Challenging the *Wigglesworthia, Sodalis, Wolbachia* symbiosis dogma in tsetse flies: *Spiroplasma* is present in both laboratory and natural populations

Doudoumis V[1+], Blow F[2+], Saridaki A[1], Augustinos A[1,3], Dyer NA[4], Goodhead I[2,#], Solano P[5], Rayaisse J-B[6], Takac P[7,9], Mekonnen S[8], Parker AG[3], Abd-Alla AMM[3], Darby, A[2*], Bourtzis K[3*], Tsiamis G[1*]

[1]Department of Environmental and Natural Resources Management, University of Patras, Agrinio, 30100, Greece

[2]Department of Functional and Comparative Genomics, Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK

[3]Insect Pest Control Laboratory, Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture, Vienna International Centre, P.O. Box 100, 1400 Vienna, Austria

[4]Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK

[5]Institut de Recherche pour le Développement, Unité Mixte de Recherche Interactions Hôtes-Vecteurs-Parasites-Environnement dans les Maladies Tropicales Négligées Dues aux Trypanosomatides, 34398 Montpellier, France

[6]Centre International de Recherche-Développement sur l'Élevage en zone Subhumide (CIRDES), N°559, Rue 5-31 Avenue du Gouverneur Louveau, 01 BP 454 Bobo Dioulasso 01, Burkina Fas

[7] Institute of Zoology, Section of Molecular and Applied Zoology, Slovak Academy of Sciences, Dubravska cesta 9, 845 06 Bratislava, Slovakia

[8] National Institute for Control and Eradication of Tsetse and Trypanosomosis (NICETT), Kality Centre, Akaki Kality, Addis Ababa, Ethiopia

[9] Scientica, Ltd, Hybesova 33, 831 06 Bratislava, Slovakia

[#] Current address: School of Environment and Life Sciences, The Crescent, University of Salford, M5 4WT, UK.

[+] DV and BF have contributed equally

*Correspondence:

Alistair.Darby@liverpool.ac.uk, K.Bourtzis@iaea.org, gtsiamis@upatras.gr

Running title: *Spiroplasma*: a new symbiont in tsetse flies

Abstract

Profiling of wild and laboratory tsetse populations using 16S *rRNA* gene amplicon sequencing allowed us to examine whether the "*Wigglesworthia-Sodalis-Wolbachia* dogma" operates across species and populations. The most abundant taxa, in wild and laboratory populations, were *Wigglesworthia* (the primary endosymbiont), *Sodalis* and *Wolbachia* as previously characterized. The species richness of the microbiota was greater in wild than laboratory populations. *Spiroplasma* was identified as a new symbiont exclusively in *Glossina fuscipes fuscipes* and *G. tachinoides*, members of the *palpalis* sub-group, and the infection prevalence in several laboratory and natural populations was surveyed. Multi locus sequencing typing (MLST) analysis identified two strains of tsetse-associated *Spiroplasma*, present in *G. f. fuscipes* and *G. tachinoides*. *Spiroplasma* density in *G. f. fuscipes* larva guts was significantly higher than in guts from teneral and 15-day old male and female adults. In gonads of teneral and 15-day old insects, *Spiroplasma* density was higher in testes than ovaries, and was significantly higher density in live versus prematurely deceased females indicating a potentially mutualistic association. Higher *Spiroplasma* density in testes than in ovaries was also detected by fluorescent *in situ* hybridization in *G. f. fuscipes*.

Introduction

Tsetse (*Glossina* spp.; Diptera: Glossinidae) are viviparous, obligate blood feeding flies found in sub-Saharan Africa. They are the only cyclical vectors of African trypanosomes, responsible for human African trypanosomosis (HAT) and animal African trypanosomosis (AAT)[1,2]. Tsetse larvae feed on milk produced in the milk glands of their mothers, pupariating less than an hour after birth. Adult flies of both sexes feed exclusively on largely sterile blood meals.

The microbiota of tsetse flies is of interest because of their unique lifestyle, highlighted by their bilateral transmission, and reproductive strategy, including the elicitation of phenotypes like cytoplasmic incompatibility, as well as its potential for vector and disease control[3-5]. So far, it is known that tsetse flies harbour three main symbiotic microbes: *Wigglesworthia*, *Sodalis* and *Wolbachia*. These three symbionts form the tsetse symbiosis dogma. The primary mutualist symbiont *Wigglesworthia* provides dietary supplements that are necessary for host fecundity as well as supporting larval development and the maturation process of the adult immune system[6-9]. The facultative symbiont *Sodalis* is present in tsetse populations with a putative role in the ability to transmit trypanosomes[10]. Finally, *Wolbachia* has been found in natural populations of tsetse flies with some species exhibiting up to 100% infection rate[11,12], while others have been found to be free of *Wolbachia*, like *G. p. palpalis* (*Gpp*)[12]. In addition, the *Wolbachia* strain present in *Glossina morsitans morsitans* (*Gmm*) can induce cytoplasmic incompatibility under laboratory conditions[11].

There have been a limited number of culture-dependent and culture-independent studies aiming to characterize the microbiota associated with tsetse flies. Using classical microbiological approaches, Geiger and colleagues isolated *Acinetobacter,*

Enterobacter, Enterococcus, *Providencia*, *Sphingobacterium*, *Chryseobacterium*, *Lactococcus*, *Staphylococcus* and *Pseudomonas* species from the guts of field collected *Gpp* in Cameroon[13-16]. They also isolated a new bacterial species, *Serratia glossinae*, from the midgut of *G. palpalis gambiensis* (*Gpg*) collected in Burkina Faso[14]. A screen for both cultivable and non-cultivable bacteria in whole *G. fuscipes fuscipes* (*Gff*) was performed with flies collected in Kenya[17]. *Firmicutes*, and particularly members of the *Bacillus* genus, were identified as the most dominant group while *Paenibacillus*, *Staphylococcus* and *Exiguobacterium* spp. were also isolated at lower density. *Gammaproteobacteria* were also present, mainly members of the *Enterobacteriaceae* family like *Morganella* and *Providencia* and to a lesser degree *Pseudomonas* spp., while *Burkholderia* was the only member of *Betaproteobacteria* detected in this study[17]. Using a culture independent approach, beyond the mutualist symbiont *Wigglesworthia,* only *Bacillus* and *Serratia* spp. were additionally detected[17]. Aksoy and colleagues sampled guts of Ugandan *Gff*, *Gmm,* and *G. pallidipes* (*Gpal*) tsetse flies, and profiled the microbiota using Illumina amplicon sequencing[18]. *Wigglesworthia* was the dominant taxon, while *Sodalis* was generally detected at low density (<0.05%). However, a small number of flies harboured high levels of *Sodalis* and *Serratia* spp. Non-*Wigglesworthia Enterobacteriaceae* together with *Halomonas* spp. and were also found at lower abundance at all field sites studied, with some bacterial taxa being unique to a sample site.

Spiroplasma is a genus of wall-less bacteria belonging to the class *Mollicutes* and it has been associated with diverse plants and arthropods[19-22]. *Spiroplasma* is grouped into three major clades as has been shown by 16S *rRNA* gene-based as well as multi locus sequence typing (MLST) studies[23-30]. *Spiroplasma* exhibits a dual life, with

capacity to live intracellularly in a variety of tissues and systemically in the haemolymph[31]. *Spiroplasma* has developed a wide range of symbiotic associations, producing diverse effects on insect evolution, ecology, reproduction and sex determination. *Spiroplasma* has been found to confer protection against a nematode in *Drosophila neotestacea*[32], against fungi in the pea aphid (*Acyrthosiphon pisum*)[33], and against a parasitoid wasp in *Drosophila hydei*[34]. *Spiroplasma* can also be pathogenic in plants[35], insects[36-38] and crustaceans[39-44]. Moreover, several species of *Spiroplasma* have been associated with reproductive alterations such as male killing[45-48,29]. Except *Spiroplasma*, other reproductive parasites that have been associated with insects are *Arsenophonus*, *Cardinium*, and *Rickettsia*. *Arsenophonus* is known to establish diverse symbiotic interactions with around 5% of insect species, with the most profound phenotype induced being the son-killer trait[49,50]. *Cardinium* has been found exclusively to Hymenoptera, Hemiptera, Diptera, and Acari and it is known to induce cytoplasmic incompatibility and feminization[51,52]. Finally, *Rickettsia* has been associated with regulating insect growth and immunity to pathogenic fungi[53-55].

In this study we employed high throughput sequencing of the 16S *rRNA* gene to unravel the diversity of tsetse associated bacteria in a wider variety of species, field and laboratory populations than any previous tsetse microbiota study. We asked whether the "*Wigglesworthia-Sodalis-Wolbachia* dogma" applies across species and populations, and whether the microbiota varies between laboratory and field individuals of the same tsetse species. *Spiroplasma* was identified as a novel symbiont of *Gff* and *G. tachinoides* (*Gt*), and infection prevalence was surveyed in laboratory and natural populations. Quantitative PCR was used to characterize its density in different developmental stages and tissues, and to quantify infection levels in

collapsing mass-rearing tsetse fly colonies. Fluorescent *in situ* hybridization (FISH) was used to localize the newly identified symbiont in tissues including the gonads.

Results

16S *rRNA* gene amplicon sequencing reveals novel interspecific diversity in natural populations of tsetse flies

Microbial community composition and diversity of thirty-two whole insects from *G. medicorum* (*Gmed*), *G. morsitans submorsitans* (*Gms*), *G. p. gambiensis* (*Gpg*), and *G. tachinoides* (*Gt*) collected in Folonzo, Burkina Faso were investigated by 16S *rRNA* gene amplicon sequencing, producing 5,761,899 reads after quality filtering. These reads were combined with a total of 8,300,515 quality-filtered reads generated from 124 whole guts of *Gff, Gmm, Gpal* from a previous study[18], which used an identical technical approach for amplicon generation and sequencing. Including the data from the above mentioned study[18] provided additional *Wigglesworthia*/co-divergence context to our dataset due to the increased host diversity. This approach enabled us to characterize low-frequency, high-abundance taxa. Whole insect samples from *Gpg*, and *Gt* were the most bacterial species-rich samples containing higher numbers of unique OTUs (Supplementary Table 1).

The primary nutritional endosymbiont of tsetse flies *Wigglesworthia glossinidia* was the most abundant taxon in all samples, and constituted between 71 and 99% of the total community in each individual. Variation in the relative abundance of *W. glossinidia* was due to the heterogeneous distribution of secondary taxa, which varied in infection frequency and abundance between individuals in both an intra- and inter-specific fashion (Figure 1a). Secondary taxa included the facultative symbionts *S. glossinidius* and *Wolbachia,* alongside *Spiroplasma*, which have not previously been

reported in tsetse flies. The relative abundance of secondary taxa was highly variable (from < 0.01% to 28%) depending upon the genus of the bacterium and the species of *Glossina* (Figure 1a). This contributed to the variation in bacterial community composition between *Glossina* species. Clustering by species is illustrated in Figure 1B, where Principal Component 1 and Principal Component 2 describe 58.53 % and 10.84 % of the variance respectively. Clustering can be partly attributed to the co-diversification of *Wigglesworthia*, which is the main component of the community, with its tsetse host[56]. For this reason, outliers are conspicuous, as is observed with the two individuals infected with *Spiroplasma* and *Rickettsia* at 13.15% and 23.72% relative abundance respectively (Figure 1b).

Sodalis was found at higher frequency and relative abundance in whole *Gmed* and *Gpal* guts (Figure 2 & Supplementary Figure 1). In the other *Glossina* species it has been detected but to a much lower abundance with a relative abundance of 0.5% or less, with *Gms* exhibiting the lowest abundance. *Wolbachia* infections were found infrequently and with low relative abundance of up to 0.04% in any wild sample, with *Gmed* exhibiting the highest infection prevalence (Supplementary Figure 1).

In addition, several other taxa previously associated with tsetse flies were detected including multiple members of the *Enterobacteriaceae*, such as *Klebsiella, Erwinia, Trabulsiella, Pantoea,* and *Serratia*. These infections occurred at low relative abundance, excluding those with *Klebsiella*, which was found to be dominant in one *Gpg* and one *Gms* whole fly at a relative abundance of 24.3% and 15.3% respectively (Figure 1a). Amplicon profiling was also able to detect taxa that had not previously been associated with the tsetse fly. Several wild individuals of *Gff* and *Gt*, which belong to the *palpalis* subgroup of the *Glossina* genus, were infected with *Spiroplasma.* Relative abundances were generally low (<1%) (Supplementary Table

2), but were found to be as high as 13.2% in one *Gt* whole fly from Burkina Faso (Figure 1a).

16S *rRNA* gene amplicon sequencing of laboratory reared tsetse flies

Gff, Gmm, and *Gpal* tissue samples from three developmental stages were sequenced, producing 2,445,369 reads after quality filtering. Similarly to wild populations, the three known taxa (*Wigglesworthia, Sodalis* and *Wolbachia*) were found in the laboratory flies. However, additional bacterial species were also detected, with members of *Flavobacterium*, *Propiniobacterium*, *Brevundimonas*, *Aeromonas*, and *Rhodospirillales* identified in *Gmm*, *Gff*, and *Gpal*. Sequences related to *Acinetobacter* and *Pantoea* were identified in *Gmm* and *Gpal*. Additionally, sequences related to *Streptococcus* were found in *Gmm*, and *Gff*, while sequences related to *Shewanella*, and *Pedobacter* were discovered only in *Gmm*. Relative abundance was influenced by tissue sample type with gut tissues being enriched for *Wigglesworthia* while reproductive tissues were characterized by the presence of *Wolbachia* and *Sodalis*.

For *Gpal*, the most bacterial species-rich samples were those associated with gonads of teneral flies while gut samples were less species-rich based on both Chao1 and ACE indices (Supplementary Table 3). Gut samples of teneral males and females displayed lower species richness (Supplementary Table 3). The same trend was observed for *Gff*. Gut samples of teneral flies exhibited the lowest species diversity and richness indices, which increased over time (Supplementary Table 3). Conversely, gonads presented a higher diversity and richness index in teneral flies and decreased in aged flies. This pattern was not observed in *Gmm*. Finally, the natural populations

exhibited a statistically significant higher species-rich index (Chao1) when compared with the laboratory populations ($p<0.016$).

We observed variation in the frequency and relative abundance of *Wolbachia* in lab populations. The mean relative abundance of *Wolbachia* was significantly higher in *Gmm* flies compared with those from the *Gff* or *Gpal* populations (ANOVA, $p\leq0.01$) (Supplementary Table 2). This was due to increased relative abundance of *Wolbachia* in reproductive tissues compared to larval or gut tissues within the *Gmm* population (ANOVA, $p\leq0.01$).

Bacterial communities were strongly clustered according to the tissue of origin separating the bacterial communities from guts from those from reproductive tissues (Figure 3a). This factor explained 81.3% of the total variance. Canonical analysis of principal coordinates (CAP), revealed distinct clustering within the gonadal tissue (Figure 3b). The bacterial communities associated with the gonadal tissue also seem to be statistically affected by the host; *Gmm*, *Gff,* and *Gpal* bacterial communities associated with the reproductive organs clustered separately (Figure 3b), with *Spiroplasma* driving the *Gff* cluster and *Wolbachia* the *Gmm*. CAP ordinations were supported by significant $trace_{Q\_m'HQ\_m}$ statistics (0.9598; $p<0.05$).

Spiroplasma infection status assessed by PCR screening of natural and laboratory tsetse populations

We used PCR-based screening methods to assay for the presence of four insect reproductive parasites: *Spiroplasma*, *Arsenophonous*, *Rickettsia*, and *Cardinium,* in four *Glossina* species from the laboratory, *Gmm* (n=19), *Gff* (n=76), *Gpal* (n=20), and *Gpg* (n=19) and wild *Gff* (n=98). Of the four examined *Glossina* species, *Spiroplasma* infections were found only in *Gff* with an infection ranging from 6.7 to 80% (Table 1),

while none of the four tsetse species examined were infected with *Arsenophonus*, *Rickettsia* or *Cardinium*.

To examine the distribution of *Spiroplasma*, six additional *Glossina* species were PCR-screened for *Spiroplasma* infection. Only *Gt* and *Gpp* were positive for *Spiroplasma*, and showed an infection rate of 26.7% and 12.5% respectively (Table 1). The PCR screening for *Spiroplasma* infection was further extended to 327 historical and contemporary samples from wild and laboratory colonies representing 10 species of tsetse fly (Table 1). Only members of the *palpalis* subgroup were found infected with *Spiroplasma*, including *Gff, Gpp* and *Gt*, with a prevalence ranging from 6% to 80%. Notably, the prevalence was higher in laboratory colonies than natural populations, and some populations demonstrated a disparity in infection between sexes (Table 1).

Genotyping of *Spiroplasma* strains

Spiroplasma strains from *Gff* flies of both sexes from laboratory colonies, a natural population from Uganda and from one natural population of *Gt* flies from Burkina Faso were genotyped by MLST analysis. Four laboratory and one field sample of *Gff* harbour *Spiroplasma* strains with identical sequences for all loci studied (Supplementary Table 4). Interestingly, the *Spiroplasma* strain present in *Gt* is distinct from the *Gff Spiroplasma* strain with sequence polymorphisms detected in all loci examined. Eight polymorphisms were observed in *fruR*, seven in the region 16S *rRNA*-23S *rRNA*-5S *rRNA*, four in 16S *rRNA*, three in *dna*A, two in *ftsZ*, and one in *rpoB* and *parE*. Both strains belong to the citri clade, which is mostly composed of plant pathogens (Figure 4 and Supplementary Figures 2-7). Most of the pathogenic *Spiroplasma* species belong to the Citri clade[57] with prominent examples including *S.*

kunkelii that causes the corn stunt disease[21], *S. phoeniceum* that infects periwinkle[58], and *S. penaei* that infects Pacific white shrimp[42]. The closest relatives of the tsetse *Spiroplasma* strains are *S. insolitum* and *S. atrichopogonis*, which were isolated from a fall flower and a biting midge (Diptera: Ceratopogonidae) respectively[59,60]. Neither *S. insolitum* or *S. atrichopogonis* have been reported to be pathogenic to plants or midges.

Spiroplasma density across developmental stages

qPCR was used to assess the density of the *Spiroplasma* infection in larval guts, and in guts and gonads of males and females collected at two developmental stages: (a) teneral and (b) 15-day-old *Spiroplasma* infection levels were significantly higher in larval guts compared to the guts of teneral or 15-day-old adults (Figure 5a). There was no significant difference in the infection levels between testes of teneral and 15-day-old adults (Supplementary Figure 8). In a similar way no significant difference was observed between ovaries of teneral and 15-day-old adults (Supplementary Figure 9). However, there was a significant difference in *Spiroplasma* infection level between testes and ovaries from teneral flies (Figure 5b).

Spiroplasma density was also examined in a mass-rearing colony where mortality was high and the colony was on the verge of collapse. Examination of live and dead insects indicated that in males *Spiroplasma* density was similar, whereas in females density was higher in live insects than in those that had recently perished (Figures 6a and 6b). When we examined exclusively females carrying a larva, we found that the live females with a larva had a higher titre of *Spiroplasma* than gravid females that died prematurely (Figure 6c). The prevalence of *Wolbachia*, *Arsenophonus*, *Cardinium*, and *Rickettsia* was also examined in whole tsetse flies from the collapsing

colony. None of the 34 individuals tested were found to harbour any of the above mentioned symbionts.

in situ hybridization of *Spiroplasma*

Dissected ovaries and testes of teneral adults from a *Gff* laboratory colony were subjected to FISH using a *Spiroplasma* specific probe. *Spiroplasma* detection was sparse and sporadic in ovaries (Figure 7a), while in testes it was observed at high densities (Figure 7b).

Discussion

The present study showed that the bacterial communities associated with tsetse flies are more complex than previously reported, thus challenging the *Wigglesworthia-Sodalis-Wolbachia* dogma[61,62,3]. Using 16S *rRNA* gene-based sequencing approaches, several additional bacterial genera with broad phylogenetic origins were discovered to be associated with the tsetse fly including *Klebsiella, Rickettsia* and *Spiroplasma*. The prevalence and infection levels observed in some tsetse species, particularly those of *Spiroplasma*, were similar to those seen for *Sodalis*, suggesting that they may play an important role in the biology and ecology of tsetse flies. The question is where these symbionts come from, and what factors determine the structure of the symbiotic communities of tsetse flies.

Previous studies have shown that the microbiota of tsetse flies is characterized by the presence of *Wigglesworthia*, *Sodalis* and *Wolbachia*. All three symbionts are maternally transmitted, while *Sodalis* can also be transmitted paternally, and colonize during the early juvenile stages: *Wigglesworthia* and *Sodalis* through milk gland secretions as larvae, and *Wolbachia* through the germ line during embryogenesis[63,3,64].

As larvae are intrauterine, the only bacteria that they encounter prior to pupation originate from within the adult female tsetse fly. Due to the obligate requirement of *Wigglesworthia*, there is high fidelity in vertical transmission from mother to offspring[65]. This makes it difficult for other bacteria to invade, as microbes occupy many of the available niches within the host from the early stages of development. Conversely, this also means that the tsetse immune system has evolved to accommodate bacteria, which could facilitate colonization by environmental microbes able to exploit deficits in the immune system. Due to the unique biology of tsetse flies, there is only a short time window for colonization between larval deposition and pupation in the soil. In addition, the colonizers would have to survive metamorphosis in order to persist.

Until recently, there was the notion that tsetse flies feed exclusively on blood, which is mostly sterile and therefore should not serve as a source of microbes. There is now evidence that *Gpg* flies deprived of a blood meal can feed on water or sugar water, and that sugar residues are detectable in wild-caught flies[66]. Therefore, it is possible that these previously unrecognized feeding habits could be a source of environmental microbes, and could be the origin of the low-frequency high-abundance infections observed in multiple individuals in this study.

Spiroplasma was detected in members of the *palpalis* sub-group (*Gff, Gpp* and *Gt*), whereas *Sodalis* was significantly more prevalent in *Gmed* (*fusca* group). Previous studies have also shown that *Sodalis* infection is more prevalent in *G. brevipalpis* (*fusca* group) than in *Gmm* and *Gpal* (both *morsitans* group)[67]. However, the relationship of *Spiroplasma* with the *palpalis* subgroup seems to be more exclusive than that of *Sodalis*, since the latter has previously been identified in individuals belonging to all tsetse sub-groups[18,67,68].

A key approach to detecting invasive taxa is to sample whole insects rather than individual tissues such as the gut, where *Wigglesworthia* is dominant and will therefore obscure the detection of lower-abundance taxa. A broad phylogenetic range of host species is important to encompass the available diversity, as there seems to be variation between sub-groups, species, and even individuals within the same species. For example, *Rickettsia* was discovered at high abundance in just one individual, despite the profiling of hundreds of insects by amplicon and PCR profiling. *Rickettsia* has been also identified in a previous study using an amplicon sequencing approach[18] but also to *G. morsitans* from Senegal during a PCR screen[69].

Spiroplasma infection was more prevalent in laboratory colonies with both males and females harbouring *Spiroplasma*, whereas in natural populations prevalence was lower and only females were infected. The lack of infection in wild individuals may be due to insufficient sampling effort, or could be due to the differences in population dynamics between laboratory-reared and wild-caught flies. It has been reported, for example, that some symbionts may be present in such low abundances that they are undetectable by conventional PCR screens[70]. MLST indicated that the strain found in wild *Gff* from Uganda was identical, based on the loci examined, to that in the colonized flies (originating from the Central African Republic), suggesting the association between *Spiroplasma* and *Gff* may be ancient. Although there have been no direct studies on the relative transmission rate of tsetse symbionts in the laboratory and field, paternal transmission during mating can occur for the secondary symbiont *Sodalis*[64]. While this study only detected *Spiroplasma* infection in *palpalis* group flies, screening more specimens from the *morsitans* and *fusca* groups should provide more detailed information on the dynamics and spread of *Spiroplasma* infection in natural populations.

374 Another potential explanation for the absence of *Spiroplasma* in the *morsitans* and
375 *fusca* groups is their frequent infection with *Wolbachia*[12,71]. In the *morsitans* group the
376 prevalence of *Wolbachia* can vary between 9.5 and 100%, while in the *fusca* group it
377 can vary from 0 to 15.6%[12,71]. An existing *Wolbachia* infection may have led to the
378 development of competitive exclusion with *Spiroplasma*, though it is not yet clear
379 whether they share an ecological niche within the host, and whether co-occurrence
380 could create evolutionary pressure strong enough to drive competitive exclusion[72]. In
381 *D. melanogaster*, coinfections between *Wolbachia* and *Spiroplasma* were
382 asymmetrical: *Spiroplasma* negatively affected the titre of *Wolbachia*, whereas
383 *Wolbachia* density did not affect *Spiroplasma* titre[73]. Similarly to *Spiroplasma* in *Gff*,
384 tissue tropism was observed in *D. melanogaster* infected with *Spiroplasma*, with the
385 ovaries showing the highest density[73]. Competitive inter- and intraspecific microbial
386 interactions have also been observed in mosquito vector species where mutual
387 exclusion between *Asaia* and *Wolbachia* has been observed in the reproductive organs
388 while native gut microbiota seems to prevent the vertical transmission of *Wolbachia*
389 in *Anopheles* mosquitoes[74,75]. *Gff* has previously been shown to harbor *Wolbachia*,
390 though prevalence in natural populations is very heterogeneous, with an average
391 infection rate of 44.3%[76]. *Spiroplasma*, on the other hand, is found at much lower
392 frequency in natural populations, but is found at higher density per individual when
393 compared with *Wolbachia*.

394 MLST analysis indicated that the *Spiroplasma* strains detected in *Gff* and *Gt*
395 populations, albeit different, both belong to the citri clade. Prominent examples of
396 taxa from this clade include *S. kunkelii*, *S. phoeniceum*, and *S. citri*, all of which are
397 plant pathogens[77,58,21]. *S. poulsonii,* which has been shown to have a protective effect
398 against parasitic wasps in *D. melanogaster,* is also a member of this clade[20].

When examining gut tissues, *Spiroplasma* titre was highest in larvae, and gradually decreased in both males and females over the course of adulthood. High larval titre indicates vertical transmission from mother to offspring, possibly via the milk gland; a mechanism already exploited by *Wigglesworthia* and *Sodalis*. High larval density is an abnormal trait in the context of other insect-associated *Spiroplasma* species. Multiple strains of *Spiroplasma* infect a number of species of *Drosophila* and are able to induce a variety of phenotypes in their insect host ranging from parasitic reproductive manipulators to protective symbionts[78,24,20]. In *D. hydei* and *D. melanogaster, Spiroplasma* titre steadily increases during larval and adult development with no differentiation between males and females[79,73]. Interestingly, *Drosophila* male killing *Spiroplasma* strains exhibit a very high titre in the haemolymph[78], a pattern not observed in the *Gff Spiroplasma* strain (data not shown). In addition, *Spiroplasma* titre in *Gff* is much lower than that described for *Drosophila* male killing strains[29,78]. *Wolbachia* is the only other maternally inherited endosymbiont found in *Drosophila,* and is also found in tsetse flies. *Wolbachia* confers density-dependent protection against insect viruses at different developmental stages in several *Drosophila* species[80-83]. Based on the above, it is possible that high *Spiroplasma* density may also play a role in larval fitness. This warrants further study, as protection against viral or bacterial pathogens during intrauterine larval development would constitute a rare phenotype for a bacterial endosymbiont. Recent studies in *D. melanogaster* showed that *Wolbachia* and *Spiroplasma* can affect immune signalling pathways in the presence of both insect pathogenic and non-pathogenic bacteria[84].

Gut infection was maintained into adulthood, particularly in males. This suggests that *Spiroplasma* is either able to maintain infection during metamorphosis, possibly due

to extracellular proliferation[73], or that it can rapidly re-colonize upon reformation of the gut. *Spiroplasma* density was also significantly higher in the testes of teneral males than in the ovaries of teneral females. Localization to the testes suggests that *Spiroplasma* may be sexually transmitted from males to females, as has already been observed with *Sodalis* in tsetse flies, and *Asaia* in *Anopheles stephensi*[64,85]. The above properties can be exploited in paratransgenic approaches in a similar way to those currently being explored for *Sodalis*[86,64] and *Asaia*[87].

In a collapsing colony of *Gff* flies, live females had a higher *Spiroplasma* density than prematurely dead females. This was true of both gravid and non-gravid females, and indicates that *Spiroplasma* may contribute to adult female fitness. It is therefore possible that *Spiroplasma* could play a protective role, as has been observed in other facultative strains of *Spiroplasma*[20,34,88] and/or a nutritional role.

Materials and Methods

Insect specimen collection and DNA isolation

All natural populations of *Glossina* specimens were collected in four countries, Burkina Faso, Uganda, United Republic of Tanzania, and South Africa (Table 1 & Supplementary Table 5). All wild flies were collected using biconical traps and collection intervals were four hours. Upon collection, flies were transferred to the main collection point and were placed in 100% acetone and stored at room temperature. Upon arrival in the lab, DNA was extracted immediately using the CTAB method (Cetyl trimethylammonium bromide)[89]. Laboratory populations were also analysed in a similar way. Samples of *Gff* suffering high mortality were collected from the mass rearing facility in Kality, Ethiopia. For a detailed description of the analysis performed see Supplementary Information.

Multiplex Illumina MiSeq Sequencing, data, and statistical analysis

The V4 region of the 16S *rRNA* gene was amplified using fusion primers F515 (5'-GTGCCAGCMGCCGCGGTAA-3'), and 805R (5'- GACTACCAGGGTATCTAAT -3') from individual wild flies of *G. medicorum* (*Gmed*), *G. m. submorsitans* (*Gms*), *G. p. gambiensis* (*Gpg*), and *G. tachinoides* (*Gt*) collected in Burkina Faso. Data generated from the wild flies were combined with the data generated from 124 whole guts of *Gff, Gmm, Gpal* from a previous study[18], which used an identical technical approach for amplicon generation and sequencing.

The V3-V4 region of the 16S *rRNA* gene was amplified using fusion primers U341F (5'-CCTACGGGRSGCAGCAG-3'), and 805R (5'- GACTACCAGGGTATCTAAT -3') from pools of tissues from larvae and adults of laboratory populations of *Gmm*, *Gff*, and *Gpal* (Supplementary Table 5).

For a detailed description of the PCR conditions please see Supplementary Information. The gene sequences reported in this study have been deposited in NCBI under Bioproject numbers PRJNA345319, and PRJNA345350-52. Statistical analyses was performed using Unifrac distances, PCoA analyses, CAP, ANOVA and Tukey-Kramer post-hoc tests as described in the Supplementary Information.

PCR screening and *Spiroplasma* multi locus genotyping

Gmm, *Gff*, *Gpg*, and *Gpal* were assayed for the presence of *Spiroplasma*, *Arsenophonus*, *Cardinium*, and *Rickettsia* symbionts by PCR. An additional six species of *Glossina* (*G. austeni (Ga), G. brevipalpis (Gb), G. m. centralis (Gmc), Gms, G. p. palpalis (Gpp)* and *Gt* were screened for *Spiroplasma* only. The primer

sequences used to detect each symbiont along with their target genes, product sizes, conditions, and annealing temperatures are listed in the Supplementary Information.

The *Spiroplasma* strains present in *Glossina* species were genotyped with a multi-locus sequence typing (MLST) approach using five marker genes (*rpo*B, *par*E, *dna*A, *fts*Z and *fru*R) and a 4,702 bp region spanning the 16S *rRNA*-23S *rRNA*-5S *rRNA* region. Details of the conditions used are presented in the Supplementary Information. Sequencing was performed as described previously[90]. All gene sequences generated in this study have been deposited into at GenBank under accession numbers KX159363-KX159393.

Phylogenetic analysis

All nucleotide sequences were manually edited with Geneious 7.1.2. Multiple alignments were generated by MUSCLE[91] and ClustalW[92] by Geneious 7.1.2, and adjusted by eye. Phylogenetic analyses were conducted for all analysed *Spiroplasma* sequences (16S *rRNA*, *rpoB*, *dnaA*, *parE*, *ftsZ* and *fruR* genes, and the region 16S *rRNA*-23S *rRNA*-5S *rRNA* region) separately by two methods: Bayesian Inference (BI) and Maximum Likelihood (for a detailed description see Supplementary Information).

Quantitative Real Time-PCR and Fluorescent *in situ* Hybridization (FISH)

Spiroplasma density was quantified by qPCR using the *dna*A *Spiroplasma* specific primers FqdnaA/RqdnaADoud for 35 cycles at 56 °C and normalized to the host *β-tubulin* gene. Primers and a detailed description used for the qPCR experiments are presented in Supplementary Table 6. qPCR data were analysed using a one-way ANOVA method, as described previously[93] using the XLSTAT program.

498 *Gff* specimens from the Seibersdorf laboratory colony were used for FISH. Teneral
499 male and female flies were dissected in PBS 2-3 days after eclosion. Dissected tissues
500 were dried on poly-L-lysine-coated glass slides (Sigma, UK) for 20 min at 65 °C and
501 kept at 4 °C until further use. Tissue samples were fixed in freshly prepared 4%
502 paraformaldehyde solution for 30 min at 4 °C. A detailed description of tissue
503 processing and image capture is included in the Supplementary Information.
504

505 **References**

506 1. Aksoy, S. Sleeping sickness elimination in sight: Time to celebrate and reflect, but
507 not relax. *PLoS Negl. Trop. Dis.* **5**, e1008 (2011).
508 2. Welburn, S. C., Fèvre, E. M., Coleman, P. G., Odiit, M. & Maudlin, I. Sleeping
509 sickness: a tale of two diseases. *Trends Parasitol.* **17**, 19–24 (2001).
510 3. Wang, J., Weiss, B. L. & Aksoy, S. Tsetse fly microbiota: form and function.
511 *Front. Cell. Infect. Microbiol.* **3**, 69 (2013).
512 4. Abd-Alla, A. M. M. et al. Improving Sterile Insect Technique (SIT) for tsetse flies
513 through research on their symbionts and pathogens. *J. Invertebr. Pathol.* **112**, S2–10
514 (2013).
515 5. Bourtzis, K., Lees, R. S., Hendrichs, J. & Vreysen, M. J. B. More than one rabbit
516 out of the hat: Radiation, transgenic and symbiont-based approaches for sustainable
517 management of mosquito and tsetse fly populations. *Acta Trop.* **157**, 115–130 (2016).
518 6. Aksoy, S. *Wigglesworthia* gen. nov. and *Wigglesworthia glossinidia* sp. nov., taxa
519 consisting of the mycetocyte-associated, primary endosymbionts of tsetse flies. *Int. J.*
520 *Syst. Bacteriol.* **45**, 848–851 (1995).
521 7. Weiss, B. & Aksoy, S. Microbiome influences on insect host vector competence.
522 *Trends Parasitol.* **27**, 514–522 (2011).

523 8. Weiss, B. L., Wang, J. & Aksoy, S. Tsetse immune system maturation requires the
524 presence of obligate symbionts in larvae. *PLoS Biol.* **9**, e1000619 (2011).
525 9. Weiss, B. L., Maltz, M. & Aksoy, S. Obligate symbionts activate immune system
526 development in the tsetse fly. *J. Immunol.* **188**, 3395–3403, (2012).
527 10. Farikou, O. et al. Tripartite interactions between tsetse flies, *Sodalis glossinidius*
528 and trypanosomes - An epidemiological approach in two historical human African
529 trypanosomiasis foci in Cameroon. *Infect. Genet. Evol.* **10**, 115–121 (2010).
530 11. Alam, U. et al. *Wolbachia* symbiont infections induce strong cytoplasmic
531 incompatibility in the tsetse fly *Glossina morsitans*. *PLoS Pathog.* **7**, e1002415
532 (2011).
533 12. Doudoumis, V. et al. Detection and characterization of *Wolbachia* infections in
534 laboratory and natural populations of different species of tsetse flies (genus *Glossina*).
535 *BMC Microbiol.* **12**, 1–13 (2012).
536 13. Geiger, A. et al. First isolation of *Enterobacter*, *Enterococcus*, and *Acinetobacter*
537 spp. as inhabitants of the tsetse fly (*Glossina palpalis palpalis*) midgut. *Inf. Genet.*
538 *Evol.* **9**, 1364–1370 (2009).
539 14. Geiger, A., Fardeau, M-L., Falsen, E., Ollivier, B. & Cuny, G. *Serratia glossinae*
540 sp. nov., isolated from the midgut of the tsetse fly *Glossina palpalis gambiensis*. *Int.*
541 *J. Syst. Evol. Microbiol.* **60**, 1261–1265 (2010).
542 15. Geiger, A. et al. Bacterial diversity associated with populations of *Glossina* spp.
543 from Cameroon and distribution within the Campo sleeping sickness focus. *Microb.*
544 *Ecol.* **62**, 632–643 (2011).
545 16. Geiger, A., Fardeau, M-L., Njiokou, F. & Ollivier, B. *Glossina* spp. gut bacterial
546 flora and their putative role in fly-hosted trypanosome development. *Front. Cell.*
547 *Infect. Microbiol.* **3**, 34 (2013).

548 17. Lindh, J.M. & Lehane, M.J. The tsetse fly *Glossina fuscipes fuscipes* (Diptera:
549 Glossina) harbours a surprising diversity of bacteria other than symbionts. *Antonie*
550 *van Leeuwenhoek* **99**, 711–720 (2011).
551 18. Aksoy, E. et al. Analysis of multiple tsetse fly populations in Uganda reveals
552 limited diversity and species-specific gut microbiota. *Appl. Environ. Microbiol.* **80**,
553 4301–4312 (2014).
554 19. Bastian, F. O., Elzer, P. H. & Wu, X. *Spiroplasma* spp. biofilm formation is
555 instrumental for their role in the pathogenesis of plant, insect and animal diseases.
556 *Exp. Mol. Pathol.* **93**, 116–128 (2012).
557 20. Paredes, J. C. et al. Genome sequence of the *Drosophila melanogaster* male-
558 killing *Spiroplasma* strain MSRO endosymbiont. *mBio* **6**, e02437–14 (2015).
559 21. Whitcomb, R. F. et al. *Spiroplasma kunkelii* sp. nov.: characterization of the
560 etiological agent of corn stunt disease. *Int. J. Syst. Evol. Microbiol.* **36**, 170–178
561 (1986).
562 22. Gasparich, G. E. The genus *Spiroplasma* and its non-helical descendants:
563 phylogenetic classification, correlation with phenotype and roots of the *Mycoplasma*
564 *mycoides* clade. *Int. J. Syst. Evol. Microbiol.* **54**, 893–918 (2004).
565 23. Nai Y-S. et al. A new spiroplasma isolate from the field cricket (*Gryllus*
566 *bimaculatus*) in Taiwan. *J. Invertebr. Pathol.* **120**, 4–8 (2014).
567 24. Haselkorn, T. S., Cockburn, S. N., Hamilton, P. T., Perlman, S. J. & Jaenike, J.
568 Infectious adaptation: potential host range of a defensive endosymbiont in
569 *Drosophila*. *Evolution* **67**, 934–945 (2013).
570 25. Heres, A. & Lightner, D. V. Phylogenetic analysis of the pathogenic bacteria
571 *Spiroplasma penaei* based on multilocus sequence analysis. *J. Invertebr. Pathol.* **103**,
572 30–35 (2010).

573 26. Haselkorn, T. S., Markow, T. A. & Moran, N. A. Multiple introductions of the
574 *Spiroplasma* bacterial endosymbiont into Drosophila. *Mol. Ecol.* **18**, 1294–1305
575 (2009).
576 27. Bi, K., Huang, H., Gu, W., Wang, J. & Wang, W. Phylogenetic analysis of
577 Spiroplasmas from three freshwater crustaceans (*Eriocheir sinensis*, *Procambarus*
578 *clarkia* and *Penaeus vannamei*) in China. *J. Invertebr. Pathol.* **99**, 57–65 (2008).
579 28. Montenegro, H., Solferini, V. N., Klaczko, L. B. & Hurst, G. D. D. Male-killing
580 *Spiroplasma* naturally infecting *Drosophila melanogaster*. *Insect. Mol. Biol.* **14**, 281–
581 287 (2005).
582 29. Anbutsu, H. & Fukatsu, T. Population dynamics of male-killing and non-male-
583 killing Spiroplasmas in *Drosophila melanogaster*. *Appl. Environ. Microbiol.* **69**,
584 1428–1434 (2003).
585 30. Fukatsu, T., Tsuchida, T., Nikoh, N. & Koga, R. *Spiroplasma* Symbiont of the Pea
586 Aphid, *Acyrthosiphon pisum* (Insecta: Homoptera). *Appl. Environ. Microbiol.* **67**,
587 1284–1291 (2001).
588 31. Schwarz, R. S. et al. Honey bee colonies act as reservoirs for two *Spiroplasma*
589 facultative symbionts and incur complex, multiyear infection dynamics. *Microbiol.*
590 *Open* **3**, 341–355 (2014).
591 32. Jaenike, J., Unckless, R., Cockburn, S. N., Boelio, L. M. & Perlman, S. J.
592 Adaptation via symbiosis: recent spread of a Drosophila defensive symbiont. *Science*
593 **329**, 212–215 (2010).
594 33. Łukasik, P., Guo, H., van Asch, M., Ferrari, J. & Godfray, H. C. J. Protection
595 against a fungal pathogen conferred by the aphid facultative endosymbionts *Rickettsia*
596 and *Spiroplasma* is expressed in multiple host genotypes and species and is not

influenced by co-infection with another symbiont. *J. Evol. Biol.* **26,** 2654–2661 (2013).

34. Xie, J., Vilchez, I. & Mateos, M. *Spiroplasma* bacteria enhance survival of *Drosophila hydei* attacked by the parasitic wasp *Leptopilina heterotoma*. *PLoS ONE* **5**, e12149 (2010).

35. Gasparich, G. E. Spiroplasmas and phytoplasmas: Microbes associated with plant hosts. *Biologicals* **38**, 193–203 (2010).

36. Mouches, C., Bové, J. M., Albisetti, J. Pathogenicity of *Spiroplasma apis* and other spiroplasmas for honey-bees in southwestern France. *Ann. Microbiol. (Paris)* **135A,** 151–155 (1984).

37. Clark, T. B. et al. *Spiroplasma melliferum*, a new species from the honeybee (*Apis mellifera*). *Int. J. Syst. Evol. Microbiol.* **35**, 296–308 (1985).

38. Humphery-Smith, I., Grulet, O., Le, Goff, F. & Chastel, C. *Spiroplasma* (Mollicutes: Spiroplasmataceae) pathogenic for *Aedes aegypti* and *Anopheles stephensi* (Diptera: Culicidae). *J. Med. Entomol.* **28**, 219–222 (1991).

39. Wang, W. et al. A spiroplasma associated with tremor disease in the Chinese mitten crab (*Eriocheir sinensis*). *Microbiology* **150**, 3035–3040 (2004).

40. Wang, W. et al. A novel *Spiroplasma* pathogen causing systemic infection in the crayfish *Procambarus clarkii* (Crustacea: Decapod), in China. *FEMS Microbiol. Lett.* **249**: 131–137 (2005).

41. Wang, W., et al. *Spiroplasma eriocheiris* sp. nov., associated with mortality in the Chinese mitten crab, *Eriocheir sinensis*. *Int. J. Syst. Evol. Microbiol.*, **61**, 703–708 (2011).

42. Nunan, L. M., Lightner, D. V., Oduori, M. A. & Gasparich, G. E. *Spiroplasma penaei* sp. nov., associated with mortalities in *Penaeus vannamei*, Pacific white shrimp. *Int. J. Syst. Evol. Microbiol.* **55**, 2317–2322 (2005).

43. Liang, T. et al. Identification and isolation of a spiroplasma pathogen from diseased freshwater prawns, *Macrobrachium rosenbergii*, in China: A new freshwater crustacean host. *Aquaculture* **318**, 1–6 (2011).

44. Ding, Z. et al. Classification of circulating hemocytes from the red swamp crayfish *Procambarus clarkii* and their susceptibility to the novel pathogen *Spiroplasma eriocheiris in vitro*. *Aquaculture* **356–357,** 371–380 (2012).

45. Williamson, D. L. et al. *Spiroplasma poulsonii* sp. nov., a new species associated with male-lethality in *Drosophila willistoni*, a neotropical species of fruit fly. *Int. J. Syst. Bacteriol.* **49 Pt 2**, 611–618 (1999).

46. Tabata, J. et al. Male Killing and incomplete inheritance of a novel *Spiroplasma* in the moth *Ostrinia zaguliaevi*. *Microb. Ecol.* **61**, 254–263 (2011).

47. Hurst, G. D. & Jiggins, F. M. Male-killing bacteria in insects: mechanisms, incidence, and implications. *Emerg. Infect. Dis.* **6**, 329–336 (2000).

48. Hayashi, M., Watanabe, M., Yukuhiro, F., Nomura, M., Kageyama, D. A nightmare for males? A maternally transmitted male-killing bacterium and strong female bias in a green lacewing population. *PLoS ONE* **11**, e0155794 (2016).

49. Duron, O. *et al.* The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol.* **6,** 27 (2008).

50. Ferree, P. M., Avery, A., Azpurua, J., Wilkes, T. & Werren, J. H. A bacterium targets maternally inherited centrosomes to kill males in *Nasonia*. *Curr. Biol.* **18,** 1409–1414 (2008).

644 51. Zhang, Y. K., Chen, Y. T., Yang, K. & Hong, X.-Y. A review of prevalence and
645 phylogeny of the bacterial symbiont *Cardinium* in mites (subclass: Acari). *Systematic*
646 *and Applied Acarology* **21,** 978–990 (2016).
647 52. Giorgini, M., Monti, M. M., Caprio, E., Stouthamer, R. & Hunter, M. S.
648 Feminization and the collapse of haplodiploidy in an asexual parasitoid wasp
649 harboring the bacterial symbiont *Cardinium*. *Heredity (Edinb)* **102,** 365–371 (2009).
650 53. Hendry, T. A., Hunter, M. S. & Baltrus, D. A. The facultative symbiont *Rickettsia*
651 protects an invasive whitefly against entomopathogenic *Pseudomonas syringae*
652 strains. *Appl. Environ. Microbiol.* **80,** 7161–7168 (2014).
653 54. Liu, Y., Shen, D., Zhou, F., Wang, G. & An, C. Identification of immunity-related
654 genes in *Ostrinia furnacalis* against entomopathogenic fungi by RNA-seq analysis.
655 *PLoS ONE* **9,** e86436 (2014).
656 55. Himler, A. G. *et al.* Rapid spread of a bacterial symbiont in an invasive whitefly is
657 driven by fitness benefits and female bias. *Science* **332,** 254–256 (2011).
658 56. Chen, X., Li, S. & Aksoy, S. Concordant evolution of a symbiont with its host
659 insect species: molecular phylogeny of genus *Glossina* and its bacteriome-associated
660 endosymbiont, *Wigglesworthia glossinidia*. *J. Mol. Evol.* **48**, 49–58 (1999).
661 57. Lo, W.-S., Chen, L.-L., Chung, W.-C., Gasparich, G. E. & Kuo, C.-H.
662 Comparative genome analysis of *Spiroplasma melliferum* IPMB4A, a honeybee-
663 associated bacterium. *BMC Genomics* **14,** 22 (2013).
664 58. Saillard, C. *et al. Spiroplasma phoeniceum* sp. nov., a new plant-pathogenic
665 species from Syria. *Int. J. Syst. Evol. Microbiol.* **37,** 106–115 (1987).
666 59. Hackett, K. J. *et al. Spiroplasma insolitum* sp. nov., a new species of group I
667 *Spiroplasma* with an unusual DNA base composition. *Int. J. Syst. Evol. Microbiol.* **43,**
668 272–277 (1993).

669 60. Koerber, R. T., Gasparich, G. E., Frana, M. F. & Grogan, W. L. *Spiroplasma*
670 *atrichopogonis* sp. nov., from a ceratopogonid biting midge. *Int. J. Syst. Evol.*
671 *Microbiol.* **55,** 289–292 (2005).
672 61. Geiger, A., Ponton, F. & Simo, G. Adult blood-feeding tsetse flies, trypanosomes,
673 microbiota and the fluctuating environment in sub-Saharan Africa. *ISME J.* **9**, 1496–
674 1507 (2015).
675 62. Wamwiri, F. N. et al. *Wolbachia*, *Sodalis* and trypanosome co-infections in
676 natural populations of *Glossina austeni* and *Glossina pallidipes*. *Parasit. Vectors* **6**,
677 232 (2013).
678 63. Balmand, S., Lohs, C., Aksoy, S., Heddi, A. Tissue distribution and transmission
679 routes for the tsetse fly endosymbionts. *J. Invertebr. Pathol.* **112**, S116–S122 (2013).
680 64. Vooght, L. D., Caljon, G., Hees, J. V. & Abbeele, J. V. D. Paternal transmission
681 of a secondary symbiont during mating in the viviparous tsetse fly. *Mol. Biol. Evol.*
682 **32**, 1977-1980 (2015).
683 65. Attardo, G. M. et al. Analysis of milk gland structure and function in *Glossina*
684 *morsitans*: Milk protein production, symbiont populations and fecundity. *J. Insect*
685 *Physiol.* **54**, 1236–1242 (2008).
686 66. Solano, P. et al. Do tsetse flies only feed on blood? *Infect. Genet. Evol.* **36**, 184–
687 189 (2015).
688 67. Dennis, J. W. *Sodalis glossinidius* prevalence and trypanosome presence in tsetse
689 from Luambe National Park, Zambia. *Parasit. Vectors* **7**, 378 (2014).
690 68. Geiger, A., Cuny, G. & Frutos, R. Two Tsetse Fly Species*, Glossina palpalis*
691 *gambiensis* and *Glossina morsitans morsitans*, carry genetically distinct populations
692 of the secondary symbiont *Sodalis glossinidius*. *Appl. Environ. Microbiol.* **71**, 8941–
693 8943 (2005).

694 69. Mediannikov, O., Audoly, G., Diatta, G., Trape, J.-F. & Raoult, D. New *Rickettsia*
695 sp. in tsetse flies from Senegal. *Comp. Immunol. Microbiol. Infect. Dis.* **35,** 145–150
696 (2012).
697 70. Schneider, D. I., Klasson, L., Lind, A. E. & Miller, W. J. More than fishing in the
698 dark: PCR of a dispersed sequence produces simple but ultrasensitive *Wolbachia*
699 detection. *BMC Microbiol.* **14**, 121 (2014).
700 71. Doudoumis, V. et al. Tsetse-*Wolbachia* symbiosis: Comes of age and has great
701 potential for pest and disease control. *J. Invertebr. Pathol.* **112**, S94–103 (2013).
702 72. Charlat, S. et al. Competing selfish genetic elements in the butterfly *Hypolimnas*
703 *bolina*. *Curr. Biol.* **16**, 2453–2458 (2006).
704 73. Goto, S., Anbutsu, H. & Fukatsu, T. Asymmetrical interactions between
705 *Wolbachia* and *Spiroplasma* endosymbionts coexisting in the same insect host. *Appl.*
706 *Environ. Microbiol.* **72**, 4805–4810 (2006).
707 74. Rossi, P. et al. Mutual exclusion of *Asaia* and *Wolbachia* in the reproductive
708 organs of mosquito vectors. *Parasit. Vectors* **8,** 278 (2015).
709 75. Hughes, G. L. et al. Native microbiome impedes vertical transmission of
710 *Wolbachia* in *Anopheles mosquitoes*. *Proc. Natl. Acad. Sci. USA* **111**, 12498–12503
711 (2014).
712 76. Alam, U. et al. Implications of microfauna-host interactions for trypanosome
713 transmission dynamics in *Glossina fuscipes fuscipes* in Uganda. *Appl. Environ.*
714 *Microbiol.* **78**, 4627–4637 (2012).
715 77. Saglio, P. et al. *Spiroplasma citri* gen. and sp. n.: A Mycoplasma-Like organism
716 associated with 'Stubborn' disease of citrus. *Int. J. Syst. Evol. Microbiol.* **23**, 191–204
717 (1973).

718 78. Anbutsu, H. & Fukatsu, T. *Spiroplasma* as a model insect endosymbiont. *Environ.*
719 *Microbiol. Rep.* **3**: 144–153 (2011).
720 79. Kageyama, D. et al. Prevalence of a non-male-killing spiroplasma in natural
721 populations of *Drosophila hydei*. *Appl. Environ. Microbiol.* **72**, 6667–6673 (2006).
722 80. Teixeira, L., Ferreira, Á. & Ashburner, M. The bacterial symbiont *Wolbachia*
723 induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol.* **6**,
724 e-1000002 (2008).
725 81. Hedges, L. M., Brownlie, J. C., O'Neill, S. L. & Johnson, K. N. *Wolbachia* and
726 virus protection in insects. *Science* **322**: 702–702 (2008).
727 82. Martinez, J. et al. Symbionts commonly provide broad spectrum resistance to
728 viruses in insects: A comparative analysis of *Wolbachia* strains. *PLoS Pathog.* **10**, e-
729 1004369 (2014).
730 83. Stevanovic, A. L., Arnold, P. A. & Johnson, K. N. *Wolbachia*-mediated antiviral
731 protection in Drosophila larvae and adults following oral infection. *Appl. Environ.*
732 *Microbiol.* **81**, 8215-8223 (2015).
733 84. Shokal, U. et al. Effects of co-occurring *Wolbachia* and *Spiroplasma*
734 endosymbionts on the Drosophila immune response against insect pathogenic and
735 non-pathogenic bacteria. *BMC Microbiol.* **16**, 16 (2016).
736 85. Damiani, C. et al. Paternal transmission of symbiotic bacteria in malaria vectors.
737 *Curr. Biol.* **18**, R1087–R1088 (2008).
738 86. Van den Abbeele, J. et al. Enhancing tsetse fly refractoriness to trypanosome
739 infection – A new IAEA coordinated research project. *J. Invertebr. Pathol.* **112,**
740 **Supplement 1**, S142–S147 (2013).
741 87. Mancini, M.V. et al. Paratransgenesis to control malaria vectors: a semi-field pilot
742 study. *Parasit. Vectors* **9**, 140 (2016).

743 88. Xie, J., Butler, S., Sanchez, G. & Mateos, M. Male killing *Spiroplasma* protects
744 *Drosophila melanogaster* against two parasitoid wasps. *Heredity* **112,** 399–408
745 (2014).
746 89. Doyle, J.J. & Doyle, J.L. Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–
747 15 (1990).
748 90. Tsiamis, G. et al. Olive-mill wastewater bacterial communities display a cultivar
749 specific profile. *Curr. Microbiol.* **64**, 197–203 (2012).
750 91. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high
751 throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).
752 92. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTALW: improving the
753 sensitivity of progressive multiple sequence alignment through sequence weighting,
754 position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**,
755 4673–4680 (1994).
756 93. Wang, J., Brelsfoard, C., Wu, Y. & Aksoy, S. Intercommunity effects on
757 microbiome and GpSGHV density regulation in tsetse flies. *Journal of Invertebrate*
758 *Pathology* 112, S32–S39 (2013).
759

760 **Acknowledgments**

761 This research was partially supported by: (a) FAO/IAEA Research Contract No
762 17752, (b) GRST grant MIS423722 “Detection and characterization of the role of
763 symbiotic bacteria in tsetse flies under mass-rearing conditions”, and (c) by the
764 Slovak Research and Development Agency under the contract No. APVV-15-0604.
765
766
767

Author Contributions

Conceived and design the study: AD, KB, GT. Conducted the experiments and analysed the results: DV, BF, AS, AA, DA, GI, SP, RB, TP, MS, PA, AbA, GT. Drafted the manuscript: DV, BF, DA, KB, GT. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper

Competing financial interests: The authors declare no competing financial interests.

Figure Legends

Figure 1 a. Relative abundance of *Klebsiella, Rickettsia, Sodalis*, and *Spiroplasma* in whole wild tsetse flies. (Gmed: *G. medicorum*; Gms: *G. morsitans submorsitans*; Gpg: *G. p. gambiensis*; Gt: *G. tachinoides*). **b.** Weighted Unifrac Principal Component Analysis of 16S *rRNA* gene MiSeq data. Each data point represents an individual tsetse fly and is coloured according to *Glossina* species. Convergent evolution between the primary endosymbiont *Wigglesworthia* and its host due to direct vertical transmission generates a tsetse species-clustering pattern that simplifies the detection of emergent taxa such as *Spiroplasma* and *Rickettsia*. All gut samples originated from the study by Aksoy et al. in 2014, and whole samples were collected in Burkina Faso. (Gmm: *Glossina morsitans morsitans*; Gff: *G. fuscipes fuscipes*; Gmed: *G. medicorum*; Gms: *G. morsitans submorsitans*; Gpal: *G. pallidipes*; Gpg: *G. palpalis gambiensis*; Gt: *G. tachinoides*)

Figure 2. *Sodalis* relative abundance in each tsetse species. Boxes denote the interquartile range, the line within the box is the median, and whiskers extend to the most extreme value within 1.5*interquartile range. Outliers are indicated as circles. Gff: *G. fuscipes fuscipes* (n=76); Gmed: *G. medicorum* (n=8); Gmm: *G. morsitans morsitans* (n=6); Gms: *G. morsitans submorsitans* (n=8); Gpal: *G. pallidipes* (n=42); Gpg: *G. p. gambiensis* (n=8); Gt: *G. tachinoides* (n=8).

Figure 3 a. Principal coordinate analysis (PCoA) of bacterial communities based on relative abundances of OTUs with ordinations from laboratory populations of gut, ovaries, testes and larvae. Variance explained by each PCoA axis is given in parentheses. **b.** Canonical analysis of principal coordinates (CAP) ordinations of gonadal bacterial communities based on relative abundances of OTUs from the laboratory populations. The constrained ordinations show maximized differences

among the different *Glossina* species, (Gmm: *Glossina morsitans morsitans*, Gff: *Glossina fuscipes fuscipes*, Gpal: *Glossina pallidipes*). ($trace_{Q\_m'HQ\_m}$ (0.9598; $p < 0.05$)).

Figure 4 Bayesian inference phylogeny based on the 16S *rRNA* gene sequence: The topology resulting from the Maximum Likelihood (ML) method was similar. Bayesian posterior probabilities and ML bootstrap values based on 1000 replicates are given at each node, with the posterior probabilities given first followed by the ML bootstrap values (only values >50% are indicated), respectively. Asterisks indicate support values lower than 50%. The *Spiroplasma* strains present in *Gff* and *Gt* are indicated in bold letters. For each *Spiroplasma* species the GenBank accession number is given to the left of the name.

Figure 5 Quantification of *Spiroplasma* titre in terms of the symbiont *dnaA* gene copies normalized by the tsetse *β-tubulin* gene. (**a**) *Gff* gut from larvae, male and female teneral and 15-day old tsetse flies (n = 3, each sample is a pool of five) $p < 0.005$, (**b**) gonads from male and female teneral tsetse flies (n = 3, each sample is a pool of five), $p < 0.05$ (Anova test was performed; statistical significant differences are indicated with an asterisk *).

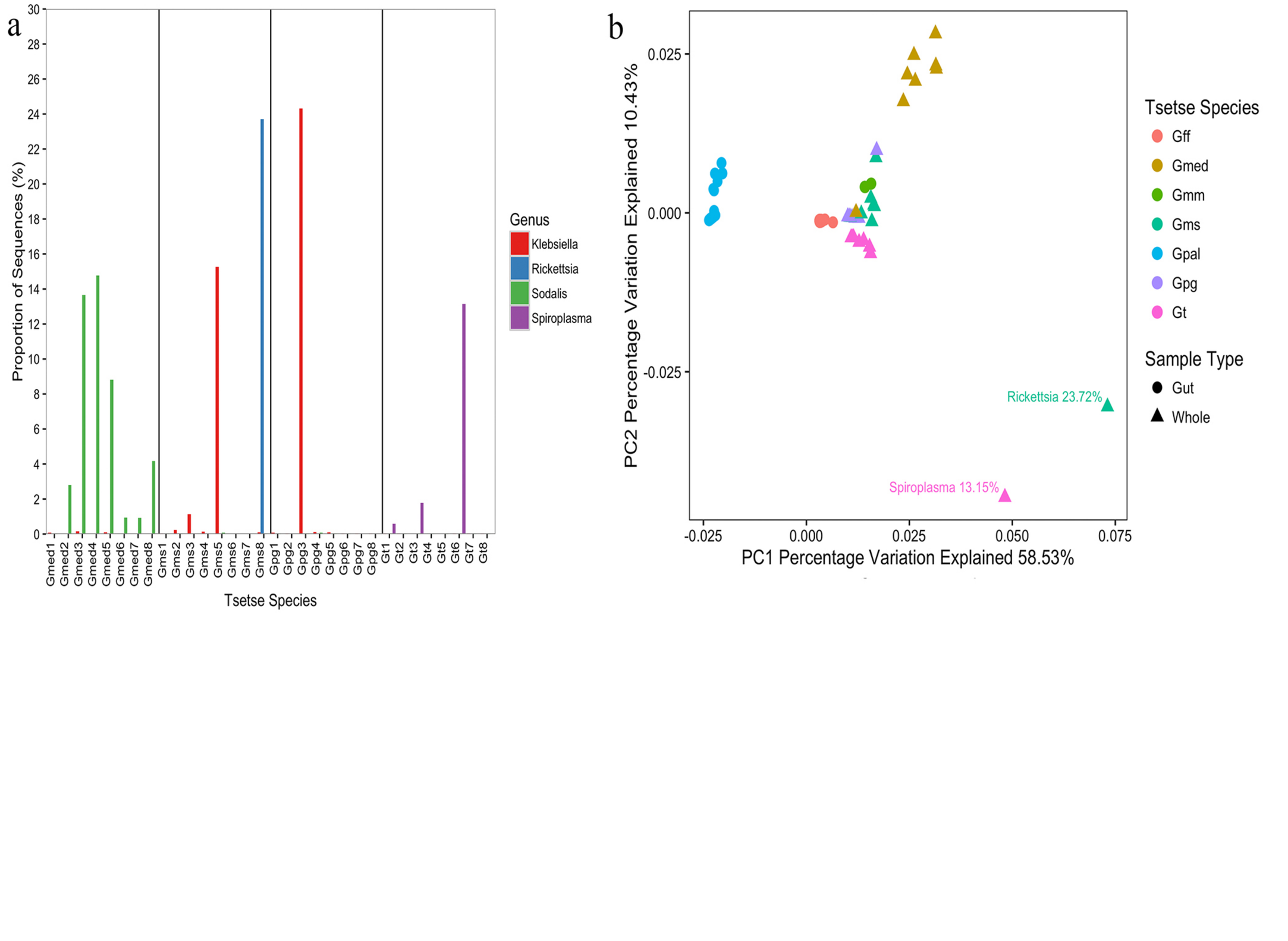
Figure 6 Quantification of *Spiroplasma* titre as *Spiroplasma dnaA* gene copy number normalized to the tsetse β-tubulin gene. (**a**) *Gff* whole insects from healthy/live males and prematurely dead males from the mass-rearing facility in Ethiopia (n = 6), (**b**) *Gff* whole insects from healthy/live females and prematurely dead females from the mass-rearing facility in Ethiopia (n = 9), $p < 0.05$. (**c**) *Gff* whole insects from healthy/live females carrying a larvae and prematurely dead females carrying a larva from the mass-rearing facility in Ethiopia (n = 6), $p < 0.05$. (ANOVA test was performed; statistical significant differences are indicated with an asterisk *).

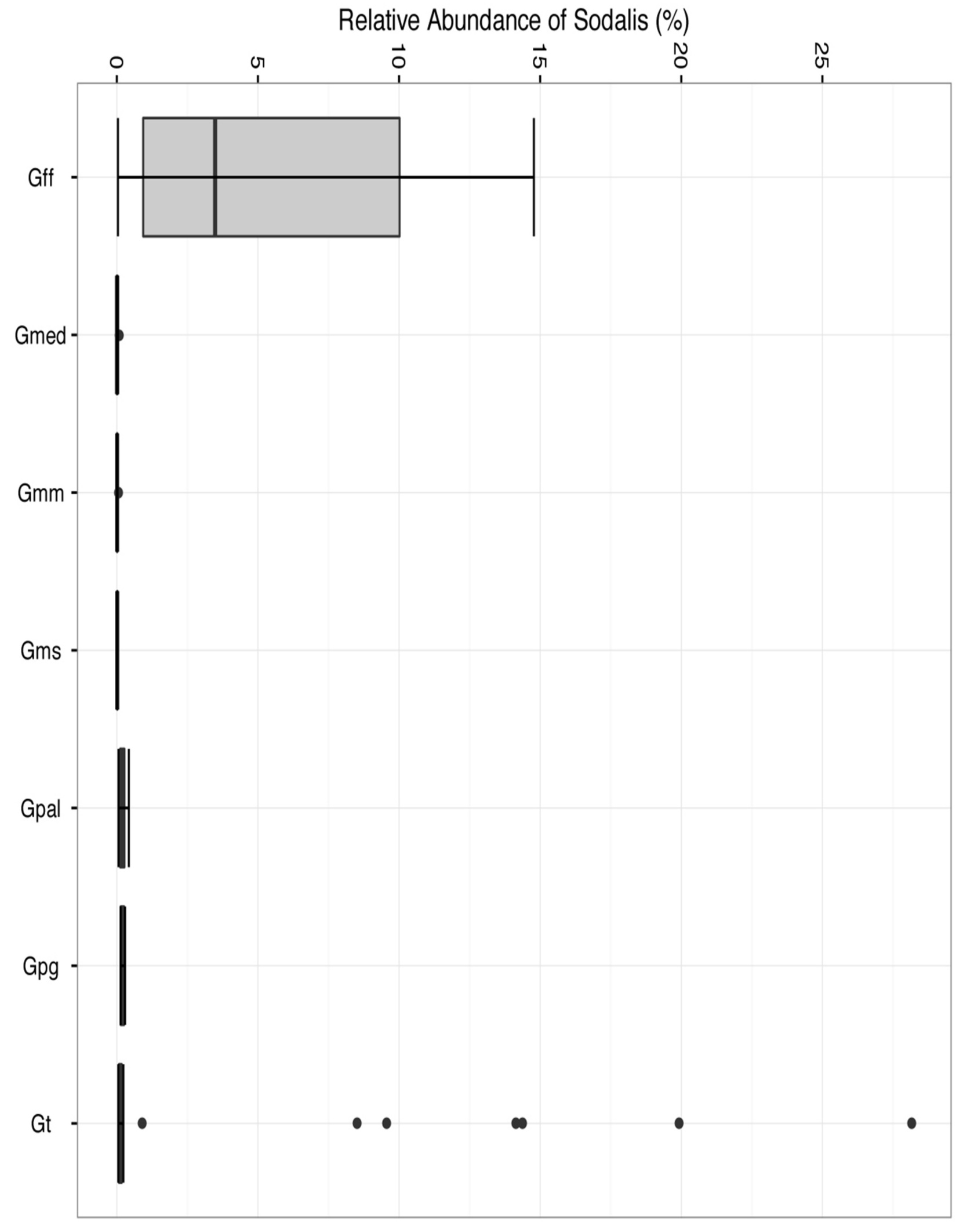
Figure 7 Localization of *Spiroplasma* in the male and female reproductive system of *Gff*. In fluorescent *in situ* hybridization (FISH) images blue and yellow indicate insect nuclear DNA and *Spiroplasma* respectively. **(a)** FISH on dissected ovaries (O), **(b)** FISH on dissected testes (T) with an inset showing a testis at a higher magnification.

Table 1 *Spiroplasma* prevalence in ten *Glossina* species

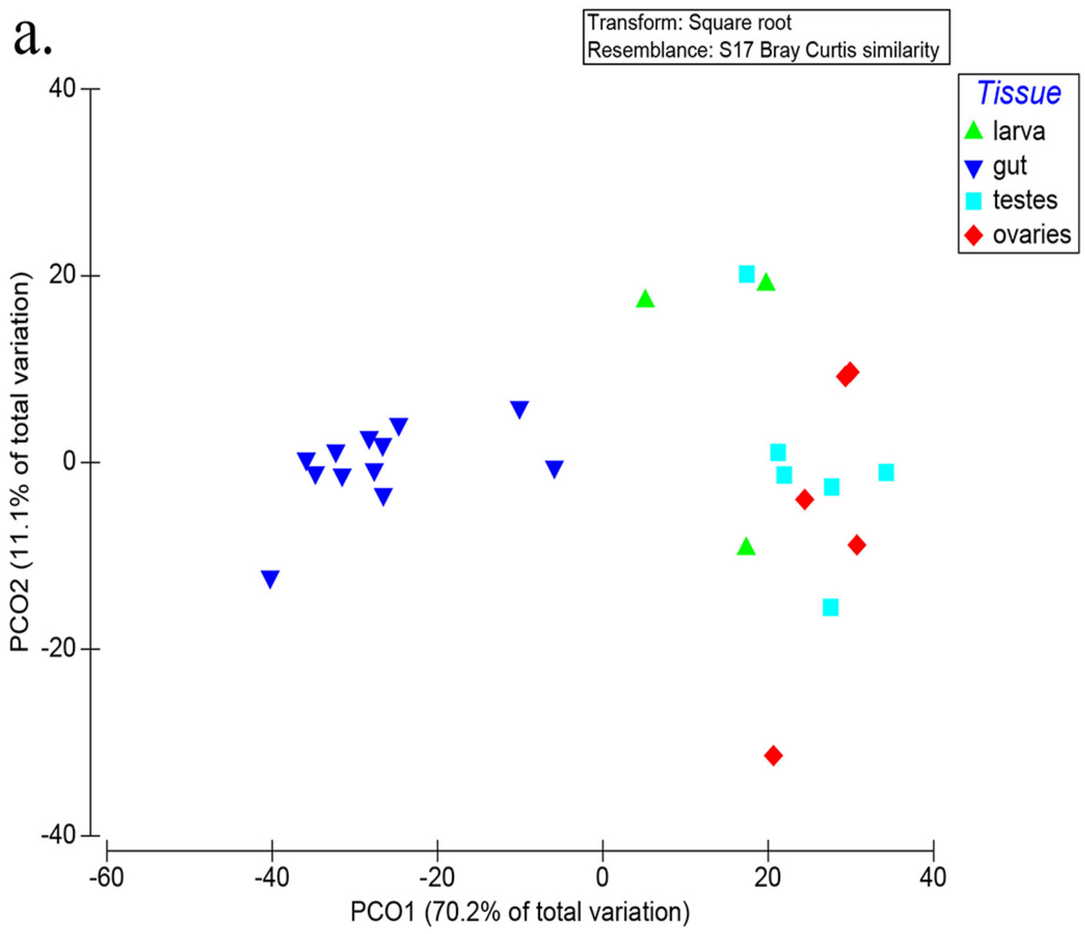
Species	Origin	Collection Date	Location (Area, Population, Sex)	Tissue	No. of Samples	Spiroplasma Infection Rate (%)
G. austeni	Field	1995	Tanzania (Zanzibar) 6 Males, 4 NA[1])	Whole	10	0
	Field	1996	Tanzania (Jozani) Females	Whole	10	0
	Field	1999	South Africa (Zululand) 3 Females, 4 Males, 3 NA	Whole	10	0
	Field	Unknown	Coastal Tanzania (Muhoro) Female	Whole	2	0
G. brevipalpis	Laboratory	1995	Seibersdorf Laboratory Colony, 8 Females and 8 Males	Whole	16	0
	Laboratory	Unknown	Coastal Tanzania (Pangani), Males	Whole	5	0
G. f. fuscipes	Field	1994	Uganda (Buvuma Island, GFTF2), NA	Whole	17	0
	Field	1994	Uganda (Buvuma Island, GFKF2), NA	Whole	5	0
	Field	1994	Uganda (Buvuma Island, GFFBUV2), NA	Whole	9	0
	Field	1994	Uganda (Buvuma Island, GFFTOR2)[2], NA	Whole	15	6.7
	Laboratory	1995	Seibersdorf Laboratory Colony[3], 18 Females, 18 Males	Whole	36	33.4
	Laboratory	2013	Bratislava Laboratory Colony[4], 20 Females, 20 Males	Whole	40	80
	Field	2014	Uganda (Lukoma-Buvuma Islands, 350)[4] 20 Females, 32 Males	Whole	52	5.8
G. m. centralis	Laboratory	2008	Yale Laboratory Colony, NA	Whole	1	0
G. m. morsitans	Laboratory	2008	KARI-TRC Laboratory Colony, NA	Whole	15	0
	Laboratory	2010	Antwerp Laboratory Colony, NA	Whole	4	0
G. m. submorsitans	Field	2010	Burkina Laboratory (Folonzo), Females	Whole	8	0
G. pallidipes	Laboratory	1999	Seibersdorf Laboratory Colony, NA	Whole	2	0
	Laboratory	2008	Seibersdorf Laboratory Colony, 6 Females, 7 Males	Whole	13	0
	Laboratory	Unknown	Uganda-UGA/IAEA, Males	Whole	5	0
G. p. gambiensis	Laboratory	1995	CIRDES Laboratory Colony, 4 Female, 5 Males	Whole	9	0
	Laboratory	2005	CIRDES Laboratory Colony, 1 Females, 9 Males	Whole	10	0
G. p. palpalis	Laboratory	1995	Seibersdorf Laboratory Colony[2], 8 Females, 8 Males	Whole	16	12.5
G. tachinoides	Laboratory	1995	Seibersdorf Laboratory Colony[2], Females	Whole	7	14.3
	Field	2010	Burkina Faso (Folonzo)[4], Females	Whole	8	37.5

[1]Sex of individuals is not known [2] Characterization of *Spiroplasma* infection was based only on 16S *rRNA* gene sequencing. [3]The Seibersdorf laboratory-colony was established from the Central African Republic in 1986. This colony was transferred to Bratislava, Slovakia in 2009. [4] Full MLST genotyping

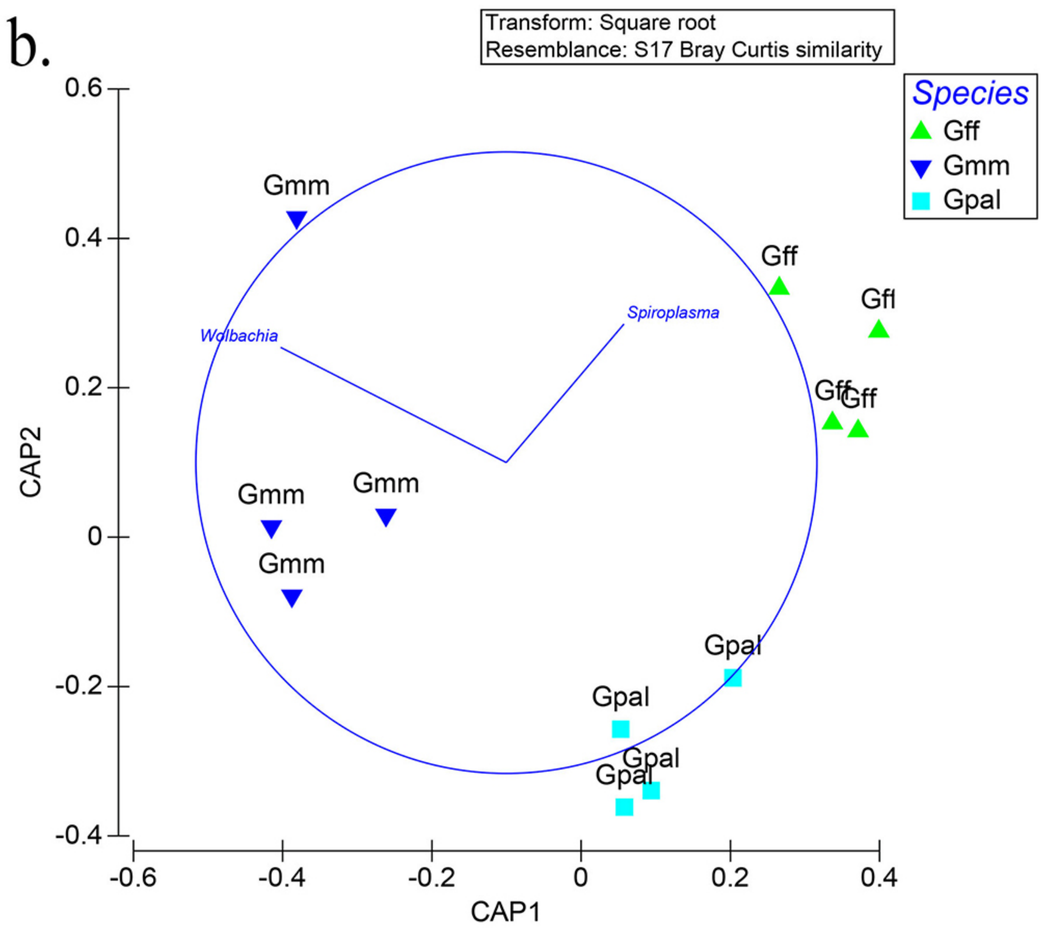

a
Proportion of Sequences (%)
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
Gmed1
Gmed2
Gmed3
Gmed4
Gmed5
Gmed6
Gmed7
Gmed8
Gms1
Gms2
Gms3
Gms4
Gms5
Gms6
Gms7
Gms8
Gpg1
Gpg2
Gpg3
Gpg4
Gpg5
Gpg6
Gpg7
Gpg8
Gt1
Gt2
Gt3
Gt4
Gt5
Gt6
Gt7
Gt8
Tsetse Species
Genus
Klebsiella
Rickettsia
Sodalis
Spiroplasma
b
PC2 Percentage Variation Explained 10.43%
0.025
0.000
-0.025
Rickettsia 23.72%
Spiroplasma 13.15%
-0.025
0.000
0.025
0.050
0.075
PC1 Percentage Variation Explained 58.53%
Tsetse Species
Gff
Gmed
Gmm
Gms
Gpal
Gpg
Gt
Sample Type
Gut
Whole

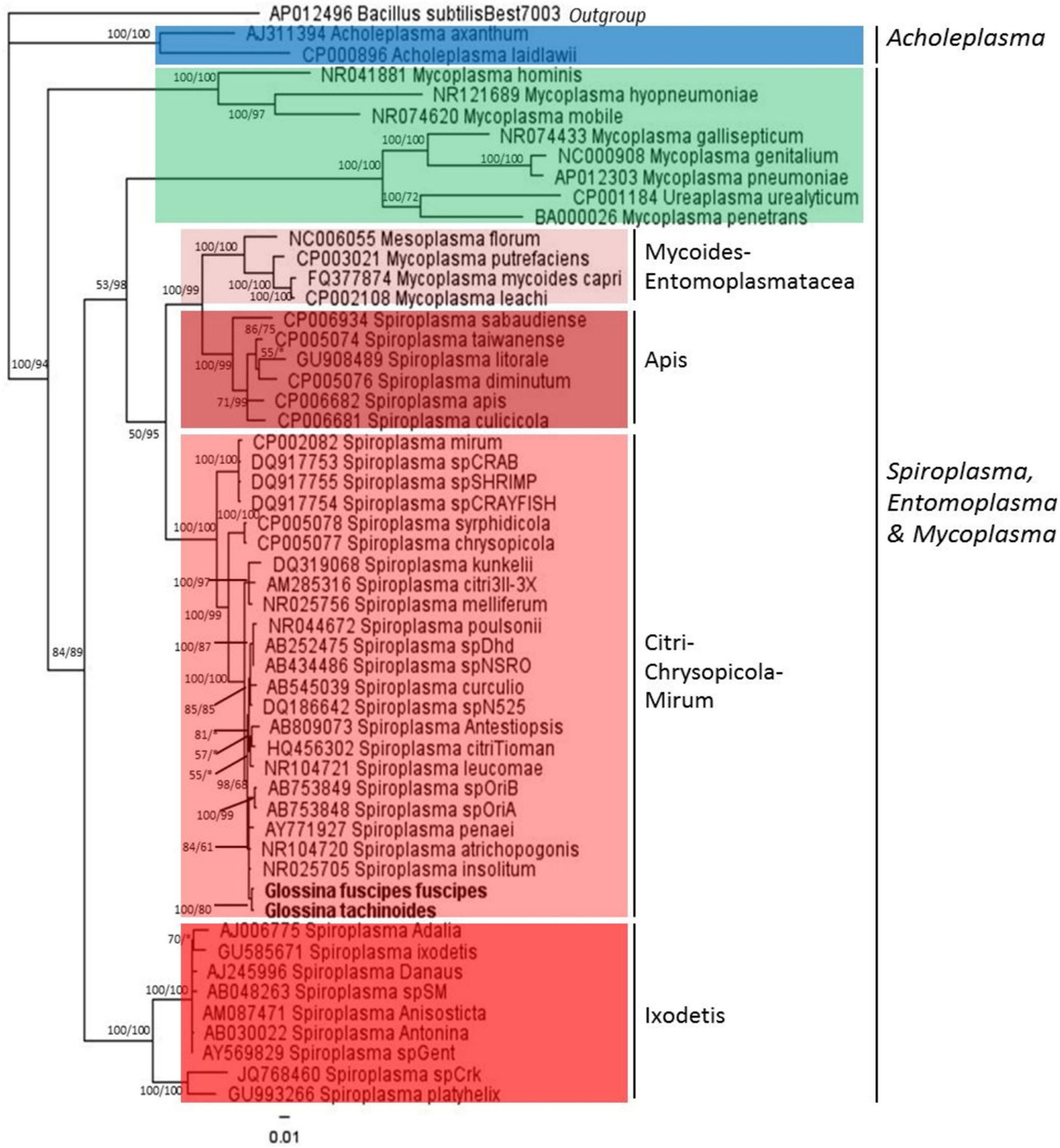

Relative Abundance of Sodalis (%)
0
5
10
15
20
25
Gff
Gmed
Gmm
Gms
Gpal
Gpg
Gt

a.

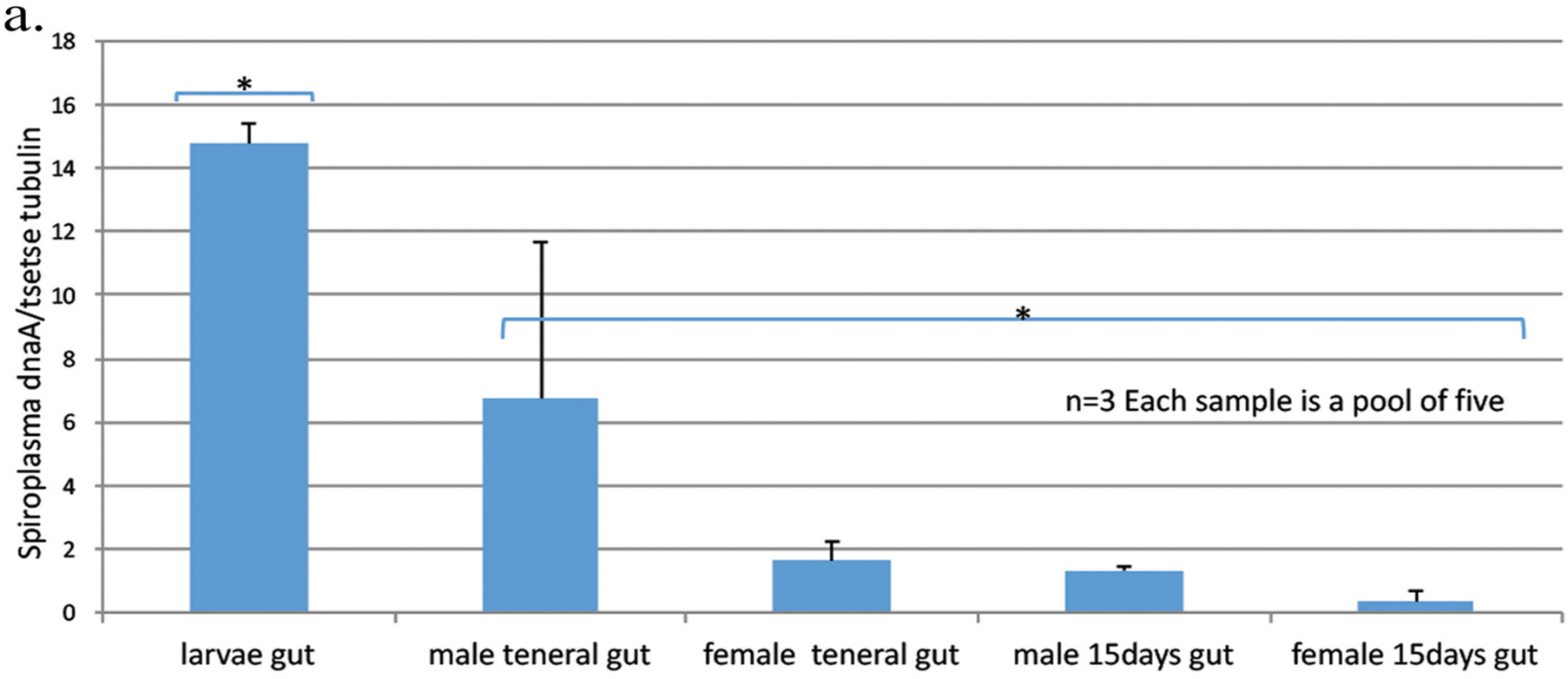

Transform: Square root
Resemblance: S17 Bray Curtis similarity
Tissue
larva
gut
testes
ovaries
PCO2 (11.1% of total variation)
PCO1 (70.2% of total variation)
40
20
0
-20
-40
-60
-40
-20
0
20
40

b.

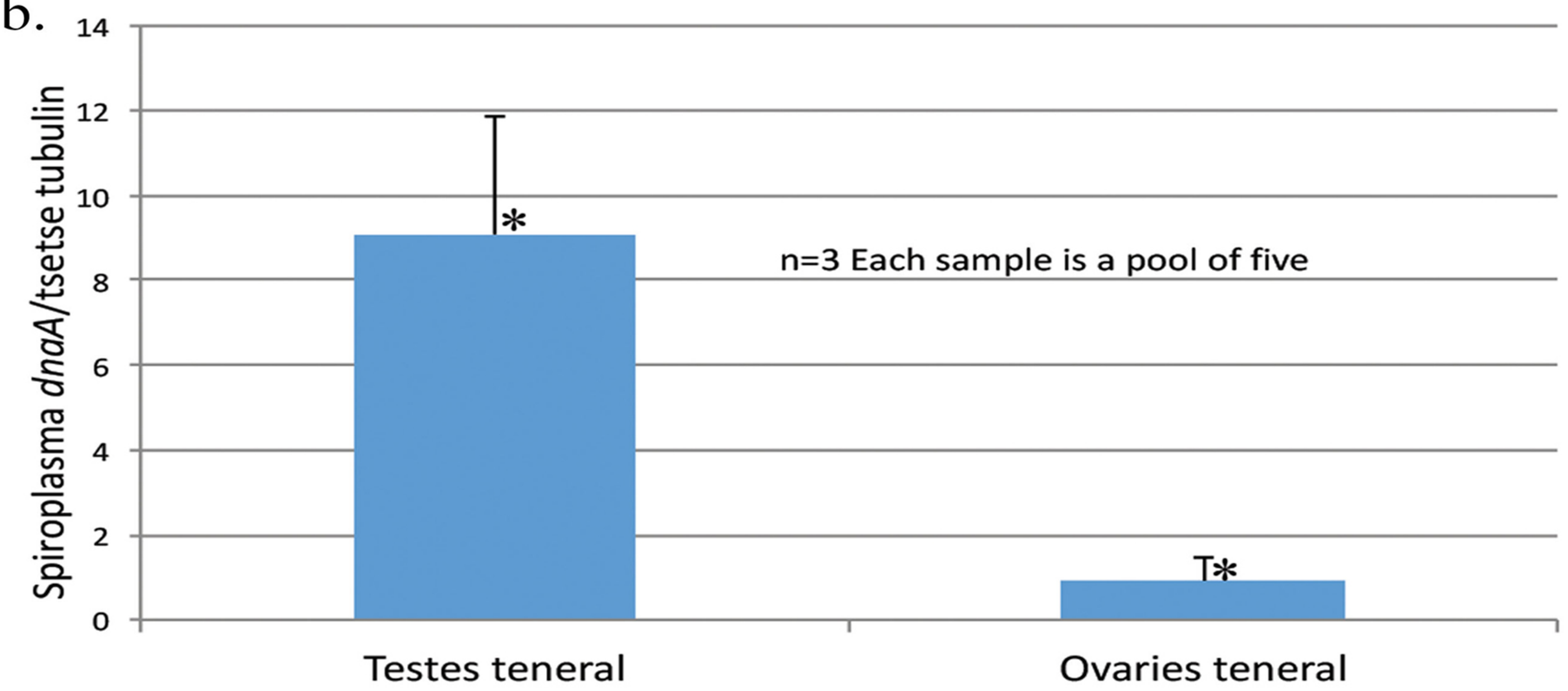

Transform: Square root
Resemblance: S17 Bray Curtis similarity
Species
Gff
Gmm
Gpal
Wolbachia
Spiroplasma
Gmm
Gff
Gfl
Gff
Gff
Gmm
Gmm
Gmm
Gpal
Gpal
Gpal
Gpal
CAP2
CAP1
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.4
-0.2
0
0.2
0.4


AP012496 Bacillus subtilisBest7003 Outgroup
AJ311394 Acholeplasma axanthum
CP000896 Acholeplasma laidlawii
Acholeplasma
NR041881 Mycoplasma hominis
NR121689 Mycoplasma hyopneumoniae
NR074620 Mycoplasma mobile
NR074433 Mycoplasma gallisepticum
NC000908 Mycoplasma genitalium
AP012303 Mycoplasma pneumoniae
CP001184 Ureaplasma urealyticum
BA000026 Mycoplasma penetrans
NC006055 Mesoplasma florum
CP003021 Mycoplasma putrefaciens
FQ377874 Mycoplasma mycoides capri
CP002108 Mycoplasma leachi
Mycoides-
Entomoplasmatacea
CP006934 Spiroplasma sabaudiense
CP005074 Spiroplasma taiwanense
GU908489 Spiroplasma litorale
CP005076 Spiroplasma diminutum
CP006682 Spiroplasma apis
CP006681 Spiroplasma culicicola
Apis
CP002082 Spiroplasma mirum
DQ917753 Spiroplasma spCRAB
DQ917755 Spiroplasma spSHRIMP
DQ917754 Spiroplasma spCRAYFISH
CP005078 Spiroplasma syrphidicola
CP005077 Spiroplasma chrysopicola
DQ319068 Spiroplasma kunkelii
AM285316 Spiroplasma citri3II-3X
NR025756 Spiroplasma melliferum
NR044672 Spiroplasma poulsonii
AB252475 Spiroplasma spDhd
AB434486 Spiroplasma spNSRO
AB545039 Spiroplasma curculio
DQ186642 Spiroplasma spN525
AB809073 Spiroplasma Antestiopsis
HQ456302 Spiroplasma citriTioman
NR104721 Spiroplasma leucomae
AB753849 Spiroplasma spOriB
AB753848 Spiroplasma spOriA
AY771927 Spiroplasma penaei
NR104720 Spiroplasma atrichopogonis
NR025705 Spiroplasma insolitum
Glossina fuscipes fuscipes
Glossina tachinoides
Citri-
Chrysopicola-
Mirum
AJ006775 Spiroplasma Adalia
GU585671 Spiroplasma ixodetis
AJ245996 Spiroplasma Danaus
AB048263 Spiroplasma spSM
AM087471 Spiroplasma Anisosticta
AB030022 Spiroplasma Antonina
AY569829 Spiroplasma spGent
JQ768460 Spiroplasma spCrk
GU993266 Spiroplasma platyhelix
Ixodetis
Spiroplasma,
Entomoplasma
& Mycoplasma
0.01

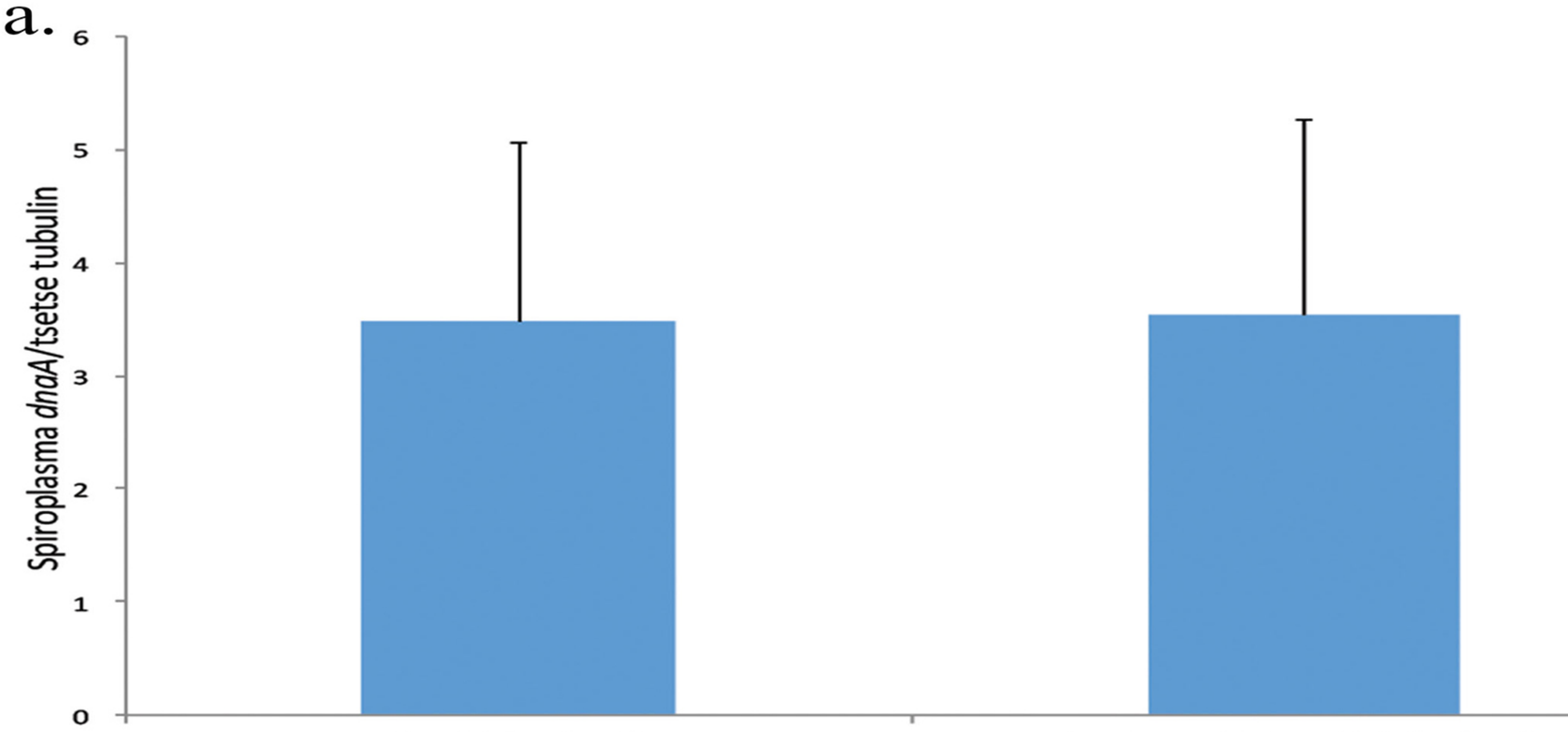
a.


Spiroplasma dnaA/tsetse tubulin
18
16
14
12
10
8
6
4
2
0
*
*
n=3 Each sample is a pool of five
larvae gut
male teneral gut
female teneral gut
male 15days gut
female 15days gut

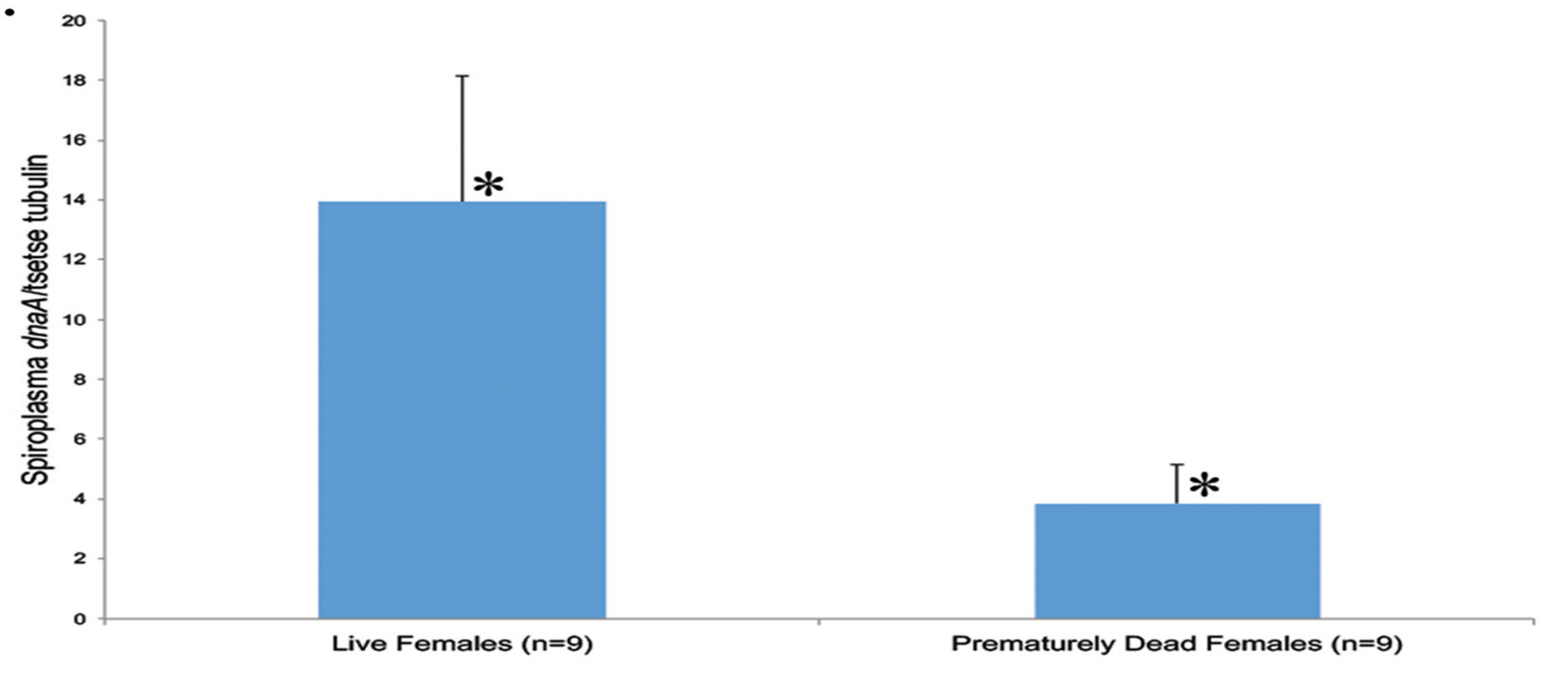
b.


Spiroplasma dnaA/tsetse tubulin
14
12
10
8
6
4
2
0
*
T*
n=3 Each sample is a pool of five
Testes teneral
Ovaries teneral

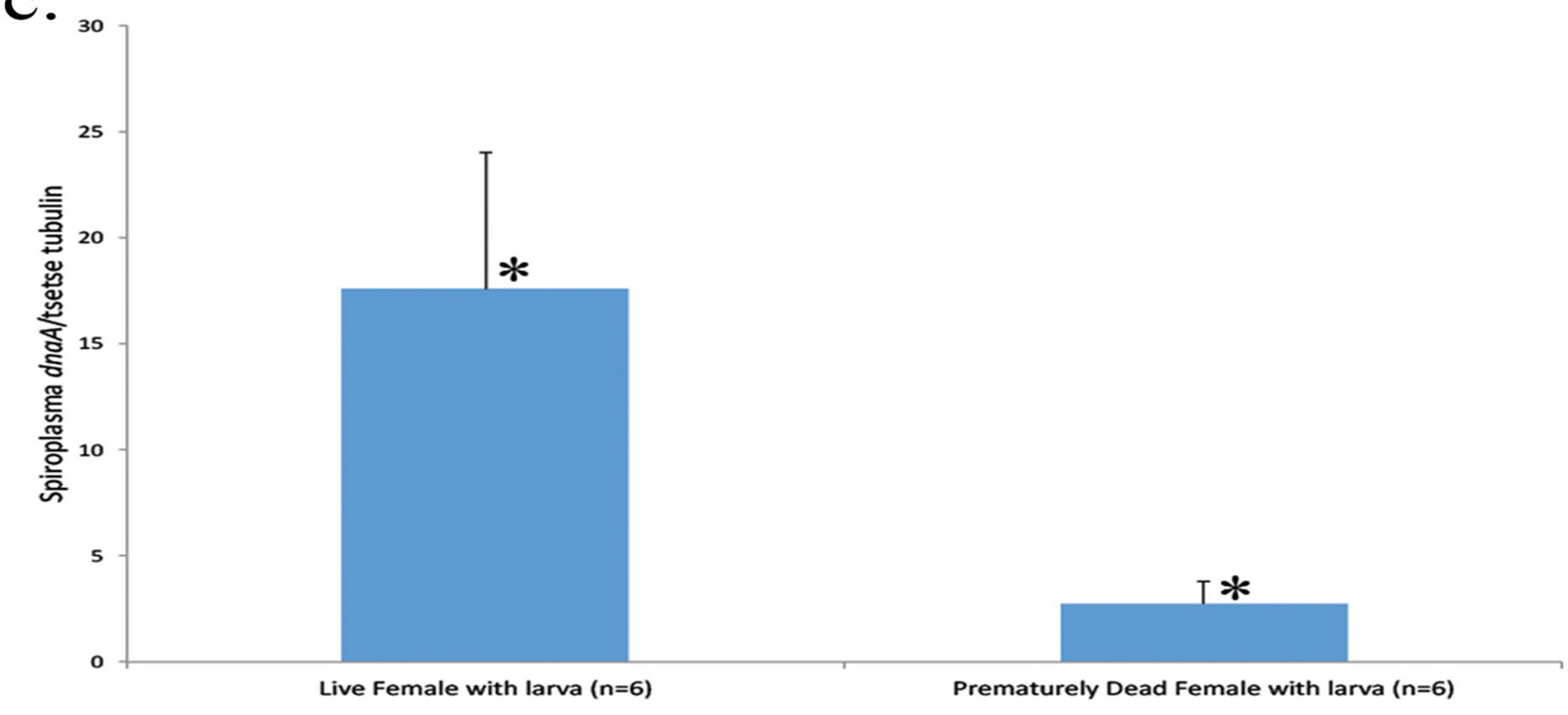
a.

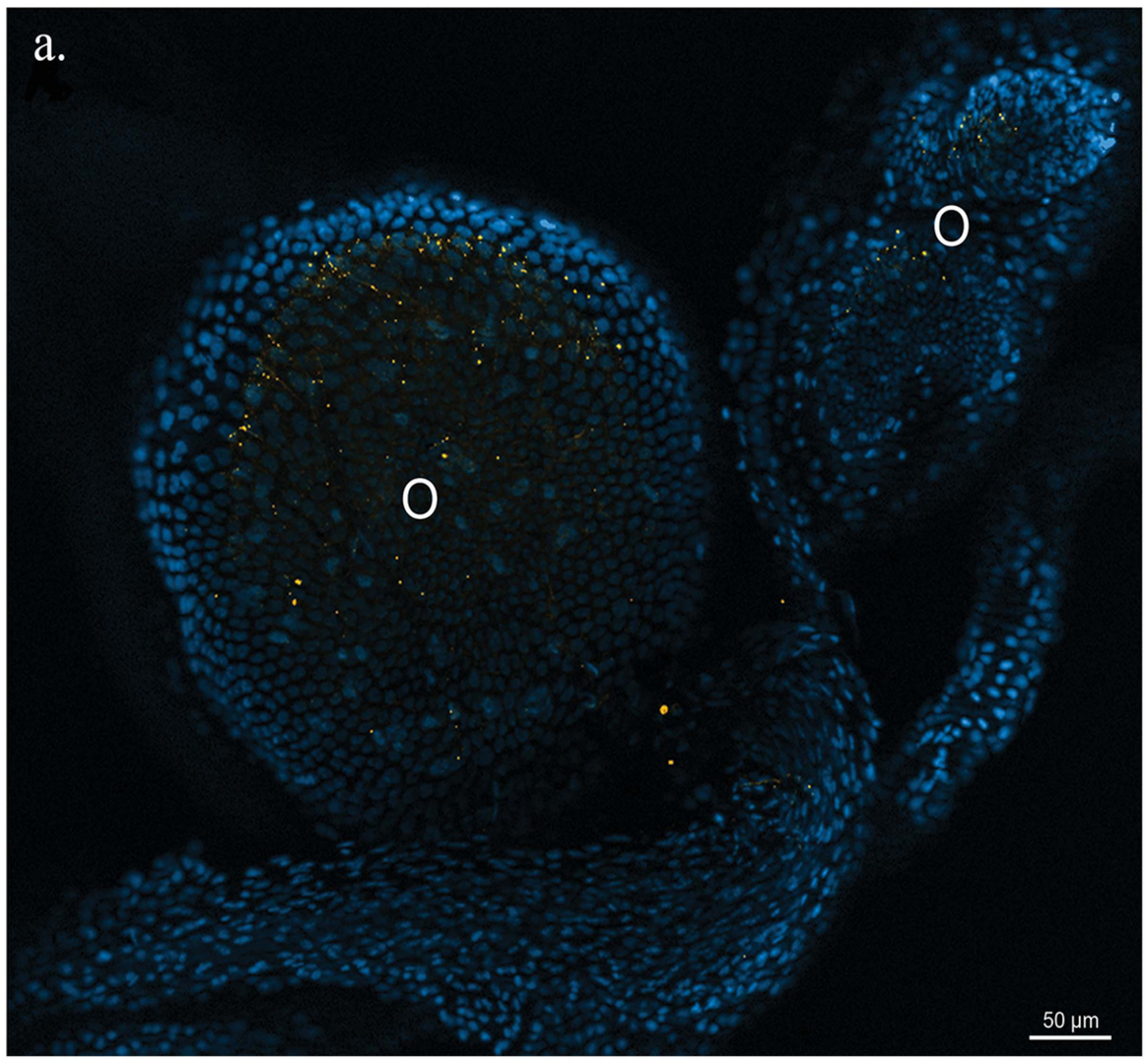

Spiroplasma dnaA/tsetse tubulin
0
1
2
3
4
5
6
Live Males (n=6)
Prematurely Dead Males (n=6)

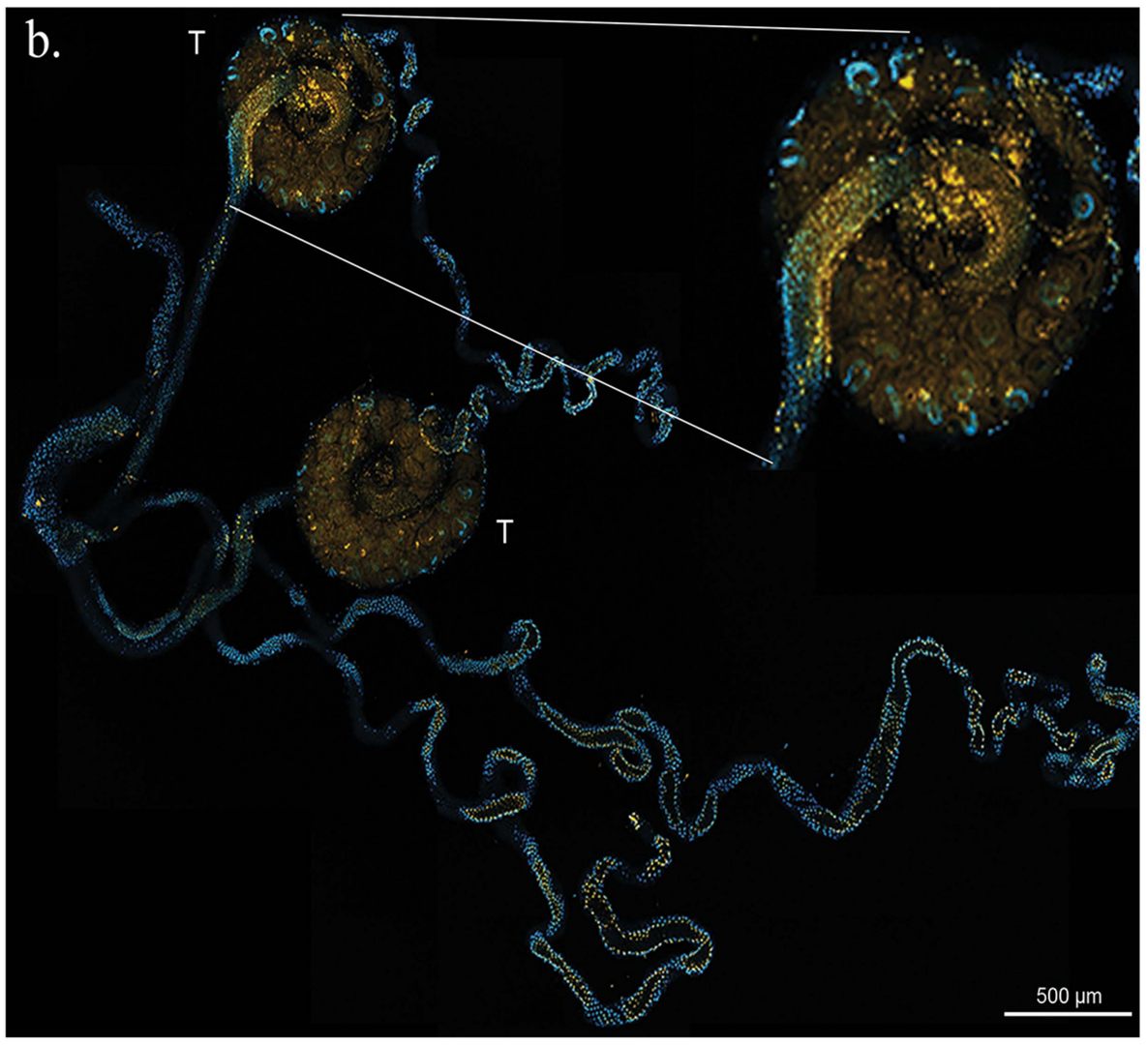
b.


Spiroplasma dnaA/tsetse tubulin
0
2
4
6
8
10
12
14
16
18
20
*
*
Live Females (n=9)
Prematurely Dead Females (n=9)

c.


Spiroplasma dnaA/tsetse tubulin
0
5
10
15
20
25
30
*
*
Live Female with larva (n=6)
Prematurely Dead Female with larva (n=6)


a.
O
O
50 µm


b.
T
T
500 µm