandaunts (siblings to the paternal grandmothers), frozen in liquid nitrogen, and stored at –80°C. To determine if embryos derived from older females had higher *Wolbachia* titers, 0-to-1-h-old embryos were collected from grape plates in batches of 50, frozen in liquid nitrogen, and stored at –80°C. Finally, to assess whether the sons of aged paternal grandmothers had higher *Wolbachia* titers, full bodies from virgin uncles (siblings of fathers used in a hatch rate assay) derived from 2-, 5-, 11-, 14-, or 18-day-old grandmothers were collected and aged 48 h at room temperature in

a standard medium vial. *Wolbachia* titers were measured in virgin uncles rather than the fathers used in the hatch rate assay because of the relationship between CI strength and male mating rate (De Crespigny et al., 2006).

Upon removal from –80°C conditions, abdomens were immediately dissected from full-body tissues, homogenized in liquid nitrogen, and mixed with 40 ul ice-cold RNase-free PBS. Each sample was split, and 30% (12 ul) was flash frozen and stored at –80°C for DNA extractions. The DNA was extracted from all tissue types using a Gentra PureGene tissue kit (Qiagen). Forty cycles of quantitative PCR (qPCR) were performed using *rp49* and *groEL* primers (Table S2) for all DNA samples as well as positive controls (infected DNA), negative controls (uninfected DNA), no-reverse-transcription controls (RNA), and notissue controls (water). Male and female abdomen samples were tested in triplicate and ovaries and embryos in duplicate under the following qPCR conditions: 50°C for 10 min; 95°C for 5 min; 40 cycles of 95°C for 10 s and 55°C for 30 s; and 95°C for 30 s. Samples were excluded from analysis if the standard deviation of results of comparisons between replicates was >0.3. Fold difference between *Wolbachia* (*groEL*) density and that of the *D. melanogaster rp49* reference gene was determined with $2^{-\Delta\Delta CT}$.

Statistical analyses.

All statistical analyses were conducted using GraphPad Prism 7. *Wolbachia* titers of embryos were analyzed using a Mann-Whitney *U* test. All other data (including data from hatch rate assays and ovary *Wolbachia* titer comparisons) were analyzed using the Kruskal-Wallis test followed by a Dunn's multiple-comparison test. Figures were created in GraphPad Prism 7 and 8. All data used in these analyses have been made publicly available.

Appendix I.

List of publications

Layton, E.M., On, J., Perlmutter, J., & Bordenstein, S.R. & **Shropshire, J.D.** (2019). Paternal grandmother age affects the strength of *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila melanogaster*. *mBio*. 10(6).

Shropshire, J. D., Leigh, B., Bordenstein, S. R., Duplouy, A., Riegler, M., Brownlie, J. C., & Bordenstein, S. R. (2019). Models and nomenclature for cytoplasmic incompatibility: Caution over premature conclusions. *Trends in genetics*. 35(6), 397-399.

Shropshire, J. D., & Bordenstein, S. R. (2019). Two-By-One model of cytoplasmic incompatibility: Synthetic recapitulation by transgenic expression of *cifA* and *cifB* in *Drosophila*. *PLOS Genetics.* 16(6), e1008221.

Shropshire, J. D., On, J., Layton, E. M., Zhou, H., & Bordenstein, S. R. (2018). One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, *115*(19), 4987-4991.

Shropshire, J. D. & A. Rokas (2017). Heredity: The gene family that cheats Mendel. *eLife* 6: e28567.

LePage, D. P., Metcalf, J. A., Bordenstein, S. R., On, J., Perlmutter, J. I., **Shropshire, J. D.**, Layton, E.M., Funkhouser-Jones, L. J., Beckmann, J.F., & Bordenstein, S. R. (2017). Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature*, *543*(7644), 243-247.

Dittmer, J., van Opstal, E. J., **Shropshire, J. D.**, Bordenstein, S. R., Hurst, G. D., & Brucker, R. M. (2016). Disentangling a holobiont–recent advances and perspectives in *Nasonia* wasps. *Frontiers in Microbiology*, *7*.

Shropshire, J. D., van Opstal, E. J., & Bordenstein, S. R. (2016). An optimized approach to germ-free rearing in the jewel wasp *Nasonia*. *PeerJ*, *4*, e2088v1.

Shropshire, J. D., & Bordenstein, S. R. (2016). Speciation by symbiosis: The microbiome and behavior. *mBio*, *7*(2), e01785-15.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinforma. Oxf. Engl. 21, 2104–2105. https://doi.org/10.1093/bioinformatics/bti263

Adachi-Hagimori, T., Miura, K., Stouthamer, R., 2008. A new cytogenetic mechanism for bacterial endosymbiont-induced parthenogenesis in Hymenoptera. Proc. Biol. Sci. 275, 2667–2673. https://doi.org/10.1098/rspb.2008.0792

Adekunle, A.I., Meehan, M.T., McBryde, E.S., 2019. Mathematical analysis of a *Wolbachia* invasive model with imperfect maternal transmission and loss of *Wolbachia* infection. Infect. Dis. Model. 4, 265–285. https://doi.org/10.1016/j.idm.2019.10.001

Akbari, O.S., Chen, C.-H., Marshall, J.M., Huang, H., Antoshechkin, I., Hay, B.A., 2014. Novel synthetic Medea selfish genetic elements drive population replacement in *Drosophila*; a theoretical exploration of Medea-dependent population suppression. ACS Synth. Biol. 3, 915–928. https://doi.org/10.1021/sb300079h

Akbari, O.S., Matzen, K.D., Marshall, J.M., Huang, H., Ward, C.M., Hay, B.A., 2013. A synthetic gene drive system for local, reversible modification and suppression of insect populations. Curr. Biol. CB 23, 671–677. https://doi.org/10.1016/j.cub.2013.02.059

Akhtar, W., de Jong, J., Pindyurin, A.V., Pagie, L., Meuleman, W., de Ridder, J., Berns, A., Wessels, L.F.A., van Lohuizen, M., van Steensel, B., 2013. Chromatin position effects assayed by thousands of reporters integrated in parallel. Cell 154, 914–927. https://doi.org/10.1016/j.cell.2013.07.018

Alexandrov, I.D., Alexandrova, M.V., Goryacheva, I.I., Rochina, N.V., Shaikevich, E.V., Zakharov, I.A., 2007. Removing endosymbiotic *Wolbachia* specifically decreases lifespan of females and competitiveness in a laboratory strain of *Drosophila melanogaster*. Russ. J. Genet. 43, 1147–1152. https://doi.org/10.1134/S1022795407100080

Aliota, M.T., Peinado, S.A., Velez, I.D., Osorio, J.E., 2016. The *w*Mel strain of *Wolbachia* reduces transmission of Zika virus by *Aedes aegypti*. Sci. Rep. 6, 28792. https://doi.org/10.1038/srep28792

Alivisatos, A.P., Blaser, M.J., Brodie, E.L., Chun, M., Dangl, J.L., Donohue, T.J., Dorrestein, P.C., Gilbert, J.A., Green, J.L., Jansson, J.K., Knight, R., Maxon, M.E., McFall-Ngai, M.J., Miller, J.F., Pollard, K.S., Ruby, E.G., Taha, S.A., 2015. A unified initiative to harness Earth's microbiomes. Science 350, 507–508. https://doi.org/10.1126/science.aac8480

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

Amato, K.R., Leigh, S.R., Kent, A., Mackie, R.I., Yeoman, C.J., Stumpf, R.M., Wilson, B.A., Nelson, K.E., White, B.A., Garber, P.A., 2015. The gut microbiota appears to compensate for seasonal diet variation in the wild black howler monkey (*Alouatta pigra*). Microb. Ecol. 69, 434–443. https://doi.org/10.1007/s00248-014-0554-7

Anderson, K.E., Russell, J.A., Moreau, C.S., Kautz, S., Sullam, K.E., Hu, Y., Basinger, U., Mott, B.M., Buck, N., Wheeler, D.E., 2012. Highly similar microbial communities are shared among related and trophically similar ant species. Mol. Ecol. 21, 2282–2296. https://doi.org/10.1111/j.1365-294X.2011.05464.x

Ant, T.H., Herd, C., Louis, F., Failloux, A.B., Sinkins, S.P., 2020. *Wolbachia* transinfections in *Culex quinquefasciatus* generate cytoplasmic incompatibility. Insect Mol. Biol. 29, 1–8. https://doi.org/10.1111/imb.12604

Ant, T.H., Herd, C.S., Geoghegan, V., Hoffmann, A.A., Sinkins, S.P., 2018. The *Wolbachia* strain *w*Au provides highly efficient virus transmission blocking in *Aedes aegypti*. PLOS Pathog. 14, e1006815. https://doi.org/10.1371/journal.ppat.1006815

Ant, T.H., Sinkins, S.P., 2018. A *Wolbachia* triple-strain infection generates self-incompatibility in *Aedes albopictus* and transmission instability in *Aedes aegypti*. Parasit. Vectors 11, 295. https://doi.org/10.1186/s13071-018-2870-0

Arai, H., Hirano, T., Akizuki, N., Abe, A., Nakai, M., Kunimi, Y., Inoue, M.N., 2018. Multiple infection and reproductive manipulations of *Wolbachia* in *Homona magnanima* (Lepidoptera: Tortricidae). Microb. Ecol. https://doi.org/10.1007/s00248-018-1210-4

Arakaki, N., Miyoshi, T., Noda, H., 2001. *Wolbachia*-mediated parthenogenesis in the predatory thrips *Franklinothrips vespiformis* (Thysanoptera: Insecta). Proc. Biol. Sci. 268, 1011–1016. https://doi.org/10.1098/rspb.2001.1628

Arbuthnott, D., Levin, T.C., Promislow, D.E.L., 2016. The impacts of *Wolbachia* and the microbiome on mate choice in *Drosophila melanogaster*. J. Evol. Biol. 29, 461–468. https://doi.org/10.1111/jeb.12788

Archie, E.A., Theis, K.R., 2011. Animal behaviour meets microbial ecology. Anim. Behav. 82, 425–436. https://doi.org/10.1016/j.anbehav.2011.05.029

Argov, Y., Gottlieb, Y., Amin-Spector, S., Zchori-Fein, E., 2000. Possible symbiont-induced thelytoky in *Galeopsomyia fausta*, a parasitoid of the citrus leafminer *Phyllocnistis citrella*. Phytoparasitica 28, 212–218. https://doi.org/10.1007/BF02981799

Asselin, A.K., Villegas-Ospina, S., Hoffmann, A.A., Brownlie, J.C., Johnson, K.N., 2018. Contrasting patterns of virus protection and functional incompatibility genes in two conspecific *Wolbachia* strains from *Drosophila pandora*. Appl Env. Microbiol AEM.02290-18. https://doi.org/10.1128/AEM.02290-18

Atyame, C.M., Delsuc, F., Pasteur, N., Weill, M., Duron, O., 2011a. Diversification of *Wolbachia* endosymbiont in the *Culex pipiens* mosquito. Mol. Biol. Evol. 28, 2761–2772. https://doi.org/10.1093/molbev/msr083

Atyame, C.M., Labbé, P., Dumas, E., Milesi, P., Charlat, S., Fort, P., Weill, M., 2014. *Wolbachia* divergence and the evolution of cytoplasmic incompatibility in *Culex pipiens*. PLOS ONE 9, e87336. https://doi.org/10.1371/journal.pone.0087336

Atyame, C.M., Pasteur, N., Dumas, E., Tortosa, P., Tantely, M.L., Pocquet, N., Licciardi, S., Bheecarry, A., Zumbo, B., Weill, M., Duron, O., 2011b. Cytoplasmic incompatibility as a means of controlling *Culex pipiens quinquefasciatus* mosquito in the islands of the south-western Indian Ocean. PLoS Negl. Trop. Dis. 5, e1440. https://doi.org/10.1371/journal.pntd.0001440

Awrahman, Z.A., Champion de Crespigny, F., Wedell, N., 2014. The impact of *Wolbachia*, male age and mating history on cytoplasmic incompatibility and sperm transfer in *Drosophila simulans*. J. Evol. Biol. 27, 1–10. https://doi.org/10.1111/jeb.12270

Baião, G.C., Schneider, D.I., Miller, W.J., Klasson, L., 2019. The effect of *Wolbachia* on gene expression in *Drosophila paulistorum* and its implications for symbiont-induced host speciation. BMC Genomics 20, 465. https://doi.org/10.1186/s12864-019-5816-9

Bakovic, V., Schebeck, M., Telschow, A., Stauffer, C., Schuler, H., 2018. Spatial spread of *Wolbachia* in *Rhagoletis cerasi* populations. Biol. Lett. 14. https://doi.org/10.1098/rsbl.2018.0161

Baldridge, G., Higgins, L., Witthuhn, B., Markowski, T., Baldridge, A., Armien, A., Fallon, A., 2017. Proteomic analysis of a mosquito host cell response to persistent *Wolbachia* infection. Res. Microbiol. 168, 609–625. https://doi.org/10.1016/j.resmic.2017.04.005

Baldridge, G.D., Baldridge, A.S., Witthuhn, B.A., Higgins, L., Markowski, T.W., Fallon, A.M., 2014. Proteomic profiling of a robust *Wolbachia* infection in an *Aedes albopictus* mosquito cell line. Mol. Microbiol. 94, 537–556. https://doi.org/10.1111/mmi.12768

Balenger, S.L., Zuk, M., 2014. Testing the Hamilton-Zuk hypothesis: past, present, and future. Integr. Comp. Biol. 54, 601–613. https://doi.org/10.1093/icb/icu059

Baton, L.A., Pacidônio, E.C., Gonçalves, D. da S., Moreira, L.A., 2013. *w*Flu: characterization and evaluation of a native *Wolbachia* from the mosquito *Aedes fluviatilis* as a potential vector control agent. PloS One 8, e59619. https://doi.org/10.1371/journal.pone.0059619

Beckmann, J., Bonneau, M., Chen, H., Hochstrasser, M., Poinsot, D., Merçot, H., Weill, M., Sicard, M., Charlat, S., 2019a. The Toxin–Antidote model of cytoplasmic incompatibility: genetics and evolutionary implications. Trends Genet. https://doi.org/10.1016/j.tig.2018.12.004

Beckmann, J., Bonneau, M., Chen, H., Hochstrasser, M., Poinsot, D., Merçot, H., Weill, M., Sicard, M., Charlat, S., 2019b. Caution does not preclude predictive and testable models of cytoplasmic incompatibility: a reply to Shropshire et al. Trends Genet. 0. https://doi.org/10.1016/j.tig.2019.03.002

Beckmann, J., Fallon, A.M., 2013. Detection of the *Wolbachia* protein WPIP0282 in mosquito spermathecae: implications for cytoplasmic incompatibility. Insect Biochem. Mol. Biol. 43, 867–878. https://doi.org/10.1016/j.ibmb.2013.07.002

Beckmann, J., Markowski, T.W., Witthuhn, B.A., Fallon, A.M., 2013. Detection of the *Wolbachia*-encoded DNA binding protein, HU beta, in mosquito gonads. Insect Biochem. Mol. Biol. 43, 272–279. https://doi.org/10.1016/j.ibmb.2012.12.007

Beckmann, J., Ronau, J.A., Hochstrasser, M., 2017. A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. Nat. Microbiol. 2, 17007. https://doi.org/10.1038/nmicrobiol.2017.7

Beckmann, J., Sharma, G.D., Mendez, L., Chen, H., Hochstrasser, M., 2019c. The *Wolbachia* cytoplasmic incompatibility enzyme CidB targets nuclear import and protamine-histone exchange factors. eLife 8, e50026. https://doi.org/10.7554/eLife.50026

Beltran-Bech, S., Richard, F.-J., 2014. Impact of infection on mate choice. Anim. Behav. 90, 159–170. https://doi.org/10.1016/j.anbehav.2014.01.026

Ben-David, E., Burga, A., Kruglyak, L., 2017. A maternal-effect selfish genetic element in *Caenorhabditis elegans*. Science 356, 1051–1055. https://doi.org/10.1126/science.aan0621

Benson, A.K., Kelly, S.A., Legge, R., Ma, F., Low, S.J., Kim, J., Zhang, M., Oh, P.L., Nehrenberg, D., Hua, K., Kachman, S.D., Moriyama, E.N., Walter, J., Peterson, D.A., Pomp, D., 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. https://doi.org/10.1073/pnas.1007028107

Ben-Yosef, M., Aharon, Y., Jurkevitch, E., Yuval, B., 2010. Give us the tools and we will do the job: symbiotic bacteria affect olive fly fitness in a diet-dependent fashion. Proc. Biol. Sci. 277, 1545–1552. https://doi.org/10.1098/rspb.2009.2102

Ben-Yosef, M., Jurkevitch, E., Yuval, B., 2008. Effect of bacteria on nutritional status and reproductive success of the Mediterranean fruit fly *Ceratitis capitata*. Physiol. Entomol. 33, 145–154. https://doi.org/10.1111/j.1365-3032.2008.00617.x

Berg, G., Rybakova, D., Grube, M., Koberl, M., 2016. The plant microbiome explored: implications for experimental botany. J. Exp. Bot. 67, 995–1002. https://doi.org/10.1093/jxb/erv466

Betelman, K., Caspi-Fluger, A., Shamir, M., Chiel, E., 2017. Identification and characterization of bacterial symbionts in three species of filth fly parasitoids. FEMS Microbiol. Ecol. 93. https://doi.org/10.1093/femsec/fix107

Bhattacharya, T., Newton, I.L.G., Hardy, R.W., 2017. *Wolbachia* elevates host methyltransferase expression to block an RNA virus early during infection. PLOS Pathog. 13, e1006427. https://doi.org/10.1371/journal.ppat.1006427

Bian, G., Joshi, D., Dong, Y., Lu, P., Zhou, G., Pan, X., Xu, Y., Dimopoulos, G., Xi, Z., 2013. *Wolbachia* invades *Anopheles stephensi* populations and induces refractoriness to Plasmodium infection. Science 340, 748–751. https://doi.org/10.1126/science.1236192

Bing, X.-L., Lu, Y.-J., Xia, C.-B., Xia, X., Hong, X.-Y., 2019. Transcriptome of *Tetranychus urticae* embryos reveals insights into *Wolbachia*-induced cytoplasmic incompatibility. Insect Mol. Biol. https://doi.org/10.1111/imb.12620

Bing, X.-L., Zhao, D.-S., Sun, J.-T., Zhang, K.-J., Hong, X.-Y., 2020. Genomic analysis of *Wolbachia* from *Laodelphax striatellus* (Delphacidae, Hemiptera) reveals insights into its "Jekyll and Hyde" mode of infection pattern. Genome Biol. Evol. https://doi.org/10.1093/gbe/evaa006

Binnington, K.C., Hoffmann, A.A., 1989. *Wolbachia*-like organisms and cytoplasmic incompatibility in *Drosophila simulans*. J. Invertebr. Pathol. 54, 344–352. https://doi.org/10.1016/0022-2011(89)90118-3

Biwot, J.C., Zhang, H.-B., Liu, C., Qiao, J.-X., Yu, X.-Q., Wang, Y.-F., 2019. *Wolbachia*-induced expression of kenny gene in testes affects male fertility in *Drosophila melanogaster*. Insect Sci. https://doi.org/10.1111/1744-7917.12730

Blagrove, M.S.C., Arias-Goeta, C., Failloux, A.-B., Sinkins, S.P., 2012. *Wolbachia* strain *w*Mel induces cytoplasmic incompatibility and blocks dengue transmission in *Aedes albopictus*. Proc. Natl. Acad. Sci. 109, 255–260. https://doi.org/10.1073/pnas.1112021108

Blekhman, R., Goodrich, J.K., Huang, K., Sun, Q., Bukowski, R., Bell, J.T., Spector, T.D., Keinan, A., Ley, R.E., Gevers, D., Clark, A.G., 2015. Host genetic variation impacts microbiome composition across human body sites. Genome Biol. 16, 191. https://doi.org/10.1186/s13059-015-0759-1

Bockoven, A.A., Bondy, E.C., Flores, M.J., Kelly, S.E., Ravenscraft, A.M., Hunter, M.S., 2019. What goes up might come down: the spectacular spread of an endosymbiont is followed by its decline a decade later. Microb. Ecol. https://doi.org/10.1007/s00248-019-01417-4

Bonneau, M., Atyame, C., Beji, M., Justy, F., Cohen-Gonsaud, M., Sicard, M., Weill, M., 2018a. *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. Nat. Commun. 9. https://doi.org/10.1038/s41467-017-02749-w

Bonneau, M., Caputo, B., Ligier, A., Caparros, R., Unal, S., Perriat-Sanguinet, M., Arnoldi, D., Sicard, M., Weill, M., 2019. Variation in *Wolbachia cidB* gene, but not *cidA*, is

associated with cytoplasmic incompatibility mod phenotype diversity in *Culex pipiens*. Mol. Ecol. 28, 4725–4736. https://doi.org/10.1111/mec.15252
Bonneau, M., Landmann, F., Labbé, P., Justy, F., Weill, M., Sicard, M., 2018b. The cellular phenotype of cytoplasmic incompatibility in *Culex pipiens* in the light of *cidB* diversity. PLOS Pathog. 14, e1007364. https://doi.org/10.1371/journal.ppat.1007364
Bordenstein, Sarah R., Bordenstein, Seth R., 2016. Eukaryotic association module in phage WO genomes from *Wolbachia*. Nat. Commun. 7, 13155. https://doi.org/10.1038/ncomms13155
Bordenstein, Sarah R., Bordenstein, Seth R., 2011. Temperature affects the tripartite interactions between bacteriophage WO, *Wolbachia*, and cytoplasmic incompatibility. PloS One 6, e29106. https://doi.org/10.1371/journal.pone.0029106
Bordenstein, S.R., 2003. Symbiosis and the origin of species. Insect Symbiosis 1, 283–304.
Bordenstein, S.R., O'Hara, F.P., Werren, J.H., 2001. *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. Nature 409, 707–710. https://doi.org/10.1038/35055543
Bordenstein, S.R., Theis, K.R., 2015. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. PLoS Biol. 13, e1002226. https://doi.org/10.1371/journal.pbio.1002226
Bordenstein, S.R., Uy, J.J., Werren, J.H., 2003. Host genotype determines cytoplasmic incompatibility type in the haplodiploid genus *Nasonia*. Genetics 164, 223–233.
Bordenstein, S.R., Werren, J.H., 2007. Bidirectional incompatibility among divergent *Wolbachia* and incompatibility level differences among closely related *Wolbachia* in *Nasonia*. Heredity 99, 278–287. https://doi.org/10.1038/sj.hdy.6800994
Bossan, B., Koehncke, A., Hammerstein, P., 2011. A new model and method for understanding *Wolbachia*-induced cytoplasmic incompatibility. PLOS ONE 6, e19757. https://doi.org/10.1371/journal.pone.0019757
Bourtzis, K., Dobson, S.L., Braig, H.R., O'Neill, S.L., 1998. Rescuing *Wolbachia* have been overlooked. Nature 391, 852–853. https://doi.org/10.1038/36017
Bourtzis, K., Nirgianaki, A., Markakis, G., Savakis, C., 1996. *Wolbachia* infection and cytoplasmic incompatibility in *Drosophila* species. Genetics 144, 1063–1073.
Boyle, L., O'neill, S., Robertson, H., Karr, T., 1993. Interspecific and intraspecific horizontal transfer of *Wolbachia* in *Drosophila*. Science 260, 1796–1799. https://doi.org/10.1126/science.8511587
Breeuwer, J., Werren, J., 1990. Microorganisms associated with chromosome destruction and reproductive isolation between 2 insect species. Nature 346, 558–560. https://doi.org/10.1038/346558a0
Brennan, L.J., Haukedal, J.A., Earle, J.C., Keddie, B., Harris, H.L., 2012. Disruption of redox homeostasis leads to oxidative DNA damage in spermatocytes of *Wolbachia*-infected *Drosophila simulans*. Insect Mol. Biol. 21, 510–520. https://doi.org/10.1111/j.1365-2583.2012.01155.x
Brennan, L.J., Keddie, B.A., Braig, H.R., Harris, H.L., 2008. The endosymbiont *Wolbachia pipientis* induces the expression of host antioxidant proteins in an *Aedes albopictus* cell line. PloS One 3, e2083. https://doi.org/10.1371/journal.pone.0002083
Bressac, C., Rousset, F., 1993. The reproductive incompatibility system in *Drosophila simulans*: dapi-staining analysis of the *Wolbachia* symbionts in sperm cysts. J. Invertebr. Pathol. 61, 226–230. https://doi.org/10.1006/jipa.1993.1044

Brucker, R.M., Bordenstein, S.R., 2013a. The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus *Nasonia*. Science 341, 667–669. https://doi.org/10.1126/science.1240659
Brucker, R.M., Bordenstein, S.R., 2013b. The capacious hologenome. Zool. Jena Ger. 116, 260–261. https://doi.org/10.1016/j.zool.2013.08.003
Brucker, R.M., Bordenstein, S.R., 2012a. Speciation by symbiosis. Trends Ecol. Evol. 27, 443–451. https://doi.org/10.1016/j.tree.2012.03.011
Brucker, R.M., Bordenstein, S.R., 2012b. In vitro cultivation of the hymenoptera genetic model, *Nasonia*. PloS One 7, e51269. https://doi.org/10.1371/journal.pone.0051269
Cai, Z., Yan, L.-J., 2013. Protein Oxidative Modifications: Beneficial Roles in Disease and Health. J. Biochem. Pharmacol. Res. 1, 15–26.
Callaini, G., Dallai, R., Riparbelli, M.G., 1997. *Wolbachia*-induced delay of paternal chromatin condensation does not prevent maternal chromosomes from entering anaphase in incompatible crosses of *Drosophila simulans*. J. Cell Sci. 110, 271–280.
Callaini, G., Riparbelli, M.G., Giordano, R., Dallai, R., 1996. Mitotic defects associated with cytoplasmic incompatibility in *Drosophila simulans*. J. Invertebr. Pathol. 67, 55–64. https://doi.org/10.1006/jipa.1996.0009
Callewaert, C., De Maeseneire, E., Kerckhof, F.-M., Verliefde, A., Van de Wiele, T., Boon, N., 2014. Microbial odor profile of polyester and cotton clothes after a fitness session. Appl. Environ. Microbiol. 80, 6611–6619. https://doi.org/10.1128/AEM.01422-14
Calvitti, M., Marini, F., Desiderio, A., Puggioli, A., Moretti, R., 2015. *Wolbachia* density and cytoplasmic incompatibility in *Aedes albopictus*: concerns with using artificial *Wolbachia* infection as a vector suppression tool. PloS One 10, e0121813. https://doi.org/10.1371/journal.pone.0121813
Cameron, D.E., Urbach, J.M., Mekalanos, J.J., 2008. A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae*. Proc. Natl. Acad. Sci. U. S. A. 105, 8736–8741. https://doi.org/10.1073/pnas.0803281105
Campbell, B.C., Steffen-Campbell, J.D., Werren, J.H., 1994. 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. https://doi.org/10.1111/j.1365-2583.1994.tb00142.x
Caputo, B., Moretti, R., Manica, M., Serini, P., Lampazzi, E., Bonanni, M., Fabbri, G., Pichler, V., Della Torre, A., Calvitti, M., 2019. A bacterium against the tiger: preliminary evidence of fertility reduction after release of *Aedes albopictus* males with manipulated *Wolbachia* infection in an Italian urban area. Pest Manag. Sci. https://doi.org/10.1002/ps.5643
Caragata, E.P., Dutra, H.L.C., Moreira, L.A., 2016. Inhibition of Zika virus by *Wolbachia* in *Aedes aegypti*. Microb. Cell 3, 293–295. https://doi.org/10.15698/mic2016.07.513
Caragata, E.P., Rancès, E., Hedges, L.M., Gofton, A.W., Johnson, K.N., O'Neill, S.L., McGraw, E.A., 2013. Dietary cholesterol modulates pathogen blocking by *Wolbachia*. PLoS Pathog. 9, e1003459. https://doi.org/10.1371/journal.ppat.1003459
Carlson, A.D., Copeland, J., Raderman, R., Bulloch, A.G.M., 1976. Role of interflash intervals in a firefly courtship (*Photinus macdermotti*). Anim. Behav. 24, 786–792. https://doi.org/10.1016/S0003-3472(76)80009-7

Carrington, L.B., Lipkowitz, J.R., Hoffmann, A.A., Turelli, M., 2011. A re-examination of *Wolbachia*-induced cytoplasmic incompatibility in California *Drosophila simulans*. PloS One 6, e22565. https://doi.org/10.1371/journal.pone.0022565

Casiraghi, M., Bordenstein, S.R., Baldo, L., Lo, N., Beninati, T., Wernegreen, J.J., Werren, J.H., 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. Microbiol.-Sgm 151, 4015–4022. https://doi.org/10.1099/mic.0.28313-0

Cator, L.J., Arthur, B.J., Harrington, L.C., Hoy, R.R., 2009. Harmonic convergence in the love songs of the dengue vector mosquito. Science 323, 1077–1079. https://doi.org/10.1126/science.1166541

Cattel, J., Nikolouli, K., Andrieux, T., Martinez, J., Jiggins, F., Charlat, S., Vavre, F., Lejon, D., Gibert, P., Mouton, L., 2018. Back and forth *Wolbachia* transfers reveal efficient strains to control spotted wing *Drosophila* populations. J. Appl. Ecol. 55, 2408–2418. https://doi.org/10.1111/1365-2664.13101

Chafee, M.E., Funk, D.J., Harrison, R.G., Bordenstein, S.R., 2010. Lateral phage transfer in obligate intracellular bacteria (*Wolbachia*): verification from natural populations. Mol. Biol. Evol. 27, 501–505. https://doi.org/10.1093/molbev/msp275

Chafee, M.E., Zecher, C.N., Gourley, M.L., Schmidt, V.T., Chen, J.H., Bordenstein, Sarah R., Clark, M.E., Bordenstein, Seth R., 2011. Decoupling of host-symbiont-phage coadaptations following transfer between insect species. Genetics 187, 203–215. https://doi.org/10.1534/genetics.110.120675

Chambers, E.W., Hapairai, L., Peel, B.A., Bossin, H., Dobson, S.L., 2011. Male mating competitiveness of a *Wolbachia*-introgressed *Aedes polynesiensis* strain under semi-field conditions. PLoS Negl. Trop. Dis. 5, e1271. https://doi.org/10.1371/journal.pntd.0001271

Champer, J., Buchman, A., Akbari, O.S., 2016. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. Nat. Rev. Genet. 17, 146–159. https://doi.org/10.1038/nrg.2015.34

Champion de Crespigny, F.E., Wedell, N., 2006. *Wolbachia* infection reduces sperm competitive ability in an insect. Proc. Biol. Sci. 273, 1455–1458. https://doi.org/10.1098/rspb.2006.3478

Charlat, S., Ballard, J.W.O., Mercot, H., 2004. What maintains noncytoplasmic incompatibility inducing *Wolbachia* in their hosts: a case study from a natural *Drosophila yakuba* population. J. Evol. Biol. 17, 322–330.

Charlat, S., Calmet, C., Merçot, H., 2001. On the mod resc model and the evolution of *Wolbachia* compatibility types. Genetics 159, 1415–1422.

Charlat, S., Reuter, M., Dyson, E.A., Hornett, E.A., Duplouy, A., Davies, N., Roderick, G.K., Wedell, N., Hurst, G.D.D., 2007. Male-killing bacteria trigger a cycle of increasing male fatigue and female promiscuity. Curr. Biol. CB 17, 273–277. https://doi.org/10.1016/j.cub.2006.11.068

Charlesworth, J., Weinert, L.A., Araujo, E.V., Welch, J.J., 2019. *Wolbachia*, *Cardinium* and climate: an analysis of global data. Biol. Lett. 15, 20190273. https://doi.org/10.1098/rsbl.2019.0273

Chatzispyrou, I.A., Held, N.M., Mouchiroud, L., Auwerx, J., Houtkooper, R.H., 2015. Tetracycline antibiotics impair mitochondrial function and its experimental use

confounds research. Cancer Res. 75, 4446–4449. https://doi.org/10.1158/0008-5472.CAN-15-1626
Chen, C.-H., Huang, H., Ward, C.M., Su, J.T., Schaeffer, L.V., Guo, M., Hay, B.A., 2007. A synthetic maternal-effect selfish genetic element drives population replacement in *Drosophila*. Science 316, 597–600. https://doi.org/10.1126/science. 1138595
Chen, D., Zheng, W., Lin, A., Uyhazi, K., Zhao, H., Lin, H., 2012. Pumilio 1 Suppresses Multiple Activators of p53 to Safeguard Spermatogenesis. Curr. Biol. 22, 420–425. https://doi.org/10.1016/j.cub.2012.01.039
Chen, H., Ronau, J.A., Beckmann, J., Hochstrasser, M., 2019. A *Wolbachia* nuclease and its binding partner provide a distinct mechanism for cytoplasmic incompatibility. Proc. Natl. Acad. Sci. 116, 22314–22321. https://doi.org/10.1073/pnas.1914571116
Chevalier, F., Herbinière-Gaboreau, J., Charif, D., Mitta, G., Gavory, F., Wincker, P., Grève, P., Braquart-Varnier, C., Bouchon, D., 2012. Feminizing *Wolbachia*: a transcriptomics approach with insights on the immune response genes in *Armadillidium vulgare*. BMC Microbiol. 12 Suppl 1, S1. https://doi.org/10.1186/1471-2180-12-S1-S1
Cho, K.-O., Kim, G.-W., Lee, O.-K., 2011. *Wolbachia* bacteria reside in host Golgi-related vesicles whose position is regulated by polarity proteins. PloS One 6, e22703. https://doi.org/10.1371/journal.pone.0022703
Christensen, S., Camacho, M., Sharmin, Z., Momtaz, A.J.M.Z., Perez, L., Navarro, G., Triana, J., Samarah, H., Turelli, M., Serbus, L.R., 2019. Quantitative methods for assessing local and bodywide contributions to *Wolbachia* titer in maternal germline cells of *Drosophila*. BMC Microbiol. 19, 206. https://doi.org/10.1186/s12866-019-1579-3
Chung, H., Pamp, S.J., Hill, J.A., Surana, N.K., Edelman, S.M., Troy, E.B., Reading, N.C., Villablanca, E.J., Wang, S., Mora, J.R., Umesaki, Y., Mathis, D., Benoist, C., Relman, D.A., Kasper, D.L., 2012. Gut immune maturation depends on colonization with a host-specific microbiota. Cell 149, 1578–1593. https://doi.org/10.1016/j.cell.2012.04.037
Claesson, M.J., Jeffery, I.B., Conde, S., Power, S.E., O'Connor, E.M., Cusack, S., Harris, H.M.B., Coakley, M., Lakshminarayanan, B., O'Sullivan, O., Fitzgerald, G.F., Deane, J., O'Connor, M., Harnedy, N., O'Connor, K., O'Mahony, D., van Sinderen, D., Wallace, M., Brennan, L., Stanton, C., Marchesi, J.R., Fitzgerald, A.P., Shanahan, F., Hill, C., Ross, R.P., O'Toole, P.W., 2012. Gut microbiota composition correlates with diet and health in the elderly. Nature 488, 178–184. https://doi.org/10.1038/nature11319
Clancy, D.J., Hoffmann, A.A., 1998. Environmental effects on cytoplasmic incompatibility and bacterial load in *Wolbachia*-infected *Drosophila simulans*. Entomol. Exp. Appl. 86, 13–24. https://doi.org/10.1046/j.1570-7458.1998.00261.x
Clark, M.E., Heath, B.D., Anderson, C.L., Karr, T.L., 2006. Induced paternal effects mimic cytoplasmic incompatibility in *Drosophila*. Genetics 173, 727–734. https://doi.org/10.1534/genetics.105.052431
Clark, M.E., O'Hara, F.P., Chawla, A., Werren, J.H., 2010. Behavioral and spermatogenic hybrid male breakdown in *Nasonia*. Heredity 104, 289–301. https://doi.org/10.1038/hdy.2009.152
Clark, M.E., Veneti, Z., Bourtzis, K., Karr, T.L., 2003. *Wolbachia* distribution and cytoplasmic incompatibility during sperm development: the cyst as the basic cellular unit of CI expression. Mech. Dev. 120, 185–198. https://doi.org/10.1016/S0925-4773(02)00424-0

Clark, M.E., Veneti, Z., Bourtzis, K., Karr, T.L., 2002. The distribution and proliferation of the intracellular bacteria *Wolbachia* during spermatogenesis in *Drosophila*. Mech. Dev. 111, 3–15. https://doi.org/10.1016/S0925-4773(01)00594-9

Conner, W.R., Blaxter, M.L., Anfora, G., Ometto, L., Rota-Stabelli, O., Turelli, M., 2017. Genome comparisons indicate recent transfer of *w*Ri-like *Wolbachia* between sister species *Drosophila suzukii* and *D. subpulchrella*. Ecol. Evol. 7, 9391–9404. https://doi.org/10.1002/ece3.3449

Cooper, B.S., Ginsberg, P.S., Turelli, M., Matute, D.R., 2017. *Wolbachia* in the *Drosophila yakuba* complex: pervasive frequency variation and weak cytoplasmic incompatibility, but no apparent effect on reproductive isolation. Genetics 205, 333–351. https://doi.org/10.1534/genetics.116.196238

Cooper, B.S., Vanderpool, D., Conner, W.R., Matute, D.R., Turelli, M., 2019. *Wolbachia* acquisition by *Drosophila yakuba*-clade hosts and transfer of incompatibility loci between distantly related *Wolbachia*. Genetics. https://doi.org/10.1534/genetics.119.302349

Cordaux, R., Pichon, S., Hatira, H.B.A., Doublet, V., Greve, P., Marcade, I., Braquart-Varnier, C., Souty-Grosset, C., Charfi-Cheikhrouha, F., Bouchon, D., 2012. Widespread *Wolbachia* infection in terrestrial isopods and other crustaceans. ZooKeys 123–131. https://doi.org/10.3897/zookeys.176.2284

Coyne, J., 2001. Reproductive isolation, in: encyclopedia of genetics. Elsevier, pp. 1679–1686. https://doi.org/10.1006/rwgn.2001.1442

Cunningham, B.C., Wells, J.A., 1989. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science 244, 1081–1085. https://doi.org/10.1126/science.2471267

Curry, M.M., Paliulis, L.V., Welch, K.D., Harwood, J.D., White, J.A., 2015. Multiple endosymbiont infections and reproductive manipulations in a linyphiid spider population. Heredity 115, 146–152. https://doi.org/10.1038/hdy.2015.2

Darwin, C., 1869. On the origin of species by means of natural selection: or the preservation of favoured races in the struggle for life. D. Appleton.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., Biddinger, S.B., Dutton, R.J., Turnbaugh, P.J., 2014. Diet rapidly and reproducibly alters the human gut microbiome. Nature 505, 559–563. https://doi.org/10.1038/nature12820

De Cock, R., Matthysen, E., 2005. Sexual communication by pheromones in a firefly, *Phosphaenus hemipterus* (Coleoptera: Lampyridae). Anim. Behav. 70, 807–818. https://doi.org/10.1016/j.anbehav.2005.01.011

De Crespigny, F.E.C., Pitt, T.D., Wedell, N., 2006. Increased male mating rate in *Drosophila* is associated with *Wolbachia* infection. J. Evol. Biol. 19, 1964–1972. https://doi.org/10.1111/j.1420-9101.2006.01143.x

Debug Fresno, 2019. Singapore collaboration achieves greater than 90% reduction in release areas. Debug Proj. URL https://blog.debug.com/2019/11/singapore-collaboration-achieves.html (accessed 1.21.20).

Debug Fresno, 2018a. Debug Fresno 2018 results in 95% suppression! Debug Proj. URL https://blog.debug.com/2018/11/debug-fresno-2018-results-in-95.html (accessed 1.21.20).

Debug Fresno, 2018b. Debug Innisfail achieves strong suppression. Debug Proj. URL https://blog.debug.com/2018/07/debug-innisfail-achieves-strong.html (accessed 1.21.20).
Degnan, P.H., Pusey, A.E., Lonsdorf, E.V., Goodall, J., Wroblewski, E.E., Wilson, M.L., Rudicell, R.S., Hahn, B.H., Ochman, H., 2012. Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park. Proc. Natl. Acad. Sci. U. S. A. 109, 13034–13039. https://doi.org/10.1073/pnas.1110994109
Demerec, M., Adelberg, E.A., Clark, A.J., Hartman, P.E., 1966. A Proposal for a Uniform Nomenclature in Bacterial Genetics. Genetics 54, 61–76.
Deng, W., Lin, H., 1997. Spectrosomes and fusomes anchor mitotic spindles during asymmetric germ cell divisions and facilitate the formation of a polarized microtubule array for oocyte specification in *Drosophila*. Dev. Biol. 189, 79–94. https://doi.org/10.1006/dbio.1997.8669
Dittmer, J., van Opstal, E.J., Shropshire, J.D., Bordenstein, S.R., Hurst, G.D.D., Brucker, R.M., 2016. Disentangling a holobiont - recent advances and perspectives in *Nasonia* wasps. Front. Microbiol. 7, 1478. https://doi.org/10.3389/fmicb.2016.01478
Dobson, S.L., Bordenstein, S.R., Rose, R.I., 2016. *Wolbachia* mosquito control: Regulated. Science 352, 526–527. https://doi.org/10.1126/science.352.6285.526-b
Dobson, S.L., Fox Charles W., Jiggins Francis M., 2002. The effect of *Wolbachia*-induced cytoplasmic incompatibility on host population size in natural and manipulated systems. Proc. R. Soc. Lond. B Biol. Sci. 269, 437–445. https://doi.org/10.1098/rspb.2001.1876
Dobzhansky, T., 1937. Genetics and the origin of species. Columbia University Press.
Dodd, D.M.B., 1989. Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. Evol. Int. J. Org. Evol. 43, 1308–1311. https://doi.org/10.1111/j.1558-5646.1989.tb02577.x
Dodson, B.L., Hughes, G.L., Paul, O., Matacchiero, A.C., Kramer, L.D., Rasgon, J.L., 2014. *Wolbachia* enhances West Nile Virus (WNV) infection in the mosquito *Culex tarsalis*. PLoS Negl. Trop. Dis. 8, e2965. https://doi.org/10.1371/journal.pntd.0002965
Donnelly, M.L., Hughes, L.E., Luke, G., Mendoza, H., ten Dam, E., Gani, D., Ryan, M.D., 2001a. The "cleavage" activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring "2A-like" sequences. J. Gen. Virol. 82, 1027–1041. https://doi.org/10.1099/0022-1317-82-5-1027
Donnelly, M.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., Gani, D., Ryan, M.D., 2001b. Analysis of the aphthovirus 2A/2B polyprotein "cleavage" mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal "skip." J. Gen. Virol. 82, 1013–1025. https://doi.org/10.1099/0022-1317-82-5-1013
Doren, M.V., Williamson, A.L., Lehmann, R., 1998. Regulation of zygotic gene expression in *Drosophila* primordial germ cells. Curr. Biol. 8, 243–246. https://doi.org/10.1016/S0960-9822(98)70091-0
Driver, C., Georgiou, A., Georgiou, G., 2004. The contribution by mitochondrially induced oxidative damage to aging in *Drosophila melanogaster*. Biogerontology 5, 185–192. https://doi.org/10.1023/B:BGEN.0000031156.75376.e3
Duffy, J.B., 2002. GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. Genes. N. Y. N 2000 34, 1–15. https://doi.org/10.1002/gene.10150
Dunn, A.M., Hatcher, M.J., Terry, R.S., Tofts, C., 1995. Evolutionary ecology of vertically transmitted parasites: transovarial transmission of a microsporidian sex ratio distorter in

Gammarus duebeni. Parasitology 111, S91–S109. https://doi.org/10.1017/S0031182000075843
Duplouy, A., Hurst, G.D.D., O'neill, S.L., Charlat, S., 2010. Rapid spread of male-killing *Wolbachia* in the butterfly *Hypolimnas bolina*. J. Evol. Biol. 23, 231–235. https://doi.org/10.1111/j.1420-9101.2009.01891.x
Duron, Olivier, Boureux, A., Echaubard, P., Berthomieu, A., Berticat, C., Fort, P., Weill, M., 2007. Variability and expression of ankyrin domain genes in *Wolbachia* variants infecting the mosquito *Culex pipiens*. J. Bacteriol. 189, 4442–4448. https://doi.org/10.1128/JB.00142-07
Duron, O., Fort, P., Weill, M., 2007. Influence of aging on cytoplasmic incompatibility, sperm modification and *Wolbachia* density in *Culex pipiens* mosquitoes. Heredity 98, 368–374. https://doi.org/10.1038/sj.hdy.6800948
Duron, O., Weill, M., 2006. *Wolbachia* infection influences the development of *Culex pipiens* embryo in incompatible crosses. Heredity 96, 493–500. https://doi.org/10.1038/sj.hdy.6800831
Dutra, H.L.C., Dos Santos, L.M.B., Caragata, E.P., Silva, J.B.L., Villela, D.A.M., Maciel-de-Freitas, R., Moreira, L.A., 2015. From lab to field: the influence of urban landscapes on the invasive potential of *Wolbachia* in Brazilian *Aedes aegypti* mosquitoes. PLoS Negl. Trop. Dis. 9, e0003689. https://doi.org/10.1371/journal.pntd.0003689
Dutra, H.L.C., Rocha, M.N., Dias, F.B.S., Mansur, S.B., Caragata, E.P., Moreira, L.A., 2016. *Wolbachia* blocks currently circulating Zika virus isolates in Brazilian *Aedes aegypti* Mosquitoes. Cell Host Microbe 19, 771–774. https://doi.org/10.1016/j.chom.2016.04.021
Dyer, K.A., Minhas, M.S., Jaenike, J., 2005. Expression and modulation of embryonic male-killing in *Drosophila innubila*: opportunities for multilevel selection. Evol. Int. J. Org. Evol. 59, 838–848.
Dyson, E.A., Hurst, G.D.D., 2004. Persistence of an extreme sex-ratio bias in a natural population. Proc. Natl. Acad. Sci. U. S. A. 101, 6520–6523. https://doi.org/10.1073/pnas.0304068101
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340
Elnagdy, S., Majerus, M.E.N., Handley, L.-J.L., 2011. The value of an egg: resource reallocation in ladybirds (Coleoptera: Coccinellidae) infected with male-killing bacteria. J. Evol. Biol. 24, 2164–2172. https://doi.org/10.1111/j.1420-9101.2011.02346.x
Emelyanov, A.V., Rabbani, J., Mehta, M., Vershilova, E., Keogh, M.C., Fyodorov, D.V., 2014. *Drosophila* TAP/p32 is a core histone chaperone that cooperates with NAP-1, NLP, and nucleophosmin in sperm chromatin remodeling during fertilization. Genes Dev. 28, 2027–2040. https://doi.org/10.1101/gad.248583.114
Engelstadter, J., Telschow, A., 2009. Cytoplasmic incompatibility and host population structure. Heredity 103, 196–207. https://doi.org/10.1038/hdy.2009.53
Ezenwa, V.O., Gerardo, N.M., Inouye, D.W., Medina, M., Xavier, J.B., 2012. Microbiology. Animal behavior and the microbiome. Science 338, 198–199. https://doi.org/10.1126/science.1227412
Ezenwa, V.O., Williams, A.E., 2014. Microbes and animal olfactory communication: Where do we go from here? BioEssays News Rev. Mol. Cell. Dev. Biol. 36, 847–854. https://doi.org/10.1002/bies.201400016

Farkas, J.Z., Hinow, P., 2010. Structured and unstructured continuous models for *Wolbachia* infections. Bull. Math. Biol. 72, 2067–2088. https://doi.org/10.1007/s11538-010-9528-1
Fast, E.M., Toomey, M.E., Panaram, K., Desjardins, D., Kolaczyk, E.D., Frydman, H.M., 2011. *Wolbachia* enhance *Drosophila* stem cell proliferation and target the germline stem cell niche. Science 334, 990–992. https://doi.org/10.1126/science.1209609
Fattouh, N., Cazevieille, C., Landmann, F., 2019. *Wolbachia* endosymbionts subvert the endoplasmic reticulum to acquire host membranes without triggering ER stress. PLoS Negl. Trop. Dis. 13, e0007218. https://doi.org/10.1371/journal.pntd.0007218
Ferguson, S.B., Blundon, M.A., Klovstad, M.S., Schüpbach, T., 2012. Modulation of gurken translation by insulin and TOR signaling in *Drosophila*. J. Cell Sci. 125, 1407–1419. https://doi.org/10.1242/jcs.090381
Ferree, P.M., Sullivan, W., 2006. A genetic test of the role of the maternal pronucleus in *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila melanogaster*. Genetics 173, 839–847. https://doi.org/10.1534/genetics.105.053272
Ferri, E., Bain, O., Barbuto, M., Martin, C., Lo, N., Uni, S., Landmann, F., Baccei, S.G., Guerrero, R., Lima, S. de S., Bandi, C., Wanji, S., Diagne, M., Casiraghi, M., 2011. New insights into the evolution of *Wolbachia* infections in filarial nematodes inferred from a large range of screened species. PLOS ONE 6, e20843. https://doi.org/10.1371/journal.pone.0020843
Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279-285. https://doi.org/10.1093/nar/gkv1344
Flores, H.A., O'Neill, S.L., 2018. Controlling vector-borne diseases by releasing modified mosquitoes. Nat. Rev. Microbiol. 1. https://doi.org/10.1038/s41579-018-0025-0
Foo, I.J.-H., Hoffmann, A.A., Ross, P.A., 2019. Cross-generational effects of heat stress on fitness and *Wolbachia* density in *Aedes aegypti* mosquitoes. Trop. Med. Infect. Dis. 4. https://doi.org/10.3390/tropicalmed4010013
Forbes, A., Lehmann, R., 1998. Nanos and Pumilio have critical roles in the development and function of *Drosophila* germline stem cells. Dev. Camb. Engl. 125, 679–690.
Foster, J., Ganatra, M., Kamal, I., Ware, J., Makarova, K., Ivanova, N., Bhattacharyya, A., Kapatral, V., Kumar, S., Posfai, J., Vincze, T., Ingram, J., Moran, L., Lapidus, A., Omelchenko, M., Kyrpides, N., Ghedin, E., Wang, S., Goltsman, E., Joukov, V., Ostrovskaya, O., Tsukerman, K., Mazur, M., Comb, D., Koonin, E., Slatko, B., 2005. The *Wolbachia* genome of *Brugia malayi*: Endosymbiont evolution within a human pathogenic nematode. Plos Biol. 3, 599–614. https://doi.org/10.1371/journal.pbio.0030121
Franzenburg, S., Walter, J., Kunzel, S., Wang, J., Baines, J.F., Bosch, T.C.G., Fraune, S., 2013. Distinct antimicrobial peptide expression determines host species-specific bacterial associations. Proc. Natl. Acad. Sci. U. S. A. 110, E3730-3738. https://doi.org/10.1073/pnas.1304960110
Fry, A.J., Palmer, M.R., Rand, D.M., 2004. Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. Heredity 93, 379–389. https://doi.org/10.1038/sj.hdy.6800514
Frydman, H.M., Li, J.M., Robson, D.N., Wieschaus, E., 2006. Somatic stem cell niche tropism in *Wolbachia*. Nature 441, 509–512. https://doi.org/10.1038/nature04756

Funkhouser, L.J., Bordenstein, S.R., 2013. Mom knows best: the universality of maternal microbial transmission. PLoS Biol. 11, e1001631. https://doi.org/10.1371/journal.pbio.1001631
Gavriel, S., Jurkevitch, E., Gazit, Y., Yuval, B., 2011. Bacterially enriched diet improves sexual performance of sterile male Mediterranean fruit flies. J. Appl. Entomol. 135, 564–573. https://doi.org/10.1111/j.1439-0418.2010.01605.x
Genty, L.-M., Bouchon, D., Raimond, M., Bertaux, J., 2014. *Wolbachia* infect ovaries in the course of their maturation: last minute passengers and priority travellers? PloS One 9, e94577. https://doi.org/10.1371/journal.pone.0094577
Geoghegan, V., Stainton, K., Rainey, S.M., Ant, T.H., Dowle, A.A., Larson, T., Hester, S., Charles, P.D., Thomas, B., Sinkins, S.P., 2017. Perturbed cholesterol and vesicular trafficking associated with dengue blocking in *Wolbachia*-infected *Aedes aegypti* cells. Nat. Commun. 8, 526. https://doi.org/10.1038/s41467-017-00610-8
Gerth, M., Rothe, J., Bleidorn, C., 2013. Tracing horizontal *Wolbachia* movements among bees (Anthophila): a combined approach using multilocus sequence typing data and host phylogeny. Mol. Ecol. 22, 6149–6162. https://doi.org/10.1111/mec.12549
Ghate, N.B., Kim, J., Shin, Y., Situ, A., Ulmer, T.S., An, W., 2019. p32 is a negative regulator of p53 tetramerization and transactivation. Mol. Oncol. 13, 1976–1992. https://doi.org/10.1002/1878-0261.12543
Giardina, T.J., Clark, A.G., Fiumera, A.C., 2017. Estimating mating rates in wild *Drosophila melanogaster* females by decay rates of male reproductive proteins in their reproductive tracts. Mol. Ecol. Resour. 17, 1202–1209. https://doi.org/10.1111/1755-0998.12661
Gilbert, S.F., 2014. A holobiont birth narrative: the epigenetic transmission of the human microbiome. Front. Genet. 5, 282. https://doi.org/10.3389/fgene.2014.00282
Gilbert, S.F., Bosch, T.C.G., Ledon-Rettig, C., 2015. Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. Nat. Rev. Genet. 16, 611–622. https://doi.org/10.1038/nrg3982
Gilbert, S.F., Sapp, J., Tauber, A.I., 2012. A symbiotic view of life: we have never been individuals. Q. Rev. Biol. 87, 325–341. https://doi.org/10.1086/668166
Gilfillan, G.D., Dahlsveen, I.K., Becker, P.B., 2004. Lifting a chromosome: dosage compensation in *Drosophila melanogaster*. FEBS Lett. 567, 8–14. https://doi.org/10.1016/j.febslet.2004.03.110
Giordano, R., Jackson, J.J., Robertson, H.M., 1997. The role of *Wolbachia* bacteria in reproductive incompatibilities and hybrid zones of *Diabrotica* beetles and *Gryllus* crickets. Proc. Natl. Acad. Sci. U. S. A. 94, 11439–11444. https://doi.org/10.1073/pnas.94.21.11439
Goodrich, J.K., Waters, J.L., Poole, A.C., Sutter, J.L., Koren, O., Blekhman, R., Beaumont, M., Van Treuren, W., Knight, R., Bell, J.T., Spector, T.D., Clark, A.G., Ley, R.E., 2014. Human genetics shape the gut microbiome. Cell 159, 789–799. https://doi.org/10.1016/j.cell.2014.09.053
Gotoh, T., Noda, H., Hong, X.-Y., 2003. *Wolbachia* distribution and cytoplasmic incompatibility based on a survey of 42 spider mite species (Acari: Tetranychidae) in Japan. Heredity 91, 208–216. https://doi.org/10.1038/sj.hdy.6800329
Gotoh, T., Sugasawa, J., Noda, H., Kitashima, Y., 2007. *Wolbachia*-induced cytoplasmic incompatibility in Japanese populations of *Tetranychus urticae* (Acari: Tetranychidae). Exp. Appl. Acarol. 42, 1–16. https://doi.org/10.1007/s10493-007-9072-3

Gottlieb, Y., Zchori-Fein, E., 2001. Irreversible thelytokous reproduction in *Muscidifurax uniraptor*. Entomol. Exp. Appl. 100, 271–278. https://doi.org/10.1046/j.1570-7458.2001.00874.x

Govind, S., 2008. Innate immunity in *Drosophila*: Pathogens and pathways. Insect Sci. Online 15, 29–43. https://doi.org/10.1111/j.1744-7917.2008.00185.x

Grognet, P., Lalucque, H., Malagnac, F., Silar, P., 2014. Genes that bias Mendelian segregation. PLoS Genet. 10, e1004387. https://doi.org/10.1371/journal.pgen.1004387

Groth, A.C., Fish, M., Nusse, R., Calos, M.P., 2004. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. Genetics 166, 1775–1782. https://doi.org/10.1534/genetics.166.4.1775

Güell, M., Yus, E., Lluch-Senar, M., Serrano, L., 2011. Bacterial transcriptomics: what is beyond the RNA horiz-ome? Nat. Rev. Microbiol. 9, 658–669. https://doi.org/10.1038/nrmicro2620

Guo, Y., Hoffmann, A.A., Xu, X.-Q., Mo, P.-W., Huang, H.-J., Gong, J.-T., Ju, J.-F., Hong, X.-Y., 2018. Vertical transmission of *Wolbachia* is associated with host vitellogenin in *Laodelphax striatellus*. Front. Microbiol. 9, 2016. https://doi.org/10.3389/fmicb.2018.02016

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

Guy, B., Krell, T., Sanchez, V., Kennel, A., Manin, C., Sodoyer, R., 2005. Do Th1 or Th2 sequence motifs exist in proteins?: Identification of amphipatic immunomodulatory domains in *Helicobacter pylori* catalase. Immunol. Lett. 96, 261–275. https://doi.org/10.1016/j.imlet.2004.09.011

Hamilton, W.D., Zuk, M., 1982. Heritable true fitness and bright birds: a role for parasites? Science 218, 384–387. https://doi.org/10.1126/science.7123238

Hamm, C.A., Begun, D.J., Vo, A., Smith, C.C.R., Saelao, P., Shaver, A.O., Jaenike, J., Turelli, M., 2014. *Wolbachia* do not live by reproductive manipulation alone: infection polymorphism in *Drosophila suzukii* and *D. subpulchrella*. Mol. Ecol. 23, 4871–4885. https://doi.org/10.1111/mec.12901

Hammond, T.M., Rehard, D.G., Xiao, H., Shiu, P.K.T., 2012. Molecular dissection of *Neurospora* Spore killer meiotic drive elements. Proc. Natl. Acad. Sci. U. S. A. 109, 12093–12098. https://doi.org/10.1073/pnas.1203267109

Hancock, P.A., Sinkins, S.P., Godfray, H.C.J., 2011. Population dynamic models of the spread of *Wolbachia*. Am. Nat. 177, 323–333. https://doi.org/10.1086/658121

Havlicek, J., Roberts, S.C., Flegr, J., 2005. Women's preference for dominant male odour: effects of menstrual cycle and relationship status. Biol. Lett. 1, 256–259. https://doi.org/10.1098/rsbl.2005.0332

He, Z., Brinton, B.T., Greenblatt, J., Hassell, J.A., Ingles, C.J., 1993. The transactivator proteins VP16 and GAL4 bind replication factor A. Cell 73, 1223–1232.

He, Z., Zheng, Y., Yu, W.-J., Fang, Y., Mao, B., Wang, Y.-F., 2019. How do *Wolbachia* modify the *Drosophila* ovary? New evidences support the "titration-restitution" model for the mechanisms of *Wolbachia*-induced CI. BMC Genomics 20, 608. https://doi.org/10.1186/s12864-019-5977-6

Hertig, M., 1936. The Rickettsia, *Wolbachia pipientis* (gen. et sp.n.) and Associated Inclusions of the Mosquito, *Culex pipiens*. Parasitology 28, 453–486. https://doi.org/10.1017/S0031182000022666
Hertig, M., Wolbach, S.B., 1924. Studies on rickettsia-like micro-organisms in insects. J. Med. Res. 44, 329-374.7.
Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., Werren, J.H., 2008. How many species are infected with *Wolbachia*? - a statistical analysis of current data. Fems Microbiol. Lett. 281, 215–220. https://doi.org/10.1111/j.1574-6968.2008.01110.x
Hill, G.E., Farmer, K.L., 2005. Carotenoid-based plumage coloration predicts resistance to a novel parasite in the house finch. Naturwissenschaften 92, 30–34. https://doi.org/10.1007/s00114-004-0582-0
Hiroki, M., Tagami, Y., Miura, K., Kato, Y., 2004. Multiple infection with *Wolbachia* inducing different reproductive manipulations in the butterfly *Eurema hecabe*. Proc. Biol. Sci. 271, 1751–1755. https://doi.org/10.1098/rspb.2004.2769
Hoffmann, A., Turelli, M., Harshman, L., 1990. factors affecting the distribution of cytoplasmic incompatibility in *Drosophila simulans*. Genetics 126, 933–948.
Hoffmann, A.A., 1988. Partial cytoplasmic incompatibility between two Australian populations of *Drosophila melanogaster*. Entomol. Exp. Appl. 48, 61–67. https://doi.org/10.1111/j.1570-7458.1988.tb02299.x
Hoffmann, A.A., Clancy, D.J., Merton, E., 1994. Cytoplasmic incompatibility in Australian populations of *Drosophila melanogaster*. Genetics 136, 993–999.
Hoffmann, A.A., Hercus, M., Dagher, H., 1998. Population dynamics of the *Wolbachia* infection causing cytoplasmic incompatibility in *Drosophila melanogaster*. Genetics 148, 221–231.
Hoffmann, A.A., Iturbe-Ormaetxe, I., Callahan, A.G., Phillips, B.L., Billington, K., Axford, J.K., Montgomery, B., Turley, A.P., O'Neill, S.L., 2014. Stability of the *w*Mel *Wolbachia* Infection following invasion into *Aedes aegypti* populations. PLoS Negl. Trop. Dis. 8, e3115. https://doi.org/10.1371/journal.pntd.0003115
Hoffmann, A.A., Montgomery, B.L., Popovici, J., Iturbe-Ormaetxe, I., Johnson, P.H., Muzzi, F., Greenfield, M., Durkan, M., Leong, Y.S., Dong, Y., Cook, H., Axford, J., Callahan, A.G., Kenny, N., Omodei, C., McGraw, E.A., Ryan, P.A., Ritchie, S.A., Turelli, M., O'Neill, S.L., 2011. Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. Nature 476, 454–457. https://doi.org/10.1038/nature10356
Hoffmann, A.A., Ross, P.A., Rašić, G., 2015. *Wolbachia* strains for disease control: ecological and evolutionary considerations. Evol. Appl. 8, 751–768. https://doi.org/10.1111/eva.12286
Hoffmann, A.A., Turelli, M., Simmons, G.M., 1986. Unidirectional incompatibility between populations of *Drosophila simulans*. Evol. Int. J. Org. Evol. 40, 692–701. https://doi.org/10.1111/j.1558-5646.1986.tb00531.x
Holden, P.R., Jones, P., Brookfield, J.F., 1993. Evidence for a *Wolbachia* symbiont in *Drosophila melanogaster*. Genet. Res. 62, 23–29.
Homa, S.T., Vessey, W., Perez-Miranda, A., Riyait, T., Agarwal, A., 2015. Reactive Oxygen Species (ROS) in human semen: determination of a reference range. J. Assist. Reprod. Genet. 32, 757–764. https://doi.org/10.1007/s10815-015-0454-x

Hornett, E.A., Charlat, S., Wedell, N., Jiggins, C.D., Hurst, G.D.D., 2009. Rapidly shifting sex ratio across a species range. Curr. Biol. 19, 1628–1631. https://doi.org/10.1016/j.cub.2009.07.071

Hornett, E.A., Duplouy, A.M.R., Davies, N., Roderick, G.K., Wedell, N., Hurst, G.D.D., Charlat, S., 2008. You can't keep a good parasite down: evolution of a male-killer suppressor uncovers cytoplasmic incompatibility. Evol. Int. J. Org. Evol. 62, 1258–1263. https://doi.org/10.1111/j.1558-5646.2008.00353.x

Hornett, E.A., Moran, B., Reynolds, L.A., Charlat, S., Tazzyman, S., Wedell, N., Jiggins, C.D., Hurst, G.D.D., 2014. The evolution of sex ratio distorter suppression affects a 25 cM genomic region in the butterfly *Hypolimnas bolina*. PLoS Genet. 10, e1004822. https://doi.org/10.1371/journal.pgen.1004822

Horowitz, H., Berg, C.A., 1996. The *Drosophila* pipsqueak gene encodes a nuclear BTB-domain-containing protein required early in oogenesis. Dev. Camb. Engl. 122, 1859–1871.

Hosokawa, T., Koga, R., Kikuchi, Y., Meng, X.-Y., Fukatsu, T., 2010. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. Proc. Natl. Acad. Sci. U. S. A. 107, 769–774. https://doi.org/10.1073/pnas.0911476107

Hu, L., Huang, M., Tang, M., Yu, J., Zheng, B., 2019. *Wolbachia* spread dynamics in multi-regimes of environmental conditions. J. Theor. Biol. 462, 247–258. https://doi.org/10.1016/j.jtbi.2018.11.009

Hu, W., Jiang, Z.-D., Suo, F., Zheng, J.-X., He, W.-Z., Du, L.-L., 2017. A large gene family in fission yeast encodes spore killers that subvert Mendel's law. eLife 6. https://doi.org/10.7554/eLife.26057

Huang, H.-J., Cui, J.-R., Chen, J., Bing, X.-L., Hong, X.-Y., 2019. Proteomic analysis of *Laodelphax striatellus* gonads reveals proteins that may manipulate host reproduction by *Wolbachia*. Insect Biochem. Mol. Biol. 103211. https://doi.org/10.1016/j.ibmb.2019.103211

Huang, M., Luo, J., Hu, L., Zheng, B., Yu, J., 2018. Assessing the efficiency of *Wolbachia* driven *Aedes* mosquito suppression by delay differential equations. J. Theor. Biol. 440, 1–11. https://doi.org/10.1016/j.jtbi.2017.12.012

Hughes, G.L., Koga, R., Xue, P., Fukatsu, T., Rasgon, J.L., 2011. *Wolbachia* infections are virulent and inhibit the human malaria parasite *Plasmodium Falciparum* in *Anopheles Gambiae*. PLOS Pathog. 7, e1002043. https://doi.org/10/bvd5pj

Hughes, G.L., Rasgon, J.L., 2014. Transinfection: a method to investigate *Wolbachia*–host interactions and control arthropod-borne disease. Insect Mol. Biol. 23, 141–151. https://doi.org/10/f5t46g

Hughes, W.O.H., Oldroyd, B.P., Beekman, M., Ratnieks, F.L.W., 2008. Ancestral monogamy shows kin selection is key to the evolution of eusociality. Science 320, 1213–1216. https://doi.org/10.1126/science.1156108

Huigens, M.E., de Almeida, R.P., Boons, P.A.H., Luck, R.F., Stouthamer, R., 2004. Natural interspecific and intraspecific horizontal transfer of parthenogenesis-inducing *Wolbachia* in *Trichogramma* wasps. Proc. R. Soc. B-Biol. Sci. 271, 509–515. https://doi.org/10.1098/rspb.2003.2640

Hurst, G.D.D., Majerus, M.E.N., Walker, L.E., 1993. The importance of cytoplasmic male killing elements in natural populations of the two spot ladybird, *Adalia bipunctata*

(Linnaeus) (Coleoptera: Coccinellidae). Biol. J. Linn. Soc. 49, 195–202. https://doi.org/10.1111/j.1095-8312.1993.tb00898.x
Hurst, G.D.D., Schilthuizen, M., 1998. Selfish genetic elements and speciation. Heredity 80, 2–8. https://doi.org/10.1046/j.1365-2540.1998.00337.x
Hurst, L.D., 1991. The evolution of cytoplasmic incompatibility or when spite can be successful. J. Theor. Biol. 148, 269–277. https://doi.org/10.1016/s0022-5193(05)80344-3
Hurst, L.D., Pomiankowski, A., 1991. Causes of sex ratio bias may account for unisexual sterility in hybrids: a new explanation of Haldane's rule and related phenomena. Genetics 128, 841–858.
Hurvich, C.M., Tsai, C.-L., 1993. A corrected akaike information criterion for vector autoregressive model selection. J. Time Ser. Anal. 14, 271–279. https://doi.org/10.1111/j.1467-9892.1993.tb00144.x
Ilinsky, Y., 2013. Coevolution of *Drosophila melanogaster* mtDNA and *Wolbachia* Genotypes. PLOS ONE 8, e54373. https://doi.org/10.1371/journal.pone.0054373
Imachi, H., Nobu, M.K., Nakahara, N., Morono, Y., Ogawara, M., Takaki, Y., Takano, Y., Uematsu, K., Ikuta, T., Ito, M., Matsui, Y., Miyazaki, M., Murata, K., Saito, Y., Sakai, S., Song, C., Tasumi, E., Yamanaka, Y., Yamaguchi, T., Kamagata, Y., Tamaki, H., Takai, K., 2020. Isolation of an archaeon at the prokaryote-eukaryote interface. Nature 577, 519–525. https://doi.org/10.1038/s41586-019-1916-6
Ishmael, N., Dunning Hotopp, J.C., Ioannidis, P., Biber, S., Sakamoto, J., Siozios, S., Nene, V., Werren, J., Bourtzis, K., Bordenstein, S.R., Tettelin, H., 2009. Extensive genomic diversity of closely related *Wolbachia* strains. Microbiol. Read. Engl. 155, 2211–2222. https://doi.org/10.1099/mic.0.027581-0
Ito, M., Baba, T., Mori, Hirotada, Mori, Hideo, 2005. Functional analysis of 1440 *Escherichia coli* genes using the combination of knock-out library and phenotype microarrays. Metab. Eng. 7, 318–327. https://doi.org/10.1016/j.ymben.2005.06.004
Iturbe-Ormaetxe, I., Howie, J., O'Neill, S., 2007. Development of *Wolbachia* transformation by homologous recombination. Progress report meeting for the Grand Challenges in Human Health Grant, Heron Island, Queensland, Australia.
Jacob, F., Perrin, D., Sanchez, C., Monod, J., Edelstein, S., 1960. The operon: a group of genes with expression coordinated by an operator. CRAcad Sci Paris 250, 1727–1729.
Jaenike, J., 2007. Spontaneous emergence of a new *Wolbachia* phenotype. Evolution 61, 2244–2252. https://doi.org/10.1111/j.1558-5646.2007.00180.x
Jaenike, J., Dyer, K.A., Cornish, C., Minhas, M.S., 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. PLoS Biol. 4, e325. https://doi.org/10.1371/journal.pbio.0040325
James, A.G., Austin, C.J., Cox, D.S., Taylor, D., Calvert, R., 2013. Microbiological and biochemical origins of human axillary odour. FEMS Microbiol. Ecol. 83, 527–540. https://doi.org/10.1111/1574-6941.12054
Jansen, V.A.A., Turelli, M., Godfray, H.C.J., 2008. Stochastic spread of *Wolbachia*. Proc. Biol. Sci. 275, 2769–2776. https://doi.org/10.1098/rspb.2008.0914
Jeffries, C.L., Walker, T., 2016. *Wolbachia* biocontrol strategies for arboviral diseases and the potential influence of resident *Wolbachia* strains in mosquitoes. Curr. Trop. Med. Rep. 3, 20–25. https://doi.org/10.1007/s40475-016-0066-2

Jeffries, C.L., Walker, T., 2015. The potential use of *Wolbachia*-based mosquito biocontrol strategies for japanese encephalitis. PLoS Negl. Trop. Dis. 9, e0003576. https://doi.org/10.1371/journal.pntd.0003576

Jia, F.-X., Yang, M.-S., Yang, W.-J., Wang, J.-J., 2009. Influence of continuous high temperature conditions on *Wolbachia* infection frequency and the fitness of *Liposcelis tricolor* (Psocoptera: Liposcelididae). Environ. Entomol. 38, 1365–1372. https://doi.org/10.1603/022.038.0503

Jiggins, F.M., Randerson, J.P., Hurst, G.D.D., Majerus, M.E.N., 2002. How can sex ratio distorters reach extreme prevalences? Male-killing *Wolbachia* are not suppressed and have near-perfect vertical transmission efficiency in *Acraea encedon*. Evol. Int. J. Org. Evol. 56, 2290–2295. https://doi.org/10.1111/j.0014-3820.2002.tb00152.x

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T.L., 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36, W5-9. https://doi.org/10.1093/nar/gkn201

Ju, J.-F., Bing, X.-L., Zhao, D.-S., Guo, Y., Xi, Z., Hoffmann, A.A., Zhang, K.-J., Huang, H.-J., Gong, J.-T., Zhang, X., Hong, X.-Y., 2019. *Wolbachia* supplement biotin and riboflavin to enhance reproduction in planthoppers. ISME J. https://doi.org/10.1038/s41396-019-0559-9

Ju, J.-F., Hoffmann, A.A., Zhang, Y.-K., Duan, X.-Z., Guo, Y., Gong, J.-T., Zhu, W.-C., Hong, X.-Y., 2017. *Wolbachia*-induced loss of male fertility is likely related to branch chain amino acid biosynthesis and iLvE in *Laodelphax striatellus*. Insect Biochem. Mol. Biol. 85, 11–20. https://doi.org/10.1016/j.ibmb.2017.04.002

Kageyama, D., Narita, S., Watanabe, M., 2012. Insect sex determination manipulated by their endosymbionts: incidences, mechanisms and implications. Insects 3, 161–199. https://doi.org/10.3390/insects3010161

Kajtoch, Ł., Kotásková, N., 2018. Current state of knowledge on *Wolbachia* infection among Coleoptera: a systematic review. PeerJ 6, e4471. https://doi.org/10.7717/peerj.4471

Kamra, D.N., 2005. Rumen microbial ecosystem. Curr. Sci. 89, 124–135.

Karr, T.L., Yang, W., Feder, M.E., 1998. Overcoming cytoplasmic incompatibility in *Drosophila*. Proc. Biol. Sci. 265, 391–395. https://doi.org/10.1098/rspb.1998.0307

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

Kent, B.N., Funkhouser, L.J., Setia, S., Bordenstein, S.R., 2011. Evolutionary genomics of a temperate bacteriophage in an obligate intracellular bacteria (*Wolbachia*). PloS One 6, e24984. https://doi.org/10.1371/journal.pone.0024984

Kittayapong, P., Kaeothaisong, N.-O., Ninphanomchai, S., Limohpasmanee, W., 2018. Combined sterile insect technique and incompatible insect technique: sex separation and quality of sterile *Aedes aegypti* male mosquitoes released in a pilot population suppression trial in Thailand. Parasit. Vectors 11, 657. https://doi.org/10.1186/s13071-018-3214-9

Klass, K.-D., Nalepa, C., Lo, N., 2008. Wood-feeding cockroaches as models for termite evolution (Insecta: Dictyoptera): *Cryptocercus* vs. *Parasphaeria boleiriana*. Mol. Phylogenet. Evol. 46, 809–817. https://doi.org/10.1016/j.ympev.2007.11.028

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

Klasson, L., Westberg, J., Sapountzis, P., Naslund, K., Lutnaes, Y., Darby, A.C., Veneti, Z., Chen, L., Braig, H.R., Garrett, R., Bourtzis, K., Andersson, S.G.E., 2009. The mosaic genome structure of the *Wolbachia w*Ri strain infecting *Drosophila simulans*. Proc. Natl. Acad. Sci. U. S. A. 106, 5725–5730. https://doi.org/10.1073/pnas.0810753106

Knizewski, L., Kinch, L.N., Grishin, N.V., Rychlewski, L., Ginalski, K., 2007. Realm of PD-(D/E)XK nuclease superfamily revisited: detection of novel families with modified transitive meta profile searches. BMC Struct. Biol. 7, 40. https://doi.org/10.1186/1472-6807-7-40

Kodaman, N., Pazos, A., Schneider, B.G., Piazuelo, M.B., Mera, R., Sobota, R.S., Sicinschi, L.A., Shaffer, C.L., Romero-Gallo, J., de Sablet, T., Harder, R.H., Bravo, L.E., Peek, R.M.J., Wilson, K.T., Cover, T.L., Williams, S.M., Correa, P., 2014. Human and *Helicobacter pylori* coevolution shapes the risk of gastric disease. Proc. Natl. Acad. Sci. U. S. A. 111, 1455–1460. https://doi.org/10.1073/pnas.1318093111

Koehncke, A., Telschow, A., Werren, J.H., Hammerstein, P., 2009. Life and death of an influential passenger: *Wolbachia* and the evolution of. PloS One 4, e4425. https://doi.org/10.1371/journal.pone.0004425

Kose, H., Karr, T.L., 1995. Organization of *Wolbachia pipientis* in the *Drosophila* fertilized egg and embryo revealed by an anti-*Wolbachia* monoclonal antibody. Mech. Dev. 51, 275–288.

Koukou, K., Pavlikaki, H., Kilias, G., Werren, J.H., Bourtzis, K., Alahiotis, S.N., 2006. Influence of antibiotic treatment and Wolbachia curing on sexual isolation among *Drosophila melanogaster* cage populations. Evol. Int. J. Org. Evol. 60, 87–96.

Kraaijeveld, K., Franco, P., Reumer, B.M., van Alphen, J.J.M., 2009. Effects of parthenogenesis and geographic isolation on female sexual traits in a parasitoid wasp. Evol. Int. J. Org. Evol. 63, 3085–3096. https://doi.org/10.1111/j.1558-5646.2009.00798.x

Kremer, N., Charif, D., Henri, H., Bataille, M., Prévost, G., Kraaijeveld, K., Vavre, F., 2009. A new case of *Wolbachia* dependence in the genus *Asobara*: evidence for parthenogenesis induction in *Asobara japonica*. Heredity 103, 248–256. https://doi.org/10.1038/hdy.2009.63

Kremer, N., Charif, D., Henri, H., Gavory, F., Wincker, P., Mavingui, P., Vavre, F., 2012. Influence of *Wolbachia* on host gene expression in an obligatory symbiosis. BMC Microbiol. 12 Suppl 1, S7. https://doi.org/10.1186/1471-2180-12-S1-S7

Kriesner, P., Hoffmann, A.A., Lee, S.F., Turelli, M., Weeks, A.R., 2013. Rapid sequential spread of two *Wolbachia* variants in *Drosophila simulans*. PLoS Pathog. 9, e1003607. https://doi.org/10.1371/journal.ppat.1003607

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580. https://doi.org/10.1006/jmbi.2000.4315

Kumar, G.A., Subramaniam, K., 2018. PUF-8 facilitates homologous chromosome pairing by promoting proteasome activity during meiotic entry in *C. elegans*. Development dev.163949. https://doi.org/10.1242/dev.163949

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054

Kuzmenkov, A.I., Peigneur, S., Chugunov, A.O., Tabakmakher, V.M., Efremov, R.G., Tytgat, J., Grishin, E.V., Vassilevski, A.A., 2017. C-Terminal residues in small potassium channel blockers OdK1 and OSK3 from scorpion venom fine-tune the selectivity. Biochim. Biophys. Acta BBA - Proteins Proteomics 1865, 465–472. https://doi.org/10.1016/j.bbapap.2017.02.001

Kyritsis, G.A., Augustinos, A.A., Livadaras, I., Cáceres, C., Bourtzis, K., Papadopoulos, N.T., 2019. Medfly-*Wolbachia* symbiosis: genotype x genotype interactions determine host's life history traits under mass rearing conditions. BMC Biotechnol. 19, 96. https://doi.org/10.1186/s12896-019-0586-7

Landmann, F., Orsi, G.A., Loppin, B., Sullivan, W., 2009. *Wolbachia*-mediated cytoplasmic incompatibility is associated with impaired histone deposition in the male pronucleus. PLOS Pathog. 5, e1000343. https://doi.org/10.1371/journal.ppat.1000343

Lassy, C.W., Karr, T.L., 1996. Cytological analysis of fertilization and early embryonic development in incompatible crosses of *Drosophila simulans*. Mech. Dev. 57, 47–58. https://doi.org/10.1016/0925-4773(96)00527-8

Laven, H., 1967a. A possible model for speciation by cytoplasmic isolation in the *Culex pipiens* complex. Bull. World Health Organ. 37, 263–266.

Laven, H., 1967b. Eradication of *Culex pipiens fatigans* through cytoplasmic incompatibility. Nature 216, 383–384. https://doi.org/10.1038/216383a0

Laven, H., 1951. Crossing experiments with *Culex* strains. Evolution 5, 370–375. https://doi.org/10.1111/j.1558-5646.1951.tb02795.x

Lax, S., Smith, D.P., Hampton-Marcell, J., Owens, S.M., Handley, K.M., Scott, N.M., Gibbons, S.M., Larsen, P., Shogan, B.D., Weiss, S., Metcalf, J.L., Ursell, L.K., Vazquez-Baeza, Y., Van Treuren, W., Hasan, N.A., Gibson, M.K., Colwell, R., Dantas, G., Knight, R., Gilbert, J.A., 2014. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science 345, 1048–1052. https://doi.org/10.1126/science.1254529

Layton, E.M., On, J., Perlmutter, J.I., Bordenstein, S.R., Shropshire, J.D., 2019. Paternal grandmother age affects the strength of *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila melanogaster*. mBio 10. https://doi.org/10.1128/mBio.01879-19

Leal, W.S., 1998. Chemical ecology of phytophagous scarab beetles. Annu. Rev. Entomol. 43, 39–61. https://doi.org/10.1146/annurev.ento.43.1.39

Leclaire, S., Nielsen, J.F., Drea, C.M., 2014. Bacterial communities in meerkat anal scent secretions vary with host sex, age, and group membership. Behav. Ecol. 25, 996–1004. https://doi.org/10.1093/beheco/aru074

Lee, J.-E., Hwang, G.-S., Lee, C.-H., Hong, Y.-S., 2009. Metabolomics reveals alterations in both primary and secondary metabolites by wine bacteria. J. Agric. Food Chem. 57, 10772–10783. https://doi.org/10.1021/jf9028442

Lee, Y.K., Mazmanian, S.K., 2010. Has the microbiota played a critical role in the evolution of the adaptive immune system? Science 330, 1768–1773. https://doi.org/10.1126/science.1195568

Lees, R.S., Gilles, J.R., Hendrichs, J., Vreysen, M.J., Bourtzis, K., 2015. Back to the future: the sterile insect technique against mosquito disease vectors. Curr. Opin. Insect Sci. 10, 156–162. https://doi.org/10.1016/j.cois.2015.05.011

Leftwich, P.T., Edgington, M.P., Harvey-Samuel, T., Carabajal Paladino, L.Z., Norman, V.C., Alphey, L., 2018. Recent advances in threshold-dependent gene drives for mosquitoes. Biochem. Soc. Trans. 46, 1203–1212. https://doi.org/10.1042/BST20180076

LePage, D., Bordenstein, S.R., 2013. *Wolbachia*: Can we save lives with a great pandemic? Trends Parasitol. 29, 385–393. https://doi.org/10.1016/j.pt.2013.06.003

LePage, D.P., Jernigan, K.K., Bordenstein, S.R., 2014. The relative importance of DNA methylation and Dnmt2-mediated epigenetic regulation on *Wolbachia* densities and cytoplasmic incompatibility. PeerJ 2, e678. https://doi.org/10.7717/peerj.678

LePage, D.P., Metcalf, J.A., Bordenstein, Sarah R., On, J., Perlmutter, J.I., Shropshire, J.D., Layton, E.M., Funkhouser-Jones, L.J., Beckmann, J., Bordenstein, Seth R., 2017. Prophage WO genes recapitulate and enhance *Wolbachi*a-induced cytoplasmic incompatibility. Nature 543, 243–247. https://doi.org/10.1038/nature21391

Letunic, I., Doerks, T., Bork, P., 2012. SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res. 40, D302-305. https://doi.org/10.1093/nar/gkr931

Lewis, Z., Champion de Crespigny, F.E., Sait, S.M., Tregenza, T., Wedell, N., 2011. *Wolbachia* infection lowers fertile sperm transfer in a moth. Biol. Lett. 7, 187–189. https://doi.org/10.1098/rsbl.2010.0605

Lewis, Z., Heys, C., Prescott, M., Lize, A., 2014. You are what you eat. Gut Microbes 5, 541–543. https://doi.org/10.4161/gmic.29153

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., Gordon, J.I., 2008. Evolution of mammals and their gut microbes. Science 320, 1647–1651. https://doi.org/10.1126/science.1155725

Li, Q., Korzan, W.J., Ferrero, D.M., Chang, R.B., Roy, D.S., Buchi, M., Lemon, J.K., Kaur, A.W., Stowers, L., Fendt, M., Liberles, S.D., 2013. Synchronous evolution of an odor biosynthesis pathway and behavioral response. Curr. Biol. CB 23, 11–20. https://doi.org/10.1016/j.cub.2012.10.047

Liang, H., Zhou, W., Landweber, L.F., 2006. SWAKK: a web server for detecting positive selection in proteins using a sliding window substitution rate analysis. Nucleic Acids Res. 34, W382–W384. https://doi.org/10.1093/nar/gkl272

Lin, H., Schagat, T., 1997. Neuroblasts: a model for the asymmetric division of stem cells. Trends Genet. TIG 13, 33–39. https://doi.org/10.1016/s0168-9525(96)10050-0

Lindsey, A., Bhattacharya, T., Newton, I., Hardy, R., 2018a. Conflict in the intracellular lives of endosymbionts and viruses: a mechanistic look at *Wolbachia*-mediated pathogen-blocking. Viruses 10, 141. https://doi.org/10.3390/v10040141

Lindsey, A., Rice, D.W., Bordenstein, Sarah R., Brooks, A.W., Bordenstein, Seth R., Newton, I.L.G., 2018b. Evolutionary genetics of cytoplasmic incompatibility genes *cifA* and *cifB* in prophage WO of *Wolbachia*. Genome Biol. Evol. 10, 434–451. https://doi.org/10/gcvmkm

Lindsley, D.L., 1980. Spermatogenesis. Genet. Biol. Drosoph. 2, 225–294.

Linn, C.J., Feder, J.L., Nojima, S., Dambroski, H.R., Berlocher, S.H., Roelofs, W., 2003. Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. Proc. Natl. Acad. Sci. U. S. A. 100, 11490–11493. https://doi.org/10.1073/pnas.1635049100

Liu, C., Wang, J.-L., Zheng, Y., Xiong, E.-J., Li, J.-J., Yuan, L.-L., Yu, X.-Q., Wang, Y.-F., 2014. *Wolbachia*-induced paternal defect in *Drosophila* is likely by interaction with the juvenile hormone pathway. Insect Biochem. Mol. Biol. 49, 49–58. https://doi.org/10.1016/j.ibmb.2014.03.014

Liu, L., Zhang, K.-J., Rong, X., Li, Y.-Y., Liu, H., 2019. Identification of *Wolbachia*-Responsive miRNAs in the Small Brown Planthopper, *Laodelphax striatellus*. Front. Physiol. 10, 928. https://doi.org/10.3389/fphys.2019.00928

Lize, A., McKay, R., Lewis, Z., 2013. Gut microbiota and kin recognition. Trends Ecol. Evol. 28, 325–326. https://doi.org/10.1016/j.tree.2012.10.013

Lo, N., Paraskevopoulos, C., Bourtzis, K., O'Neill, S.L., Werren, J.H., Bordenstein, S.R., Bandi, C., 2007a. Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. Int. J. Syst. Evol. Microbiol. 57, 654–657. https://doi.org/10.1099/ijs.0.64515-0

Lo, N., Paraskevopoulos, C., Bourtzis, K., O'Neill, S.L., Werren, J.H., Bordenstein, S.R., Bandi, C., 2007b. Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. Int. J. Syst. Evol. Microbiol. 57, 654–657. https://doi.org/10.1099/ijs.0.64515-0

Loew, O., 1900. A new enzyme of general occurrence in organisms. Science 11, 701–702. https://doi.org/10.1126/science.11.279.701

Lorenzen, M.D., Gnirke, A., Margolis, J., Garnes, J., Campbell, M., Stuart, J.J., Aggarwal, R., Richards, S., Park, Y., Beeman, R.W., 2008. The maternal-effect, selfish genetic element Medea is associated with a composite Tc1 transposon. Proc. Natl. Acad. Sci. U. S. A. 105, 10085–10089. https://doi.org/10/cs2nw4

Lu, M.-H., Zhang, K.-J., Hong, X.-Y., 2012. Tripartite associations among bacteriophage WO, *Wolbachia*, and host affected by temperature and age in *Tetranychus urticae*. Exp. Appl. Acarol. 58, 207–220. https://doi.org/10.1007/s10493-012-9578-1

Lubke, K.T., Pause, B.M., 2015. Always follow your nose: the functional significance of social chemosignals in human reproduction and survival. Horm. Behav. 68, 134–144. https://doi.org/10.1016/j.yhbeh.2014.10.001

Macdonald, P.M., 1992. The *Drosophila* pumilio gene: an unusually long transcription unit and an unusual protein. Dev. Camb. Engl. 114, 221–232.

Madhav, M., Parry, R., Morgan, J.A.T., James, P., Asgari, S., 2020. *Wolbachia* endosymbiont of the horn fly *Haematobia irritans irritans* : a supergroup A strain with multiple horizontally acquired cytoplasmic incompatibility genes. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.02589-19

Mains, J.W., Brelsfoard, C.L., Rose, R.I., Dobson, S.L., 2016. Female adult *Aedes albopictus* suppression by *Wolbachia*-infected male mosquitoes. Sci. Rep. 6, 33846. https://doi.org/10.1038/srep33846

Mains, J.W., Kelly, P.H., Dobson, K.L., Petrie, W.D., Dobson, S.L., 2019. Localized control of *Aedes aegypti* (Diptera: Culicidae) in Miami, FL, via inundative releases of *Wolbachia*-infected male mosquitoes. J. Med. Entomol. 56, 1296–1303. https://doi.org/10.1093/jme/tjz051

Maistrenko, O.M., Serga, S.V., Vaiserman, A.M., Kozeretska, I.A., 2016. Longevity-modulating effects of symbiosis: insights from *Drosophila-Wolbachia* interaction. Biogerontology 17, 785–803. https://doi.org/10.1007/s10522-016-9653-9

Margulis, L., 1967. On the origin of mitosing cells. J Theor Biol 14, 225–274.
Maroja, L.S., Clark, M.E., Harrison, R.G., 2008. *Wolbachia* plays no role in the one-way reproductive incompatibility between the hybridizing field crickets *Gryllus firmus* and *G. pennsylvanicus*. Heredity 101, 435–444. https://doi.org/10.1038/hdy.2008.75
Marshall, J.F., 1938. The British mosquitoes. British Museum, London.
Martinez, J., Longdon, B., Bauer, S., Chan, Y.-S., Miller, W.J., Bourtzis, K., Teixeira, L., Jiggins, F.M., 2014. Symbionts Commonly Provide Broad Spectrum Resistance to Viruses in Insects: A Comparative Analysis of *Wolbachia* Strains. PLOS Pathog. 10, e1004369. https://doi.org/10.1371/journal.ppat.1004369
Martinez-Rodriguez, P., Bella, J.L., 2018. *Chorthippus parallelus* and *Wolbachia*: Overlapping orthopteroid and bacterial hybrid zones. Front. Genet. 9, 604. https://doi.org/10.3389/fgene.2018.00604
Matsuura, K., 2001. Nestmate recognition mediated by intestinal bacteria in a termite, *Reticulitermes speratus*. Oikos 92, 20–26. https://doi.org/10.1034/j.1600-0706.2001.920103.x
Mayr, E., 1963. Animal species and evolution. Belknap Press of Harvard University Press.
Mayr, E., 1942. Systematics and the origin of species from the viewpoint of a zoologist. New York: Columbia University.
McCutcheon, J.P., von Dohlen, C.D., 2011. An interdependent metabolic patchwork in the nested symbiosis of mealybugs. Curr. Biol. CB 21, 1366–1372. https://doi.org/10.1016/j.cub.2011.06.051
McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H.V., Domazet-Lošo, T., Douglas, A.E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A.H., Kremer, N., Mazmanian, S.K., Metcalf, J.L., Nealson, K., Pierce, N.E., Rawls, J.F., Reid, A., Ruby, E.G., Rumpho, M., Sanders, J.G., Tautz, D., Wernegreen, J.J., 2013. Animals in a bacterial world, a new imperative for the life sciences. Proc. Natl. Acad. Sci. 110, 3229–3236. https://doi.org/10.1073/pnas.1218525110
McFall-Ngai, M., Heath-Heckman, E.A.C., Gillette, A.A., Peyer, S.M., Harvie, E.A., 2012. The secret languages of coevolved symbioses: insights from the *Euprymna scolopes-Vibrio fischeri* symbiosis. Semin. Immunol. 24, 3–8. https://doi.org/10.1016/j.smim.2011.11.006
McFall-Ngai, M.J., 2015. Giving microbes their due--animal life in a microbially dominant world. J. Exp. Biol. 218, 1968–1973. https://doi.org/10.1242/jeb.115121
McKnite, A.M., Perez-Munoz, M.E., Lu, L., Williams, E.G., Brewer, S., Andreux, P.A., Bastiaansen, J.W.M., Wang, X., Kachman, S.D., Auwerx, J., Williams, R.W., Benson, A.K., Peterson, D.A., Ciobanu, D.C., 2012. Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits. PloS One 7, e39191. https://doi.org/10.1371/journal.pone.0039191
Meany, M.K., Conner, W.R., Richter, S.V., Bailey, J.A., Turelli, M., Cooper, B.S., 2019. Loss of cytoplasmic incompatibility and minimal fecundity effects explain relatively low *Wolbachia* frequencies in *Drosophila mauritiana*. Evolution. https://doi.org/10.1111/evo.13745
Meisel, J.D., Panda, O., Mahanti, P., Schroeder, F.C., Kim, D.H., 2014. Chemosensation of bacterial secondary metabolites modulates neuroendocrine signaling and behavior of *C. elegans*. Cell 159, 267–280. https://doi.org/10.1016/j.cell.2014.09.011

Mereschkowsky, C., 1910. Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. Biol. Cent. 30, 278-288,289-303,322-374,353–367.
Metcalf, J.A., Jo, M., Bordenstein, Sarah R., Jaenike, J., Bordenstein, Seth R., 2014. Recent genome reduction of *Wolbachia* in *Drosophila recens* targets phage WO and narrows candidates for reproductive parasitism. PeerJ 2, e529. https://doi.org/10.7717/peerj.529
Miller, P.B., Obrik-Uloho, O.T., Phan, M.H., Medrano, C.L., Renier, J.S., Thayer, J.L., Wiessner, G., Bloch Qazi, M.C., 2014. The song of the old mother: reproductive senescence in female *Drosophila*. Fly (Austin) 8, 127–139. https://doi.org/10.4161/19336934.2014.969144
Miller, W.J., Ehrman, L., Schneider, D., 2010. Infectious speciation revisited: impact of symbiont-depletion on female fitness and mating behavior of *Drosophila paulistorum*. PLoS Pathog. 6, e1001214. https://doi.org/10.1371/journal.ppat.1001214
Mizanur, R.M., Frasca, V., Swaminathan, S., Bavari, S., Webb, R., Smith, L.A., Ahmed, S.A., 2013. The C terminus of the catalytic domain of type A botulinum neurotoxin may facilitate product release from the active site. J. Biol. Chem. 288, 24223–24233. https://doi.org/10.1074/jbc.M113.451286
Mohanty, I., Rath, A., Mahapatra, N., Hazra, R.K., 2016. *Wolbachia*: A biological control strategy against arboviral diseases. J. Vector Borne Dis. 53, 199–207.
Molloy, J.C., Sommer, U., Viant, M.R., Sinkins, S.P., 2016. *Wolbachia* modulates lipid metabolism in *Aedes albopictus* mosquito cells. Appl. Environ. Microbiol. 82, 3109–3120. https://doi.org/10.1128/AEM.00275-16
Moreau, J., Bertin, A., Caubet, Y., Rigaud, T., 2001. Sexual selection in an isopod with *Wolbachia*-induced sex reversal: males prefer real females. J. Evol. Biol. 14, 388–394. https://doi.org/10.1046/j.1420-9101.2001.00292.x
Moreira, L.A., Iturbe-Ormaetxe, I., Jeffery, J.A., Lu, G., Pyke, A.T., Hedges, L.M., Rocha, B.C., Hall-Mendelin, S., Day, A., Riegler, M., Hugo, L.E., Johnson, K.N., Kay, B.H., McGraw, E.A., van den Hurk, A.F., Ryan, P.A., O'Neill, S.L., 2009. A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and Plasmodium. Cell 139, 1268–1278. https://doi.org/10.1016/j.cell.2009.11.042
Moretti, R., Marzo, G.A., Lampazzi, E., Calvitti, M., 2018. Cytoplasmic incompatibility management to support Incompatible Insect Technique against *Aedes albopictus*. Parasit. Vectors 11, 649. https://doi.org/10.1186/s13071-018-3208-7
Morgan, J.A.W., Bending, G.D., White, P.J., 2005. Biological costs and benefits to plant-microbe interactions in the rhizosphere. J. Exp. Bot. 56, 1729–1739. https://doi.org/10.1093/jxb/eri205
Morita, M., Wang, H.L., 2001. Association between oral malodor and adult periodontitis: a review. J. Clin. Periodontol. 28, 813–819. https://doi.org/10.1034/j.1600-051x.2001.028009813.x
Mouton, L., Henri, H., Bouletreau, M., Vavre, F., 2006. Effect of temperature on *Wolbachia* density and impact on cytoplasmic incompatibility. Parasitology 132, 49–56. https://doi.org/10.1017/S0031182005008723
Murakawa, G.J., Kwan, C., Yamashita, J., Nierlich, D.P., 1991. Transcription and decay of the lac messenger: role of an intergenic terminator. J. Bacteriol. 173, 28–36. https://doi.org/10.1128/jb.173.1.28-36.1991

Mustafa, M.S., Rastogi, V., Gupta, R.K., Jain, S., Singh, P.M.P., Gupta, A., 2016. *Wolbachia*: the selfish trojan horse in dengue control. Med. J. Armed Forces India 72, 373–376. https://doi.org/10.1016/j.mjafi.2015.07.002

Najarro, M.A., Sumethasorn, M., Lamoureux, A., Turner, T.L., 2015. Choosing mates based on the diet of your ancestors: replication of non-genetic assortative mating in *Drosophila melanogaster*. PeerJ 3, e1173. https://doi.org/10.7717/peerj.1173

Nakanishi, K., Hoshino, M., Nakai, M., Kunimi, Y., 2008. Novel RNA sequences associated with late male killing in *Homona magnanima*. Proc. Biol. Sci. 275, 1249–1254. https://doi.org/10.1098/rspb.2008.0013

Narita, S., Kageyama, D., Nomura, M., Fukatsu, T., 2007. Unexpected mechanism of symbiont-induced reversal of insect sex: feminizing *Wolbachia* continuously acts on the butterfly *Eurema hecabe* during larval development. Appl. Environ. Microbiol. 73, 4332–4341. https://doi.org/10.1128/AEM.00145-07

Narita, S., Shimajiri, Y., Nomura, M., 2009. Strong cytoplasmic incompatibility and high vertical transmission rate can explain the high frequencies of *Wolbachia* infection in Japanese populations of *Colias erate* poliographus (Lepidoptera: Pieridae). Bull. Entomol. Res. 99, 385–391. https://doi.org/10.1017/S0007485308006469

Negri, I., Pellecchia, M., Mazzoglio, P.J., Patetta, A., Alma, A., 2006. Feminizing *Wolbachia* in *Zyginidia pullula* (Insecta, Hemiptera), a leafhopper with an XX/X0 sex-determination system. Proc. Biol. Sci. 273, 2409–2416. https://doi.org/10.1098/rspb.2006.3592

Newton, I.L.G., Rice, D.W., 2019. The Jekyll and Hyde symbiont: could *Wolbachia* be a nutritional mutualist? J. Bacteriol. https://doi.org/10.1128/JB.00589-19

Nguyen, D.T., Morrow, J.L., Spooner-Hart, R.N., Riegler, M., 2017. Independent cytoplasmic incompatibility induced by *Cardinium* and *Wolbachia* maintains endosymbiont coinfections in haplodiploid thrips populations. Evol. Int. J. Org. Evol. 71, 995–1008. https://doi.org/10.1111/evo.13197

Ni, J.-Q., Zhou, R., Czech, B., Liu, L.-P., Holderbaum, L., Yang-Zhou, D., Shim, H.-S., Tao, R., Handler, D., Karpowicz, P., Binari, R., Booker, M., Brennecke, J., Perkins, L.A., Hannon, G.J., Perrimon, N., 2011. A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. Nat. Methods 8, 405–407. https://doi.org/10.1038/nmeth.1592

Niang, E.H.A., Bassene, H., Fenollar, F., Mediannikov, O., 2018. Biological Control of Mosquito-Borne Diseases: The Potential of *Wolbachia*-Based Interventions in an IVM Framework. J. Trop. Med. 2018, 1470459. https://doi.org/10.1155/2018/1470459

Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., Pettersson, S., 2012. Host-gut microbiota metabolic interactions. Science 336, 1262–1267. https://doi.org/10.1126/science.1223813

Nikolouli, K., Colinet, H., Renault, D., Enriquez, T., Mouton, L., Gibert, P., Sassu, F., Cáceres, C., Stauffer, C., Pereira, R., Bourtzis, K., 2018. Sterile insect technique and *Wolbachia* symbiosis as potential tools for the control of the invasive species *Drosophila suzukii*. J. Pest Sci. 91, 489–503. https://doi.org/10/gc55cr

Nishanth, M.J., Simon, B., 2020. Functions, mechanisms and regulation of Pumilio/Puf family RNA binding proteins: a comprehensive review. Mol. Biol. Rep. 47, 785–807. https://doi.org/10.1007/s11033-019-05142-6

Nougue, O., Gallet, R., Chevin, L.-M., Lenormand, T., 2015. Niche limits of symbiotic gut microbiota constrain the salinity tolerance of brine shrimp. Am. Nat. 186, 390–403. https://doi.org/10.1086/682370

Nuckolls, N.L., Bravo Núñez, M.A., Eickbush, M.T., Young, J.M., Lange, J.J., Yu, J.S., Smith, G.R., Jaspersen, S.L., Malik, H.S., Zanders, S.E., 2017. *wtf* genes are prolific dual poison-antidote meiotic drivers. eLife 6. https://doi.org/10.7554/eLife.26033

Ochman, H., Worobey, M., Kuo, C.-H., Ndjango, J.-B.N., Peeters, M., Hahn, B.H., Hugenholtz, P., 2010. Evolutionary relationships of wild hominids recapitulated by gut microbial communities. PLoS Biol. 8, e1000546. https://doi.org/10.1371/journal.pbio.1000546

O'Connor, L., Plichart, C., Sang, A.C., Brelsfoard, C.L., Bossin, H.C., Dobson, S.L., 2012. Open release of male mosquitoes infected with a *Wolbachia* biopesticide: field performance and infection containment. PLoS Negl. Trop. Dis. 6, e1797. https://doi.org/10.1371/journal.pntd.0001797

Oliver, K.M., Campos, J., Moran, N.A., Hunter, M.S., 2008. Population dynamics of defensive symbionts in aphids. Proc. Biol. Sci. 275, 293–299. https://doi.org/10.1098/rspb.2007.1192

Olsen, K., Reynolds, K.T., Hoffmann, A.A., 2001. A field cage test of the effects of the endosymbiont *Wolbachia* on *Drosophila melanogaster*. Heredity 86, 731–737. https://doi.org/10.1046/j.1365-2540.2001.00892.x

O'Mahony, S.M., Clarke, G., Dinan, T.G., Cryan, J.F., 2017. Early-life adversity and brain development: Is the microbiome a missing piece of the puzzle? Neuroscience 342, 37–54. https://doi.org/10.1016/j.neuroscience.2015.09.068

O'Neill, S.L., 2018. The use of *Wolbachia* by the World Mosquito Program to interrupt transmission of *Aedes aegypti* transmitted viruses. Adv. Exp. Med. Biol. 1062, 355–360. https://doi.org/10.1007/978-981-10-8727-1_24

O'Neill, S.L., Karr, T.L., 1990. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. Nature 348, 178–180. https://doi.org/10.1038/348178a0

O'Neill, S.L., Ryan, P.A., Turley, A.P., Wilson, G., Retzki, K., Iturbe-Ormaetxe, I., Dong, Y., Kenny, N., Paton, C.J., Ritchie, S.A., Brown-Kenyon, J., Stanford, D., Wittmeier, N., Anders, K.L., Simmons, C.P., 2018. Scaled deployment of *Wolbachia* to protect the community from dengue and other *Aedes* transmitted arboviruses. Gates Open Res. 2, 36. https://doi.org/10.12688/gatesopenres.12844.2

Ote, M., Ueyama, M., Yamamoto, D., 2016. *Wolbachia* protein tomo targets nanos mrna and restores germ stem cells in *Drosophila* sex-lethal mutants. Curr. Biol. CB 26, 2223–2232. https://doi.org/10.1016/j.cub.2016.06.054

Pannebakker, B.A., Schidlo, N.S., Boskamp, G.J.F., Dekker, L., van Dooren, T.J.M., Beukeboom, L.W., Zwaan, B.J., Brakefield, P.M., van Alphen, J.J.M., 2005. Sexual functionality of *Leptopilina clavipes* (Hymenoptera: Figitidae) after reversing *Wolbachia*-induced parthenogenesis. J. Evol. Biol. 18, 1019–1028. https://doi.org/10.1111/j.1420-9101.2005.00898.x

Papafotiou, G., Oehler, S., Savakis, C., Bourtzis, K., 2011. Regulation of *Wolbachia* ankyrin domain encoding genes in *Drosophila* gonads. Res. Microbiol. 162, 764–772. https://doi.org/10.1016/j.resmic.2011.06.012

Parisi, M., Lin, H., 1999. The *Drosophila* pumilio gene encodes two functional protein isoforms that play multiple roles in germline development, gonadogenesis, oogenesis and embryogenesis. Genetics 153, 235–250.

Peña-Cardeña, A., Grande, R., Sánchez, J., Tabashnik, B.E., Bravo, A., Soberón, M., Gómez, I., 2018. The C-terminal protoxin region of *Bacillus thuringiensis* Cry1Ab toxin has a

functional role in binding to GPI-anchored receptors in the insect midgut. J. Biol. Chem. 293, 20263–20272. https://doi.org/10.1074/jbc.RA118.005101
Penn, D., Potts, W.K., 1998. Chemical signals and parasite-mediated sexual selection. Trends Ecol. Evol. 13, 391–396. https://doi.org/10.1016/s0169-5347(98)01473-6
Petrella, L.N., Smith-Leiker, T., Cooley, L., 2007. The Ovhts polyprotein is cleaved to produce fusome and ring canal proteins required for *Drosophila* oogenesis. Dev. Camb. Engl. 134, 703–712. https://doi.org/10.1242/dev.02766
Pianotti, R., Pitts, G., 1978. Effects of an antiseptic mouthwash on odorigenic microbes in the human gingival crevice. J. Dent. Res. 57, 175–179. https://doi.org/10.1177/00220345780570020201
Pijls, J.W.A.M., Steenbergen, H.J. van, Alphen, J.J.M. van, 1996. Asexuality cured: the relations and differences between sexual and asexual *Apoanagyrus diversicornis*. Heredity 76, 506–513. https://doi.org/10.1038/hdy.1996.73
Pinto, S.B., Stainton, K., Harris, S., Kambris, Z., Sutton, E.R., Bonsall, M.B., Parkhill, J., Sinkins, S.P., 2013. Transcriptional regulation of *Culex pipiens* mosquitoes by *Wolbachia* influences cytoplasmic incompatibility. PLoS Pathog. 9, e1003647. https://doi.org/10.1371/journal.ppat.1003647
Poinsot, D., Bourtzis, K., Markakis, G., Savakis, C., Mercot, H., 1998. *Wolbachia* transfer from *Drosophila melanogaster* into *D. simulans*: Host effect and cytoplasmic incompatibility relationships. Genetics 150, 227–237.
Poinsot, D., Charlat, S., Mercot, H., 2003. On the mechanism of *Wolbachia*-induced cytoplasmic incompatibility: confronting the models with the facts. Bioessays 25, 259–265. https://doi.org/10.1002/bies.10234
Polin, S., Simon, J.-C., Outreman, Y., 2014. An ecological cost associated with protective symbionts of aphids. Ecol. Evol. 4, 826–830. https://doi.org/10.1002/ece3.991
Presgraves, D.C., 2000. A genetic test of the mechanism of *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila*. Genetics 154, 771–776.
Prout, T., 1994. Some evolutionary possibilities for a microbe that causes incompatibility in its host. Evolution 48, 909–911. https://doi.org/10.2307/2410496
Puggioli, A., Calvitti, M., Moretti, R., Bellini, R., 2016. *w*Pip *Wolbachia* contribution to *Aedes albopictus* SIT performance: Advantages under intensive rearing. Acta Trop. 164, 473–481. https://doi.org/10.1016/j.actatropica.2016.10.014
Ramírez-Puebla, S.T., Ormeño-Orrillo, E., Vera-Ponce de León, A., Lozano, L., Sanchez-Flores, A., Rosenblueth, M., Martínez-Romero, E., 2016. Genomes of *Candidatus Wolbachia bourtzisii w*DacA and *Candidatus Wolbachia pipientis w*DacB from the Cochineal Insect *Dactylopius coccus* (Hemiptera: Dactylopiidae). G3 Bethesda Md 6, 3343–3349. https://doi.org/10.1534/g3.116.031237
Randerson, J.P., Jiggins, F.M., Hurst, L.D., 2000. Male killing can select for male mate choice: a novel solution to the paradox of the lek. Proc. Biol. Sci. 267, 867–874. https://doi.org/10.1098/rspb.2000.1083
Rasgon, J.L., 2008. Using Predictive Models to Optimize *Wolbachia*-Based Strategies for Vector-Borne Disease Control, in: Aksoy, S. (Ed.), Transgenesis and the management of vector-borne disease. Springer New York, New York, NY, pp. 114–125. https://doi.org/10.1007/978-0-387-78225-6_10

Raychoudhury, R., Werren, J.H., 2012. Host genotype changes bidirectional to unidirectional cytoplasmic incompatibility in *Nasonia longicornis*. Heredity 108, 105–114. https://doi.org/10.1038/hdy.2011.53

Reynolds, K.T., Hoffmann, A.A., 2002. Male age, host effects and the weak expression or non-expression of cytoplasmic incompatibility in *Drosophila* strains infected by maternally transmitted *Wolbachia*. Genet. Res. 80, 79–87.

Reynolds, K.T., Thomson, L.J., Hoffmann, A.A., 2003. The effects of host age, host nuclear background and temperature on phenotypic effects of the virulent *Wolbachia* strain popcorn in *Drosophila melanogaster*. Genetics 164, 1027–1034.

Riegler, M., Stauffer, C., 2002. *Wolbachia* infections and superinfections in cytoplasmically incompatible populations of the European cherry fruit fly *Rhagoletis cerasi* (Diptera, Tephritidae). Mol. Ecol. 11, 2425–2434.

Rigaud, T., Juchault, P., 1993. Conflict between feminizing sex ratio distorters and an autosomal masculinizing gene in the terrestrial isopod *Armadillidium vulgare Latr.* Genetics 133, 247–252.

Rigaud, T., Juchault, P., 1992. Genetic control of the vertical transmission of a cytoplasmic sex factor in *Armadillidium vulgare Latr.* (Crustacea, Oniscidea). Heredity 68, 47–52. https://doi.org/10.1038/hdy.1992.6

Rigaud, T., Moreau, J., 2004. A cost of *Wolbachia*-induced sex reversal and female-biased sex ratios: decrease in female fertility after sperm depletion in a terrestrial isopod. Proc. Biol. Sci. 271, 1941–1946. https://doi.org/10.1098/rspb.2004.2804

Ringo, J., Sharon, G., Segal, D., 2011. Bacteria-induced sexual isolation in *Drosophila*. Fly (Austin) 5, 310–315. https://doi.org/10.4161/fly.5.4.15835

Riparbelli, M.G., Giordano, R., Callaini, G., 2007. Effects of *Wolbachia* on sperm maturation and architecture in *Drosophila simulans* Riverside. Mech. Dev. 124, 699–714. https://doi.org/10.1016/j.mod.2007.07.001

Ritchie, S.A., Townsend, M., Paton, C.J., Callahan, A.G., Hoffmann, A.A., 2015. Application of *w*MelPop *Wolbachia* Strain to Crash Local Populations of *Aedes aegypti*. PLoS Negl. Trop. Dis. 9, e0003930. https://doi.org/10.1371/journal.pntd.0003930

Ritchie, S.A., van den Hurk, A.F., Smout, M.J., Staunton, K.M., Hoffmann, A.A., 2018. Mission accomplished? We need a guide to the 'post release' world of *Wolbachia* for *Aedes*-borne disease control. Trends Parasitol. 34, 217–226. https://doi.org/10.1016/j.pt.2017.11.011

Rohou, A., Nield, J., Ushkaryov, Y.A., 2007. Insecticidal toxins from black widow spider venom. Toxicon 49, 531–549. https://doi.org/10.1016/j.toxicon.2006.11.021

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Rørth, P., 1998. Gal4 in the *Drosophila* female germline. Mech. Dev. 78, 113–118. https://doi.org/10/bftzng

Rosenberg, E., Zilber-Rosenberg, I., 2013. Introduction: symbioses and the hologenome concept, in: Rosenberg, E., Zilber-Rosenberg, I. (Eds.), The hologenome concept: human, animal and plant microbiota. Springer International Publishing, Cham, pp. 1–8. https://doi.org/10.1007/978-3-319-04241-1_1

Ross, P.A., Axford, J.K., Yang, Q., Staunton, K.M., Ritchie, S.A., Richardson, K.M., Hoffmann, A.A., 2020. Heatwaves cause fluctuations in *w*Mel *Wolbachia* densities and frequencies

in *Aedes aegypti*. PLoS Negl. Trop. Dis. 14, e0007958. https://doi.org/10.1371/journal.pntd.0007958
Ross, P.A., Hoffmann, A.A., 2018. Continued susceptibility of the *w*Mel *Wolbachia* infection in *Aedes aegypti* to heat stress following field deployment and selection. Insects 9. https://doi.org/10.3390/insects9030078
Ross, P.A., Ritchie, S.A., Axford, J.K., Hoffmann, A.A., 2019. Loss of cytoplasmic incompatibility in *Wolbachia*-infected *Aedes aegypti* under field conditions. PLoS Negl. Trop. Dis. 13, e0007357. https://doi.org/10.1371/journal.pntd.0007357
Round, J.L., Mazmanian, S.K., 2009. The gut microbiota shapes intestinal immune responses during health and disease. Nat. Rev. Immunol. 9, 313–323. https://doi.org/10.1038/nri2515
Ruhmann, H., Wensing, K.U., Neuhalfen, N., Specker, J.-H., Fricke, C., 2016. Early reproductive success in *Drosophila* males is dependent on maturity of the accessory gland. Behav. Ecol. 27, 1859–1868. https://doi.org/10.1093/beheco/arw123
Ryan, S., Saul, G., 1968. Post-fertilization effect of incompatibility factors in *Mormoniella*. Mol. Gen. Genet. 103, 29-. https://doi.org/10.1007/BF00271154
Sacchi, L., Genchi, M., Clementi, E., Negri, I., Alma, A., Ohler, S., Sassera, D., Bourtzis, K., Bandi, C., 2010. Bacteriocyte-like cells harbour *Wolbachia* in the ovary of *Drosophila melanogaster* (Insecta, Diptera) and *Zyginidia pullula* (Insecta, Hemiptera). Tissue Cell 42, 328–333. https://doi.org/10.1016/j.tice.2010.07.009
Sadowski, I., Ma, J., Triezenberg, S., Ptashne, M., 1988. GAL4-VP16 is an unusually potent transcriptional activator. Nature 335, 563–564. https://doi.org/10.1038/335563a0
Sampson, T.R., Mazmanian, S.K., 2015. Control of brain development, function, and behavior by the microbiome. Cell Host Microbe 17, 565–576. https://doi.org/10.1016/j.chom.2015.04.011
Sanders, J.G., Powell, S., Kronauer, D.J.C., Vasconcelos, H.L., Frederickson, M.E., Pierce, N.E., 2014. Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. Mol. Ecol. 23, 1268–1283. https://doi.org/10.1111/mec.12611
Saxton, T.K., Lyndon, A., Little, A.C., Roberts, S.C., 2008. Evidence that androstadienone, a putative human chemosignal, modulates women's attributions of men's attractiveness. Horm. Behav. 54, 597–601. https://doi.org/10.1016/j.yhbeh.2008.06.001
Schebeck, M., Feldkirchner, L., Stauffer, C., Schuler, H., 2019. Dynamics of an ongoing *Wolbachia* spread in the european cherry fruit fly, *Rhagoletis cerasi* (Diptera: Tephritidae). Insects 10, 172. https://doi.org/10.3390/insects10060172
Schmidt, T.L., Barton, N.H., Rašić, G., Turley, A.P., Montgomery, B.L., Iturbe-Ormaetxe, I., Cook, P.E., Ryan, P.A., Ritchie, S.A., Hoffmann, A.A., O'Neill, S.L., Turelli, M., 2017. Local introduction and heterogeneous spatial spread of dengue-suppressing *Wolbachia* through an urban population of *Aedes aegypti*. PLOS Biol. 15, e2001894. https://doi.org/10.1371/journal.pbio.2001894
Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675. https://doi.org/10.1038/nmeth.2089
Schultz, M.J., Isern, S., Michael, S.F., Corley, R.B., Connor, J.H., Frydman, H.M., 2017. Variable inhibition of Zika virus replication by different *Wolbachia* strains in mosquito cell cultures. J. Virol. 91. https://doi.org/10.1128/JVI.00339-17

Sears, C.L., Buckwold, S.L., Shin, J.W., Franco, A.A., 2006. The C-terminal region of *Bacteroides fragilis* toxin is essential to its biological activity. Infect. Immun. 74, 5595–5601. https://doi.org/10.1128/IAI.00135-06

Seidel, H.S., Ailion, M., Li, J., van Oudenaarden, A., Rockman, M.V., Kruglyak, L., 2011. A novel sperm-delivered toxin causes late-stage embryo lethality and transmission ratio distortion in *C. elegans*. PLoS Biol. 9, e1001115. https://doi.org/10.1371/journal.pbio.1001115

Serbus, L.R., Casper-Lindley, C., Landmann, F., Sullivan, W., 2008. The genetics and cell biology of *Wolbachia*-host interactions. Annu. Rev. Genet. 42, 683–707. https://doi.org/10.1146/annurev.genet.41.110306.130354

Sharon, G., Segal, D., Ringo, J.M., Hefetz, A., Zilber-Rosenberg, I., Rosenberg, E., 2010. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U. S. A. 107, 20051–20056. https://doi.org/10.1073/pnas.1009906107

Shaw, W.R., Catteruccia, F., 2019. Vector biology meets disease control: using basic research to fight vector-borne diseases. Nat. Microbiol. 4, 20–34. https://doi.org/10.1038/s41564-018-0214-7

Shoemaker, D.D., Katju, V., Jaenike, J., 1999. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. Evolution 53, 1157–1164. https://doi.org/10.1111/j.1558-5646.1999.tb04529.x

Shropshire, J.D., Bordenstein, S.R., 2019. Two-by-one model of cytoplasmic incompatibility: synthetic recapitulation by transgenic expression of *cifA* and *cifB* in *Drosophila*. PLOS Genet. 15, e1008221. https://doi.org/10.1371/journal.pgen.1008221

Shropshire, J.D., Bordenstein, S.R., 2016. Speciation by symbiosis: the microbiome and behavior. mBio 7, e01785-15. https://doi.org/10.1128/mBio.01785-15

Shropshire, J.D., Leigh, B., Bordenstein, Sarah R., Duplouy, A., Riegler, M., Brownlie, J.C., Bordenstein, Seth R., 2019. Models and nomenclature for cytoplasmic incompatibility: caution over premature conclusions – a response to Beckmann et al. Trends Genet. 0. https://doi.org/10.1016/j.tig.2019.03.004

Shropshire, J.D., On, J., Layton, E.M., Zhou, H., Bordenstein, S.R., 2018. One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U. S. A. 115, 4987–4991. https://doi.org/10.1073/pnas.1800650115

Shuster, S.M., Lonsdorf, E.V., Wimp, G.M., Bailey, J.K., Whitham, T.G., 2006. Community heritability measures the evolutionary consequences of indirect genetic effects on community structure. Evol. Int. J. Org. Evol. 60, 991–1003.

Sicard, M., Bouchon, D., Ceyrac, L., Raimond, R., Thierry, M., Le Clec'h, W., Marcadé, I., Caubet, Y., Grève, P., 2014. Bidirectional cytoplasmic incompatibility caused by *Wolbachia* in the terrestrial isopod *Porcellio dilatatus*. J. Invertebr. Pathol. 121, 28–36. https://doi.org/10.1016/j.jip.2014.06.007

Siddharth, J., Holway, N., Parkinson, S.J., 2013. A Western diet ecological module identified from the "humanized" mouse microbiota predicts diet in adults and formula feeding in children. PloS One 8, e83689. https://doi.org/10.1371/journal.pone.0083689

Siegmund, T., Lehmann, M., 2002. The *Drosophila* Pipsqueak protein defines a new family of helix-turn-helix DNA-binding proteins. Dev. Genes Evol. 212, 152–157. https://doi.org/10.1007/s00427-002-0219-2

Sin, Y.W., Buesching, C.D., Burke, T., Macdonald, D.W., 2012. Molecular characterization of the microbial communities in the subcaudal gland secretion of the European badger

(*Meles meles*). FEMS Microbiol. Ecol. 81, 648–659. https://doi.org/10.1111/j.1574-6941.2012.01396.x
Smadja, C., Butlin, R.K., 2009. On the scent of speciation: the chemosensory system and its role in premating isolation. Heredity 102, 77–97. https://doi.org/10.1038/hdy.2008.55
Smith-White, S., Woodhill, A.R., 1955. The Nature and significance of non-reciprocal fertility in *Aedes scutellaris* and other mosquitoes. Proc. Linn. Soc. New South Wales 79.
Snook, R.R., Cleland, S.Y., Wolfner, M.F., Karr, T.L., 2000. Offsetting effects of *Wolbachia* infection and heat shock on sperm production in *Drosophila simulans*: analyses of fecundity, fertility and accessory gland proteins. Genetics 155, 167–178.
Solignac, M., Vautrin, D., Rousset, F., 1994. Widespread occurrence of the proteobacteria *Wolbachia* and partial cytoplasmic incompatibility in *Drosophila melanogaster*. Comptes Rendus Acad. Sci. - Ser. III 317, 461–470.
Song, S.J., Lauber, C., Costello, E.K., Lozupone, C.A., Humphrey, G., Berg-Lyons, D., Caporaso, J.G., Knights, D., Clemente, J.C., Nakielny, S., Gordon, J.I., Fierer, N., Knight, R., 2013. Cohabiting family members share microbiota with one another and with their dogs. eLife 2, e00458. https://doi.org/10.7554/eLife.00458
Southall, T.D., Elliott, D.A., Brand, A.H., 2008. The GAL4 system: a versatile toolkit for gene expression in *Drosophila*. CSH Protoc. 2008, pdb.top49. https://doi.org/10.1101/pdb.top49
Spor, A., Koren, O., Ley, R., 2011. Unravelling the effects of the environment and host genotype on the gut microbiome. Nat. Rev. Microbiol. 9, 279–290. https://doi.org/10.1038/nrmicro2540
Steczkiewicz, K., Muszewska, A., Knizewski, L., Rychlewski, L., Ginalski, K., 2012. Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily. Nucleic Acids Res. 40, 7016–7045. https://doi.org/10.1093/nar/gks382
Steinway, S.N., Dannenfelser, R., Laucius, C.D., Hayes, J.E., Nayak, S., 2010. JCoDA: a tool for detecting evolutionary selection. BMC Bioinformatics 11, 284. https://doi.org/10.1186/1471-2105-11-284
Stennett, M.D., Etges, W.J., 1997. Premating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. III. Epicuticular hydrocarbon variation is determined by use of different host plants in *Drosophila mojavensis* and *Drosophila arizonae*. J. Chem. Ecol. 23, 2803–2824. https://doi.org/10.1023/A:1022519228346
Stevens, D., Cornmell, R., Taylor, D., Grimshaw, S.G., Riazanskaia, S., Arnold, D.S., Fernstad, S.J., Smith, A.M., Heaney, L.M., Reynolds, J.C., Thomas, C.L.P., Harker, M., 2015. Spatial variations in the microbial community structure and diversity of the human foot is associated with the production of odorous volatiles. FEMS Microbiol. Ecol. 91, 1–11. https://doi.org/10.1093/femsec/fiu018
Stouthamer, R., 1997. *Wolbachia*-induced parthenogenesis. Inherit. Microb-Org. Arthropod Reprod. Influ. Passeng. 102–124.
Stouthamer, R., Breeuwer, J.A.J., Hurst, G.D.D., 1999. *Wolbachia pipientis*: Microbial manipulator of arthropod reproduction. Annu. Rev. Microbiol. 53, 71–102. https://doi.org/10.1146/annurev.micro.53.1.71
Stouthamer, R., Luck, R.F., Hamilton, W.D., 1990. Antibiotics cause parthenogenetic *Trichogramma* (Hymenoptera/Trichogrammatidae) to revert to sex. Proc. Natl. Acad. Sci. U. S. A. 87, 2424–2427. https://doi.org/10.1073/pnas.87.7.2424

Sumi, T., Miura, K., Miyatake, T., 2017. *Wolbachia* density changes seasonally amongst populations of the pale grass blue butterfly, *Zizeeria maha* (Lepidoptera: Lycaenidae). PloS One 12, e0175373. https://doi.org/10.1371/journal.pone.0175373

Sutton, E.R., Harris, S.R., Parkhill, J., Sinkins, S.P., 2014. Comparative genome analysis of *Wolbachia* strain *w*Au. BMC Genomics 15, 928. https://doi.org/10.1186/1471-2164-15-928

Tan, J.M.M., Wong, E.S.P., Kirkpatrick, D.S., Pletnikova, O., Ko, H.S., Tay, S.-P., Ho, M.W.L., Troncoso, J., Gygi, S.P., Lee, M.K., Dawson, V.L., Dawson, T.M., Lim, K.-L., 2008. Lysine 63-linked ubiquitination promotes the formation and autophagic clearance of protein inclusions associated with neurodegenerative diseases. Hum. Mol. Genet. 17, 431–439. https://doi.org/10.1093/hmg/ddm320

Taylor, M.J., Bordenstein, S.R., Slatko, B., 2018. Microbe Profile: *Wolbachia*: a sex selector, a viral protector and a target to treat filarial nematodes. Microbiology 164, 1345–1347. https://doi.org/10.1099/mic.0.000724

Taylor, M.W., Radax, R., Steger, D., Wagner, M., 2007. Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. Microbiol. Mol. Biol. Rev. MMBR 71, 295–347. https://doi.org/10.1128/MMBR.00040-06

Teixeira, L., Ferreira, A., Ashburner, M., 2008. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. PLoS Biol. 6, e2. https://doi.org/10.1371/journal.pbio.1000002

Telschow, A., Flor, M., Kobayashi, Y., Hammerstein, P., Werren, J.H., 2007. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and speciation: mainland-island model. PloS One 2, e701. https://doi.org/10.1371/journal.pone.0000701

Telschow, A., Yamamura, N., Werren, J.H., 2005. Bidirectional cytoplasmic incompatibility and the stable coexistence of two *Wolbachia* strains in parapatric host populations. J. Theor. Biol. 235, 265–274. https://doi.org/10.1016/j.jtbi.2005.01.008

Terradas, G., McGraw, E.A., 2017. *Wolbachia*-mediated virus blocking in the mosquito vector *Aedes aegypti*. Curr. Opin. Insect Sci. 22, 37–44. https://doi.org/10.1016/j.cois.2017.05.005

Theis, K.R., Venkataraman, A., Dycus, J.A., Koonter, K.D., Schmitt-Matzen, E.N., Wagner, A.P., Holekamp, K.E., Schmidt, T.M., 2013. Symbiotic bacteria appear to mediate hyena social odors. Proc. Natl. Acad. Sci. U. S. A. 110, 19832–19837. https://doi.org/10.1073/pnas.1306477110

Thiem, S., 2014. A genetic manipulation system for *Wolbachia* in mosquitoes. USDA, Michigan State University.

Thompson, J.N., 1987. Symbiont-induced speciation. Biol. J. Linn. Soc. 32, 385–393. https://doi.org/10.1111/j.1095-8312.1987.tb00439.x

Tolley, S.J.A., Nonacs, P., Sapountzis, P., 2019. *Wolbachia* horizontal transmission events in ants: what do we know and what can we learn? Front. Microbiol. 10, 296. https://doi.org/10.3389/fmicb.2019.00296

Toomey, M.E., Panaram, K., Fast, E.M., Beatty, C., 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. https://doi.org/10.1073/pnas.1301524110

Tortosa, P., Charlat, S., Labbe, P., Dehecq, J.-S., Barre, H., Weill, M., 2010. *Wolbachia* age-sex-specific density in *Aedes albopictus*: a host evolutionary response to cytoplasmic incompatibility? PloS One 5, e9700. https://doi.org/10.1371/journal.pone.0009700

Tracey, W.D., Ning, X., Klingler, M., Kramer, S.G., Gergen, J.P., 2000. Quantitative analysis of gene function in the *Drosophila* embryo. Genetics 154, 273–284.

Tram, U., Fredrick, K., Werren, J.H., Sullivan, W., 2006. Paternal chromosome segregation during the first mitotic division determines *Wolbachia*-induced cytoplasmic incompatibility phenotype. J. Cell Sci. 119, 3655–3663. https://doi.org/10.1242/jcs.03095

Tram, U., Sullivan, W., 2002. Role of delayed nuclear envelope breakdown and mitosis in *Wolbachia*-induced cytoplasmic incompatibility. Science 296, 1124–1126. https://doi.org/10.1126/science.1070536

Trpis, M., Perrone, J.B., Reissig, M., Parker, K.L., 1981. Control of cytoplasmic incompatibility in the *Aedes scutellaris* complex incompatible crosses become compatible by treatment of larvae with heat or antibiotics. J. Hered. 72, 313–317. https://doi.org/10.1093/oxfordjournals.jhered.a109513

Truitt, A.M., Kapun, M., Kaur, R., Miller, W.J., 2018. *Wolbachia* modifies thermal preference in *Drosophila melanogaster*. Environ. Microbiol. https://doi.org/10.1111/1462-2920.14347

Tsuchida, T., Koga, R., Horikawa, M., Tsunoda, T., Maoka, T., Matsumoto, S., Simon, J.-C., Fukatsu, T., 2010. Symbiotic bacterium modifies aphid body color. Science 330, 1102–1104. https://doi.org/10.1126/science.1195463

Tsuchiya, Y., Ohta, J., Ishida, Y., Morisaki, H., 2008. Cloth colorization caused by microbial biofilm. Colloids Surf. B Biointerfaces 64, 216–222. https://doi.org/10.1016/j.colsurfb.2008.01.028

Tung, J., Barreiro, L.B., Burns, M.B., Grenier, J.-C., Lynch, J., Grieneisen, L.E., Altmann, J., Alberts, S.C., Blekhman, R., Archie, E.A., 2015. Social networks predict gut microbiome composition in wild baboons. eLife 4. https://doi.org/10.7554/eLife.05224

Turelli, M., 2010. Cytoplasmic incompatibility in populations with overlapping generations. Evol. Int. J. Org. Evol. 64, 232–241. https://doi.org/10.1111/j.1558-5646.2009.00822.x

Turelli, M., 1994. Evolution of incompatibility-inducing microbes and their hosts. Evol. Int. J. Org. Evol. 48, 1500–1513. https://doi.org/10.1111/j.1558-5646.1994.tb02192.x

Turelli, M., Barton, N.H., 2017. Deploying dengue-suppressing *Wolbachia*: Robust models predict slow but effective spatial spread in *Aedes aegypti*. Theor. Popul. Biol. 115, 45–60. https://doi.org/10.1016/j.tpb.2017.03.003

Turelli, M., Cooper, B.S., Richardson, K.M., Ginsberg, P.S., Peckenpaugh, B., Antelope, C.X., Kim, K.J., May, M.R., Abrieux, A., Wilson, D.A., Bronski, M.J., Moore, B.R., Gao, J.-J., Eisen, M.B., Chiu, J.C., Conner, W.R., Hoffmann, A.A., 2018a. Rapid global spread of *w*Ri-like *Wolbachia* across multiple *Drosophila*. Curr. Biol. 28, 963-971.e8. https://doi.org/10.1016/j.cub.2018.02.015

Turelli, M., Cooper, B.S., Richardson, K.M., Ginsberg, P.S., Peckenpaugh, B., Antelope, C.X., Kim, K.J., May, M.R., Abrieux, A., Wilson, D.A., Bronski, M.J., Moore, B.R., Gao, J.-J., Eisen, M.B., Chiu, J.C., Conner, W.R., Hoffmann, A.A., 2018b. Rapid global spread of *w*Ri-like *Wolbachia* across multiple *Drosophila*. Curr. Biol. 28, 963-971.e8. https://doi.org/10.1016/j.cub.2018.02.015

Turelli, M., Hoffmann, A.A., 1999. Microbe-induced cytoplasmic incompatibility as a mechanism for introducing transgenes into arthropod populations. Insect Mol. Biol. 8, 243–255. https://doi.org/10/brk6mb

Turelli, M., Hoffmann, A.A., 1995. Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. Genetics 140, 1319–1338.
Turelli, M., Hoffmann, A.A., 1991. Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature 353, 440–442. https://doi.org/10.1038/353440a0
Turelli, M., Hoffmann, A.A., McKechnie, S.W., 1992. Dynamics of cytoplasmic incompatibility and Mtdna variation in natural *Drosophila simulans* populations. Genetics 132, 713–723.
Turner, T.R., James, E.K., Poole, P.S., 2013. The plant microbiome. Genome Biol. 14, 209. https://doi.org/10.1186/gb-2013-14-6-209
Vala, F., Egas, M., Breeuwer, J.A.J., Sabelis, M.W., 2004. *Wolbachia* affects oviposition and mating behaviour of its spider mite host. J. Evol. Biol. 17, 692–700. https://doi.org/10.1046/j.1420-9101.2003.00679.x
Vala, F., Weeks, A., Claessen, D., Breeuwer, J. a. J., Sabelis, M.W., 2002. Within- and between-population variation for *Wolbachia*-induced reproductive incompatibility in a haplodiploid mite. Evol. Int. J. Org. Evol. 56, 1331–1339.
Vallenet, D., Engelen, S., Mornico, D., Cruveiller, S., Fleury, L., Lajus, A., Rouy, Z., Roche, D., Salvignol, G., Scarpelli, C., Médigue, C., 2009. MicroScope: a platform for microbial genome annotation and comparative genomics. Database J. Biol. Databases Curation 2009, bap021. https://doi.org/10.1093/database/bap021
van den Hurk, A.F., Hall-Mendelin, S., Pyke, A.T., Frentiu, F.D., McElroy, K., Day, A., Higgs, S., O'Neill, S.L., 2012. Impact of *Wolbachia* on infection with chikungunya and yellow fever viruses in the mosquito vector *Aedes aegypti*. PLoS Negl. Trop. Dis. 6, e1892. https://doi.org/10.1371/journal.pntd.0001892
van Opijnen, T., Breeuwer, J.A., 1999. High temperatures eliminate *Wolbachia*, a cytoplasmic incompatibility inducing endosymbiont, from the two-spotted spider mite. Exp. Appl. Acarol. 23, 871–881.
van Opstal, E.J., Bordenstein, S.R., 2015. Rethinking heritability of the microbiome. Science 349, 1172–1173. https://doi.org/10.1126/science.aab3958
Vavre, F., Dedeine, F., Quillon, M., Fouillet, P., Fleury, F., Bouletreau, M., 2001. Within-species diversity of *Wolbachia*-induced cytoplasmic incompatibility in haplodiploid insects. Evol. Int. J. Org. Evol. 55, 1710–1714.
Vavre, F., Fleury, F., Varaldi, J., Fouillet, P., Bouletreau, M., 2000. Evidence for female mortality in *Wolbachia*-mediated cytoplasmic incompatibility in haplodiploid insects: epidemiologic and evolutionary consequences. Evol. Int. J. Org. Evol. 54, 191–200.
Vector Control Advisory Group, W.H.O., 2016. Mosquito (vector) control emergency response and preparedness for Zika virus. World Health Organ. Website. URL http://www.who.int/neglected_diseases/news/mosquito_vector_control_response/en/ (accessed 12.3.19).
Veneti, Z., Bentley, J.K., Koana, T., Braig, H.R., Hurst, G.D.D., 2005. A functional dosage compensation complex required for male killing in *Drosophila*. Science 307, 1461–1463. https://doi.org/10.1126/science.1107182
Veneti, Z., Clark, M.E., Karr, T.L., Savakis, C., Bourtzis, K., 2004. Heads or tails: Host-parasite interactions in the *Drosophila-Wolbachia* system. Appl. Environ. Microbiol. 70, 5366–5372. https://doi.org/10.1128/AEM.70.9.5366-5372.2004
Veneti, Z., Clark, M.E., Zabalou, S., Karr, T.L., Savakis, C., Bourtzis, K., 2003. Cytoplasmic incompatibility and sperm cyst infection in different *Drosophila-Wolbachia* associations. Genetics 164, 545–552.

Venu, I., Durisko, Z., Xu, J., Dukas, R., 2014. Social attraction mediated by fruit flies' microbiome. J. Exp. Biol. 217, 1346–1352. https://doi.org/10.1242/jeb.099648
Versace, E., Nolte, V., Pandey, R.V., Tobler, R., Schlotterer, C., 2014. Experimental evolution reveals habitat-specific fitness dynamics among *Wolbachia* clades in *Drosophila melanogaster*. Mol. Ecol. 23, 802–814. https://doi.org/10.1111/mec.12643
Vizcaíno, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.-W., Wang, R., Hermjakob, H., 2016. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 44, 11033. https://doi.org/10.1093/nar/gkw880
Walker, T., Johnson, P.H., Moreira, L.A., Iturbe-Ormaetxe, I., Frentiu, F.D., McMeniman, C.J., Leong, Y.S., Dong, Y., Axford, J., Kriesner, P., Lloyd, A.L., Ritchie, S.A., O'Neill, S.L., Hoffmann, A.A., 2011. The *w*Mel *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. Nature 476, 450–453. https://doi.org/10.1038/nature10355
Wallin, I.E., 1927. Symbionticism and the origin of species. Рипол Классик.
Wang, G.H., Jia, L.-Y., Xiao, J.-H., Huang, D.-W., 2016a. Discovery of a new *Wolbachia* supergroup in cave spider species and the lateral transfer of phage WO among distant hosts. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 41, 1–7. https://doi.org/10.1016/j.meegid.2016.03.015
Wang, G.H., Sun, B.F., Xiong, T.L., Wang, Y.K., Murfin, K.E., Xiao, J.H., Huang, D.-W., 2016b. Bacteriophage WO can mediate horizontal gene transfer in endosymbiotic *Wolbachia* genomes. Front. Microbiol. 7, 1867. https://doi.org/10.3389/fmicb.2016.01867
Wang, J., Kalyan, S., Steck, N., Turner, L.M., Harr, B., Kunzel, S., Vallier, M., Hasler, R., Franke, A., Oberg, H.-H., Ibrahim, S.M., Grassl, G.A., Kabelitz, D., Baines, J.F., 2015. Analysis of intestinal microbiota in hybrid house mice reveals evolutionary divergence in a vertebrate hologenome. Nat. Commun. 6, 6440. https://doi.org/10.1038/ncomms7440
Wang, N., Jia, S., Xu, H., Liu, Y., Huang, D.-W., 2016. Multiple horizontal transfers of bacteriophage WO and host *Wolbachia* in fig wasps in a closed community. Front. Microbiol. 7, 136. https://doi.org/10.3389/fmicb.2016.00136
Wangen, J.R., Green, R., 2020. Stop codon context influences genome-wide stimulation of termination codon readthrough by aminoglycosides. eLife 9, e52611. https://doi.org/10.7554/eLife.52611
Weeks, A.R., Turelli, M., Harcombe, W.R., Reynolds, K.T., Hoffmann, A.A., 2007. From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*. Plos Biol. 5, 997–1005. https://doi.org/10.1371/journal.pbio.0050114
Wegner, K.M., Volkenborn, N., Peter, H., Eiler, A., 2013. Disturbance induced decoupling between host genetics and composition of the associated microbiome. BMC Microbiol. 13, 252. https://doi.org/10.1186/1471-2180-13-252
Weidmann, C.A., Goldstrohm, A.C., 2012. *Drosophila* Pumilio Protein Contains Multiple Autonomous Repression Domains That Regulate mRNAs Independently of Nanos and Brain Tumor. Mol. Cell. Biol. 32, 527–540. https://doi.org/10.1128/MCB.06052-11
Weinert, L.A., Araujo-Jnr, E.V., Ahmed, M.Z., Welch, J.J., 2015. The incidence of bacterial endosymbionts in terrestrial arthropods. Proc R Soc B 282, 20150249. https://doi.org/10.1098/rspb.2015.0249
Werren, J.H., 1997. Biology of *Wolbachia*. Annu. Rev. Entomol. 42, 587–609. https://doi.org/10.1146/annurev.ento.42.1.587

Werren, J.H., Baldo, L., Clark, M.E., 2008. *Wolbachia*: master manipulators of invertebrate biology. Nat. Rev. Microbiol. 6, 741–751. https://doi.org/10.1038/nrmicro1969
Werren, J.H., Jaenike, J., 1995. *Wolbachia* and cytoplasmic incompatibility in mycophagous *Drosophila* and their relatives. Heredity 75 (Pt 3), 320–326.
Werren, J.H., Loehlin, D.W., 2009. Rearing *Sarcophaga bullata* fly hosts for *Nasonia* (parasitoid wasp). Cold Spring Harb. Protoc. 2009, pdb.prot5308. https://doi.org/10.1101/pdb.prot5308
Wertz, I.E., Dixit, V.M., 2010. Signaling to NF-kappaB: regulation by ubiquitination. Cold Spring Harb. Perspect. Biol. 2, a003350. https://doi.org/10.1101/cshperspect.a003350
West, S.A., Diggle, S.P., Buckling, A., Gardner, A., Griffin, A.S., 2007. The social lives of microbes. Annu. Rev. Ecol. Evol. Syst. 38, 53–77. https://doi.org/10.1146/annurev.ecolsys.38.091206.095740
White-Cooper, H., 2012. Tissue, cell type and stage-specific ectopic gene expression and RNAi induction in the *Drosophila* testis. Spermatogenesis 2, 11–22. https://doi.org/10.4161/spmg.19088
Wiwatanaratanabutr, I., Kittayapong, P., 2009. Effects of crowding and temperature on *Wolbachia* infection density among life cycle stages of *Aedes albopictus*. J. Invertebr. Pathol. 102, 220–224. https://doi.org/10.1016/j.jip.2009.08.009
Woese, C.R., 1998. Default taxonomy: Ernst Mayr's view of the microbial world. Proc. Natl. Acad. Sci. U. S. A. 95, 11043–11046. https://doi.org/10.1073/pnas.95.19.11043
Wong Sak Hoi, J., Dumas, B., 2010. Ste12 and Ste12-like proteins, fungal transcription factors regulating development and pathogenicity. Eukaryot. Cell 9, 480–485. https://doi.org/10.1128/EC.00333-09
Wright, J.D., Barr, A.R., 1981. *Wolbachia* and the normal and incompatible eggs of *Aedes polynesiensis* (Diptera: Culicidae). J. Invertebr. Pathol. 38, 409–418. https://doi.org/10.1016/0022-2011(81)90109-9
Wright, J.D., Wang, B.-T., 1980. Observations on *Wolbachiae* in mosquitoes. J. Invertebr. Pathol. 35, 200–208. https://doi.org/10.1016/0022-2011(80)90185-8
Wu, M., Sun, L.V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J.C., McGraw, E.A., Martin, W., Esser, C., Ahmadinejad, N., Wiegand, C., Madupu, R., Beanan, M.J., Brinkac, L.M., Daugherty, S.C., Durkin, A.S., Kolonay, J.F., Nelson, W.C., Mohamoud, Y., Lee, P., Berry, K., Young, M.B., Utterback, T., Weidman, J., Nierman, W.C., Paulsen, I.T., Nelson, K.E., Tettelin, H., O'Neill, S.L., Eisen, J.A., 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis w*Mel: A streamlined genome overrun by mobile genetic elements. Plos Biol. 2, 327–341. https://doi.org/10.1371/journal.pbio.0020069
Xi, Z., Gavotte, L., Xie, Y., Dobson, S.L., 2008. Genome-wide analysis of the interaction between the endosymbiotic bacterium *Wolbachia* and its *Drosophila* host. BMC Genomics 9, 1. https://doi.org/10.1186/1471-2164-9-1
Yajjala, V.K., Widhelm, T.J., Endres, J.L., Fey, P.D., Bayles, K.W., 2016. Generation of a transposon mutant library in *Staphylococcus aureus* and *Staphylococcus epidermidis* using *Bursa aurealis*. Methods Mol. Biol. Clifton NJ 1373, 103–110. https://doi.org/10.1007/7651_2014_189
Yamada, R., Floate, K.D., Riegler, M., O'Neill, S.L., 2007. Male development time influences the strength of *Wolbachia*-induced cytoplasmic incompatibility expression in *Drosophila melanogaster*. Genetics 177, 801–808. https://doi.org/10.1534/genetics.106.068486

Yamada, R., Iturbe-Ormaetxe, I., Brownlie, J.C., O'Neill, S.L., 2011. Functional test of the influence of *Wolbachia* genes on cytoplasmic incompatibility expression in *Drosophila melanogaster*. Insect Mol. Biol. 20, 75–85. https://doi.org/10.1111/j.1365-2583.2010.01042.x

Yang, J., Zhao, X., Cheng, K., Du, H., Ouyang, Y., Chen, J., Qiu, S., Huang, J., Jiang, Y., Jiang, L., Ding, J., Wang, J., Xu, C., Li, X., Zhang, Q., 2012. A killer-protector system regulates both hybrid sterility and segregation distortion in rice. Science 337, 1336–1340. https://doi.org/10.1126/science.1223702

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., Heath, A.C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J.G., Lozupone, C.A., Lauber, C., Clemente, J.C., Knights, D., Knight, R., Gordon, J.I., 2012. Human gut microbiome viewed across age and geography. Nature 486, 222–227. https://doi.org/10.1038/nature11053

Yeap, H.L., Rasic, G., Endersby-Harshman, N.M., Lee, S.F., Arguni, E., Le Nguyen, H., Hoffmann, A.A., 2016. Mitochondrial DNA variants help monitor the dynamics of *Wolbachia* invasion into host populations. Heredity 116, 265–276. https://doi.org/10.1038/hdy.2015.97

Yen, J.H., Barr, A.R., 1973. The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. J. Invertebr. Pathol. 22, 242–250. https://doi.org/10.1016/0022-2011(73)90141-9

Yen, J.H., Barr, A.R., 1971. New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens L.* Nature 232, 657–658. https://doi.org/10.1038/232657a0

Yoshida, K., Sanada-Morimura, S., Huang, S.-H., Tokuda, M., 2019. Influences of two coexisting endosymbionts, CI-inducing *Wolbachia* and male-killing *Spiroplasma* , on the performance of their host *Laodelphax striatellus* (Hemiptera: Delphacidae). Ecol. Evol. https://doi.org/10.1002/ece3.5392

Yu, B., Huang, Z., 2015. Variations in antioxidant genes and male infertility. BioMed Res. Int. 2015. https://doi.org/10.1155/2015/513196

Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J., Brinkman, F.S.L., 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinforma. Oxf. Engl. 26, 1608–1615. https://doi.org/10.1093/bioinformatics/btq249

Yuan, L.-L., Chen, X., Zong, Q., Zhao, T., Wang, J.-L., Zheng, Y., Zhang, M., Wang, Z., Brownlie, J.C., Yang, F., Wang, Y.-F., 2015. Quantitative proteomic analyses of molecular mechanisms associated with cytoplasmic incompatibility in *Drosophila melanogaster* induced by *Wolbachia*. J. Proteome Res. 14, 3835–3847. https://doi.org/10.1021/acs.jproteome.5b00191

Zabalou, S., Apostolaki, A., Pattas, S., Veneti, Z., Paraskevopoulos, C., Livadaras, I., Markakis, G., Brissac, T., Mercot, H., Bourtzis, K., 2008. Multiple rescue factors within a *Wolbachia* strain. Genetics 178, 2145–2160. https://doi.org/10.1534/genetics.107.086488

Zabalou, S., Charlat, S., Nirgianaki, A., Lachaise, D., Mercot, H., Bourtzis, K., 2004. Natural *Wolbachia* infections in the *Drosophila yakuba* species complex do not induce cytoplasmic incompatibility but fully rescue the *w*Ri modification. Genetics 167, 827–834. https://doi.org/10.1534/genetics.103.015990

Zachar, I., Boza, G., 2020. Endosymbiosis before eukaryotes: mitochondrial establishment in protoeukaryotes. Cell. Mol. Life Sci. https://doi.org/10.1007/s00018-020-03462-6

Zala, S.M., Bilak, A., Perkins, M., Potts, W.K., Penn, D.J., 2015. Female house mice initially shun infected males, but do not avoid mating with them. Behav. Ecol. Sociobiol. 69, 715–722. https://doi.org/10.1007/s00265-015-1884-2

Zchori-Fein, E., Faktor, O., Zeidan, M., Gottlieb, Y., Czosnek, H., Rosen, D., 1995. Parthenogenesis-inducing microorganisms in *Aphytis* (Hymenoptera: Aphelinidae). Insect Mol. Biol. 4, 173–178. https://doi.org/10.1111/j.1365-2583.1995.tb00023.x

Zhang, D., Zheng, X., Xi, Z., Bourtzis, K., Gilles, J.R.L., 2015. Combining the sterile insect technique with the incompatible insect technique: I-impact of *Wolbachia* infection on the fitness of triple- and double-infected strains of *Aedes albopictus*. PloS One 10, e0121126. https://doi.org/10.1371/journal.pone.0121126

Zhang, Y., 2009. I-TASSER: fully automated protein structure prediction in CASP8. Proteins 77 Suppl 9, 100–113. https://doi.org/10.1002/prot.22588

Zhang, Y., Skolnick, J., 2004. Scoring function for automated assessment of protein structure template quality. Proteins 57, 702–710. https://doi.org/10.1002/prot.20264

Zhang, Y.-K., Ding, X.-L., Rong, X., Hong, X.-Y., 2015. How do hosts react to endosymbionts? A new insight into the molecular mechanisms underlying the *Wolbachia*-host association. Insect Mol. Biol. 24, 1–12. https://doi.org/10.1111/imb.12128

Zheng, X., Zhang, D., Li, Y., Yang, C., Wu, Y., Liang, X., Liang, Y., Pan, X., Hu, L., Sun, Q., Wang, X., Wei, Y., Zhu, J., Qian, W., Yan, Z., Parker, A.G., Gilles, J.R.L., Bourtzis, K., Bouyer, J., Tang, M., Zheng, B., Yu, J., Liu, J., Zhuang, J., Hu, Zhigang, Zhang, M., Gong, J.-T., Hong, X.-Y., Zhang, Z., Lin, L., Liu, Q., Hu, Zhiyong, Wu, Z., Baton, L.A., Hoffmann, A.A., Xi, Z., 2019. Incompatible and sterile insect techniques combined eliminate mosquitoes. Nature. https://doi.org/10.1038/s41586-019-1407-9

Zheng, Y., Ren, P.-P., Wang, J.-L., Wang, Y.-F., 2011. *Wolbachia*-induced cytoplasmic incompatibility is associated with decreased Hira expression in male *Drosophila*. PloS One 6, e19512. https://doi.org/10.1371/journal.pone.0019512

Zheng, Y., Shen, W., Bi, J., Chen, M.-Y., Wang, R.-F., Ai, H., Wang, Y.-F., 2019. Small RNA analysis provides new insights into cytoplasmic incompatibility in *Drosophila melanogaster* induced by *Wolbachia*. J. Insect Physiol. 118, 103938. https://doi.org/10.1016/j.jinsphys.2019.103938

Zilber-Rosenberg, I., Rosenberg, E., 2008. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. FEMS Microbiol. Rev. 32, 723–735. https://doi.org/10.1111/j.1574-6976.2008.00123.x

Zug, R., Hammerstein, P., 2015. *Wolbachia* and the insect immune system: what reactive oxygen species can tell us about the mechanisms of *Wolbachia*–host interactions. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.01201

Zug, R., 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. https://doi.org/10.1371/journal.pone.0038544