EXPLORING THE PERSISTENCE OF A NON-MALE-KILLING S*PIROPLASMA* INFECTION IN *DROSOPHILA* AND THE POTENTIAL HOST RANGE OF *DROSOPHILA*-INFECTING SPIROPLASMAS

A Dissertation

by

IGOR VILCHEZ-RAMIREZ

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Norman O. Dronen
Co-Chair of Committee,	Spencer J. Johnston
Committee Members,	Deborah A. Siegele
	Aaron M. Tarone
Head of Department,	David Caldwell

August 2018

Major Subject: Wildlife and Fisheries Sciences

ABSTRACT

This work addresses the persistence of a non-male-killing *Spiroplasma* strain infecting the fly *Drosophila hydei* under conditions where there are no apparent fitness benefits to the host. This work also addresses the potential host range and the factors affecting those potential host ranges of distinctly related *Spiroplasma* strains on a group of *Drosophila* hosts that are distinctly related. We began by exploring if the infection modifies a set of characters in the host as compared to the same characters in uninfected hosts. The characteristics explored were sex ratio, larval stage mortality, longevity and fecundity. These factors were chosen because if modified by the infection, they could promote its long-term persistence. We then reexamined longevity and fecundity and expanded to the exploration of the effect of the infection on age at maturity and preferential mating to infected females by all males. Our results show that longevity and early fecundity are affected by the infection, with infected females having a higher early fecundity and shorter life-span than uninfected ones; together, these factors can promote the long-term persistence of the infection. These results suggest that fitness effects have to be reexamined on other host-parasite pairs where no benefits to the host seem to occur along with the long-term persistence of parasites. We also report the results of the potential host range experiments, for which we cross-infected 5 *Spiroplasma* strains into each other's 5 natural *Drosophila* hosts, generating 25 pairs of infections. The Rohde Index was used to evaluate the host range. The factors evaluated for their effect on host ranges were parasite and host phylogenies, intrinsic host and parasite characteristics, and the interaction of all of these factors. The potential host ranges of *Spiroplasma* strains are varied and not predictable simply based on phylogenetic relatedness of either bacteria or host. Only the host and parasite intrinsic characteristics and their interaction

were found to affect the infection success rate of *Spiroplasma* when introduced to new and old hosts.

To my wife Sophie

To my parents, Rosa and Herman, united on this page as they should be,

To my brothers and sisters, Lery, Ruth, Herman, Rosa Maria, Martin, Carlos, and Jaime

ACKNOWLEDGEMENTS

First and foremost, to Sophie, who brings meaning to the passing of the time and the new spaces we occupy every instant.

My deep gratitude to Dr. Norman Dronen and Dr. Spencer J. Johnston, my co-advisors, who welcomed me into their laboratories when I needed a place to perform my work. Their unconditional support at every step of this work was vital to the completion of my research. Dr. Johnston, through what was an unquestioningly colossal undertaking, helped me to organize the final product of my work.

I thank Dr. Deborah Siegele for providing her invaluable mentorship, support, and gracious humanity when I needed it the most.

I give thanks to Dr. Charles Criscione and Dr. Aaron Tarone for their critical input in strengthening this work.

I am grateful to Mr. Marek and Banka Cieplucha, from Pracownia Architektury Formart for their funding and support to this entire work.

Finally, I want to thank to my friend Dr. Alejandra Gonzalez for her reliability and friendship through this experience.

CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a dissertation committee consisting of Professor Norman O. Dronen, my advisor, along with my co-advisor Professor Spencer J. Johnston, and Professor Deborah A. Siegele of the Department of Biology, and Professor Aaron M. Tarone of the Department of Entomology. All work for the dissertation was completed independently by the student.

Graduate study was supported by Mr. Marek Cieplucha of Pracownia Architektury Format, A Private Foundation from the Republic of Poland.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

This work addresses two issues related to the evolution of host-parasite relationships, namely, 1) the long-term persistence of maternally heritable endosymbionts and 2) the potential host range of parasites. These two topics were studied by using the *Drosophila-Spiroplasma* association.

Spiroplasma (Procaryotae: Mollicutes: Entomoplasmatales: Spiroplasmataceae) are commonly associated with organisms such as plants, shrimps, honey bees, ticks, the fruit fly *Drosphila*, and humans, among others (Regassa & Gasparich, 2006; Yu, Liu, Ruan, Chen, &Wand, 2009; Lorenz, Schroeder, & Reischl, 2002; Aquilino, et al., 2015; Mueller et al., 2015). Spiroplasmas act as pathogens or mutualists, although their effect on many of the hosts is currently unknown (Regassa & Gasparich, 2006; Yu et al., 2009).

Spiroplasma is a maternally transmitted parasite (vertically transmitted) in *Drosophila* and produces two overt phenotypes. One phenotype is male-killing, which causes infected females to produce almost exclusively female progeny (Williamson & Poulson, 1979). Death of males typically occurs during embryogenesis (Counce & Poulson, 1962) (early male-killing) or at the larval/pupal stage or early age (late male-killing) (Kageyama, Anbutsu, Shimada, & Fukatsu, 2007). The other phenotype can be called non-male-killing (i.e., it does not kill male offspring); this was first observed in a population of *Drosophila hydei* from Japan (Ota, Kawabe, Oishi, & Poulson, 1979). Eight strains of male-killing *Spiroplasma* and 10 strains of non-male-killing *Spiroplasma* have been reported infecting a total of 16 diversely related species of *Drosophila* (*D. melanogaster* can be infected with both, a male-killing strain and a non-male-killing strain

and *D. hydei* with 2 non-male killing strains) (Mateos et al., 2006; Fig. 1 in Haselkorn, 2010a). These spiroplasmas are grouped into 4 clades, namely poulsonii, citri, ixodetis, and tenebrosa and they are more closely related to non-*Drosphila*-associated spiroplasmas within their clades than to the *Drosophila*-associated spiroplasmas from other closely related clades (Haselkorn, Markow, & Moran, 2009). The *Spiroplasma* strains and their natural *Drosophila* hosts (*Drosophila-Spiroplasma* pairs) produce host-parasite phylogenies that are incongruent (Haselkorn et al., 2009). The diverse association of the *Spiroplasma* strains and the incongruent host-parasite trees are the consequence of multiple introductions of the bacteria into *Drosophila* (Haselkorn et al., 2009).

Both, male-killing and non-male-killing *Spiroplasma-Drosophila* associations are optional for the host (*Drosophila* can reproduce and survive without the presence of *Spiroplasma*).
To date, the male-killing strains of *Spiroplasma* have been reported to occur in the tropical regions of America and Africa (Williamson & Poulson, 1979; Montenegro, Solferini, Klaczko, & Hurst, 2005; Montenegro, Hatadani, Medeiros, & Klaczko, 2006; Pool, Wong, & Aquadro, 2006; Mateos et al., 2006). The non-male-killing strains have been mainly reported from regions near the Tropic of Cancer and the temperate zone of the northern hemisphere, with some strains occurring near the equator (e.g., *Spiroplasma* infecting *D. ananassae* from Hawaii) (Ota et al., 1979; Mateos et al., 2006; Serga, Rozhok, Protsenko, Kozeretska, & Mousseau, 2010; Haselkorn et al., 2009). The *Drosophila-Spiroplasma* association is a newly discovered relationship and no systematic worldwide survey has been attempted so far; therefore, it is expected that the discovery of new *Spiroplasma* strains infecting *Drosophila* species will continue in the foreseeable future.

The long-term persistence of maternally heritable endosymbionts can be achieved by one or a combination of several mechanisms. Perfect vertical transmission would, in theory, guarantee the persistence of any symbiont for many generations to come. Endosymbionts without perfect transmission will be lost from the host unless there is horizontal transfer or an increase in fitness of the host due to the infection (Werren & O'Neill, 1997).

The *Drosophila-Spiroplasma* association is suited to study the long-term persistence of host-parasite relationships, because within this system of diversely related host-parasite pairs, the male-killing and non-male killing strains seem to bring different consequences for the persistence of the relationship.

Male-killing is believed to confer fitness advantage to the infected females through the reduction of inbreeding and the freeing of resources that would be utilized by males to be used by the all-female progeny (Werren & O'Neill, 1997). In addition to this, male-killing spiroplasmas show a high vertical transmission rate Xie, Butler, Sanchez, and Mateos (2014). So, a combination of high transmission rates and direct fitness benefits to the infected females contribute to the long-term persistence of male-killing spiroplasmas in *Drosophila* populations.

Contrary to male-killing spiroplasmas, the non-male-killing spiroplasmas reportedly do not produce direct fitness effects on the host (e.g., Kageyama et al., 2006; Osaka, Nomura, Watada, & Kageyama, 2008) and do not show perfect transmission (Osaka, Ichizono, Kageyama, Nomura, & Watada, 2013a). Under these conditions, the infection would disappear in the long term from the populations. For example, in a population in which all members can reproduce

without fail, starting with a 100% infection frequency and assuming a constant 92% transmission rate, at generation 30 the infection frequency in the population would be 8.2% (i.e., infection frequency, IF, at generation "30" is $IF_{30}=0.92^{30}=0.082$); this situation would lead to the loss of the infection in the long term. These predictions do not correspond to the observations that the infection has persisted for many years in the wild (e.g., Ota et al., 1979; Osaka et al., 2013a) and in laboratory populations (Watts, Haselkorn, Moran, & Markow, 2009; Osaka et al., 2013a).

Context-dependent fitness benefits of non-male-killing *Spiroplasma* infection on *Drosophila* (Jaenike, Unckless, Cockburn, Boelio, & Perlman, 2010; Xie, Vilchez, & Mateos, 2010; Xie, Winter, Winter, & Mateos, 2015) and intraspecific transmission of the infection by mites (Jaenike, Polak, Fiskin, Helou, & Minhas, 2007) would explain the persistence of *Spiroplasma* in the absence of perfect vertical transmission or direct fitness benefits to the host. However, these two phenomena have not been confirmed for a wide variety of these host-bacteria associations in the wild (e.g., Jaenike et al., 2007 on horizontal transfer and Xie et al., 2010; Xie et al., 2015 on protection against wasps), therefore, their importance in the maintenance of the infection in field populations needs still to be assessed. Anyway, neither of these two phenomena would explain the persistence of the infection in laboratory populations, where culture methods keep the flies free of natural enemies and mites. Non-male killing *Spiroplasma* infections have persisted in laboratory *Drosophila* populations for decades (Watts et al., 2009) and have been stably maintained in laboratory conditions for more than 50 generations (Osaka et al., 2013a).

This suggest that the non-male-killing infections have unidentified direct fitness benefits that promote the long-term persistence of the infection. This is likely because reports on the apparent

lack of direct fitness benefits to the host of the non-male-killing *Spiroplasma* infection have not been properly addressed. Kageyama et al. (2006) and Osaka et al. (2008) do not describe the experiments conducted that support their contention of no direct effect on fitness of the host by these spiroplasmas and cannot be authenticated.

To properly address the issue of the long-term persistence of the non-male killing *Spiroplasma* infection in *Drosphila* due to direct fitness benefits, *D. hydei*, a fly that is naturally infected by two closely related strains of *Spiroplasma* (Mateos el al., 2006; Haselkorn et al., 2009) without co-infection (Halselkorn et al., 2009) was used. The two haplotypes, named haplotype 1 (*Hyd1* hereafter) and haplotype 2 (*Hyd2* hereafter) belong to the clades poulsonii and citri, respectively. The poulsonii clade, to which *Hyd1* belongs, groups together male-killing and non-male killing *Drosophila*-infecting spiroplasmas, while the citri clade, to which *Hyd2* belongs only groups together non-male-killing spiroplasmas. The haplotype *Hyd1* was chosen to be used in these experiments because it is the most studied non-male-killing *Spiroplasma*, therefore, the results of this study can be compared to and discussed along with a growing number of reports from different areas of the world and render the conclusions of this study more generalizable. Furthermore, as indicated earlier, *Hyd1* is closely related to the male-killing strains of spiroplasmas and all are part of the poulsonii clade; by focusing on *Hyd1* the results of this work could help bridge studies on both-male killing and non-male killing phenotypes of the bacteria.

An exploratory study was first conducted in chapter 1 to achieve two general goals: 1) to identify characteristics of the infected flies that would allow for the persistence of the infection 2) to identify sources of variation in the experimental protocols so they can be optimized in follow up

experiments. There were four fitness related characters studied in chapter 1. 1) Ratio of larvae-to-emerged-flies; this character, which is an indicator of the survival from first instar larvae to emerged fly, was used to evaluate whether or not the infection produces any pathology at this interval of development, particularly late male-killing (Kageyama et al., 2007), when combined with the next character. 2) Sex ratio; this is an indicator of male-killing. 3) Longevity. 4) Fecundity. These two latter characters are part of a common life history trade-off (Agnew, Koella, & Michalakis, 2000). Chapter 2 applied the adjustments to the experimental protocols to confirm and re-examine findings of chapter 1 and to expand the exploration of fitness related characters to include 1) age at maturity and 2) preferential mating by all males to infected females.

The second topic addressed by this work is the potential host range of parasites. The potential host range of a parasite is the range of species that it could infect barring extrinsic barriers (De Vienne, Hood, & Giraud, 2009) as well as intrinsic barriers. The potential host range is a key determinant of, among other things, how likely parasites are to become established following their introduction into new areas (Taraschewski, 2006), a pressing issue in a constantly changing world (e.g., Dobson & Carper, 1992). The potential host range is a property of the parasites than can be evaluated under standardized conditions in laboratory experiments. A simple method to quantify the potential host range are the indexes of host specificity, a concept borrowed from the field of parasitology. These indexes score the number of species that a parasite can infect and also take into account how many individuals within each species are infected by the parasite (e.g., Rohde & Rohde, 2008). These scores allow the classification of the parasite on a continuum scale, from a score of 1, or specialist (for our purpose, the parasites having the

smallest potential host ranges) to a score of 0, or generalists (the parasites having the largest potential host range).

To explore the potential host range of parasites, the *Drosophila-Spiroplasma* association was used. As indicated earlier, this association is characterized by diversely related spiroplasmas infecting diversely related flies, which is a consequence of multiple introductions of *Spiroplasma* into *Drosophila* (Haselkorn et al., 2009). Consequently, the host-parasite trees for this associations are incongruent. (Haselkorn et al., 2009). This diverse host-parasite phylogenetic relationships allows for the exploration of several factors determining the potential host range of the parasites. The factors tested are: 1) host and parasite phylogeny (these phylogenies predict the potential host range of parasites, see De Vienne et al., 2009), and 2) host and parasite strains (i.e., intrinsic characteristics of the hosts and parasites).

The issues outlined above are of importance in the general frame of the evolution of the parasite-host relationship. Regarding the first issue, it is important to clarify if it is possible that with neither high transmission rates nor direct fitness benefits to the host, parasites can persist in the long-term, because this situation is not in line with the expectations from the theory. Regarding the second issue, the diverse phylogenetic relationships between spiroplasmas, between flies, and their incongruent host-parasite trees, create a particular situation in which the exploration of the role of host and parasite phylogeny in determining vertical transmission of the parasites is possible.

2. EXPLORING FITNESS CHARACTERISTICS OF *DROSOPHILA HYDEI* INFECTED WITH *SPIROPLASMA*

2.1 Introduction

This chapter addresses issues related to the long-term persistence of maternally heritable endosymbionts under conditions where there are no apparent fitness benefits to the host. These issues are explored by using the association of a non-male-killing *Spiroplasma* with the fruit fly *Drosophila*.

Spiroplasma (Procaryotae: Mollicutes: Entomoplasmatales: Spiroplasmataceae) are commonly associated with many organisms such as plants, invertebrates, such as the shrimp, honey bees, ticks, the fruit fly *Drosphila*; and vertebrates, such as humans (Regassa & Gasparich, 2006; Yu et al., 2009; Lorenz et al., 2002; Aquilino at al., 2015; Mueller et al., 2015). Spiroplasmas act as pathogens or mutualists, although their effect on many of the insect hosts is currently unknown (Regassa & Gasparich, 2006; Yu et al. 2009).

Eighteen strains of *Spiroplasma* infect at least 16 species of *Drosophila* (see Fig. 1 in Haselkorn, 2010a). An overt manifestation of the infection by a group of those *Drosphila*-infecting *spiroplasmas* is male-killing (Williamson & Poulson, 1979), whereby the male progeny is killed during embryogenesis or at the larval or pupal stages (Counce & Poulson, 1962; Kageyama et al., 2007). The other group of spiroplasmas does not kill male progeny of their *Drosophila* hosts (e.g., Ota et al., 1979). Both, male-killing and non-male-killing *Spiroplasma-Drosophila* associations are optional for the host (*Drosophila* can reproduce and survive without the

presence of *Spiroplasma*). The main mode of transmission of spiroplasmas in *Drosophila* is from mother to progeny through the germ plasm (vertical transmission). Routine laboratory maintenance of infected lines of *Drosophila* in labs around the world for many years has helped to consolidate this view.

A phylogenetic analysis using 16S rRNA of 12 *Drosophila*-infecting *Spiroplasma* strains shows that those spiroplasmas are separated in four clades (poulsonii, citri, tenebrosa and ixodetis) (Haselkorn et al., 2009). One clade, the poulsonii, includes male-killing and non-male-killing spiroplasmas, while the other three clades only include non-male-killing strains (Haselkorn et al., 2009).

The long-term persistence of maternally heritable endosymbionts can be achieved by perfect vertical transmission or the combination of several mechanisms, as described below. Perfect vertical transmission (i.e., 100% of the progeny inherits the infection) would, in theory, guarantee the persistence of any symbiont for many generations to come (i.e., infection frequency, IF, at generation "n" can be defined by the formula $IF_n=1.00^n$, where "1.00" is 100% vertical transmission and "n" is the number of generations). Endosymbionts without perfect vertical transmission will be lost from the host unless there is horizontal transfer or an increase in fitness of the host due to the infection (e.g., *Spiroplasma* protection against parasitic nematodes as reported for *D. neotestacea* by Jaenike et al., 2010.) Other mechanisms that seem to help with the persistence of endosymbionts with imperfect transmission are biasing the sex ratio of infected hosts towards females (e.g., male-killing), decreasing fitness of uninfected hosts, and biasing the sex ratio of uninfected hosts toward females (e.g., through cytoplasmic

incompatibility) (Werren & O'Neill, 1997). Additionally, environmental factors such as temperature may play a role in the persistence of endosymbionts in natural populations (Werren & O'Neill, 1997). For example, temperature seems to affect the transmission rates of *Spiroplasma* in laboratory experiments (Anbutsu, Goto, & Fukatsu, 2008; Osaka et al., 2008) and it seems to do so in field populations (Osaka, Watada, Kageyama, & Nomura, 2010).

Applying the mechanisms outlined above, the persistence of the male-killing strains of *Spiroplasma* in *Drosophila* populations can be explained by the fitness benefits conferred to the host through male-killing. The elimination of the male progeny contributes to the fitness of the infected flies because it frees resources to be used by the female-only progeny, as well as avoids inbreeding. In addition to this, because only female flies transmit the bacteria to their progenies, the death of males has no direct negative effect on the transmission of *Spiroplasma*.

Contrary to the male-killing strains, the non-male killing strains infecting *Drosophila* species do not distort the sex ratio of the host (Ota et al., 1979), do not show perfect (100%) vertical transmission (pers. obs.) and do not offer direct fitness benefits to their hosts (e.g., Kageyama et al., 2006; Osaka et al., 2008, referenced as unpublished data). Under these conditions, the infection would be lost in the long-term. For example, in a hypothetical population where all members are able to reproduce without fail, and it starts with an infection frequency of 100% (all population members are infected), with a constant 0.91 vertical transmission, and no fitness benefits to the host, the frequency of the infection at generation 50 would be about 0.9% (0.91^{50}=0.009). This low frequency will allow loss of the infection from hosts' populations due to random events (e.g., dying off of infected mothers, failure of transmission of the infection,

etc.). These predictions are at odds with the fact that the infection has persisted for many years in the wild (e.g., Ota et al., 1979; Osaka et al., 2013a) and in laboratory populations (Watts et al., 2009; Osaka et al., 2013a).

In the absence of perfect vertical transmission and direct fitness benefits to the host, horizontal transfer could play a role on the persistence of non-male killing bacteria in *Drosophila*. Jaenike et al. (2007) has experimentally shown that mites can transmit *Spiroplasma* between and within species of *Drosophila*; furthermore, Osaka et al. (2013b) have reported that mites collected from field-trapped *Drosphila* flies in Japan carried a *Spiroplasma* whose *p58* gene shared 100% identity with a 756 bp of the *p58* gene of the *Spiroplasma* carried by the fly. Horizontal transfer of *Drosophila*-infecting spiroplasmas by mites has not been evaluated in the field, but its contribution to the persistence of the infection could be negligible (Osaka et al., 2013b). Furthermore, transmission by mites would not explain the persistence of the infection in laboratory population of *Drosophila* carrying non-male-killing *Spiroplasma*, because laboratory fly cultures are kept mite-free. Carvalho and Da Cruz (1962) reported horizontal transfer through ingestion of the bacteria in a species of *Drosophila* but attempts to replicate this type of transmission have been unsuccessful (e.g., Williamson & Poulson, 1979; Ebbert,1991).

Another mechanism that could explain the persistence of the infection under no direct fitness benefits to the host is provision by the former of context-dependent fitness benefits to the latter. It has been shown that some spiroplasmas do confer protection against parasitic worms, as reported for *D. neotestacea* (Jaenike et al., 2010), or against parasitoid wasps in *D. hydei* (Xie et al., 2010), although the effect on *D. hydei* has not been shown to occur in the wild. The above

would explain the persistence of the infection in hosts populations exposed to those natural enemies (i.e., in the field), but not in laboratory populations, which are sheltered from such natural enemies.

The motivations for this work are fourfold. First, none of the explanations above that would allow *Spiroplasma* with non-perfect transmission to persist in the long-term can be applied to laboratory populations. Second, the direct fitness benefits for the host of non-male killing spiroplasmas have not been satisfactorily addressed; for example, Kageyama et al. (2006) and Osaka et al. (2008) do not describe the experiments conducted to reach their conclusions of no effect on fitness of the host. Third, the protection conferred by *Spiroplasma* against parasitoid wasps to *D. hydei* (Xie et al., 2010; Xie et al., 2015) has not been shown to occur in the wild. Fourth, the contribution of horizontal the transfer (Jaenike et al., 2007; Osaka et al., 2013b) to the long-term persistence of the infection has not been directly evaluated in the wild, but Osaka et al. (2013b) hints that the contribution of this factor may be negligible.

Because this is an exploratory study, it will not be limited to exploring positive fitness effects on the infection only, but also will investigate negative effects. This is because a negative effect of the infection on a given trait can prompt positive changes in another trait. For example, a shortening of the lifespan of the host due to the infection can result in an increase of its reproductive activity at earlier ages (for examples of these effects on *Daphnia magna* due to their parasites see Ebert, 2005).

To explore if any type of effects on fitness exist in the relationship of non-male killing spiroplasmas and their fly hosts, the association between *D. hydei* and *Spiroplasma* was used. *Drosophila hydei* is naturally infected by two non-male-killing *Spiroplasma* strains, one belonging to the poulsonii clade (haplotype 1), and the other to the citri clade (haplotype 2) (Mateos et al., 2006), without co-infection (Haselkorn et al., 2009). This work will focus on the *Spiroplasma* haplotype 1 (named hereafter *Hyd1*). This strain was chosen because it is widely distributed and has been reported to naturally infect *D. hydei* from Japan (Kageyama et al., 2006), Mexico (Mateos et al., 2006), and Mexico and the USA (Haselkorn et al., 2009). It follows that the results of this work can be compared to and discussed along with a growing number of reports from different areas of the world and this will render its conclusions more generalizable. Furthermore, as indicated earlier, *Spiroplasma Hyd1* is closely related to the male-killing strains of spiroplasmas as part of the poulsonii clade. This study focusing on *Hyd1* could help bridge studies on both-male killing and non-male killing phenotypes. Examples of some questions that can be addressed on this system are: Is *Spiroplasma Hyd1* a mild male-killing strain? If so, are the mechanisms of male-killing similar or different from the other male-killing strains? At what stage does male killing occur?

Based on the above, preliminary experiments were performed to explore if *Spiroplasma* haplotype 1 that infects *D. hydei* affects various fitness related characteristics of its host. The characteristics chosen to study were: 1) ratio of larvae-to-emerged-flies (i.e., survival from larvae to emergence, as an indicator of male-killing), 2) sex ratio (as another indicator of male-killing), 3) fecundity and 4) longevity; these two last characteristics are part of a common trade-off on life history in infected hosts (Agnew et al., 2000).

The ultimate goals of this exploratory section are twofold. The first one is to identify any fitness effects that show potential for promoting the persistence of *Spiroplasma* in *D. hydei*; the results will lead to perform confirmatory experiments and expand the study into related host characteristics that may act in conjunction with the newly identified ones that could promote the persistence of the bacteria. The second ultimate goal is to optimize the experimental protocols by identifying sources of variation that may obscure the effect of some of the characters measured; this will lead to the re-examination of some of the effects of the infection identified in this section and the examination of new characteristics by using improved experimental protocols.

2.2 Materials and Methods

2.2.1 Detection of Endosymbionts

We screened the experimental flies to confirm the presence (in infected flies) or absence (in uninfected flies and wild-caught females) of *Spiroplasma* and to confirm the absence of *Wolbachia* in wild-caught females. We screened wild-caught females because Mateos et al. (2006) reported that *Drosophila* flies are naturally infected by *Spiroplasma* and *Wolbachia*. Although Mateos et al. (2006) also reported that *D. hydei* is only infected by *Spiroplasma*, we share the opinion of some authors (Xie et al., 2010; Xie et al., 2015) that non-infection by *Wolbachia* in wild-caught females should be confirmed. *Wolbachia* affects the fitness of *D. melanogaster* (e.g., Fry et al., 2004), and if these bacteria were present in *D. hydei*, the fitness effects of *Spiroplasma* (the bacteria of interest in this study) could be obscured by *Wolbachia*.

To detect bacterial infection, we ran 12.5-μL PCR reactions using 1.5 μL of crude DNA extract as a template. The template was obtained from whole flies crushed in 50 μL of buffer (10 mM

Tris-HCl, 25 mM NaCl, 1 mM EDTA, pH 8) containing 2 µL of proteinase K (20 mg/mL). For *Spiroplasma* we used two primer sets, *p58* (Ye, Melcher, & Fletcher, 1997) ran under the conditions reported in Montenegro et al. (2005) and p58IV, designed by the author, which targets a fragment of about 360 bp that is a subset of the *p58* primer set target. The annealing temperature for p58IV is 53°C and the sequences are p58IV_F 5'-AAAGGTTTACATTCACCAAGTCG-3' and p58IV_R 5'-AATTGTTCATTAACTTTATCTTGTGG-3'. For *Wolbachia* we used the *wsp* primer set (Jeyaprakash & Hoy, 2000) under the conditions reported in Mateos et al. (2006). No infection by *Spiroplasma* or *Wolbachia* was detected in uninfected flies and wild-caught females.

2.2.2 Isofemale Lines (Isolines)

To maximize the amount of data available for analysis, we used isolines available in the laboratory and newly cultured isolines. This strategy to maximize analyzable data resulted in two different groups of flies: 1) flies that were naturally infected with *Spiroplasma* with their antibiotic-treated *Spiroplasma*-free counterparts as controls, and 2) flies artificially infected with *Spiroplasma* with their naturally *Spiroplasma*-free counterparts as controls. These two groups were considered "populations" (singular, "population") in the first step of the analyses.

2.2.2.1 "Population" 1 (naturally infected flies)

Infected flies were generated from two isolines collected in Mexico in 2004 (Mateos et al., 2006), SPE and TEN (which we will call SPE-1 and TEN-1), which were kept under laboratory conditions for about 5 years prior to the experiments. These isolines are naturally infected with the *Spiroplasma* haplotype 1 (hereafter, *Hyd1*), described by Mateos et al. (2006). These stock

flies were kept in half pint bottles during their more than 5 years of laboratory culture prior to being used in this study, but they were cultured in narrow vials during the experiments carried out in this work. We randomly chose ~100 females from the main culture bottles of each isoline and placed each group in separate pint milk bottles with banana-opuntia food and allowed them to reproduce for one generation. Next, we divided each group's new progeny into two halves. One half was cultured in regular banana-opuntia food and the other half was placed in food to which we added a combination of tetracycline and erythromycin (final concentration = 0.2 and 0.16 mg/ml, respectively) to cure the flies of the *Spiroplasma* infection. The progeny of each group of flies were kept in half pint bottles with their respective types of food (i.e., antibiotic added and antibiotic-free) for two generations. The antibiotic-cured flies and the flies that were not treated with antibiotic (i.e., flies that carry the *Spiroplasma* infection) were maintained separately in half pint bottles on regular food for one more generation, but we added a saline solution of crushed dead flies that were naturally *Spiroplasma*-free to the antibiotic-cured group to allow for recovery of the normal gut flora. At this point we had two treatment groups for the isolines SPE-1 and TEN-1: infected, and the antibiotic-cured group or control (uninfected). This generation was considered to be generation 0 or G0 for both groups. Biological repeats for the experiment were obtained by randomly choosing six infected females and three uninfected females per isoline and individually housing them in labelled separated vials. The experimental treatment applied to G1 and G2 and the experimental data collected at G2 are described in the section "Experimental Treatments and Measuring of Fitness Components" below. *Spiroplasma* infection status was assessed for uninfected and infected flies by PCR screening (as described in "Detection of Endosymbionts") and/or examination of hemolymph under dark field microscopy for subsets of mothers from G0, G1 and G2.

2.2.2.2 "Population" 2 (wild-caught females)

Seven field-mated *Drosophila hydei* female flies were trapped in College Station, TX, USA between April and November 2009 using banana baits. Each female was used to establish an independent isoline by allowing mating among its descendants only. These flies were cultured in narrow vials prior and during the experiments carried in this study. The originally collected females were PCR-screened for infection by *Spiroplasma* and *Wolbachia*. The screening confirmed the absence of these two bacteria.

After 12 generations in the lab, six randomly chosen females per isoline were artificially infected with the *Hyd1* strain TEN-1 from a Mexican *D. hydei* (Mateos et al., 2006). Infection was achieved by transferring hemolymph from adult infected females into adult uninfected females by microinjection using pulled glass microcapillaries. Each of these artificially infected females were housed in separated vials where they were allowed to recover for 48 hours. These six flies per isoline constituted the infected group at generation 0 (G0) and were later used to produce the next generation of infected flies. Parallel to the generation of the infected females, the G0 control or uninfected group was generated by randomly choosing 3 females per isoline and housing each of these females in separate vials. The six infected and three uninfected flies per isoline were considered biological repeats. The experimental treatment applied to G1 and G2 of infected and uninfected flies and the experimental data collected at G2 are described in the next section, "Experimental Treatments and Measuring of Fitness Components".

Spiroplasma infection status was assessed for uninfected and infected flies by PCR screening and/or examination of hemolymph under dark field microscopy for mothers from G0, G1 and G2.

2.2.3 Why and How Repeats Were Generated for Isolines Used in the Experiments

Laboratory experience with *Spiroplasma* infected *Drosophila* has shown us that several factors interact to reduce the number of experimental flies per isoline that are left to analyze at the end of an experiment. We identified 5 potential causes of or for? this, namely 1) some reproductive females or their progeny may produce a small number of progeny that will die off for reasons beyond the experimenters' control (this may be caused by the old age or the bad health status of the female in question); 2) some of the randomly chosen infected mothers necessary to produce the next generation may not carry the infection due to imperfect vertical transmission; 3) the progeny of an infected mother may screen as negative for the infection; 4) the randomly chosen infected and uninfected mothers may not produce eggs in the time frame required for the flow of the experiment or may fail to produce eggs at all; and 5) staff may not be available for handling the flies in due time for the flow of the experiment.

We aimed to maximize the number of flies per isoline available for analysis at the end of the experiment. To achieve this, we inbred a large number of individuals from the progeny of each founder female, expecting to lose many of those flies in the course of the breeding process due to the above-mentioned factors.

For our experiment we started with nine isolines at G0, each of them with an infected and an uninfected group. The infected group for each isoline started with six biological repeats and the uninfected group with 3. As explained in the sections describing “Population” 1 and “Population” 2 above, the repeats were generated by randomly chosen adult females and placing each of them in a separated individually labelled vial. We attempted to keep all biological repeats up to G2, with various degrees of success for each isoline (see Table S1.1 and Table S1.2 in Supplementary Material (SM) S1). G2 was the generation at which measurements for several fitness related characteristics were taken for the available biological repeats within each isoline. To measure one of these fitness related characteristics, fecundity, we treated the G2 biological repeats differently than for the other characteristics, always aiming to increase our final sample size. Therefore, for fecundity, we set up 3 vials per biological repeat for uninfected females and 3 for infected females. The degree of success at increasing the sample size for fecundity analysis varied among isolines. SM S1 summarizes the number of isolines, biological repeats and vials remaining at the end of the experiment for the analysis of the various fitness related characteristics measured in these exploratory experiments.

2.2.4 Experimental Treatments and Measuring of Fitness Related Components

Throughout this study, flies were cultured at 25°C under a 12h: 12h light/dark regime and fed on banana-opuntia food. Flies were kept in narrow vials.

At G0, for all the starting biological repeats per isoline, adult females were mated to mature (older than 10 days) uninfected males from their isolines. Mated females were allowed to oviposit on fresh media (changed every 24 h) over five days; after this period females were

frozen at -80°C for later PCR screening. To control for density, thirty first-instar larvae were collected from the eggs laid by each G0 female and transferred to a fresh vial to produce the G1 flies. The emerging male and female G1 flies were collected every 24 hours and kept in separate vials sorted by emergence day and sex. The G1 females flies were set up at ages 6 to 10 days old and were treated as the G0 flies to produce the G2 flies. All fitness related measures were taken at G2.

The larvae-to-emerged-flies survival rate was measured as the number of emerging flies per vial divided by the number of larvae placed into each vial. Sex ratio was estimated as the number of emerged males divided by the total number of emerged flies (measured for each vial).

To measure early fecundity at G2, we had to treat these flies in a different manner from the ones used to measure larvae survival and sex ratio. For these females we set up 3 vials per biological repeat for uninfected females and 3 vials per biological repeat for infected females. Each of these vials contained a 4-6 day old randomly chosen virgin female mated to 1-2 uninfected mature males (about 10 days old) from their respective isoline. The females were allowed to oviposit over five days on fresh media (replaced every 24 h), and the number of eggs laid per day was recorded over this period. Because female size is known to correlate with fecundity (Tantawy & Vetukhiv, 1960), we measured thorax length as the length of the scutellum along the midline from a dorsal view to assess its use as a potential covariate. Thorax length was measured with an ocular micrometer (1 mm scale) with a precision of 2 decimals at the maximum magnification of the microscope used (45X). The age at which the females were entered in the experiment was recoded (hereafter “Age at Setup”). Those ages were chosen to ensure maturity. Markow (1985)

reported that 85.7% of *D. hydei* males reach maturity at 10 days old, and 91.3% of females do so at 4 days old.

To assess longevity for experimental females, a subset of randomly chosen infected and uninfected G2 females that were originally used to measure early fecundity were followed until their death (49 infected and 45 uninfected flies). Females to be scored for longevity were transferred individually to a fresh vial after removal of males on day 6 post pairing. During this period, female flies were transferred to fresh vials every 2–3 days, but vials were examined daily for dead females and date and age of death were recorded.

All females used for measuring these fitness components (larvae-to-emerged-fly survival, sex ratio, early fecundity and longevity) were collected for subsequent PCR screening of infection status and all males were discarded. For the group of infected flies, only females that screened positive for *Spiroplasma* (some flies from the infected group may not carry the infection due to the imperfect transmission of the bacteria) were used in the analyses of fitness related effects of infection and those screening negative were discarded from the analysis. All uninfected females were used in the analyses of fitness effects.

For a diagram of the flow of the steps for the treatments applied from G0 to G2 in these exploratory experiments, see Fig. S2 in SM S2.

2.2.5 Statistical Analysis

To maximize the data available for analysis, we experimented on all isolines available in the laboratory and the ones generated from wild-caught flies. Thus, for all analyses of fitness effects of infection on the fly hosts, we had data originating from two "populations" of isolines (described in the "Isofemale Lines" section). To assess if all data could be analyzed as originating from a single "population", we tested if this variable had an effect on the dependent variable (if the "populations" had no effect on the dependent variable, then we could group the two "populations" into a single group for subsequent analyses). To this end, for each fitness measure addressed in this work, we ran one-way ANOVA analysis by using PROC ANOVA as implemented in SAS Enterprise Guide® 7.1. (SAS EG® 7.1 hereafter) (SAS Institute Inc., 2014). For these analyses, we used the fitness measure as the dependent variable and "population" as the independent factor.

2.2.5.1 Statistical analysis for larvae-to-emerged-flies survival ratio

To test if the different "populations" had an effect on the larvae-to-emerged-flies ratio, we transformed this response variable into the arc sin square root values and used "population" as the independent factor in a preliminary ANOVA analysis (see "Statistical Analysis" section).

Next, we tested for difference in the odds of flies emerging from collected larvae due to Infection Status. The larvae-to-emerged-flies ratio was analyzed by fitting a generalized linear mixed model to the data by using the procedure for fitting Generalized Linear Mixed Models implemented in SAS 9.4® (PROC GLIMMIX hereafter); the PROC GLIMMIX code was run in SAS EG 7.1® (SAS Institute Inc., 2014). The PROC GLIMMIX was run with the response

(larvae-to-emerged-flies) entered in the form of "events/trials" (number of flies emerged/number of larvae in vial). The variable of interest, Infection Status, which could be either Infected (females that screened positive for *Spiroplasma*) or Uninfected (control females), was entered as a fixed factor in the model. Random variables in the model were Isoline, Repeat nested within Isoline, and the interaction "Infection Status*Isoline" (see Fig. S3 in SM S3 showing our understanding of what constitute the factors, (biological) repeats and units of analysis for this experiment). We specified a binomial distribution of the response with a logit link function. The odds ratio estimates table was generated by requesting it in the model statement. All other options were set at default. Extreme observations (outliers and influential values) were detected by requesting conditional residuals plots. The goodness of fit of the model was assessed by using the General Chi-Square divided by its degrees of freedom (a value of 1 indicates good fit).

2.2.5.2 Statistical analysis of sex ratio

To test if "population" had an effect on sex ratio, we transformed this response variable into the arc sine square root values and used "population" as the independent factor in a preliminary ANOVA analysis.

Two tests were used to assess whether or not *Spiroplasma* kills male progeny and lowers the sex ratio in *D. hydei*. First, we tested if the sex ratio of infected females was different from the sex ratio of uninfected females; second, we tested if both groups of females produce sex ratios different from 50%. This was done because the sex ratio can vary due to several factors (e.g., genetic, environmental, experimental conditions); even though infected and uninfected females may produce different sex ratios, both sex ratios could be biased due to the factors mentioned

above, but not due to male-killing. If male-killing would be the cause for sex ratio differences between infected and uninfected females, then we expected to see the following results: 1) the first test would show a significant difference on the sex ratio between the two groups of females, and 2) the sex ratio of infected females would be significantly lower than 50% while the uninfected females would produce a 50% sex ratio. Assuming that the uninfected females sex ratio (control group) should be 50% is justified under the expectation that the sex ratio tends to the 1:1 equilibrium ratio (Hamilton, 1967).

The sex ratios of infected and uninfected females were compared by fitting a generalized linear mixed model to the data with PROC GLIMMIX ran in SAS EG 7.1® (SAS Institute Inc., 2014). The PROC GLIMMIX was run with Sex Ratio (response) entered in the form of "events/trials" (number of male progeny/total number of emerged flies); the variable of interest, Infection Status, was entered as a fixed factor in the model. Random variables entered in the model were the following: Isoline, Repeat nested within Isoline, and the interaction "Infection Status*Isoline". We specified a binomial distribution of the response with a logit link function. The link function specifies a nonlinear transformation between the linear predictor for Y and the assumed distribution function. We generated the odds ratio estimates table by requesting it in the model statement. All other options were set at default. The goodness of fit of the model was assessed using the General Chi-Square divided by its degrees of freedom. Extreme observations were detected by requesting conditional residuals plots.

We tested if infected females produced a sex ratio different from 50% by running an Exact Binomial Test by using R version 3.1.0 (R Core Team, 2013). Uninfected flies were tested separately for 50% sex ratio for control purposes by using the same test.

2.2.5.3 Statistical analysis of longevity (survival) analysis

To test if "population" had an effect on longevity, we used this variable as the response and "population" as the independent factor in a preliminary ANOVA analysis.

The analysis was conducted by using the SAS procedure designed to compute nonparametric estimates of the survival distribution function, PROC LIFETEST. The PROC LIFETEST was ran as implemented in SAS EG 7.1® (SAS Institute Inc., 2014) to obtain the survival function for infected and uninfected females as estimated by the Product-limit (Kaplan-Meier) method (Kaplan & Meier, 1958). The PROC LIFETEST uses three tests to determine if there are significant differences in the survival distributions for the control and experimental groups and we report here the results of two of them: the log-rank test and the Wilcoxon test.

All isolines were pooled together for this analysis. Twenty-one females were part of "population" 1 (10 uninfected and 11 infected flies) and 73 were part of "population" 2 (35 uninfected and 38 infected flies).

2.2.5.4 Statistical analysis for fecundity

To test if "population" had an effect on fecundity we used the sum of the 5-day oviposition counts per each female as the response variable and "population" as the independent factor in a preliminary ANOVA analysis.

Because larger females live longer and produce more eggs (Tantawy & Vetukhiv, 1960), we wanted to use longevity and female size (continuous variables) as covariates added to the above model to run an ANCOVA as defined by Wildt and Ahtola (1978). To evaluate the potential covariate "size", we first tested if "population" had an effect on it by running a one-way ANOVA using the PROC ANOVA of SAS as implemented in SAS EG 7.1® (SAS Institute Inc., 2014). We used Female Size as the dependent variable and "population" as a factor. Because "population" had a significant effect on size, we tested next if size was linearly related to the dependent variable (fecundity) at each level of the independent variable ("uninfected" and "infected" groups of the variable Infection Status) (Wildt & Ahtola, 1978) for each "population". The other potential covariate, longevity, was affected by the treatment as shown by the survival analysis performed earlier; therefore, it did violate the recommendation for the use of a covariate in an ANCOVA, which states that the covariate should be independent of the experimental treatments (Huitema, 2011, pp. 209-213; Wildt & Ahtola, 1978, p. 19).

We ran a linear mixed effects model by using the SAS procedure for mixed linear models as implemented in SAS EG 7.1® (SAS Institute Inc., 2014) (hereafter PROC MIXED) to analyze fecundity. We used the variable of interest, Infection Status, as well as Age at Setup (see methods section) and the interaction between these two variables, as fixed factors. We removed

nonsignificant interactions from the model. “Age at Setup” was entered in the model as a categorical variable. We entered three random factors in the model, namely Isoline, Repeats nested within Isoline, and the interaction Infection Status*Isoline. We analyzed the data in the same way as West, Welch, and Galecki (2014).

Because longevity and fecundity are part of a common life history trade-off (Agnew et al., 2000) and the survival analysis indicated that longevity is affected by Infection Status, we reran the model above including longevity as a continuous variable, to assess if there was an interaction between longevity and Infection Status.

To check the assumptions for the final model we looked for 1) the normal distribution of the residuals, 2) constant residual variances across observations, and 3) presence/absence of outliers and influential values. We used the conditional raw residuals and the conditional studentized residuals to test for the normal distribution and constant variance of the residuals (West et al., 2014). We obtained a panel of residual graphics by using the option “plots = (residualpanel boxplot influencestatpanel)” of PROC MIXED. We also obtained tables with the residuals by using the option “residual outpred =”. We used those plots and residual tables to visually assess normality and constant variance for uninfected and infected flies’ fecundity and the random factors in the model. If non-normality was suspected, a Shapiro-Wilk test for normality was run to confirm the visual assessment.

Because likelihood-based methods, such as restricted maximum likelihood (REML), that were used in this analysis are sensitive to unusual observations, we used influence diagnostics to

identify observations that may heavily influence the parameter estimates of fixed and random effects in the model (West et al., 2014). The goal of influence diagnostics for a given observation (or subset of observations) is to quantify the effect of the omission of those observations on the results of the analysis. We assessed the influence of individual observations by using the "influence" option and its sub-option "iter=" in SAS PROC MIXED. These options produce graphics and an influence diagnostics table that allow visualization of the observations that stand out from the rest of the data points, enabling a further examination of such observations as possible influential points. Two types of diagnostics can be performed using these graphic and table outputs: "global" and "specific" diagnostics. Both types of outputs allow for "global" diagnostics that assess the influence of individual observations on the overall fit of the model, on the fixed factors as a whole, and on the covariance parameters (random factors) as a whole. The graphics output allows for "specific" diagnostics addressing influences for each of the fixed effect estimates and for each of the covariance parameter estimates in the model.

The "global" measures selected for the diagnostic (after West et al., 2014) are the following: the "Restricted Likelihood Distance" (a measure of influence on the overall fit of the model), "Cook's D" and "Cook's D Covariance Parameters" (a measure of influence on the point estimate of parameter estimates for fixed and random covariance parameters, respectively) and "Covratio" and "Covratio Covariance Parameters" (a measure of influence on the precision of the parameter estimates for fixed and random covariance parameters, respectively).

The "specific" diagnostic allowed us to explore the influences of individual observations on the fixed factors "intercept" and "Infection Status" as well as on the covariance parameters "Isoline",

"Repeat(Isoline)" and "Residual". Two ways to deal with influential values is to explore if the model can be improved by addition of other variables or by running the model without the influential values (Littell, Milliken, Stroup, Wolfinger, & Schabenberger 2006; West et al., 2014).

2.2.5.5 Repeated measures analysis of age at setup and daily fecundity

Because we set up females aged from 4 to 6 days old, we wanted to test if age of flies affected the observed pattern of daily fecundity. This was done because upon completion of the experiments we realized that we had accidentally set up different sample sizes of infected and uninfected females at ages 4, 5 and 6 (Fig. 2.1). We reasoned that this may adversely affect the daily fecundity patterns of the experimental groups because it is known that the percentage of mature (able to lay eggs) *D. hydei* females increases with age starting from 2 days old (Markow, 1985).

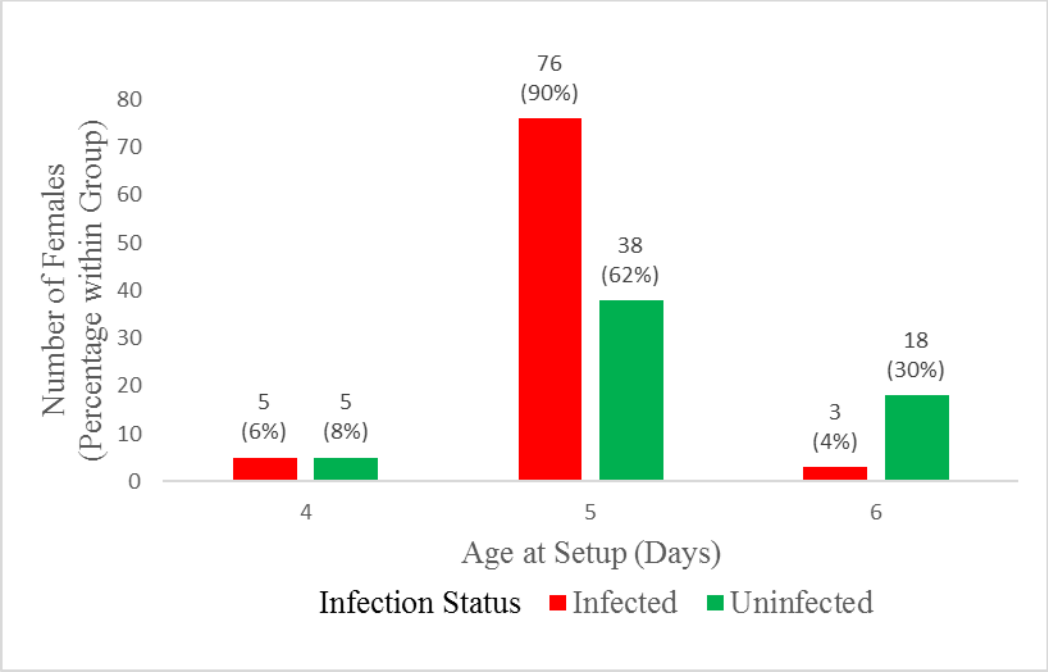

Fig. 2.1 Age at setup for uninfected and infected groups for the fecundity experiment.

We used repeated measures analysis to evaluate the patterns of oviposition for infected and uninfected females during the 5 days in which we recorded fecundity. We ran PROC GLIMMIX in SAS EG 7.1® (SAS Institute Inc., 2014) to perform this analysis. The 5 daily recorded numbers of eggs of each female served as the dependent variable. There were three fixed factors entered in the model: 1) Infection Status (with two values, infected or uninfected), 2) Age at Setup (age at which females were entered in the experiment), which was entered in the model as a categorical factor with values of 4, 5 or 6 days old, and 3) Time (time of oviposition, representing the days when egg were counted, was entered as a categorical factor with values of 1, 2, 3, 4, and 5). Isoline and the Repeats of females nested within isoline were entered as random factors. For the fixed factors, we entered all possible interactions and single factor effects in the analysis. We indicated that the fecundity measures were taken multiple times on

individual flies in the SAS code. The goodness of fit was assessed by using the General Chi-Square divided by its degrees of freedom. Extreme observations were detected by requesting conditional residuals plots.

2.3 Results

2.3.1 Descriptive Parameters for Larvae-to-Emerged-Flies Survival Ratio

The progeny of 57 females from nine isolines were available for analysis (21 uninfected females and 36 infected females). The mean ratio of larvae-to-emerged-flies was 0.84±0.14 standard deviation for infected females and 0.76±0.18 standard deviation for uninfected females (Table 2.1). The minimum and maximum ratios of larvae-to-emerged flies were 0.47 to 1.00 for infected females and 0.20 to 0.97 for uninfected females.

Table 2.1 Ratio of larvae-to-emerged-flies survival for infected and uninfected flies.

Infection Status	Number of Observations	Mean	Standard Deviation	Minimum Ratio of Emerged	Maximum Ratio of Emerged
Infected	36	0.8417	0.1435	0.4667	1
Uninfected	21	0.7635	0.1847	0.02	0.9667

2.3.2 Results of the Statistical Analysis for Larvae-to-Emerged-Flies Survival Ratio

"Population" had no effect on the ratio of the larvae-to-emerged-flies (p=0.1099), so the two groups of flies were analyzed together. The diagnostic for the PROC GLIMMIX based on the General Chi-Square divided by its degrees of freedom (value of 3.12) indicated a low level of

overdispersion. Carruthers, Lewis, McCue, and Westley (2008) considers a value of 3 to be typical once one has found an adequate model structure, and sophisticated modeling of overdispersion may well be unnecessary at these low levels.

The examination of conditional residuals showed that two "number of events / number of trials" sets (1 uninfected and 1 infected) could be considered outliers. Upon removal of these rows, the conclusions of this outlier-free analysis were similar to the full data set's conclusions. The diagnostic based on the General Chi-Square divided by its degrees of freedom (value 2.96) indicated a level of over-dispersion that seemed trivial for the reduced data. Carruthers et al. (2008) lists the presence of outliers as one of the causes for overdispersion. This seems to be the case with the current data, where the low level overdispersion is reduced by the removal of outliers. Since all observations were taken with the utmost care and are to our knowledge free of errors of measurement, we believe that the observations must be kept in the analysis. Therefore, we are reporting the results from the full data analysis.

The analysis of the full data indicates that Infection Status does not have a significant effect on the number of flies emerged from the larvae collected in the experiment (p=0.3859) (Table 2.2). This means that the odds of a fly emerging from the uninfected larvae collected in this experiment are not different from the odds of a fly emerging from the infected larvae collected for this experiment (i.e., the odds ratio varies between 0.337 and 1.599, an interval that includes 1) (Table 2.3). The covariance parameter estimates (random factors) for the interaction Infection Status*Isoline were significant (p=0.0222). (Table 2.4).

Table 2.2 Fixed effects significance test for the generalized linear mixed model predicting the likelihood of a fly emerging from a collected larva based on infection status of flies (n=57).

Fixed Effect	Numerator Degrees of Freedom	Denominator Degrees of Freedom	F Value	Pr > F
Infection Status (uninfected & infected)	1	8	0.84	0.3859

Table 2.3 Generalized linear mixed model predicting the likelihood of a fly emerging from a collected larva based on infection status of flies (n=57). Odds ratio estimates are shown.

Infection Status Odds Numerator / Odds Denominator	Odds Ratio Estimate	Degrees of Freedom	95% Confidence Limits for Odds Ratio Estimates
Uninfected / Infected	0.734	8	0.337 - 1.599

Table 2.4 Covariance parameter estimates for the generalized linear mixed model predicting the likelihood of a fly emerging from a collected larva based on infection status of flies (n=57).

Covariance Parameter	Estimate	Standard Error	Z Value	Pr > Z
Isoline	0	0.2447	0	0.5
Infection Status*Isoline	0.4053	0.2017	2.01	0.0222
Repeat(Isoline)	0.01745	0	.	.

To explore the interaction Infection Status*Isoline, we ran an additional PROC GLIMMIX (SAS Institute Inc., 2014) with Infection Status, Isolines and the interaction Infection Status*Isoline as fixed factors and Repeat(Isoline) as a random factor. The random factor for this new analysis was removed because its covariance parameters estimates were zero. The fixed effects

interaction Infection Status*Isoline was significant (p=0.0081), as it is Isoline (p=0.0112), but our variable of interest, Infection Status, is not significant (p=0.5189) (Table 2.5).

Table 2.5 Fixed effects significance test for the generalized linear mixed model predicting the likelihood of a fly emerging from a collected larva based on infection status, isoline and their interaction (n=57).

Effect	Numerator Degrees of Freedom	Denominator Degrees of Freedom	F Value	Pr > F
Infection Status	1	39	0.42	0.5189
Isoline	8	39	2.95	0.0112
Infection Status*Isoline	8	39	3.12	0.0081

Fig. 2.2 shows the Infection Status*Isoline interaction obtained from running the new PROC GLIMMIX for a model where Infection Status, Isolines, and the interaction of these two factors were set as fixed factors, with no random factors in the model. Contrasts were used to compare the likelihood of a fly emerging from pupae in infected vs. uninfected flies for each isoline. Table 2.6 shows only the significant contrasts from the analysis. Notice that for isolines SEVENTEEN and TWENTY, the negative estimates indicate that a fly has lower odds of emerging from an uninfected larva than from an infected one (Fig. 2.2).

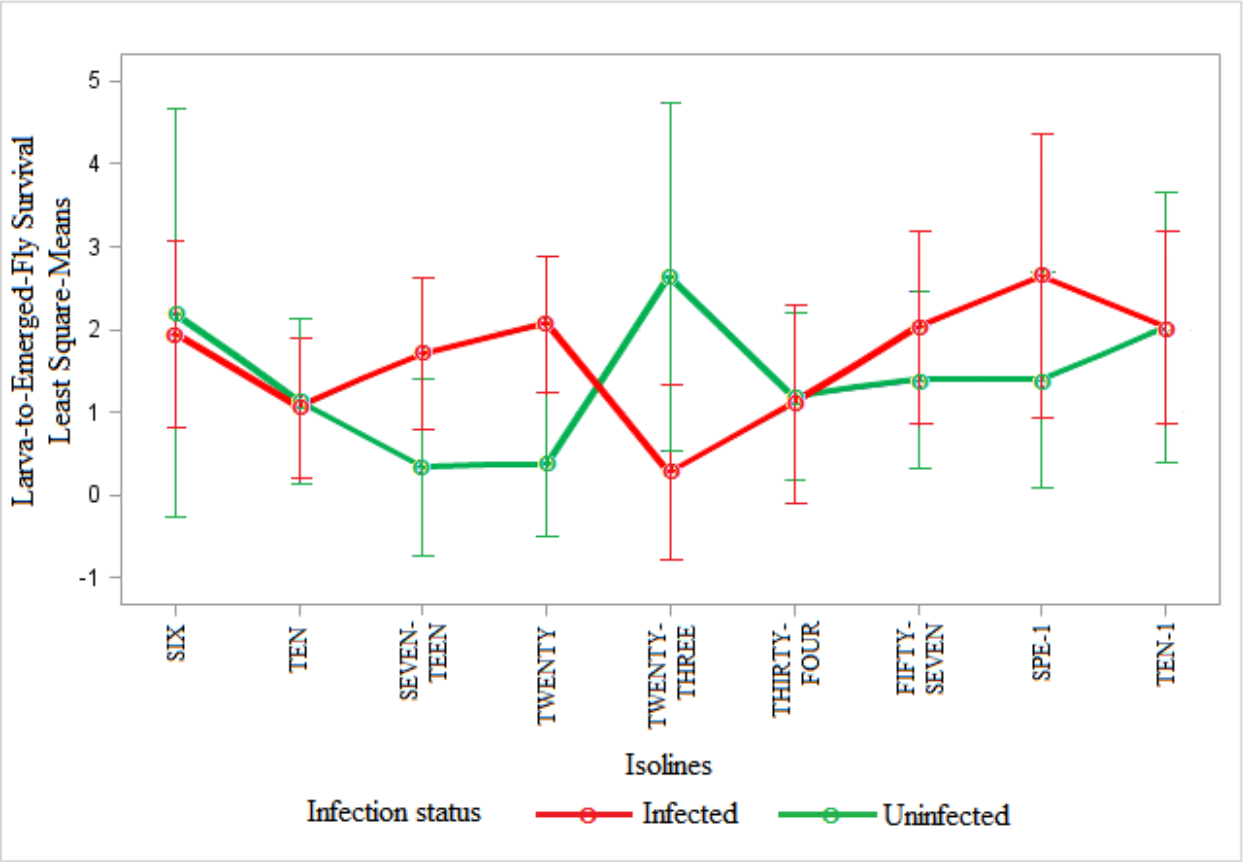

Fig. 2.2 Interaction of fixed factors infection status*isoline for a generalized linear mixed model predicting the likelihood of a fly emerging from a collected larva (n=57) (the vertical bars represent the 95% confidence intervals for the estimates).

Table 2.6 Significant contrasts for infected females vs uninfected females from 4 isolines from an analysis of fixed effects for a generalized linear mixed model predicting the likelihood of a fly emerging from a pupa based on infection status, isoline and their interaction (n=57).

Effect	Contrast Label	Estimate	Standard Error	Degrees of Freedom	t Value	*p* value
Infection Status*Isoline	Infected from Isoline SEVENTEEN vs. Uninfected from Isoline SEVENTEEN	-1.3722	0.6941	39	-1.98	0.0551
	Infected from Isoline TWENTY vs. Uninfected from Isoline TWENTY	-1.7061	0.5933	39	-2.88	0.0065
	Infected from Isoline TWENTYTHREE vs. Uninfected from Isoline TWENTYTHREE	2.3708	1.1613	39	2.04	0.048

2.3.3 Descriptive Parameters for Sex Ratio

The progeny of 57 females from nine isolines were available for analysis (21 uninfected and 36 infected females). The mean ratio of male-to-total emerged flies (sex ratio) was 0.51±0.09 standard deviation for infected females and 0.59±0.11 standard deviation for uninfected females (Table 2.7). The minimum and maximum sex ratio was 0.33 to 0.7143 for infected females and 0.33 to 0.8400 for infected females.

Table 2.7 Ratio of male-to-total emerged flies (sex ratio) for infected and uninfected flies.

Infection Status	Number of Observations	Mean	Standard Deviation	Minimum Sex Ratio	Maximum Sex Ratio
Uninfected	21	0.58804	0.114787	0.333333	0.84
Infected	36	0.51382	0.091392	0.333333	0.7143

2.3.4 Results from Statistical Analysis of Sex Ratio

"Population" had no effect on the sex ratio (p=0.8078), so the two groups of flies were analyzed together. For the PROC GLIMMIX, the covariance parameter estimates for all random factors in the model were zero. This indicates that all the random factors could be removed from the model (SAS Institute Inc., 2010). We removed the random factors and ran a model with the fixed factor "Infection Status" only. For the later analysis, the residual graphics showed an extreme value (uninfected set number 7). We re-ran the model after removing the outlier, and the general conclusion of the analysis was not changed; therefore, we are reporting the results for the analysis for the full data. The analysis shows that Infection Status has a significant effect on sex ratio, $F_{(1, 55)} = 9.29$, $p = 0$ .0035. The odds of birthing a male are 1.127 to 1.780 times higher for uninfected females than for the infected ones (Table 2.8). The diagnostic based on the General Chi-Square divided by its degrees of freedom (value 0.92) indicated a level of under-dispersion that seems to be trivial.

Table 2.8 Generalized linear mixed model predicting the likelihood of birthing a male based on infection status of flies (n=57). Odds ratio estimates are shown.

Infection Status Odds Numerator / Odds Denominator	Odds Ratio Estimate	Degrees of Freedom	95% Confidence Limits for Odds Ratio Estimates
Uninfected / Infected	1.377	55	1.127 - 1.780

The results of an exact binomial test for the sex ratio of infected females shows that they produce a ratio of males no different from 50% (p = 0.6190), and another exact binomial test found that the uninfected females produce a sex ratio significantly different from 50% (p=0.00003), which is 0.59 or 59% (Table 2.9).

Table 2.9 Exact binomial test for the null hypothesis that the total number of males born to all uninfected or infected fly mothers is 50%.

Infection Status	Number of Mothers	Total Number of Male Progeny (Sex Ratio)	Total Number of Female Progeny	p value	Ho: Sex Ratio = 50%
Infected	36	463 (0.51)	447	0.619	Not rejected
Uninfected	21	286 (0.59)	195	0.00003	Rejected

2.3.5 Descriptive Parameters for Longevity

For ”population” 1, 21 females were available for analysis (11 infected and 10 uninfected ones); the mean longevity was 59.82 days for infected females and 71.60 days for uninfected flies (Table 2.10). Infected flies lived a minimum of 45 days and a maximum of 89 days; while uninfected females lived a minimum of 48 days and a maximum of 100 days (Table 2.10).

Table 2.10 Descriptive statistics for longevity for ”population” 1.

Infection Status	Number of Females	Mean	Standard Deviation	Minimum Longevity	Maximum Longevity
Infected	11	59.82	13.75	45	89
Uninfected	10	71.6	18.8	48	100

For ”population” 2, 73 females were available for analysis (38 infected and 35 uninfected ones); the mean longevity was 54.74 days for infected females and 60.83 days for uninfected flies (Table 2.11). Infected flies lived a minimum of 41 days and a maximum of 72 days, while uninfected females lived a minimum of 48 days and a maximum of 88 days (Table 2.11).

Table 2.11 Descriptive statistics for longevity for ”population” 2.

Infection Status	Number of Females	Mean	Standard Deviation	Minimum Longevity	Maximum Longevity
Infected	38	54.74	8.33	41	72
Uninfected	35	60.83	10.54	48	88

2.3.6 Results of Statistical Analysis for Survival Analysis (Longevity)

No censored records were present in the dataset for both “populations” (longevity was known for 100% of the individuals studied). For each “population” described in the methods section, to the best of our knowledge, no external factor other than the infection status could have affected the survival, because all flies were studied at the same time of the year and all of them were treated the same way; so no known source of bias was introduced at this step of the experiment for each “population”.

The one-way ANOVA indicated that “population” had an effect on survival. Flies from ”population” 1 lived longer than their ”population” 2 counterparts (see Tables 2.10 and 2.11 for the descriptive statistics). Likely causes of these lifespan differences between “populations” include differences in the length of time that the two “populations” have been cultured in the lab as well as the method of laboratory culture (Sgro & Partridge, 2000). This research was not designed to answer this question, so it will not be addressed any further.

Because the longevity of the two “populations” was different, they were analyzed separately. For ”population” 1, the survival distributions for the treatments (uninfected and infected flies) were not statistically significantly different, $\chi^2(1) = 2.3216$, $p > .1276$ (log-rank test) and $\chi^2(1) = 1.2943$, $p > .2553$ (Wilcoxon test). For ”population” 2, these distributions were statistically significantly different: $\chi^2(1) = 7.0477$, $p > .0079$ (log-rank test) and $\chi^2(1) = 5.5983$, $p > .0180$ (Wilcoxon test).

A closer examination of the data for ”population” 1 shows that there are three observations that may be the cause of the non-significance of the different survival distributions for infected and uninfected flies. Three uninfected flies died before the age of 52 days: 1 at 48 and 2 at 49. In contrast, among infected flies only one died at 45 and a second died at 52 days of age. Removing the three uninfected flies that died very early results in significantly different survival distributions for uninfected and infected females for “population” 1, and uninfected flies live longer (p=0.01 for the log-rank test and p=0.007 for the Wilcoxon test). Furthermore, the removal of the two uninfected flies dying at age 49 produces a weak (albeit statistically significant) result (p=0.03 for the log-rank test and p=0.03 for the Wilcoxon test). This suggested to us that the non-significant result for the data analysis for ”population” 1 is caused by the dispersion that those uninfected fly death observations add to the estimates for this treatment group and the small sample size of ”population” 1, which diminishes the power of the test to detect the statistical differences. Because we wanted to maximize the data available for analysis and given that both “populations” showed a common trend (uninfected flies lived longer than infected ones, Table 2.10 and Table 2.11), we decided to combine the data for the two “populations” for the survival analysis. For the full data set, 45 uninfected females and 49 infected females were available for analysis (Table 2.12).

Table 2.12 Descriptive statistics for longevity for the full data set.

Infection Status	Number of Females	Mean	Standard Deviation	Minimum Longevity	Maximum Longevity
Infected	49	55.88	9.87	41	89
Uninfected	45	63.22	13.33	48	100

The survival analysis shows that uninfected flies live an average of 63.22 (95% CI, 59.33 to 67.12 days), which is longer than the average lifespan for infected flies (55.88, 95% CI, 53.11 to 58.64 days). The survival distributions for the two treatments are shown in Fig. 2.3. These distributions were statistically significantly different, $\chi^2(1) = 8.3412$, $p < .004$ (log-rank test) and $\chi^2(1) = 7.6211$, $p < .006$ (Wilcoxon test).

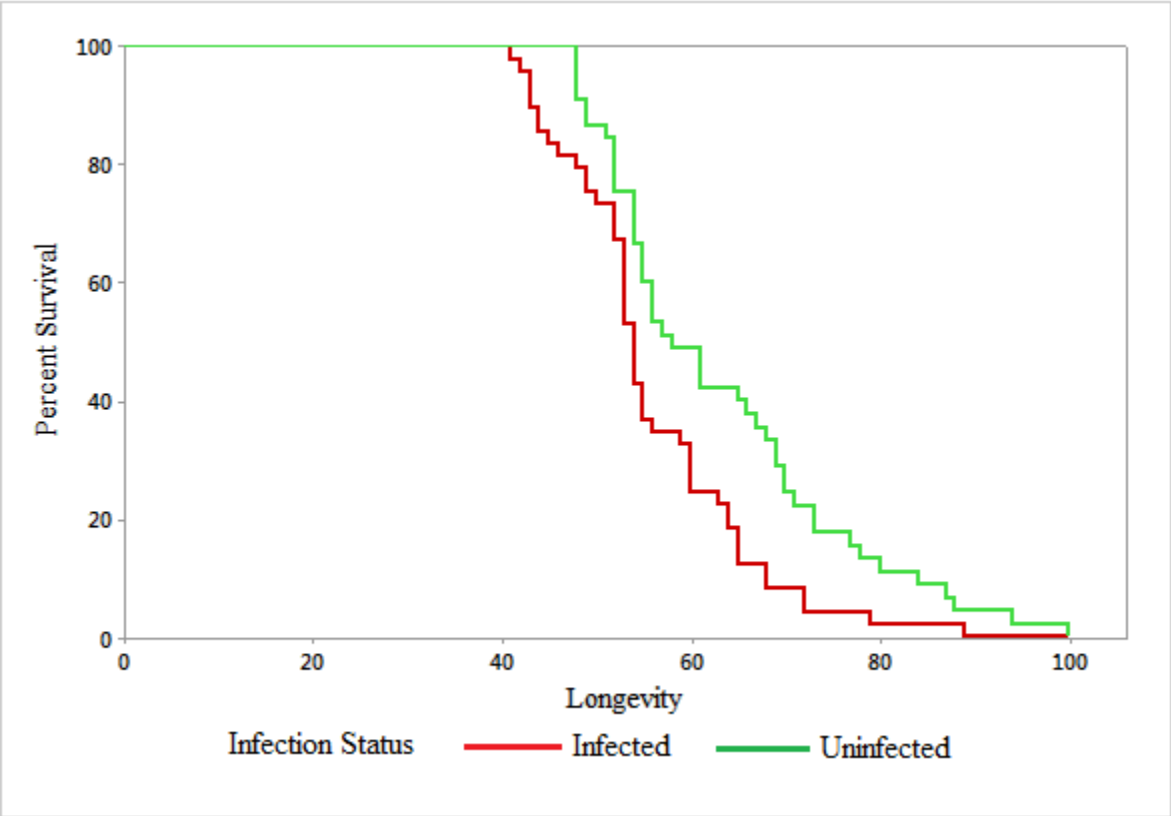

Fig. 2.3 Survival of *Hyd1 Spiroplasma*-infected *D. hydei* females is lower than the survival of uninfected females (number of females tested: uninfected=45, infected=49; $p < .004$).

2.3.7 Descriptive Parameters for Fecundity

Nine fly strains were analyzed in total. The final analyses were performed on 145 individuals (61 uninfected flies and 84 infected ones; SM S1). Descriptive statistics show that infected flies laid an average of 179.62±67.05 eggs during the 5 days and uninfected flies laid 171.79±69.45 eggs on average for the equivalent time (Table 2.13). The minimum number of eggs laid by infected and uninfected flies were 23 and 0, respectively, while the maximum number of eggs laid by those flies were 295 and 341, respectively (Table 2.13).

Table 2.13 Descriptive statistics for (early) fecundity of flies.

Infection Status	Number of Females	Mean	Standard Deviation	Minimum Number of Eggs Laid	Maximum Number of Eggs Laid
Infected	84	179.62	67.05	23	295
Uninfected	61	171.79	69.45	0	341

When the flies were separated by “population”, ”population” 1 (31 females) laid an average of 137.90±68.91 eggs per female during the 5 days and ”population” 2 (114 females) laid 186.77±64.09 eggs per female during the same time period (Table 2.14). The minimum number of eggs laid by ”population” 1 females and ”population” 2 females were 23 and 0, respectively, while the maximum number of eggs laid by those flies were 264 (”population” 1 females) and 341 (”population” 2 females) (Table 2.14).

Table 2.14 Descriptive statistics for (early) fecundity of flies by “population.”

"Population"	Number of Females	Mean	Standard Deviation	Minimum number of eggs laid	Maximum number of eggs laid
1	31	137.9032	68.9122	23	264
2	114	186.7719	64.0855	0	341

2.3.8 Results of Statistical Analysis for Fecundity

“Population” had an effect on fecundity, with ”population” 1 having significantly lower fecundity than the flies from ”population” 2 (p=0.0003). We ran a two-way ANOVA for the full set of observations with Infection Status, “population” and their interaction as factors, and the results showed that Infection Status and the interaction with “population” had no significant effect on fecundity (p>0.6587 for Infection Status and p>0.5719 for the interaction), but “population” still had a significant effect on fecundity (p>0.0007). Separate one-way ANOVA tests for each “population” indicated that Infection Status had no effect on fecundity (p=0.9335 for ”population” 1 and p=0.3528 for “population” 2). We concluded that the fecundity difference between “populations” was likely due to the difference in the length of time that the two “populations” had been cultured in the lab as well as difference in method of laboratory culture (i.e., bottle vs vials) as suggested by Sgro and Partridge (2000). “Population” 1 flies that were collected in Mexico were kept in the laboratory for more than 5 years in half-pint bottles, while ”population” 2 was field-collected in College Station, TX and were cultured in vials for less than a year. This research was not designed to answer the above question, so it will not be addressed any more in this paper.

Given the exploratory nature of these experiments and the intention to maximize the data available for analysis, we decided to run the fecundity analysis for all data grouped together. We had several reasons to do this: 1) we wanted to test if the treatment of interest (the fixed factor "Infection Status") had an effect on fecundity, 2) because the percentage of mature females gradually increases, starting at an age of 2 days old (Markow, 1985), we wanted to test if the age at which the experimental females were introduced into this experiment (i.e., "Age at Setup" variable) had an effect on fecundity, and 3) we were interested to know if the interaction between Age at Setup and the Infection Status had an effect on fecundity. Because "population" had an effect on fecundity, we ran the analysis both with and without "population" as a factor. The results of the analyses with and without "population" were similar to analyses with other fixed and random factors included, so we decided, to report the results without "population" as a factor.

The evaluation of female size as a potential covariate indicated that "population" had an effect on female size ($p<0.0001$), so this potential covariate was evaluated separately for each "population". Female size did not show a linear relationship to the fecundity of the uninfected and infected groups for each "population" ("population" 1, uninfected $p=0.0649$, infected $p=0.4587$; "population" 2, uninfected $p=0.0898$, infected $p=0.7741$) and we discarded the use of female size in the analysis.

We used Infection Status, Age at Setup and the interaction between those two as fixed factors in the analysis. We first compared a model with the fixed factors to a model that includes these factors plus the random factors as defined in the methods section. A likelihood ratio test

indicated that the model with the fixed factors plus the random factors fitted the data better than the model with the fixed factors only ($\chi^2_{(2)} = 7.2$, $p = 0.0136$); therefore, we kept the model with random factors for the next step of the analysis. Next, we tested whether we needed a model having homogeneous residual variances (above selected model) or heterogeneous residual variances. A non-significant likelihood ratio test ($\chi^2_{(1)} = 0.1$, $p = 0.7518$) indicated that none of the models were a better fit for the data, so we opted for keeping the simpler model, namely, the model with the fixed factors plus the nested random factors and interaction, with homogeneous residual variances. The interaction Isoline*Infection Status had a covariance parameter estimate of zero and was removed from the final model.

The diagnostics criteria used for this analysis indicated that the residuals met the assumption of normality (the Shapiro-Wilk *p*-value was 0.1680 for the combined data) and equality of variances (the standard deviations of the residuals were similar for the infected (53.8560) and uninfected (55.0474) groups, indicating no strong evidence of non-homogeneity of variance); the box plots of the conditional studentized residuals show that the distribution is approximately homogeneous across “isolines” and “repeats nested within fly strains”. Seven outliers were observed, identified by IDs 29, 31, 32, 35, 36, 75 and 88. The graphics used for influence diagnostics indicated that a combined set of 10 observations identified by ID numbers 01, 02, 36, 65, 73, 75, 88, 125 and 135 stand out from the other observations and could be considered influential values for all the criteria affecting the model, as described in the methods section. However, when the set of 13 observations was removed and the analyses rerun without it, the conclusions from the results did not change, so we report here the results of the full data analysis.

The final analysis shows that fecundity is not significantly affected by any of the main effects tested (Table 2.15), namely Infection Status ($F_{(1,43.8)} = 0.62$, $p = 0.4351$), Age at Setup ($F_{(1,44.7)} = 2.27$, $p = 0.1148$) and their interaction ($F_{(1,46.6)} = 0.60$, $p = 0.5509$). Table 2.16 shows the average differences in fecundity for individuals by each fixed effect relative to the reference level for each factor. The reference level of a factor is the level to which all other levels of that particular factor are compared to. For Infection Status, the control level was "infected", for age at setup, 6 days (old), and for the interaction Infection Status*Age at Setup, "Infected * 6 days (old)". By default, those reference levels are not shown in the table.

Table 2.15 Test of the main effects of infection status and age at setup on fecundity.

Effect	Numerator Degrees of Freedom	Denominator Degrees of Freedom	F Value	Pr > F
Infection Status	1	43.8	0.62	0.4351
Age at Setup	2	44.7	2.27	0.1148
Infection Status*Age at Setup	2	46.6	0.6	0.5509

Table 2.16 shows that that the estimated effect of the Infection Status "uninfected" is negative ("b" estimate = -5.04) and not significant, which means that the estimated fecundity of uninfected flies is lower than that of infected flies (for infected flies it is 135.97 and for uninfected it is 135.97-5.04=130.93) but is not statistically significant (p=0.9173). The average difference in fecundity of the flies' setup at age 4 relative to the control level (Age at Setup of 6 days) is 10.49 and not significant (p=0.8523). A similar explanation applies to all other factors in Table 2.16. Table 2.16 also shows the results for the random effects analysis, and it can be seen

that the variances of isoline and repeat(isoline) are not significant (*p* values are 0.1710 and 0.0814, respectively).

Table 2.16 Fixed effects and random effects for fecundity analysis.

Fixed Effects	Estimate (b)	Standard Error	*p*
Intercept	135.97	44.92	0.0042
Infection Status (UN)	-5.04	48.31	0.9173
Age at Setup: 4 Days	10.49	56.03	0.8523
Age at Setup: 5 Days	48.25	45.59	0.296
Uninfected x Age at Setup 4 Days	60.45	67.08	0.3722
Uninfected x Age at Setup 5 Days	8.03	51.54	0.8769
Random Effects (Covariance Parameter Estimates)	Estimate (b)	Standard Error	*p*
Isoline	367.07	386.27	0.171
Repeat(Isoline)	639.93	458.57	0.0814
Residual	3448.76	500.58	<0.0001

The results of the mixed model analysis exploring the interaction between Longevity and Infection Status indicate that this interaction does not affect fecundity (p=0.3574), nor does Infection Status (p=0.2041) or longevity (p=0.3587) on their own. Next, we wanted to explore the effect of 5 factors recorded in our study (Infection Status, Age at Setup, “population”, Longevity, and Size) on fecundity. To this end we used the PROC GLMSELECT as implemented in SAS 9.4® by running the SAS code for the procedure in SAS EG ® 7.1. (SAS Institute Inc., 2014) to select variables that would produce the “best” model (SAS Institute Inc., 2010; Beal, 2007). The variables selected were: Age at Setup, “population”, and the interaction

Longevity*"population" We did not run the model because the variable of interest, Infection Status, was not selected.

2.3.9 Repeated Measures Analysis of Fecundity

We ran a repeated measure analysis to explore whether the 5-day fecundity patterns observed for infected and uninfected females were affected by the age at which the females were entered in the experiment (i.e., Age at Setup). The results of this analysis indicate that the interaction Age at Setup*Time*Infection Status has a significant effect on fecundity (p=0.0313) (Table 2.17). Also, the 2-way interaction Age at Setup*Time and the main effect Time are significant (p=0.0002 and p=0.0255, respectively) (Table 2.17), but the main effect of the factor of interest, Infection Status, is not significant (p=0.4799). The 3-way interaction indicates that infected and uninfected females showed different daily fecundity patterns for the 5 days we recorded this parameter, and these patterns changed with the "Age at Setup" of said females. Figs. 2.4, 2.5 and 2.6 show these changing patterns of fecundity. Fig. 2.4 shows that when females were set up at age 4, infected females start day 1 with lower fecundity than uninfected females. Additionally, it shows that their fecundity rises higher than the uninfected ones at day 2 and continues rising at day 3 (but it is lower than the uninfected flies' fecundity this day). At day 4, both infected and uninfected flies showed a reduction of fecundity compared to day 3, but uninfected females had higher fecundity than infected ones. Finally, day 5 showed infected females with higher fecundity that uninfected ones, but both groups had lower fecundity than at day 4.

Table 2.17 Repeated measures analysis of early fecundity.

Effect	Numerator Degrees of Freedom	Denominator Degrees of Freedom	F Value	Pr > F
Infection Status	1	59.45	0.51	0.4799
Age at Setup	2	59.33	2.64	0.0795
Age at Setup*Infection Status	2	62.03	0.59	0.5576
Time	4	139	2.87	0.0255
Time*Infection Status	4	139	1.8	0.1326
Age at Setup*Time	8	139	4.03	0.0002
Age at Setup*Time*Infection Status	8	139	2.2	0.0313

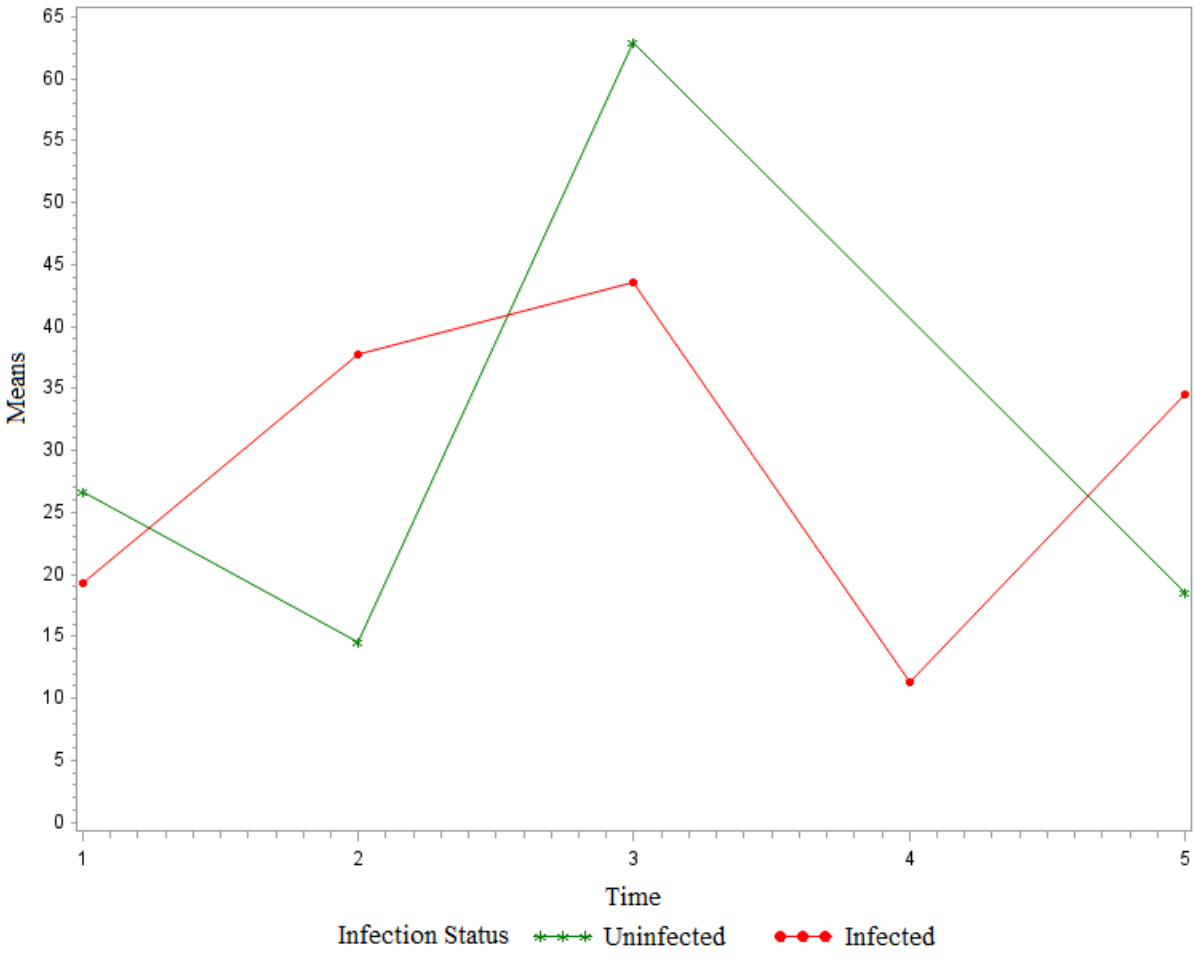

Fig. 2.4 Repeated measures analysis of fecundity. Infection status vs time at age at setup 4 days.

Fig. 2.5 shows that when females were set up at age 5, infected females started day 1 with lower fecundity than uninfected females, and their fecundity rose higher than the uninfected ones at day 2 (accompanied by a reduction in fecundity of uninfected females). At day 3 infected females diminished in fecundity parallel to a reduction of the fecundity of uninfected files, which continued to be lower than the first. At day 4, both infected and uninfected fecundities rose, but infected fecundity was lower than uninfected fecundity. The series ended at day 5 with infected females showing lower fecundity than uninfected flies.

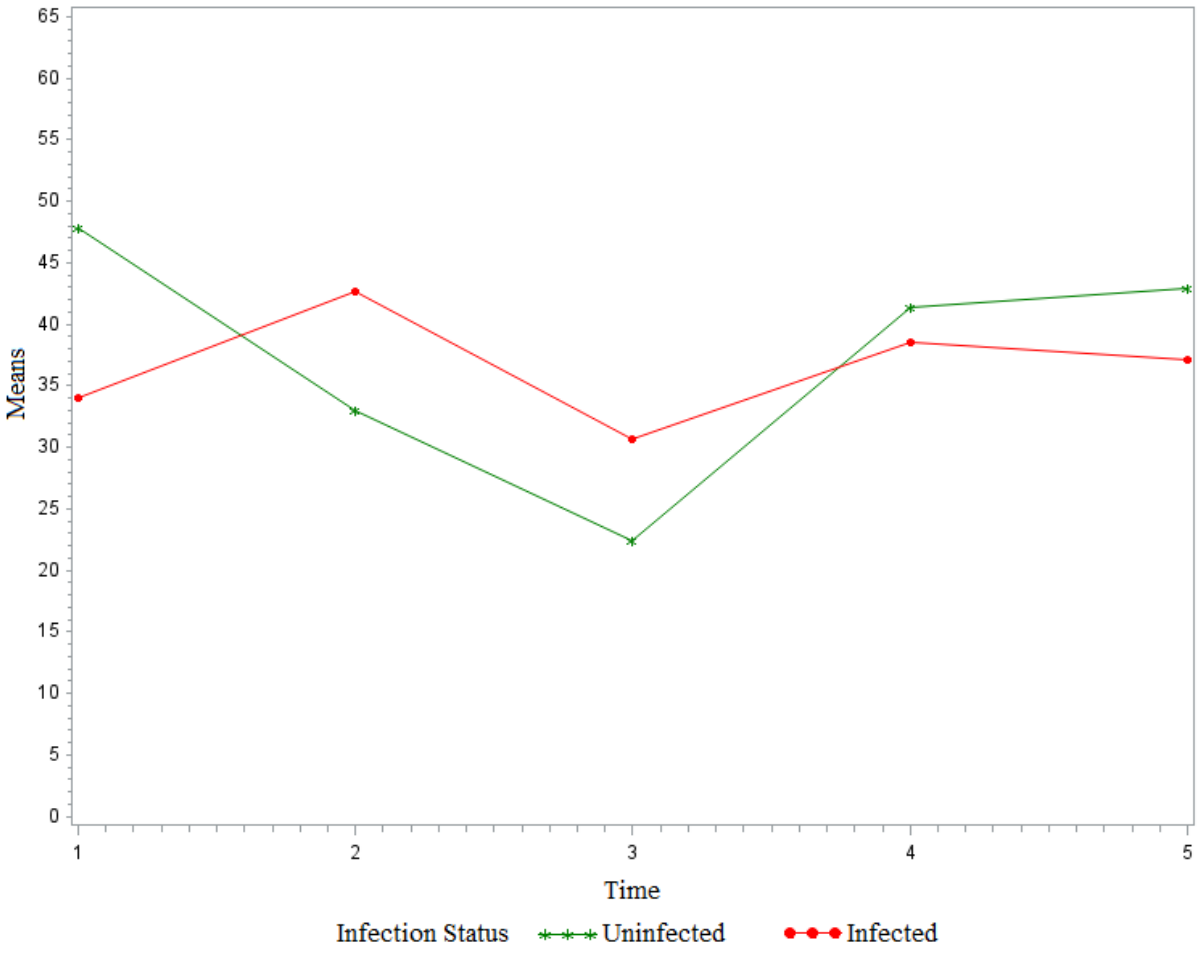

Fig. 2.5 Repeated measures analysis of fecundity. Infection status vs time at age at setup 5 days.

Fig. 2.6 shows that when females were set up at 6 days of age, infected females started day 1 with lower fecundity than uninfected females, and their fecundity rose slightly higher than the uninfected ones at day 2 (accompanied by a reduction in fecundity of uninfected females). From day 2 on, both groups showed parallel changes in fecundity (day 3 fecundity up, day 4 fecundity down, day 5 fecundity up again) with infected flies showing a pattern of higher fecundity than the uninfected ones.

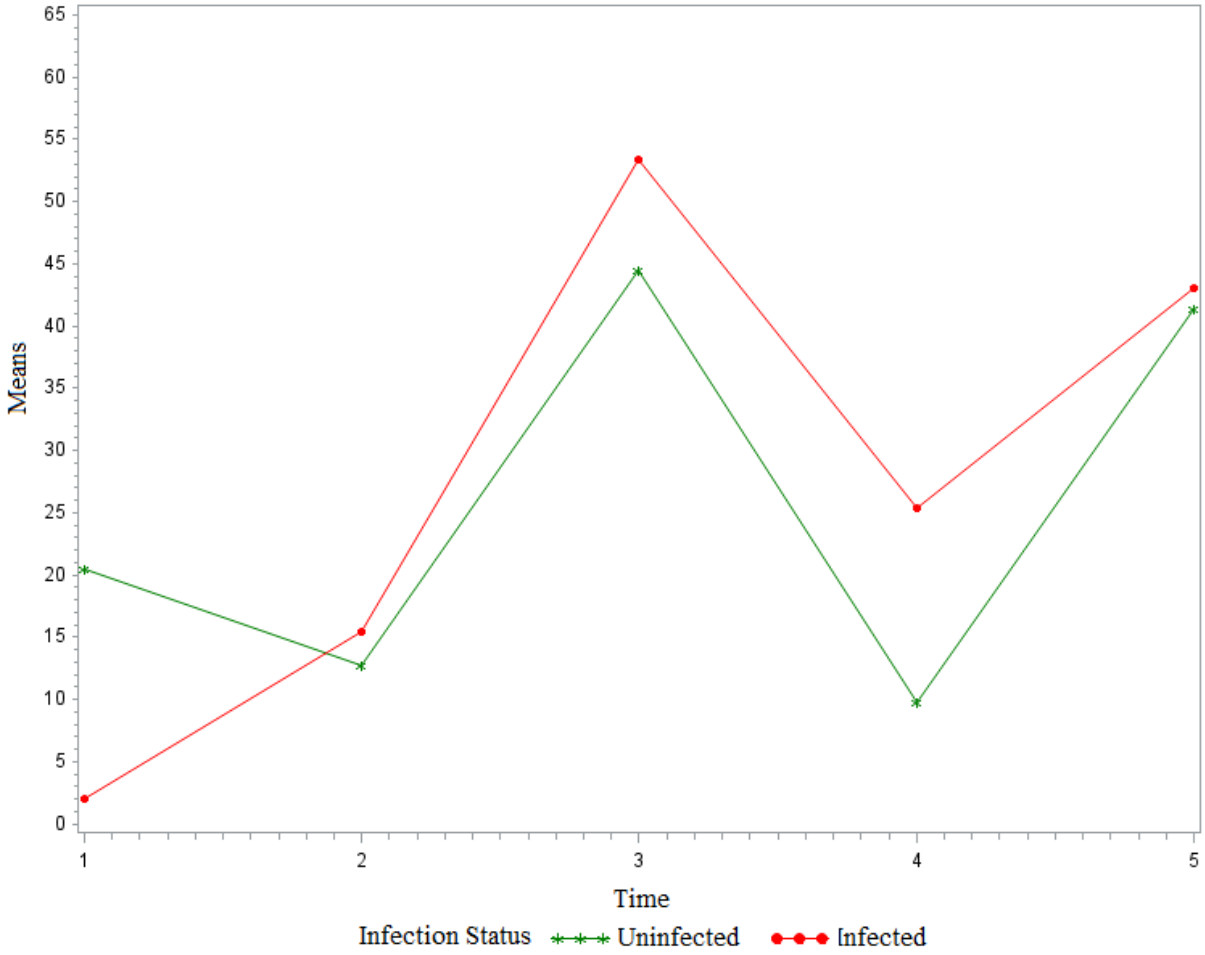

Fig. 2.6 Repeated measures analysis of fecundity. Infection status vs time at age at setup 6 days.

The patterns of variation described above indicate that Age at Setup interacts with the other factors included in the analyses. Age at Setup must be uniform in order to avoid the extra source of variation.

2.4 Discussion

We want to reiterate the exploratory nature of this study. As such, the discussion and conclusions arrived at here will lead to recommendations for the re-examining of some of our results in further experiments and to address the weaknesses of the current work in future studies as well as the formulation of questions that will lead to clarification of how infection persists in *D. hydei*.

2.4.1 Discussion for Ratio of Larvae-to-Emerged-Flies Survival

The full data analysis indicates that the ratio of larvae-to-emerged-flies is not statistically different for uninfected and infected flies. This agrees with Xie et al. (2010), who, working on the same strain of *Spiroplasma* and fly species as we did, reported that mortality for two complimentary intervals that together represent a measure of the ratio of larvae-to-emerged-flies (i.e., “larva-to-pupa survival rate” and “pupa-to-adult fly survival rate” in Xie et al., 2010), are not significantly different for *Spiroplasma*-infected and uninfected flies. Furthermore, Osaka et al. (2013a), working on a similar *Spiroplasma* strain and fly species as ours, reported that the “survival rate between hatched larvae and emerged adults” is not significantly different between infected and uninfected *D. hydei* females.

The literature also offers indications that *Hyd1* does not kill the progeny of *D. hydei* prior to the larval stage. Osaka et al. (2013a) reported that the “egg hatch rates” are not significantly

different between *Hyd1* infected and uninfected *D. hydei* females. Although Osaka et al. (2013a) did not describe the method to estimate the "egg hatch rates", it is logical to assume this was estimated as the number of larvae hatched divided by the number of eggs laid (per brood). Although *D. hydei* can lay unfertilized eggs (Markow, 1985), mated females most likely lay fertilized ones, and as such, they should contain the embryos in development, which are not killed by *Spiroplasma*.

This work, Xie et al. (2010), and Osaka et al. (2013a) show that *Spiroplasma* strain *Hyd1* does not kill *D. hydei* flies at the pre-emergence stages (embryogenesis, larval and pupal stages). We can conclude that *Spiroplasma Hyd1* does not affect the fitness of *D. hydei* through the early death of their progeny.

2.4.2 Discussion of Sex Ratio

We tested whether *Spiroplasma* infection kills the sons of *D. hydei* in two steps: 1) comparing the sex ratio of infected and uninfected females and 2) testing for deviations from the 50% sex ratio, because the sex ratio normally tends towards this equilibrium in the long term (Hamilton, 1967). The GLIMMIX analysis shows that uninfected females produce significantly more males than infected females did (p=0.0035). However, this is not due to a lower ratio of males to females born to infected females, as would be expected from a male-killing *Spiroplasma* strain, but it is due to a sex ratio higher than 50% for males born to uninfected females (Table 2.7, Table 2.9, and Fig. 2.7). The exact binomial test indicated that the 51% sex ratio (raw data) produced by infected females is not statistically different from the expected 50% ratio (p=0.619),

while uninfected females produced a male-biased sex ratio (59%, raw data) (p=0.00003) (Table 2.9 and Fig. 2.7).

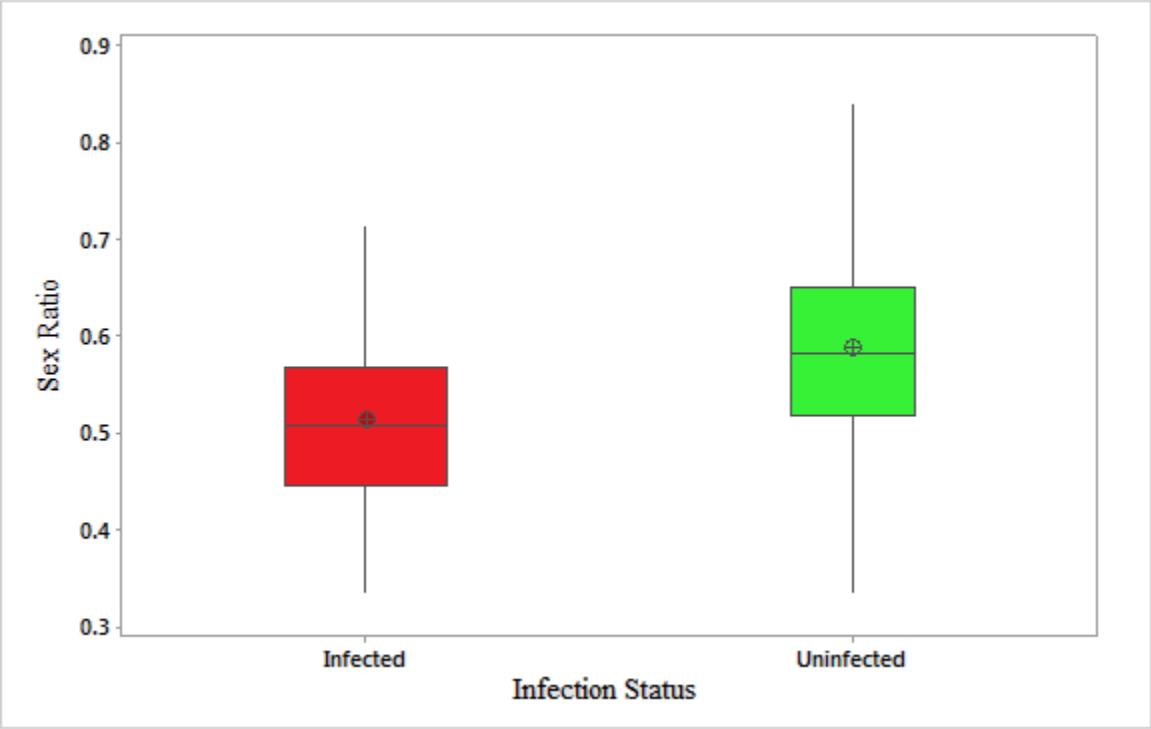

Fig. 2.7 Sex ratios (number of males/number of males plus females) for the progeny of *Hyd1 Spiroplasma*-infected and uninfected *D. hydei* females.

The literature reporting the sex ratio of *D. hydei* populations is scant. We were able to find reports of sex ratios for field-collected populations of *D. hydei* (Malogolowkin-Cohen, 1980; Pitnick, 1993). However, these reports likely include a mix of the secondary sex ratio (sex ratio at hatching) and tertiary sex ratio (sex ratio at maturity), and our data represents only the secondary sex ratio, so no valid comparison can be drawn between this study and those reports.

Although the results show a biased sex ratio at emergence for uninfected females, the sex ratio for infected females is no different from 50%. This does not support the hypothesis of male-killing by *Hyd1*in *D. hydei* females. Male killing would be supported if infected females would produce a sex ratio significantly lower than 50%.

Furthermore, our conclusion in the previous section that *Spiroplasma* does not kill the progeny of *D. hydei* females from the embryonic-to-the-emergence-of-immature-flies interval is another strong argument against male killing by *Spiroplasma* in our experiments. To understand this argument, it is necessary to restate here the characteristics of male-killing by *Spiroplasma.* Male-killing occurs at the embryonic stage (early male killing) or at the larval or pupal stages (late male killing) (Kageyama et al., 2007). When male-killing occurs, an excess of dead embryos/larvae/pupae for the infected females is accompanied by a lower (or 0) sex ratio for the emerging progeny of these infected females (pers. obs.; Kageyama et al., 2007). From this, it is logical to infer that the *Spiroplasma*-killed embryos/larvae/pupae are males. If *Hyd1* would kill males, we would expect to see the same pattern in *Hyd1* infected *D. hydei* females in our work. However, there is no statistical evidence for an excess of dead embryo-to-emerged-flies for infected females (p=0.3859) for our experiments, and a binomial analysis indicates that the sex ratio for the progeny of infected females is no different from 50% (p = 0.6190). All of the above allows for the parsimonious conclusion of no male killing by *Hyd1* in *D. hydei*. This means that the male-biased sex ratio reported in this work for uninfected females could be caused by factors other than *Spiroplasma.* We propose that one likely cause for our results on sex ratio is biased larvae collection, unintentionally caused by collecting infected and uninfected larvae at different times post-hatching. Another likely explanation is that the excess of males could be caused by

factors that bias the sex ratio without embryo mortality, such as meiotic drive (Committee on Gene Drive Research, 2016). We address these two arguments next.

2.4.2.1 Male-biased sex ratio for uninfected females is an artifact unintentionally caused by biased larvae collection

First, *D. hydei* females can lay unfertilized eggs (pers. obs.), making it difficult to control larval density as required by this experiment. Because of this, we had to rely on collection of first instar larvae (the first post-eclosion larval stage) to set up an equal number of larvae for infected and uninfected flies. We argue that at this stage, bias in larvae collection may have involuntarily been introduced. We collected larvae for the infected group right after eclosion and sometimes as they were hatching, but due to logistic reasons, the collection of larvae from uninfected females was mostly performed 24-72 hours post-eclosion. Collecting larvae that "catches the eye" of the collector (e.g., more active or mobile larvae, larvae roaming on the food surface) after 24 hours post-eclosion, would bias the sex ratio towards males if male larvae show behaviors that make them more noticeable to the collector. Larval behavior did not bias the sex ratio of larvae collected right after eclosion (i.e., larvae from infected females), because at the time of collection there was not a surplus of larvae to choose from, but sometimes we had to pause the collection for a few hours until more larva hatched. We then restarted the collection to complete the acquisition of the 30 larvae per females required by our protocol.

We do not have direct proof that the behavior of larvae biased the sex ratio of the uninfected flies, but we can hypothesize that this is possible by using published data regarding the locomotive activity of *D. melanogaster* larvae. Aleman-Meza, Jung, and Zhong (2015) reported

no statistical differences in several indicators of locomotive activity between males and females at the third instar larvae for *D. melanogaster*; however, an inspection of Aleman-Meza's "Table S2 "Factors affecting locomotive parameters" reveals that although not statistically significant, males seem to show measurement values that are related to faster locomotion. In fact, one measure, "run count (count/minute)," is higher for males than for females (males=2.34±0.95, females=2.17±0.91), although weakly significant p= 0.03). We propose that if the general trend observed by Aleman-Meza et al. (2015) for *D. melanogaster* is also observable for second instar larvae in *D. hydei*, then males would be more mobile than females at the time of collection. If the experimenter is prompted to collect the more active flies due to their higher "visibility", this would effectively bias the number of larvae collected for the samples towards males, thus biasing the sex ratio of the emerging flies from such a collected sample of 30 larvae (in this case, for the uninfected females).

A second line of support to our argument above comes from a study that did not use larvae collection to control for larval density, thus avoiding the potential for the bias described for our study. This study reported that the sex ratio did not vary significantly between uninfected and infected females (Xie et al., 2015).

2.4.2.2 Thee excess of males observed in uninfected flies is caused by factors other than Spiroplasma infection

A potential mechanism that would bias the sex ratio is meiotic drive, which has been reported for several fruit fly species, where it produces a female biased sex ratio (e.g., Committee on Gene Drive Research, 2016; Cazemajor, Landre, & Montchamp-Moreau, 1997). This phenomenon is

not associated with postzygotic mortality, but it is characterized by spermiogenic failure (Cazemajor et al., 1997). The gene complex responsible for this phenomenon occurs at low frequencies in the population (e.g., 1-5% in *D. melanogaster*; Committee on Gene Drive Research, 2016). Although this phenomenon has not been reported in *D. hydei*, meiotic drive could be a plausible cause of the male biased sex ratio without embryonic mortality reported in this work. This is because meiotic drive can also produce a male biased sex ratio due to the loss of sperm for species such as the Mediterranean fruit fly *Ceratitis capitate* (Shahjahan, Rendon, Cook, & Wood, 2006).

For our data, one out of the nine isolines used in the work (isoline TWENTY) causes the significant result ($p = 0.0159$) for the sex ratio analysis that compares infected vs uninfected female progenies; when isoline TWENTY is removed from the analysis, the difference in sex ratio becomes not significant ($p = 0.0656$), but the trend of higher sex ratio for uninfected flies still remains. If the male biased sex ratio of the uninfected isolines in the full data analysis (and isoline TWENTY specifically) is due to meiotic drive, then the difference in sex ratio between the infected and uninfected flies has to be explained by the presence of *Spiroplasma* in the infected group, but not through male-killing by the bacteria (we did not find a significant mortality difference between the uninfected and infected group at the stages male killing occurs). At this moment we do not have an explanation for the sex ratio difference reported here, under the meiotic drive hypothesis.

Another mechanism biasing the sex ratio but not involving male killing by *Spiroplasma* is the one reported by Long and Pischedda (2005), who, studying *D. melanogaster* in a context other

than *Spiroplasma* infection, reported that females produce more male progeny when mated to younger males than when mated to older ones.

Meiotic drive and mating to younger males as cause of male biased sex ratio in *D. hydei* are open questions that require further research. However, because the goal of our experiment was to find out if *Spiroplasma* affects the fitness of its host through male killing, we will not pursue those lines of research in this work.

2.4.2.3 Final word for sex ratio

Overall, the results of this section showing that the sex ratio of the infected flies' progeny is not different from 50%, with support from the finding that there is not differential early mortality for infected and uninfected females, as well as with support from other researchers' works, leads to the conclusion that *Spiroplasma Hyd1* is not a male killer in its natural host, *D. hydei.* We can conclude that *Spiroplasma* does not offer a fitness advantage to *D. hydei* through the phenomenon of male-killing, hence male-killing does not contribute to the persistence of the infection in this fly.

This work illustrates the difficulty of working with flies that can lay unfertilized eggs, because this situation does not allow for readily controlling density to equalize the rearing conditions of each female and their progeny in a given experiment. Larvae collection, as discussed above, may introduce bias. The best method to study the sex ratio of such species is perhaps to study equal-sized groups of females in pint-sized bottles, and, after a period of egg-laying, to remove the mothers and sex the progeny as they emerge. This method was used by Xie et al. (2015) to obtain

the sex ratio of *Spiroplasma*-infected and uninfected *D. hydei* females, arriving to the conclusion that *Spiroplasma* does not kill males in this fly host.

2.4.2.4 Future directions for sex ratio

The apparent male-biased sex ratio of the uninfected females prompted us to consider issues such as how males, regardless of sex ratio in *D. hydei* populations, may contribute to the persistence of *Spiroplasma*. We reasoned that *Spiroplasma* infection may prompt males (both uninfected and infected) to mate preferentially with infected females. Infected females would have the advantage of a larger pool of males to fertilize their eggs, so they would produce more progeny than uninfected females. The possibility of males mating preferentially with infected females, coupled to other effects of bacterial infection such as increased fertility or early fecundity advantage, should be further explored.

2.4.3 Discussion for Fecundity and Longevity

2.4.3.1 Longevity

Drosophila hydei females infected with the *Hyd1 Spiroplasma* lived shorter lives than their uninfected counterparts. To the best of our knowledge, this is the first time that this effect has been documented for this species. Until now, the prevailing wisdom was that the native *Spiroplasma* strains do not produce detectable direct fitness effects in *D. hydei* (Kageyama et al., 2006; Osaka et al., 2008; referenced as unpublished data for both). However, the reports above do not offer any details on how the experiments were conducted or what type of data were collected, because those reports refer to unpublished data.

The finding of reduced longevity of *D. hydei* females infected by *Spiroplasma* is in agreement with other studies that show that the bacterial infection shortens the life span of the host. For example, *Spiroplasma* reduces the life span of female *Harmonia axyridis* (a coccinellid coleopteran) (Majerus, 2003), in the Pea Aphid *Acyrthosiphon pisum* (a homopteran) (Fukatsu, Tsuchida, Nikoh, & Koga, 2001), in the leafhoppers *Dalbulus* and *Baldulus* (hemipterans) (Madden & Nault, 1983) and in *Drosophila melanogaster* (Herren et al., 2014).

The report by Herren et al. (2014) is directly relevant to this work because the *Spiroplasma* strain that infects *D. melanogaster* is closely related to the strain that infects *D. hydei*. Certainly, these two strains belong to the same clade (the poulsonii clade), as determined by a phylogenetic analysis using 16s rDNA (Haselkorn et al., 2009). This may indicate that spiroplasmas from the poulsonii clade have a detrimental effect on the fitness of their hosts. Shorter lifespan will not favor the persistence of *Spiroplasma* in *D. hydei*. However, if a diminished lifespan due to infection prompts an increased reproductive activity of the host (Agnew et al., 2000), thereby increasing fitness, then this would be a mechanism that promotes the persistence of the bacteria in *D. hydei*. Contrary to what is expected from that theory, this work shows that early fecundity (at least the fecundity for the first 9 to 11 days studied here) is not increased in *D. hydei* for infected females as a response to the shortening of longevity.

2.4.3.2 Fecundity

Fecundity is not different between *Spiroplasma*-infected and uninfected flies (p=0. 4351), and this may indicate that *Spiroplasma* infection is not associated with a fecundity benefit to female *D. hydei*, and that fecundity is not involved in the persistence of infection in *D. hydei*. We will

argue next that these results cannot be taken as definite because we introduced sources of variation in the experiment that may obscure the effect of the variable of interest, Infection Status, on Fecundity. Therefore, we may need to repeat this experiment with a better design.

A repeated measures analysis shows that there is a significant effect of the three-way interaction of the fixed effects Infection Status, Age at Setup and Time (of oviposition) on the pattern of fecundity of the experimental flies (p=0.0313, Table 2.17). The objective of this repeated measures analysis was not to go into details on the repeated measures, but to show that a controllable factor, time at which the females were entered in the experiment (Age at Setup), can obscure the effect of the factor of interest in the experiment.

Figs. 2.4, 2.5 and 2.6 show that in this study *D. hydei* females lay eggs in "waves", with troughs every 3 to 4 days. Infected females, regardless of the Age at Setup, started day 1 of the 5-day oviposition series at a lower level of fecundity than uninfected flies. Whether this is due to many of the infected females being immature or them being at a wave's trough at setup cannot be determined from the data, because we did not record their fecundities before we set them up. Regardless of the cause of the described pattern, this situation seems to put the infected females at a disadvantage in fecundity when compared to uninfected flies. This view is reinforced by the fact that more infected females in the sample started day 1 with 0 eggs laid when compared to uninfected ones. Table 2.18 shows the percentage of females laying 0 eggs at day 1 within their age group per infection status. Infected females starting with 0 eggs laid at day 1 are: 50% (4 days old at setup), 50% (5 days old at setup) and 100% (6 days old at setup), while for uninfected females the equivalent percentages are: 60% (4 at setup), 34% (5 at setup) and 56% (6 at setup).

Table 2.18 also allows to reiterate the information presented in Fig. 2.1, about a source of variation that was introduced in the experiment: the unequal number of infected and uninfected females at each age of setup. A small percentage of infected (5/84, or 6%) and uninfected (5/61, or 8%) females were set up at 4 days age old. It is at 5- and 6-day-old setup ages that there are large differences between the uninfected and infected groups: 90% (76/84) of infected females and 62% (38/61) of uninfected females for 5 days, and 4% (3/84) of infected females and 30% (18/61) of uninfected females for 6 days of Age at Setup.

Table 2.18 Number of experimental females in the fecundity experiment and percentage of females that laid zero eggs at day 1 within their age group per infection status.

Age at Setup (Days)	Number of Infected Females	Percentage of Infected Females Laying Zero Eggs at Day 1	Number of Uninfected Females	Percentage of Uninfected Females Laying Zero Eggs at Day 1
4	5	50.00%	5	60.00%
5	76	50.00%	38	34.00%
6	3	100.00%	18	56.00%
Total	84		61	

Table 2.19 shows the results of an ANOVA comparing the fecundity of infected and uninfected flies by Age at Setup. The differences in fecundity for the infected and uninfected flies are not significant (p=0.35, p=0.65 and p=0.98 for ages 4, 5 and 6 days old, respectively) (Table 2.19). Coupling the above to data from Table 2.18 that show that at ages at setup of 5 and 6 days old, a larger percentage of infected females started with zero eggs laid at day 1, we propose that there is an argument for the disadvantage in infected females' fecundity being due to the setup of the

experiment. In summary, we consider that the results of the fecundity analysis are not conclusive due to the bias introduced in the experiment by variance in age at which the females were set up for the data collection.

Table 2.19 Selected output from the ANOVA comparing fecundity of infected and uninfected flies by age at setup.

Age at Setup (Days)	Infection Status (# of flies at setup)	Mean	Standard Deviation	*p*-value
4	Infected (5)	144.80	65.19	0.35
	Uninfected (5)	183.00	55.64	
5	Infected (76)	183.91	67.3	0.65
	Uninfected (38)	190.11	68.09	
6	Infected (3)	129.00	26	0.98
	Uninfected (18)	130.00	59.73	

2.4.3.3 Future directions for fecundity

To address the bias introduced by variation in the age at which females were entered in the experiment (i.e., variable "Age at Setup" in this work), in future studies females should be set up at a uniform age, before they mature. Due to the variability of the ages at which females attain maturity (Markow, 1985), we recommend starting with females at the age of 0 days old and sampling fecundity for periods no shorter than 15 days (this way, the fecundity of flies will be sampled for more than a week, allowing most of the females to reach maturity). Alternatively, one could start them at 10 days old, to make sure that sampling occurs when most of the females are mature and laying eggs.

Other sources of variation identified in this study were the length of time that the flies were cultured in the laboratory (more than 5 years of laboratory culture for “population”1 vs recently wild caught for “population” 2 of this study), as well as the type of container in which the flies were kept during their time in culture prior to being used in our experiments (half pint bottles for “population” 1 vs narrow vials for population 2 of this study). These two culture conditions affect the lifespan and fecundity of flies (Sgro & Partridge, 2000), two of the various factors of interest in our study. To control for the variability in biological characteristics of experimental flies produced by different times and container types of culture, we recommend using only flies that underwent one type of culturing condition in future experiments. We recommend using recently wild-caught females because they will have the natural variability of wild flies. The type of container should be restricted to either half pint bottles or narrow vials.

2.4.3.4 Fecundity and longevity future experiments

Because we found that infected females lived shorter lives than uninfected ones, we consider that further experiments should be conducted to assess if the shorter life span of infected females leads to a diminution of lifetime fecundity and an increase in reproductive activity at early ages, as predicted by Agnew et al. (2000). Although our data analysis shows that there is no difference between uninfected and infected females’ early fecundity, we have argued above that these results are not conclusive due to bias in the age at which the flies were set up for the experiment. To address the prediction by Agnew et al. (2000), we recommend repeating the longevity and early fecundity experiments in order to confirm or reject the present results. we also recommend adding another variable, lifetime fecundity, to explore if the reduced lifetime caused by the infection results in a reduction of the lifetime reproductive output. Lifetime and early fecundity

should be obtained from the same females used for the longevity experiment. An increase in early fecundity in infected females may be accompanied by an earlier age of maturity for these females, so future experiments should also address this issue.

2.5 Conclusions

2.5.1 Findings

This exploratory research found that male killing or early mortality of infected flies (as a proxy of early male killing) do not occur in the association of *Hyd1 Spiroplasma* and *D. hydei* females, at least under the conditions in which these experiments were conducted. *Hyd1 Spiroplasma* infection has a mild detrimental effect on the longevity of the females of its natural host *D. hydei*. This finding is similar to a report that the male-killing MSRO (Melanogaster Sex-Ratio Organism) shortens the lifespan of infected *D. melanogaster* females (Herren et al., 2014). This report and Herren et al. (2014) suggest that shortening the lifespan of their *Drosophila* female hosts may be a common characteristic of *Drosophila*-associated spiroplasmas. We also report that infection by *Spiroplasma* does not affect the fecundity of infected females. Overall, the shorter lifespan of infected females indicates that *Spiroplasma* infection has a negative effect on the fitness of *D. hydei*, and this effect does not help to explain the persistence of the infection in this host.

2.5.2 Future Research

In line with the ultimate goal of this exploratory work, which is to identify any fitness effects on the host that could promote the persistence of *Spiroplasma* in *D. hydei*, we recommend following this work up with further research on several topics. We recommend that the lifespan reduction

of infected *D. hydei* females reported here be strengthened with confirmatory experiments. This is because the shortening of the lifespan of *D. hydei* females by *Spiroplasma* is reported here for the first time, despite it being generally accepted that *Hyd1 Spiroplasma* is not detrimental to the female *D. hydei*.

One theory predicts that when a parasite diminishes the fitness of the host through shortening of the lifespan, this should be accompanied by increased reproductive activity of the host (Agnew et al., 2000). We report here that no increased early fecundity occurs in infected flies, which contradicts the prediction by the theory. Because the shortening of the lifespan of the host does not promote the persistence of the bacteria, further research must be conducted to discern the causes of the persistence of *Spiroplasma* in *D. hydei*. Based on the theory's predictions, we recommend that the study of characteristics that relate to the host's reproductive activity linked to longevity be prioritized. One of those characteristics is early fecundity, because the results of this work on early fecundity are inconclusive. Another characteristic that merits study is age of maturity (whether infected females mature earlier). Total fecundity (whether infected females are more fecund than uninfected flies) should also be addressed. Those three characteristics that result in increased fitness of infected females would promote the persistence of the bacteria in spite of the decreased lifespan of infected females.

The initial results on the biased sex ratio of uninfected *D. hydei* in this work (which upon deeper analysis was deemed not to be correct) prompted us to ask if preferential mating could play a role in the persistence of *Spiroplasma* in this host. This is plausible because endosymbionts have been shown to be involved with preferential mating in other insects, including *D. melanogaster*

(e.g., Sharon et al., 2010). We suggest that preferential mating of uninfected and *Spiroplasma* infected males towards *Spiroplasma* infected females be investigated. Although preferential mating with infected females per se would not promote the persistence of the infection, it could act in conjunction with other factors to this end.

3. EFFECTS OF *SPIROPLASMA* INFECTION ON *DROSOPHILA HYDEI* AND ITS POSSIBLE CONSEQUENCES ON THE PERSISTENCE OF THE INFECTION

3.1 Introduction

This chapter is the second of two addressing the long-term persistence of maternally heritable endosymbionts. Two fitness related characteristics previously explored in chapter 1 are readdressed in this chapter. One characteristic is readdressed to confirm the exploratory results from chapter 1, while the other is re-examined with improved protocols to avoid the source of variation identified in chapter 1. This chapter also addresses questions generated from the exploration of fitness characteristics in chapter 1, by expanding the study into several characteristics of the host that may be affected by the infection, thus affecting its fitness and promoting the long-term persistence of the bacteria.

The specific problem this work addresses is the persistence of an endosymbiont, the *Spiroplasma* strain *Hyd1* infecting its host *D. hydei*, that does not produce fitness benefits to the host (Kageyama et al., 2006; Osaka et al., 2008, referenced as unpublished data), or do not have 100% vertical transmission, and as such, according to our interpretation of Werren and O'Neill (1997), it should have long-since disappeared. However, this association has persisted in the long-term in the wild (Ota et al., 1979; Osaka et al., 2010) and in laboratory populations (mentioned by Watts et al., 2009; Osaka et al., 2013a). Although context-dependent fitness benefits and horizontal transfer may be possible in this relationship (Xie et al., 2010; Xie et al., 2015; Osaka et al., 2013b), the contributions of these factors to the persistence of the endosymbiont still have to be demonstrated and evaluated in the wild. One reason to choose the

association of *Spiroplasma* strain *Hyd1* and *D. hydei* is that we had used it already in exploratory experiments and the information from that research helped to focus the questions for the current experiments. Another reason for our interest on this association is that the *Hyd1* belongs to the poulsonii clade, which includes male-killing strains as well as some non-male-killing strains that infect *Drosophila* (Haselkorn et al., 2009), and the results obtained from studying *Hyd1* will help bridge studies on both-male killing and non-male killing phenotypes. Another reason to focus on *Hyd1* and its host in this study is that our results can be compared to and discussed along with a growing number of reports from different areas of the world, a fact that may render our conclusions more generalizable.

We collected preliminary data that initially seemed to indicate that *Spiroplasma* reduced the sex ratio of infected *D. hydei* females, but further analysis revealed this was not the case. However, the initially apparent biased sex ratio caused by the infection prompted us to think of other ways males may contribute to the persistence of the infection. We reasoned that if both infected and uninfected males mated preferentially with infected females, then there would be an excess of mating partners available for infected females, which would mean infected females mate more times than the uninfected ones, thus contributing a larger proportion of individuals to the population. Preferential mating prompted by *Spiroplasma* is plausible because parasites/commensals can affect the mate choice of males and females. For example, in *D. melanogaster* commensal gut bacteria have been reported to produce positive assortative mating in populations grown in two different media (a corn-molasses-yeast-based media vs. a starch-based media), with flies mating preferentially within their own food group as early as two

generations after the switch of the groups onto separate media (Sharon et al., 2010). The consideration above prompted us to test the mating preference of males in this work.

We must clarify here that we are not proposing that preferential mating alone would cause the persistence of *Spiroplasma* in *D. hydei*. As shown later, this research will address several other potential effects of the infection on the host that may contribute to the persistence of the bacteria. If preferential mating occurs side by side to other effects of the infection on the host, such as a modest increase in fecundity, then the combination of those factors may contribute to the persistence of *Spiroplasma* in *D. hydei*.

Our preliminary data also indicate that *Spiroplasma* infection shortens the lifespan of *D. hydei*. When hosts have their lifespan shortened by parasites, theoretically predicts that the host will alter their reproductive schedule by placing increased reproductive efforts earlier in life than non-parasitized individuals (Agnew et al., 2000). Earlier higher reproductive activity/effort of infected females would allow them to compensate for the loss of reproductive output due to early death**.** In our system under study, this advantage could contribute to the persistence of the infection in laboratory and field-dwelling populations. This hypothesis is in line with the idea that endosymbionts can benefit from the host compensatory responses triggered by their presence (i.e., infection), because such response can ensure the persistence of the infection in future generations (reviewed by Lefèvre et al., 2009). An alternative interpretation of the phenomena is that the endosymbionts manipulate the fitness of their hosts to ensure their persistence (ibid.).

However, our preliminary data indicates that the infected females with reduced longevity do not have higher fecundity than the uninfected ones, which contradicts the theory above. A more careful look at those preliminary results suggested that the effect of infection may have been obscured by the unintentional introduction of variability, such as setting up the experimental females at several ages (4, 5 and 6 days old) and within those age groups, the number of infected and uninfected females included in the experiments were different. It is known that in populations of *D. hydei* the number of females reaching maturity increases with age (Markow, 1985). Unequal numbers of infected and uninfected females maturing at different rates within the age at setup groups may have created a situation where the effect of infection status on fecundity could not be distinguished from other factors.

Under the considerations above, the issue of whether or not there is an early or lifetime fecundity advantage for infected females in our experimental flies merits another look. This is especially true when considering that the reports of lack of benefits for infected females (Kageyama et al., 2006; Osaka et al., 2008) do not offer any details on how the experiments were conducted or what type of data were collected, because those reports refer to unpublished data.

The preliminary data also showed that the young infected females referred to above seemed to start a 5-day reproductive series with lower fecundity than uninfected females. This brought to the forefront whether the infected females mature later than the uninfected ones (hence, their consistent lower fecundity) or earlier (hence, the lower fecundity was in reality a trough from a wavelike egg laying pattern shown by the flies). This issue warrants further investigation given that (NSRO), a *Spiroplasma* strain that naturally infects *D. nebulosa*, has been associated with

infected female hosts maturing 1 day earlier than uninfected ones (Malogolowkin-Cohen & Rodriguez-Pereira, 1975). Early maturity would, in theory, increase the fitness of infected females and thus, could be a factor promoting the long-term persistence of *Spiroplasma* infection in *D. hydei*.

We decided to address several potential benefits of the infection that are amenable to experimental study in the laboratory. These potential benefits are the effects of infection on age of maturity, mating preference (i.e., if the infection produces preferential mating to infected *Hyd1* females by both infected and uninfected males), fecundity, and lifespan of infected females. Our ultimate goal is to determine if there is a factor, or a set of factors acting together, that contribute to the long-term persistence of *Spiroplasma* in *D. hydei*. This research was undertaken with four objectives: 1) to determine if infected and uninfected males mate preferentially with infected females; 2) to determine if infected flies mature faster when compared to uninfected flies; 3) to confirm preliminary information that infected females are shorter-lived than uninfected flies; and 4) to determine if infected females have higher early fecundity or higher lifetime fecundity than uninfected flies.

3.2 Materials and Methods

3.2.1 Spiroplasma Infected Strains

D. hydei fly strain TEN 104-106 infected with haplotype 1 *Spiroplasma* (from now on *Hyd1*) was donated by Dr. Tamara Haselkorn. This strain was maintained for more than 10 generations at Rochester University and other institutions before it was used as the source of infectious hemolymph to create infected fly isolines for our experiments.

3.2.2 Fly Rearing Conditions

Flies were reared on banana opuntia food at 12 hours light/12 dark cycles. The daily temperature varied from 22 to 24°C in the fly rearing room used for our experiments. Experimental flies generated from fly stock populations were reared at a density of 50-100 females to 150 males in 75 ml of food in half pint milk bottles and experimental flies generated from field-trapped females at 1-10 females to 2-10 males in narrow vials containing approximately 7.40 ml of food.

3.2.3 Trans-infections

We use this term to indicate the laboratory-accomplished process of transferring hemolymph from a *Spiroplasma*-infected donor fly into an uninfected recipient fly, with the purpose of establishing the infection in the latter. We performed this procedure on CO2- anesthetized flies. To transfer the hemolymph, we used manual micro-injectors consisting of micro tips prepared from pulled micro-capillaries, a plunger prepared from a tailor's pins, and a viscous phase (mineral oil), to carry the pressure of the plunger through the micro-tip. The oil does not come in contact with the hemolymph during the injection process and is not injected into the recipient fly. We punctured the infected donor and the recipient flies at the soft juncture of the mesopleuron and scutellum to minimize the magnitude of the wound on the recipient flies. After trans-infection, the flies were left to recover from the wound for about two days before further manipulation. Typically, the survival rate is very high (~80% to 100%) and the success of transferring the infection is ~100% (at least for the *Spiroplasma-Drosophila hydei* system used in the present research).

3.2.4 Extracting DNA and PCR Screening of Bacterial Infection

In this study, a crude DNA extraction was performed by incubating a crushed fly on a mixture of 96 uL buffer solution (10 mM Tris.Hcl, 25 mM NaCl, 1 mM EDTA pH 8) and 4 uL of proteinase K solution (10 mg/mL, Denville Scientific). The incubation was carried on at 50° C for 60 minutes followed by denaturation of the proteinase K at 85° C for 30 minutes. To detect bacterial infection, we run 15 uL PCR reactions using as a template 2 uL of the crude DNA extract above. For *Spiroplasma* we used two primer sets 1) Spoul (Montenegro et al., 2005) run at an annealing temperature of 55° C; 2) p58IV, designed by the author, targeting a fragment of about 360 bp that is a subset of the *p58* primer set product; the annealing temperature for p58IV is 53°C and the sequences are p58IV_F 5'-AAAGGTTTACATTCACCAAGTCG-3' and p58IV_R 5'-AATTGTTCATTAACTTTATCTTGTGG-3'.

3.2.5 Generating Uninfected Flies from Infected Stocks

Uninfected flies were generated from the infected stocks by antibiotic curing a subset of the infected females. A combination of tetracycline and erythromycin (final concentration = 0.2 and 0.16 mg/ml, respectively) was mixed with the fly food. The flies were kept on this antibiotic-food mix for two generations. Later generations were kept in antibiotic-free food. PCR screening was used to confirm that the third-generation flies were cured. Both infected and cured (uninfected) groups were kept separately in half pint milk bottles (75 mL of antibiotic free food) at 50-females to 100-males density for eight generations before using them in the experiments, including the male mating preference experiment.

3.2.6 Generating Isofemale Lines (Isolines) from Field-trapped Flies

Five field-dwelling *D. hydei* females were trapped with banana baits in College Station, TX on May 2013. These flies were used as the founders to generate the isolines to be used in our experiment. Due to the relatively small number of field-collected individuals, we aimed to maximize the number of flies per isoline available for analysis at the end of our experiments. To achieve this, we inbred a large number of progeny from each founder female because we expected to lose many of those flies in the course of the breeding process due to a wide variety of factors. We identified 5 potential causes for losing isolines in our experiments, namely 1) some field-caught females or their progeny may produce a small number of progeny that will die off for reasons beyond the experimenters' control (this may be caused by the old age or the bad health status of the female in question); 2) any of the randomly chosen infected mothers necessary to produce the next generation may not carry the infection due to chance alone (the vertical transmission of *Spiroplasma* is not perfect in *D. hydei*); 3) the progeny of an infected mother may screen as negative for the infection; 4) the randomly chosen infected and uninfected mothers may not produce eggs in the time frame required for the flow of the experiment or may fail to produce eggs 5) staff may not be available for handling the flies in due time at every step of the experiment.

With the above in mind, each field-caught female (founder female) was placed in a separate vial containing banana-opuntia food labelled 1, 2, 3, 4, or 5. The females were allowed to oviposit for five days and the eggs of each founder female were incubated in their respective vials and the progenies allowed to breed with their siblings for four post-field-capture generations. The emerging flies from the fifth generation were divided into nine repeats (1, 2, 3, 4, 5, 6, 7, 8, and

9). Isoline 4 was lost at the fourth generation, leaving four isolines with nine biological replicates (or repeats) of each of the isofemale lines. These 36 total biological replicates were allowed to inbreed for four more generations (post-field-captured generations five to eight), using five females and ten males per vial each generation to avoid overcrowding and to control density.

Flies were pulled from the 9th inbred generation to create the "generation zero" (G0) experimental groups "infected" and "uninfected." For each repeat within the fly isolines the infected group was generated by trans-infecting six 10-day old females (later placed in vials labelled 1 to 6). The uninfected group was generated by injecting 10-day old females with hemolymph from uninfected siblings to replicate the stress of puncturing caused by trans-infections. For the uninfected group, only 4 flies were injected and placed individually in vials labelled 1 to 4. At this stage, G0 was constituted by 2 groups, one formed by infected flies and the other formed by uninfected flies; each group contained "duplicates" of the 4 isofemale lines originating from the wild-caught females. Each isoline within the infected and uninfected groups was represented by nine repeats, with each repeat constituted by 6 (infected group) and 4 (uninfected group) vials, respectively. Each of the G0 females was cultured with two uninfected males in a vial as the founders of infected and uninfected sets for each strain. The progeny of these uninfected and potentially infected females was allowed to inbreed for four more post-treatment generations (G1-G4) in sets of one 10-day old female and two 10-day old males per vial before taking the step of density control for the cultures at G5. Control of density of the cultures was implemented by collecting 30 G5 first instar larvae for each experimental vial and incubating them to collect the emerging G5 flies. Emerging G5 flies were mated at 10 days old (one female two males per vial). G6 to G7 flies were treated as G5. Emerging G7 flies were used

in the age of maturity experiment and those emerging at G8 were used in the fecundity-longevity experiments (see Table S4 in SM S4 for a summary of the inbreeding process followed to create the infected and uninfected isolines). Through losses of females due to the factors 1 through 5 outlined earlier, at the end of the experiment we obtained 47 vials of infected flies for each of the four isolines and 27 vials for their uninfected counterparts (see Table S5 in SM S5 for a summary of number of isolines, biological repeats and vials remaining for the analysis in these confirmatory experiments).

3.2.7 Screening Wild-caught Founding Females and Post-injection Females for Spiroplasma Infection

After oviposition, the founder females were screened for *Spiroplasma* and *Wolbachia*. Based on previous experience trapping *D. hydei* in College Station, no *Spiroplasma* or *Wolbachia* infections were expected to be found in this species population, a fact confirmed in our experiment. The G1 to G8 descendants of the infected females were screened for *Spiroplasma* by randomly choosing five flies per infected vial. Vials from which all five flies produced a negative PCR reaction for the bacteria were discarded.

3.2.8 Male Mating Preference Experiment

Infected fly stocks and their cured counterparts described in the section “Generating uninfected flies from infected stocks” were used for this experiment. The first generation of emerging cured flies (the progeny of flies after the 2 generations of antibiotic curing) were collected every 24 hours for 5 days. Males and females were kept separately in vials and allowed to age to 10-days-old and then mated in a new bottle at a density of 50-females to 100-males. This was repeated for

eight generations. Emerging flies from the eighth generation were separated by sex and aged to 10 days old to be used in the male mating preference experiment. Sets of 2 females and 1 male constituted each replicate for testing male mating preference.

Two 10-day-old virgin females (1 infected and 1 uninfected) were placed into a modified vial ("mating chamber", Fig. S6 in SM S6) with 1 10-day-old virgin male. Half of tests used an infected male and half used a non-infected male. Wings of females were clipped for identification. To control for any clipping effects, clipping alternated between infected and uninfected females. Flies were transferred into the mating chambers by using a funnel, without use of an anesthetic. Males were introduced first, held in half of the chamber by a movable panel, and then, the females were added, the chamber was closed and the panel removed. The flies were observed under a dissecting microscope until the first mating occurred. After mating, the expected infection status of the mating male and female were recorded and all flies were collected and frozen at -20°C for later PCR screening to confirm their infection status.

Mating experiments were carried out from 1:00 PM to 5:00 PM. at room temperature (24°C). Twenty-five mating chambers were utilized, ensuring that each day a chamber was used only once. At the end of the day, the used chambers were rinsed with ethanol and left to dry until the next day.

3.2.9 Age at Maturity of Males and Females Experiments

To determine the age at maturity of infected and non-infected flies, an isoline was chosen at random from the available G3 post-treatment trans-infection isolines that are described in the

section titled “Generating Fly isolines from field trapped flies.” The chosen isoline was identified as Number 2; both infected and uninfected groups of this isoline were used for this experiment. The process of inbreeding these sets of flies was described above and summarized in SM S4. We choose to test maturity of males at ages 6, 7, 8, 9, 10, and 11 days, and for females the ages tested were 0, 1, 2, 3, 4, and 5 days. This was done because a report on age of maturity of *D. hydei* shows that males start to mature at 7 days old (3.4% of examined males) and at 12 days old 95% of examined males have reached maturity; females start to mature at 2 days old (24.3% of examined females), and at 5 days old 95.9% of examined females have reached maturity (Markow, 1985). A portion of newly emerged virgin infected and uninfected G7 males were collected every 12 hours for 72 hours, separated by sex and allowed to reach the ages of 6, 7, 8, 9, 10, and 11 days old. These aged males were individually placed into a vial with 2 10-day-old uninfected females and allowed to interact for 1 hour. After 1 hour the male was removed. The vials containing the mated females were checked after 72 hours of the removal of the male for the presence of larvae. If larvae were present, the male was recorded as mature at the age of mating. To test for age at maturity of females, infected and uninfected flies aged 0, 1, 2, 3, 4 and 5 days old, were placed individually into a vial with 2 10-days old uninfected males and these mating sets were allowed to interact for 1 hour. After 1 hour males were removed. The vials with the mated females were checked after 72 hours of removal of males for the presence of larvae. If larvae were present, the female was recorded as mature at the age of mating. For both, males and females, different sets of flies were use per each age tested. All infected flies used in the experiment were screened after reproduction. If infected flies screened negative for *Spiroplasma*, their records were discarded. The reason for discarding the latter is that the status of those flies is uncertain. They could be truly uninfected (true negative) or they could be infected, but for

diverse reasons (e.g., densities of *Spiroplasma* they carried were too low to be detected, or the presence of PCR inhibitors) the reactions failed (false negative). The situation could be more complicated if said group contained both false negative and true negative individuals.

3.2.10 Longevity and Fecundity Experiments

Infected and uninfected flies from all available isolines, 1, 2, 3, and 5, were used for this experiment. As mentioned in the previous section, part of isoline 2 was used in the age at maturity experiment; another part was used for the experiment described in this section. Also, as mentioned earlier, all G8 isolines were utilized for this experiment. Emerging G8 virgin females were collected every 12 hours for 72 hours. Each collected 0-day-old virgin female was mated with 2 uninfected 10-day old virgin males. Infected and uninfected groups were mated similarly.

The number of eggs produced by each female was counted every 24, 48, and rarely at 72 hours (due to logistic limitations eggs were not counted at fixed periods) until the death of the females. The days when eggs were counted were scheduled to make sure that 7-day egg count intervals were obtained for each experimental female. All vials were checked every 24 hours for survival of males and females. Only female death age was recorded. Dead males were replaced with uninfected virgin males of equivalent age at time of death. It is important to reiterate here that males do not transmit the infection and they were placed with females in this particular step of the experiment to replicate the conditions throughout the entire inbreeding process of the flies, where females were always mated to coeval males of contemporary age. Thorax length (dorsal view) of females was measured at 1 week old to use as covariate for fecundity and survival analysis. All dead females were collected and frozen at -80° C for later screening. SM S5

summarizes the isolines, number of repeats and number of vials per repeats obtained at the end of this experiment.

We used the data gathered in this experiment to analyze the longevity, early fecundity, and lifetime fecundity of flies. Early fecundity was analyzed because we wanted to test if the reduced longevity of infected flies observed in our preliminary experiment was associated with increased early reproductive activity (e.g., early fecundity or early maturation time), as predicted by the theory (Agnew et al., 2000). In addition to that, it is unlikely that *D. hydei* females in the field live up to the older ages of our experimental flies. Factors such as predation, food availability and other environmental factors may cut short the lives of field-dwelling flies, thus their contribution to the population fecundity may not be as high as the total or lifetime fecundity that we anticipated to obtain in our experiments. To address this issue, we wanted to take into account the likely age flies may contribute their maximum reproductive output to the population. Using our data, we chose the fecundity of our flies at an age of 28 days to analyze the early fecundity (as opposed to the fecundity at the age of death for lifetime fecundity). We used two criteria to choose this age for analysis: 1) this age is close to the time when mortality in our experimental populations begins to occur in both treatment groups (uninfected flies begin to die at age 33 days and infected ones at age 32 days); 2) this is the age at which our experimental groups have achieved, on average, about 50% of their overall lifetime fecundity (uninfected flies 49.5%, and infected ones 55.1%).

3.2.11 Statistical Analysis

3.2.11.1 Statistical analysis for male mating preference

Male mating preference was analyzed by using a contingency table chi-square test to test if male mating preference deviated from 1:1 random expectation. The test was used as implemented in IBM® SPSS® v. 21 (IBM Corp, 2012).

3.2.11.2 Statistical analysis for age maturity of males and females

Data on age at which infected and uninfected males and females reach maturity was analyzed by using a two-way ANOVA as implemented in IBM® SPSS® v. 21 (IBM Corp, 2012). Two factors were included in the analysis. One was age of flies divided into two age classes. For males, one age class was named “young” formed by data point measurements for ages six, seven, and eight days old; the other age class was named “old” and it was formed by data point measurements for nine, ten and 11 days old); for females, the age class “young” was constituted by data point measurements for ages zero, one, and two days old and the age class “old” by data for three, four and five days old). The second variable, our variable of interest, was the Infection Status of the flies (uninfected and infected). The dependent variable was the arcsine square root transformation of the proportion of mature flies at each sampled time. Males and females were analyzed separately, because as indicated earlier, males and females reach maturity at different ages (Markow, 1985). Residual analysis was performed to test for the assumptions of the two-way ANOVA. Outliers were assessed by inspection of a boxplot, normality was assessed using Shapiro-Wilk's normality test for each cell of the design and homogeneity of variances was assessed by Levene's test.

3.2.11.3 Statistical analysis for survival

All fly longevity data recorded for our experiment were used for the analysis. The analysis was conducted by using the SAS procedure designed to compute nonparametric estimates of the survival distribution function, PROC LIFETEST. The PROC LIFETEST was ran as implemented in SAS EG 7.1® (SAS Institute Inc., 2014) to obtain the survival function for infected and uninfected females as estimated by the Product-limit (Kaplan-Meier) method (Kaplan & Meier, 1958). The PROC LIFETEST uses three tests to determine if there are significant differences in the survival distributions for the control and experimental groups and we report here the results of two of them, the log-rank test and the Wilcoxon test.

3.2.11.4 Statistical analysis for fecundity

We ran a linear mixed effects model by using the SAS procedure for mixed linear models as implemented in SAS EG 7.1® (SAS Institute Inc., 2014) (hereafter PROC MIXED) to analyze fecundity. The PROC MIXED was used to run separated mixed ANOVAs with “Lifetime Fecundity” or “Early Fecundity” as the response and Infection Status as the predictor; “Isoline” and “Repeat nested within Isolines” were entered in the model as the random factors; the interaction between Isoline and Infection Status was also investigated by entering this interaction as a random factor; the individual “Vials” within the repeats were used as units of analysis (see SM S3). We followed West et al. (2014) to perform the analysis of our data. Because larger females live longer and produce more eggs (Tantawy & Vetukhiv, 1960) we wanted to use Longevity and Female Size (continuous variables) as covariates added to the above model to run an ANCOVA after Wildt and Ahtola (1978). However, upon evaluation of the variables, we found out that Longevity was affected by the treatment (infection) as shown by a one-way

ANOVA ($F_{(1, 72)} = 5.20$, $p = 0.0255$), therefore it did violate the recommendation for the use of a covariate in an ANCOVA, which states that the covariate should be independent of the experimental treatments (Huitema 2011, pp. 209-213; Wildt & Ahtola 1978, p.19.). In addition to this, there are proper ways to analyze longevity data, survival analysis (used in this work) being one of them. The other potential covariate, Female Size, was not used in the analysis because no association (i.e., slopes were not different from zero) was detected between this variable and early or total fecundity for any of our treatment groups (analysis not shown). Because of the above we decided to run first the model without covariates (i.e., not an ANCOVA), to assess the effect of the factor of interest, Infection Status, on fecundity.

Because longevity and fecundity are part of a common life history trade-off (Agnew et al., 2000), we reran the model above including Longevity as a continuous variable to assess if there was an interaction between Longevity and Infection Status.

To check the assumptions for our final model we looked at 1) the normal distribution of the residuals, 2) whether the residual variances are constant across observations, 3) presence/absence of outliers and influential values. I used the conditional raw residuals and the conditional studentized residuals to test for the normal distribution and constant variance of the residuals (West et al., 2014). We obtained a panel of residual graphics by using the option “plots = (residualpanel boxplot influencestatpanel) of proc mixed. We also obtained tables with the residuals by using the option “residual outpred =” We used those plots and residual tables to visually assess normality and constant variance for uninfected and infected flies’ fecundity and the random factors in the model; if non-normality was suspected, a Shapiro-Wilk test for

normality was run to confirm the visual assessment. If non-normality was detected, we used Minitab 17 Statistical Software (2010) to assess and apply a normalizing transformation to the data (Stone, Scibilia, Pammer, Steele, & Keller, 2012) and the fecundity analysis was run again in SAS EG 7.1® (SAS Institute Inc., 2014) with the transformed data. To report confidence intervals, data were back-transformed with Wolfram Alpha V11 (Wolfram Alpha LLC., 2009).

Because likelihood-based methods, such as restricted maximum likelihood (REML) used in this analysis are sensitive to unusual observations, we used influence diagnostics to identify observations that may heavily influence the parameter estimates of fixed and random effects in the model (West et al., 2014). The idea of influence diagnostics for a given observation (or subset of observations) is to quantify the effect of the omission of those observations on the results of the analysis.

We assessed the influence of individual observations by using the “influence” option and its sub-option “iter=” in SAS Proc Mixed. These options produce graphics and an influence diagnostics table that allow visualizing the observations that stand out from the rest of the data points, granting a further examination of such observations as possible influential points.
Two types of diagnostics can be performed by using the graphic and table output above, a “global” and a “specific” diagnostic. The graphics and the table output allow for “global” diagnostics of the influence of individual observations on the overall fit of the model, on the fixed factors as a whole, and on the Covariance parameters (i.e., random factors) as a whole. The graphics output allows for “specific” diagnostic addressing influences for each of the fixed effect estimates and for each of the Covariance Parameter estimates in the model.

The “global” measures selected for our diagnostic (after West et al., 2014) are the following: the “Restricted likelihood distance” (a measure of influence on the overall fit of the model), “Cook’s D” and “Cook’s D Covariance Parameters” (a measure of influence on the point estimate of parameter estimates for fixed and random covariance parameters, respectively) and “Covratio” and “Covratio Covariance Parameters” (a measure of influence on the precision of the parameter estimates for fixed and random covariance parameters, respectively). The “specific” diagnostic (after West et al., 2014) allowed us to explore the influences of individual observations on the fixed factors Infection Status as well as on the Covariance Parameters Isoline, Repeat nested within Isoline and Residual. Two ways to deal with influential values are to explore if the model can be improved by addition of other variables or by running the model without the influential values (Littell et al., 2006, West et al., 2014).

3.3 Results

3.3.1 Results for Male Mating Preference

A total of 32 infected males chose to mate with infected females while 30 chose to mate with uninfected females. In the complementary test, a total of 34 uninfected males chose to mate with infected females, whereas 31 chose to mate with uninfected females. A table of the results of the experiment was entered in IBM® SPSS® v. 21 (IBM Corp, 2012) with the infection status of the males in the first column, and the infection status of the females they chose to mate in the second column; the number of mating of both types was entered in the third column. Because we had total count data, prior to analysis, the data was weighted by the frequencies of types of mating.

A chi-square contingency table test was conducted between male infected status and preference for mating with infected or uninfected females. All expected cell frequencies were greater than five. There was no statistically significant association between infection status of the male and the infection status of the females they choose to mate, $\chi^2_{(1)} = 0.006$, $p = 0.540$. The association was small $\varphi = -0.007$, $p = 0.938$. Fig. 3.1 shows the distribution of mating preferences for infected and uninfected males.

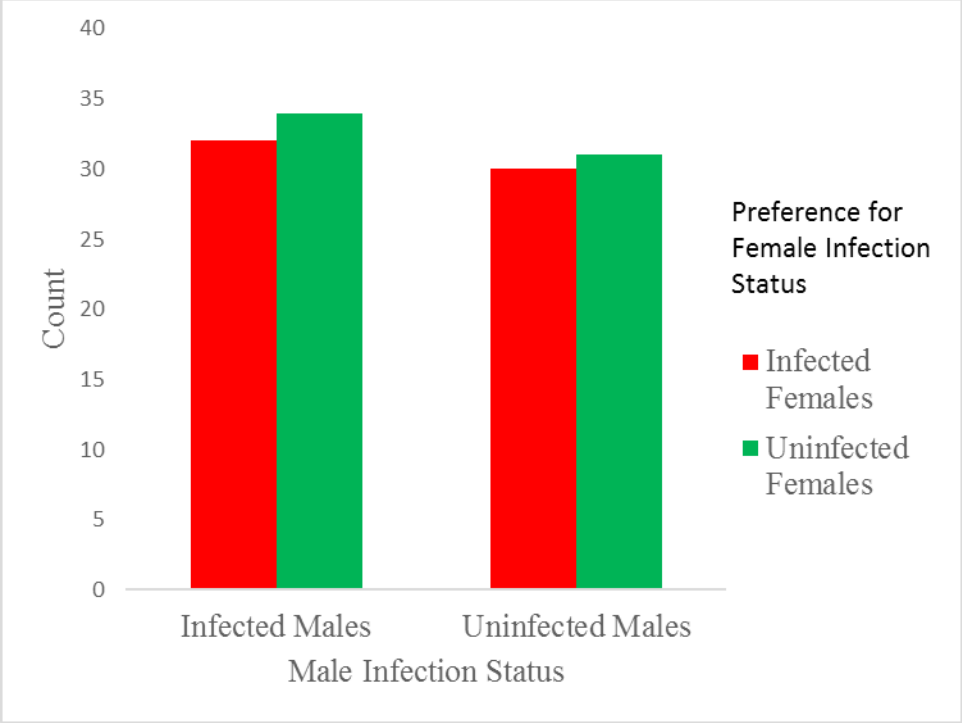

Fig. 3.1 Mating preference of infected and uninfected males for females based on their infection status.

3.3.2 Results for the Age at Maturity for Males and Females

3.3.2.1 Brief description of the age at maturity for males and females

No mature males were recorded at an age of six days, regardless of their infection status. Mature males, both infected and uninfected, were first reported at 7 days of age, and the maximum ratio of mature males for uninfected males was reported at age 9 and for infected males at age 8 days old (Table 3.1, Fig. 3.2). No mature females were reported at zero days of age, regardless of their infection status. The first mature females were reported at one day old for uninfected flies, and at two days old for infected females. The ratio of mature uninfected females reached its maximum at three days of age while the ratio of infected females reached its maximum at five days of age (Table 3.2, Fig. 3.3).

Table 3.1 Age of maturity of males.

Age (Days)	Age (Class)	Infection Status	Number of Mature Males/Total Number of Males	Ratio of Mature Males	Arcsine of Square Root of Ratio of Mature Males
6	Young	Uninfected	0/10	0	0
7	Young	Uninfected	4/14	0.2857	0.5639
8	Young	Uninfected	16/18	0.8889	1.231
9	Old	Uninfected	24/24	1	1.5708
10	Old	Uninfected	20/22	0.9091	1.2645
11	Old	Uninfected	16/18	0.8889	1.231
6	Young	Infected	0/10	0	0
7	Young	Infected	8/18	0.4444	0.7297
8	Young	Infected	21/21	1	1.5708
9	Old	Infected	15/17	0.8824	1.2207
10	Old	Infected	19/20	0.95	1.3453
11	Old	Infected	15/19	0.7895	1.0941

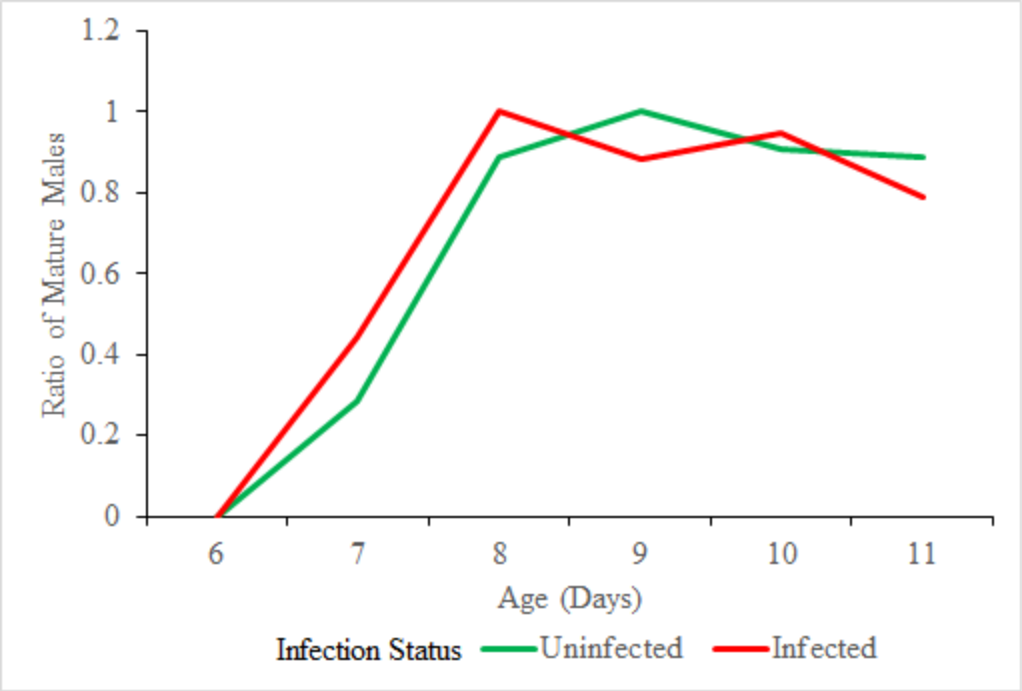

Fig. 3.2 Ratio of mature males from 6 to 11 days old.

Table 3.2 Age of maturity of females.

Age (Days)	Age Class	Infection Status	Number of Mature Females/Total Number of Males	Ratio of Mature Females	Arcsine of Square Root of Ratio of Mature Females
0	Young	Uninfected	0/18	0	0
1	Young	Uninfected	1/16	0.0625	0.2527
2	Young	Uninfected	6/24	0.25	0.5236
3	Old	Uninfected	19/24	0.7917	1.0968
4	Old	Uninfected	19/27	0.7037	0.9952
5	Old	Uninfected	18/24	0.75	1.0472
0	Young	Infected	0/16	0	0
1	Young	Infected	0/27	0	0
2	Young	Infected	3/27	0.1111	0.3398
3	Old	Infected	20/32	0.625	0.9117
4	Old	Infected	18/24	0.75	1.0472
5	Old	Infected	22/26	0.8462	1.1677

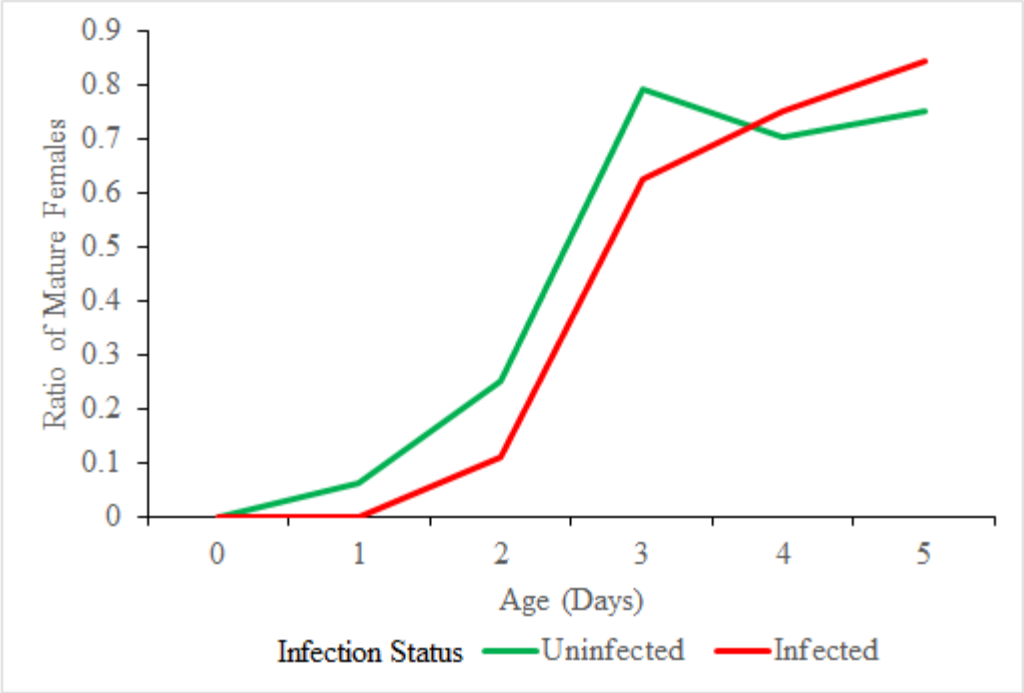

Fig. 3.3 Ratio of mature females from 0 to 5 days old.

3.3.2.2 ANOVA results of the effect of infection status on age at which males reach maturity

Table 3.1 above shows the data gathered for age maturity of males. It is important to clarify that a different set of flies was used for each Age shown in Table 3.1, therefore, the measurements reported in this table do not represent measures taken on the same individuals repeatedly. We analyzed the full data set by running a two-way Anova as implemented in SPSS, with Infection Status and Age Class ("young" coded a "0" and "old" coded as "1") as factors and the arcsine transformation of the ratio of mature flies as the dependent variable. There were no outliers, residuals were normally distributed ($p > .05$), and there was homogeneity of variances ($p =$ 0.219). The interaction effect between Infection Status and Age Class on the arcsine square root transformation of the ratio of mature males was not statistically significant, $F(1, 8) = .264$, $p =$

0.621. There was no statistically significant difference in arcsine square root transformation of the ratio of mature males for uninfected and infected males according to Age Class, $F(1, 8) = 0.003$, $p = 0.957$.

3.3.2.3 ANOVA results of the effect of infection status on age at which females reach maturity

Table 3.2 above shows the data gathered for age maturity of females. It is important to clarify that a different set of flies was used for each age shown in Table 3.2, therefore, the measurements reported in this Table do not represent measures taken on the same individuals repeatedly. There were no outliers, and there was homogeneity of variances ($p = 0.276$). The arcsine square root of proportion of mature females was normally distributed for most of the factor combinations as assessed by Shapiro-Wilk's test ($p > .05$), except for the infected young females, where Shapiro-Wilk's test is significant ($p<.0001$). We decided to continue with the analysis because the dependent variable was already the arcsine square root transformation of the proportion of females and further transformation would obscure the interpretation of the results. Furthermore, it has been argued that ANOVA is robust to violations to the assumption of normality and that even when data are non-normal the actual Type I error rate is usually close to the desired 0.05 alpha level value (Maxwell & Delaney 2004, pp 112-114). The interaction effect between Infection Status and Age Class on the arcsine square root transformation of the proportion of mature females was not statistically significant, $F(1, 8) =.475$, $p = 0.510$. There was no statistically significant difference in arcsine square root transformation of the proportion of mature females for uninfected and infected flies according to age, $F(1, 8) = 0.533$, $p = 0.486$.

3.3.2.4 Analysis of truncated data

After analyzing the age of maturity for both sexes and upon examining Fig. 3.2 (age of maturity of males) and Fig. 3.3 (age of maturity of females), it was evident that more infected males seem to consistently mature at all daily ages than the uninfected, while the reverse is true of females, where it is the uninfected that consistently seem to mature earlier. Males reach maximum maturity by day 8. Females reach maximum maturity by day 3. A question arose on whether the significance of the observed difference was obscured by the identity among those above the age at maximum maturity (ages 9, 10 and 11 for males and ages 4, 5 and 6 for females). In order to clarify this issue, we reanalyzed the data by truncating the male data at age 8 and the female data at age 3.

We used PROC GLM as implemented in SAS EG 7.1® (SAS Institute Inc., 2014) to run a two-way anova, with the arcsine square root transformation of the proportion of mature flies as the dependent variable, and Age (Days) and Infection Status and their interaction as categorical factors. The inclusion of the interaction Age (Days) *Infection Status produced a saturated model, so the interaction was dropped and the analysis was run with only main effects. Our variable of interest, Infection Status, had no effect on age or maturity of males (p=0.2280) or females (p=0.0643). Age did have a significant effect on age of maturity of males (p=0.0145) and females (p=0.0029). For both, male and female age of maturity analysis the assumption of normality was not violated (Shapiro-Wilk test p = 0.1930 for males and p = 0.9686 for females). Evaluation of boxplots for our variable of interest, Infection Status, for males and females separately, indicated the absence of outliers as well as that the equality of variances assumption was approximately met.

3.3.3 Results for Longevity.

3.3.3.1 Descriptive parameters for longevity

As stated in the "Methods" section, we only measured longevity for female flies. Infected flies have a shorter lifespan on average (56.87±15.80 days) than uninfected flies (65.30±19.89 days). The ranges of the lifetimes for the two groups are shown in Table 3.3 (minimum 32 days for uninfected and 31 for uninfected, and maximum 107 days for infected flies and 109 for the uninfected ones).

Table 3.3 Descriptive statistics for longevity (survival) of infected females.

Infection Status	Number of Females	Mean Longevity	Standard Deviation	Minimum Longevity	Maximum Longevity
Infected	47	56.87	15.8	32	107
Uninfected	27	65.3	19.89	31	109

3.3.3.2 Results of statistical analysis for survival analysis

The longevity data for the four fly isolines were combined for the analysis. The analysis was performed on 74 individuals (47 infected flies and 27 uninfected ones). No censored records were present in the data set (Longevity was known for 100% of the individuals studied). Our experimental conditions aimed to avoid deaths due to factors other than the infection status of our flies (all individuals entered in the study at the same time of the year and all of them were kept under the same experimental conditions). The survival distributions for the two treatments are shown in Fig. 3.4. These distributions were statistically significantly different, $\chi^2_{(1)} = 4.4546$, $p = 0.035$ (log rank test).

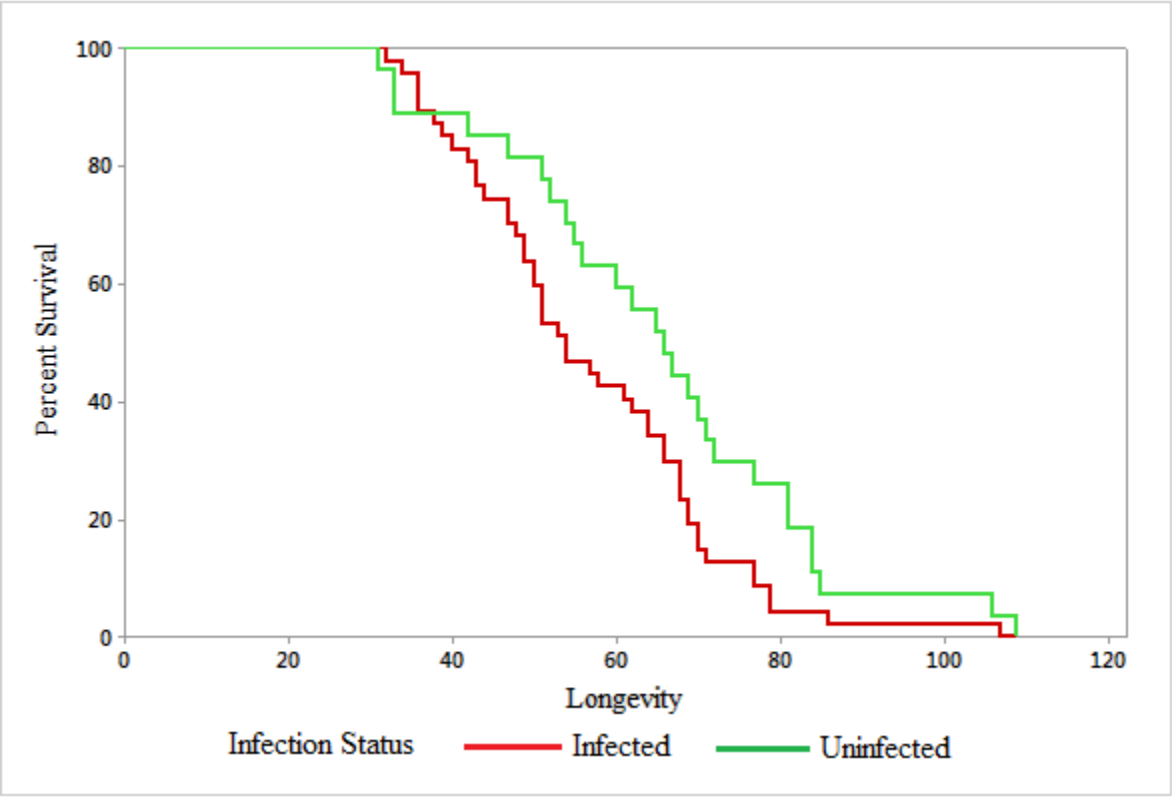

Fig. 3.4 Survival of *Spiroplasma*-infected (n=47) and uninfected (n=27) *D. hydei* females (survival data for all isolines were combined for this analysis, n=74, $p = 0.035$).

3.3.4 Results for Fecundity

3.3.4.1 Description of fecundity tables

Four fly isolines were analyzed in total. The final analyses were performed on 74 individuals (47 infected flies and 27 uninfected ones). Descriptive statistics show that infected flies lay an average of 1697.55±595.45 eggs during their lifetime and uninfected flies lay 1475.37±871.27 eggs in average for the equivalent time (Table 3.4). The minimum number of eggs laid by infected and uninfected flies were 382 and 0, respectively, while the maximum number of eggs laid by those flies were 3537 (infected flies) and 3076 (uninfected flies).

Table 3.4 Descriptive statistics for lifetime and early fecundity (lifetime fecundity measured the number of eggs laid by a female during its lifetime. Early fecundity measured the number of eggs laid from 0 to 28 days old).

Fecundity	Infection Status	Number of Females	Mean	Standard Deviation	Minimum Number of Eggs Laid	Maximum Number of Eggs Laid
Lifetime Fecundity	Infected	47	1697.55	595.45	382	3537
	Uninfected	27	1475.37	871.27	0	3076
Early Fecundity	Infected	47	935.6	264.33	201	1434
	Uninfected	27	717.19	319.56	0	1121

The cumulative fecundity at 28 days of age (Early Fecundity, Table 3.4, Fig. 3.5) for infected flies averaged 935.60±264.33 eggs and for uninfected flies 717.19±319.56 eggs (Table 3.4). The minimum number of eggs laid by infected and uninfected flies were 201 and 0, respectively, while the maximum number of eggs laid were 1434 (infected flies) and 1121 (uninfected flies).

We collected data that allowed us to divide fecundity for our two experimental groups of flies into 7-day intervals (Fig. 3.5). Fig. 3.5 shows that there is a general trend for infected flies to lay more eggs than uninfected ones. Examining this figure, we find that there is a period of higher fecundity from [08-14] up to the [22-28]-day interval for both experimental groups. The low fecundity from [00-07] days old is explained by the fact that females mature at different ages, so they do not start laying eggs synchronously from day zero. For our data, the first four intervals in Fig. 3.5 (from 0 to 28 days old age) represent the time when, on average, flies have contributed around 50% of the 77-day truncated lifetime fecundity of their respective treatment groups (infected 55% and uninfected 50%). This is the period that we defined as Early Fecundity in our analysis. From age 29 to the last interval analyzed [71-77] (i.e., 77 days old) there is a period of

lower fecundity for uninfected and infected flies (Fig. 3.5). This means that most flies tend to lay fewer eggs as they age.

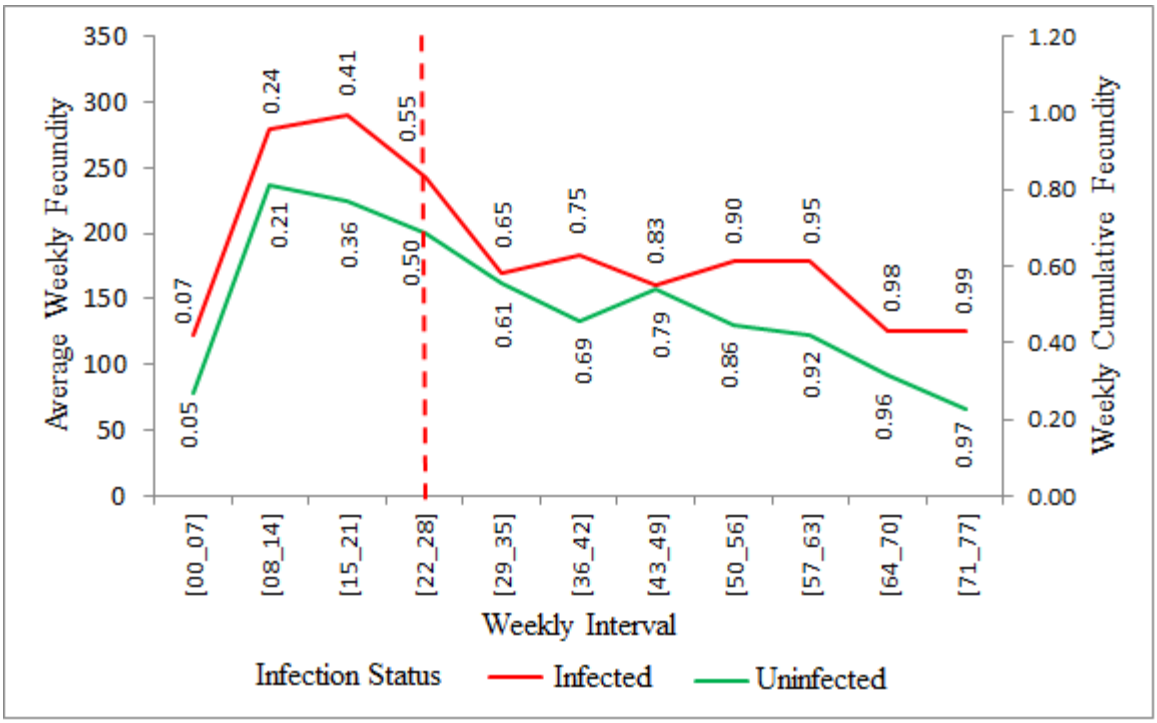

Fig. 3.5. Weekly average fecundity (primary axis) and percentage cumulative fecundity (secondary axis) of infected and uninfected flies. The labels show the percentage cumulative fecundity at each weekly interval (estimated as: overall $week_n$ fecundity/lifetime fecundity truncated at 77-days). The red vertical broken line shows the point in time when flies have cumulatively contributed to the population about 50% of their lifetime reproductive output truncated at 77-days (early fecundity).

3.3.4.2 Results of statistical analysis for fecundity

3.3.4.2.1 Total or lifetime fecundity analysis

For lifetime fecundity, the data met the assumption of normality and equality of variances and no obvious outliers or influential values were diagnosed upon using the diagnostic tools described in

the methods section (not shown). There was no statistically significant difference in total fecundity between infected and uninfected flies ($F_{(1,16)} = 0.47$, $p = 0.5030$) (Table 3.5). For the random factors, the interaction of Isoline with Infection Status was removed (as recommended in SAS manual) because the parameter estimate was zero, leaving only Isoline and Repeat(Isoline) (a nested factor) in the analysis, which are not significant ($p = 0.208$ and $p = 0.0652$, respectively) (Table 3.5).

Table 3.5 Total or lifetime fecundity analysis.

Fixed Factor	Estimate (b)	Standard Error	*p*
Intercept	1627.1	197.45	0.0005
Infection Status (Uninfected)	-146.64	214.02	0.503
Random Factor (Covariance Parameter Estimates)	Estimate (b)	Standard Error	*p*
Isoline	76383	93919	0.208
Repeat(Isoline)	123475	81618	0.0652
Residual	321696	63194	<0.0001

3.3.4.2.2 Early fecundity analysis

For Early Fecundity (cumulative fecundity at 28 days old), the combined data for the two groups (uninfected and infected) failed the normality test for the residuals (Shapiro-Wilk $p<0.029$). Upon a separate test for each group, infected flies were found not to be normal (Shapiro-Wilk $p<0.0058$). We explored a normalizing transformation for the whole data by using Minitab 17 Statistical Software (2010), and as a result of it, we transformed the data using the unbounded form (SU) Johnson transformation [AD = 0.258, $p = 0.709$; transforming equation is "transformed values" = 0.470930 + 1.05689 *(Asinh ((x-995.866) /197.807)), where "x" is the

original untransformed value]. Because our main interest was to test if our treatment (Infection Status) had an effect on Early Fecundity, we first compared a model with only fixed factors versus a model that includes our nested random factors defined in the methods section. We ran a likelihood ratio test that indicated that the model including the random factors had a significantly reduced -2 Res Log Likelihood as compared to a fixed-factor-only model ($\chi^2_{(3)} = 12$, $p = 0.0037$), therefore, we kept the model with fixed and random factors for further analysis. Next, we tested whether we needed a model having homogeneous residual variances (above model) or heterogeneous residual variances. A non-significant likelihood ratio test ($\chi^2_{(1)} = 1.4$, $p = 0.2367$) indicated that none of the models were better, so we opted for keeping the simpler model (homogeneous residual variance) for further analysis. The covariance parameter estimates for Isoline and the interaction of Isoline with Infection Status were zero and we removed them from the final model. The final analysis indicates that Infection Status has a statistically significant effect on Early Fecundity ($F_{(1,16)} = 0.47$, $p = 0.0281$) (Table 3.6).

Table 3.6 Early fecundity analysis of Johnson transformed data.

Fixed Factor	Estimate (b)	Standard Error	*p*
Intercept	0.3193	0.2062	0.1378
Infection Status (Uninfected)	-0.7723	0.342	0.0281
Random Factor (Covariance Parameter Estimates)	Estimate (b)	Standard Error	*p*
Repeat(Isoline)	0.4047	0.1912	0.0172
Residual	0.5707	0.1117	<0.0001

Table 3.6 shows that the effect of the infection status “uninfected” is negative (“b” estimate = -0.7723) and significant, which means that the Early Fecundity of uninfected flies is significantly lower than the Early Fecundity of infected flies. The variance of the random effect “Repeat(Isoline)” is significantly greater than zero ($p = 0.0172$), which indicates that there is an amount of random variation in the response values at this nested factor that remains unexplained (West et al., 2104; Dickey, 2007). Table 3.7 reports the Johnson transformed and back transformed estimated Early Fecundity means and 95% CI for our two study groups.

Table 3.7 95% CI for least square means estimates for the early fecundity of infected and uninfected females. (*The estimate mean shows the mean number of eggs laid by a female from 0 to 28 days of age).

Infected	Estimate Mean*	Lower	Upper
Johnson Transformed	0.3193	-0.1118	0.7504
Back Transformed	967.39	884. 29	1045.4
Uninfected	Estimate Mean*	Lower	Upper
Johnson Transformed	-0.4531	-1.0239	0.1178
Back Transformed	800.04	622.48	923.89

The diagnostics criteria used for this analysis indicated that the residuals met the assumption of normality (Graphics not shown and Shapiro-Wilk test, uninfected flies $p = 0.4480$, infected flies $p = 0.0871$, combined $p = 0.2938$) and equality of variances (the standard deviations of the residuals were similar for the infected (0.6221) and uninfected (0.7733) groups, indicating no strong evidence of non-homogeneity of variance); the box plots of the conditional studentized residuals show that the distribution is approximately homogeneous across fly isolines and repeats

nested within fly isolines. The graphics used for influence diagnostics seemed to indicate two values (observations identified by ID numbers 10, from the uninfected groups, and 14, from the infected group) stand out from the rest of observations. However, when those observations were removed and the analyses rerun without them (removing with replacement and removing the two observations at the time) the results general conclusions from the analysis did not change and we report here the results of the full data analysis.

3.3.4.2.3 Interaction between infection status and longevity and its effect on fecundity

For lifetime fecundity the linear mixed model analysis shows the interaction Infection Status* Longevity does not have a significant effect on fecundity, but longevity has a significant effect on lifetime fecundity (Table 3.8). The test results for covariance parameter estimates for "Repeat nested within Isoline" ($p>0.0307$) and "Residual" ($p<0.0001$) indicate that, after controlling for the fixed factors Infection Status, Longevity and their interaction, significant variation between and within the random factor "Repeat(Isoline)" lifetime fecundity still remains to be explained (Table 3.9). A regression analysis shows that longevity and lifetime fecundity are positively associated ($p<0.0001$) for the full set of flies, which indicates that flies that live longer lay more eggs during their lifetime than flies that live shorter times. Separate regression analyses for the lifetime fecundity of infected and uninfected flies vs. Longevity indicates that they are also positively associated (infected has a $p = 0.0001$ and uninfected has a $p = 0.0001$).

Table 3.8 Fixed effects test for lifetime fecundity vs infection status and longevity for the full set of flies used in this study.

Effect	Num DF	Den DF	F Value	Pr > F
Infection Status	1	66.8	1.69	0.20
Longevity	1	70	85.3	<.0001
Longevity*Infection Status	1	70	0.02	0.88

Table 3.9 Covariance parameter estimates for lifetime fecundity vs infection status and longevity for the full set of flies used in this study.

Covariance Parameter	Estimate	Standard Error	Z Value	Pr > Z
Repeat(Isoline)	82125	43896	1.87	0.0307
Residual	151387	30177	5.02	<.0001

For Early Fecundity we used the Johnson transformed fecundity because the residuals of the untransformed fecundity were not normal (see previous section “Early Fecundity analysis”). The linear mixed model analysis shows that there is a significant interaction between Infection Status and Longevity ($p=0.0314$) (Table 3.10 and Fig. 3.6). The test results for covariance parameter estimates for “Repeat nested within Isoline” ($p>0.0231$) and “Residual” ($p<0.0001$) indicate that, after controlling for the fixed factors Infection Status, Longevity and their interaction, significant variation between and within the random factor means Early Fecundity still remains to be explained (Table 3.11). Fig. 3.6 shows that the interaction of Infection Status and Longevity is determined by the fact that infected flies that have higher Early Fecundity live shorter than infected flies that have lower Early Fecundity, while the uninfected flies show the opposite trend, with flies that live longer lives being more fecund than flies that live shorter lives.

Table 3.10 Fixed effects test for early fecundity (fecundity at 28 days of age) vs infection status and longevity for the full set of flies used in this study.

Effect	Num DF	Den DF	F Value	Pr > F
Infection Status	1	67.4	9.16	0.0035
Longevity	1	69.6	1.3	0.258
Longevity*Infection Status	1	69.6	4.82	0.0314

Table 3.11 Covariance parameter estimates for early fecundity (fecundity at 28 days of age) vs infection status and longevity for the full set of flies used in this study.

Covariance Parameter	Estimate	Standard Error	Z Value	Pr > Z
Repeat(Isoline)	0.3572	0.1792	1.99	0.0231
Residual	0.5483	0.1095	5.01	<.0001

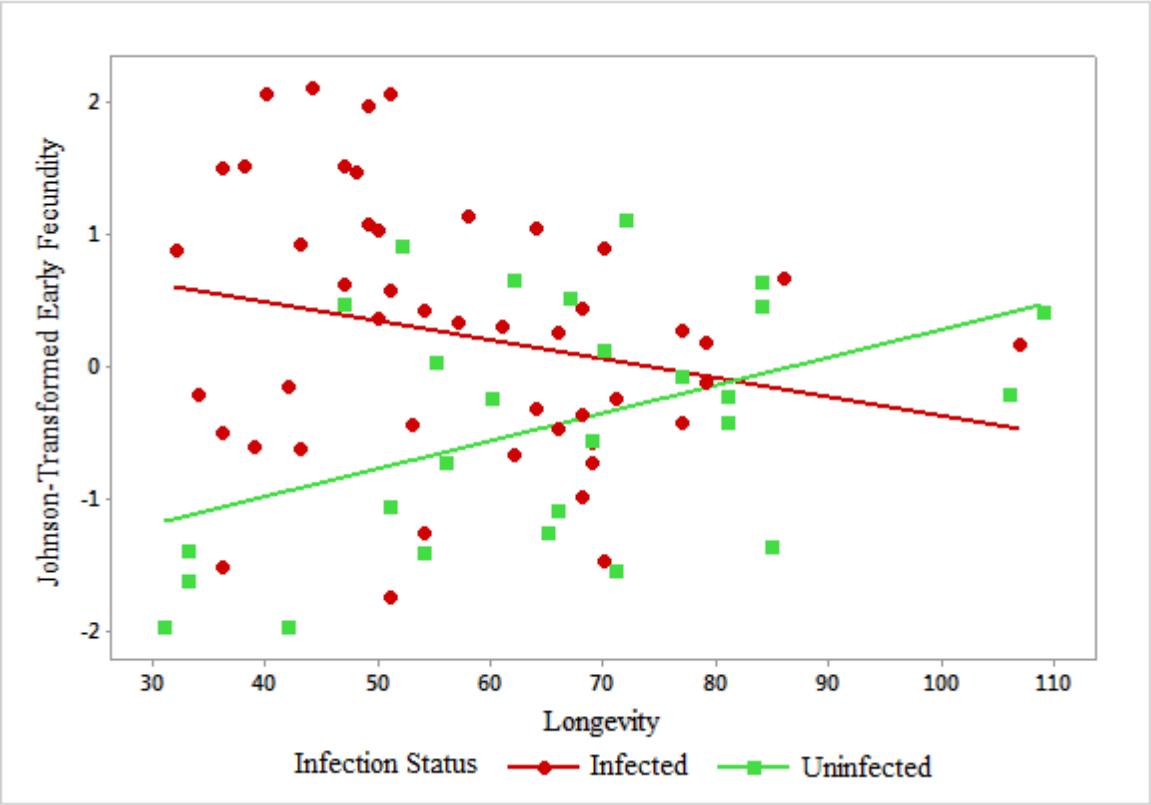

Fig. 3.6 Interaction between the fixed factors infection status (infected flies and uninfected flies) and longevity as related to early fecundity of flies (Johnson transformed cumulative fecundity at 28 days of age).

To explore the interaction between Infection Status and Longevity, we ran simple linear regressions to predict Early Fecundity based on Longevity for infected and uninfected flies separately. A linear regression of the full data for infected females established that the Longevity of this group could not predict their Early Fecundity, $F(1, 45) = 8.16$, $p = 0.1195$. For the full data of the uninfected group, a linear regression analysis established that Longevity could statistically significantly predict Early Fecundity, $F(1, 25) = 6.46$, $p = 0.0176$. An analysis of the assumptions of regression detected 4 observations that were outliers and leverages for the infected group (Fly ID = 14, 32, 39, 52) and 4 for the uninfected group (Fly ID = 03, 07, 22, 23).

Upon removal of those observations, separated simple regression analyses for infected and uninfected flies indicate that Longevity could statistically significantly predict Early Fecundity for infected flies, $F(1, 41) = 8.16$, $p = 0.0067$ and for uninfected flies, $F(1, 21) = 5.12$, $p = 0.0344$. Longevity accounted for 16.60% of the explained variability in Early Fecundity for infected females and 19.60% for uninfected females. The regression equation for the infected group was: predicted Early Fecundity = 1.9238 – 0.0286 x (Longevity); the regression equation for the uninfected group was: predicted Early Fecundity = -2.1016 + 0.0253 x (Longevity).

3.4 Discussion

3.4.1 Discussion of Mating Preference

Neither infected nor uninfected males discriminate between infected and uninfected females at time of mating in a paired mate choice test. Thus, we reject our hypothesis that infected and uninfected males preferentially mate with infected females. Our rationale to test this and all hypothesis in this work was that, in general, maternally transmitted endosymbionts increases the fitness of their hosts, therefore increasing the likelihood of being passed to the next generation. We had preliminary information that showed that uninfected females had a biased sex ratio (this later was found out to likely be a bias introduced by the data collection method), so there was a likely excess of males in a population where infected and uninfected flies existed together. Consequently, if infected and uninfected males mated preferentially with infected females, then, the latter would have most of the males available to mate and obtain an advantage at the time of reproduction over uninfected females, which would have a diminished access to mating partners. For this scenario to happen there is no need for all males to exclusively mate with infected females or for uninfected females to go permanently unmated in the population. This is because

D. hydei males and females re-mate several times per day (Markow, 1985) and Markow, Beall, and Castrezana (2012) have argued that this elevated female re-mating is critical to maintain female fertility. In addition, *D. hydei* males do not discriminate between virgin and non-virgin females for mating purposes (Markow, 1985), so a single female can be mated multiple times per day. We propose that under these circumstances, if any level of mating preference to infected females exist in the *D. hydei* populations, it would reduce (but not permanently eliminate) the mating chances of uninfected females. Because *D. hydei* females can lay large number of unfertilized eggs (Markow, 1985), it is not out of the realm of possibility that unmated uninfected females lay a clutch of unfertilized eggs and will not mate again until more eggs mature. This would reduce, albeit for short periods of time, the number of breeding females in the population, and this reduced number of reproductive females available for mating would be made up of more infected than uninfected females. This would effectively increase the frequency of infected flies (# infected flies/# of flies sampled), produced in such a population, although the total number of progeny born into that population would be reduced.

We did not statistically detect mating preference in our experiment. However, we can present a simple example to clarify what we propose would happen in a population under mating preference for infected females. Assume a population composed of 5 infected females with a transmission rate of 0.90 and 5 uninfected females (0% transmission). Furthermore, assume that each female, irrespective of its infection status, produce 5 female daughters. With no mating preference they would produce 50 daughters in total, from which 25x0.9= 22.5 (45%) are infected and 27.5 (55%) are uninfected flies. If, at any given time, we assume a mating preference for infected females, all 5 infected females would be mated, but not all uninfected

females. For the sake of this example, let's assume that only 4 uninfected females are mated. Under these new conditions, the 5 infected females will produce 25 daughters (2.5 uninfected and 22.5 of the daughters infected as before) and the 4 uninfected females will produce 20 uninfected daughters (making up a total of 45 daughters). In this new set of daughters, the 22.5 infected females will make up 50% of the population (a 5% increase of infection frequency from 45% to 50%). If increased fecundity would be an additional effect of the infection on the host, the 5-daughter deficit described above will be compensated by the increased number of daughters produced by infected females.

Because our results do not support our hypothesis we will briefly explore several reasons that can explain this. The default explanation for these results is that the bacteria do not have an effect on male mate choice as suggested by our result.

3.4.1.1 Spiroplasma cannot modify the levels of fly sex hormones to elicit mating preference

In *D. melanogaster*, commensal gut bacteria have been reported to produce positive assortative mating in populations grown in two different media (a corn-molasses-yeast-based media vs. a starch-based media), with flies mating preferentially within their own group as early as after two generations of the groups growing on separated media (Sharon et al., 2010). The "at least in part" bacteria responsible for the assortative mating was identified as the *Lactobacillus plantarum*, a bacterial species that lives extracellularly in mammals, forming a biofilm firmly attached to the epithelium of the gastrointestinal tract (Walter, 2008). Sharon et al. (2010) hypothesized that the mechanism for the assortative mating could be a volatile compound secreted by the host and modified by the bacteria or a compound released by the bacteria *per se*. Further experiments

showed that the two groups of flies showed significant differences in the levels of five pheromones and those differences were reduced after antibiotic treatment, as did the assortative mating (ibid). The lack of mate choice by males in our experiment may be explained by the fact that although *Spiroplasma* occupies the guts of *Drosophila* (Anbutsu & Fukatsu, 2006), like the *Lactobacillus* does in the study by Sharon et al. (2010), the former bacteria is not able to modify the levels of fly sex pheromones production, hence it is not able to prompt male mate choice through this mechanism.

3.4.1.2 Culture conditions eliminated body size as a cue used for males to choose mates

D. hydei seems to be a good candidate for the existence of male mate choice. However, it is possible that infection does not produce the cues involve in mate choice. According to the theory of male mate choice evolution, mate choice is likely to arise in promiscuous systems of reproduction such *as D. hydei*'s, where a fly has multiple mating partners with no permanent association of male-female pairs. In such systems, males are expected to choose females based on their higher fecundity, assessed through their larger body size (Bonduriansky, 2001). Larger body size is an indicator of higher fecundity in *Drosophila* (for *D. melanogaster* see Lefranc and Bungaard, 2000; for *D. psuedoobscura* see Tantawy and Vetukhiv, 1960). However, in our experiments conducted in parallel with the mate choice test, body size was not correlated to fecundity, as shown by a set of flies cultured at low density (t-test comparing sizes of infected and uninfected females is not significant, $t_{(40)}=0.82$, $p = 0.415$). We attribute this lack of size-fecundity correlation to the fact that our culture methods (controlling the number of breeding flies to control density) were effective at standardizing female size across our treatments. This may have deprived the males of the cue they use to assess fecundity of females.

3.4.1.3 Cryptic mate choice

Our experiment did not address cryptic mate choice by males. Males can allocate variable amount of resources (e.g., amount of ejaculate) according to the female quality and/or when sperm competition exists between males (Bonduriansky, 2001). *D. hydei* males experience sperm competition because females (and males) mate repeatedly and there is mixing of the sperm from mating males in the female's reproductive tract (sperm storage structures) (Markow, 1985). Given that under those circumstances males do not deposit all their sperm in a female (Bonduriansky, 2001), there is the possibility that both infected and uninfected males depositing larger amounts of sperm on the more fecund females (infected ones at age of 28 days) as a mechanism to outcompete their rivals. If more sperm are deposited in more fecund females (the infected ones) at higher frequencies than in uninfected females, then more infected eggs will be fertilized at the population level. This could contribute to the persistence of the infection in *D. hydei* populations through the mechanisms proposed at the beginning of this section and increased fertility of infected females. Male cryptic mate choice in this system could be addressed in future investigations in the lab or in the field. For such a study, lab or wild-caught mated females should be screened for the presence or absence of the infection. If males mate preferentially with infected females, the count of the number of sperm present in the female reproductive tract should be higher in infected females when compared to uninfected females.

3.4.1.4 Alternate methodological approaches might have allowed us to detect male mate choice

For example, the effect of infection in male mate choice could be weak and a large sample size may be needed to detect it. We were able to use a sample size of 127 mating data points. The variability in infection frequencies of *Spiroplasma* in our *D. hydei* laboratory isolines (from

<0.45 to ~0.9, pers. obs.) made it logistically difficult to gather a larger data set for this study. The difficulty is illustrated by the requirement that the infected male and female used in each set of 4 mating flies (see Methods section) should screen positive for *Spiroplasma*. If any of the potentially infected flies (screened post-experiment) failed to produce a positive result for the bacteria, the whole set had to be eliminated from the analysis.

If the mate preference is weak or age dependent, our experimental design could have made it even more difficult to detect it. Our sampling focused on 10-day-old males and females, and at this age, *Spiroplasma* density inside *D. hydei* is lower with respect to the levels it will reach later (Haselkorn, Watts, & Markow, 2013; Kageyama et al., 2006). Initially, we planned to use 3-day-old females and 10-day-old males, which could have made mate choice more difficult to detect if it is age dependent. However, we tested 10-day-old flies due to logistic considerations. The mate choice tests were conducted parallel to several other experiments reported in this document. Ten-day-old flies of both sexes were readily available as part of the regular maintenance of laboratory flies, whereas, handling several experiments at the same time made it difficult to match flies of any other desired ages.

After analyzing our data, we recommend a combination of two criteria for re-testing mating preference in future studies. First, mating preference should be tested when fecundity is at its maximum in the population. This may maximize the advantage of preferential mating by ensuring the maximum output of eggs after those couplings. Second, *Spiroplasma* densities within the hosts should be at high, so that any effect on mating preference would be strong enough to be detected experimentally.

Considering the above, we propose to test for mating preference between 3 and 4 weeks of age, when females show a higher peak of fecundity (see Fig. 3.5). Also, at that age, *Spiroplasma* densities would be higher than at our tested age of 10 days old. Certainly, Haselkorn et al. (2013) measured *Spiroplasma* density in laboratory *D. hydei* flies from third instar larvae to 13 days after emergence, and at the latter age *Spiroplasma* densities were still on the rise. Similarly, Kageyama et al. (2006) showed that in females 4 and 5 weeks of age *Spiroplasma* density was still increasing. Our data on the fecundity experiment shows that 21 to 28-day-old flies are at the peak of their reproductive activity. The longevity of flies used in our experiments was about twice that age (i.e., flies lived 57 and 65 days on average if they were infected and uninfected, respectively, Table 3.3). Testing male mating preference at 28 days of age seems a reasonable female age that would meet the two criteria prosed above, maximum fecundity of females and high *Spiroplasma* density inside the flies. In the field, it is likely that 28-day-old females may still be able to compete for mates with younger females, because laboratory experiments show that males do not have a preference for virgin or recently mated females (Markow, 1985). This suggests that 28-day-old females are part of the pool of females with whom males can choose to mate.

3.4.2 Discussion of Infection and Age of Maturity for Males and Females

The age of maturity of males is not statistically significantly different regardless of their infection status, although 6 to 8-day-old infected males show a small increase in rate of maturity with respect to uninfected males (Fig. 3.2). The same holds true for females, although here the trend is reversed, with a small increase in time to maturity for infected females (Fig. 3.3). We tested the hypothesis that the infection status of the flies affected time to maturity because

preliminary data showed that *Spiroplasma* infection shortens the lifespan of infected females. The theory predicts that when the host cannot suppress an infection through the action of their immunological system, infected organisms that accelerate their schedule of reproductive investment will be favored by natural selection (Agnew et al., 2000). This is especially true in systems where the parasite reduces the future reproductive success of the host, in which case the host will try to gain as much reproductive success as they can before the cost of the infection renders reproductive success null (ibid). In the *Drosophila-Spiroplasma* system, it has been reported that *Spiroplasma* is a parasite that does not elicit an immunological response from the host. Hurst, Anbutsu, Kutsukake, and Fukatsu (2003) reported that a *Spiroplasma* strain, NSRO seems to survive undetected by the immunological system of *D. melanogaster*. The authors of the above study argued that the lack of cell wall (whose components are elicitors of immune responses) in *Spiroplasma* was the cause for this phenomenon. The fact that the immunological system of the fly does not hinder the proliferation of the bacteria was confirmed by the observation that the growth of the *Spiroplasma* strain MSRO is apparently self-regulated inside its natural host *D. melanogaster* (Herren & Lemaitre, 2011; Herren et al., 2014) and a later study showed that the proliferation of the bacteria is limited by the availability of host hemolymph lipids, but not by the immunological system of the host (Herren et al., 2014). The proliferation the bacteria decreases in old (and nutrient deprived) flies, which show decreased hemolymph lipids concentrations (Herren et al., 2014). It is likely that the same phenomenon occurs in *D. hydei*, where *Spiroplasma* has been reported to continually increase inside the female host until 5 weeks of age (Figure 3 of Kageyama et al., 2006), which is the last age reported in that study. Kageyama et al. (2006) seems to indicate that like *D. melanogaster*, *D. hydei* cannot suppress *Spiroplasma* infection, which makes it a candidate to fulfill Agnew et al. (2000) prediction about

bringing forward its reproductive schedule in response to the infection. Here, the slight delay to maturity in females (which is not statistically significant according to our analysis) would seem inconsistent with increased early age reproduction; however, we observed a significant increase in early age fecundity, which fits the prediction well. Infection rates in males would seem to have no effect, since males do not transmit infection. However, if infection rates among males is a function of transmission rates in females, they would contribute indirectly favoring lines (genotypes) more susceptible to transmission, passing more copies of the high transmission rate genotype to the next generation through the early age to maturity (a non-statistically significant trend observed in our data) and (untested) increased fertility of infected males. Infected females will produce more offspring, with the proportion of infected offspring dependent upon the transmission rate, while infected males will pass on more of the genotype associated with higher rates of transmission.

There are some studies that seem to support the shift forward of reproductive schedule in *Spiroplasma* infected *Drosophila*. These studies focused on the effect of the infection on development time or reproductive activity of infected flies, but not on longevity. For example, *D. nebulosa* females infected by its naturally occurring strain of male-killing *Spiroplasma* (NSRO) have been reported to mature a day earlier than uninfected ones (Malogolowkin-Cohen & Rodriguez-Pereira, 1975). Faster development time of *D. melanogaster* larvae infected with its native male-killer *Spiroplasma* at two different densities was reported by Martins, Ventura, and Klaczko (2010). Martins et al. (2010). The authors of the latter study argued that density reduction through killing half of the progeny (sons) could have caused those differences, because more resources were available to female larvae; however, they could not rule out a direct effect

of the male killing bacterial infection in the differences observed (ibid.). From the reports mentioned above, it seems that the effect of *Spiroplasma* infection on *Drosophila* age of maturity and developmental time is not very strong. Infected flies seem to be faster in maturing and developing by only about a day compared to uninfected flies (Malogolowkin-Cohen & Rodriguez-Pereira, 1975; Martins et al., 2010). Contrary to those studies, our results indicate that *Hyd1* does not significantly affect the age of maturity of infected *D. hydei* flies by even this amount. This suggests that *Spiroplasma* may not drastically reduce the future reproductive success of infected *D. hydei* flies, which in turn means that infected flies do not need to prominently accelerate their reproductive schedule (i.e., bring forward the age of maturity).

3.4.3 Discussion of Longevity and Fecundity

The survival analysis shows that *Spiroplasma*-infected *D. hydei* females live shorter than uninfected ones, with a log rank test significance of $p < 0.0348$). In our preliminary study (chapter 1 of this document), we obtained a similar result, but with a smaller significance ($p = .004$, log rank test). In the previous experiment we had a larger sample size for analysis, which undoubtedly reduced the variability in the data set. An analysis of the combined data set for the current and previous experiment shows that the survival distributions for the two treatments are statistically significantly different, $\chi^2_{(1)} = 11.1827$, $p < .0008$ (log rank test) (Table 3.12, Fig. 3.7). Taken together, our results for survival analysis strongly prove that *Hyd1 Spiroplasma* infection shortens the length of life of females of their natural host *D. hydei*.

Table 3.12 Descriptive statistics for the survival data for our two experiments.

Experiment	Infection Status	Number of Females	Mean	Uninfected -Infected	95% CI	*p*-value (log rank)
Chapter 1	Infected	49	55.88	7.34	53.11 to 58.64	0.004
	Uninfected	45	63.22		59.33 to 67.12	
Chapter 2	Infected	47	56.87	8.43	52.36 to 61.40	0.035
	Uninfected	27	65.3		57.80 to 72.80	
Pooled	Infected	96	56.36	7.64	53.75 to 58.98	0.0008
	Uninfected	72	64.00		60.30 to 67.70	

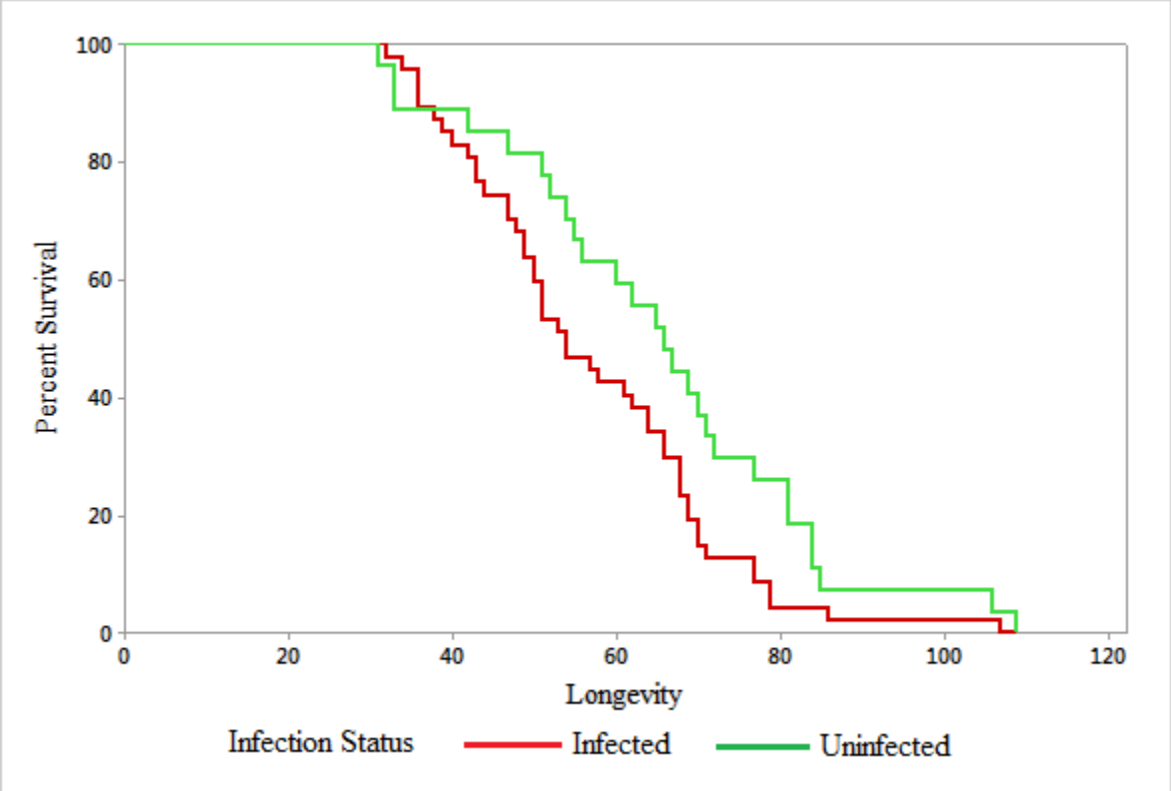

Fig. 3.7 Survival analysis for the combined data set of our two experiments showing that the survival of *Hyd1 Spiroplasma*-infected *D. hydei* females (infected) is lower than the survival of uninfected females (number of females tested: infected=96; uninfected=72; p = 0 .0008).

Our mixed effects analysis shows that infected females albeit shorter lived, have a significantly. (p=0.0281) higher Early Fecundity than uninfected females (Table 3.6), while lifetime fecundity is not different between infected and uninfected females (p = 0.503) (Table 3.5). Taken together, the longevity and fecundity analyses indicate that *Spiroplasma* infection shortens the lifespan of *D. hydei* females and increases their Early Fecundity, with little if any difference in lifetime fecundity. The analysis of the interaction effect between Infection Status and Longevity on lifetime and Early Fecundity further supports this view (Table 3.8, Table 3.10 and Fig. 3.6). There is no effect of the interaction between Infection Status and Longevity on lifetime fecundity (p = 0.88) (Table 3.8), but there is a significant effect of such interaction on Early Fecundity (p = 0.0314) (Table 3.10 and Fig. 3.6), with infected females living shorter than uninfected females but laying more eggs than uninfected flies at 28 days old (Fig. 3.6).

It can be argued that the shorter lifespan of flies showing higher early fecundity could be due to their increased reproduction, and not *Spiroplasma* infection, given that reproduction has negative effects on longevity (e.g., Partridge, 1986). The survival of the females used in chapter 1 offer an insight into this question. Those females were set up at young ages (4, 5 and 6 days old) and mated for 5 days to collect data on fecundity. After 5 days, the males were removed and the females were followed to collect data on their longevity. Those females lived on average 55 days (infected females) and 61 days (uninfected females). Five days represent 1/10 of the average lifespan of the infected females and about 1/12 of the lifespan of the uninfected. Being unmated for a large part of their lives would, in theory, have removed the cost of reproductive activity on longevity, but still, those unmated infected females lived shorter lifespans than the unmated uninfected ones.

It is important to notice that the difference in means for Longevity of infected and uninfected females is not very large. For any of our experiments or the pooled data for the two experiments; the difference in survival is not very large and only about 7 to 8 days (Uninfected - Infected in Table 3.12). A similar situation occurs with the difference of fecundity means at each of the 7-day intervals, where the differences in favor of infected females vary from 7% or 12 eggs (Fecundity Interval [29-35] days) to an exceptional 42% or 52 eggs ([00-07] days interval) (Table 3.13). For Early Fecundity (cumulative fecundity at 28 days old), the difference is about 30% in favor of the infected females ((935-717)/717, raw data) (Table 3.4). The rather small differences found in our experiments suggest that the effect of *Hyd1* on *D. hydei* fitness is not strong.

Table 3.13 Mean fecundity comparison at 7-day intervals from 0 to 70 days between *Hyd1 Spiroplasma*-infected *D. hydei* females and uninfected controls.

Fecundity Interval (Days)	Infected Flies Fecundity Mean	Uninfected Flies Fecundity Mean	Difference Fecundity Means Infected-Uninfected	Percentage Difference Means Infected-Uninfected
[00-07]	122	70.33	51.67	42.35
[08-14]	280.09	236.44	43.64	15.58
[15-21]	290	217.15	72.85	25.12
[22-28]	243.51	193.26	50.25	20.64
[29-35]	170.02	158.11	11.91	7.01
[36-42]	183.6	135.04	48.56	26.45
[43-49]	160.58	174.44	-13.86	-8.63
[50-56]	178.67	148	30.67	17.16
[57-63]	179.27	137.06	42.21	23.55
[64-70]	125.22	105	20.22	16.15

Our results seem to be in line with the theory that vertically transmitted parasites should have low harmful effects on the host lifetime reproductive success because the reproduction, and therefore, survival of the host contributes to the parasite reproduction (Lipsitch, Siller, & Nowak, 1996). If hosts can adjust their life history traits to negative effects on their fitness caused by parasites, then it follows that those small fitness effects of infection will produce small fitness increases in the host. Small or not very large effects on host fitness have been reported for other *Spiroplasma-Drosophila* associations. Indeed, Haselkorn (2010b) found "many small and statistically significant effects of *Spiroplasma* infection" on several fitness components of *D. mojavensis*. Martins et al. (2010) found that *D. melanogaster* infected by a male-killing strain showed a 1-day shorter significant difference in developmental time from egg to emergence when compared to uninfected flies. Malogolowkin and Rodrigues (1975) reported that *D. nebulosa* infected with its natural *Spiroplasma* NSRO mature and mate at least 1 day earlier than uninfected females. Herren et al. (2014) upon studying fecundity and longevity of MSRO-infected *D. melanogaster* concluded that the infection has a low impact on the general fitness of its host.

To contrast what we call strong effects of parasites on the fitness of the host we look to systems other than the *Spiroplasma-Drosophila*. For example, Minchela and Loverde (1981) reported that the females from the snail *Biomphlaria glabrata* infected by *Schistosoma mansoni* show marked decreased longevity and fecundity. Indeed, infected female snails experience a sharp longevity reduction at about 10 weeks of age and by 15 weeks of age all have died; similarly, the fecundity of infected females peaks at 2 weeks of age and sharply decreases from week 3 (about

55 eggs/snail) to 0 eggs/snail at week 8. Contrarily to what is observed in infected females, the uninfected ones show a minimal decrease in longevity and at week 20 about 90% of females is still alive, while their fecundity drops slowly, continuing beyond week 20, at which time is about 30 eggs/snail. In an experiment conducted in *Daphnia magna* infected by *Pasteuria ramosa* and cultured at high food level, the infected females experienced high mortality starting at 30 days of age and all of them died at 50 days of age, while all controls were alive at age 50 days, when the last infected host died, and at 70 days of age 80% of the controls were still alive; similarly, infected females produced about 20 offspring/female and uninfected controls produce about 150 offspring/female, both fecundities expressed in log-transformed offspring numbers (although no base for the log transformation is given in the paper) (Ebert, Carius, Little, & Decaestecker, 2004).

3.4.3.1 Persistence of male killing vs non-male killing strains

Literature review shows that other host-parasite systems show comparable trends to our results, a fact that lends support to our hypotheses that the differences on fecundity and longevity observed between our experimental groups is a consequence of the infection, which in turn may contribute to the persistence of *Spiroplasma* in field and laboratory *D. hydei* populations.

We can compare our results to Herren et al. (2014) report on *D. melanogaster* infected with its native male-killing *Spiroplasma* strain (MSRO), a closely related *Spiroplasma* to *Hyd1* (Haselkorn et al., 2009). Herren et al. (2014) performed a series of experiments with virgin and mated females kept on cornmeal rich media and cornmeal poor media. We did not have conditions comparable to the cornmeal poor media. However, our experiment setup allows for

comparing our results to the mated flies in rich media of Herren's study. In Herren et al. (2014) study, the longest-lived flies died at around 37 days of age and the number of eggs laid on a 14-day period is around 100; in our study the longest lived uninfected fly died at 109 days and at 14 days our flies lay about 380 eggs; therefore, *D. hydei* in our study were longer lived and more fecund than Herren et al. (2104) *D. melanogaster*.

Herren et al. (2014) reports that infected flies have a shorter lifespan than uninfected ones (see Figure 1-figure supplement 1 of Herren's work), which is in agreement with our results for survival analysis. Herren et al. (2014) also report that the cumulative fecundity of infected and uninfected flies from zero to 14 days is not statistically different. Herren et al. (2014) were able to divide fecundity into 2-days intervals and found that fecundity is significantly higher in infected females at age interval [0-2] days old but not at later age intervals. Taking Herren's 14-days fecundity as an analog to our "total fecundity" and their [0-2] days old as an analog to our [00-07] fecundity (which is higher for infected females in our study, $p = 0.0295$) we have remarkably similar results for our fecundity analysis. Overall, although Herren et al. (2014) did not report cumulative fecundity, both studies show an agreement on the effect of *Spiroplasma* infection in their hosts, namely, shorter lifespan and increased fecundity at early ages.

However, a study in the same system examined by us seems to both support and contradict our results. Xie et al. (2015) published a report addressing issues on the defensive function against parasitoid wasps of *Hyd1* infection on *D. hydei*, and part of their results are contrary to our report on fecundity of infected flies. Two of their treatment groups were S+Lh- (infected flies with no wasp attack) and S-Lh- (uninfected flies with no wasp attack), which are the equivalent to our

infected and uninfected groups and can be compared as such. They report the fecundity of infected and uninfected flies for 10 generations. The effects of *Spiroplasma* infection on the host "realized fecundity" can be worked out from the figures and tables they present in their paper. According to their analysis, there is a high degree of intergenerational variation on fecundity in their experimental flies. For most generations, fecundity of infected flies is not significantly different from uninfected flies, but at generations 2 uninfected flies are significantly more fecund than infected ones, whereas at generation 5, they saw the opposite, similar to our findings, where infected flies are significantly more fecund than uninfected flies. Based on the inter-generation variability of fecundity reported by Xie et al. (2015), it could be argued that our results, based on data from a single generation, could be due to chance alone. After all, Xie et al. (2015) showed that only 1 generation out of 10 shows statistically significant higher fecundity on infected females when compared to uninfected ones. However, at the very least, it could be argued that their generation 5 and our experiment show that conditions exist wherein infection increases fecundity. Our experiment could be repeated looking into total and early cumulative fecundity of infected and uninfected flies for various generations to address this issue.

Although it could be argued that our results are a product of chance, as previously mentioned, there are other factors that may explain the difference between our results and Xie et al. (2015). It is not our intention to produce an exhaustive comparison of our study versus Xie et al. (2015), but to address what might be some of the most important factors producing the difference in our results. It is possible that the differences we find on the compared parameters were caused by differences in methods of fly culturing, on the type of data recorded, and in the time the sampling was performed.

Methods of fly culture during the experiments are different. Indeed, in Xie et al. (2015) data was gathered from experimental flies that were cultured in banana-opuntia media for 10 generations in half-pint glass bottles at a density of 50 females and 50 males per bottle; in our study, data were obtained from a single fly generation from flies that were kept in polypropylene narrow vials at a density of one female and two males per vial, using the same media formulas as Xie et al. (2015). Because there were significant differences in rearing numbers, a factor that could affect population size and crowding, comparisons of results between both studies should be taken with much caution. Indeed, Ebbert (1995) reported that crowding may affect fitness of *Spiroplasma*-infected populations of the host *D. pseudoobscura*. We recommend that the issue of rearing densities should not be ignored and could be further explored in other studies.

The flies' fecundity reported by Xie et al. (2015) for the S+Lh- (*Spiroplasma* infected) group in their experiment was obtained by averaging the "realized fecundity" of the flies within the group, which was composed of varying percentages of infected and uninfected flies (infection frequency varied from 41.2 to 67.6 throughout their experiment). This means that the average fecundity of infected flies reported by Xie et al. (2015) is not equivalent to our reported fecundity because our infected group was constituted of 100% infected flies. Therefore, one important source for the difference in the infected flies' fecundity reported by both studies may be that the Xie et al. (2015) reported fecundity may have been lowered by the inclusion of uninfected flies in their estimates.

Another source of difference in our results may be found in the fecundity-sampling period in the two studies. Our data was gathered during the lifetime of the experimental and control flies, and

we were able to separate the fecundity into seven-day periods, while Xie et al. (2015) appear to have collected data on fertility of flies for a single period of ~two or ~three days (not clear from their methods section) for flies aged ~15 days old. Averages of progeny/per fly for each bottle sampled at this 2 or 3-day period were used for fecundity analysis for each generation. The Herren et al. (2014) study discussed earlier shows that fecundity of *Spiroplasma*-infected *D. melanogaster* was significantly different from uninfected ones in the first two-day period of oviposition, while the remaining time (two-day periods up to14 days) and overall fecundity at 14 days were not significantly different. This suggests that a single short-time fertility sampling, as conducted by Xie et al. (2015), may miss the sampling period when differences between treatment groups can be detected. If Herren et al. (2014) sampled their flies for anything other than a 0-2-day-old time period, they would not have found any difference between their treatment groups. In our experiment, the non-cumulative fecundity (fecundity for each closed interval) is significantly different between infected and uninfected flies in four intervals [0-7], [15-21], [22-28] and [36-42] days old but was not significant for seven other 7-day intervals (analysis not shown). Had our sampling of fecundity (and statistical analysis) been restricted to any other of the non-significant intervals (e.g., [8-14] days old, [29-35] days old) we would not have found a significant difference. This tells us that appropriate sampling time is not a trivial issue for this type of research; if the appropriate sampling period is not included in the experiment, then erroneous conclusion may be reached regarding the parameter studied.

It is important to notice in this discussion that the assumption in Xie et al. (2015) work is that the short time interval sampling they used can capture true differences in fecundity between experimental flies and no consideration is given to the cumulative early fecundity. In our

experiment, the cumulative early fecundity of infected flies is significantly higher than uninfected ones at 7, 14, 21, 28, 35 and 42 days (analysis not shown), but as mentioned above only the non-cumulative early fecundity at the [0-7], [15-21], [22-28] and [36-42] day-old intervals show significant difference between infected and uninfected flies. Considering the above, cumulative early fecundity, rather than early fecundity at shorter time intervals, would be more likely to capture real differences between experimental groups.

It is also important to address the unpublished reports referenced by Kageyama et al. (2006) and Osaka et al. (2008) on the lack of fitness effects of *Spiroplasma* infection on *D. hydei*. Although we do not have any details on those experiments, their conclusions are clear and do not agree with ours. We suggest that such differences may be due to the fact that the host and *Spiroplasma* strains used in our respective works are genetically different. Genetic background plays a role on how the infection may affect the host and the response of the host to the infection (e.g., Decaestecker, Vergote, Ebert, & Meester, 2003; Ebert, 1994; Vale & Little, 2012; for studies on *Daphnia*).

We cannot perform an exhaustive comparison of the genetic difference between our bacteria *Hyd1* strain (carried by *D. hydei* strain TEN 104-106 collected in Tenancingo, Estado de Mexico, Mexico) and the strain used by Osaka et al. (2008), collected in Southern Japan, because we do not have the adequate genetic material or published gene sequences for such purpose. However, sequences for two genes have been published and can be used for a preliminary comparison of the *Spiroplasma* strains in question. Indeed, Kageyama et al. (2006) published partial sequences for the 16s rDNA (GenBank accession number AB252475) and p58 (GenBank accession number

AB252477) genes from a *Spiroplasma* strain infecting *D. hydei* collected in Tsukuba, Southern Japan; those gene sequences show a 100% identity to 1292 bp of the 16s rDNA gene from our *Hyd1* strain (GenBank accession number FJ657202, Haselkorn et al., 2009) and a 360 bp fragment of the p58 gene for the *Hyd1 Spiroplasma* carried by *D. hydei* strain TEN104-102 (GenBank accession number pending) from the same location as TEN104-106. However, the genes above may not have enough variability to produce a reliable diagnosis of the genetic variation between our *Spiroplasma* strains. For example, the 16s rDNA gene is highly conserved and a phylogenetic tree constructed from this gene groups together spiroplasmas known to be different (i.e., NSRO, MSRO, *Spiroplasma poulsonii*, a *Spiroplasma* carried by *D. hydei* similar to our *Hyd1*, and the *Spiroplasma* strain carried by *D. simulans*) (Haselkorn et al., 2009). The p58 fragment above produces 0% pairwise nucleotide distance between MSRO, NSRO and *Spiroplasma poulsonii*, which are known to be different, as well as between the strain carried by *D. mojavensis* and a strain similar to *Hyd2* (carried by *D. hydei*).

With respect to the host, we do not have any gene sequence for *D. hydei* from Japan to compare it to TEN-104-102 *D. hydei*, but this issue can be addressed by a collaborative research with the Japanese researchers interested in the issue.

A good first step to address the issue of genetic variability of host and endosymbiont and the effects of such variability on the infection effects would be to use *D. hydei* flies from several locations from Japan, and *D. hydei* from several location from Mexico/USA and their respective *Spiroplasma* strains in experiments conducted under the standardized experimental conditions

and data collection regimes. These experiments may be difficult to carry on because of the difficulties involved in importing flies from different countries into the USA.

After consideration of other studies, we are taking Herren et al. (2014) as a positive confirmation of our results. Herren et al. (2014) used 5 females per vial in their study and, as explained before, their methodology in general is more similar to ours. Further, there are other studies on other host-parasite systems that can be used to support our findings. Certainly, examination of other parasite-host relationships reveals that infection by maternally inherited bacteria increase the reproductive output of their hosts. For example, in the psocid *Liposcelis tricolor*, *Wolbachia*-infected females produce more eggs than uninfected females (Dong, Wang, & Zhao, 2006). Similarly, a *Cytophaga*-like organism (CLO) (an intracellular bacterium from the Cytophaga–Flavobacterium–Bacteroides phylum) that has been reported to increase the fecundity of infected females of the mite *Metaseiulus occidentalis* (Weeks & Stouthamer, 2003).

3.4.3.2 Is increased fecundity an effect of the bacteria or an adaptive response of the host to increased mortality?

There are some recent works suggesting that the reduction of longevity as well as the increased early fecundity of infected flies is caused by the presence of *Spiroplasma*, although the mechanism through which this phenomenon is caused is not known with certainty. Regarding longevity, Herren et al. (2014) hypothesized that the shorter lifespan of MSRO-infected *D. melanogaster* females, compared to their uninfected counterparts, could be related to the production of cardiolipins by *Spiroplasma*. MSRO and *Hyd1* are closely related and grouped into the same clade (Haselkorn et al., 2009), therefore it is reasonable to believe that a similar

mechanism affecting longevity may be at play in the *D. hydei*-*Hyd1* and *D. melanogaster*-MSRO systems.

Endosymbionts can directly affect the fecundity of their hosts and can occur in the system studied here. An example of the direct effect of bacteria on the reproductive process of the host comes from Fast et al. (2011) who have shown that a strain of endosymbiont bacteria *Wolbachia*, *w*Mau, can increase egg production in its natural host, *D. mauritania*. Herren, Paredes, Schüpfer, and Lemaitre (2013) and Herren et al. (2014) seem to indicate that the mechanism whereby *Spiroplasma* affects fecundity in *D. hydei* is likely to be different from *Wolbachia*'s; however, the mechanism is unknown (Herren et al., (2014).

A clue to another potential mechanism responsible for the effect of *Spiroplasma* in the fecundity of *D. hydei* comes from Hutchence, Fischer, Paterson, and Hurst (2011), who artificially introduced *Hyd1* originating from *D. hydei* strain TEN 104-106 (the same strain used in our work) into a non-natural host, *D. melanogaster*. The strain *Hyd1* downregulates clusters of genes associated with fertilization and egg production in this new host (ibid), and this suggests *Hyd1* may be capable of acting upon similar clusters of genes in *D. hydei*. However, contrarily to what was observed in *D. melanogaster*, the bacteria may upregulate those genes in its native host, a hypothesis that was not tested in this work. This hypothesized upregulation in the native host implies that the negative effect on fecundity in *D. melanogaster* is a consequence of maladaptation of the bacteria to the distantly related new host (See Figure 1 in Haselkorn, 2010a for a view of the relationships among *Drosophila* hosts of *Spiroplasma*).

An alternative explanation to our results could be based on nutrition. Nutrition has been shown to affect longevity and early fecundity in *D. melanogaster*, with different optimal protein to carbohydrate ratios favoring early fecundity and longevity (e.g., Lee at al., 2008; Lee, 2015). Given that diet affects early fecundity and longevity in *D. melanogaster* (and if the same holds true for *D, hydei*), it could be argued that there is the possibility that with a different type of food composition (i.e., one favoring early fecundity or another favoring higher longevity), we would have obtained a different response for those two parameters in our experiments. We cannot test for the role of nutrition in our results because we tested our flies under one type of food only, the "banana-opuntia" recipe, which is one of the media used to successfully grow *D. hydei* in laboratory cultures (Markow & O'Grady, 2006).

Insights into how nutrition may modify the fecundity and longevity response of flies infected by *Spiroplasma* can be gained from Herren et al. (2014), previously cited. Herren et al. (2014) used two types of diets in his research, "rich media" and "poor media," with the "rich media" having higher fat and protein content than the "poor" one. Mated infected females cultured in rich and poor media and virgin infected females cultured in rich media had higher fecundity and lived shorter than their uninfected counterparts, regardless of the absolute value of fecundity and longevity for each treatment (Herren et al., 2014). Only for virgin females cultured in poor media was no difference on fecundity or longevity reported for infected vs uninfected females (Herren et al., 2014). The Herren et al. (2014) report shows that under conditions likely to be found in the field (mated flies under poor food and rich food conditions), the presence of *Spiroplasma* modifies longevity and fecundity of the infected flies in a way that is congruent to our results. However, the potential role of role of nutrition should be controlled in future *Spiroplasma*

experiments addressing longevity and fecundity. We recommend that in future experiments, a minimum of two types of food be tested: one favoring early fecundity and another favoring higher longevity.

Another alternative explanation for the early fecundity increase reported for infected females in this work and Herren et al. (2014) is an adaptive response to the infection's shortening of life of the infected hosts. Many researchers have proposed that increased early fecundity or reproductive effort is an adaptive response to the host to infection (e.g., Gérard & Theron, 1997; Adamo, 1999; Ebert et al., 2004; Chadwick & Little, 2005). Adult snails *Biomphalaria glabrata* infected under laboratory conditions with the trematode *Schistosoma mansoni* laid an increased number of eggs from one to three weeks after the infection and showed a reduced eggs output after week five, when compared to uninfected snails (Gérard & Theron, 1997). Adamo (1999) reported that crickets *Acheta domesticus* experimentally infected with the bacteria *Serratia marcescens* (that can kill the infected host) produced more eggs that the control group after 24 hours post infection. In the *Daphnia* – parasite systems, two parasites seem to elicit early fecundity in the host as a response to the reduced longevity or reproductive output later in life due to the infection; an obligate endoparasite of *Daphnia*, the bacteria *Pasteuria ramosa* that invariably produces castration of its host (drastically reducing its total fecundity) also reduces longevity of infected *Daphnia*, and the infected host shows an earlier age of reproduction when compared to uninfected ones (Ebert et al., 2004).

Ultimately, our results do not support the hypothesis that increased early fecundity for infected females is an adaptive response of the host to the infection. Although one of the conditions for

eliciting this response in the host seems to occur (i.e., *Spiroplasma* reduces the longevity of *D. hydei* females), one other condition is not reported in our results (i.e., the total fecundity of infected females is not reduced, as shown by the non-significant difference between infected and uninfected females' fecundity). In addition, the compensatory response of the infected hosts usually does not result in their fecundity being higher than that of their uninfected counterparts (Agnew et al., 2000). With our results, we are inclined to assume that the increased early fecundity of infected females is a direct effect of the bacteria on the fecundity on the fly, as proposed early.

3.4.3.3 Consequences of increased early fecundity in infected females for persistence of infection

The early fecundity advantage conferred by the infection might increase the fitness of the field-dwelling infected flies, where intra and interspecific interactions are guaranteed to occur. In turn, higher early fecundity of infected females might be one of the mechanisms contributing to maintaining the infection in the field and laboratory populations. Protection against parasitoid wasps produced by infection, as reported by Xie et al. (2015) could be important in maintaining the infection in field populations, but this hypothesis fails to explain the long-term persistence of the infection in *D. hydei* laboratory cultures where no wasps are present. While an early fecundity advantage for infected flies was confirmed by the present work and Herren et al. (2014), this advantage was not confirmed in the field, and it remains an open research question. In this section we address how early fecundity advantage could play a role in the long-term persistence of *Hyd1* infection in laboratory and natural populations.

We will use simple simulations in excel spreadsheets to show scenarios where infections frequencies change over generations in hypothetical *D. hydei* populations. A baseline simulation for a population initially composed of 100% infected females with a constant transmission rate of 0.9, with infected and uninfected females producing the same number of progeny (we are using the least squares mean for early fecundity for uninfected females obtained in this work, that back converted from the Johnson transformation is ~800 eggs) is shown in SM S7; the infection frequency in the population after 10 generations would be 0.9^10 = 0.3487, after 20 generations 0.9^20 = 0.1216, and after 36 generations 0.9^36 = 0.0225.

Applying our simulation technique to our fly fecundity data using reports on transmission rate from the literature, we can hypothesize how those factors contribute to the persistence of the infection in laboratory populations. Xie et al. (2015) reported that the vertical transmission rate of *Spiroplasma* infection in 10 generations of *D. hydei* was almost 100% (varying from 0.9875 to 1.00); similar results were reported by Osaka et al. (2008), who showed that the transmission rate in *D. hydei* at 25 degrees Celsius is 0.99 (varying from 0.95 to 1.00). In laboratory cultures flies are usually maintained in half pint bottles and we are assuming that fly-keeping is performed every 15-28 days. Parent flies are discarded and the bottles, containing larvae/pupae, are kept to collect the newly emerging flies that are allowed to reproduce before being discarded after oviposition (after 15-28 days) to produce the next generation. This fly-keeping cycle is continuously repeated. A 28 days fly-keeping cycle matches the time we have defined as “early fecundity” in this report. Table 3.4 shows that infected flies lay 935 eggs on average and uninfected ones lay 717, which represents an approximate of 30% advantage for infected flies, raw data. Assuming that the percentage of fly emergence success is not affected by the infection

status of the flies and that fly rearing conditions affects all flies equally, a fly culture starting with an infection frequency of 50% (i.e., 50% infected and 50% uninfected flies) will not lose the infection in the long term given a transmission rate of 0.925 (SM S8). Such a population will keep the infection frequency above 60% for many generations; at generation 10 the infection frequency will be about 0.6444, at generations 30, 0.6739, and at generations 75, 0.6795 (see SM S8). Notice that the transmission rate used in this simulation is much lower than the ones reported by Xie et al. (2015) and Osaka et al. (2008). The simulations above show that an early fecundity advantages of about 30% and 18% for infected females accompanied by a 0.925 transmission rate, and a starting infecting frequency of 50% seem to explain the persistence of *Spiroplasma* infection on a hypothetical *D. hydei* laboratory cultures in the long-term.

In field populations, in addition to fecundity and transmission rates, we have to consider environmental factors that affect the fly breeding season, population abundance of the host fly, transmission rates of the infection, and infection frequencies of *Spiroplasma*, among other parameters. For this part of the discussion, we will focus on temperature, an environmental factor that in laboratory experiments has been shown to affect the transmission and infection frequencies of NSRO (a *Spiroplasma* strain native to *D. nebulosa*) in its native host and in a nonnative host *D. melanogaster* (Anbutsu et al., 2008) as well as the transmission rate (Osaka et al., 2008) and the infection frequencies of *Hyd1* in *D. hydei* (Osaka et al., 2010).

Southern Japan is a geographical area well suited for exploring the role of temperature in the persistence of *Spiroplasma* infections in field populations of *D. hydei*. This is because there is information about monthly temperatures and seasonal changes on fly population levels for this

area. Furthermore, as explained earlier the *Spiroplasma* strain infecting *D. hydei* in Japan is similar to the bacteria strain we studied in this work.

Southern Japan has a period of cold temperatures and another one of hot temperatures (see Table S9 in SM S9). Average monthly temperatures from 2005 to 2008 for Chiba, Southern Japan. Data from Japan Meteorological Agency: http://www.data.jma.go.jp/obd/stats/etrn/view/monthly_s3_en.php?block_no=47682&view=1). The SM S9 shows that temperatures higher than 10°C occur from March or April to November and lower than 10°C occur from December to February or March.

In Southern Japan *D hydei* is present in the field all year round (Katoh &Watada, 2011; Toda, 1979; Yamamoto, 1992) and they seem to reproduce even during winter times as determined by the ovarian condition of wild-caught flies (Toda, 1979), but the numbers of flies collected during winter times are almost nonexistent (see tables reported in Katoh &Watada, 2011 and Toda, 1979). *D. hydei* is abundant during the hotter times of the year (Yamamoto, 1992; Osaka et al., 2010). Yamamoto (1992) reported that in 1979 in Oiso (35°19'N, 139°20'E), *D. hydei* appeared in collection traps from June to the first week of August, with larger numbers trapped in June, much fewer in July, and almost none the first half of August. The temperatures registered during the collection times were between 20.9 to 26.8 in June, 26.1 to 27.8 in July, and 28.7 (beginning of August) to 23.9 (end of August) (Yamamoto 1992). Osaka et al. (2010) collected flies in two localities of southern Japan, Chiba (35°47′ N, 139° 54′ E), and Ehime, but only Chiba was sampled intensively (Osaka et al., 2010) and it is used in this discussion. They collected flies on May, June, July and part of August for the years 2006, 2007 and 2008. Their data show that *D.*

hydei was captured from May to July 2006-2008, as well as in August for 2007, with a peak in June for all years sampled. The minimum and maximum temperatures for June varied from approximately 16°C to 21°C and 23°C to 27.5°C, respectively (Osaka et al., 2010). As in Yamamoto's work, Osaka et al. (2010) collected many fewer flies in July as compared to June, and the lowest number of flies was collected in August, the hottest month during their study, with temperatures as high as 29.86 C in 2007. The combined information reported in Yamamoto (1992) and Osaka et al. (2010) indicates that in Southern Japan *D. hydei* first appears in large numbers in May, the numbers peak in June, and sharply drop in July and are virtually nonexistent on the second week of August (see SM S10). In summary, seasonal changes of temperature determine the abundance of field populations of *D. hydei*, which are at their minimum levels during the cold winter season and increase in abundance as the temperatures rise on Spring-Summer (e.g., Yamamoto, 1992; Toda, 1979; Osaka et al., 2010; Katoh & Watada, 2011).

We will continue this discussion by focusing on Chiba, because this is a location where published information on infection frequencies of *Spiroplasma* on *D. hydei* field populations have been systematically documented for three years during the season when *D. hydei* is abundant in the field (Osaka et al., 2010). SM S11 shows the infection frequencies and air temperatures for periods from May to August 10 for the years 2006, 2007, and 2008 (data extracted and modified from Osaka et al., 2010.)

The period when temperatures falls below 10°C in Chiba is from December to February- March. For the years 2005 to 2008, only March 2007 and March 2008 had average temperatures just

above 10°C (10.6 and 10.3°C, respectively); for the same period, the lowest temperature recorded was 4.7°C (Jan 2006) (SM S9). The highest temperatures of this cold period are close to temperatures that have been reported to hinder the development of *D. hydei* embryos. Certainly, laboratory experiments published by Harisanova and Ralchev (1985) show that maintaining at low temperature embryos that were collected within one hour of oviposition almost completely suppressed their development but not their viability. Harisanova and Ralchev (1985) were able to keep *D. hydei* embryos at 10°C for 4, 8, 12, 16, 20, and 24 and later place them at 25°C where they developed normally. It is therefore likely that the environmental temperatures below and just above 10°C occurring from December to March in Chiba, could suppress the development of field *D. hydei* embryos, therefore no flies seem to be produced during those cold months in the field.

Assuming that the low temperatures registered in Chiba are an impediment for the expansion of populations of *D. hydei* in open areas, we can further assume that this species will start reproducing in high numbers around the end of April to beginning of May (as can be deduced from Osaka et al, 2010 collection data), when temperatures are much warmer than 10°C (SM S9), with the number of collected flies increasing from that time until June (Table S10.1 and Fig. S10.1 within SM S10, Chiba). Note that there is a sharp drop on the number of flies captured in July and August (Fig. S10.1 within SM S10, Chiba), which suggest that there are fewer and fewer flies that are reproductively active as time progresses from July to August. Therefore, it seems that there is a discrete time from the end of April to the end of July when *D. hydei* is actively reproducing in large numbers in open areas in Southern Japan.

It is important to remember here the effect of temperature on the transmission of *Spiroplasma*. Laboratory experiments on the strain of *Spiroplasma* naturally infecting *D. hydei* show that 15°C totally suppress the transmission of the bacteria, 18°C may partially do so, and 25°C and 28°C seem to be optimal temperature for the transmission of the bacteria (Osaka et al., 2008); the authors concluded that the lower threshold temperature for vertical transmission lies somewhere between 15°C and 18°C. However, data on transmission rates of field-caught infected flies that were allowed to breed at 25°C show that even at that optimal temperature the transmission rate for some individual flies was as low as 0.364 (Osaka et al., 2013a); therefore, the assertion that 18°C may partially suppress the transmission rate of the bacteria based on the observation of two females showing 0.38 and 0.83 transmission rates (the three other sampled females at this temperature had a transmission rate of 1.0) (Osaka et al., 2008) needs to be corroborated with larger sample sizes, and for this discussion, this assertion will be considered incorrect. Based on the above, we will consider in this discussion that transmission rates from 18°C to 28°C are similar, and that somewhere between temperatures > 10°C and <18°C the transmission rate is partly suppressed.

In Chiba, for the years 2006, 2007, and 2008, temperatures reached 18°C sometime in May (on the way up to hotter temperatures) and it is again reached sometime between October and November (on the way down to cooler temperatures) (see SM S9). This means that there is a discrete time when female flies can transmit *Spiroplasma* to their progeny and in Southern Japan and this period overlaps with most of the discrete time of the year when we propose *D. hydei* females reproduce in larger numbers in the field (May, June and July).

We ran a simulation whose results help to illustrate how the early fecundity advantage of infected females may contribute to the persistence of *Spiroplasma* infection in *D. hydei* populations in Chiba. Based on Harisanova and Ralchev (1985), Osaka et al. (2008) and Osaka et al. (2010) and Yamamoto (1992) findings we can assume that there is a limited number of days (~100 -~120, from middle of April to Middle of July) in which *D. hydei* can reproduce in open areas (and when females can transmit the infection to their progeny). Therefore, there are a limited number of fly generations that can be produced in the open field during the year.

Starting with a cohort of freshly deposited eggs and another cohort of mature males and females and allowing about 10 days for males to mature and 11 days for larval development, it takes about 21 days, (starting from eggs), and 11 days (starting from mature males and females) to produce a new generation for *D. hydei*. If the first eggs and mature males and females appear at the end of April, this would allow for the production of about five and six generations with 28-day generation intervals. Overall, an estimate of 11 to 12 28-day generation flies would be produced during the high reproductive period in the field in Chiba. However, because matures flies may lay eggs every day, their reproductive activity will produce overlapping 28 days-old generations, (i.e., more than 11 or 12 generations will be produced during the high reproductive period of the flies). SM S12 shows the results of such a simulation and it makes a strong case that a population with early fecundity advantage for infected females will parallel the increase of the infection frequencies reported for 2007 in Chiba by Osaka et al. (2010). We used a 30% starting infection frequency to emulate the starting infection frequency for 2007 (Osaka et al., 2010) and a 0.97 transmission rate, which is a realistic rate, based on Osaka et al. (2008), Osaka et al. (2013a) and Xie et al. (2015). With these constrains, we tweaked the early fecundity data to

represent about 18% fecundity advantage for infected females (935.6 average eggs for infected females and 793.8 average eggs for uninfected females; these values for fecundity are realistic based on our experimental data), to approximate the trend for 2007 in Chiba. The infection frequencies from the simulation described for SM S12 shows that the infection frequency increases steadily from 0.30 at generation zero to 0.5524 at generation 9. The table and figure in Appendix A12 shows that the increases of infection frequencies from the simulation roughly parallel the increased infection frequencies in Chiba for 2007 (Osaka et al., 2010). The values for these parameters for Chiba do not need to be the same as the ones used in the simulation above, because there are alternative values that they can take and produce similar results. Therefore, we propose that in the Chiba *D. hydei* population, *Spiroplasma* infection dynamics for 2007 and 2008 are likely determined by the early fecundity advantage for infected flies, high transmission rates, and low starting infection frequencies.

Much remains to be done to fully understand how *Spiroplasma* infection persists in *D. hydei* and other *Drosophila* species. It is not known what happens to the infection transmission rates or infection frequencies during the cold months of winter in southern Japan. As discussed above, the transmission of *Spiroplasma* is suppressed at 15°C in the laboratory (Osaka et al., 2008) and sometime between the beginning of November and the end of April from 2006-2008 the temperature fell below that value (SM S9). Also, *D. hydei* embryos stop their development at 10°C (Harisanova & Ralchev, 1985) and from December to March 2006-2008 the temperature fell below that value (SM S9). We do not know what happens to the bulk of *D. hydei* populations and the *Spiroplasma* infection they carry during the days those cold temperatures occur. As indicated earlier, *D. hydei* still occur in the field and seem to reproduce during the cold months in

small numbers (Katoh & Watada, 2011; Toda, 1979), but no reports on the infection frequencies for those times of the year have been published for Southern Japan. Yamamoto (1984), Yamamoto (1992), and Yamamoto and Ohba (1984) state that *D. hydei* is considered a domestic species in Japan, and as such, these flies may survive the winter by inhabiting locations near human dwellings where they may be partly protected against cold weather. Because cold temperatures are reported to prevent the transmission of *Spiroplasma* in *D. hydei*, *Spiroplasma* may survive at low infection frequencies during those periods. *Spiroplasma* seem to have an imperfect but high transmission rate at warm temperatures in *D. hydei*, so when the temperatures in Southern Japan cross an unknown threshold optimal temperature for the transmission of *Spiroplasma* (>18 C), the infection frequencies will rise generation after generation. The above implies that the infection frequencies reported for Chiba from 2006 to 2008 (Osaka et al., 2010) may drop and stay low during the cold weather season and rise during the warmer period of the year, starting April-May.

As addressed earlier, there are yearly variations on the behavior of the infection frequencies in Chiba for 2006 and 2008. For example, data for the infection frequencies for the year 2007 shows a pattern compatible to with the scenario drawn in the previous paragraph, but not the data for 2006 and 2008 (Osaka et al., 2010). These variations may be explained by environmental stochasticity, as we will attempt to show next.

The role played by environmental stochasticity (e.g., yearly variations of winter severity) on the infection frequencies observed in Chiba can be addressed tentatively in this discussion. Reportedly, rising temperatures significantly correlates with infection frequencies in field

collected *D. hydei* during 2007, but not in 2006 and 2008 (Osaka et al., 2010, SM S11). In searching for an explanation to this observation, we took a closer look to the lowest temperatures during the cold season for each of those years (see SM S13). SM S13 clearly shows that from November of the previous year to March of the current year, the temperatures were colder for 2006 and 2008 than for 2007; for example, the minimum temperature for 2006 was 4.7°C (January) and 5.4°C for 2008 (February), while in 2007, it was 7.4°C (January); snow fell in 2006 and 2008, but not in 2007. We propose that on 2006 and 2008, before the massive reproductive season that started in April-May, the flies went through a more stressful period than in 2007, due colder temperatures. SM S10 (Table S10.1 and Fig. S10.1) shows that the number of flies collected for Chiba in May 2007 was larger than in May 2006 and May 2008, which seems to corroborate that the latter years were more stressful than the former for the fly populations. Furthermore, the infection frequencies for all years were low at the beginning of May, which indicates that the cold temperatures impaired the transmission of *Spiroplasma* from 2006 to 2008. However, the effect of low temperatures on transmission rates and infection frequencies may have lasted longer in 2006 and 2008 than in 2007. Overall then, in 2007, the fly populations start the massive production period in better conditions than 2006 and 2008. The correlation observed for 2007 between the increasing temperature and the increasing infection frequencies reported by Osaka et al. (2010) may not indicate a causal relationship. Instead, the rising trend of the infection frequencies in field populations observed in 2007, according to our SM S12, is determined by the high transmission rates of the infection, low initial infection frequencies and early fecundity advantage for infected females, all of which were, in turn, influenced by variations in temperatures in the ways discussed above.

There may be other unknown factors that play a role on the *Spiroplasma* infection persistence. The fact our simulations produce infection frequencies that are not a perfect match to the values reported for Chiba (Osaka et al., 2010) or for laboratory populations (Xie et al., 2015) suggests that those other factors may be intrinsic to the fly or the bacteria as well as environmental in nature.

3.4.3.4 Early fecundity advantage and protective function

An important outcome of this study is that the higher early fecundity of *Hyd1*-infected females reported here represents and additional mechanism that may promote the long-term persistence of the infection in *D. hydei*, in addition to the defensive function of this *Spiroplasma* strain against parasitoids reported by Xie et al. (2010). However, the higher early fecundity for infected females reported here may be more broadly applicable to the problem of persistence of the infection than the protective function. For example, higher early fecundity may be used to explain the persistence of the infection in laboratory populations, but not the defensive function. The best proof for the latter is that laboratories around the world (with no wasp attacks present) have been able to maintain for many generations *Spiroplasma*-infected stocks of *D. hydei*.

The early fecundity advantage and the protective function of *Spiroplasma* may be used to explain the persistence of the infection in the wild. However, these mechanisms were observed in laboratory experiments, therefore they need confirmation in the wild. Through simulations, we attempted to show how early fecundity may promote the persistence of the infection in wild populations in Japan, but we do not have proof of this mechanism working in wild populations. To confirm if early fecundity advantage of infected females exists in the wild it would be

necessary to find a population where the infection has persisted for a long time and capture infected and uninfected flies and bring them to the lab and measure their fecundity and fertility. The problem with this experiment would be that we would have control on the age of the flies captured. To address this issue, we would collect the flies after the beginning of a season of abundant fly populations. For example, in Chiba, Japan we would collect flies on the middle of May or beginning of June, when it is likely we would have a large population of younger flies.

To study the defensive function of *Hyd1* in the wild, it is necessary to find field populations of *D. hydei* where attack by parasitoid wasp occurs. If such populations are found, long term monitoring of parasitoid attack rates and *Spiroplasma* frequencies are necessary to clarify if the former is a factor contributing to the persistence of *Spiroplasma* infections.

While the early fecundity of infected flies has not been addressed in publications (this is the first report for a non-male-killing *Spiroplasma*), we were able to find reports that may help us to, at least in theory and preliminarily, evaluate the potential of the protective function of *Hyd1* in wild *D. hydei*. Parasitoid attacks to *D. hydei* in the field have not been confirmed in the literature. Laboratory experiments show that *D. hydei* is an intermediately suitable host for the parasitoid *Leptopilina*, unlike *D. melanogaster* and *D. simulans*, which are best suited for it (Fleury, Gibert, Ris, & Allemand, 2009). These two species are sympatric with *D. hydei* throughout most of its range, and parasitoids may favor these two optimal hosts instead of a mediocre one (as indeed reported for South-eastern France by Fleury, Gibert, Ris, and Allemand 2004, who shows that parasitoids attack preferentially *D. melanogaster* and *D. simulans.*) A look into the preferences of fly hosts by parasitoids in another region of the world, Japan (localities of Sapporo, Sendai,

Tokyo, and Kagoshima), comes from a report by Mitsui, Achterberg, Nordlander, and Kimura (2007). The authors collected pupae of 18 species of flies (*D. hydei* not included among them) in banana traps and report that the most abundant pupae collected belonged to the species *D. melanogaster*, *D. simulans*, *D. lutescens*, and *D. auraria* and they were parasitized by parasitoids from the genera *Asobara*, *Ganaspis*, *Leptopilina*, *Leptolamina*, *Trichopria* and *Pachycrepoideus*. Mitsui et al. (2007) report that, *L. heterotoma*, the species studied in Xie et al. (2015) seems to prefer *D. melanogaster* and *D. simulans* as well as the four species from the *D. auraria* complex; *Asobara japonica* seems to prefer *D. melanogaster* and *D. simulans* as well as *D. lutescens*, D. *takahashii* and *Scaptodrosophila subtilis* (Mitsui et al., 2007.) We mention *A. japonica* because this parasitoid can attack *D. hydei* in laboratory experiments, but it is not known if it does so in the field (Masahito T. Kimura, pers. comm.) Although Mitsui et al. (2007) data does not include parasitism on *D. hydei*, their report confirms that parasitoids prefer to attack *D. melanogaster* and *D. simulans*, as in south-eastern France. All of above suggests that if attack by parasitoids on *D. hydei* occurs in the field, the high attack rates used in the experiments by Xie et al. (2015) are not supported by published data.

Should *Hyd1* infection defend *D. hydei* against parasitoid attacks in field populations, the benefit is likely to be limited for the persistence of *Spiroplasma*. *D. hydei* surviving wasp attacks due to protection by *Hyd1* show lower survival and lower fertility than untacked ones (Xie, Tiner, Vilchez, & Mateos, 2011; Xie at al., 2015). In Xie et al. (2015), infected flies exposed to the parasitoid attacks show significantly lower fecundity than infected non-exposed and uninfected non-exposed flies in the first two generations. In addition to the above, a significantly low number of infected flies survive to maturity after the wasp attack even though they are protected

by *Spiroplasma* (Xie et al., 2011). Another thing to consider is that none of the publications about *Hyd1* protection of *D. hydei* against parasitoid wasps report the effects of the wasp attack on the development and behavior of the fly larvae (Xie et al., 2010; Xie et al., 2011; Xie et al., 2015). It is reasonable to assume that the wasp may have an effect on the development and behavior of the fly larvae and reduce its competitive ability when developing in the field in parallel with conspecific and interspecific larvae. Assuming that there is an unreported effect of the wasp on *Spiroplasma* infected larvae and considering the report by Xie et al. (2015) that the emerging flies belonging to this group have a lower fecundity than other flies for at least 2 generations post-attack, the overall effect of the wasp on the flies they attack may be the reduction of their competitiveness in a setting where they have to undergo intra and interspecific competition, (i.e., in field populations).

In summary, both potential mechanisms that would promote the long-term persistence of *Hyd1* in wild *D. hydei* populations, early fecundity advantage and protective function, await for confirmation on wild populations.

3.5 Conclusions

Spiroplasma Hyd1 infection reduces the longevity of *D. hydei* and increases the early fecundity of infected females, and the latter represents a direct fitness benefit for the host, which may promote the long-term persistence of the infection. However, this is likely a complex issue, with the protective function of *Spiroplasma* against parasitoid wasps, and some degree of intraspecific horizontal transfer of the bacteria through mites, as an additional potential mechanism involved

in the process. Field research is urgently needed to address the contribution of each of these three factors to the persistence of the *Hyd1* infection in *D. hydei* populations.

Simulations of early fecundity advantage of infected females (infected flies were simulated as having 30% higher fecundity than the uninfected ones), coupled with a high, but not perfect transmission rate of the infection (0.925) acting on a population with 50% infection frequency shows that this mechanism may contribute to the persistence of the infection in laboratory (and field populations) in the long term. By modifying any or all of the three parameters above, we were able to simulate populations where the infection frequencies increase or decrease with time. For field populations that live in variable environments, the persistence of a Japanese strain similar to the *Hyd1* strain used in our experiments helped to illustrate, by means of our simulations, the interplay of early fecundity advantage, transmission rates, starting infection frequencies, and a seasonally changing temperature to determine the persistence of *Spiroplasma*. The low temperature experienced by field fly populations during part of the year in Japan impairs the reproductive activity of large numbers of *D. hydei*; it may impair *Spiroplasma* transmission and may reduce the infection frequencies of naturally infected flies. This may explain why *Hyd1*, which produces early fecundity advantage in the host and has a high transmission rate, does not reach frequency fixation in field populations that experiment seasonal changes in temperatures, such as in Japan. Needlessly to say, the behavior of the infection depicted in our simulations needs to be corroborated or disproved by field research.

The effect of the *Spiroplasma* infection in *Drosophila*, excepting male-killing, may be weak as a general rule; this work on *D. hydei* and Herren et al. (2014) on *D. melanogaster*, point in that

direction, because both reports indicate that longevity and fecundity are weakly affected by the infection in these two species. Because the effect is weak, it might be difficult to detect experimentally, especially if sampling times and sample sizes are not appropriate.

Our results can be interpreted from the perspective of parasite manipulation of the host. From this point of view, increasing the early fecundity of the host may constitute a common strategy of *Spiroplasma* infecting *Drosophila* species to ensure the bacteria's spread in the populations and their long-term persistence. A potential fitness advantage via early fecundity and a weak effect on longevity is in line with the observation that maternally transmitted symbionts are unlikely to evolve characteristics detrimental to the host because the wellbeing of their host is to their advantage (Douglas, 1994).

4. POTENTIAL HOST RANGE OF *SPIROPLASMA* ASSOCIATED WITH *DROSOPHILA*

4.1 Introduction

This section explores the potential host range of distinctly related parasites on a set of hosts that are also distinctly related, which is a situation that arose due to multiple introductions of the parasites into the hosts, producing incongruent host-parasite relationships (Haselkorn et al.; 2009). This is an appropriate setting for the exploration of the factors that determine the success of the transmission of a parasite into new hosts.

The evaluation of the potential host range of parasites and the factors affecting those ranges are necessary in a world undergoing accelerated changes. This is because changes in distribution of species in response to environmental changes depend not only on the latter but also on the biological characteristics of the implicated species (Chen, Hill, Ohlemüller, Roy, & Thomas, 2011).

The term “host” refers to those species in which a parasite can successfully survive, grow and/or reproduce (Poulin, Krasnov, & Mouillot, 2011). “Host range” is the number of host species infected by a parasite (Lymbery, 1989; Rohde & Rohde, 2005). A parasite that is restricted to one host species is considered a highly host-specific or a specialist parasite, and a parasite that lives in many host species is considered a generalist parasite (Poulin, 2007). The host range of a parasite is influenced by a variety of factors, such as extrinsic factors (e.g., geographical barriers) and intrinsic factors (e.g., physiological characteristics of the host) (De Vienne et al., 2009).

The range of hosts a parasite uses can be seen as a consequence of both historical events and current ecological and physiological characteristics of the host and parasites (Poulin, 2007). The study of host-parasite phylogenies provides insights into the historical processes that have shaped their relationships. For example, congruent host-parasite phylogenetic trees can be interpreted as a consequence of high host specificity and indicates that the parasite speciates when the host speciates (Page, 1994). Incongruent phylogenetic trees can be due to a number of reasons such as independent speciation of hosts and parasites, extinction of parasites, and the transfer of parasites from one host to another (Page, 1994). Within host range studies, both host phylogenies (Perlman & Jaenike, 2003; Gilbert & Webb, 2007) and host-parasite phylogenies (De Vienne et al., 2009), were analyzed as the genetic distance between pairs of hosts and pairs of parasites, have been shown to be suited to predict potential host ranges for parasites (Gilbert & Webb, 2007; Perlman & Jaenike, 2003; De Vienne et al., 2009). Host phylogeny is understood to affect host ranges because hosts that share a recent common ancestor are likely to share an internal environment that a given parasite can utilize, as well as a lack of defenses against that parasite. A similar rationale can be applied to parasite phylogeny: if closely related parasites share biological requirements to thrive inside a host, and/or mechanisms to avoid or overcome host defenses, then they should be able to exploit the same kind of host (De Vienne et al., 2009). This way, a parasite may be able to thrive on a host that currently harbors its close relative, independent of the phylogenetic relationship between the pathogen's host-of-origin and the new host (De Vienne et al., 2009).

The "filter concept" (Combes, 2001) is a simple model that can be used to address the physiological and ecological factors affecting the evolution of host specificity. Two filters

determine how many animal species can be used as hosts by a parasite: an encounter filter and a compatibility filter. The “encounter filter” excludes hosts the animals that the parasite cannot meet for geographical, ecological, and ethological reasons. The “compatibility filter” eliminates the animals that, although encountered by the parasite, represent environments where the parasite cannot survive and develop due to morphological, physiological, or immunological reasons (Combes, 2001).

Potential host range is the species that a parasite could infect barring extrinsic barriers (De Vienne et al., 2009). Potential host range is a fundamental property of parasites and needs to be determined to understand a parasite’s potential durability as a species. For example, it is a key determinant of how likely parasites are to persist following host extinction (Dunn, Harris, Colwell, Koh, & Sohdi, 2009; Moir et al., 2010), or to become established following their introduction into new areas (Taraschewski, 2006). Potential host range of parasites is usually evaluated under controlled conditions, as in laboratory experiments. Under this condition, parasites are introduced into new hosts and the success or failure of the infection is recorded. Nevertheless, it is a useful concept to predict the expansion of host range or extinction of parasites as stated above.

A simple method to quantify the potential host range are the indexes of host specificity, a concept borrowed from the field of parasitology. These indexes score the number of species that a parasite can infect and also take into account how many individuals within each species are infected by the parasite (e.g., Rohde & Rohde, 2005; Rohde & Rohde, 2008). These scores allow the classification of the parasite on a continuum scale, from a score of 1 (the highest degree of

host specificity) to 0 (the lowest degree of host specificity). The parasites showing high host specificity could be considered as having the smallest potential host range, and the ones showing the lowest host specificity, as having the largest potential host range.

To study the potential host range of distinctly related parasites on a set of distinctly related hosts, we chose the *Spiroplasma-Drosophila* association. At least 16 species of *Drosophila* are infected by at least 18 different *Spiroplasma* strains (see Fig. 1 in Haselkorn, 2010a). A phylogenetic analysis using 16S rRNA of 12 *Drosophila*-infecting *Spiroplasma* strains shows that those spiroplasmas are separated in four clades (poulsonii, citri, tenebrosa and ixodetes) (Haselkorn et al., 2009). The poulsonii includes male-killing as well as non-male-killing spiroplasmas, while the other three clades only include non-male-killing strains (Haselkorn et al., 2009). *D. hydei* is infected by two non-male killing *Spiroplasma* strains, one belonging to the poulsonii clade and the other to the citri clade, without co-infection (Haselkorn et al., 2009). The 4 clades defined in the phylogenetic tree of *Drosophila* spiroplasmas can be divided into two groups (or larger-inclusive clades): the citri-poulsonii group and the tenebrosa-ixodetis group (Halselkorn et al., 2009; Halselkorn, 2010a; Regassa & Gasparich, 2006; Lo, Chen, Chung, Gasparich, & Kuo, 2013). These two groups are separated by at least two clades that do not contain any *Drosophila*-infecting spiroplasmas (Lo et al., 2013), which suggests that they are highly divergent.

It has been reported that the host-parasite phylogenetic tree in the *Spiroplasma-Drosophila* system is incongruent (Haselkorn et al., 2009). Haselkorn et al. (2009) proposed that this incongruence is due to multiple introductions of *Spiroplasma* into *Drosophila* via horizontal transfer, because the clades including *Drosophila* spiroplasmas also include spiroplasmas that

infect organisms other than *Drosophila* flies (Haselkorn et al., 2009). This implies a situation in which *Drosophila* species shared diverse habitats with non-*Drosophila* hosts that carried diversely related spiroplasmas, and thus, some *Drosophila* species acquired those spiroplasmas though horizontal transfers from other *Drosophila* and non-*Drosophila* hosts. If the *Drosophila* species that acquired those infections were distinctly related (i.e., some closely, other distantly), then this may explain the incongruent parasite-host tree observed currently in the system.

Although spiroplasmas infecting *Drosphila* are maternally transmitted (Williamson & Poulson, 1979), experimental intra- and interspecific horizontal transfer of *Spiroplasma* through mites has been reported in *Drosophila* (Jaenike et al., 2007); therefore, the above scenario may be plausible. Under a scenario of horizontal transfer of *Spiroplasma* among *Drosophila* species, several pertinent questions are brought to mind regarding the *Drosophila-Spiroplasma* system. For example, are the host ranges of *Spiroplasma* that infect the diverse species of *Drosophila* broader than observed? Are host and parasite phylogeny predictors of success of new host infections? What factors influence the potential host range of *Spiroplasma*? The answers to these questions are important for better understanding the evolution of the host-parasite relationship in the *Drosophila-Spiroplasma* system and to understand why some species are able to expand their host range.

These questions can be addressed using the cross-infection method, which consists of experimentally infecting pairs of different hosts with the native parasites of the other (Lively, 1989; Ruiz-Gonzalez & Brown., 2006; De Vienne et al, 2009). These experiments can be designed to include variables such as phylogenic, ecological, or geographical relationships

among hosts and among parasites in the analysis (Poulin & Keeney, 2008). Experimental cross-infections can artificially overcome the "encounter filter" by forcing the introduction of a parasite into a host that it would not meet under normal conditions, opening the possibility of testing the "compatibility filter" in such host-parasite pairs that would otherwise be impossible to test. Experimental infections into non-native hosts of parasites can be used to assess the potential host ranges of organisms (Poulin & Keeney, 2008). Studies in host parasite systems have shown that some parasites have a much broader potential host range than their actual ones in nature (e.g., King & Cable, 2007), whereas others are highly specific (e.g., Molnar, Ostoros, & Baska, 2005).

Under experimental conditions, for vertically transmitted parasites such as in the *Drosophila-Spiroplasma* system, transmission of the infection to progeny is a key factor for the establishment of a new infection after introduction of the bacteria into a new host. The fact that a parasite can proliferate inside the recipient individual does not guarantee that it will transfer to the progeny at frequencies that will establish infection in the new host, or that it will be transmitted at all (pers. obs.), because the parasite may not be able to overcome the "compatibility filter." For example, *Spiroplasma* can thrive in the hemolymph of *Musca domestica* upon artificial transmission of the infection, but this species fails to transmit the infection to their progeny (Williamson & Poulson, 1979). For this reason, in the *Drosophila-Spiroplasma* system, maternal transmission to at least the first-generation post-infection must be a requirement to consider the infection in a new host a success.

There are examples of successful experimental interspecific infection in the *Drosophila-Spiroplasma* system, although the experiments reporting them were not set up as cross-infections (e.g., Williamson & Poulson, 1979; Hutchence et al., 2011; Haselkorn & Jaenike, 2015). These reports hint that the potential host range of *Spiroplasma* could be larger than observed in nature. However, no reciprocal cross-infections were performed in those experiments, therefore, it is not known if infection in the opposite direction from the tested one would show similar results. The current study will attempt to fill this gap.

The objectives of this work are to: 1) experimentally evaluate the potential host range among *Drosophila*-associated *Spiroplasma* and 2) determine whether phylogenetic constraints play a role in the potential host range in the *Drosophila-Spiroplasma* system.

4.2 Materials and Methods

4.2.1 General Methods

4.2.1.1 Transinfections

This is the process of transferring hemolymph from a *Spiroplasma*-infected donor fly into an uninfected recipient fly, with the purpose of establishing the infection in the latter. It is performed by using micro tips prepared from pulled capillaries and puncturing the flies on the thorax or abdomen. For this experiment, both donor and recipient flies were punctured at the soft juncture of the mesopleuron and scutellum to minimize damage. After transinfection, the flies were left to recover from the wound for about two days before further manipulation. Typically, the survival rate is very high (~80% to 100%, pers. obs.).

4.2.1.2 PCR screening (of infection)

PCR is routinely used to screen flies for *Spiroplasma* infection. Primers specific for *Spiroplasma* have been developed. DNA extraction is performed by incubating a crushed fly on a buffer solution containing proteinase K.

4.2.1.3 Fly rearing

Flies are maintained in vials or bottles containing banana opuntia or cornmeal. Bottles must be changed every two weeks and vials every week.

4.2.1.4 Antibiotic curing

This is a technique routinely used to cure flies from infections. A combination of tetracycline and erythromycin (final concentration = 0.2 and 0.16 mg/ml, respectively) is mixed with the fly food. Usually, flies are kept on this antibiotic-food mix for two generations. The third and further generations are kept with antibiotic-free food. PCR screening is used to confirm that the flies are cured.

4.2.1.5 Definition of host

This research will use artificial infections of parasites into novel recipient organisms. Not all new recipients of the infection may allow the persistence of the infection in the long term, and long-term persistence of infection is implied in the objectives of this research. Therefore, the term "host" will be used to mean an organism that is able to transmit the parasite to further generations (at least one generation post-infection). This distinction is necessary because it has been shown that although *Spiroplasma* can be experimentally transmitted to new fly species,

these newly infected flies may be unable to transmit the infection to their progeny. For example, Williamson and Poulson (1979) report that the strain of *Spiroplasma* that naturally infects *D. willistoni* is able to thrive in the hemolymph of *Musca domestica* upon transinfection of the infection, but this species fails to transmit the infection to its progeny.

4.2.2 Methods

4.2.2.1 Fly Strains

Infected flies used in this experiment were obtained from the Rochester University Fly Collection thanks to Dr. Tamara Haselkorn. Each strain of infected flies was divided into two groups; one group was cured from the infection with antibiotics (see below) and the other group was allowed to keep the infection. Infected and uninfected fly cultures were expanded to two bottles and were allowed to reproduce for at least eight generations before using them in the experiments. All females in the experiment were mated to uninfected males from their own strain. This was done to standardize the mating process and to prevent any issue that could potentially confound the results of the experiment (e.g., cytoplasmic incompatibility, although this phenomenon has not been reported for the *Spiroplasma-Drosophila* system).

4.2.2.2 Curing of Infected Flies to Generate Uninfected Flies

A combination of tetracycline and erythromycin (final concentration = 0.2 and 0.16 mg/ml, respectively) was mixed with the fly food. Flies were kept on this antibiotic-food mix for two generations. The third and further generations were fed antibiotic-free food. PCR screening was used to confirm that the flies were cured.

4.2.2.3 Extracting DNA and PCR Screening of Bacterial Infection

In this study, a crude DNA extraction was performed by incubating a crushed fly in a mixture of 96 μL buffer solution (10 mM Tris.Hcl, 25 mM NaCl, 1 mM EDTA pH 8) and 4 μL of proteinase K solution (10 mg/mL, Denville Scientific). The incubation was carried out at 50° C for 60 minutes followed by denaturation of the proteinase K at 85° C for 30 minutes. To detect *Spiroplasma* infection, we ran 15 μL PCR reactions using 2 μL of the crude DNA extract as a template. We used three primer sets: 1) Spoul (Montenegro et al. 2005), run at an annealing temperature of 55° C; 2) p58IV, designed by the author, targeting a fragment of about 360 bp that is internal to the *p58* gene targeted by the *p58* primer set (Montenegro et al., 2005); with an annealing temperature of 53°C and the sequences being p58IV_F 5'-AAAGGTTTACATTCACCAAGTCG-3' and p58IV_R 5'-AATTGTTCATTAACTTTATCTTGTGG-3' ; and 3) SpoulIXO, designed by the author and targeting an inner region of the 16s rDNA gene targeted by Spoul; the annealing temperature for SpoulIXO is 50° C and the sequences are SpoulIXO_F 5'-GCTCAACCCYTAACCGCC-3' and SpoulIXO_R 5'-CCTGTATCCTTGTTAACCTC-3. Spoul and p58IV were used to diagnose infection in *D. melanogaster*, *D. hydei* and *D. mojavensis*, while SpoulIXO was used to diagnose infection in *D. atripex*.

4.2.2.4 Cross-Infections

The transfer of *Spiroplasma* was done from infected fly donors to uninfected fly recipients. This procedure was performed on CO_2-anesthetized flies. To transfer the hemolymph, the experimenter used a manual micro-injector constituted of micro tips prepared from pulled micro-capillaries with a plunger prepared from a tailor's pin, and a viscous phase (mineral oil) to carry

the pressure through the micro-injector when the plunger is pushed through the shaft of the micro-injector. The oil did not come in contact with the hemolymph during the injection process and it is not injected into the recipient fly. The infected donor and the recipient flies were punctured at the soft juncture of the mesopleuron and scutellum to minimize the magnitude of the wound on the recipient flies. After transinfection, the flies were left to recover from the wound for about two days before further manipulation. Typically, the survival rate is very high (~80% to 100%) and the success of transferring the infection is ~100% (at least for the *Spiroplasma-Drosophila* system used in the present research).

4.2.2.5 Sampling Scheme

The experimenter reciprocally transinfected 5 fly strains with the native *Spiroplasma* strains carried by each fly strain. The fly strains used in this study belong to two species groups from two subgenera: 1) the repleta group from the subgenus *Drosophila* (two *D. hydei* strains and one *D. mojavensis* strain) and 2) the melanogaster group from the subgenus *Sophophora* (*D. melanogaster* and *D. atripex*, one strain each). The sampling encompassed three of the four *Spiroplasma* clades reported by Haselkorn et al. (2009): 1) from the poulsonii clade, the male-killing MSRO strain (*Mel*, in our study) that infects *D. melanogaster*, and a closely related non-male-killing strain (*Hyd1* in our study) that infects *D. hydei*; 2) from the citri clade, the *Spiroplasma* strain infecting *D. mojavensis* (*Moj*, in this study) and *Hyd2* that infects *D. hydei*; and 3) the strain infecting *D. atripex* (*Atri*, in this study) was sampled from the ixodetis clade.

For each pair of *Spiroplasma*-fly recipient strains used in a cross-infection, two groups of flies were injected with the bacteria: the “fly recipients” and the “fly donors.” The “fly recipients”

were the flies injected with the *Spiroplasma* strain taken from the "fly donor," the natural host of the *Spiroplasma* being injected. The "fly donors" were injected with the *Spiroplasma* hosted by the "fly recipient." As can be seen, the group denominations are interchangeable. For each fly strain, a recipient group for each foreign *Spiroplasma* strain from fly donors was composed of nine sets of 10 antibiotic-cured, 3-day-old virgin females (9 repeats of 10 individuals each). A set of nine repeats was injected with hemolymph from infected conspecifics (positive control) once per each species. SM S14 shows an example of these cross-infections for 3 groups of flies. The injected females constituted Generation zero (G0).

Injected flies were allowed to incubate the infection for 10 days. After 10 days, each female was placed in a separate vial with two mature uninfected males. After 24 hours the mating pair was placed into a fresh vial and the old vial discarded. Vials containing mating pairs were checked every 48 hours for eggs and larvae. Once the vials were detected to contain larvae, the males were discarded and the females were screened for *Spiroplasma*. The vials containing larvae from G0 females testing negative for *Spiroplasma* were discarded. The infection success rate of G0 females is usually high, and if needed, more G0 females were infected to complete each set.

The vials containing eggs and larvae of females testing positive were incubated at room temperature (24° C) at 12 h light - 12 h dark. Emerging virgin females (G1) were collected and placed in vials for 10 days to allow for *Spiroplasma* growth inside the fly. After 10 days, females were collected and frozen at 80° C for later screening. The G1 transmission rate was used in further analyses.

4.2.2.6 Estimating Transmission Success Rates at G1

Transmission success rates (or transmission rate or transmission fidelity) at G1 for a given donor or recipient strain was estimated by dividing the number of female progeny testing positive for the infection by the number of females screened (in our sampling scheme, a set of 90 females).

4.2.2.7 Potential Host Range

The potential host range for the *Spiroplasma* strains was estimated by using the Rohde index of specificity corrected for number of species (normalized) (Rohde & Rohde, 2005; Rohde & Rohde, 2008). For these estimates, the transmission rate at G1 post-cross-infection was used as prevalence, because the formulas are similar (# of infected individuals / # of hosts surveyed). Also, we included the self-infection and the zero values of prevalence for these estimations. The Rohde index of specificity (Rohde &Rohde, 2005; Rohde &Rohde, 2008) takes into account the uneven distribution of parasites across different hosts. In the numerator of this index, prevalence is weighted by the inverse rank (hosts which are less used contribute less to the sum). Thus, the value of the index is more stable and less affected by accidental or ephemeral occurrences of parasite species. The Rohde index of specificity (Rohde &Rohde, 2005; Rohde &Rohde, 2008) is not sensitive to the number of host species evaluated and can be applied for the comparison of parasites using different numbers of host species. This index ranges from 1 (the highest degree of host specificity, or most specialist parasite) to 0 (the lowest degree of host specificity, or most generalist parasite). For our purpose, we defined a parasite as a specialist if it produces a score of 1 (i.e., highly host specific), which indicates that the parasite is able to infect a large number of individuals from a few of the available host species, while infecting a small number of individuals from other available species. In contrast, we define a parasite as a generalist if it

produces a score of 0 (i.e., lowest host specific), which indicates that the parasite is able to infect a consistent number of individuals from several host species. For our work, the specialist parasites have the smallest potential host ranges and the generalist parasites have the largest potential host ranges.

The Rohde index was complimented in the "discussion" section by examining studies where spiroplasmas were introduced into *Drosphila* flies. Those studies were examined to find patterns of how the new infection behaved in new hosts. For example, can the low or high transmission fidelity at G1 predict the success or failure of the infection in future generations? Does infection by some *Spiroplasma* strains show the capacity to persist in new hosts in the long-term? Does infection by some *Spiroplasma* strains show the tendency to disappear from new hosts in the long-term? The Rohde index was further complimented by examining the patterns of infection success observed on the current experiment. The patterns from the literature and from this work were used to refine the classification of parasite specialist or generalist obtained from the Rohde index.

4.2.2.8 Estimating Genetic Distances: Obtaining Sequences for Bacterial Strains and Fly Strains

Genetic distances between hosts and between parasites were used in the statistical analysis as indicators of phylogenetic relationships between hosts and between parasites, respectively (De Vienne et al., 2009). GenBank was surveyed to obtain *Spiroplasma* and fly hosts' gene sequences for measuring genetic distances between the strains under study. The experimenter was able to acquire 16s rDNA sequences for all *Spiroplasma* strains used in this work (deposited

at GenBank by Haselkorn et al., 2009). The GenBank accession numbers for the acquired sequences are: FJ657180 (for MSRO or *Mel*, a strain carried by *D. melanogaster*), FJ657202 (for TEN104106, a strain of *Hyd1* carried by *D. hydei*), FJ657220 (for CI4011830, a strain of *Moj* carried by *D. mojavensis*), FJ657237 (for OPNM0407A4, a Hyd2 strain carried by *D. hydei*), and FJ657246 (for the *Atri* strain carried by *D. atripex*).

For measuring the fly distance, the Cytochrome c oxidase gene subunit I (COI) was used, because it has been adopted by the Barcoding of Life Consortium as the standard gene for identifying species of flies and other animal groups (http://www.barcodeoflife.org/content/about/what-dna-barcoding). The experimenter attempted to obtain the COI gene sequence for all the flies in the study and was able to find records for *D. Mojavensis* strain CI1, a fly from Catalina Island, California (Reed, Nyboer, & Markow, 2007), which carries the infection by what is referenced in this work as *Spiroplasma* strain *Moj*. The experimenter downloaded a fly sequence identified by accession number DQ383710 and used it in our analysis. The experimenter also found a COI gene sequence for a *D. melanogaster* strain Canton-S, which is able to maintain *Spiroplasma* MSRO infection (e.g., Martin, Chong, & Ferree, 2013); its GenBank accession number is KJ767244 (Yeh et al., 2014). No other fly had publicly available COI gene sequences; therefore, the experimenter decided to sequence COI for our remaining three fly strains (the two *D. hydei* carrying the two different *Spiroplasma* strains and *D. atripex*).

Genomic DNA extraction of *D. hydei* and *D. atripex* was carried out with a Terra™ PCR Direct Polymerase Mix kit from Takara. The primers used to PCR-amplify the COI gene for two strains of *D. hydei* were HydCOI_1F (forward) 5'-TGA-GCG-GGA-ATA-GTG-GGA-ACA-3' (a

modified version of primer C1-N-1560 by Bonacum, et al., 2001) and HCO2198_R (reverse) 5'-TAA-ACT-TCA-GGG-TGA-CCA-AAA-AAT-CA-3' (Flores, Evans, & McAllister, 2008). To amplify the COI gene for *D. atripex* we used a different forward primer, HydCOI_2F, 5'- TTT-CAA-CAA-ATC-ACA-AAG-ATA-TTG-G-3' (a modified version of primer LCO1490 by Flores et al., 2008) and the same reverse primer used for *D. hydei*.

All primer sets were used at an annealing temperature of 48° C. Exonuclease-I was used to digest unincorporated primers from the PCR products, and the sequencing in both directions was performed by Eton Bioscience by using the PCR amplification primers. The sequences were edited with Sequencher 4.8 (Gene Codes Corporation, http://www.genecodes.com/).

4.2.2.9 Estimating Genetic Distances: Measuring the Genetic Distances Between Spiroplasma Strains and Between Fly Strains

To measure the genetic distance between the *Spiroplasma* strains that naturally infect the fly strains, the 16s rDNA gene sequences were aligned using the online SINA Alignment Service at https://www.arb-silva.de/aligner/ (Pruesse, Peplies, & Glöckner, 2012), and the resulting alignment was downloaded in FASTA format and opened in the Alignment Explorer of MEGA 7 (Kumar, Stecher, & Tamura, 2016), where the gap-only sites and the ends that contained gaps were deleted. The file was saved in MEGA format for further analysis. To measure the genetic distance between the fly strains, the fly COI sequences were aligned by using MUSCLE (Edgar, 2004) as implemented in MEGA 7 (Kumar et al., 2016). The ends of the aligned sequences that contained gaps were trimmed and the file was saved in MEGA format for further analysis.

The “Distances” function of MEGA 7 was ran to estimate pairwise distances (nucleotide differences between sequences, *sensu* Nei & Kumar, 2000) between the aligned 16s rDNA sequences of spiroplasmas, using the Kimura 2-parameter method (Kimura, 1980), including Transitions + Transversions substitution types, with uniform rates among sites and partial deletions for gaps and missing data. The distances between the COI aligned sequences for the fly strains were estimated in a similar fashion.

4.2.2.10 Phylogenetic Tree of Spiroplasma and Drosophila Strain Used in this Work

Phylogenetic trees were built for the hosts and for the parasites to show the relationships between the individuals of each group (i.e.; between hosts and between parasites) as well as to show the host-parasite relationships. Separated phylogenetic trees for the bacterial strains and fly strains used in this work were constructed from the same 16S rDNA sequences (for *Spiroplasma)* and the COI sequences (for *Drosophila*) used for measuring genetic distances. MEGA 7 (Kumar et al., 2016) was used to run a model selection analysis of substitution patterns, with default values, for each set of DNA sequences. For the bacterial 16S rDNA, the HKY (Hasegawa, Kishino, & Yano, 1985) model was selected, and for the flies’ COI sequences, the general time-reversible model (GTR hereafter; Lanave, Preparata, Saccone, & Serio, 1984; Tavaré, 1986) + Gamma parameter model was selected. These two models were run separately in MEGA 7 (Kumar et al., 2016) by using the statistical method of maximum likelihood and 500 bootstrap replications, with the other options set at default. The trees were edited to show the bacteria and fly phylogenies side-by-side, to show the relationships between the bacteria and their host flies.

4.2.2.11 Statistical Analysis to Determine the Factors that Affect Infection Success Rate (Transmission Rate) at G1

A logistic regression was run by using the PROC LOGISTIC as implemented in the SAS EG® 7.1 (SAS Institute Inc., 2014). The infection success rate at G1 was used as the dependent variable and entered in the analysis as "# of events / # of trials". This format was used for the response because many trials were collected at once and each set of these trials was conducted under similar conditions.

It was planned to use the genetic distance between *Spiroplasma* strains (genetic distance between the inoculated *Spiroplasma* strain and the *Spiroplasma* strain originating from the recipient *Drosophila* species, labelled DistSpir) and genetic distance between fly strains (genetic distance between the fly strains receiving the inoculated strain and the fly strains donating the inoculated *Spiroplasma* strain, labelled DistFly) as predictors. In order to do this, first it was tested if the genetic distances between *Spiroplasma* strains were correlated to the genetic distances between fly strains. This is important because highly correlated predictors can give rise to the problem of multicollinearity, which makes it difficult to obtain reliable estimates of their distinct effects on the dependent variable (Allison, 2012). Even though multicollinearity doesn't bias the coefficients, it does make them more unstable; standard errors may get large, and variables that appear to have weak effects, individually, may actually have quite strong effects as a group (Allison, 2012). The Mantel test was used to assess the correlation between the bacteria and fly distances, because the bacteria and fly strains are involved in multiple pairs, creating pseudo replications. The null hypothesis that the bacterial strains' distances matrix and the fly strains'

distances matrix are not correlated could not be rejected (p=0.178), so both distances were used as predictors in the analysis, as originally intended.

Additional factors included in the analysis are described next. Fly recipient strain (FlyRecip) is the fly strain where the *Spiroplasma* strains were introduced (the "new host"). The *Spiroplasma* strain to be injected into a new host was considered to be the "spiroplasma donor" (SpirDonor). These two variables represent the" intrinsic characters" of the hosts and the bacteria. The species groups for the recipient flies were entered as FlyRecipGROUP and the clades for spiroplasmas were entered as SpirDonorGROUP, and these two variables represent the intrinsic characteristics of the higher classes of both the fly and the bacteria.

The selection of variables in the model was automated by using stepwise selection with a p-value of 0.3 to enter the model and 0.35 to stay in the model. The initial model included all the main factors described above and all possible two-way interactions between these factors. The resulting model from the stepwise selection was further simplified by first removing non-significant interactions and then non-significant main factors. The categorical factors in the model were parametrized by using GLM coding, where the beta estimates the difference in the effect of each non-reference level compared to the effect of the reference level, which is the last level (in our data, these were *Atr* and Atripe, for the SpirDonor and FlyRecip variables, respectively).

Under- and over-dispersion were evaluated using Pearson Goodness-of-fit statistics (Value/DF close to 1 indicates no serious problem with under- or over-dispersion in the model). Goodness

of fit for the model was tested with the Hosmer-Lemeshow test (Hosmer, Lemeshow, & May, 2008). Influential and leverage values were evaluated by using the plots option. Leverage and influential values were removed once. We used the SAS LSMESTIMATE statement to test hypotheses about the relationships between main factors by using non-positional syntax for contrasts. The LSMEANS statement and diff=control options from SAS were used in conjunction to compare the success rates between the interactions of interest.

4.3 Results

4.3.1 Genetic Distances

The length of the 16S rDNA sequences used in this analysis was 1297 sites. The estimated distances are shown in Tables 4.1 (spiroplasmas) and 4.2 (flies). Table 4.1 shows that *Atri*, the strain infecting *Drosophila atripex*, is the most genetically distant from all the other *Spiroplasma* strains, and *Mel* (hosted by *D. melanogaster*) and *Hyd1* (hosted by *D. hydei*) are the closest in genetic distance. Excluding *Atri*, the other four *Spiroplasma* strains show less than 2% genetic distance between them (Table 4.1).

Table 4.1 Pairwise genetic distance (nucleotide differences between sequences) for *Spiroplasma* strains based on partial sequences of the 16s rDNA gene.

Accession number	*Spiroplasma* strain (Fly host strain)	*Spiroplasma* strain				
		Atr	*Mel*	*Hyd1*	*Moj*	*Hyd2*
FJ657246	*Atr (D. atripex)*	0.000				
FJ657180	*Mel (D. melanogaster)*	0.153	0.000			
FJ657202	*Hyd1 (D. hydei)*	0.153	0.000	0.000		
FJ657220	*Moj (D. mojavensis)*	0.154	0.013	0.013	0.000	
FJ657237	*Hyd2 (D. hydei)*	0.157	0.015	0.015	0.006	0.000

The length of COI sequences used in this analysis was 621 sites. *D. mojavensis* (strain Mojave) is the most genetically distant of all the fly hosts, with more than 15% nucleotide differences from all other fly strains (Table 4.2). The closest strains (with less than 0.5% nucleotide difference) are Hydei1 and Hydei2, which belong to the same species, *D. hydei* (Table 4.2). The two strains from the melanogaster species group (Atripe and Melano) are less distant (9.79% nucleotide difference) than the strains from the repleta species group (Mojave vs Hydei1 and Hydei2, with 15.17% and 15.77% nucleotide difference; Table 4.2).

Table 4.2 Pairwise genetic distance (nucleotide differences between sequences) for host fly strains based on partial sequences of the COI gene.

Accession number	Fly strain (Fly species)	Fly strain				
		Atripe	Melano	Hydei2	Hydei1	Mojave
	Atripe (*D. atripex*)	0.000				
KJ767244	Melano (*D. melanogaster*)	0.098	0.000			
	Hydei2 (*D. hydei*)	0.145	0.126	0.000		
	Hydei1 (*D. hydei*)	0.147	0.132	0.005	0.000	
DQ383710	Mojave (*D. mojavensis*)	0.167	0.155	0.158	0.152	0.000

4.3.2 Phylogenetic Relationship Between Spiroplasmas and Between Their Hosts

The phylogenetic relationship between the spiroplasmas and between the *Drosophila* flies is known, as well as the relationships between these parasites and their hosts (Seetharam & Stuart, 2013; Haselkorn et al., 2009; Haselkorn, 2010a). Table 4.3 summarizes these relationships.

Table 4.3 *Spiroplasma*-infected fly strains (Genus *Drosophila*, Family Drosophilidae) and their natural bacterial strains used in this study. The naming of the fly and *Spiroplasma* strains for this study shown in their respective columns; the names in parenthesis indicate the naming on Dr. Haselkorn's collection. Table based on Seetharam and Stuart (2013); Haselkorn et al. (2009); Haselkorn (2010a).

Fly species group	Fly species	Fly strain	*Spiroplasma* strain	*Spiroplasma* clade
Repleta	*D. hydei*	Hydei1 (TEN104-102)	*Hyd1*	poulsonii
	D. hydei	Hydei2 (ABHS)	*Hyd2*	citri
Repleta	*D. mojavensis*	Mojave (CI-1)	*Moj*	citri
Melanogaster	*D. melanogaster*	Melano (288)	*Mel* (MSRO)	poulsonii
Melanogaster	*D. atripex*	Atripe	*Atr*	ixodetis

The phylogenetic trees in Fig. 4.1 confirms the relationships shown in Table 4.3. The incongruence of both trees is also graphically depicted in Fig. 4.1. The *Spiroplasma* strain *Hyd1* (clade poulsonii) parasitizes *D. hydei* strain 1 (Hydei1) (Subgenus *Drosophila*, group repleta), but the other two flies from the repleta group (*D. hydei* strain 2 or Hydei2 and *D. mojavensis* or Mojave) are parasitized by spiroplasmas from the citri clade (*Hyd2* and *Moj*, respectively). The other poulsonii *Spiroplasma*, the strain *Mel*, parasitizes *D. melanogaster* (strain Melano) (Subgenus *Sophophora*, group melanogaster), while the other *Sophophora* fly, *D. atripex* (Atripe) is parasitized by *Atr*, a *Spiroplasma* that belongs to the ixodetis clade.

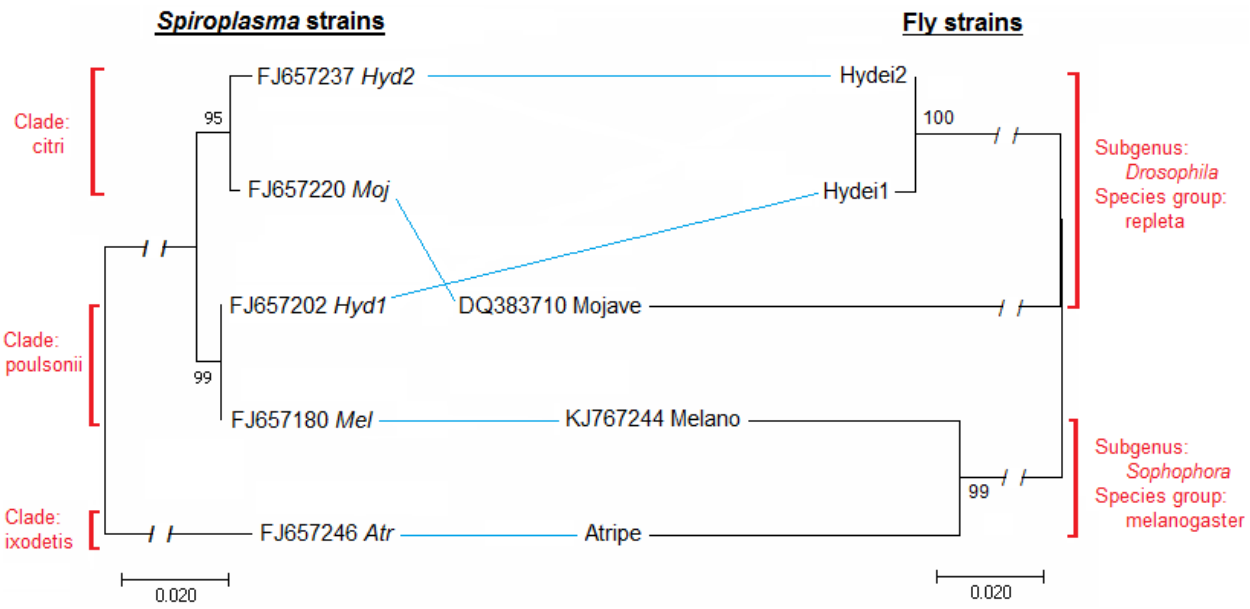

Fig. 4.1 Composite phylogenetic trees of *Spiroplasma* and *Drosophila* fly strains used in this work. The blue lines match the bacteria to their natural fly hosts.

4.3.3 Spiroplasma Transmission Success Rates and Transmission Rates

The pairwise transinfections (cross-infections) can be visualized from two perspectives based on the direction of the flow of the infection: 1) *Spiroplasma* strain-to-recipient fly perspective, and 2) recipient-fly-from-*Spiroplasma* strain perspective. For example, for *D. atripex* (Atripe) and their *Atr* spiroplasmas cross-infecting with *D. melanogaster* (Melano) and their *Mel* spiroplasmas, the relationships formed are (from Atripe, *Atr* perspective): 1) *Spiroplasma* strain-to-recipient fly, *Atr*-Melano and 2) recipient fly-from-*Spiroplasma* strain, Melano-*Atr*. These two transinfections are actually the same, only the direction of the relationship has changed. The reciprocal transinfection (from Melano, *Mel* perspective) comprises 1) *Mel*-Atripe and 2) Atripe-*Mel*. The pairwise transmission rates resulting from all cross-infections in this study are shown in detail in Fig. 4.2 and Fig. 4.3.

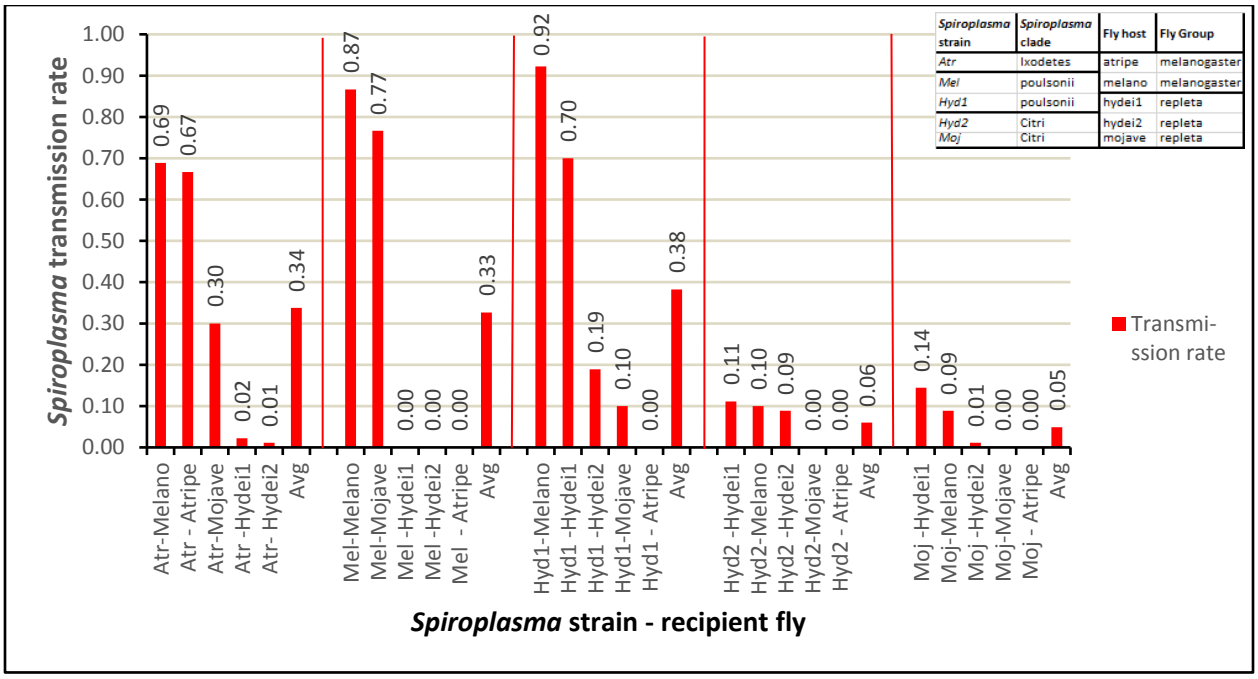

Fig. 4.2 Donor-to-fly-recipient perspective**.** *Spiroplasma* strains transmission rates (# of infected progeny/total progeny #) from mother to progeny at generation 1 post-infection. Each set of bacteria-recipient flies is separated by a red line and arranged according to the transmission rates from highest to lowest.

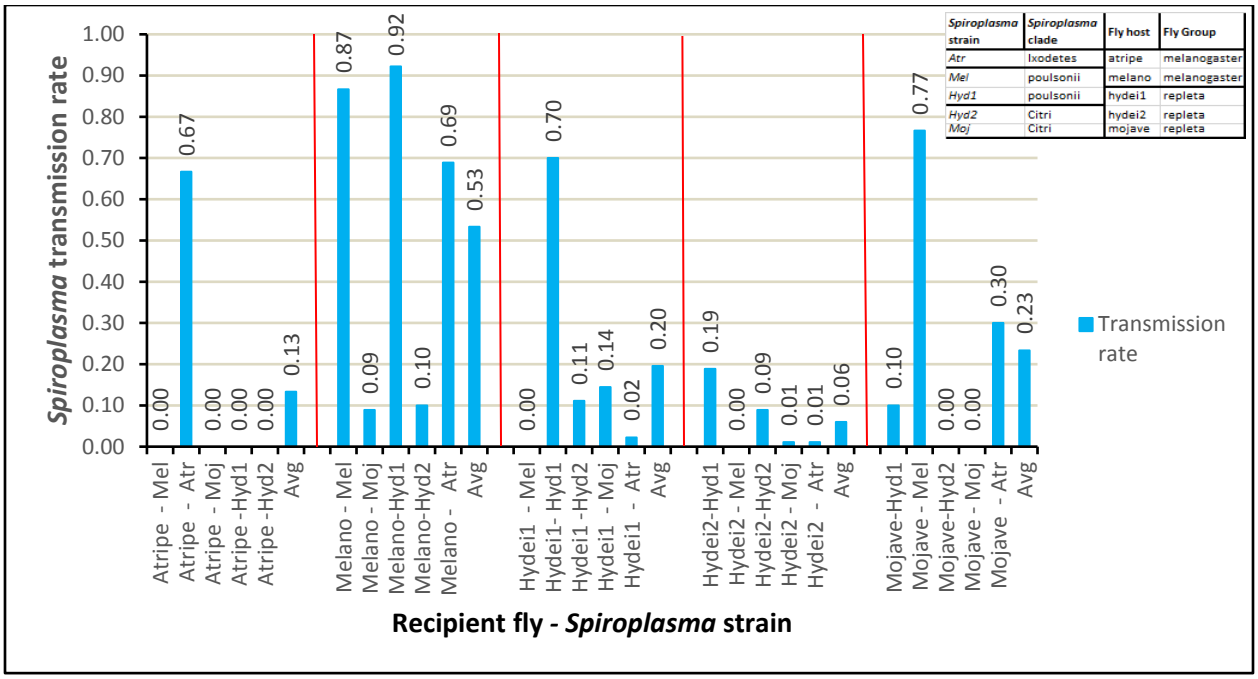

Fig. 4.3 Recipient-from-donor perspective. *Spiroplasma* strains transmission rates (# of infected progeny/total progeny #) from mother to progeny at generation 1 post-infection. Each set of bacteria-recipient flies is separated by a red line.

4.3.4 Infection Success of Spiroplasmas at G1

The success of *Spiroplasma* infection at G1 in the new hosts is varied. SM S15 shows the success or failure of *Spiroplasma* infection at generation 1 post-cross-infection, from the donor-to-recipient perspective.

The *Spiroplasma* strain *Atr* (ixodetis clade) carried by the fly *D. atripex* (strain Atripe, Atripe hereafter, melanogaster species group) is successful at producing infection in all recipient flies at generation 1 post-infection. The bacterial strain *Mel* (a male killer from the poulsonii clade) produces infection in only two flies, its natural host, *D. melanogaster* (strain Melano, Melano hereafter), from the melanogaster species group, and *D. mojavensis* (strain Mojave, hereafter

Mojave), from the repleta group. The strain *Hyd1* (a non-male killer from the poulsonii clade), whose host is *D. hydei* strain 1 (strain Hydei1, hereafter Hydei1, from the repleta group) produces infection in all flies except Atripe. The strain *Hyd2* (citri clade) hosted by *D. hydei* strain 2 (strain Hydei2, hereafter Hydei2, from the repleta group) failed to produce infection in the G1 progeny of Mojave and Atripe. The strain *Moj* (citric clade) fails to produce infection in its native host Mojave and Atripe.

SM S16 shows the success or failure of *Spiroplasma* infection at generation 1 post-cross-infection, from the perspective of the recipient flies being able to transmit the infection to their progeny after acquiring the bacteria through trans-infection (recipient-from-donor perspective). Atripe (melanogaster group) mothers are unable to transmit the infection from any strain of spiroplasmas except their own (*Atr*, ixodetis clade). Melano (melanogaster group) are able to transmit the infection from all five *Spiroplasma* strains tested, which belong to three different clades: poulsonii, citri, and ixodetis. Both Hydei1 and Hydei2 (repleta group) are only unable to transmit the infection caused by *Mel*, the male-killing strain from the poulsonii clade. Mojave (repleta group) is unable to transmit the infection caused by *Hyd2* and *Moj* (its native strain), both from the citri clade.

4.3.5 Transmission Rates of Spiroplasmas

Fig. 4.2 and SM S15 show the transmission rate of *Spiroplasma* at generation 1 post-cross-infection from the donor-to-recipient perspective.

The transmission rate (# of progeny infected/Total progeny # tested) at G1 of *Spiroplasma* strains by their native and new hosts are highly variable (Fig. 4.2 and SM S15). The highest average transmission rate (across all hosts used in the experiment) is for *Hyd1* (0.38), even though one host, Atripe, does not transmit it to its progeny (Fig. 4.2 and SM S15). *Hyd1* also has the highest transmission rate data point (0.92), which is achieved by Melano, a fly that belongs to the melanogaster group, which is distinct from the repleta group of the fly host of *Hyd1*, Hydei1. *Hyd1* natural host (Hydei1) achieves a high transmission rate (0.70) of this strain, but other flies that belong to its own fly group (repleta), transmit *Hyd1* at low rates (Hydei2 at 0.19 and Mojave at 0.10 (Fig. 4.2 and SM S15).

The next highest transmission rate is a tie for *Atr* and *Mel* (0.34 and 0.33, respectively) (Fig. 4.2 and SM S15). However, on the one hand, *Atr* is transmitted by all hosts tested, namely, Melano (0.69) and its natural host Atripe (0.67) of the melanogaster group, and Mojave (0.30), Hydei1 and Hydei2 (both with low transmission rates (0.02 and 0.01, respectively) of the repleta group (Fig. 4.1). On the other hand, *Mel* is transmitted only by two flies, its natural host, Melano (0.87) (melanogaster group) and Mojave (0.77) (repleta group) (Fig. 4.2 and SM S15).

The two lowest average transmission rates are for *Hyd2* and *Moj* (0.06 and 0.05, respectively) (Fig. 4.2 and SM S15). Both strains are not transmitted by Mojave (which means that under our experimental conditions, *Moj* is not transmitted by its natural host) and Atripe. They are transmitted by Hydei1, Melano and Hydei2 (in this order); these three flies transmit the two spiroplasmas at low rates, with *Moj* being transmitted at the highest (0.14 by Hydei1) and the lowest (0.01 by Hydei 2) rates of the two of them (Fig. 4.2 and SM S15).

Overall, patterns that can be extracted from the data are that *Atr* is transmitted by all hosts used in this experiment, but its natural host Atripe cannot transmit any other spiroplasmas except its own (Fig. 4.2 and Fig. 4.3). *Mel* is transmitted by only two hosts, but its host Melano can transmit all strains used in the experiment (Fig. 4.2 and Fig. 4.3). *Moj* is not transmitted by its own natural host (Mojave). *Hyd2* and *Moj*, when transmitted by new hosts, are transmitted at low rates (Fig. 4.2 and Fig. 4.3)

4.3.6 Spiroplasma Potential Host Range

In the current study, we performed a cross-inoculation experiment, using 5 different strains of *Spiroplasma* (*Spiroplasma* donor) and 5 fly host strains (Fly recipient) to assess the potential host range of the bacteria.

In theory, the normalized Rohde index (Rohde & Rohde, 2005; Rohde & Rohde, 2008) ranges from 1 (highest degree of host specificity) to 0 (lowest degree of specificity). The index values obtained in our experiments range from 0.5448 to 0.7348 (Table 4.4). For the bacterial strains studied here, *Moj* shows the highest Rohde index value (0.7348). Strains *Hyd2* (0.5448) and *Atr* (0.5863) show the lowest value of Rohde index. Strain *Hyd1* (0.6538) shows a value that is in the middle of the range of the index values obtained in this study (0.5448 – 0.7348). Strain *Mel*, with an index value of 0.7066, shows the second highest Rohde index value after Moj.

Table 4.4 Rohde (normalized) index of host specify for *Spiroplasma* strains introduced into *Drosophila* strain hosts.

Spiroplasma strain	Rohde Index (normalized)
Hyd2	0.545
Atr	0.586
Hyd1	0.640
Mel	0.707
Moj	0.735

We interpret the Rohde index values obtained in this study as indicating that *Hyd2* and *Atr* are the most generalist from the 5 strains tested in our study, while *Moj* and *Mel* are the most specialized, and *Hyd1* is intermediate. For our ends, a generalist parasite possesses the ability to infect a more consistent number of individuals from a larger number of host species than the specialist ones; therefore, our results show that *Hyd2* and *Atr* have larger potential host ranges than the rest of the strains studied here. Conversely, a specialized parasite infects a disproportionate number of individuals from a (sometimes) reduced number of hosts as compared to a generalist parasite (i.e., it disproportionately infects a larger number of individuals from a few of its host species); therefore, *Moj* and *Mel* have the smallest potential host range when compared to the rest of the strains studied in our experiment. *Hyd1* has an intermediate potential host range. This means its range lays somewhere between the generalist (*Hyd2* and *Atr*) and the specialist (*Moj* and *Mel*) *Spiroplasma* strains used in this study).

4.3.7 Factors Affecting the Success and Transmission Rate of New Spiroplasma Infections.

A binomial logistic regression was performed to ascertain what variables have an effect on the likelihood of success of new *Spiroplasma* infections being transmitted to the progeny of new hosts. Two-hundred and twenty-five sets of Event/Trials, totaling 2250 observations, were available for analysis. Two variables and their interaction were included in the final model: FlyRecip (Fly recipient, the fly strain receiving the *Spiroplasma* infection, or new host) and SpirDonor (*Spiroplasma* donor, the *Spiroplasma* strain being introduced into the new host), and their interaction was called FlyRecip*SpirDonor. All of these were significant (Table 4.5, Fig. 4.4). The model showed a quasi-complete separation of data points, so we used the model option Firth to perform Firth's penalized maximum likelihood estimation to reduce bias in the parameter estimates (Heinze & Schemper, 2002; Firth, 1993). The model did not show evidence of over-dispersion (the value of Pearson Goodness-of-Fit statistics divided by its degrees of freedom was 1.12, close to 1). The Hosmer-Lemeshow Goodness-of-Fit test was not significant ($p = 0.47$), indicating that the model is a good fit to the data. We detected sets of observations that were influential or leverages, and we ran the analysis after their removal, but the general conclusions of the analysis did not change; therefore, we are reporting the results for the full data set (Table 4.5, Table 4.6).

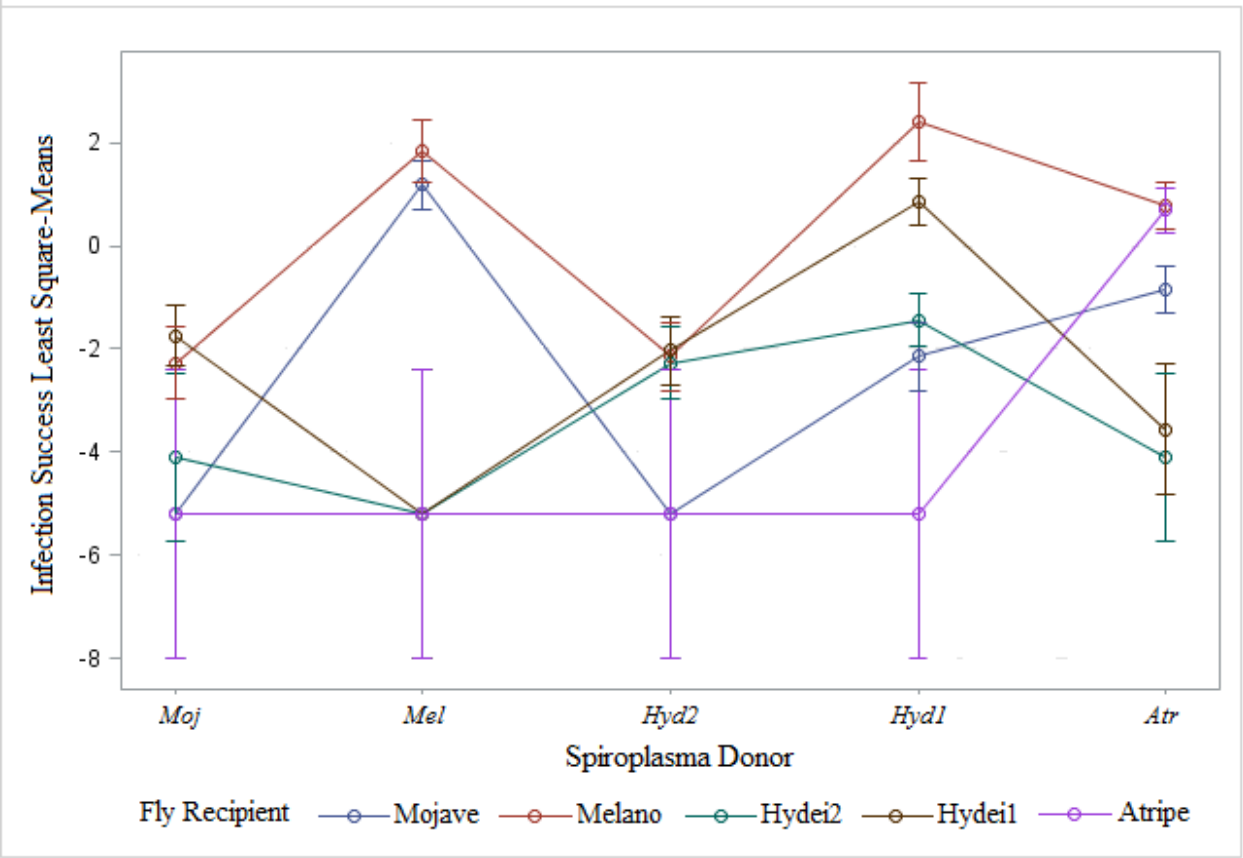

Fig. 4.4 Interaction plot for fly recipient vs *Spiroplasma* donor infection success rates (the vertical lines show the 95% confidence interval of estimates).

Table 4.5 Logistic regression for the factors predicting *Spiroplasma* transmission fidelity at generation 1 post artificial infection.

Effect	DF	Wald Chi-Square	Pr > ChiSq
FlyRecip	4	154.867	<.0001
SpirDonor	4	41.368	<.0001
FlyRecip*SpirDonor	16	204.309	<.0001

Table 4.6 Firth maximum likelihood estimates of beta (coefficients) for *Spiroplasma* donor and fly recipient and their Interaction (reference categories are *Atr* and Atripe, which are set to zero).

Parameter			DF	Beta Estimate	Standard Error	Wald Chi-Square	Pr > Chi Sq	Exp(Est)
Intercept			1	0.685	0.223	9.408	0.002	1.984
Fly recipient	Mojave		1	-1.522	0.320	22.581	<.0001	0.218
	Melano		1	0.100	0.319	0.099	0.753	1.106
	Hydei2		1	-4.774	0.858	30.994	<.0001	0.008
	Hydei1		1	-4.252	0.682	38.813	<.0001	0.014
Spiroplasma donor	*Moj*		1	-5.883	1.443	16.617	<.0001	0.003
	Mel		1	-5.883	1.443	16.617	<.0001	0.003
	Hyd2		1	-5.883	1.443	16.617	<.0001	0.003
	Hyd1		1	-5.883	1.443	16.617	<.0001	0.003
Fly recipient * *Spiroplasma* donor	*Moj*	Mojave	1	1.522	2.042	0.556	0.456	4.580
	Moj	Melano	1	2.825	1.505	3.523	0.061	16.866
	Moj	Hydei2	1	5.883	1.858	10.022	0.002	358.989
	Moj	Hydei1	1	7.702	1.608	22.934	<.0001	2213.687
	Mel	Mojave	1	7.893	1.482	28.356	<.0001	2679.590
	Mel	Melano	1	6.935	1.493	21.585	<.0001	1028.029
	Mel	Hydei2	1	4.774	2.191	4.746	0.029	118.355
	Mel	Hydei1	1	4.252	2.129	3.989	0.046	70.220
	Hyd2	Mojave	1	1.522	2.042	0.556	0.456	4.580
	Hyd2	Melano	1	2.949	1.501	3.858	0.050	19.081
	Hyd2	Hydei2	1	7.699	1.703	20.444	<.0001	2206.874
	Hyd2	Hydei1	1	7.413	1.615	21.074	<.0001	1657.592
	Hyd1	Mojave	1	4.571	1.502	9.267	0.002	96.625
	Hyd1	Melano	1	7.508	1.511	24.706	<.0001	1822.515
	Hyd1	Hydei2	1	8.537	1.685	25.663	<.0001	5099.920
	Hyd1	Hydei1	1	10.287	1.597	41.472	<.0001	29344.390

Table 4.6 shows the Maximum Likelihood Estimates of Beta coefficients for the logistic regression with a Firth correction (Heinze & Schemper, 2002; Firth, 1993). Most of main effects and interactions are significant. Among the main effects, only Fly strain “Melano” (p>0.7528) is

not significant and among the interactions, only Moj*Mojave, Moj*Melano and Hyd2*Mojave are not significant ($p>0.4561$, $p>0.0605$ and $p>0.4561$, respectively).

The column Exp(Est) provides the odds ratio for the parameters in the first column in Table 4.6. For example, when the main effect level for the *Spiroplasma* donor is *Hyd1*, it shows lower odds of producing infection than *Atr*, the reference level for *Spiroplasma* donor (coefficient for *Hyd1* is negative, -5.883). In fact, the odds of *Hyd1* producing infection in a host are 0.003 times as small as the odds of *Atr* producing infection in a host (Table 4.6). The interaction *Hyd1**Mojave is determined by the difference *Hyd1* – Mojave = (*Hyd1-Atr* when Fly recipient is Mojave) – (*Hyd1-Atr* when Fly recipient is Atripe). This means that the odds of *Hyd1* producing infection in Mojave are 95.62 times as large as the odds of *Atr* producing an infection in Mojave (Table 4.6).

4.3.8 Patterns of Infection Success Rates or Transmission Fidelity Observed on the Current Experiment

4.3.8.1 Atr (ixodetis clade) infects the largest number of new hosts than all other spiroplasmas and has higher odds of producing infection than most other Spiroplasma strains used in the study

The *Atr* strain is the only *Spiroplasma* that produced successful infections in all 5 of the hosts tested. The pairwise comparisons (contrasts) of *Atr* infection success rates versus all other *Spiroplasma* strains show that *Atr* has higher odds of producing infection in new hosts than all other strains tested. *Atr* has 9.9417 times higher odds of producing infection in new recipients than *Moj* ($p<0.0001$), 7.1471 times higher odds than *Hyd2* ($p<0.0001$), and 3.0418 times higher

odds than *Mel* (p=0.0424). *Atr* has no significantly different odds of producing infection in new hosts than *Hyd1* (p=0.4384).

4.3.8.2 Atr is the "universally transmitted Spiroplasma"

The ability of *Atr* to be transmitted by all hosts tested in this study makes it the "universally transmitted *Spiroplasma*" for this work.

4.3.8.3 Mel (poulsonii clade) infects the smallest number of new hosts and has varied, but mostly lower odds of infection success when compared to other Spiroplasma strains used in the study

The *Mel* strain produced successful infections in only one new host (Mojave=*D. mojavensis*) and its own (Melano=*D. melanogaster*). The pairwise comparisons (contrasts) of *Mel* infection success versus all other *Spiroplasma* strains showed that *Mel* has mostly lower odds of producing infection than the other bacterial strains tested. *Mel* has 0.2442 times lower odds of producing infection in new recipients flies than its closest relative in the poulsonii clade, *Hyd1* (p<0.0167), and 0.3287 times lower odds than *Atr* (p<0.0424). *Mel* has no significantly different odds of producing infection in new hosts than *Moj* and *Hyd2* (p=0.0772 and p=0.1912, respectively).

4.3.8.4 Moj and Hyd2 (citri clade) have the lowest odds of producing infection in new hosts

The *Spiroplasma* strains *Moj* and *Hyd2* have the lowest infection success rates of all *Spiroplasma* donor strains (Fig. 4.2 and Fig. 4.3). These two strains do not produce infection in Mojave (*D. mojavensis*) and Atripe (*D. atripex*). Testing for a difference in success rate of the infection between *Moj* and *Hyd2* indicated a non-significant difference between the odds of producing infection by these two strains (*p*=0.5903).

Pairwise comparisons (contrasts) of *Moj* infection success versus the success of the *Spiroplasma* strains outside the citri clade shows that *Moj* produces the lowest odds of infection among all the strains studied. *Moj* has 0.07472 times lower odds to produce infection in new recipients than *Hyd1* ($p<0.0001$), and 0.1006 times lower odds than *Atr* ($p<0.0001$). *Moj* has no significantly different odds of producing infection in new hosts than *Mel* ($p=0.0772$).

Similarly, pairwise comparisons (contrasts) of *Hyd2* odds of producing infection versus the odds of producing infection of the *Spiroplasma* strains outside the citri clade shows that *Hyd2* has similar low odds as *Moj* and is lower than most of the strains studied. *Hyd2* has 0.1039 times lower odds of producing infection in new recipients than *Hyd1* ($p<0.0001$), and 0.1399 times lower odds than *Atr* ($p<0.0001$). *Hyd2* has no significantly different odds of producing infection in new hosts than *Mel* ($p=0.1912$).

4.3.8.5 Hyd1 (poulsonii clade) has higher odds of infection success than most other Spiroplasma strains used in the study

The *Hyd1* strain produced successful infections in all new hosts, except for the host of *Atr*, Atripe (*D. atripex*), the most divergent *Spiroplasma* strain in this study. The pairwise comparisons (contrasts) of *Hyd1* infection success versus all other *Spiroplasma* strains showed that Hyd1 has mostly higher odds of producing infection in new hosts than the other bacterial strains tested. *Hyd1* has 4.0947 times higher odds of producing infection in new recipients than its closest relative in the poulsonii clade, *Mel* ($p<0.0167$), 9.6209 times higher odds than *Hyd2* (citri clade) ($p=0.0001$) and 13.3828 times higher odds than *Moj* (citri clade) ($p<0.0001$). *Hyd1* has no significantly different odds of producing infection in new hosts than *Atr* ($p=0.4384$).

4.3.8.6 The Melano fly recipient allows for infection from more Spiroplasma strains and has higher odds of allowing infections from new parasite strains than the other recipient strains used in the study

Melano is the only fly recipient that allowed infections from 5 out of 5 spiroplasmas tested in this study. Pairwise comparisons of Melano allowing infection success versus all other fly strains show that Melano has higher odds of allowing infection than all other fly strains tested. Melano has 12.9881 times higher odds of allowing infection by new spiroplasmas than Mojave, 34.4274 times higher odds than Hydei2, 11.7593 times higher odds than Hydei1 and 63.04 higher odds than Atripe, with a *p*-value <0.0001 for each comparison.

4.3.8.7 Melano (D. melanogaster) is the "universal recipient of Spiroplasma"

The host of *Mel*, *D. melanogaster*, could be considered the "universal recipient of *Spiroplasma*" because it allows infection by all *Spiroplasma* strain tested. Although the sample size in this study was small, support to this conclusion comes from Williamson and Poulson (1979) report that *D. melanogaster* offers a favorable background for maintaining several strains of male-killing spiroplasmas (SM S17).

4.3.8.8 The Atripe fly recipient allows for infection success from fewer Spiroplasma strains and has lower odds of allowing infections from new parasite strains than other recipient strains used in the study

Atripe is the only fly recipient that only allowed infection from its own *Spiroplasma* strain and no others. Pairwise comparisons (contrasts) of Atripe allowing infection success versus all other

fly strains show that Atripe has lower odds of allowing infection than all other strains tested. Atripe has 0.2060 times lower odds of allowing infection by new spiroplasmas than Mojave ($p=0.0254$), 0.01586 times lower odds than Melano ($p<0.00010$), and 0.1865 times lower odds than Hydei1 ($p=0.0109$). Atripe has no significantly different odds of producing infection in new hosts than Hydei2 ($p=0.3783$).

4.4 Discussion

This discussion will address issues that will help to understand the potential host range of spiroplasmas.

4.4.1 Phylogenetic Relationship and Genetic Distances Between Spiroplasmas and Between Their Hosts

Drosophila-infecting spiroplasmas are grouped into 4 clades: poulsonii, citri, tenebrosa and ixodetis (Haselkorn et al., 2009). The above clades can be grouped into two extended clades or inclusive groups, according to the relatedness or divergence among said clades: the citri-poulsonii extended clade, or inclusive group, and the tenebrosa-ixodetis extended clade, or inclusive group (Haselkorn et al., 2009; Haselkorn, 2010a; Lo et al., 2013; Regassa & Gasparich, 2006). We studied five *Spiroplasma* strains from these clades (Fig. 4.1) and these clades can be divided into two extended clades or groups: 1) the ixodetis group, represented by the strain *Atr* (from the ixodetis clade; we did not have a representative strain from the tenebrosa clade) and 2) the citri-poulsonii group represented by the strains *Hyd2* and *Moj* (from the citri clade) and *Hyd1, Mel* (from the poulsonii clade). This grouping reflects the close relationship between *Mel, Hyd1, Hyd2* and *Moj*, while also reflecting the fact that *Atr* is the most divergent strain. The

genetic distances between these bacterial strains mirror the phylogenetic relationships, with *Mel*, *Hyd1*, *Hyd2* and *Moj* showing small distances among each other (the distances between these strains range from 0.0055 to 0.015), while *Atr* is the most distant one from the rest of the strains (these distances range from 0.1533 to 0.1571; Table 4.1). The set of fly strains used in this work belongs to two groups, the *Melanogaster* group (*D. melanogaster* or Melano and *D. Atripex* or Atripe) and the repleta group (*D. hydei* or Hydei and *D. mojavensis* or Mojave) (Bächli, 2017; Van Der Linde, Houle, Spicer, & Steppan, 2010). The genetic distance between the flies mirrors these phylogenetic relationships (Table 4.2, Fig. 4.1). The bacteria and fly trees are incongruent, suggesting multiple introductions of the bacteria into the *Drosophila* hosts (Haselkorn et al., 2009).

4.4.2 Proliferation of Spiroplasma in the Hemolymph of G0 (Infection Recipient) Flies and How the Bacteria Reach the Oocytes

In this work, injection of the bacteria into the new host flies enabled the study of infections that may not occur in nature by effectively bypassing the "encounter filter" (Combes, 1991) for these parasites and hosts. This would leave the components of the "compatibility filter" (environments where the parasite cannot survive and develop due to morphological, physiological, or immunological reasons; Combes, 1991) as the set of factors determining the success or failure of the new infections being investigated. This is reflected by the results of the statistical analysis, showing that the intrinsic characteristics of the bacteria, hosts, and the interaction of these characteristics, affect the transmission fidelity of the bacteria into the progeny of new hosts.

Drosophila-infecting spiroplasmas live predominantly in the hemolymph of their hosts (Anbutsu & Fukatsu, 2006) and are transmitted vertically, from mother to progeny. This work measured transmission rates at G1 post-infection flies, which are the progeny of G0 females (infection recipients). Only the progeny of infected G0 females (flies where *Spiroplasma* thrived, as assessed by PCR screening) were used to estimate the transmission rates reported in this work. This means that all *Spiroplasma* strains tested, regardless of their clade of origin, were able to proliferate in the hemolymph of all *Drosophila* hosts tested.

The fact that all bacteria were able to thrive in all G0 females indicates that all *Spiroplasma* strains found favorable environments in all new hosts and were able to avoid the immunological systems of the flies. Support for these results comes from reports on the ability of NSRO and MSRO *Spiroplasma* strains (poulsonii clade) to go undetected by the immunological system of *D. melanogaster*, and hence be able to proliferate inside this host (Hurst et al., 2003; Herren & Lemaitre, 2011). Other works have also shown that *Drosophila*-infecting spiroplasmas can thrive inside non-native flies (e.g., Williamson & Poulson, 1979; Haselkorn & Jaenike, 2015; Hutchence et al., 2011). All of this suggests that being able to thrive in non-native fly hosts may be a common characteristic of all *Spiroplasmas* that infect *Drosophila*.

Given that all strains tested were able to proliferate inside all hosts tested, the next step for the establishment of the infection by these maternally transmitted bacteria in new hosts should be the transmission of the bacteria to the progeny of the fly that received the foreign *Spiroplasma*. To do this, the bacteria have to cross from the hemolymph into the oocytes through the cooption of the yolk transport and yolk uptake machinery (Herren et al., 2013).

This means that the bacterial strains that were able to produce infection in new hosts were able to proliferate in the hemolymph of those new hosts, while also reaching the oocytes in those new hosts. A good example of bacteria that were able to reach the oocytes of all new hosts is the *Spiroplasma* strain *Atr* (native to *D. atripex*), which was able to be transmitted by all the host to which it was introduced. Although the infections were not followed for more than 1 generation, so it is unknown if these new infections would be successful in the long term, the transmission of *Atr* to G1 progenies makes this strain a good candidate to establish infections in all the new hosts tested in this work.

Meanwhile, the lack of transmission reported for some spiroplasmas by new hosts is not caused by the lack of proliferation of the bacteria in the new hosts upon artificial infection, but by their inability to reach the oocytes. Certainly, the results of this work show that bacteria thriving in new hosts upon artificial introduction are not always accompanied by transmission to progeny, resulting in failure to establish infection in new hosts. For example, although *Mel* thrives in all new hosts tested, it is not transmitted by *D. hydei* or *D. atripex*; similarly, *D. mojavensis* does not transmit *Hyd2*. This means that some component of the "compatibility filer" may still act against the successful establishment of an infection upon the introduction of the bacteria into a new host.

4.4.3 Complimenting and Refining the Estimates of Spiroplasma Potential Host Range

Our results indicate that all *Spiroplasma* strains tested have potential host ranges larger than those observed in nature, because they were able to infect other hosts than their natural ones (Fig. 4.1). Through artificial infections, our experiments rendered null the "encounter filter" because

we created associations of host-bacteria that may not occur in nature. This left the effect of the "compatibility filter" as the only filter influencing the host range of the bacteria in our study.

According to our results, *Hyd2* and *Atr* (Rohde indexes of 0.5448 and 0.5863, respectively) are the most generalist *Spiroplasma* strains in relation to the other strains studied here, and they show the largest potential host range. This suggests that these two strains are the most likely candidates to switch hosts or expand their host ranges if the opportunity arises (Poulin & Keeney, 2008).

The interpretation of Rohde index also suggests that *Moj* (0.7348) and *Mel* (0.7066) are the most specialist spiroplasmas that we studied, suggesting that these two strains are the least likely candidates to switch hosts or expand their host ranges if the opportunity arises (Poulin & Keeney, 2008).

Finally, the interpretation of Rohde index suggests that *Hyd1*, with an intermediate index value (0.6538) among the other strains tested, shows an intermediate potential to switch hosts or expand their host ranges if the opportunity arises.

4.4.3.1 Searching complimentary information from other studies and modifying the definition of generalist and specialist parasite

For an experimental study such as this, the host specificity indexes can be considered to be an incomplete exploratory approach to measure the potential host range of parasites. This is because the indexes will only give a picture of parasite ranges under the specific conditions in which the

experiments were run. For example, there may be some stochasticity in the *Spiroplasma* transmission rates across the G0 females that could be caused by how much spiroplasma was initially injected into these females, which in turn could influence how much bacteria reached the oocytes to be transmitted to the next generation (Haselkorn, pers. comm.). This could have caused the uneven distribution of the transmission rates for the donors and the recipients observed in Fig.4.2 and Fig. 4.3 (although other explanations for these patterns will be discussed later). To validate and refine the interpretation of the potential host ranges, information coming from other studies addressing the transmissibility of the parasites in question to other hosts must be added. We also used to this end the patterns of transmission rates observed in this study.

Several studies exist that can be used to this end, at least partly. Although not designed as cross-infection studies, three recent works reported the introduction of foreign spiroplasmas into new hosts in the *Drosophila-Spiroplasma* system (Nakayama et al., 2015; Hutchence et al., 2011; Haselkorn & Jaenike, 2015). Table 4.7 summarizes those studies and part of our results.

Table 4.7 Experiments used to compare this work.

Spiroplasma (natural host)	Recipient Fly	Initial Transmission Fidelity	End Transmission Fidelity	Reference
Atr (*D. atripex*)	*D. melanogaster*	0.69	NA	This work
Mel or MSRO (*D. melanogaster*)	*D. melanogaster*	0.87		
Hyd1 (*D. hydei*)	*D. melanogaster*	0.92		
Hyd2 (*D. hydei*)	*D. melanogaster*	0.1		
Moj (*D. mojavensis*)	*D. melanogaster*	0.09		
Mel or MSRO (*D. melanogaster*)	*D. hydei* strain 1	0	NA	This work
	D. hydei strain 2	0		
	D. mojavensis	0.77		
	D. atripex	0		
MSRO (*D. melanogaster*)	*D. melanogaster*	No Information	>0.98	Hutchence et al., (2011)
NSRO (*D. nebulosa*)	*D. melanogaster*		>0.95	
HY1 (*D. hydei*)	*D. melanogaster*		c. 0.80	
MOJ (*D. mojavensis*)	*D. melanogaster*		c. 0.25	
HY1 (*D. hydei*) - Replicate 1	*D. melanogaster*	0.76	0.89	Nakayama et al., (2015)
HY1 (*D. hydei*) - Replicate 2	*D. melanogaster*	0.89	0.45	
HY1 (*D. hydei*) - Replicate 3	*D. melanogaster*	0.78	0.85	
HY1 (*D. hydei*) - Replicate 4	*D. melanogaster*	0.78	0.5	
sHyd-p (*D. hydei*)	*D. neotestacea*	High	0.98	Haselkorn & Jaenike, (2015)
sHyd-c (*D. hydei*)	*D. neotestacea*	0.44	Lost Infection	
MSRO (*D. melanogaster*)	*D. neotestacea*	0.54	Lost Infection	
sAna (*D. ananassae*)	*D. neotestacea*	High	0.85	

Hutchence et al. (2011) introduced several *Spiroplasma* strains (MSRO, NSRO, HY1, MOJ) into *D. melanogaster* and reared them for a year before using them in experiments. They reported that the strains that naturally infect *D. mojavensis* (MOJ=*Moj* in our work), *D. hydei* (HY1=*Hyd1* in our work), *D. nebulosa* (NSRO) and *D. melanogaster* (MRSO=*Mel* in our work) achieved transmission rates of about 25%, 80%, 95%, and 98%, respectively (Table 4.7).

Nakayama et al. (2015) introduced HY1 (equivalent to our *Hyd1*) into *D. melanogaster* and the infection persisted until generation 25 (the last generation screened in their experiment) (Table 4.7).

Haselkorn and Jaenike (2015) introduced 4 strains of *Spiroplasma* into *D. neotestacea*. This work included three strains similar to the ones we used in our study: *sHyd-p* (equivalent to our *Hyd1* from the poulsonii clade, carried by *D. hydei*), *sHyd-c* (equivalent to our *Hyd2* from the citri clade, carried by *D. hydei*) and MSRO (equivalent to our *Mel* from the poulsonii clade, carried by *D. melanogaster*). In addition to those strains, Haselkorn and Jaenike (2015) used the strain *sAna* (ixodetis clade), carried by *D. ananassae*; this strain is closely related to our strain *Atr* from the ixodetis clade, carried by *D. atripex*. The control used in Haselkorn and Jaenike (2015) was *sNeo* (poulsonii clade) carried by *D. neotestacea* (Table 4.7). The strains *sHyd-p* (=*Hyd1*) and the control *sNeo* were transmitted by *D. neotestacea* females to their progenies until generation 10, with high transmission fidelity (0.98). *sAna* (closely related to our *Atr*) was transmitted at an 0.85 fidelity (Haselkorn & Jaenike, 2015). The other two strains, *sHyd-c* (=*Hyd2*) and MSRO (=*Mel*) showed low transmission rates (0.44 and 0.54, respectively, reported at generation 1) and were lost before generation 10 (Haselkorn & Jaenike, 2015; Table 4.7). Based on our results shown in Figs. 4.2 and Fig. 4.3 and the results of *Spiroplasma* artificial infections shown in Table 4.7, we can refine, to some extent, our first conclusions based on the Rohde Index.

4.4.3.2 Refining potential host range estimates based on information extracted from other studies as shown in Table 4.7 and the patterns observed in this study

One pattern identified is that *Spiroplasma* strains that show low transmission success at the first generations post-introduction of the infection into new hosts are lost later on. This is the case for sHyd-c (=*Hyd2*) and MSRO (=*Mel*) (Haselkorn & Jaenike, 215). Contrarily, the strains that start with high transmission rates, persist later on, such as the case of sHyd-p or HY1 (*Hyd1*) (Nagayama et al., 2015; Haselkorn & Jaenike, 2015).

Another pattern is that *D. melanogaster* allows for the persistence of the infection by its own *Spiroplasma* strain MSRO as well as NSRO, HY1 and MOJ (Hutchence et al., 2011; Nakayama et al.; 2015), confirming this work results that Melano allows for the infection of all new strain tested (Fig. 4.3).

Table 4.8 shows the refined estimated of the potential hoist ranges of the *Spiroplasma* strains used in this study. This was done by using information from Table 4.7 and the information presented in the section “Patterns of Infection Success Observed on the Current Experiment” from the results of this work. Table 4.8 also shows that the Rohde index did not reflect well the potential host range of the spiroplasmas. Therefore, the ordering from generalist to specialist strains had to be modified from Table 4.4.

Table 4.8 Refined estimation of the potential host range of *Spiroplasma* strains used in the present study. The first row shows the most generalist strain (*Atr*) and the last row the most specialist one (*Mel*), with specialization growing from the first to the last row. [1] Hutchence et al., (2011); [2] Nakayama et al., (2015); [3] Haselkorn and Jaenike (2015).

Clade (clade group)	*Spiroplasma* strain	Rohde Index (normalized)	Odds Infection at G1	Number of hosts infected (including own's)	Tendency to persists in new host	Likely to Expand Host Range?
ixodetis (ixodetis)	*Atr*	0.586	Highest	5	No information	Yes
poulsonii (citri-poulsonii)	*Hyd1*	0.64	medium	4	Yes[1,2,3]	Yes, in hosts of own clade group
citri (citri-poulsonii)	*Moj*	0.735	Lowest	3	In *D. melanogatser*[1]	No, but perhaps in hosts of own clade group?
citri (citri-poulsonii)	*Hyd2*	0.545	Lowest	3	No[3]	
poulsonii (citri-poulsonii)	*Mel*	0.707	Lower	2	No[3]	No

4.4.3.2.1 Atr is the most likely strain to opportunistically expand its range

This is based on: 1) *Atr* (ixodetis clade group) is able to produce infection in the G1 progeny of its own host and all the other 4 new hosts tested, which encompasses hosts parasitized by spiroplasmas outside the ixodetis group clade (i.e., it is able to infect the largest number of hosts); 2) *Atr* has higher odds of producing infection after artificial introduction into the hosts tested than most of the other *Spiroplasma* strains studied here (Table 4.8).

To the best of our knowledge, there are no published studies on the persistence of *Atr* (ixodetis group clade) in new hosts after an artificial infection. The lack of additional information on *Atr* persistence in new hosts allows for an alternative interpretation of this work's results. *Atr* is transmitted at 0.02 and 0.01 fidelity by *D. hydei* strains Hydei1 and Hydei2, respectively (Fig. 4.2), which may indicate that the infection will be eventually lost from said hosts. *Atr* may also be eventually lost by Mojave, which transmits it at a low 0.30 fidelity (Fig. 4.2). If this is correct, then, *Atr* may not be able to expand its range as stated above.

However, the relationships between the host flies also has to be taken into account when interpreting the results of host ranges, because the statistical analysis indicates that the intrinsic characteristics of the bacteria and the flies (as well as their interaction) affect the transmission fidelity of the bacteria. *Atr* produces a high transmission rate in *D. melanogaster*, from the melanogaster species group, which is closely related to its host, *D. atripex*. In this context, the low transmission fidelity observed in the *D. hydei* and *D. mojavensis,* from the repleta group, distantly related to *D. atripex*, may be interpreted as an indication that *Atr* needs a period of adaptation, albeit short (Haselkorn & Jaenike, 2015), to these new hosts.

Based on the information gathered in this section, *Atr* can be considered the most generalist strain and it may be able to expand its host range to the hosts of the distantly related citri-poulsonii clade-group bacterial strains, after an adaptation period in these new hosts.

4.4.3.2.2 Mel is the least likely strain to opportunistically expand its range

This is based on: 1) *Mel* is able to produce infection in the G1 progeny of only two hosts tested (its native host and *D. mojavensis*), which are parasitized by spiroplasmas from its own group clade, the citri-poulsonii (i.e., it is able to infect the smallest number of hosts); 2) *Mel* has lower odds of producing infection after artificial introduction into the hosts tested than most of the other *Spiroplasma* strains studied here (Table 4.8).

Table 4.7 shows that *D. melanogaster* transmits its native strain *Mel* at ≥0.80. This lends support to the results of this work showing that *Mel* is transmitted at high fidelity post-infection by its natural host. Although *Mel* is transmitted at high fidelity by *D. mojavensis* (>0.70), it could be eventually lost. This is because Table 4.7 shows that a non-natural host of *Mel*, *D. neotestacea*, transmits *Mel* at >0.50 fidelity in the initial post-infection generations, but the infection is eventually lost (see Table 4.7). This would leave *D. melanogaster* as the only host able to transmit *Mel* to its G1 progeny.

Based on the information gathered in this section, *Mel* can be considered the most specialist strain and the least likely to expand its range.

4.4.3.2.3 Moj and Hyd2 are the second least likely (after Mel) to expand their host range

This is based on: 1) *Moj* and *Hyd2* (citri-poulsonii group clade) are able to produce infection at low rates in the G1 progeny of three out of five hosts tested, which are the ones parasitized by spiroplasmas from their own clade-group, the citri-poulsonii; 2) *Moj* and *Hyd2* have the lowest odds of producing infection after artificial introduction into the hosts tested here (Table 4.8). *Moj* can persist in *D. melanogaster* at ~0.25 transmission fidelity after a year of being experimentally introduced (Hutchence et al., 2011 in Table 4.7 of our work). This work shows that *Moj* has 0.09 transmission fidelity post-infection in *D. melanogaster* (Table 4.7) and even lower than that in other hosts tested (e.g., 0.0 in Mojave, its natural host and Atripe), except in *D. hydei* strain Hydei1, where it shows a slightly higher transmission rate 0.014) (Fig. 4.2). The persistence of *Moj* in *D. melanogaster,* for about a year after its introduction in this host (Hutchence et al., 2011), may indicate a property of the host, not of the *Spiroplasma* strain. This is because *D. melanogaster* seems to allow the persistence of all *Spiroplasma* strains introduced into it, namely NSRO, the 2 strains of spiroplasmas that infect *D. hydei*, *Moj* and *Atr* (Table 4.7 and Fig. 4.2).

This work shows that *Hyd2* exhibits low transmission fidelity by all new hosts tested, including 0.0 transmission rate by Mojave and Atripe (Fig. 4.2). The low post-infection transmission fidelity shown by the 2 citri clade spiroplasmas suggests that these spiroplasmas may eventually be lost in most new hosts (as seen in *sHyd-c* (=*Hyd2*) in *D. neotestacea* in Table 4.7), except in Melano (*D. Melanogaster*), the "universal recipient of *Spiroplasma*" (see section "Patterns of Infection Success Rates or Transmission Fidelity Observed on the Current Experiment").

However, if low transmission rates are a natural characteristic of the citri clade spiroplasmas, then they may be able to expand their host ranges, after all.

Based on the information gathered in this section, *Moj* and *Hyd2* can be considered the second least specialized strains after *Mel.* As such, they may not be able to expand their host range due to their low transmission rates in all hosts tested. However, if low transmission rates are natural characteristic of these spiroplasmas, then they may expand their host range into hosts of spiroplasmas of the own clade-group, the citri-poulsonii.

4.4.3.2.4 Hyd1 is the second most likely to expand its range (after Atr)

This is based on: 1) *Hyd1* (citri-poulsonii group clade) is able to produce infections at low to high transmission rates in the G1 progeny of four out of five hosts tested (including its own), which are parasitized by spiroplasmas from its own group clade, the citri-poulsonii, but not by *Atr* (ixodetis clade group); 2) *Hyd1* has higher odds of producing infection after artificial introduction into the hosts tested than most of the other *Spiroplasma* strains studied here (Table 4.8).

Table 4.7 shows that *Hyd1* persists post-infection in *D. melanogaster* at transmission fidelities ranging from ~0.50 (Nakayama et al., 2015) to ≥0.80 (Hutchence et al., 2011) and at 0.98 transmission fidelity in *D. neotestacea* (Haselkorn & Jaenike, 2015). However, the results of this study indicate that although *Hyd1* is transmitted at 0.92 fidelity by *D. melanogaster* and 0.70 fidelity by its natural host *D. hydei*, it is transmitted poorly by Hydei2 and Mojave, the hosts of two closely related spiroplasmas (citri-poulsonii clade group), and not at all by *D. atripex*, the

host of the distantly related *Atr*, from the ixodetis clade group (Fig. 4.2). This suggests that although *Hyd1* may be able to expand its host range opportunistically, this ability is limited due to its low transmission fidelity by some new hosts for strains belonging to its citri-poulsonii extended group and to its inability to produce infection outside its citri-poulsonii extended group.

Based on the information gathered in this section, *Hyd1* (citri-poulsonii group clade) can be considered the second most generalist strain and the second likely to expand its host range, limited to hosts of spiroplasmas belonging to its own clade-group, the citri-poulsonii.

4.4.4 Factors Affecting Transmission Rates of the Bacteria and Hypothesizing the Mechanisms Determining the Potential Host Range

4.4.4.1 Bacteria and host strains and their interaction, but not phylogeny, affect the transmission rate (or transmission success rate) of the infection

Under the experimental conditions of this work, the infection rates of *Spiroplasma* at generation 1 post-cross-infection in a new host depend on the fly host and the *Spiroplasma* strain as well as the interaction between these two factors. This means that infection success rates vary depending on the combination of host and the bacteria intrinsic characteristics. Other factors examined in this work, such as genetic distances between bacterial strains and between fly strains as well as higher phylogenetic classes (*Spiroplasma* clade and fly species group), had no effect on the transmission rate of the infection. One possible explanation for this is that we have a small sample size (number of strains used in this work) and most of the genetic distances are small, especially with regard to the *Spiroplasma* strains. A cross-infection experiment with a larger

number of *Spiroplasma* strains, including representatives from all clades described in Haselkorn et al. (2009), would help to clarify this issue.

No effect of phylogeny in transmission rates of the bacteria in new hosts contradicts previous works where researchers found that parasite transmission rates declined with the phylogenetic distance of the parasite strains (e.g., Hutchence et al., 2011) or the phylogenetic proximity of the experimentally inoculated host and the pathogen's host-of-origin (e.g., Gilbert & Webb, 2007) or both (e.g., Perlman & Jaenike, 2003; De Vienne et al., 2009).

For this discussion we will take a closer look at research concerning the *Drosophila-Spiroplasma* system because this is the same one we worked on for our experiments. Hutchence et al. (2011) introduced several diversely related *Spiroplasma* strains into *D. melanogaster*. The most divergent bacterial strain, the one infecting *D. mojavensis* (which belongs to the citri clade), achieved a transmission rate of about 25%. The other strains (all of them belonging to the poulsonii clade) achieved vertical transmission efficiencies of 80% (*Hyd 1*), >95% (NSRO), and >98% (the native MSRO). Although Hutchence et al. (2011) did not report actual genetic distances, the classification of genetic relatedness of the spiroplasmas they used for their experiments accurately represents the genetic distances between these flies. Our results with *D. melanogaster* as a recipient (Melano strain) roughly matches the Hutchence et al. (2011) trend for the transmission of *Spiroplasma* strains *Moj* (9% transmission efficiency), *Hyd1* (92%) and the self-inoculation of *Mel* (MSRO) (87%). Our work includes a more distantly related strain, *Atr*, and its transmission rate is 69%, which does not match the trend of lower transmission rates as genetic distance increases as reported by Hutchence et al. (2011). The results of this work for

other *Spiroplasma* strains introduced into other fly hosts further show that genetic distance of the bacteria does not correlate with infection success rates in general. For example, *Moj* is not transmitted by its native host, Mojave (*D. mojavensis*); *Mel* is transmitted by Mojave, but not by the host of its closest relative, *D. hydei* (strain 1 named Hydei1 in our work, which is the host of *Hyd1*) (Fig. 4.1).

From the perspective of the fly strains that received the infection, the genetic distance between the host of origin of the *Spiroplasma* strain and the host of the newly introduced strain did not predict transmission rate very well. For example, *D. mojavensis* (Mojave) transmits strains *Mel* (77%) and *Atr* (30%) better than strains *Hyd1* (10%) and *Hyd2* (0%), both hosted by *D. hydei*, which is more closely related to *D. mojavensis* than *D. melanogaster* or *D atripex* (Table 4.2, Fig. 4.1). As mentioned previously, Mojave does not transmit its own native strain *Moj* (0%). Similar results of transmission success of spiroplasmas into new, diversely related hosts were reported by Williamson and Poulson (1979). For example, SM S17, modified from Williamson and Poulson (1979), shows that the *Spiroplasma* carried by *D. equinoxialis* shows incomplete transmission (U) when introduced into *D. melanogaster* (a co-subgeneric *Sophophora*), but its transmission is excellent (+++) when introduced into *D. robusta* (a species from the genus *Drosophila*). Contrarily, *D. willistoni*'s WSRO and *D. nebulosa*'s NSRO have excellent transmission in all species tested regardless of their relatedness (excepting WSRO in *D. robusta*).

The lack of a correlation between the phylogeny of the hosts and the transmission fidelity of *Spiroplasma* has been also reported by Tinsley and Majerus (2007), who found no association

between genetic distance and vertical transmission of the *Spiroplasma* from *Adalia bipunctata* into several other coccinellid beetles.

4.4.4.2 Multiple mechanisms may affect the transmission rates of the bacteria in the hosts

It is expected that multiple mechanisms be at play affecting the transmission rates of the bacteria to the progeny of the new hosts tested in our work. This is because the bacterial strains and the fly strains used in our experiment include highly divergent as well as closely related ones, so the peculiar characteristics of each host and parasite affecting host ranges have to reflect those diverse relationships. We will address each bacterial strain below.

4.4.4.2.1 The case of Mel from the poulsonii clade introduced into D. atripex

It is worth pointing out that *D. atripex* (from the melanogaster fly group) does not transmit to its progeny any of the foreign *Spiroplasma* strains tested in this study. Therefore, not transmitting any new strain (*Mel* included) seems to be part of the general pattern observed for this host (Fig. 4.3). The *Spiroplasma* that *D. atripex* carries (*Atr*, ixodetis clade) are the most divergent ones from the rest of the strains studied here (Haselkorn et al., 2009; Haselkorn, 2010a; Table 4.1). It is likely that the cause of the lack of transmission of *Mel* by Atripe is that while the bacteria can thrive in the hemolymph of this fly, there may be differences in the yolk transport and yolk uptake machinery of *D. melanogaster* and *D. atripex*, and *Mel* cannot use *D. atripex* machinery to reach the oocytes.

4.4.4.2.2 The case of Mel from the poulsonii clade introduced into repleta group flies D. hydei and D. mojavensis

The repleta group flies that were tested as new hosts for *Mel* show mixed results for the transmission rate to their progenies (G1). *D. mojavensis* (strain Mojave from the Repleta fly group that carries *Moj*, a *Spiroplasma* from the Citri clade) transmitted the infection at 0.77 fidelity. The two *D. hydei* strains, Hydei1 and Hydei2, that carry the *Spiroplasma* strains *Hyd1* (poulsonii clade) and *Hyd2* (Citri clade), respectively, failed to transmit *Mel* at all. It is important to reiterate here, as indicated earlier, that *Mel* was able to proliferate in the hemolymph of the G0 females and can cross from the hemolymph into the oocytes by cooption of the yolk transport and yolk uptake machinery (Herren et al., 2013) of these hosts, as suggested by the 77% transmission rate of *Mel* by Mojave (Fig. 4.2). We propose that the lack of transmission success of *Mel* by *D. hydei* may be explained by the fact that *Mel* is quickly eliminated from the progeny because it reduces the fitness of the hosts through male killing. SM S18 shows that *Mel* significantly reduces the sex ratio of the G1 progeny on all repleta flies (but not in *D. atripex*, a melanogaster group fly). The no detection of *Spiroplasma* through screening, but the observation of the male killing phenotype in the G1 progeny of repleta flies allows us to hypothesize that all infected progeny died off at the larval stage and only uninfected flies emerged. This early elimination of *Mel* from the progeny may be due to the fact that *D. hydei* males are expensive to produce because of their gigantic sperm cells and their related delayed maturity with respect to females (Pitnick, Markow, & Spicer, 1995).

The successful transmission of Mel to the G1 Mojave progeny could be considered part of the norm, because *Mel* is able to be transmitted to the G1 of *D. neotestacea* (Haselkorn & Jaenike,

2015) (Table 4.7). Therefore, the no-transmission of *Mel* by *D. hydei* might be considered an exception brought about by the reason explained earlier. At this moment we cannot propose a mechanism for the rapid elimination of *Mel* from *D. hydei*.

4.4.4.2.3 The case for horizontal transmission impaired in the citri clade spiroplasmas and their repleta group hosts

The *Spiroplasma* strains from the citri clade show a striking pattern of low infection success rates across all hosts tested at G1 (Fig. 4.2). Furthermore, the *D. mojavensis'* natural *Spiroplasma* strain (*Moj*) used in this work is not transmitted by its own native host, the fly strain Mojave. The other citri strain (*Hyd2*) is also not transmitted by Mojave. As clarified earlier, the poor transmission rate of the citri clade spiroplasmas does not seem to be due to the lack of proliferation of the bacteria inside the maternal generation (G0) or their inability to reach the oocytes (the various low transmission rates of citri clade spiroplasmas at G1 show that these spiroplasmas are able to use the yolk transport and yolk uptake machinery of some hosts from the melanogaster and repleta fly groups).

To explain the low transmission rate of the citri clade spiroplasmas in our experiments, we propose that citri clade spiroplasmas can produce pathogenicity in their native or new hosts, and this pathogenicity is a common characteristic in the interaction of citri spiroplasmas and their repleta (and non-repleta) hosts; in turn, this pathogenicity can produce the low citri *Spiroplasma* transmission rate from repleta (and non-repleta) host mothers to their progenies.

We have not studied the pathogenicity of all spiroplasmas in this work, but this characteristic can be hypothesized based on published works by other authors. The phylogenetic tree of *Spiroplasma* shows that both the poulsonii and citri clade spiroplasmas are closely related to pathogenic strains within their clades (Haselkorn et al., 2009), which suggests that these bacteria have the potential to harm their hosts. The two poulsonii clade spiroplasmas tested here, *Hyd1* and *Mel* (or MSRO), have been shown to affect longevity and fecundity in their natural hosts, *D. hydei* (Chapter 2 of this work) and *D. melanogaster* (Herren et al., 2014), respectively. Additionally, *Mel* kills the male progeny in *D. melanogaster*. The *Spiroplasma* strains from the citri clade used in this work, *Hyd2* and *Moj*, may also be pathogenic to their hosts (the Repleta group flies *D. hydei* and *D. mojavensis*, respectively) although the mechanisms involved may be different from the poulsonii clade. For example, Herren and Lemaitre (2011) reported that *Spiroplasma citri*, a plant pathogen after which the citri clade is named, is pathogenic to the Melanogaster group fly *D. melanogaster*. *S. citri* grows to high density in the hemolymph of *D. melanogaster*, and this is related to its pathogenicity in this fly (Herren and Lemaitre, 2011). That the pathogenicity of *S. citri* is related to its density in the host was also reported in the plant host *Citrus sinensis* (Mello, Yokomi, Melcher, & Fletcher, 2010).

This allows us to propose that pathogenicity to the hosts (or the potential for it) due to high densities of the bacteria could be a common characteristic of the citri clade spiroplasmas. This implies that a minimally harmful, continuous association of the citri clade spiroplasmas with their hosts requires the bacteria to exist at low densities in the latter, to avoid overt pathogenicity. We propose that this is the case for the relationship between citri clade spiroplasmas and their natural host, the *Drosphila* repleta flies. Our study was not designed to explore the density of the

bacteria inside the *Drosophila* hosts or the pathogenic effects of the infection, but Haselkorn et al. (2013) reported that citri clade spiroplasmas (*Hyd2* and *Moj* in our study) show lower titers inside their natural hosts (*D. hydei* strain 2, hydei2 in our study, and *D. mojavensis*) when compared to poulsonii clade spiroplasmas infecting their natural hosts, *D. hydei* and *D. melanogaster*. Furthermore, from the two poulsonii clade spiroplasmas studied here, Haselkorn et al. (2013) reported that *Hyd1*, which infects *D. hydei* strain 1 (Hydei1) shows lower titers than the male-killing *Mel* in *D. melanogaster*. We argue that this lends support to our hypothesis that low-density infection is required for the long-term association of citri spiroplasmas with their repleta *Drosophila* hosts. We propose that the low density of the citri clade spiroplasmas in their natural hosts is related to the low the transmission fidelity of the bacteria. This is feasible because transmission fidelity seems to depend on endosymbiont density (Jaenike, 2009). At this moment, we do not know what mechanism or mechanisms are involved in this host-bacteria interplay. As stated earlier, this work was not designed to investigate the density or pathogenicity of *Hyd1*, *Hyd2*, *Moj*, and *Mel* to their hosts, or the mechanisms involved in the low (or high) transmission fidelity of spiroplasmas. Further research may address this issue.

4.4.4.2.4 Moj from the citri clade is not transmitted by its native host D. mojavensis

Mojave failed to transmit *Moj*, its native *Spiroplasma*, to its progeny. This peculiar result may be due to our experimental conditions. We say this because Watts et al. (2009) reported that the prevalence of *D. mojavensis* in natural populations varies from 15% to 85%, which requires a transmission rate above 0%. To explain these results let's recall that that the statistical interaction between *Spiroplasma* and host affects the transmission rate from mother to progeny in experimental hosts. Based on this, we propose that the lack of transmission reported here is due

to the pathogenicity of *Moj* and the susceptibility of Mojave to *Spiroplasma* and non-*Spiroplasma* infections. It is plausible that the wound caused by the micro injector needle could have become a way for the entry of pathogens that may have synergistically acted with opportunistic pathogens to produce disease in these flies. Hurst et al. (2003) reported a reduction of titers of NSRO infecting *D. melanogaster* after the flies underwent septic shock. If a similar behavior of the immune system occurred in our experimental *D. mojavensis,* then the density of *Moj* may have been greatly reduced inside Mojave. If transmission fidelity in Mojave depends on *Spiroplasma* density, as seems to be the case for some endosymbionts (Jaenike, 2009), then the low *Moj* densities postulated above may explain the non-transmission of this bacteria in our study. All of this led us to hypothesize that *Moj* (as well as *Hyd2*, the other citri clade strain) are not well transmitted to G1 after horizontal transfer (at least under our experimental conditions) due in part to their characteristic of producing pathogenicity in the host, and in part to the characteristics of their repleta hosts, who are susceptible to opportunistic infections. This could be tested by culturing these flies under strict aseptic conditions and performing horizontal cross-infections under those same conditions. A parallel experiment should be conducted under the standard conditions used in the experiment we are reporting.

4.4.4.2.5 D. atripex is not a good transmitter of new Spiroplasma strains, and their bacteria Atr is transmitted by all fly hosts tested

D. atripex (melanogaster fly group) strain Atripe does not transmit to its progeny any of the *Spiroplasma* strains tested in this study, but their natural *Spiroplasma* (*Atr*, ixodetis clade) is transmitted by all new hosts tested. A closer look at the phylogenetic relationships of *Atr* to the other strains may help us to explain this pattern. As presented earlier, *Atr* is the most divergent

strain from the rest of the strains used here. A phylogenetic tree based on the 16S rDNA indicates that the ixodetis clade is separated from the citri-poulsonii group by two clades of bacteria that do not infect *Drosophila* (Regassa & Gasparich, 2006; Lo et al., 2013). The ixodetis clade diverges at the base of the *Spiroplasma* evolutionary tree (Gasparich, 2002; Regassa & Gasparich, 2006). The species type for the ixodetis clade, *Spiroplasma ixodetis* (a tick-infecting bacteria) is serologically distinct from other *Spiroplasma* species and phylogenetically unique (Williamson et al., 2015). *S. ixodetis* was not compared to *Atr*, the ixodetis clade strain used in our study as seen in "Fig. 3" in Williamson et al. (2015). *Atr* has not been serologically studied, but it is reasonable to assume that it is similar to *S. ixodetis* and is as distinct from the other spiroplasmas as *S. ixodetis* is, judging by the genetic distance between the latter and *Atr* (0.39%).

Based on the above, it is reasonable to propose that *Atr* possesses a "generalized" mechanism to utilize the yolk transport and uptake machinery of the host that allows the bacteria reach the oocytes. This enables the bacteria to reach the oocytes of all hosts tested in this work, regardless of their phylogenetic relationships. A hypothesis to explain the non-transmission of the citri-poulsonii spiroplasmas by *D. atripex* can be proposed along similar lines. The citri-poulsonii spiroplasmas possess a "specialized" mechanism that allows them to reach the oocyte of the hosts that carry other closely related citri-poulsonii spiroplasmas; because of that specialization, the citri-poulsonii spiroplasmas cannot reach the oocytes of the ixodetis clade *Spiroplasma* host.

4.5 Conclusions

The ability of 5 *Drosophila*-infecting *Spiroplasma* strains to expand their host range was evaluated in this work. The potential host ranges were evaluated by experimental cross-

infections. These cross-infections were analyzed by using the Rohde index of host range. Following this analysis, we used our results and published studies to validate and refine on the results from the first procedure. The results show that the host range of most of the spiroplasmas studied here are larger than in nature.

Our refined results indicate that the *Spiroplasma* strain *Mel* or MSRO that naturally infects *D. melanogaster* is the most specialized bacteria because it can produce infection in the progeny of the least number of hosts studied (members of the citri-poulsonii extended group only), and we conclude that this strain is the least likely to expand its host range. The *Spiroplasma* strain *Atr* that naturally infects *D. atripex* is the most generalist bacteria (it can infect all the hosts studied, which include strains from the citri-poulsonii and ixodetes extended groups), and it can be concluded that this strain is the most likely to expand its host range. The strains *Hyd2* and *Moj* are less specialized than *Mel* but produce low or zero infection rates in the progeny of all hosts tested. New host that do not transmit *Hyd2* and *Moj* include the host of the distantly related *Atr* (the fly *D. atripex* strain Atripe) and Mojave, the host of *Moj*. This suggests that these two stains will have a limited success in expanding their host range, if at all. The strain *Hyd1* comes second to *Atr* regarding the number of host it can infect, but its ability to infect new hosts is restricted to hosts of closely related *Spiroplasma* strains (bacterial stains from the citri-poulsonii extended group), indicating that it may be able to expand its host range if the opportunity arises, but to a lesser degree than *Atr*.

The results of the host-bacteria cross-infections (transmission fidelity or success rate in new hosts' progenies) were analyzed to elucidate the factors affecting the potential host range of the

bacterial strains. A logistic regression analysis indicates that the infection success rate depends on the *Spiroplasma* strain, host strain and the interaction between these 2 factors (i.e., the interplay between the intrinsic characteristics of the bacteria and flies) but does not depend on phylogeny (genetic distance between *Spiroplasma* strains or genetic distance between fly hosts). Therefore, it can be said that with the samples studied, the potential host range of these bacteria are not predictable based on phylogeny. However, these ranges can be understood based on the interplay of the intrinsic characteristic of the bacterial strains and hosts species tested, as well as the relationship between their clades.

Based on the results of the statistical analysis, we hypothesize several mechanisms to explain the success and failure of the infection in new hosts, namely, negative fitness effects of the infection on new hosts not adapted to cope with such effects, pathogenicity of *Spiroplasma*, susceptibility of some hosts to infection by opportunistic bacteria, and co-adaptation of bacteria and host.

We also report that all *Spiroplasma* strains tested were able to proliferate inside all hosts tested, which is in line with other reports showing spiroplasmas that infect *Drosophila* can proliferate in hosts other that their natural ones (Hutchence et al., 2011; Nakayama et al., 2015; Haselkorn & Jaenike, 2015). This suggests that the reported capacity of two male-killing and a non-male-killing *Spiroplasma* strains to elude the immunological system of *D. melanogaster* may be a common characteristic of all *Drosophila*-infecting spiroplasmas (Hurst et al., 2003; Anbutsu & Fukatsu, 2010). The ability of all *Spiroplasma* strains tested to proliferate in new hosts, and the failure of several strains to produce infection in the progeny on some of those new hosts

indicates that the “compatibility filter” can still hinder the establishment of a new infection even when the “encounter filter” is overcome.

Finally, because the potential host range of *Spiroplasma* depends on the interplay of intrinsic characteristics of the bacteria and the hosts, these factors have to be taken into account, in addition to environmental factors, when evaluating the expansion of the host range of these bacteria.

It can be argued that the potential host range, as studied here, is of limited value because it was obtained under ideal conditions from host-parasite pairs that have distinct geographic origins in diverse environments. This argument can be refuted by indicating that the ideal nature of this concept is already acknowledged in its definition (see the “introduction” section), but this ideal nature does not prevent this concept to have practical applications. To cite a few examples, studies of potential host ranges have been used to gain insights on what factors have to be further investigated to understand better why parasites do not use some available hosts in the wild or why they are specialists (Shimizu-kaya et al., 2013; Morehead & Feener, 2000), or to evaluate the potential of parasites as biological controls (Hiroyoshi et al., 2017; Wang & Messing, 2006). Furthermore, the current work has produced a practical recommendation, which is that the intrinsic characteristics of the bacteria and the hosts and the interplay of those characteristics have to be taken into account, in addition to environmental factors, when evaluating (or trying to predict) the expansion of the host range of *Drosophila*-infecting *Spiroplasma*.

5. CONCLUSIONS

This work addressed two issues related to the evolution of host-parasite relationships: long-term persistence of maternally heritable endosymbionts and the specificity of host-parasite associations, i.e. host range.

The long-term persistence of maternally heritable endosymbionts can be achieved by a single mechanism or a combination of several others. Perfect vertical transmission would, in theory, guarantee the persistence of any symbiont for many generations to come. Endosymbionts without perfect transmission will be lost from the host unless there is horizontal transfer or an increase in fitness of the host due to the infection (Werren & O'Neill, 1997).

However, long-term associations between bacterial endosymbionts and fly hosts have been reported that appear to not meet these criteria (Mateos et al., 2006; Haselkorn et al., 2009).

One such association is between the bacterial parasite *Spiroplasma* strain *Hyd1*and its natural host, the fruit fly *Drosophila hydei*. *Spiroplasma* strain *Hyd1* reportedly does not produce direct fitness benefits to its host (Kageyama et al., 2006; Osaka et al., 2008, referenced as unpublished data) and does not show perfect or nearly perfect transmission (Osaka et al., 2013a). Yet, this parasite-host association has persisted long-term in the wild (e.g., Ota et al., 1979; Osaka et al., 2013a) and in laboratory populations (Watts et al., 2009; Osaka et al., 2013a). These observations suggest that the infection has fitness benefits that have yet to be identified.

The experiments described in Chapters 1 and 2 looked for possible effects of *Spiroplasma* infection on host characters that could account for long-term persistence of the host-parasite association. Chapter 1 describes preliminary experiments to examine characters of the host that if modified by the infection, would give infected hosts a fitness advantage and allow for long-term persistence of the infection. Measurements were made on populations of infected and uninfected female flies. The characters studied were sex ratio, larval-to-emerged-fly survival and fecundity. None of the characters measured appeared to be affected by infection status, except for longevity, which was found to be reduced in infected females. Longevity reduction by itself would have a negative effect on fitness of the infected host. However, in combination with increased fecundity, longevity reduction is predicted to promote long-term persistence (Agnew et al., 2000).

An additional goal of the preliminary studies was to optimize the experimental protocols. During the course of the initial experiments, sources of variation were identified that affected some of the characters that were being measured. For this reason, the effects of infection by *Spiroplasma* strain *Hyd1* were re-examined in a more controlled way. These experiments are described in Chapter 2. The studies were expanded to also look at age at maturity and preferential mating to infected females by all males.

The results presented in chapter 2 confirmed that infected females have a shorter life span than uninfected ones and also showed that the lifetime fecundity of infected and uninfected females is the same. However, unlike in chapter 1, infected females were found to have higher early fecundity than uninfected females. This is in line with the theoretical prediction of the behavior

of these two life-history traits (Agnew et al., 2000). I used simple simulations to show that the higher early fecundity of infected females promotes the persistence of the infection in the long-term. A similar effect on longevity and early fecundity of MSRO-infected *D. melanogaster* (a natural *Spiroplasma-Drosophila* pair) was reported (Herren et al., 2014), suggesting that these effects on fitness may be common in the *Spiroplasma-Drosphila* system. These findings also suggest that the long-term persistence of the *Spiroplasma* infection in *Drosophila* is a complex problem, because there may be other mechanisms implicated in this phenomenon, such as context-dependent fitness benefits to the host (Xie et al. 2011; Xie et al., 2015; Jaenike et al., 2010) and horizontal transfer (Jaenike et al., 2007). The findings of this work as well as those of Xie et al. (2010) and Jaenike et al. (2007) need to be evaluated in the field to assess their real contribution to the persistence of the infection in wild populations. For laboratory populations, where flies are sheltered from natural enemies and the mites that are reported to be responsible for horizontal transfer of *Spiroplasma* infection (Jaenike et al., 2007), the advantage in early reproductive output of infected females reported here seem to explain the long-term persistence of the infection.

Chapter 3 explored the potential host range and the factors affecting those potential host ranges of distinctly related *Spiroplasma* strains on a group of their *Drosophila* natural hosts that are also distinctly related, showing an incongruent host-parasite tree (Haselkorn et al., 2009). The ability of a parasite to expand its range to other hosts beyond their own (i.e., potential host range) can help determine how likely a parasite is to persist, following host extinction (Dunn et al. 2009; Moir et al., 2010) or to become established following its introduction into new areas (Taraschewski, 2006). The evaluation of the potential host range of parasites and the factors

affecting those ranges are necessary in a world undergoing accelerated changes. This is because changes in distribution of a species in response to environmental changes may depend not only on the latter but also on the biological characteristics of the species implicated (Chen et al., 2011).

To examine the potential host range of *Spiroplasma* and the factors affecting it, I used 5 natural *Spiroplasma-Drosophila* pairs: *Mel-D. melanogaster*, *Hyd1-D. hydei*, *Hyd2-D. hydei*, *Moj-D. mojavensis*, and *Atr-D. atripex*. Those 5 *Spiroplasma* strains were cross-infected into each other's uninfected 5 natural hosts, generating 25 sets of different infections (including a self-infection for each host-parasite pair) with sets of 90 repeats per set and a total of 2250 observations. I measured the infection rate (= transmission rate) at generation 1 post-cross-infection to quantify the potential host rage of the 5 spiroplasmas. I then evaluated several factors for their effect on the host ranges of spiroplasmas, namely, parasite and host phylogenies, intrinsic host and parasite characteristics, and the interaction of all of these factors.

The potential host ranges of *Spiroplasma* strains studied are varied. Extreme cases are represented by *Mel*, which is not able to produce infection in the progeny of any of the new hosts tested but its own, a fact that qualifies it as a specialist, which strongly suggests that this bacterial strain will not be able to expand its host range even if the opportunity arises. The opposite happens with *Atr*, which can infect all hosts tested, thus making it a generalist, suggesting that this strain will be able to expand its host range if the opportunity arises. The strains *Hyd2* and *Moj* are less specialized than *Mel* but produce low or zero infection rates in the progeny of new hosts, which suggest that they will have limited success in expanding their host range, if at all.

The strain *Hyd1* comes second to *Atr* regarding the number of hosts it can infect, but its ability to infect new hosts is restricted to hosts of closely related *Spiroplasma* strains, indicating that it may be able to expand its host range if the opportunity arises, but to a lesser degree than *Atr*.

Only the host and parasite intrinsic characteristics and their interaction were found to affect the infection success rate of *Spiroplasma* when introduced to new and old hosts.

Finally, the fact that *Spiroplasma* strain *Hyd1* confers fitness advantage to its host *D. hydei* suggests that other non-male-killing *Spiroplasma-Drosophila* pairs may show a similar relationship, and this may lead to the establishment of a mutualistic host-bacteria relationship in the long run. Other final conclusions are that the potential host range of *Drosophila*-associated *Spiroplasma* are not predictable based simply on the phylogenetic relatedness of *Spiroplasma* strains and *Drosophila* hosts. Host range appears to be dependent on the interplay of biological characteristics of the host and the parasite. Moreover, any evaluation of the potential expansion of the host ranges of these bacteria must take into account both biological and environmental factors.

REFERENCES

Agnew. P., Koella J. C., Michalakis, Y. (2000). Host life history responses to parasitism. *Microbes and Infection*, *2*(8), 891−896.

Adamo, S. A. (1999). Evidence for adaptive changes in egg laying in crickets exposed to bacteria and parasites. *Animal Behavior*, *57*(1), 117-124.

Aleman-Meza, B., Jung, S., & Zhong, W. (2015). An automated system for quantitative analysis of Drosophila larval locomotion. *BMC Developmental Biology* 15: 11. https://doi.org/10.1186/s12861-015-0062-0.

Allison P. D. (2012). *Logistic regression using the SAS system: Theory and application* (2nd ed.). Cary, NC: SAS Institute.

Anbutsu, H., & Fukatsu, T. (2006). Tissue-specific infection dynamics of male-killing and nonmale-killing spiroplasmas in *Drosophila melanogaster*. *FEMS Microbiology Ecology*, *57*(1): 40-46. https://doi.org/10.1111/j.1574-6941.2006.00087.x.

Anbutsu, H., Goto, S., & Fukatsu, T. (2008). High and Low Temperatures Differently Affect Infection Density and Vertical Transmission of Male-Killing *Spiroplasma* Symbionts in *Drosophila* Hosts. *Applied and environmental microbiology*, *74*(19), 6053-6059.

Anbutsu, H., & Fukatsu, T. (2010). Evasion, suppression and tolerance of *Drosophila* innate immunity by a male-killing *Spiroplasma* endosymbiont. *Insect Molecular Biology*, 19, 481-488. doi:10.1111/j.1365-2583.2010.01008.x.

Aquilino, A., Masiá, M., López, P., Galiana, A. J., Tovar, J., Andrés, M., & Gutiérrez, F. (2015). First Human Systemic Infection Caused by *Spiroplasma*. *Journal of Clinical Microbiology*, *53*(2), 719–721. http://doi.org/10.1128/JCM.02841-14.

Bächli, G. (1999–2017). TaxoDros: The Database on Taxonomy of Drosophilidae. Available at http://taxodros.unizh.ch.

Beal, D. J. (2007). SA05: Information Criteria Methods in SAS for Multiple Linear Regression Models. SESUG 2007: The Proceedings of the South East SAS Users Group, Hilton Head, SC, 2007. Retrieved from http://analytics.ncsu.edu/sesug/2007/SA05.pdf.

Bonacum, J., DeSalle, R., O'Grady, P., Olivera, D., Wintermute, J., & Zilversmit, M. (2001). New nuclear and mitochondrial primers for systematics and comparative genomics in Drosophilidae. *Drosophila Information Service*, 84, 201-204.

Bonduriansky, R. (2001). The evolution of male mate choice in insects: a synthesis of ideas and evidence. *Biological reviews of the Cambridge Philosophical Society*, *76*(3), 305-339.

Carvalho, G. G., & Da Cruz, M. P. (1962). Transfer of the "Sex-Ratio" Factor in Drosophila willistoni by ingestion. *Science, 138*(3536), 51-51. Retrieved from http://www.jstor.org/stable/1709846.

Carruthers, E., Lewis, K., McCue, T., Westley, P. (2008). Generalized linear models: Model selection, diagnostics, and overdispersion. Memorial University of Newfoundland. Retrieved from http://www.mun.ca/biology/dschneider/b7932/B7932Final4Mar2008.pdf.

Cazemajor, M., Landre, C., & Montchamp-Moreau, C. (1997). The Sex-Ratio Trait in *Drosophila Simulans*: Genetic Analysis of Distortion and Suppression. *Genetics*, *147*(2), 635–642.

Chadwick, W., & Little, T. J. (2005). A parasite-mediated life-history shift in *Daphnia magna*. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1562), 505–509. http://doi.org/10.1098/rspb.2004.2959.

Chen, I. C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid Range Shifts of Species Associated with High Levels of Climate Warming, *Science*, 19, 1024-1026.

Combes, C. (2001). *Parasitism: The Ecology and Evolution of Intimate Interactions*. (I. de Buron & V. A. Connors, Trans.). Chicago, ILL: University of Chicago Press. (Original work published 1995).

Committee on Gene Drive Research in Non-Human Organisms: Recommendations for Responsible Conduct; Board on Life Sciences; Division on Earth and Life Studies; National Academies of Sciences, Engineering, and Medicine. Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values. Washington (DC): National Academies Press (US); 2016 Jul 28. 2, The State of Knowledge of the Molecular Biology, Population Genetics, and Ecology of Gene-Drive Modified Organisms. Available from: https://www.ncbi.nlm.nih.gov/books/NBK379272/.

Counce, S. J., & Poulson, D. F. (1962). Developmental Effects of the Sex-ratio Agent in Embryos of *Drosophila willistoni*. *Journal of Experimental Zoology*, 151, 17-31. doi:10.1002/jez.1401510103.

De Vienne, D. M., Hood, M. E., & Giraud, T. (2009). Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology*, 22, 2532-2541. doi:10.1111/j.1420-9101.2009.01878.x.

Decaestecker, E., Vergote, A., Ebert, D., & Meester, L. D. (2003), Evidence for Strong Host Clone-Parasite Species Interactions in the Daphnia Microparasite System. *Evolution, 57*(4), 784-792. doi:10.1111/j.0014-3820.2003.tb00290.x.

Dickey, D. A. (2007). PROC MIXED: Underlying Ideas with Examples. Retrieved from URL: http://analytics.ncsu.edu/sesug/2007/SA07.pdf.

Dobson, A.P., & Carper, R. J. (1992). Global warming and potential changes in host-parasite and disease-vector relationships. In R. Peters & T. Lovejoy (Eds.). *Consequences of Global Warming for Biodiversity* (*pp*. 201-207). New Haven, Connecticut: Yale University Press.

Dong, P., Wang, J. -J., & Zhao, Z. -M. (2006). Infection by *Wolbachia* bacteria and its influence on the reproduction of the stored-product psocid, *Liposcelis tricolor*. *Journal of Insect Science, 6*, 24. http://doi.org/10.1673/2006_06_24.1.

Douglas, A. E. (1994). *Symbiotic interactions*. Oxon, GB: Oxford University Press.

Dunn, R. R., Harris, N. C., Colwell, R. K., Koh, L. P., & Sodhi, N. S. (2009). The sixth mass coextinction: are most endangered species parasites and mutualists? *Proceedings of the. Royal Society of London, B: Biological Sciences*, 276, 3037–3045.

Ebert, D. (1994). Virulence and Local Adaptation of a Horizontally Transmitted Parasite. *Science, 265(*5175), 1084-1086. https://doi.org/10.1126/science.265.5175.1084.

Ebert, D., Carius, H. J., Little, T., & Decaestecker, E. (2004). The Evolution of Virulence When Parasites Cause Host Castration and Gigantism. *The American Naturalist, 164*(S5), S19-S32. doi:10.1086/424606.

Ebert, D. (2005). Ecology, Epidemiology, and Evolution of Parasitism in Daphnia [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology. Information. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Books.

Ebbert, M. A. (1991). The Interaction Phenotype in the *Drosophila willistoni*-Spiroplasma Symbiosis. *Evolution, 45*(4), 971-988.

Ebbert, M. A. (1995). Variable Effects of Crowding on *Drosophila* Hosts of male-lethal and non-male-lethal spiroplasmas in Laboratory Populations. *Heredity*, 74, 227-240.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32*(5), 1792-1797.

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*(6058), 990–992. http://doi.org/10.1126/science.1209609.

Firth, D. (1993). Bias reduction of maximum likelihood estimates, *Biometrika*, *80*(1), 27–38. https://doi.org/10.1093/biomet/80.1.27.

Fleury, F., Gibert, P., Ris, N., & Allemand, R. (2009). Ecology and life history evolution of frugivorous *Drosophila* parasitoids. *Advances in parasitology*, 70, 3-44.

Fleury, F., Ris, N., Allemand, R., Fouillet, P., Carton, Y., & Boulétreau, M. (2004). Ecological and Genetic Interactions in *Drosophila*-parasitoids communities: A Case Study with *D. Melanogaster*, *D. Simulans* and their Common *Leptopilina* Parasitoids in South-eastern France. *Genetica*, 120, 181-194.

Flores, S. V., Evans, A. L., & McAllister, B. F. (2008). Independent Origins of New Sex-Linked Chromosomes in the *melanica* and *robusta* Species Groups of *Drosophila*. *BMC Evolutionary Biology*, 8, 33. http://doi.org/10.1186/1471-2148-8-33.

Fry, A.J., Palmer, M.R., & Rand, D.M. (2004). Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. *Heredity*, 93, 379-389.

Fukatsu, T., Tsuchida, T., Nikoh, N., & Koga, R. (2001). *Spiroplasma* Symbiont of the Pea Aphid, *Acyrthosiphon pisum* (Insecta: Homoptera). *Applied and Environmental Microbiology*, *67*(3), 1284–1291. http://doi.org/10.1128/AEM.67.3.1284-1291.2001.

Gasparich, G. E. (2002). Spiroplasmas: Evolution, adaptation and diversity. *Frontiers in bioscience*, 7, 619-640.

Gérard, C., & Théron, A. (1997). Age/Size- and Time-Specific Effects of *Schistosoma mansoni* on Energy Allocation Patterns of Its Snail Host *Biomphalaria glabrata*. *Oecologia*, *112*(4), 447-452. Retrieved from http://www.jstor.org/stable/4221799.

Gilbert, G. S., & Webb, C. O. (2007). Phylogenetic signal in plant pathogen–host range. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 4979–4983. http://doi.org/10.1073/pnas.0607968104.

Hamilton, W. D. (1967). Extraordinary sex ratios. *Science*, *156*(3774), 477-488. http://dx.doi.org/10.1126/science.156.3774.477.

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

Harisanova, N. T., & Ralchev, K. H. (1985). The effect of 10°C temperature on the embryonic development of Drosophila hydei. *Drosophila Information Service*, 61, 88-89.

Haselkorn, T. S., Markow, T. A., & Moran, N. A. (2009). Multiple introductions of the *Spiroplasma* bacterial endosymbiont into *Drosophila*. *Molecular Ecology*, 18, 1294-1305. doi:10.1111/j.1365-294X.2009.04085.x.

Haselkorn, T. S. (2010a). The Spiroplasma heritable bacterial endosymbiont of Drosophila. *Fly*, *4*(1), 80-87. https://doi.org/10.4161/fly.4.1.10883.

Haselkorn, T. (2010b). Understanding the distribution of the *Spiroplasma* heritable bacterial endosymbiont in *Drosophila* (Doctoral dissertation). *UC San Diego*. ProQuest ID: Haselkorn_ucsd_0033D_11126. Merritt ID: ark:/20775/bb2322073z. Retrieved from https://escholarship.org/uc/item/82r2t37b.

Haselkorn, T. S., Watts, T. D., & Markow, T. A. (2013). Density dynamics of diverse *Spiroplasma* strains naturally infecting different species of *Drosophila*. *Fly*, *7*(3), 204–210. http://doi.org/10.4161/fly.25469.

Haselkorn, T. S. & Jaenike, J. (2015). Macroevolutionary persistence of heritable endosymbionts: acquisition, retention and expression of adaptive phenotypes in *Spiroplasma*. *Molecular Ecology*, 24, 3752-3765. doi:10.1111/mec.13261.

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. Statistics in Medicine, 21, 2409-2419. https://doi.org/10.1002/sim.1047.

Herren, J. K., & Lemaitre, B. (2011). *Spiroplasma* and host immunity: activation of humoral immune responses increases endosymbiont load and susceptibility to certain Gram-negative bacterial pathogens in *Drosophila melanogaster*. *Cellular Microbiology*, 13: 1385-1396. doi:10.1111/j.1462-5822.2011.01627.x.

Herren, J. K., Paredes, J. C., Schüpfer, F., & Lemaitre, B. (2013). Vertical Transmission of a *Drosophila* Endosymbiont Via Cooption of the Yolk Transport and Internalization Machinery. *mBio*, *4*(2), e00532–12. http://doi.org/10.1128/mBio.00532-12.

Herren, J. K., Paredes, J. C., Schüpfer, F., Arafah. K., Bulet, P., & Lemaitre, B. (2014). Insect endosymbiont proliferation is limited by lipid availability. *eLife*, 3:e02964. DOI: 10.7554/eLife.02964.

Hiroyoshi, S., Harvey, J. A., Nakamatsu, Y., Nemoto, H., Mitsuhashi, J., Mitsunaga, T., & Tanaka, T. (2017). Potential Host Range of the Larval Endoparasitoid *Cotesia vestalis* (=*plutellae*) (Hymenoptera: Braconidae). *International Journal of Insect Science*, 9, 1-12. https://doi.org/10.1177/1179543317715623.

Hosmer, D. W., Lemeshow, S., May, S. (2008). *Applied Survival Analysis: Regression Modeling of Time-to-Event Data* (2nd Ed.). New York, NY: Wiley-Interscience.

Huitema, B. D. (2011). *The Analysis of Covariance and Alternatives: Statistical Methods for Experiments, Quasi-Experiments, and Single-Case Studies* (2nd ed.) Hoboken: John Wiley & Sons, Inc.

Hurst, G. D., Anbutsu, H., Kutsukake, M., & Fukatsu, T. (2003). Hidden from the host: *Spiroplasma* bacteria infecting *Drosophila* do not cause an immune response, but are suppressed by ectopic immune activation. *Insect Molecular Biology*, 12, 93-97. doi:10.1046/j.1365-2583.2003.00380.x.

Hutchence, K. J., Fischer, B., Paterson, S., & Hurst, G. D. (2011). How do insects react to novel inherited symbionts? A microarray analysis of *Drosophila melanogaster* response to the presence of natural and introduced *Spiroplasma*. *Molecular Ecology*, 20, 950-958. doi:10.1111/j.1365-294X.2010.04974.x.

IBM Corp. (2012). IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

Jaenike, J., Polak, M., Fiskin, A., Helou, M., & Minhas, M. (2007). Interspecific transmission of endosymbiotic *Spiroplasma* by mites. *Biology Letters, 3*(1), 23–25. http://doi.org/10.1098/rsbl.2006.0577.

Jaenike, J. (2009). Coupled population dynamics of endosymbionts within and between hosts. *Oikos*, 118, 353- 62. 10.1111/j.1600-0706.2008.17110.x.

Jaenike, J., Unckless, R., Cockburn, S. N., Boelio, L. M., & Perlman, S. J. (2010). Adaptation via Symbiosis: Recent Spread of a *Drosophila* Defensive Symbiont. *Science*, 329, 212-215. https://doi.org/10.1126/science.1188235.

Jeyaprakash, A., &Hoy, M. A. (2000), Long PCR improves *Wolbachia* DNA amplification: *wsp* sequences found in 76% of sixty-three arthropod species. *Insect Molecular Biology*, 9, 393-405. doi:10.1046/j.1365-2583.2000.00203.x.

Kageyama, D., Anbutsu, H., Watada, M., Hosokawa, T., Shimada, M., & Fukatsu, T. (2006). Prevalence of a Non-Male-Killing Spiroplasma in Natural Populations of *Drosophila hydei*. *Applied and Environmental Microbiology*, *72*(10), 6667–6673. http://doi.org/10.1128/AEM.00803-06.

Kageyama, D., Anbutsu, H., Shimada, M., & Fukatsu, T. (2007). Spiroplasma infection causes either early or late male killing in *Drosophila*, depending on maternal host age. *Naturwissenschaften*, *94*(4), 333-337. https://doi.org/10.1007/s00114-006-0195-x.

Kaplan, E., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, *53*(282), 457-481. doi:10.2307/2281868.

Katoh, T. K., Watada, M. (2011). Faunal and seasonal surveys on drosophilid flies by net sweeping in Ehime Prefecture, Shikoku, Japan. *Low Temperature Science*, 69, 121-127.

King, T. A., & Cable, J. (2007). Experimental infections of the monogenean *Gyrodactylus turnbulli* indicate that it is not a strict specialist. *International Journal of Parasitology*, 37, 663–672.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Molecular. Evolution*, 16, 111-120.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33,1870-1874.

Lanave, C., Preparata, G., Saccone, C., & Serio, G. (1984). A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution*, 20, 86-93.

Lee, K. P., Simpson, S. J., Clissold, F. J., Brooks, R., Ballard, J. W. O., Taylor, P. W., … Raubenheimer, D. (2008). Lifespan and reproduction in *Drosophila*: New insights from nutritional geometry. Proceedings of the National Academy of Sciences of the United States of America, 105(7), 2498–2503. http://doi.org/10.1073/pnas.0710787105.

Lee, K. P. (2015). Dietary protein:carbohydrate balance is a critical modulator of lifespan and reproduction in *Drosophila melanogaster*: A test using a chemically defined diet. *Journal of Insect Physiology*, 5, 2-19. htps://doi.org/10.1016/j.jinsphys.2015.02.007.

Lefèvre, T., Adamo, S. A., Biron, D. G., Missé, D, Hughes, D., & Thomas, F. (2009). Invasion of the Body Snatchers: The Diversity and Evolution of Manipulative Strategies in Host–Parasite Interactions. *Advances in Parasitology*, 68, 45-83.

Lefranc, A., & Bungaard, J. (2000). The influence of male and female body size on copulation duration and fecundity in *Drosophila melanogaster*. *Hereditas*, 132, 243-247. doi:10.1111/j.1601-5223.2000.00243.x.

Lipsitch, M., Siller, S., & Nowak, M. A. (1996). The Evolution of Virulence in Pathogens with Vertical and Horizontal Transmission. *Evolution*, 50, 1729-1741. doi:10.1111/j.1558-5646.1996.tb03560.x.

Littell, R.C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., Schabenberger, O. (2006). *SAS for Mixed Models* (2nd ed.). Cary NC: SAS Institute Inc.

Lively, C. M. (1989). Adaptation by a parasitic trematode to local populations of its snail host. *Evolution*, 43, 1663-1671.

Lo, W.-S., Chen, L.-L., Chung, W.-C., Gasparich, G. E., & Kuo, C.-H. (2013). Comparative genome analysis of *Spiroplasma melliferum* IPMB4A, a honeybee-associated bacterium. *BMC Genomics*, 14, 22. http://doi.org/10.1186/1471-2164-14-22.

Long, T. A.., & Pischedda, A. (2005). Do female *Drosophila melanogaster* adaptively bias offspring sex ratios in relation to the age of their mate? *Proceedings of the Royal Society B: Biological Sciences*, *272*(1574), 1781–1787. http://doi.org/10.1098/rspb.2005.3165.

Lorenz, B., Schroeder, J., & Reischl, U. (2002). First evidence of an endogenous *Spiroplasma* sp. infection in humans manifesting as unilateral cataract associated with anterior uveitis in a premature baby. *Graefe's Archives for Clinical and Experimental Ophthalmology*, 240, 348-353. https://doi.org/10.1007/s00417-002-0453-3.

Lymbery, A. J. (1989). Host specificity, host range and host preference. *Parasitology Today*, *5*(9), 298.

Madden, L.Y., & Nault, L.R. (1983). Differential pathogenicity of corn stunting mollicutes to leafhopper vectors in *Dalbulus* and *Baidulus* species. *Phytopathology*, 73, 1608-1614. http://doi**.org/**10.1094/Phyto-73-1608.

Majerus, M. E. N. (2003). *Sex wars: genes, bacteria, and biased sex ratios*. Princeton, NJ: Princeton University Press.

Malogolowkin-Cohen, Ch. (1980). The distribution of Drosophila subobscura in relation to other species in Israel. *Drosophila Information Service*, 55, 91. Retrieved from http://www.ou.edu/journals/dis/DIS55/Malogolowkin-Cohen%2090.pdf.

Malogolowkin-Cohen, Ch., Rodriguez-Pereira, M. A. (1975). Sexual drive of normal and SR flies of *Drosophila nebulosa*. *Evolution*, *29*(3), 579-580.

Markow, T. A. (1985). A comparative investigation of the mating system of *Drosophila hydei*. *Animal Behaviour*, *33*(3), 775-781. http://dx.doi.org/10.1016/S0003-3472(85)80010-5.

Markow, T. A., Beall, S., & Castrezana, S. (2012). The wild side of life. *Fly*, *6*(2), 98-101. DOI: 10.4161/fly.19552.

Markow, T. A., O'Grady, P. M. (2006). *Drosophila: a guide to species identification and use*. Oxford, UK: Elsevier.

Martin, J., Chong, T., & Ferree, P. M. (2013). Male Killing *Spiroplasma* Preferentially Disrupts Neural Development in the *Drosophila melanogaster* Embryo. *PLoS ONE*, *8*(11): e79368. https://doi.org/10.1371/journal.pone.0079368.

Martins, A. B., Ventura, I. M., Klaczko, L. B. (2010). *Spiroplasma* infection in *Drosophila melanogaster*: What is the advantage of killing males? *Journal of Invertebrate Pathology*, 105, 145-150. https://doi.org/10.1016/j.jip.2010.06.002.

Mateos, M., Castrezana, S. J., Nankivell, B. J., Estes, A. M., Markow, T. A., & Moran, N. A. (2006). Heritable Endosymbionts of Drosophila. *Genetics*, *174*(1), 363–376. http://doi.org/10.1534/genetics.106.058818.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). New York, NY: Psychology Press.

Mello, A. F. S., Yokomi, R. K., Melcher, U., Chen, J. C., & Fletcher, J. (2010). Citrus Stubborn Severity Is Associated with *Spiroplasma citri* Titer But Not with Bacterial Genotype. *Plant Disease*, *94*(1), 75-82. https://doi.org/10.1094/PDIS-94-1-0075.

Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com).

Mitsui, H., Achterberg, V. K., Nordlander, G., & Kimura, M. T. (2007). Geographical distributions and host associations of larval parasitoids of frugivorous Drosophilidae in Japan. *Journal of Natural History*, 41, 1731-1738. https://doi.org/10.1080/00222930701504797.

Moir, M. L., Vesk, P. A., Brennan, K. E. C., Keith, D. A., Hughes, L., & McCarthy, M. A. (2010). Current constraints and future directions in estimating coextinction. *Conservation Biology*, 24, 682–690. https://doi.org/10.1111/j.1523-1739.2009.01398.x.

Montenegro, H., Solferini, V. N., Klaczko, L. B., & Hurst, G. D. (2005), Male-killing *Spiroplasma* naturally infecting *Drosophila melanogaster*. *Insect Molecular Biology*, 14, 281-287. doi:10.1111/j.1365-2583.2005.00558.x.

Montenegro, H., Hatadani, L. M., Medeiros, H. F., & Klaczko, L. B. (2006). Male killing in three species of the *tripunctata* radiation of *Drosophila* (Diptera: Drosophilidae). *Journal of Zoological Systematics and Evolutionary Research*, *44*, 130-135. doi:10.1111/j.1439-0469.2006.00353.x.

Molnar, K., Ostoros, G., & Baska, F. (2005). Cross-infection experiments confirm the host specificity of *Goussia* spp. (Eimeriidae: Apicomplexa) parasitizing cyprinid fish. Acta Protozoologica, 44, 43-49.

Morehead, S. A., & Feener, D. H. (2000). An experimental test of potential host range in the ant parasitoid *Apocephalus paraponerae*. *Ecological Entomology*, 25, 332-340. doi:10.1046/j.1365-2311.2000.00261.x.

Mueller, N. J., Tini, G. M., Weber, A., Gaspert, A., Husmann, L., Bloemberg, G., Boehler, A., & Benden, C. (2015). Hepatitis from *Spiroplasma* sp. in an Immunocompromised Patient. *American Journal of Transplantation*, 15, 2511-2516. doi:10.1111/ajt.13254.

Nakayama, S., Parratt, S.R., Hutchence, K.J., Lewis, Z., Price, T.A., & Hurst, G.D. (2015). Can maternally inherited endosymbionts adapt to a novel host? Direct costs of *Spiroplasma* infection, but not vertical transmission efficiency, evolve rapidly after horizontal transfer into *D. melanogaster*. *Heredity*, 114, 539-543. doi:10.1038/hdy.2014.112.

Nei, M., & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. Oxford, NY: Oxford University Press.

Osaka, R., Nomura, M., Watada, M., Kageyama, D. (2008). Negative Effects of Low Temperatures on the Vertical Transmission and Infection Density of a Spiroplasma endosymbiont in *Drosophila hydei. Current Microbiology*, 57, 335-339.

Osaka, R., Watada, M., Kageyama, D., Nomura, M. (2010). Population dynamics of a maternally-transmitted *Spiroplasma* infection in *Drosophila hydei*. *Symbiosis*, 52, 41-45. https://doi.org/10.1007/s13199-010-0100-4.

Osaka, R., Ichizono, T., Kageyama, D., Nomura, M., & Watada, M. (2013a). Natural variation in population densities and vertical transmission rates of a *Spiroplasma* endosymbiont in *Drosophila hydei*. *Symbiosis*, 60, 73-78. https://doi.org/10.1007/s13199-013-0242-2.

Osaka, R., Watada, M., Kageyama, D., & Nomura, M. (2013b). Detection of *Spiroplasma* from the mite *Macrocheles* sp. (Acari; Macrochelidae) ectoparasitic to the fly *Drosophila hydei* (Diptera; Drosophilidae): a possible route of horizontal transmission? *Symbiosis*, 60, 79-84. https://doi.org/10.1007/s13199-013-0241-3.

Ota, T., Kawabe, M., Oishi, K., Poulson, D. F. (1979). Non-male-killing spiroplasmas in *Drosophila hydei*. *The Journal of Heredity*, *70*(3), 211-213. https://doi.org/10.1093/oxfordjournals.jhered.a109235.

Page, R. D. M. (1994). *Parallel phylogenies: Reconstructing the history of host-parasite assemblages. Cladistics*, *10*(2), 155-173.

Perlman, S. J. & Jaenike, J. (2003). Infection success in novel hosts: an experimental and phylogenetic study of drosophila-parasitic nematodes. *Evolution*. 57(3), 544–557. https://doi.org/10.1111/j.0014-3820.2003.tb01546.x.

Partridge, L. (1986). Sexual activity and life span. In: K. G. Collatz & R. S. Sohal (Eds.). *Insect Aging. Strategies and Mechanisms (pp.* 45–54). New York, NY: Springer-Verlag.

Pitnick, S. (1993). Operational Sex Ratios and Sperm Limitation in Populations of *Drosophila pachea*. *Behavioral Ecology and Sociobiology*, *33*(6), 383-391. Retrieved from http://www.jstor.org/stable/4600897.

Pitnick, S., Markow, T. A., & Spicer, G. S. (1995). Delayed male maturity is a cost of producing large sperm in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(23), 10614–10618.

Pool, J. E., Wong, A., & Aquadro, C. F. (2006). Finding of male-killing *Spiroplasma* infecting *Drosophila melanogaster* in Africa implies transatlantic migration of this endosymbiont. *Heredity*, *97*(1), 27–32. http://doi.org/10.1038/sj.hdy.6800830.

Poulin, R. (2007). *Evolutionary Ecology of Parasites* (2nd ed.). Princeton NJ: Princeton University Press. Retrieved from http://www.jstor.org/stable/j.ctt7sn0x.

Poulin, R., & Keeney, D. B. (2008). Host specificity under molecular and experimental scrutiny. *Trends in Parasitology*, 24, 24–28. https://doi.org/10.1016/j.pt.2007.10.002.

Poulin, R., Krasnov, B. R., & Mouillot, D. (2011). Host specificity in phylogenetic and geographic space. *Trends in Parasitology, 27*(8): 355-361. https://doi.org/10.1016/j.pt.2011.05.003.

Pruesse, E., Peplies, J., & Glöckner, F. O. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*, *28*(14), 1823–1829. http://doi.org/10.1093/bioinformatics/bts252.

R Core Team (2013). R: A language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. URL http://www.R-project.org/.

Reed, L. K., Nyboer, M., & Markow, T. (2007). Evolutionary relationships of *Drosophila mojavensis* geographic host races and their sister species *Drosophila arizonae*. *Molecular Ecology*, 16, 1007-1022. doi:10.1111/j.1365-294X.2006.02941.x).

Regassa, L. B., & Gasparich, G. E. (2006). Spiroplasmas: evolutionary relationships and biodiversity. *Frontiers in Bioscience*, 11, 2983-3002.

Rohde, K. & Rohde, P. P. (2005). The ecological niches of parasites. In K. Rohde (Ed.), *Marine Parasitology* (pp. 286-293). Collingwood: CSIRO Publishing.

Rohde, K., & Rohde, P. P. (2008). How to measure ecological host specificity. *Vie et Milieu-Life and Environment*, ***58*(2)**: 121–124.

Ruiz-González, M. X., & Brown, M. J. (2006). Honey bee and bumblebee trypanosomatids: specificity and potential for transmission. *Ecological Entomology*, 31, 616-622. doi:10.1111/j.1365-2311.2006.00823.x.

Sgrò, C. M., Partridge, L. (2000). Evolutionary Responses of the Life History of Wild-Caught *Drosophila melanogaster* to Two Standard Methods of Laboratory Culture. *The American Naturalist, 156*(4), 341-353. doi:10.1086/303394.

SAS Institute Inc. (2010). SAS/STAT® 9.22 User's Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. (2014). SAS Enterprise Guide® 7.1: Reference. Cary, NC: SAS Institute Inc.

Serga, S. V., Rozhok, A. I., Protsenko, O. V., Kozeretska, I, A. & Mousseau, T. A. (2010). Spiroplasma in natural populations of *Drosophila melanogaster* from Ukraine. *Drosophila Information Service*, 93, 148-154.

Shahjahan, R. M., Rendón, P. A., Cook, L. M., & Wood, R. J. (2006). Male biased sex ratio in the Mediterranean fruit fly *Ceratitis capitata*, an example of Y-chromosome meiotic drive. *Heredity*, 96, 464-470. https://doi.org/10.1038/sj.hdy.6800824.

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*. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(46), 20051–20056. http://doi.org/10.1073/pnas.1009906107.

Shimizu-kaya, U., Okubo, T., Inui, Y., & Itioka, T. (2013). Potential host range of myrmecophilous *Arhopala* butterflies (Lepidoptera: Lycaenidae) feeding on *Macaranga* myrmecophytes. *Journal of Natural History*, 47, 43-44, 2707-2717, DOI: 10.1080/00222933.2013.791943.

Seetharam, A. S., & Stuart, G. W. (2013). Whole genome phylogeny for 21 *Drosophila* species using predicted 2b-RAD fragments. PeerJ 1:e226 https://doi.org/10.7717/peerj.226.

Stone, B. K., Scibilia, B., Pammer, C., Steele, C., & Keller, D. (2012, March 8). How to Identify the Distribution of Your Data using Minitab. [Blog Post]. Retrieved from http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-identify-the-distribution-of-your-data-using-minitab.

Tantawy, A., & Vetukhiv, M. (1960). Effects of Size on Fecundity, Longevity and Viability in

Populations of *Drosophila pseudoobscura*. *The American Naturalist, 94*(879), 395-403. Retrieved from http://www.jstor.org/stable/2458805.

Taraschewski, H. (2006). Hosts and parasites as aliens. *Journal of Helminthology, 80*(2), 99-128. doi:10.1079/JOH2006364.

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. In R. M. Miura (Ed.), *Lectures on Mathematics in the Life Sciences* (pp. 57-86). Providence RI, American Mathematical Society.

Tinsley, M. C., & Majerus, M. E. (2007). Small steps or giant leaps for male-killers? Phylogenetic constraints to male-killer host shifts. *BMC Evolutionary Biology*, *7*, 238. http://doi.org/10.1186/1471-2148-7-238.

Toda, M. J. (1979). A Preliminary Note on winter Drosophilid Flies in Southern Japan, with Special Reference to Reproductive Conditions. *Low temperature science, Series B, Biological sciences*, 37, 39-45.

Vale, P. F., & Little, T. J. (2012). Fecundity compensation and tolerance to a sterilizing pathogen in *Daphnia*. *Journal of Evolutionary Biology*, *25*(9), 1888–1896. http://doi.org/10.1111/j.1420-9101.2012.02579.x.

Van Der Linde, K., Houle, D., Spicer, G., & Steppan, S. (2010). A supermatrix-based molecular phylogeny of the family Drosophilidae. *Genetics Research, 92*(1), 25-38. doi:10.1017/S001667231000008X.

Walter, J. (2008). Ecological Role of Lactobacilli in the Gastrointestinal Tract: Implications for Fundamental and Biomedical Research. *Applied and Environmental Microbiology*, *74*(16), 4985–4996. http://doi.org/10.1128/AEM.00753-08.

Wang, X., & Messing, R.H. (2007). Potential Host Range of the Newly Introduced Aphid

Parasitoid *Aphidius transcaspicus* (Hymenoptera : Braconidae) in Hawaii. *Proceedings of the Hawaian Entomological Socity*, 38, 81-86.

Watts, T., Haselkorn, T. S., Moran, N. A., & Markow, T. A. (2009). Variable Incidence of *Spiroplasma* Infections in Natural Populations of *Drosophila* Species. *PLoS ONE*, *4*(5), e5703. http://doi.org/10.1371/journal.pone.0005703.

Weeks, A. R., & Stouthamer, R. (2004). Increased fecundity associated with infection by a *Cytophaga*-like intracellular bacterium in the predatory mite, *Metaseiulus occidentalis*. *Proceedings of the Royal Society B: Biological Sciences*, *271*(Suppl 4), S193–S195.

West, B. T., Welch, K. B., Galecki, A. T. (2014). *Linear Mixed Models: A Practical Guide using Statistical Software* (2nd ed.). Boca Raton, FL: Chapman Hall / CRC Press.

Werren, J. H., O'Neill, S. L. (1997). The evolution of heritable symbionts. In S. L. O'Neill, A. A. Hoffmann, & J. H. Werren (Eds.), *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction* (pp. 1-41). Oxford: Oxford University Press.

Wildt, A. R., & Ahtola, O. T. (1978). *Analysis of Covariance*. Iowa City: Sara Miller McCune, Sage Publications, Inc.

Williamson, D. L., & Poulson, D. F. (1979). Sex ratio organisms (*Spiroplasma*s) of *Drosophila*. In R. F. Whitcomb & J. G. Tully (Eds.). *The Mycoplasmas* (Vol. 3) (pp. 175-208). New York: Academic Press. https://doi.org/10.1016/B978-0-12-078403-5.50012-8.

Williamson, D. L., Gasparich, G. E., Regassa, L. B., Saillard, C., Renaudin, J., Bové, J. M., &Whitcomb, R. F. (2015). *Spiroplasma*. In W. B. Whitman, F. Rainey, P. Kämpfer, M. Trujillo, J. Chun, P. DeVos, B. Hedlund, & S. Dedysh (Eds.). Bergey's Manual of Systematics of Archaea and Bacteria (pp. 1-48). doi:10.1002/9781118960608.gbm01262.

Wolfram Alpha LLC. 2009. Wolfram|Alpha. http://www.wolframalpha.com/input/?i=2%2B2 (access January 30, 2018).

Xie, J., Vilchez, I., & Mateos, M. (2010). *Spiroplasma* Bacteria Enhance Survival of *Drosophila hydei* Attacked by the Parasitic Wasp *Leptopilina heterotoma*. *PLoS ONE*, *5*(8), e12149. http://doi.org/10.1371/journal.pone.0012149.

Xie, J., Tiner, B., Vilchez, I., & Mateos, M. (2011). Effect of the *Drosophila* endosymbiont *Spiroplasma* on parasitoid wasp development and on the reproductive fitness of wasp-attacked fly survivors. *Evolutionary Ecology*, *53*(5), 1065–1079. http://doi.org/10.1007/s10682-010-9453-7.

Xie, J., Butler, S., Sanchez, G., & Mateos, M. (2014). Male killing *Spiroplasma* protects *Drosophila melanogaster* against two parasitoid wasps. Heredity, 112(4), 399–408. http://doi.org/10.1038/hdy.2013.118.

Xie, J., Winter, C., Winter, L. A., & Mateos, M. (2015). Rapid spread of the defensive endosymbiont *Spiroplasma* in *Drosophila hydei* under high parasitoid wasp pressure. *FEMS microbiology ecology, 91(2),* 1-11. https://doi.org/10.1093/femsec/fiu017.

Yamamoto, A., Ohba, S. (1984). Temperature Preferences of Eleven *Drosophila* Species from Japan: The Relationship between Preferred Temperature and Some Ecological Characteristics in Their Natural Habitats. *Zoological Science*, 1, 631-640.

Yamamoto, A. H. (1992). Niche differentiation of drosophilids in Oiso, Japan. *Memories of Konan University, Science Series*, *39*(2), 287–299.

Ye, F., Melcher, U., & Fletcher, J. (1997). Molecular characterization of a gene encoding a membrane protein of *Spiroplasma citri*. *Gene*, 189, 95-100. https://doi.org/10.1016/S0378-1119(96)00840-2.

Yeh, S., Grotthuss, M.V., Gandasetiawan, K.A., Jayasekera, S.U., Xia, X., Chan, C., Jayaswal, V., & Ranz, J.M. (2014). Functional divergence of the miRNA transcriptome at the onset of *Drosophila* metamorphosis. *Molecular biology and evolution*, *31*(10), 2557-72.

Yu, H., Liu, S., Ruan, K., Chen, Y., & Wand, Z. (2009). Advances in classification and biodiversity of spiroplasmas – A review. *Acta Microbiologica Sinica*, *49*(5), 567-572.