bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Infection with endosymbiotic *Spiroplasma* disrupts tsetse (*Glossina fuscipes fuscipes*) metabolic and reproductive homeostasis

Short title: *Spiroplasma* infection detrimentally impacts tsetse reproductive phenotypes

Jae Hak Son[1,*], Brian L. Weiss[1,*,#], Daniela I. Schneider[1,a], Kisweda-sida M. Dera[2,3], Fabian Gstöttenmayer[2], Robert Opiro[4], Richard Echodu[4], Norah P. Saarman[5,b], Geoffrey M. Attardo[1,c], Maria Onyango[1,d], Adly M.M. Abdalla[2], Serap Aksoy[1,#]

[1]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA

[2]Insect Pest Control Laboratory, Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture, P.O. Box 100, Wagrammer Straße 5, A-1400, Vienna, Austria

[3]Insectarium de Bobo-Dioulasso - Campagne d'Eradication de la mouche Tse´-tse´ et de la Trypanosomiase (IBD-CETT), Bobo-Dioulasso, Burkina Faso.

[4]Department of Biology, Faculty of Science, Gulu University, Uganda

[5]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

* Contributed equally to this work

Corresponding authors

Current addresses:

[a]Moffit Cancer Center, Tampa, Florida, USA

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[b]Department of Biology, Utah State University, Logan, Utah, USA

[c]Department of Entomology and Nematology, University of California, Davis, USA

[d]Wadsworth Centre, New York State Department of Health, Slingerlands, New York, USA

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Tsetse flies (*Glossina* spp.) house a population-dependent assortment of microorganisms that can include pathogenic African trypanosomes and maternally transmitted endosymbiotic bacteria, the latter of which mediate numerous aspects of their host's metabolic, reproductive, and immune physiologies. One of these endosymbionts, *Spiroplasma*, was recently discovered to reside within multiple tissues of field captured and laboratory colonized tsetse flies grouped in the Palpalis subgenera. In various arthropods, *Spiroplasma* induces reproductive abnormalities and pathogen protective phenotypes. In tsetse, *Spiroplasma* infections also induce a protective phenotype by enhancing the fly's resistance to infection with trypanosomes. However, the potential impact of *Spiroplasma* on tsetse's viviparous reproductive physiology remains unknown. Herein we employed high-throughput RNA sequencing and laboratory-based functional assays to better characterize the association between *Spiroplasma* and the metabolic and reproductive physiologies of *G. fuscipes fuscipes* (*Gff*), a prominent vector of human disease. Using field-captured *Gff*, we discovered that *Spiroplasma* infection induces changes of sex-biased gene expression in reproductive tissues that may be critical for tsetse's reproductive fitness. Using a *Gff* line composed of individuals heterogeneously infected with *Spiroplasma*, we observed that the bacterium and tsetse host compete for finite nutrients, which negatively impact female fecundity by increasing the length of intrauterine larval development. Additionally, we found that when males are infected with *Spiroplasma*, the motility of their sperm is compromised following transfer to the female spermatheca. As such, *Spiroplasma* infections appear to adversely impact male reproductive fitness by decreasing the

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

51 competitiveness of their sperm. Finally, we determined that the bacterium is maternally
52 transmitted to intrauterine larva at a high frequency, while paternal transmission was
53 also noted in a small number of matings. Taken together, our findings indicate that
54 *Spiroplasma* exerts a negative impact on tsetse fecundity, an outcome that could be
55 exploited for reducing tsetse population size and thus disease transmission.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Author Summary

Endosymbiotic bacteria regulate numerous aspects of their host's reproductive physiology. Natural populations of the tsetse fly, *Glossina fuscipes fuscipes* (*Gff*), house heterogeneous infections with the bacterium *Spiroplasma glossinidia*. Infection with the bacterium results in the presentation of several phenotypes in both male and female *Gff* that would put them at a significant reproductive disadvantage when compared to their counterparts that do not house the bacterium. These *Spiroplasma* induced phenotypes include changes in sex–biased gene expression in the reproductive organs, a depletion in the availability of metabolically critical lipids in pregnant females that results in delayed larval development, and compromised sperm fitness. These findings indicate that *Spiroplasma* exerts an overall negative impact on both male and female reproductive fitness and thus likely has a profound effect on fly population structure. This outcome, in conjunction with the fact that *Spiroplasma* infected tsetse are unusually refractory to infection with pathogenic African trypanosomes, indicates that the bacterium could be experimentally exploited to reduce disease transmission through the fly.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Tsetse flies (*Glossina spp*.), which vector pathogenic African trypanosomes, reproduce via a process called adenotrophic viviparity. Following mating, female tsetse ovulate one oocyte per gonotrophic cycle (GC). The oocyte is fertilized in the maternal uterus by sperm that are released from the spermathecae. Unlike in most arthropods, tsetse embryos and larvae develop exclusively *in utero*, and larvae receive nourishment in the form of maternal milk secretions. Following the completion of larvigenesis, the mother gives birth to a fully developed 3rd instar larva that pupates within 30 minutes. This process repeats itself approximately every 10 days [1]. During the mating process, male tsetse transfer seminal fluid (SF), which contains sperm as well as numerous male accessory gland (MAG) derived proteins, into the reproductive tract of receptive females. Once in the female's uterus, this mixture forms into a proteinaceous spermatophore that facilitates successful transfer of sperm into the spermathecae for long-term storage [2-6]. Because tsetse's viviparous reproductive strategy results in the production of relatively few offspring, targeting reproduction can be a most effective approach to reduce tsetse populations. Methods to inhibit tsetse fecundity can be highly effective in reducing disease transmission given that the fly is an obligate vector for parasite transmission [1,7]. In fact, the application of sterile male programs has been very successful with tsetse [8] and is currently endorsed to eliminate tsetse populations on the African continent [8,9].

Tsetse flies have evolved long-term associations with several vertically transmitted endosymbiotic bacteria that impact numerous aspects of their host's nutritional, developmental and reproductive physiology. All tsetse flies house the obligate endosymbiont *Wigglesworthia*, which provides nutrients absent from the fly's vertebrate

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

95 blood-specific diet [10-13]. In addition to *Wigglesworthia*, laboratory reared and natural
96 tsetse populations can also harbor facultative *Sodalis* as well as parasitic *Wolbachia* and
97 *Spiroplasma* [14,15]. To date *Spiroplasma glossinidia* (hereafter referred to as
98 *Spiroplasma*), which is closely related to the *Spiroplasma poulsonii* strain MRSO, has
99 been found exclusively within tsetse flies of the subgenus Palpalis [14,15]. In both field
100 captured and laboratory reared *Glossina fuscipes fuscipes* (*Gff*), the bacterium resides in
101 reproductive and digestive tissues as well as hemolymph [14]. *Spiroplasma* induces an
102 immune protective effect in laboratory reared *Gff*, as flies that house the bacterium are
103 significantly more resistant to infection with trypanosomes than are flies that do not [15].
104 The mechanism that underlies *Spiroplasma* enhanced refractoriness to trypanosome
105 infection in *Gff* is currently unknown. However, in other insects the bacterium confers
106 pathogen (e.g., nematodes, fungi and parasitoid wasps) resistant phenotypes through the
107 production of immune effector molecules and/or through nutrient scavenging that limits
108 metabolically critical nutrients for other pathogens [16-19].

109 Infection with *Spiroplasma* can also impact host reproductive fitness and lead to
110 reproductive abnormalities. In *Drosophila,* females infected with *Spiroplasma* are less
111 fecund and produce fewer eggs, which may be a consequence of nutritional competition
112 between the fly and bacterium for metabolically important free (hemolymph-borne) lipids
113 [20]. In several insect taxa, the bacterium induces a male killing phenotype [21-23]. The
114 potential functional role of *Spiroplasma* in tsetse reproductive physiology is currently
115 unknown. Herein we performed a detailed investigation of the molecular and physiological
116 associations that characterize reproductive aspects of the *Gff-Spiroplasma* symbiosis.
117 Specifically, we performed RNA sequencing (RNA-seq) of reproductive tissues from field

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

118 captured (Uganda) *Spiroplasma* infected and uninfected male and female *Gff* and report
119 on genes and pathways that are differentially regulated in the presence of the bacterium.
120 We also made use of a *Gff* lab line that carries a heterogenous infection with *Spiroplasma*
121 to characterize the trans-generational transmission dynamics of this endosymbiont and
122 to characterize *Spiroplasma*-induced metabolic and reproductive phenotypes in tsetse.
123 Knowledge obtained from this study provides insight into the physiological mechanisms
124 that underlie the tsetse-*Spiroplasma* symbiosis and may have translational implications
125 with respect to controlling tsetse populations and the ability of the fly to transmit African
126 trypanosomes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Results

Overall gene expression profile

To understand the potential effects of *Spiroplasma* infection on tsetse's reproductive physiology, we compared the RNA-seq data obtained from the male reproductive organs (testis, accessory gland and ejaculatory duct) of *Spiroplasma* infected (*Gff*$^{Spi+}$) and uninfected (*Gff*$^{Spi-}$) individuals obtained from wild populations in Uganda. We similarly obtained and compared RNA-seq data from *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ female reproductive organs (ovaries and oocytes) isolated from the same populations. The percentage of reads from the different biological replicates that mapped to the *Gff* reference genome ranged from 49% to 81% (S1 Table).

We used principal component (PC) and hierarchical cluster (HC) analyses to compare the overall gene expression profiles across all biological replicates according to sex and *Spiroplasma* infection status (Fig 1A and 1B). We found that PC1 and PC2 accounted for 80% and 7% of variance across all biological replicates, respectively. The variance in PC1 could be explained by differences in sex, with females clustering on the left side along the PC1 axis, and males clustering on the right side along the PC1 axis (Fig 1A). HC analyses also demonstrated a similar clustering result among different biological replicates (Fig 1B). Female and male clusters were distant from each other, and within the male samples, *Gff*$^{Spi+}$ were separated from the *Gff*$^{Spi-}$ samples. However, within the female samples, one *Gff*$^{Spi+}$ (female 4) and one *Gff*$^{Spi-}$ (female 3), were distant to the other three female samples (one *Gff*$^{Spi+}$ and two *Gff*$^{Spi-}$). This difference we noted within the female dataset could be due to the presence of greater variation in gene

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

149 expression within the female reproductive tissues when compared to males, which is
150 similar to the results shown in our PCA.
151
152 ***The impact of Spiroplasma on sex-biased gene expression in males and females***
153 As our biological replicates clustered by sex, we first quantified the proportion of
154 differentially expressed (DE) genes between females and males (sex-biased genes)
155 using sex as a factor in the model we created in DESeq2 (see Materials and Methods for
156 details). Of the 15,247 genes annotated in the *Gff* genome, we detected 10,540 genes
157 expressed in our transcriptomes, with an adjusted *p*-value created by FDR correction [24]
158 (**S1 Data)**. We observed that 21.7% of these genes (2,288/10,540) were preferentially
159 expressed in females (female-biased genes; $\log_2$male/female<0 and adjusted p<0.05),
160 while 20.5% (2,164/10,540) were preferentially expressed in males (male-biased genes;
161 $\log_2$male/female>0 and adjusted p<0.05). Using pairwise comparisons for *Spiroplasma*
162 infection status within each sex from the DESeq2 model, we found that a total of 194 and
163 299 genes were DE upon *Spiroplasma* infection in males and females, respectively.
164 Within the male dataset, we determined that 117 (60.3%) and 77 (39.7%) of the DE genes
165 were up- and down-regulated upon infection (adjusted p<0.05), respectively (Fig 2A, S2
166 Table). A similar analysis of the 299 DE genes in females showed that 61 (21.4%) and
167 238 (79.6%) were up- and down-regulated in the presence of *Spiroplasma,* respectively
168 (Fig 2B, S3 Table).
169 *Spiroplasma* infections can manipulate the reproductive physiology of their host
170 insects [25]. In *Drosophila*, infections with the *S. poulsonii* strain MSRO, which is closely
171 related to the *Spiroplasma* strain that infects *Gff,* affects gene regulation and dosage
172 compensation machinery in males, resulting in male-killing during embryogenesis [22,26].

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Because genes displaying sex-biased expression profiles in reproductive tissues can confer sexual dimorphic phenotypes [27,28], we first identified the sex-biased genes expressed in male and female gonads, and next evaluated their DE status based on *Spiroplasma* infection. We determined that 41 of the 117 genes (35%) up-regulated, and 41 of the 77 genes (53%) down-regulated in Gff^{Spi+} males are male-biased (Fig 2A, S2 Table). We also noted that the DE male-biased genes (n=82) identified within the male DE gene set (n=194) are enriched when compared to genome-wide male-biased genes (n=2164) (Fig 2A, S2 Table; $p<0.001$ for up-regulation and $p<10^{-9}$ for down-regulation in Fisher's exact test). We next examined the transcript abundance of the 41 up-regulated male-biased genes and determined their contribution to be 0.85% of the total transcriptome in the infected group, up from the 0.06% in the uninfected group (Fig 2C). We similarly analyzed the 41 down-regulated male-biased genes and determined their contribution to be 0.29% of the total transcriptome in the infected group, down from the 0.51% in the uninfected group (Fig 2C).

We next performed the same analysis with the female DE gene dataset. We determined that 48 of the 61 genes (78%) up-regulated in the Gff^{Spi+} females were female-biased and enriched within the up-regulated genes when compared to genome-wide female-biased genes (n=2288) (S3 Table; $p<10^{-20}$ in Fisher's exact test). Conversely, only one of the 238 genes down-regulated in Gff^{Spi+} females was female-biased (Fig 2B, S3 Table), while 234 (98%) were male-biased (Fig 2B, S3 Table). The male-biased genes were thus overrepresented within the down-regulated genes (S3 Table; $p<10^{-149}$ in Fisher's exact test). We also examined the transcript abundance of the 48 up-regulated female-biased genes and determined their contribution to be 14.11% of the total

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

196 transcriptome in the infected group, up from the 4.02% transcript abundance in the
197 uninfected group (Fig 2D). The transcript abundance of the 234 down-regulated male-
198 biased genes in the female dataset represented 1.16% of the total transcriptome in the
199 infected group, down from 17.21% transcript abundance in the uninfected group (Fig 2D).
200 In both sexes, unbiased (non-sex biased) genes were underrepresented in the DE
201 datasets when compared to the genome-wide unbiased genes (Fig 2; S2 and S3 Table;
202 $p<10^{-4}$ for males and $p<10^{-9}$ for females in Fisher's exact test). Collectively, these results
203 suggest that *Spiroplasma* infections strongly influence the expression of sex-biased
204 genes in tsetse's reproductive tissues. Given that sexual dimorphisms are largely shaped
205 by sex-biased gene expressions [27,28], the significant number of sex-biased genes and
206 transcript abundance affected by *Spiroplasma* in females relative to males suggests that
207 the endosymbiont may affect female reproductive physiology to a greater extent.
208
209 **_Spiroplasma_ infection effects on genes that encode spermatophore constituents**
210 When tsetse copulate, the ejaculate, which is composed of sperm and secretory
211 products (i.e., seminal fluid) derived from the male testes and accessory glands, are
212 transferred to the female uterus where they are encapsulated into a spermatophore.
213 The spermatophore functions as a protective container for the ejaculate and ensures
214 that sperm and seminal fluid are delivered to the female spermathecal ducts, which in
215 turn modifies female behavior, including potential inhibition of further matings by
216 competing males [3-5]. Within 24 h of the commencement of copulation, sperm are
217 transferred to the female's spermatheca and the spermatophore is discharged from the
218 uterus [3]. We had previously identified the products detected in the spermatophore in
219 *G. morsitans morsitans (Gmm)* through proteomics analysis of spermatophores

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

collected shortly after completion of copulations [2]. We first identified the homologs of these *Gmm* proteins in *Gff* and then evaluated their expression status in the presence of *Spiroplasma*. We found that of the 287 genes whose products were detected in the *Gmm* spermatophore, seven were DE in Gff^{Spi+} males and 12 were DE in Gff^{Spi+} females relative to Gff^{Spi-} datasets (S1 Data). Of the seven genes encoding spermatophore proteins contributed by male gonads, six were up-regulated and one was down-regulated in the presence of *Spiroplasma*. Of the 12 genes encoding spermatophore proteins contributed by female reproductive organs, four were up-regulated and eight were down-regulated in the presence of *Spiroplasma*. Among the DE genes in females were two that encode transcription factors and three that encode serine protease inhibitors, while the up-regulated genes in males primarily encoded cuticle related proteins.

Immunity genes up-regulated in Gff^{Spi+} males

Long term persistence of endosymbionts requires evasion of, and/or resistance against, host immune responses. Although the absence of metabolically costly immune responses would benefit host fitness, induced immunity can also be beneficial to the symbiosis as it confers resistance to other pathogens that could compete for limited nutritional resources. In *Drosophila*, *Spiroplasma* infections do not activate host immune responses, and the bacterium is not susceptible to either the cellular or humoral arms of the fly's immune system [29]. To understand tsetse-*Spiroplasma* immune dynamics, we investigated whether DE genes in Gff^{Spi+} males and females encoded immunity related functions. We found that in the male gonads, the presence of *Spiroplasma* significantly induces Toll pathway constituents, including *toll-like receptor 7* (GFUI009072), the antimicrobial

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

peptide *defensin* (GFUI031425), and *easter* (GFUI050711), which encodes a serine-protease required to process the extracellular Toll ligand Spätzle. In *Spiroplasma* infected *Drosophila*, experimental activation of antimicrobial immune pathways actually increased *Spiroplasma* titers. We also noted that in addition to the Toll signaling pathway, a number of abundantly expressed Mucin encoding genes (GFUI039642, GFUI016405, GFUI054349, GFUI017943) were induced in *Gff*$^{Spi+}$ males. Mucins are produced by epithelial tissues where they function in different roles from lubrication to cell signaling to forming chemical barriers as well as binding and entrapping pathogens [30].

Gene ontology (GO) analysis of DE gene products

To understand the major putative function(s) for the 493 DE gene products associated with *Spiroplasma* infection in *Gff* reproductive physiology, we performed GO (gene ontology) term enrichment analysis [31]. We found that within the up-regulated genes in *Gff*$^{Spi+}$ males, the significantly enriched GO terms are associated with cuticle development, chitin process, cell adhesion, defense response and receptor activity (S1 Fig A; S3 Data). The down-regulated genes in *Gff*$^{Spi+}$ males were enriched for lipid catabolic process and peptidyl-dipeptidase activity (S1 Fig B; S3 Data). The up-regulated genes in *Gff*$^{Spi+}$ females were enriched for lipid metabolic process and defense response (S1 Fig C; S3 Data) while the down-regulated genes in *Gff*$^{Spi+}$ females were enriched for peptidase activity, proteolysis, and chitin metabolic process (S1 Fig D; S3 Data).

Spiroplasma infection effects on female gene expression

The up-regulated female-biased genes in *Gff*$^{Spi+}$ females compromised over 14% of the entire transcriptome, while these same genes comprised only 4% of the total

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

transcriptome in *Gff*$^{Spi-}$ females. Among the up-regulated gene dataset, we noted the presence of several highly abundant tsetse milk proteins, including Mgp1 (GFUI006902), Mgp10 (GFUI050429), Mgp4 (GFUO050451). In addition, we detected high level expression of a gene annotated as Apolipoprotein (GFUI006901) located adjacent to the *mgp1* (GFUI006902) locus. Milk proteins (Mgp1-10) represent a family of tsetse-specific proteins that are abundantly expressed in the female accessory gland and are a prominent component of tsetse milk [32]. When pregnant, female tsetse lactate, and the intrauterine larvae acquire nourishment in the form of milk secreted by cells of the accessory gland [1]. These proteins are thought to be mostly lipid carriers, and experimental evidence suggests that they act in this capacity as lipid emulsification agents and possible phosphate carrier molecules [33]. Our field collected female gonad samples may have been contaminated with accessory gland tissue, which is composed of an extensive tubular network that is difficult to separate from the female reproductive tissues. Hence, the transcripts that encode Mgp1, Mgp10 and Mgp4, which are abundant in our transcriptomes, may originate from the accessary gland. Furthermore, the pregnancy status of the female samples collected in the field are unknown, and hence the variable expression levels we observed of milk protein encoding genes may reflect the different pregnancy stages of the analyzed samples. Nevertheless, the higher levels of milk protein production in infected females may reflect an adaptive response to compensate for the nutrients scavenged by the bacterium that are necessary for proper larval development. Future laboratory experiments with carefully dissected age and pregnancy matched *Spiroplasma* infected and uninfected females obtained from a

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

heterogenous laboratory line will help minimize tissue contamination and help evaluate the nutritional impact of *Spiroplasma* on tsetse milk production.

We also noted that 234 male-biased genes were dramatically down-regulated in *Spiroplasma* infected females. These genes contributed a total of 1% of the transcriptome in the infected state, down from 17% in the uninfected state. Among the abundant and highly reduced genes were ones that encode multiple proteins annotated with digestive functions, such as those that had signatures of midgut trypsins (GFUI029963, GFUI029966, GFUI14886, GFUI024126, GFUI049688, GFUI006483, GFUI026212), serine proteases (GFUI026465, GFUI009869, GFUO026202), trypsin-like serine proteases (GFUI028730, GFUI028738), zinc-carboxypeptidases (GFUI030548, GFUI007477), chymotrypsin-like proteins (GFYI032994, GFUI032998, GFUI010855), and a carboxypeptidase (GFUI022995). The physiological consequences of this reduced proteolytic activity on *Gff*$^{Spi+}$ females remain to be determined.

Impact of Spiroplasma infection on female metabolism and reproductive fitness

We next investigated whether *Spiroplasma* induced differences in gene expression regulate critical metabolic processes relevant for fecundity. To do so we measured several reproductive fitness parameters in *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ pregnant females and males. We began by measuring the impact of *Spiroplasma* infection on the length of tsetse's gonotrophic cycle (GC), which includes oogenesis, embryogenesis and larvigenesis. We observed that the1st, 2nd, and 3rd GCs of *Gff*$^{Spi+}$ females (23.0 ± 0.24, 12.6 ± 0.19, and 13.0± 0.2 days, respectively) were significantly longer than that of their age-matched counterparts that did not harbor *Spiroplasma* (21.0 ± 0.36, 11.0 ± 0.29, and 11.2± 0.33 days, respectively) (GC1, p=0.0001; GC2, p<0.0001; GC3, p=0.0009; log rank test) (Fig

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

3A). We next weighed pupae-stage offspring from *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ females and observed no significant difference between groups across all three GCs (GC1, *Gff*$^{Spi+}$ = 26.0 ± 1.4 mg, *Gff*$^{Spi-}$ = 24.6 ± 1.1 mg; GC2, *Gff*$^{Spi+}$= 25.7 ± 1.0 mg, *Gff*$^{Spi-}$ = 27.4 ± 1.8 mg; GC3, *Gff*$^{Spi+}$ = 25.4 ± 1.7 mg, *Gff*$^{Spi-}$ = 26.4 ± 1.6 mg) (GC1, p=0.34; GC2, p=0.08; GC3, p=0.15; multiple t tests) (Fig 3B). These findings indicate that an infection with *Spiroplasma* does not impact pupal weight but does increase the length of tsetse's GC. As such, infection with this bacterium may result in the production of fewer offspring over the course of the female's lifespan, thus exerting a detrimental impact on the fly's overall reproductive fitness.

The protracted GC we observed in *Spiroplasma* infected tsetse suggests that these flies are nutritionally deficient, and this phenotype may reflect a competition for resources between the bacterium and pregnant female flies. This theory is supported by the fact that in *Drosophila*, *Spiroplasma* proliferation is limited by the availability of circulating lipids [20]. The gut of 3rd instar tsetse larvae contains high levels of triacylglyceride (TAG) [34], which originate in tsetse's fat body and are transferred through hemolymph to the milk gland where they are incorporated into maternal milk secretions [33,35]. To determine if *Spiroplasma* hijacks tsetse TAG, at the nutritional expense of developing intrauterine larvae, we compared TAG levels in the hemolymph of pregnant *Spi*$^{+}$ and *Spi*$^{-}$ females. We found that pregnant *Spi*$^{-}$ females had significantly more TAG circulating in their hemolymph (19.2 ± 0.4 μg/μl) than did their age and pregnancy stage-matched *Spi*$^{+}$ counterparts (11.5 ± 0.4 μg/μl) (p<0.0001; t test) (Fig 3C). These findings suggest that, like in *Drosophila*, *Spiroplasma* uses tsetse lipids as an energy source. This competition between tsetse and *Spiroplasma* for dietarily limited yet metabolically

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

337 important lipids may result in the relatively long GC we observed in *Spi*$^{+}$ moms because
338 less of this nutrient is available for incorporation into milk secretions.
339
340 ***Impact of Spiroplasma infection on Gff male reproductive fitness***
341 Our transcriptomic data informed us that infection with *Spiroplasma* impacts the
342 expression of several genes that encode spermatophore-associated proteins. These
343 proteins, many of which arise from the male accessory gland and are transferred in
344 seminal fluid, play a prominent role in sperm fitness [2,36]. We thus compared the fitness
345 of sperm derived from *Spi*$^{+}$ compared to *Spi*$^{-}$ males to determine if infection with the
346 bacterium impacts the fitness off *Gff* sperm. To do so we quantified the transcript
347 abundance of *sperm-specific dynein intermediate chain* (*sdic*; VectorBase gene ID
348 GFUI025244) in the spermathecae of 7 day old female flies two days after having mated
349 with either *Spi*$^{+}$ or *Spi*$^{-}$ males. The *Drosophila* homologue of this gene is expressed
350 exclusively in sperm cells [37] and has been used to quantify sperm abundance and
351 competitiveness (as a function of motility) [38-40]. We observed that *sdic* transcript
352 abundance in the spermathecae of females that mated with *Spi*$^{-}$ males was significantly
353 higher than in that of their counterparts that mated with *Spi*$^{+}$ males (Fig 4). This
354 conspicuous difference in the abundance of *sdic* transcripts expressed by sperm within
355 the spermathecae of females that mated with *Spi*$^{-}$ versus *Spi*$^{+}$ males was surprising
356 considering the fact that expression of this gene in the reproductive tract of field captured
357 *Spi*$^{-}$ and *Spi*$^{+}$ males was not significantly different when measured by RNAseq
358 (Supplemental Data1). To validate this RNAseq data, we used RT-qPCR to measure *sdic*
359 expression in sperm housed in the reproductive tract of 7 days old *Spi*$^{-}$ and *Spi*$^{+}$ lab reared
360 *Gff* males (these sperm were the same age as those used to measure *sdic* transcript

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

361 abundance in the spermathecae of mated females). Similar to what we observed via
362 RNAseq, *sdic* expression was not significantly different in sperm housed in the male
363 reproductive tract (Fig 4).

364 The above data indicate that *sdic* expression decreases significantly in sperm that
365 originate from *Spi*$^{+}$ males following transfer to the female spermatheca, thus suggesting
366 that the bacterium plays a role in either regulating the number of sperm transferred during
367 mating and/or sperm motility (and thus competitiveness) after mating. To investigate
368 further we quantified both the quantity and motility of sperm in the spermathecae of
369 females 24 hrs after having mated with either *Spi*$^{+}$ or *Spi*$^{-}$ males. Manual counting
370 indicated no significant difference in the number of stored sperm when females had mated
371 with either *Spi*$^{+}$ (5375 ± 537 sperm per spermathecae) or *Spi*$^{-}$ (5358 ± 632 sperm per
372 spermathecae) males (Fig 4B). We additionally quantified the abundance of stored sperm
373 as a reflection of spermathecal fill, which measures the quantity of sperm and seminal
374 fluid transferred from males to the spermatheca of their female mates. Male *Spiroplasma*
375 infection status had no impact on spermathecal fill. However, when female mates were
376 infected with *Spiroplasma* we observed a significant reduction in the size of their
377 spermatheca compared to that of *Spi*$^{-}$ females (Kruskal-Wallis, X^2 = 6.1763, df=1,
378 p=0.0129) (Fig 4C). We also investigated whether the precipitous drop in *sdic* expression
379 in sperm transferred from *Spi*$^{+}$ compared to *Spi*$^{-}$ males was representative of a decrease
380 in the motility of sperm derived from males in the latter group. We observed that sperm
381 derived from *Spi*$^{+}$ males exhibited a mean beat frequency of 13.4 (±1.7) hertz (Hz), while
382 those derived from *Spi*$^{-}$ males exhibited a mean beat frequency of 22.1 (±1.4) Hz (Fig
383 4D).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Taken together, these results indicate that the *Spiroplasma* infection status of male *Gff* does not impact the number of sperm they transfer during mating, but the female infection status of the mating pairs does impact the number of sperm stored in their spermatheca. Additionally, *Spiroplasma* exerts a significant impact on the competitiveness of sperm stored in the spermatophore, as cells that originate from infected males are significantly less motile than are those that originate from their uninfected counterparts.

Vertical transmission of Spiroplasma

Spiroplasma is maternally transmitted in several insect systems, including the tsetse fly, as evidenced by the presence of the endosymbiont in the intrauterine larva [14]. However, not all tsetse populations, nor individuals within distinct populations, house the bacterium [14,15]. This heterogeneity in infection prevalence suggests that vertical transmission of the bacterium may be imperfect and thus occur at a frequency of less than 100%. Three distinct experiments were performed as a means of quantitating the fidelity of *Spiroplasma* vertical transmission (see Materials and Methods for experimental details). For experiment 1, one random female and one random male were allowed to mate in a small mating tube, while for experiments 2 and 3, females were given a choice and an equal number of females and males (n=40 of each) were allowed to mate in large mating cages. Mating pairs from both experiments (2 and 3) were subsequently separated, and females from all three groups were then held in individual tubes over the course of three GCs. In the case of experiment 1, we determined the *Spiroplasma* infection prevalence of both moms and their corresponding adult offspring using our PCR assay. For experiment 1, we observed that 59% of moms (20/34) harbored *Spiroplasma* infections,

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

as well as did 95% (19/20), 95% (19/20), and 90% (18/20) of GC1, GC2 and GC3 offspring from infected moms, respectively. For experiments 2 and 3, the fidelity of *Spiroplasma* vertical transmission was determined using newly deposited pupal offspring. Under these conditions, in experiment 2, we determined that 58% of moms (15/26), and 73% (11/15) and 100% (9/9) of their pupal offspring from GCs 1 and 2, respectively, harbored the bacterium. No pregnant *Spi*$^{+}$ moms survived long enough to deposit GC3 offspring. Similarly, in experiment 3, we found that 71% (12/17) of moms within the mating cage housed the bacterium, as did 83% (10/12), 63% (5/8), 67% (2/3) of pupal offspring from GCs 1, 2, and 3 from these matings (Table 1; S4 data).

Table 1. Maternal transmission of *Spiroplasma*.

	Spiroplasma infection prevalence (# infected/total sample size)			
experiment #[a]	mom	GC1[b]	GC2	GC3
1	59% (20/34)	95% (19/20)	95% (19/20)	90% (18/20)
2	58% (15/26)	73% (11/15)	100% (9/9)	N/A[c]
3	71% (12/17)	83% (10/12)	63% (5/8)	67% (2/3)

[a]In experiment 1, *Spiroplasma* infection prevalence was determined using adult offspring, while in experiments 2 and 3 *Spiroplasma* infection prevalence was determined using pupal offspring.

[b](# infected offspring/infected mom for GC1, GC2, and GC3)

[c]No *Spi*$^{+}$ pregnant females survived to produced GC3 offspring.

In experiments 1 and 3, we also monitored the *Spiroplasma* infection status of *Gff* dads, as well as in the offspring that resulted from matings between *Spi*$^{-}$ moms and *Spi*$^{+}$ dads, in order to determine whether the bacterium can also be paternally transmitted. In experiment 1, nine matings between *Spi*$^{-}$ moms and *Spi*$^{+}$ dads produced 27 progeny over the three gonotrophic cycles. When tested as newly eclosed adults, none of the progeny

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

was found to carry *Spiroplasma* infections, indicating no evidence of paternal transmission (Table 2). In experiment 3, four matings between *Spi*$^{-}$ moms and *Spi*$^{+}$ dads collectively resulted in the deposition of nine pupae, four of which carried *Spiroplasma* infections (Table 2). Interestingly, in three out of four mating pairs, the second offspring (GC2) was positive for the *Spiroplasma* infection while first offspring (GC2) was negative (S4 Data), indicating imperfect transmission at best. The varying results we obtained between the first and third experiment could reflect the different developmental stages of the offspring we tested (experiment 1, newly eclosed adults; experiment 3, newly deposited pupae). *Spiroplasma* density decreases during development as tsetse age from larvae through pupation and eclosion to adulthood [14]. Hence, it remains to be seen whether the *Spiroplasma* present in the pupae analyzed in experiment 3 would persist through the lengthy pupal stage (~ 30 days) to the adulthood. Taken together, these data indicate that *Spiroplasma* is vertically transmitted via maternal and paternal *Gff* lineages, although maternal transmission occurs at a higher frequency than does paternal transmission.

Table 2. Paternal transmission of *Spiroplasma*.

	Spiroplasma infection prevalence (# infected/total sample size)			
experiment #[a]	dad[b]	GC1[c]	GC2	GC3
1	9	0% (0/9)	0% (0/9)	0% (0/9)
3	4	25% (1/4)	100% (3/3)	0% (0/2)

[a]In experiment 1, *Spiroplasma* infection prevalence was determined using adult offspring, while in experiment 3 *Spiroplasma* infection prevalence was determined using pupal offspring.

[b]The number of *Spi*+ dads that were mated with *Spi*- moms is presented.

[c]GC1, GC2, and GC3 offspring were derived from *Spi*- moms mated with *Spi*+ dads.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Discussion

Reproductive tissue-associated heritable endosymbionts affect their arthropod host's physiology to facilitate their transmission. In tsetse, both natural populations and laboratory *Gff* lines are found to house heterogenous infections with parasitic *Spiroplasma*, but little is known about the physiological impact(s) of this microbe on tsetse reproduction. Our transcriptomic analyses revealed that *Spiroplasma* infection significantly alters gene expression in reproductive tissues of both male and female *Gff*. In particular, amongst the genes impacted by *Spiroplasma* infection in females, significantly more female-biased genes are up-regulated, and male-biased genes are down-regulated when compared to their uninfected counterparts. Using a laboratory line of *Gff* that carries a heterogenous infection with *Spiroplasma*, we discovered that infection with the bacterium results in a reduction in fecundity as evidenced by a significantly longer gonotrophic cycle in infected females compared to their uninfected counterparts. Loss of fecundity likely results from the decreased levels of hemolymph lipids in *Spiroplasma* infected females, which suggests that the bacterium competes with its host during pregnancy for nutrients that are critical for larval development. Additionally, we observed a dramatic reduction in the abundance of a sperm-specific *sdic* transcripts following transfer of spermatozoa from *Spi*$^{+}$ males to their mate's spermatheca. Upon further investigation we determined that these sperm cells exhibit compromised motility and thus likely decrease competitiveness. Finally, our *Spiroplasma* transmission studies suggest maternal transmission of the bacterium occurs with high fidelity from infected mothers to each of their offspring. However, we also observed evidence of paternal transmission of the bacterium, which could explain the heterogenous infections we observed in this

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

462 laboratory line. Collectively, our findings significantly enhanced fundamental knowledge
463 on the tsetse-*Spiroplasma* symbiosis with implications for other arthropods. Additionally,
464 the information obtained in this study may be applicable to the development of novel
465 tsetse control strategies for population reduction.

466 Sex-biased gene expression strongly influences sexual dimorphisms in most
467 animals [27,28], and a large proportion of sex-biased genes are expressed in reproductive
468 tissues [41,42]. With this in mind, the differences in sex-biased gene expression we
469 observed in between *Gff*$^{Spi-}$ and *Gff*$^{Spi+}$ females suggest that housing the bacterium could
470 induce beneficial reproductive phenotypes. Our transcriptomic data in *Gff*$^{Spi+}$ females
471 showed that female-biased genes are up-regulated while male-biased genes are down-
472 regulated relative to their uninfected (*Gff*$^{Spi-}$) counterparts. However, it remains to be
473 determined whether these *Spiroplasma* induced changes in the sex-biased gene
474 expressions confer fitness benefits in *Gff*$^{Spi+}$ females. The decrease in circulating TAG
475 levels in *Gff*$^{Spi+}$ females suggests that *Spiroplasma* infection may confer a metabolic cost
476 that reduces female fecundity. This lipid and its catabolized byproducts represent
477 prominent components of tsetse milk [32,34], which serves as the sole food source for
478 developing intrauterine larva. Thus, depletion of this nutrient by hemolymph borne
479 *Spiroplasma* likely accounts for the increase we observed in GC length in pregnant moms
480 that harbor the bacterium. Despite the longer larval development period we observed for
481 progeny of *Spi*$^{+}$ compared to *Spi*$^{-}$ moms, the weight of pupal offspring deposited was
482 similar between the two groups. Because female tsetse produce unusually few offspring
483 (6-8) over the course of their lifespan (compared to other insects), an increase in GC
484 length would likely result in a significant reduction in population size over time. In fact,

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

485 infection with a trypanosome strain that induces a metabolically costly immune response
486 also lengthens tsetse's GC by a duration similar to that (approximately 2 days) induced
487 by infection with *Spiroplasma* [43]. Mathematical modelling indicates that this increase in
488 GC length would theoretically reduce tsetse fecundity by approximately 30% over the
489 course of a female's reproductive lifespan [43]. Thus, a moderate trypanosome infection
490 prevalence of 26%, which is similar to the *Spiroplasma* infection prevalence in field
491 captured *Gff* (5-34%, depending on population geographic location) [14,15], would thus
492 significantly decrease fly population size [43]. The models also predict that infection
493 prevalence above these frequencies would result in a population crash.

494 Our results indicate that infection with *Spiroplasma* also exerts a significant impact
495 on reproductive processes in laboratory reared *Gff* males. We observed that sperm from
496 both *Spi*$^{+}$ and *Spi*$^{-}$ males expressed the same abundance of sperm-specific *sdic*
497 transcripts prior to mating, but that following insemination and transfer to the female
498 spermatheca, sperm from *Spi*$^{+}$ males expressed conspicuously fewer transcripts of this
499 gene than did sperm that had originated from *Spi*$^{-}$ males. The *Drosophila* homolog of this
500 gene encodes a cytoplasmic dynein intermediate chain that is necessary for the proper
501 function of the cytoplasmic dynein motor protein complex [37], which is involved in sperm
502 motility [44]. Accordingly, one prominent phenotype we observed was that following
503 insemination and transfer to the female spermatheca, the motility of sperm that had
504 originated from *Spi*$^{+}$ males was impaired when compared to that of sperm that had
505 originated from *Spi*$^{-}$ males. This finding implies that *sdic* exhibits the same function in *Gff*
506 as it does in *Drosophila*. Motility is of paramount importance to sperm competitiveness
507 [45], and sperm with impaired motility have less fertilization success than heathy sperm

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

508 [46]. Field data indicate that wild populations of *Gff* females can exhibit polyandry and
509 store sperm from more than one mate [7]. Our data suggest that sperm transferred by
510 *Spi*$^{+}$ males may be at a competitive disadvantage in comparison to sperm derived from
511 *Spi*$^{-}$ males because they exhibit reduced motility. It remains to be determined if, as a
512 means of promoting polyandry, females seek additional mates if they first copulate with a
513 *Spi*$^{+}$ male that transfers motility compromised sperm.

514 *Spiroplasma* mediated mechanisms that influence the regulation of *sdic*
515 expression are currently unknown. [47]. *Spiroplasma* could mark sperm in the
516 reproductive tract of *Gff* males, possibly via post-transcriptional modification of *sdic,* such
517 that sperm motility is compromised following insemination and transfer to the female
518 spermatheca. Alternatively, female contributions to the spermatophore could also
519 influence *sdic* expression after sperm are packaged into the structure. In *Spiroplasma*
520 infected females, we have observed 12 DE genes that encode products identified as
521 constituents spermatophore. Further experiments are necessary to determine if any of
522 these *Spiroplasma* modified female products, including two transcription factors and three
523 serine protease inhibitors, play a role in the sperm modifications we observed here.

524 Infections with heritable endosymbionts can come at a significant metabolic cost
525 to their host [48]. This outcome appears to also be the case with the *Spiroplasma* and
526 tsetse symbiosis, and results in a significant reduction in the reproductive fitness of both
527 female and male flies. However, from an evolutionary perspective, these microbes,
528 including *Spiroplasma*, must also provide an overall fitness advantage in order for
529 infections to be maintained. In the tsetse model system, the fly, the commensal symbiont
530 *Sodalis*, as well as pathogenic trypanosomes, typically compete for host nutrients as they

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

531 are auxotrophic for metabolically critical B vitamins [12,49-51]. Interestingly, these
532 nutrients are present in low quantities in tsetse's vertebrate blood specific diet, and they
533 are instead supplemented via the mutualistic endosymbiont, *Wigglesworthia* [11-13].
534 Despite the negative impacts of *Spiroplasma* on tsetse's fecundity, this bacterium also
535 benefits its host by creating an environment within the fly that is hostile to parasitic
536 trypanosomes [15]. The parasite resistance phenotype conferred by *Spiroplasma*
537 infections could arise from induced host immune responses, or could reflect competition
538 for critical nutrients (i.e., found in the bloodmeal and/or produced by tsetse) that both
539 organisms require to sustain their metabolic needs [52-54]. This effect of inhibiting
540 metabolically costly trypanosome infections could thus offset the cost of housing similarly
541 detrimental *Spiroplasma* infections. Additionally, we demonstrate that while *Spiroplasma*
542 can be both matrilineally and patrilineally transmitted, vertical transmission of the
543 bacterium within our *Gff* colony does not occur at a frequency of 100%. This data
544 corroborates observations from the field, which reveal that individual flies between and
545 within distinct populations also present heterogeneous *Spiroplasma* infections [14,15]. In
546 our study, while we observed near perfect maternal transmission efficiency, paternal
547 transmission occurred infrequently. This outcome could contribute to the heterogenous
548 infection prevalence we observe in the *Gff* lab line. The absence of steadfast vertical
549 transmission of *Spiroplasma* within *Gff* populations may represent another advantageous
550 factor that facilitates the bacterium's ability to sustain infections within the fly across
551 multiple generations.

552 *Gff* is the prominent vector of pathogenic African trypanosomes in east and central
553 Africa, and reducing the size of fly populations is currently the most effective method for

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

554 controlling parasite transmission. One means of accomplishing this is through the use of
555 sterile insect technique (SIT), which involves the sequential release of a large number of
556 sterilized males into a target environment [55,56]. Because of their large number, these
557 infertile males outcompete wild males for resident female mates, and the population size
558 drops significantly, or the population is completely eliminated [57,58]. Knowledge gained
559 from this study, in conjunction with that from previous studies, will have a significant
560 impact on tsetse rearing efficiency for the SIT programs. Specifically, decreased fecundity
561 can have a negative impact on the success of colony maintenance. However, enhanced
562 resistance to trypanosomes is advantageous for SIT applications as the released males,
563 which also feed exclusively on vertebrate blood, have the potential to serve as vectors of
564 disease. Thus, *Spiroplasma* has the potential to reduce the impact of this shortcoming by
565 inducing a trypanosome refractory phenotype in released males such that they would be
566 less likely vector disease, thus increasing the overall safety and efficacy of the control
567 program.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Materials and Methods

Field sampling

Glossina fuscipes fuscipes (*Gff*) were collected from the Albert Nile river drainage in Northwest Uganda. Sampling sites included three villages from the Amuru district: Gorodona (GOR; 3°15'57.6"N, 32°12'28.8"E), Okidi (OKS; 3°15'36.0"N, 32°13'26.4"E), and Toloyang (TOL; 3°15'25.2"N, 32°13'08.4"E). Flies were sampled using biconical traps and wing fray data were recorded (stage 2 for the majority of samples, stage 3 for a small subset of samples, corresponding approximately to 3 and 4 week-old adults, respectively). All females analyzed had been mated based on the visual presence of intrauterine larva or a developing oocyte in their ovaries. The *Spiroplasma* infection prevalence in flies from the Goronda and Okidi regions was about 46% [Schneider et al 2019]. Reproductive tissues from all flies were dissected in the field in sterile 1x phosphate buffered saline (PBS) and flash frozen in liquid nitrogen for later RNA extraction.

Data generation (Illumina HiSeq), transcriptome assembly and quality assessment

Total RNA from *Gff* female (ovaries, spermathecae, uterus and oocysts) and male (testes, accessory glands and ejaculatory duct) reproductive tracts was extracted according to TRIzol® reagent manufacturer's instructions (Thermo Fisher Scientific, USA). Total RNA was treated with Ambion TURBO DNA-free DNase (Thermo Fisher Scientific, USA) and RNA quality was evaluated using an Agilent 2100 Bioanalyzer. An aliquot from each RNA was reverse transcribed into cDNA and screened for *Spiroplasma* using PCR amplification assay with symbiont-specific primers as described [15]. Prior to pooling RNA for sequencing (see below), the same cDNAs were used to confirm the sex of the field

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

collected tissues by performing PCR using primers that amplify sex specific genes (see S4 Table for gene names/IDs and amplification primers). RNA from three female or male reproductive tracks were pooled into distinct biological replicates for RNA-seq analysis. We analyzed two biological replicates from *Gff*$^{Spi+}$ females and three biological replicates from *Gff*$^{Spi+}$ males, and three biological replicates from *Gff*$^{Spi-}$ females and two biological replicates from *Gff*$^{Spi-}$ males. Ribosomal reduction libraries were prepared with the Ribo-Zero Gold rRNA removal kit (human/mouse/rat) MRZG12324 (Illumina, USA) and sequenced at the Yale University Center of Genome Analysis (YCGA) on an Illumina HiSeq 2500 machine (75bp paired-end reads). In order to obtain information on *Spiroplasma* transcripts as well as the tsetse host, we chose Ribosomal Reduction libraries over *m*RNA (poly A) libraries.

RNA-seq data analysis

RNA-seq sequencing produced an average of 96 million reads across all biological samples. Reads were mapped to the *Gff* reference genome (accession # JACGUE000000000 from NCBI BioProject PRJNA596165) using HISAT2 v2.1.0 with default parameters [59,60]. The annotations for the assembly were ported over from the annotation of the previous assembly (https://vectorbase.org/vectorbase/app/) using the UCSC liftOver software suite of programs [61]. The function 'htseq-count' in HTSeq v0.11.2 [62] was then used to count the number of reads aligned to the annotated genes in the reference genome with the option "-s reverse" because we generated strand-specific forward reads in RNA-seq libraries. We evaluated the number of DE genes between groups of female and male reproductive organs from *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ using the HTSeq output as input into DESeq2 v1.22.1 [63]. We first developed a DESeq2 model to

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

616 predict differential gene expression as a function of sex and *Spiroplasma*-infection status,
617 and the interaction between sex and *Spiroplasma*-infection status. We then extracted the
618 $\log_2$ fold changes ($\log_2$FCs) for each gene with false discovery rate (FDR) adjusted *p*-
619 values [24]. Finally, we defined genes as DE if the $\log_2$ fold change ($\log_2$FC) was
620 significantly different from 0, and we only considered genes with adjusted *p*-values as
621 expressed for downstream analysis.

622 We next conducted a PC and HC analysis to analyze normalized expression count
623 data from the DESeq2 program. We used a regularized log transformation of the
624 normalized data created by "rlog" function DESeq2 [63]. For the HC by sample-to-sample
625 distance, we used "pheatmap" function in R. We measured the expression abundance as
626 the sum of TPM that were calculated using the TPM Calculator [64].

627 Because the function of genes in *Gff are* not well annotated, we performed 'tblastx'
628 in CLC Genomics Workbench (CLC Bio, Cambridge, MA) with *Gff* transcript to search for
629 predicted protein sequences bearing the closest homology to those from *G. m. morsitans*,
630 *Musca domestica*, and *Drosophila melanogaster*. Using gene-associated GO terms [31],
631 we performed enrichment analysis of gene ontology (GO) terms with the topGO R
632 package [65].

633

634 *Laboratory maintenance of Gff and determination of Spiroplasma infection status*

635 *Gff* pupae were obtained from Joint FAO/IAEA IPCL insectary in Seibersdorf, Austria and
636 reared at Yale University insectary at 26°C with 70-80% relative humidity and a 12 hr
637 light:dark photo phase. Flies received defibrinated bovine blood every 48 hr through an
638 artificial membrane feeding system [66].

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

639 *Spiroplasma* infection status of all lab-reared female, male and pupal *Gff* was
640 determined by extracting genomic DNA (gDNA) from individual whole organisms or fly
641 legs using a DNeasy Blood and Tissue kit according to the manufacturer's (Qiagen)
642 protocol. To confirm that intact gDNA was successfully extracted, all samples were
643 subjected to PCR analysis using primers that specifically amplify *Gff tubulin* or
644 microsatellite region *GpCAG* (as controls for genomic DNA quality) and *Spiroplasma 16s*
645 *rRNA*. All PCR primers used in this study are listed in S4 Table. The *Spiroplasma 16s*
646 *rRNA* locus was amplified by PCR (as described in [14]) using the following parameters:
647 initial denaturation at 94°C for 5 min; 34 cycles of 94°C for 45 s, 59°C for 45 s, and 72°C
648 for 1 min; and a final extension at 72°C for 10 min. *Spiroplasma* PCR reactions were
649 carried out in a volume of 25 µl containing 1.5 µl gDNA. PCR products were analyzed on
650 2% agarose gels, and samples were considered infected with *Spiroplasma* if the expected
651 PCR product of 455 bp was detected.
652
653 *Impact of Spiroplasma infection on Gff metabolism and reproductive fitness*
654 *Hemolymph triacylglyceride (TAG) assay*: Hemolymph (3 µl/fly) was collected (as
655 described in [67]) from two-week-old pregnant female *Gff*, centrifuged (4°C, 3000xg for 5
656 minutes) to remove bacterial cells, diluted 1:10 in PBS containing 1.2 µl/ml of 0.2%
657 phenylthiourea (to prevent hemolymph coagulation) and immediately flash frozen in liquid
658 nitrogen. Hemolymph TAG levels were quantified colorimetrically by heating samples to
659 70°C for 5 min followed by a 10 min centrifugation 16,000xg. Five µl of supernatant was
660 added to 100 µl of Infinity Triglycerides Reagent (Thermo Scientific) and samples were
661 incubated at 37C for 10 min. Absorbance was measured at 540nm using a BioTek

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

662 Synergy HT plate reader [68]. All *Gff* sample spectra data were compared to that
663 generated from a triolein standard curve (0-50 μg, 10 μg increments; S2 Fig).
664
665 *Impact on fecundity*: The effect of *Spiroplasma* infection on *Gff* fecundity was measured
666 by quantifying the length of three gonotrophic cycles (GC) and by weighing pupal
667 offspring. To measure GC length, *Gff* females were mated as five days old adults and
668 thereafter maintained in individual cages. All females were monitored daily to determine
669 when they deposited larvae, and all deposited larvae were weighed.
670
671 *Sperm-specific gene expression*: Individual five-day old male and female flies (each fed
672 twice) were placed together into tubular cages (height, 12.7 cm; diameter 6 cm) and
673 allowed to mate for two days. Individual mating pairs were subsequently separated, and
674 RNA (using Trizol reagent) was isolated from male and female reproductive tracts. RNA
675 was then DNase treated and reverse transcribed into cDNA (using a Bio-Rad iScript
676 cDNA synthesis kit) by priming the reaction with random hexamers. All males were then
677 screened by PCR as described above to determine their *Spiroplasma* infection status.
678 *Gff sperm-specific dynein intermediate chain* (*sdic*; GFUI025244) transcript
679 abundance (measured via RT-qPCR, primers used are listed in S4 Table) was used as a
680 proxy measurement of sperm density and sperm fitness in male reproductive tracts and
681 the spermathecae of pregnant females [37,44]. All RT-qPCR results were normalized to
682 tsetse's constitutively expressed *pgrp-la* gene, and relative expression of *sdic* was
683 compared between male reproductive tracts and sperm (in the female spermathecae) that
684 originated from Spi^+ and Spi^- males. All RT-qPCR assays were carried out in duplicate,

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and replicates quantities are indicated as data points on corresponding figures. Negative controls were included in all amplification reactions.

Sperm quantification and spermathecal fill assays: Sperm abundance in the spermathecae of females that had mated with either *Spi*$^{+}$ and *Spi*$^{-}$ males was quantified by both direct cell counting and by measuring spermathecal fill. For direct counting, spermathecae were excised from mated females 24 hrs post-copulation with either *Spi*$^{+}$ or *Spi*$^{-}$ males and placed into 10 μl of HEPES-buffered saline solution [145 mM NaCl, 4 mM KCl, 1 mM MgCl2, 1.3 mM CaCl2, 5 mM D-glucose, 10 mM 4-(2-hydroxyethyl)-1-piperazineethane- sulfonic acid (HEPES), pH 7.4] in the well of a concave glass microscope slide. Spermathecae were gently poked with a fine needle and sperm were allowed to exude from the organ for 5 minutes. Samples were subsequently diluted 1:100 in HEPES-buffered saline and counted using a Neubauer counting chamber. To measure spermathecal fill, spermathecae from mated females were microscopically dissected 24h post-copulation in physiological saline solution (0.9% NaCl) and assessed subjectively at 100x magnification. Spermathecal fill of each individual organ was scored to the nearest quarter as empty (0), partially full (0.25, 0.50, or 0.75), or full (1.0) (S3 Fig), and the amount of sperm transferred was then computed as the mean spermathecal filling values of the spermathecae pairs [69,70].

Sperm motility assays: Spermathecae were excised from females 24 hrs post-copulation with either *Spi*$^{+}$ and *Spi*$^{-}$ males and placed into 10 μl of HEPES-buffered saline solution [145 mM NaCl, 4 mM KCl, 1 mM MgCl2, 1.3 mM CaCl2, 5 mM D-glucose, 10 mM 4-(2-

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

hydroxyethyl)-1-piperazineethane- sulfonic acid (HEPES), pH 7.4] in the well of a concave glass microscope slide. Spermathecae were gently poked with a fine needle and sperm were allowed to exude from the organ for 5 minutes. Sperm beating was recorded using an inverted microscope (10x phase contrast, Zeiss Primovert) that housed a charge-coupled camera (Zeiss Axiocam ERc 5s). Two sperm tails per sample were analyzed. Recordings were acquired at a rate of 30 frames per second, and beat frequency was analyzed using FIJI and the ImageJ plugin SpermQ [71].

Maternal transmission efficiency: Three separate experiments, performed out by different researchers in different insectaries, were implemented to monitor the dynamics of *Spiroplasma* transmission.

Experiment 1: Individual five-day old male and female flies (each fed twice) were placed together into small tubular cages (height, 12.7 cm; diameter 6.0 cm) and allowed to mate for two days. Males were subsequently removed from the cages (and frozen for future analysis) and pregnant females were fed every other day throughout the course of three GCs. Pupae were collected from each female and housed separately until adult emergence. Genomic DNA from legs of parent flies and all offspring was purified and subjected to PCR analysis to determine *Spiroplasma* infection status, as indicated above in subsection '*Laboratory maintenance of Gff and determination of Spiroplasma infection status*'.

Experiment 2: Forty 7-9 day old male *Gff* were released into large mating cages (45x45x45 cm), and 15 minutes later an equal number of 3-5 day old females were introduced. Mating was observed under standard rearing conditions. As soon as

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

731 copulation began, mating pairs were collected a placed into individual cages (4 cm
732 diameter x 6 cm height) where they were kept together for 24 h. Males were subsequently
733 removed (and conserved in absolute ethanol at -20°C for further analysis), and mated
734 females were pooled together in a larger cage and maintained under normal rearing
735 conditions for 10 days. Finally, pregnant females were again separated into individual
736 cages and maintained under normal conditions over the course of three GCs (~40 days).
737 Individual pupa were collected from each female following each GC and conserved in
738 ethanol at -20°c for further analysis 24h and 72 h post deposition.
739 Experiment 3: This experiment was performed the same in the same manner as was
740 experiment 2, with the exception that, after the removal of males, females were
741 maintained individually over the course of three GCs so as to later be able to determine
742 the *Spiroplasma* infection status of their specific mate.
743
744 *Statistical analyses*
745 All statistical analyses were carried out using GraphPad Prism (v.9), Microsoft Excel or
746 RStudio (v.1.2.5033 and v.1.3.1073). All statistical tests used, and statistical significance
747 between treatments, and treatments and controls, are indicated on the figures or in their
748 corresponding legends. All sample sizes are provided in corresponding figure legends or
749 are indicated graphically as points on dot plots. Biological replication implies distinct
750 groups of flies were collected on different days.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgements

This work was generously supported with funding to Serap Aksoy from the Li Foundation, Ambrose Monell Foundation, and NIH/NIAID (RO1AI068932). We thank Dr. Güler Demirbas-Uzel (Insect Pest Control Laboratory, Joint FAO/IAEA Programme of Nuclear Techniques in Food & Agriculture, Vienna, Austria) for technical assistance with spermathecal fill assays. We are thankful for support from the United Nations, International Atomic Energy Association sponsored Coordinated Research Project entitled 'Improvement of Colony Management in Insect Mass-Rearing for SIT Applications'. We thank Dr. Peter Takac, at the Slovakian Academy of Sciences, for supplying tsetse pupae.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1. Benoit JB, Attardo GM, Baumann AA, Michalkova V, Aksoy S: **Adenotrophic viviparity in tsetse flies: potential for population control and as an insect model for lactation**. *Annu Rev Entomol* 2015, **60**:351-371.
2. Scolari F, Benoit JB, Michalkova V, Aksoy E, Takac P, Abd-Alla AM, Malacrida AR, Aksoy S, Attardo GM: **The Spermatophore in Glossina morsitans morsitans: Insights into Male Contributions to Reproduction**. *Sci Rep* 2016, **6**:20334.
3. Pollock JN: **Sperm transfer by spermatophores in Glossina austeni Newstead**. *Nature* 1970, **225**:1063-1064.
4. Saunders DS, Dodd, C.W.H.: **Mating, insemination, and ovulation in the tsetse fly, Glossina morsitans**. *J. Insect Physiol.* 1972, **18**:187-198.
5. Odhiambo T, Kokwaro, E., Sequeira, L. : **Histochemical and ultrastructural studies of the male accessory reproductive glands and spermatophore of the tsetse, Glossina morsitans morsitans Westwood**. *Insect Science and Its Application* 1983, **4**:227-236.
6. Attardo GM, Tam N, Parkinson D, Mack LK, Zahnle XJ, Arguellez J, Takac P, Malacrida AR: **Interpreting Morphological Adaptations Associated with Viviparity in the Tsetse Fly Glossina morsitans (Westwood) by Three-Dimensional Analysis**. *Insects* 2020, **11**.
7. Bonomi A, Bassetti F, Gabrieli P, Beadell J, Falchetto M, Scolari F, Gomulski LM, Regazzini E, Ouma JO, Caccone A, et al.: **Polyandry is a common event in wild populations of the Tsetse fly Glossina fuscipes fuscipes and may impact population reduction measures**. *PLoS Negl Trop Dis* 2011, **5**:e1190.
8. Vreysen MJ, Saleh K, Mramba F, Parker A, Feldmann U, Dyck VA, Msangi A, Bouyer J: **Sterile insects to enhance agricultural development: the case of sustainable tsetse eradication on Unguja Island, Zanzibar, using an area-wide integrated pest management approach**. *PLoS Negl Trop Dis* 2014, **8**:e2857.
9. Kabayo JP: **Aiming to eliminate tsetse from Africa**. *Trends Parasitol* 2002, **18**:473-475.
10. Bing X, Attardo GM, Vigneron A, Aksoy E, Scolari F, Malacrida A, Weiss BL, Aksoy S: **Unravelling the relationship between the tsetse fly and its obligate symbiont Wigglesworthia: transcriptomic and metabolomic landscapes reveal highly integrated physiological networks**. *Proc Biol Sci* 2017, **284**.
11. Akman L, Yamashita A, Watanabe H, Oshima K, Shiba T, Hattori M, Aksoy S: **Genome sequence of the endocellular obligate symbiont of tsetse flies, Wigglesworthia glossinidia**. *Nat Genet* 2002, **32**:402-407.
12. Snyder AK, Rio RV: **"Wigglesworthia morsitans" Folate (Vitamin B9) Biosynthesis Contributes to Tsetse Host Fitness**. *Appl Environ Microbiol* 2015, **81**:5375-5386.
13. Rio RV, Symula RE, Wang J, Lohs C, Wu YN, Snyder AK, Bjornson RD, Oshima K, Biehl BS, Perna NT, et al.: **Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: glossinidae) obligate symbiont Wigglesworthia**. *mBio* 2012, **3**.
14. Doudoumis V, Blow F, Saridaki A, Augustinos A, Dyer NA, Goodhead I, Solano P, Rayaisse JB, Takac P, Mekonnen S, et al.: **Challenging the Wigglesworthia, Sodalis, Wolbachia symbiosis dogma in tsetse flies: Spiroplasma is present in both laboratory and natural populations**. *Sci Rep* 2017, **7**:4699.
15. Schneider DI, Saarman N, Onyango MG, Hyseni C, Opiro R, Echodu R, O'Neill M, Bloch D, Vigneron A, Johnson TJ, et al.: **Spatio-temporal distribution of Spiroplasma infections in the tsetse fly (Glossina fuscipes fuscipes) in northern Uganda**. *PLoS Negl Trop Dis* 2019, **13**:e0007340.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

16. Lukasik P, Guo H, van Asch M, Ferrari J, Godfray HC: **Protection against a fungal pathogen conferred by the aphid facultative endosymbionts Rickettsia and Spiroplasma is expressed in multiple host genotypes and species and is not influenced by co-infection with another symbiont**. *J Evol Biol* 2013, **26**:2654-2661.
17. Hamilton PT, Peng F, Boulanger MJ, Perlman SJ: **A ribosome-inactivating protein in a Drosophila defensive symbiont**. *Proc Natl Acad Sci U S A* 2016, **113**:350-355.
18. Ballinger MJ, Perlman SJ: **The defensive Spiroplasma**. *Curr Opin Insect Sci* 2019, **32**:36-41.
19. Perlmutter JI, Bordenstein SR: **Microorganisms in the reproductive tissues of arthropods**. *Nat Rev Microbiol* 2020, **18**:97-111.
20. Herren JK, Paredes JC, Schupfer F, Arafah K, Bulet P, Lemaitre B: **Insect endosymbiont proliferation is limited by lipid availability**. *Elife* 2014, **3**:e02964.
21. Hurst GD, Bandi C, Sacchi L, Cochrane AG, Bertrand D, Karaca I, Majerus ME: **Adonia variegata (Coleoptera: Coccinellidae) bears maternally inherited flavobacteria that kill males only**. *Parasitology* 1999, **118 (Pt 2)**:125-134.
22. Harumoto T, Lemaitre B: **Male-killing toxin in a bacterial symbiont of Drosophila**. *Nature* 2018, **557**:252-255.
23. Harumoto T, Fukatsu T, Lemaitre B: **Common and unique strategies of male killing evolved in two distinct Drosophila symbionts**. *Proc Biol Sci* 2018, **285**.
24. Benjamini Y, Hochberg, Y.: **Controlling the false discovery rate: a practical and powerful approach to multiple testing**. *Journal of the Royal Statistical Society, Series B* 1995, **57**:289-300.
25. Engelstadter J, Hurst GDD: **The Ecology and Evolution of Microbes that Manipulate Host Reproduction**. *Annual Review of Ecology Evolution and Systematics* 2009, **40**:127-149.
26. Harumoto T, Anbutsu H, Lemaitre B, Fukatsu T: **Male-killing symbiont damages host's dosage-compensated sex chromosome to induce embryonic apoptosis**. *Nat Commun* 2016, **7**:12781.
27. Ellegren H, Parsch J: **The evolution of sex-biased genes and sex-biased gene expression**. *Nat Rev Genet* 2007, **8**:689-698.
28. Mank JE: **The transcriptional architecture of phenotypic dimorphism**. *Nat Ecol Evol* 2017, **1**:6.
29. Herren JK, Lemaitre B: **Spiroplasma and host immunity: activation of humoral immune responses increases endosymbiont load and susceptibility to certain Gram-negative bacterial pathogens in Drosophila melanogaster**. *Cell Microbiol* 2011, **13**:1385-1396.
30. Dias RO, Cardoso C, Pimentel AC, Damasceno TF, Ferreira C, Terra WR: **The roles of mucus-forming mucins, peritrophins and peritrophins with mucin domains in the insect midgut**. *Insect Mol Biol* 2018, **27**:46-60.
31. Attardo GM, Abd-Alla AMM, Acosta-Serrano A, Allen JE, Bateta R, Benoit JB, Bourtzis K, Caers J, Caljon G, Christensen MB, et al.: **Comparative genomic analysis of six Glossina genomes, vectors of African trypanosomes**. *Genome Biol* 2019, **20**:187.
32. Benoit JB, Attardo GM, Michalkova V, Krause TB, Bohova J, Zhang Q, Baumann AA, Mireji PO, Takac P, Denlinger DL, et al.: **A novel highly divergent protein family identified from a viviparous insect by RNA-seq analysis: a potential target for tsetse fly-specific abortifacients**. *PLoS Genet* 2014, **10**:e1003874.
33. Benoit JB, Yang G, Krause TB, Patrick KR, Aksoy S, Attardo GM: **Lipophorin acts as a shuttle of lipids to the milk gland during tsetse fly pregnancy**. *J Insect Physiol* 2011, **57**:1553-1561.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

34. Cmelik SH, Hurrell DP, Lunat M: **Lipid content and composition of the tse-tse fly, Glossina morsitans Westwood**. *Comp Biochem Physiol* 1969, **31**:65-78.
35. Attardo GM, Benoit JB, Michalkova V, Yang G, Roller L, Bohova J, Takac P, Aksoy S: **Analysis of lipolysis underlying lactation in the tsetse fly, Glossina morsitans**. *Insect Biochem Mol Biol* 2012, **42**:360-370.
36. Sepil I, Hopkins BR, Dean R, Bath E, Friedman S, Swanson B, Ostridge HJ, Harper L, Buehner NA, Wolfner MF, et al.: **Male reproductive aging arises via multifaceted mating-dependent sperm and seminal proteome declines, but is postponable in Drosophila**. *Proc Natl Acad Sci U S A* 2020, **117**:17094-17103.
37. Nurminsky DI, Nurminskaya MV, De Aguiar D, Hartl DL: **Selective sweep of a newly evolved sperm-specific gene in Drosophila**. *Nature* 1998, **396**:572-575.
38. Yeh SD, Do T, Chan C, Cordova A, Carranza F, Yamamoto EA, Abbassi M, Gandasetiawan KA, Librado P, Damia E, et al.: **Functional evidence that a recently evolved Drosophila sperm-specific gene boosts sperm competition**. *Proc Natl Acad Sci U S A* 2012, **109**:2043-2048.
39. Weiss BL, Maltz MA, Vigneron A, Wu Y, Walter KS, O'Neill MB, Wang J, Aksoy S: **Colonization of the tsetse fly midgut with commensal Kosakonia cowanii Zambiae inhibits trypanosome infection establishment**. *PLoS Pathog* 2019, **15**:e1007470.
40. Jayaswal V, Jimenez J, Magie R, Nguyen K, Clifton B, Yeh S, Ranz JM: **A species-specific multigene family mediates differential sperm displacement in Drosophila melanogaster**. *Evolution* 2018, **72**:399-403.
41. Ingleby FC, Flis I, Morrow EH: **Sex-biased gene expression and sexual conflict throughout development**. *Cold Spring Harb Perspect Biol* 2014, **7**:a017632.
42. Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, Eastman S, Oliver B: **Paucity of genes on the Drosophila X chromosome showing male-biased expression**. *Science* 2003, **299**:697-700.
43. Hu C, Rio RV, Medlock J, Haines LR, Nayduch D, Savage AF, Guz N, Attardo GM, Pearson TW, Galvani AP, et al.: **Infections with immunogenic trypanosomes reduce tsetse reproductive fitness: potential impact of different parasite strains on vector population structure**. *PLoS Negl Trop Dis* 2008, **2**:e192.
44. Yeh SD, Do T, Abbassi M, Ranz JM: **Functional relevance of the newly evolved sperm dynein intermediate chain multigene family in Drosophila melanogaster males**. *Commun Integr Biol* 2012, **5**:462-465.
45. Civetta A, Ranz JM: **Genetic Factors Influencing Sperm Competition**. *Front Genet* 2019, **10**:820.
46. Singh SR, Singh BN, Hoenigsberg HF: **Female remating, sperm competition and sexual selection in Drosophila**. *Genet Mol Res* 2002, **1**:178-215.
47. Shropshire JD, Leigh B, Bordenstein SR: **Symbiont-mediated cytoplasmic incompatibility: what have we learned in 50 years?** *Elife* 2020, **9**.
48. Mathe-Hubert H, Kaech H, Ganesanandamoorthy P, Vorburger C: **Evolutionary costs and benefits of infection with diverse strains of Spiroplasma in pea aphids**. *Evolution* 2019, **73**:1466-1481.
49. Rio RVM, Jozwick AKS, Savage AF, Sabet A, Vigneron A, Wu Y, Aksoy S, Weiss BL: **Mutualist-Provisioned Resources Impact Vector Competency**. *mBio* 2019, **10**.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

50. Snyder AK, Deberry JW, Runyen-Janecky L, Rio RV: **Nutrient provisioning facilitates homeostasis between tsetse fly (Diptera: Glossinidae) symbionts**. *Proc Biol Sci* 2010, **277**:2389-2397.
51. Snyder AK, Rio RV: **Interwoven biology of the tsetse holobiont**. *J Bacteriol* 2013, **195**:4322-4330.
52. McCutcheon JP, Boyd BM, Dale C: **The Life of an Insect Endosymbiont from the Cradle to the Grave**. *Curr Biol* 2019, **29**:R485-R495.
53. Paredes JC, Herren JK, Schupfer F, Lemaitre B: **The Role of Lipid Competition for Endosymbiont-Mediated Protection against Parasitoid Wasps in Drosophila**. *mBio* 2016, **7**.
54. Sudakaran S, Kost C, Kaltenpoth M: **Symbiont Acquisition and Replacement as a Source of Ecological Innovation**. *Trends Microbiol* 2017, **25**:375-390.
55. Percoma L, Sow A, Pagabeleguem S, Dicko AH, Serdebeogo O, Ouedraogo M, Rayaisse JB, Bouyer J, Belem AMG, Sidibe I: **Impact of an integrated control campaign on tsetse populations in Burkina Faso**. *Parasit Vectors* 2018, **11**:270.
56. Vreysen MJ, Seck MT, Sall B, Bouyer J: **Tsetse flies: their biology and control using area-wide integrated pest management approaches**. *J Invertebr Pathol* 2013, **112 Suppl**:S15-25.
57. Lees RS, Gilles JR, Hendrichs J, Vreysen MJ, Bourtzis K: **Back to the future: the sterile insect technique against mosquito disease vectors**. *Curr Opin Insect Sci* 2015, **10**:156-162.
58. McGraw EA, O'Neill SL: **Beyond insecticides: new thinking on an ancient problem**. *Nat Rev Microbiol* 2013, **11**:181-193.
59. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements**. *Nat Methods* 2015, **12**:357-360.
60. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL: **Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype**. *Nat Biotechnol* 2019, **37**:907-915.
61. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, et al.: **The UCSC Genome Browser Database: update 2006**. *Nucleic Acids Res* 2006, **34**:D590-598.
62. Anders S, Pyl PT, Huber W: **HTSeq--a Python framework to work with high-throughput sequencing data**. *Bioinformatics* 2015, **31**:166-169.
63. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. *Genome Biol* 2014, **15**:550.
64. Vera Alvarez R, Pongor LS, Marino-Ramirez L, Landsman D: **TPMCalculator: one-step software to quantify mRNA abundance of genomic features**. *Bioinformatics* 2019, **35**:1960-1962.
65. Alexa A, Rahnenfuhrer J.: **topGO: Enrichrichment analysis for gene ontology. R package version 2.42.0.** Edited by; 2020.
66. Moloo SK: **An artificial feeding technique for Glossina**. *Parasitology* 1971, **63**:507-512.
67. Benoit JB, Vigneron A, Broderick NA, Wu Y, Sun JS, Carlson JR, Aksoy S, Weiss BL: **Symbiont-induced odorant binding proteins mediate insect host hematopoiesis**. *Elife* 2017, **6**.
68. Agbu P, Cassidy JJ, Braverman J, Jacobson A, Carthew RW: **MicroRNA miR-7 Regulates Secretion of Insulin-Like Peptides**. *Endocrinology* 2020, **161**.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

943 69. Mutika GN, Kabore I, Parker AG, Vreysen MJ: **Storage of male Glossina palpalis**
944 **gambiensis pupae at low temperature: effect on emergence, mating and survival**.
945 *Parasit Vectors* 2014, **7**:465.
946 70. Mutika GN, Marin C, Parker AG, Boucias DG, Vreysen MJ, Abd-Alla AM: **Impact of**
947 **salivary gland hypertrophy virus infection on the mating success of male Glossina**
948 **pallidipes: consequences for the sterile insect technique**. *PLoS One* 2012, **7**:e42188.
949 71. Hansen JN, Rassmann S, Jikeli JF, Wachten D: **SpermQ(-)A Simple Analysis Software to**
950 **Comprehensively Study Flagellar Beating and Sperm Steering**. *Cells* 2018, **8**.
951 72. Mangiafico S: **Rcompanion: Functions to Support Extension Education Program**
952 **Evaluation**. **R package version 2.3.26.** 2020.
953 73. Ogle DH, Wheeler, P, Dinno, A: **FSA: Fisheries Stock Analysis. R package version 0.8.32**.
954 2021.
955

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

957 **Figure legends**
958
959 **Fig 1.** Principal component analysis (A) and Hierarchical clustering (B) of expression
960 data. The Principal Component Analysis (PCA) is based on differentially expressed
961 genes from the female and male samples, while the Hierarchical Clustering (HC)
962 reflects all genes within the dataset. *Spiroplasma*-infection status and sex are color-
963 coded in the PCA.
964
965 **Fig 2.** Enrichment of genes with sex-biased expression that are differentially expressed
966 between *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ flies. (A) The number of genes with sex-biased expression that
967 are up- and down-regulated in *Gff*$^{Spi+}$ males. (B) The number of genes with sex-biased
968 expression profiles that are up- and down-regulated in *Gff*$^{Spi+}$ females. (C) Transcript
969 abundance (%) of up- and down-regulated genes relative to total transcript abundance in
970 *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ males. (D) Transcript abundance of up- and down-regulated genes
971 relative to total transcript abundance in *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ females. Female-biased genes
972 are represented as red blue and male-biased genes as blue. Genes with no sex-bias are
973 indicated as gray.
974
975 **Fig 3**. Impact of *Spiroplasma* infection on the reproductive and nutritional fitness of female
976 tsetse flies. (A) Gonotrophic cycle (GC) length of offspring from *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ females.
977 Age-matched, pregnant females from each group (n = 34 per group) were housed in
978 individuals cages and monitored daily to observe frequency of pupal deposition. Statistical
979 significance was determined via log-rank test. (B) The weight of pupae deposited by
980 *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ females. Each dot represents an individual pupa, bars represent median

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

values of pupae form each GC. Statistical significance was determined via multiple t-tests. (C) Amount of triacylglyceride (TAG) circulating in the hemolymph of *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ females. Three microliters of hemolymph was extracted from two-week-old pregnant females and TAG was quantified colorimetrically via comparison to triolein standard curve. Each dot represents an individual pupa, bars represent median values of pupae form each GC. Statistical significance was determined via unpaired t-test.

Fig 4. The impact of *Spiroplasma* infection on male *Gff* reproductive fitness. (A) Relative expression of *sperm-specific dynein intermediate chain* (*sdic*) in sperm located within the reproductive tract of *Gff*$^{Spi+}$ and *Gff*$^{Spi-}$ males or within the female spermatheca following copulation. *Sdic* expression in each sample was normalized relative to tsetse's constitutively expressed *pgrp-la* gene. Each dot represents one biological replicate, and bars indicate median values. Statistical significance was determined via one-way ANOVA followed by Tukey's HSD post-hoc analysis. (B) Quantification of sperm within the spermatheca of female flies that mated with either *Spi*$^{+}$ or *Spi*$^{-}$ males. Measurements were made using a Neubauer counting chamber. Each dot represents one spermatheca, and bars indicate median values. Statistical significance was determined via student's t-test. (C) Spermathecal fill of female flies that mated with either *Spi*$^{+}$ or *Spi*$^{-}$ males. Spermatheca fill data were analyzed using Kruskal-Wallis rank test using R companion and FSA software packages [72,73]. (D). Motility, as a measure of flagellar beat frequency in hertz (Hz), of sperm within the spermatheca of female flies that mated with either *Spi*$^{+}$ or *Spi*$^{-}$ males. Video recordings of sperm were acquired at a rate of 50 frames per second, and beat frequency was analyzed using FIJI and the

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ImageJ plugin SpermQ. Each dot on the graph represents the mean beat frequency of two sperm tails from each spermatheca. Statistical significance was determined via student's t-test.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supporting Information

S1 Appendix. Supporting tables.

S1 Fig. Distribution of differentially expressed transcript products in functional classes analyzed by Gene Ontology (GO) enrichment analysis. The bar diagrams show the significantly enriched GO terms among **(A)** the up-regulated and **(B)** the down-regulated genes in *Gff*$^{Spi+}$ males, and among **(C)** the up-regulated and **(D)** the down-regulated genes in *Gff*$^{Spi+}$ females. The number of genes associated with the corresponding GO terms to the number of genes belonging to that GO term within the entire set of genes in the genome is shown for each bar. The colors associated with the different bars denote the two different GO categories; BP: Biological Process, MF: Molecular Function.

S2 Fig. Triolein standard curve used to determine the concentration of triacylglyceride circulating in the hemolymph of pregnant *Gff*$^{Spi+}$ compared to *Gff*$^{Spi-}$ females. 0-50 μg aliquots of triolein were mixed with 100 μl of Infinity Triglycerides Reagent (Thermo Scientific) and samples were incubated at 37C for 10 min. Absorbance was measured at 540nm using a BioTek Synergy HT plate reader.

S3 Fig. Diagrammatic guide used to visually quantify spermathecal fill in females 24 h post-copulation. Image generated by Dr. Güler Demirbas-Uzel.

S1 Data. Gene expression data created in DESeq2 analysis, homologies in other Diptera of DE genes, and DE genes that constitute spermatophore proteins.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

S2 Data. Transcripts Per Million (TPM) for each sample.

S3 Data. GO enrichment analysis results.

S4 Data. *Spiroplasma* vertical transmission data

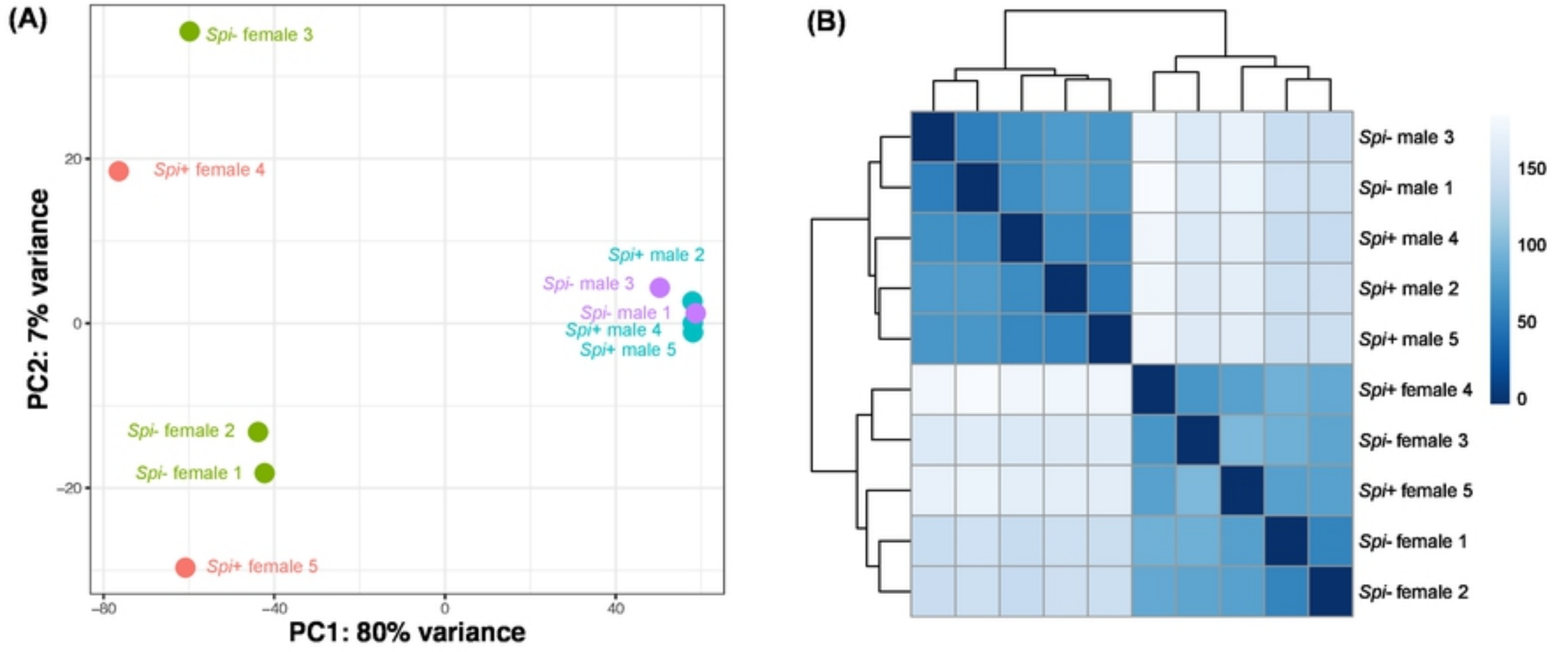

(A)
Spi- female 3
Spi+ female 4
Spi+ male 2
Spi- male 3
Spi- male 1
Spi+ male 4
Spi+ male 5
Spi- female 2
Spi- female 1
Spi+ female 5
PC2: 7% variance
PC1: 80% variance
20
0
−20
−80
−40
0
40
(B)
Spi- male 3
Spi- male 1
Spi+ male 4
Spi+ male 2
Spi+ male 5
Spi+ female 4
Spi- female 3
Spi+ female 5
Spi- female 1
Spi- female 2
150
100
50
0

Fig_1

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

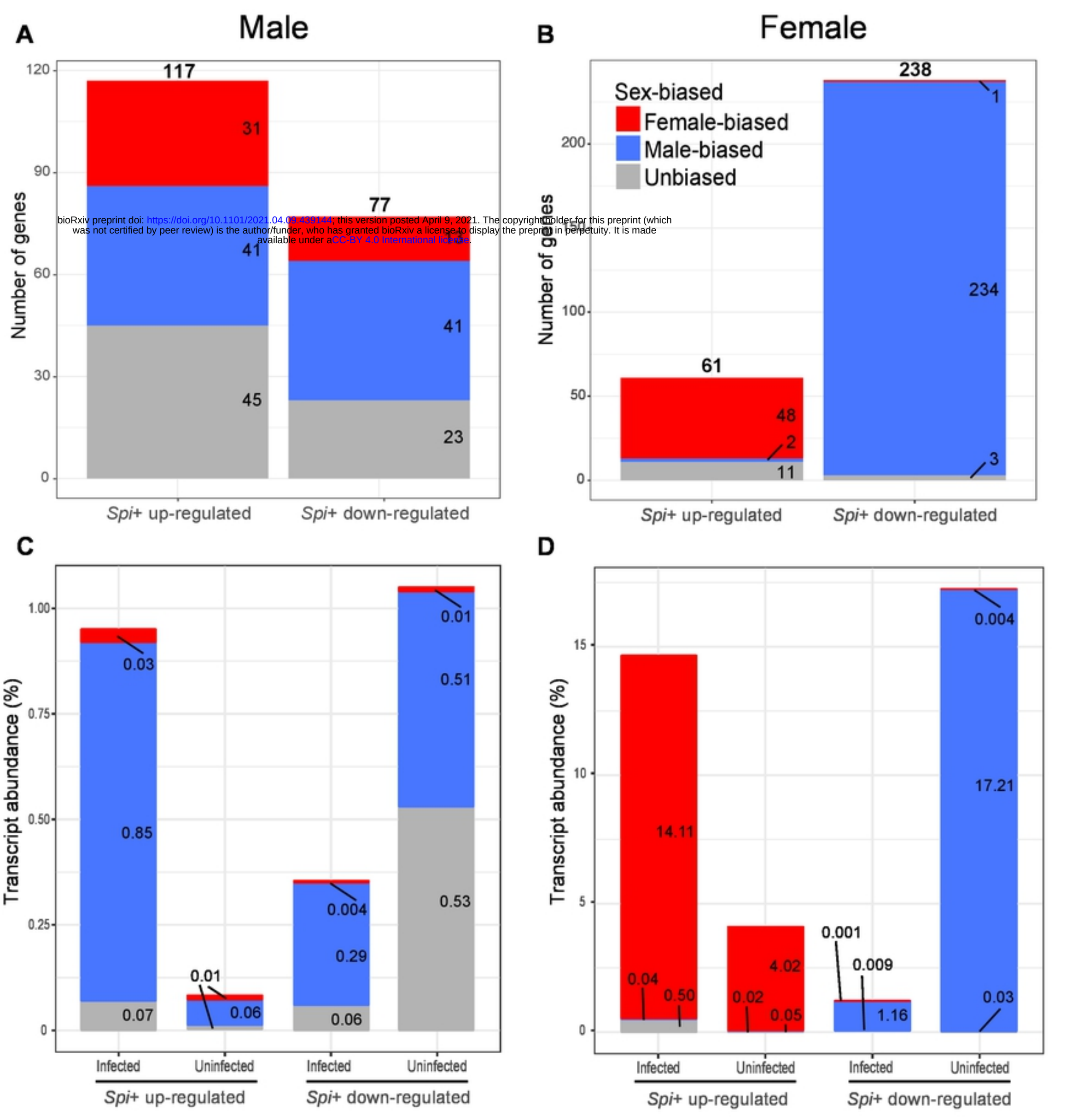


Fig_2

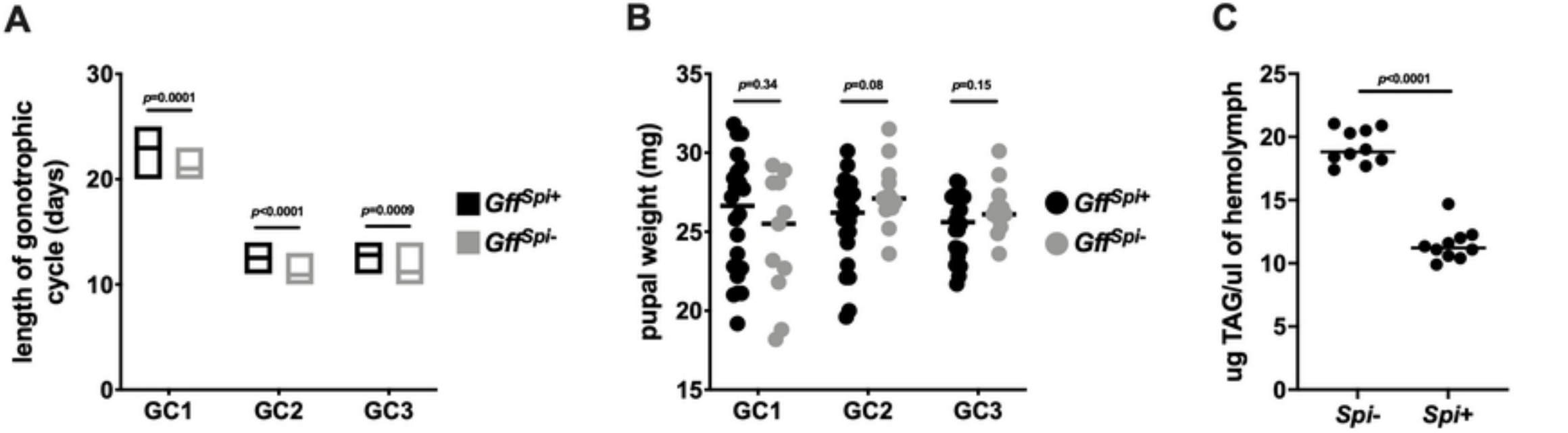
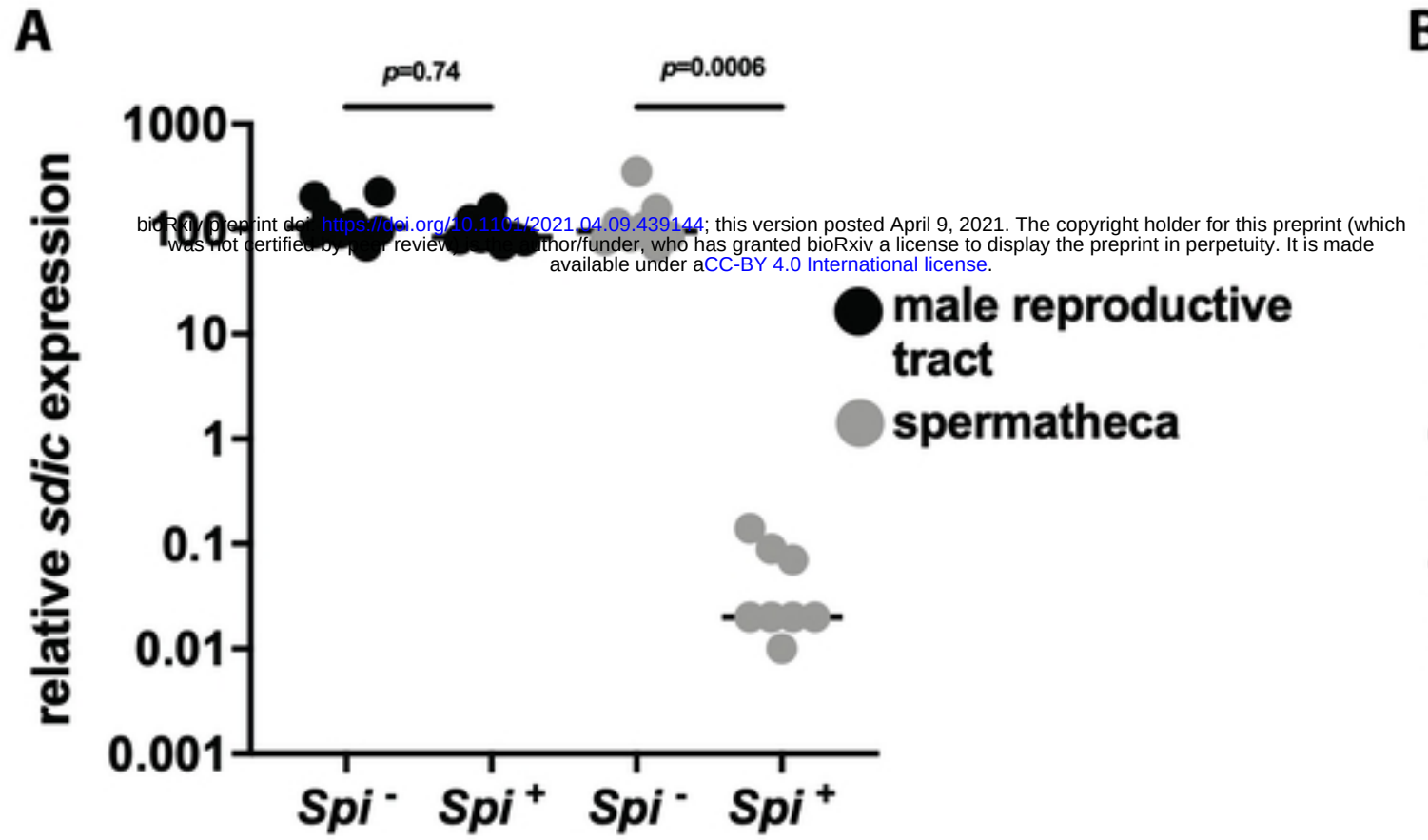
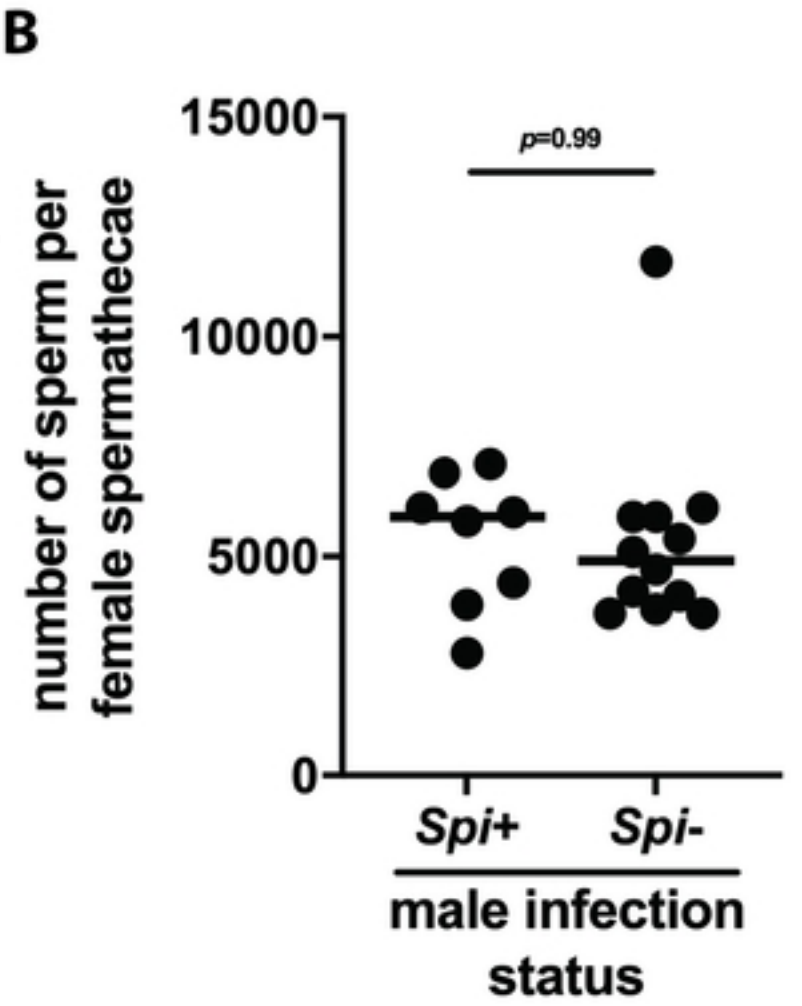
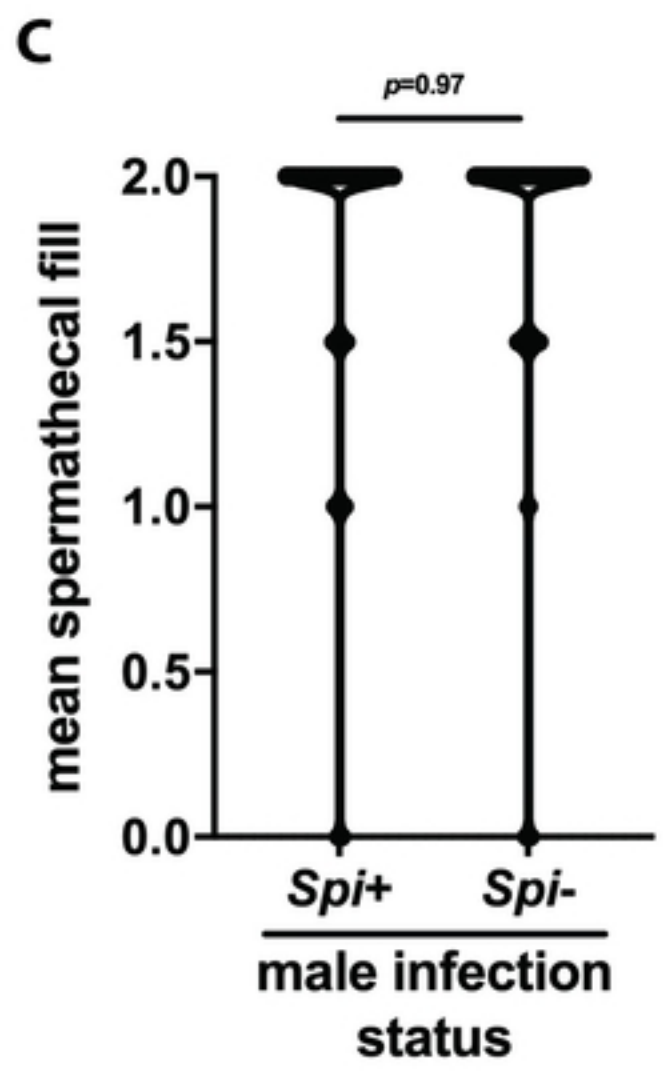

A
length of gonotrophic cycle (days)
30
20
10
0
p=0.0001
p<0.0001
p=0.0009
GC1
GC2
GC3
GffSpi+
GffSpi-
B
pupal weight (mg)
35
30
25
20
15
p=0.34
p=0.08
p=0.15
GC1
GC2
GC3
GffSpi+
GffSpi-
C
ug TAG/ul of hemolymph
25
20
15
10
5
0
p<0.0001
Spi-
Spi+

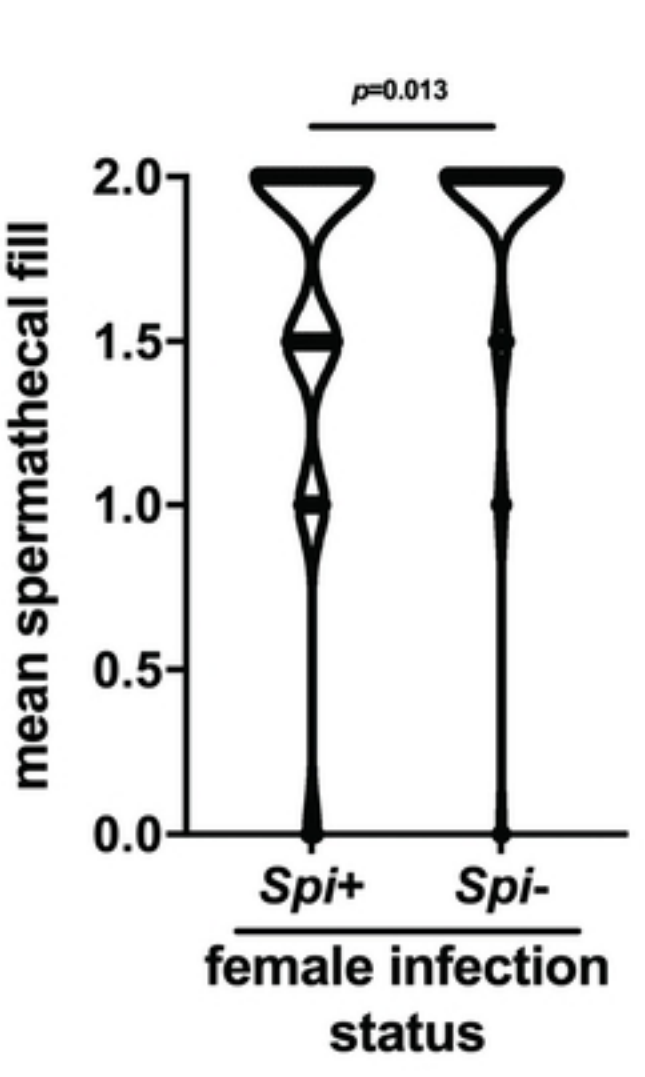
Fig_3

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.09.439144; this version posted April 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.









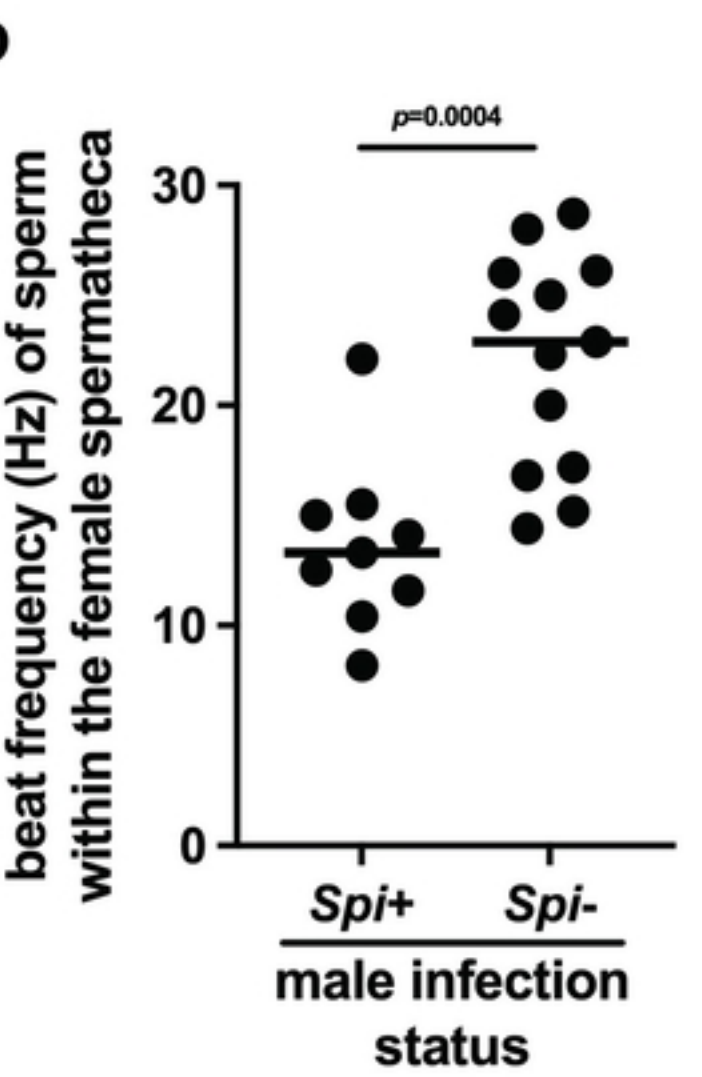


Fig_4