TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

DISSERTATION

Population genetics and research on host-symbiont interactions of the tsetse fly (*Glossina* spp.)

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der Naturwissenschaften unter der Leitung von

Univ. Prof. Dr. Robert. L. Mach

E166

Institut für Verfahrenstechnik, Umwelttechnik und technische Biowissenschaften

eingereicht an der Technischen Universität Wien

Fakultät für Technische Chemie

von

Fabian Gstöttenmayer

[redacted]

Wien, 2024

Acknowledgements

I want to extend my sincere gratitude to my supervisor at the Insect Pest Control Laboratory, Prof. Dr. Adly M. M. Abdalla. Your mentorship and guidance have been a cornerstone of my PhD journey, providing me with both professional scientific expertise and encouragement to pursue my interests in this field of research. I am particularly thankful for your open attitude to explore new ideas and methods, as well as pushing me further to achieve important milestones for my doctorate and my scientific career. Moreover, I am truly thankful for the supervision of Prof. Dr. Robert L. Mach from the Technical University of Vienna, the consistent guidance and thoughtful advice that has proven to be essential for the progress of my studies.

A heartfelt thank you goes to Dr. Chantel de Beer for the fruitful discussions and guidance on tsetse fly rearing and development of experimental protocols to precisely address the biological questions of my dissertation. My gratitude also goes to Dr. Marc J.B. Vreysen and Dr. Polychronis Rempoulakis, the laboratory heads of the Insect Pest Control Library, for hosting me during my time at the laboratory and establishing an exceptional working environment. My appreciation furthermore goes to the professional and diverse team that I had the pleasure to be a part of and not only learn about their scientific work, but also to have such rich cultural exchange. From the bottom of my heart, thank you Julija, Carmen, Arooj, Henry, Hasim, Olga, Anibal, Montse, Carlos, Steph, Barro, Tankoano, Herbert, Marilena, Chrysa, Amanda, Antonia, Hager, Giovanni, Germano, Gogo, Emma, Davor, Jhonatan, Lilly, Ina, Witness and Vasilis.

A special mention goes to the PhD dream team Dera, Hannah and Aristide – it was nothing short but incredible to navigate through this journey together with you, supporting each other through the highs and lows and always sharing moments of joy and laughter regardless how difficult our day to day challenges have been. I am truly happy to have shared these experiences with you and it is safe to say we have transformed from colleagues to family. I wish you all the happiness, health and success in your life and I am looking forward to the next time we will meet again. I also want to thank my dear friends and colleagues from the Medical University that I had the pleasure to work with and spend time outside of the laboratory. I am truly happy to have you in my life, all the support and encouragement that I have received from you and the laughter that we shared. Thank you Martin, Wolfgang, Anton, Elina, Marlene and Jasmin.

To my family, I can not express enough how truly grateful I am for all your support over the years. From supporting my curiosity about nature and living beings from a very young age to enabling me to study and further pursue my interests, you were always there for me and supported me in all possible ways. You gave me the freedom and trust to follow my path, wherever it would take me, as long as it

made me happy. This accomplishment is as much yours as it is mine. Thank you to my brother Simon who was always open to discuss my work and continuously encouraged me to keep going and to dive into the world of bioinformatics.

Lastly, but most importantly, I want to express my deepest gratitude to my beloved girlfriend, Helena. Your love, support and understanding have been a constant source of strength throughout this journey. Your steady belief in me has given me the courage to keep pushing forward. Thank you for your patience, for listening to my worries and for celebrating my achievements. I am endlessly grateful to have you by my side.

Table of contents

List of abbreviations

Abbreviation	Definition
AAT	African Animal Trypanosomosis
AFLP	Amplified Fragment Length Polymorphism
AW-IPM	Area-Wide Integrated Pest Management
CATT	Card Agglutination Test for Trypanosomosis
CI	Cytoplasmic Incompatibility
COI	Cytochrome c oxidase subunit I
DNA	Deoxyribonucleic Acid
dsDNA	double-stranded Deoxyribonucleic Acid
ELISA	Enzyme-linked Immunosorbent Assay
FAO	Food and Agriculture Organisation
Gff	*Glossina fuscipes fuscipes*
Gmm	*Glossina morsitans morsitans*
GmmIV	*Glossina morsitans morsitans* Iflavirus
GmmNegeV	*Glossina morsitans morsitans* Negevirus
GpSGHV	*Glossina pallidipes* Salivary Gland Hypertrophy Virus
HAT	Human African Trypanosomosis
IAEA	International Atomic Energy Agency
IPCL	Insect Pest Control Laboratory
Kb	Kilobase
km²	square Kilometer
Mb	Megabase
MOZ	Mozambique
mtDNA	mitochondrial DNA
NGS	Next-Generation Sequencing
NAFA	Nuclear Applications in Food and Agriculture
OVI	Onderstepoort Veterinary Institute
PCR	Polymerase Chain Reaction
RAD-seq	Restriction-Site Associated DNA Sequencing
RNA	Ribonucleic Acid

SAT	Sequential Aerosol Technique
*s*Gff	*Spiroplasma* of *Glossina fuscipes fuscipes*
SIT	Sterile Insect Technique
SNPs	Single Nucleotide Polymorphisms
Spaid	*Spiroplasma poulsonii* androcidin
USD	United States Dollar
*w*Gmm	*Wolbachia* of *Glossina morsitans morsitans*

Abstract

Tsetse flies (*Glossina* spp.) are the vectors of African trypanosomes, causing Human African Trypanosomosis (HAT) and African Animal Trypanosomosis (AAT), diseases with profound medical and socioeconomic impacts in sub-Saharan Africa. Despite extensive research, vaccines against these diseases are not available to date and medication against Trypanosomosis can have severe side effects. Vector control is therefore considered as a promising strategy to suppress fly populations and break the cycle of disease transmission. A wide array of vector control approaches is currently available, ranging from chemical and biotechnological control to more species-specific and sustainable methods like the Sterile Insect Technique (SIT). This technique includes the mass-rearing of the target insect species, sterilization of males with ionizing radiation and subsequent field releases in the target area. Released sterile males mate with wild virgin females and as a result, no offspring are produced, leading to a gradual decline of the insect population. SIT is particularly effective when it is applied in the context of an area-wide integrated pest management approach (AW-IPM), utilizing multiple pest control techniques in a systematic manner. For effective vector control, a fundamental understanding of the field populations is a necessity, which includes considerations on the ecological habitats in the target area, as well as the identification of population structure and potentially isolated populations. Molecular population genetics approaches can provide valuable insights into the structure and gene flow of field populations, aiding in the design and implementation of appropriate control strategies. In addition, the effectiveness of SIT can be influenced by different factors regarding the productivity in mass-rearing settings or the influence of bacterial symbionts on vector competence. It is therefore of importance to study the symbiotic relationship of tsetse flies and their heritable bacteria to identify their consequences for the SIT. Tsetse flies host several bacteria that can affect the flies' physiology in terms of reproduction, nutrition and vector competence for transmitting trypanosome parasites, namely *Wigglesworthia glossinidia, Sodalis glossinidius, Wolbachia pipientis* and *Spiroplasma*. In particular, there is growing evidence that certain symbionts can modulate tsetse flies' ability to transmit trypanosomes, opening up possibilities to improve existing vector control strategies by leveraging the natural biological processes and interactions with their symbiotic bacteria.

The objectives of this dissertation were firstly, to develop and characterize novel microsatellite markers for the tsetse fly species *Glossina brevipalpis* to enable population genetics studies and secondly, to investigate interactions between tsetse flies, trypanosomes and symbiotic bacteria, specifically *Sodalis* and *Spiroplasma*. The research aimed to assist vector control strategies, including the Sterile-Insect-Technique (SIT), by providing insights into population structure, gene flow, and the influence of bacterial symbionts on vector competence and the efficiency of mass-rearing settings.

Microsatellite markers were successfully developed and validated, revealing genetic differentiation between laboratory and field populations from South Africa and Mozambique. Moreover, studies on the prevalence of trypanosomes and *Sodalis* across various tsetse species and regions highlighted significant variability, species-specific and location-dependent interactions. Investigations into *Spiroplasmas* effects on *G. fuscipes fuscipes* demonstrated significant shifts in reproductive and metabolic physiology, including prolonged gonotrophic cycles and reduced sperm motility, potentially impacting SIT efficacy. *In vitro* cultivation and genome sequencing of *Spiroplasma* provided genomic data, revealing core metabolic genes, potential virulence factors and interactions with host cells and trypanosomes.

Overall, this dissertation contributes to the scientific understanding of tsetse fly population genetics, symbiont interactions and vector control, supporting efforts to reduce the disease burden in Africa. These findings have significant implications for improving vector control programs, ultimately aiming to reduce the incidence of HAT and AAT, enhancing public health and economic stability in affected regions.

Keywords: *Glossina*, tsetse flies, *Sodalis*, *Spiroplasma*, Trypanosomosis, SIT

Zusammenfassung

Tsetsefliegen (*Glossina* spp.) sind die Überträger von Afrikanischen Trypanosomen, die die Krankheiten Afrikanische Trypanosomose in Menschen (HAT) und Afrikanische Tier-Trypanosomose (AAT) in Tieren verursachen. Diese Krankheiten haben erhebliche medizinische und sozioökonomische Auswirkungen in afrikanischen Ländern südlich der Sahara. Trotz umfangreicher Forschung sind bisher keine Impfstoffe gegen diese Krankheiten verfügbar, darüber hinaus kann die Medikation gegen Trypanosomose schwere Nebenwirkungen auslösen. Deshalb wird die Kontrolle des Vektors als vielversprechende Strategie angesehen, um Fliegenpopulationen zu reduzieren und somit den Übertragungszyklus der Krankheit zu durchbrechen. Eine breite Palette von Vektorkontrollstrategien steht derzeit zur Verfügung, die von chemischen und biotechnologischen Methoden bis hin zu umweltfreundlicheren Methoden wie der Sterile Insect Technique (SIT) reichen. Der Ansatz der SIT umfasst die Massenaufzucht von Insekten, Sterilisation der Männchen mittels ionisierender Strahlung und die anschließende Freisetzung dieser sterilen Männchen im Zielgebiet. Sterile Männchen paaren sich daraufhin mit wildlebenden, unbefruchteten Weibchen, was zu keinen Nachkommen und in weiterer Folge zu einem Rückgang der Insektenpopulationsgröße führt. Die SIT ist besonders wirksam, wenn sie im Rahmen eines flächendeckenden integrierten Schädlingsbekämpfungskonzepts (AW-IPM) angewandt wird, bei dem mehrere Methoden systematisch kombiniert werden. Für eine wirksame Vektorkontrolle ist außerdem ein grundlegendes Verständnis der wilden Populationen erforderlich. Dazu gehören Erkenntnisse zu den ökologischen Lebensräumen im Zielgebiet sowie die Ermittlung der Populationsstruktur und potenziell isolierter Populationen. Molekulare populationsgenetische Ansätze können hierfür wertvolle Einblicke in die Struktur und den Genfluss von Feldpopulationen liefern und so bei der Entwicklung und Umsetzung geeigneter Kontrollstrategien helfen. Außerdem kann die Wirksamkeit der SIT durch verschiedene Faktoren beeinflusst werden, wie die Produktivität in Massenaufzuchtsanlagen oder den Einfluss bakterieller Symbionten auf die Vektorkompetenz. Daher ist es besonders wichtig, die Beziehung zwischen Tsetsefliegen und ihren symbiotischen Bakterien zu untersuchen, um deren Auswirken auf die SIT zu identifizieren. Tsetsefliegen beherbergen eine Vielzahl von Bakterien, die die Physiologie der Fliegen in Bezug auf Fortpflanzung, Ernährung und Vektorkompetenz für die Übertragung von Trypanosomen beeinflussen können, und zwar *Wigglesworthia glossinidia, Sodalis glossinidius*, *Wolbachia pipientis* und *Spiroplasma*. Insbesondere mehren sich die Hinweise darauf, dass bestimmte Symbionten die Fähigkeit der Tsetsefliegen zur Übertragung von Trypanosomen modifizieren können, was Möglichkeiten zur Verbesserung bestehender Vektorkontrollstrategien eröffnet, indem die natürlichen biologischen Prozesse und Symbiont-Interaktionen genutzt werden.

Die Ziele dieser Dissertation waren erstens: die Entwicklung und Charakterisierung neuartiger populationsgenetischer Microsatellite-Marker für die Tsetsefliegenart *Glossina brevipalpis*, und zweitens: die Untersuchung der Wechselwirkungen zwischen Tsetsefliegen, Trypanosomen und symbiotischen Bakterien, insbesondere *Sodalis* und *Spiroplasma*. Ziel der Forschung war es, Strategien zur Vektorkontrolle, einschließlich der SIT zu unterstützen, indem Einblicke in die Populationsstruktur, Genfluss und den Einfluss bakterieller Symbionten auf die Vektorkompetenz und die Effizienz der Massenaufzucht gewonnen wurden. Microsatellite-Marker wurden erfolgreich validiert, wobei eine genetische Differenzierung zwischen Labor- und Feldpopulationen und weiters zwischen Feldpopulationen in Mosambik und Südafrika festgestellt wurde. Darüber hinaus zeigten Studien zur Prävalenz von Trypanosomen und *Sodalis* bei verschiedenen Arten der Tsetsefliege und in verschiedenen Regionen eine erhebliche Variabilität sowie artspezifische und ortsabhängige Wechselwirkungen. Untersuchungen zu den Auswirkungen von *Spiroplasma* auf *G. fuscipes fuscipes* zeigten signifikante Veränderung der Fortpflanzungs- und Stoffwechselphysiologie, einschließlich verlängerter gonotropher Zyklen und verringerter Spermienmotilität, was sich auf die Massenaufzucht von Tsetsefliegen und SIT Programme auswirken kann. Die *in vitro* Kultivierung und Genomsequenzierung von *Spiroplasma* lieferte genomische Daten, die zentrale Gene des Stoffwechsels und potentielle Gene identifizierte, welche sich auf die Interaktion mit dem Wirt und Trypanosomen auswirken könnten.

Insgesamt trägt diese Dissertation zum wissenschaftlichen Verständnis der Populationsgenetik der Tsetsefliege, der Interaktionen mit Symbionten und der Vektorkontrolle bei und unterstützt damit die Bemühungen um eine Reduzierung der Krankheitslast in Afrika. Diese Erkenntnisse haben Auswirkungen auf die Verbesserung der Vektorkontrollprogramme, die letztlich darauf abzielen, das Auftreten von HAT und AAT zu verringern und sowohl die öffentliche Gesundheit, als auch die wirtschaftliche Stabilität in den betroffenen Regionen zu verbessern.

Schlüsselwörter: *Glossina*, Tsetsefliegen, *Sodalis*, *Spiroplasma*, Trypanosomose, SIT

Declaration

I hereby declare and affirm that this dissertation titled “Population genetics and research on host-symbiont interactions of the tsetse fly (*Glossina* spp.)” was authored by me to obtain the Doctorate degree in Natural Sciences under the supervision of Univ. Prof. Mag. Dr. rer. nat. Robert Mach from the Institute of Chemical, Environmental and Biological Engineering and Prof. Dr. Adly Abdalla from the Insect Pest Control Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture in Seibersdorf, Austria. I furthermore certify that this dissertation is a product of original research performed at the Insect Pest Control Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture. I have taken appropriate measures to ensure the originality of this work, strictly adhering to copyright laws and citation rules when referencing other work. This dissertation will not be submitted for any other degree or professional qualification. The experimental work performed for this dissertation was performed by me and any collaborative contributions of myself and colleagues are clearly indicated in the publications contained in this dissertation. The findings presented in this dissertation corresponds to the listed publications in the following section named “List of publications”.

List of publications

1. **Gstöttenmayer, F.**, Moyaba, P., Rodriguez, M., Mulandane, F.C., Mucache, H.N., Neves, L., De Beer, C., Ravel, S., De Meeûs, T., Mach, R.L., Vreysen, M.J.B., Abd-Alla, A.M.M., 2023. Development and characterization of microsatellite markers for the tsetse species *Glossina brevipalpis* and preliminary population genetics analyses. Parasite 30, 34. https://doi.org/10.1051/parasite/2023038

2. Dieng, M.M., Dera, K.M., Moyaba, P., Ouedraogo, G.M.S., Demirbas-Uzel, G., **Gstöttenmayer, F.**, Mulandane, F.C., Neves, L., Mdluli, S., Rayaisse, J.-B., Belem, A.M.G., Pagabeleguem, S., De Beer, C.J., Parker, A.G., Van Den Abbeele, J., Mach, R.L., Vreysen, M.J.B., Abd-Alla, A.M.M., 2022. Prevalence of *Trypanosoma* and *Sodalis* in wild populations of tsetse flies and their impact on sterile insect technique programmes for tsetse eradication. Sci Rep 12, 3322. https://doi.org/10.1038/s41598-022-06699-2

3. Son, J.H., Weiss, B.L., Schneider, D.I., Dera, K.M., **Gstöttenmayer, F.**, Opiro, R., Echodu, R., Saarman, N.P., Attardo, G.M., Onyango, M., Abd-Alla, A.M.M., Aksoy, S., 2021. Infection with endosymbiotic *Spiroplasma* disrupts tsetse (*Glossina fuscipes fuscipes*) metabolic and reproductive homeostasis. PLoS Pathog 17, e1009539. https://doi.org/10.1371/journal.ppat.1009539

4. Bruzzese, D., **Gstöttenmayer, F.**, Weiss, B., Abd-Alla, A.M.M., Aksoy, S., 2024. *In vitro* cultivation and genomic insights into the *Spiroplasma* symbiont *of Glossina fuscipes fuscipes*. (draft, to be submitted to PloS Pathogens)

Chapter 1: General Introduction

1. General Introduction

This general introduction provides a background for the work performed in this dissertation. It begins with an overview of the systematics, reproductive biology and geographic distribution of tsetse flies (*Glossina* spp.). Their vector status for African trypanosomes and the resulting medical and economic impacts are outlined, followed by strategies of vector control to manage tsetse fly populations, ranging from conventional control methods to more refined approaches such as the Sterile Insect Technique (SIT). The introduction subsequently explores the role of population genetics as a tool in designing and implementing effective vector control programs, emphasizing how genetic studies can reveal insights into population structure and gene flow of tsetse populations, which are essential insights for targeted control efforts. Lastly, the symbiotic relationships between tsetse flies and their heritable bacterial symbionts and viruses are investigated. These symbionts may significantly influence tsetse biology, reproduction and vector competence. Understanding these interactions can provide further insights into the vector competence of tsetse flies and have implications for mass-rearing activities for SIT programs. This general introduction thus frames a context for the subsequent chapters, in which the performed work is presented and discussed.

1.1. Tsetse flies: Systematics, Reproductive Biology and Distribution

Tsetse flies are dipteran insects comprising of a single genus *Glossina*, part of the *Glossinidae* family which belongs to the super family of blood feeding Hippoboscoidea [1]. They are distributed throughout sub-Saharan Africa in 37 countries, spanning over a vast area of 10 million km^2 and diverse ecological habitats including savannah, forests and riverine systems. The genus *Glossina* comprises of over 30 species and their subspecies, classified into three subgenera according to their habitat preferences: the subgenus *morsitans* is adapted to the woodland savannahs in relatively dry and open environments, the *fusca* group inhabits lowland rain forests and the *palpalis* group occurs in lowland rain forests in coastal regions, extending to the riverine systems of the savannah [2–6]. The distribution is dependent on a multitude of environmental factors such as temperature, humidity, flora and the presence of hosts [7–10]. A distribution map of tsetse fly species across sub-Saharan Africa is presented in Figure 1.

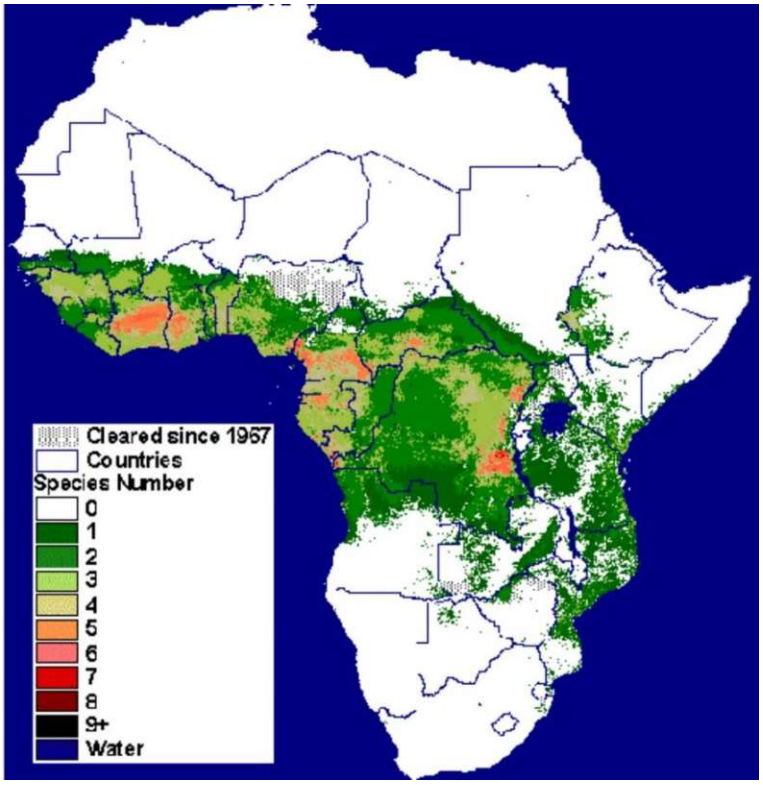

Figure 1: Tsetse fly geographic distribution on continental scale indicating the spread of different tsetse fly species. Colours indicate occurrence of >1 species [11]

Tsetse flies exhibit a remarkable reproductive strategy that is unique among Dipterans and is known as adenotrophic viviparity, characterized by the development of a single offspring inside the females' uterus, where it is nourished with milk gland secretions and given live birth at a late larval stage [12,13]. Notably, the reproduction of tsetse flies is very slow as compared to other insects, which stems from the energy-intensive process of adenotrophic viviparity [7,12]. In principal, one successful mating with a male provides enough sperm for the entire lifespan of the female, with the sperm being stored in the females' spermathecae [2]. Female tsetse flies produce their first offspring after approximately 20 days post-emergence. This period includes the adult maturation, mating and the first gonotrophic cycle. The gonotrophic cycle contains embryonic development, intrauterine larval development through three larval stages and larviposition (Figure 2). After giving birth to the first offspring, females typically produce one offspring at intervals of approximately 10 days for up to 8 gonotrophic cycles throughout their lifespan, depending on environmental conditions and nutritional state [14,15]. An inadequate nutritional state or other environmental factors might also lead to aborted larvae, in which case not enough nutrients can be provided to the offspring [16]. After successful larviposition in the third instar larval stage, larvae burrow approximately 2-3 cm into the ground and pupate for a period of up 30 days [14]. A teneral adult fly ecloses from the puparium, takes the first blood meal within 12 to 24 hours and the life cycle repeats [14,17].

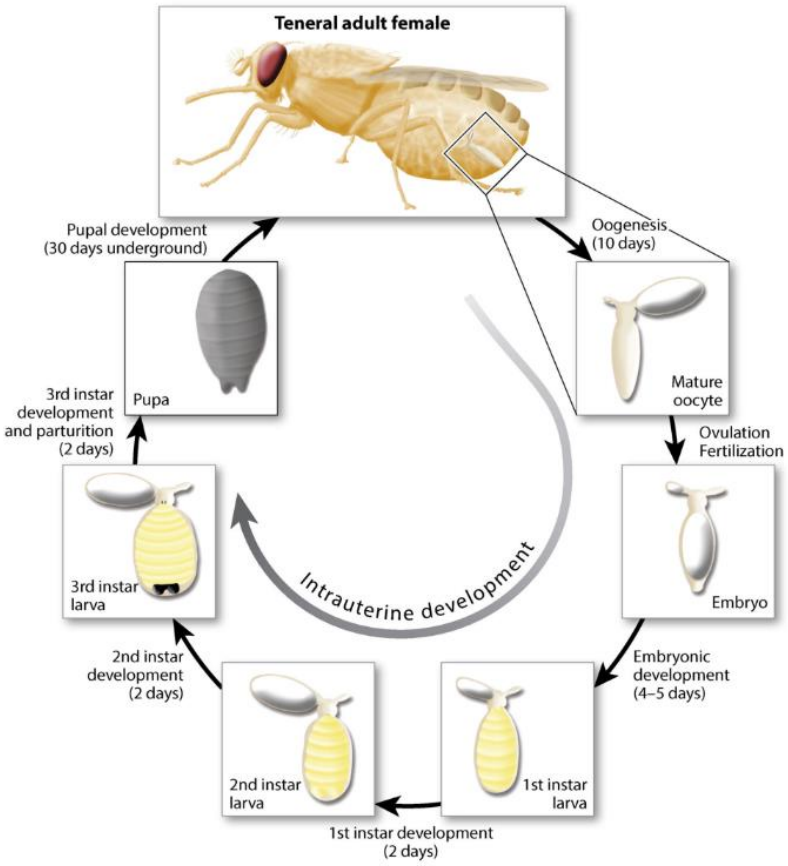

Figure 2: Schematic representation of the female tsetse fly gonotrophic cycle [12]

1.1.1. African trypanosomes and the associated diseases

Tsetse flies are strictly hematophagous insects [7], both female and male adults feed on blood of humans and animals. Tsetse flies are the only cyclic vectors of African trypanosomes, leading to the devastating diseases Human African Trypanosomosis (HAT), also referred to as sleeping sickness in humans and African Animal Trypanosomosis (AAT) or nagana in animals [3,18,19]. Trypanosomes are protozoan hemoflagellated parasites of the genus *Trypanosoma*, which are ingested by tsetse flies through the blood feeding on infected hosts. The trypanosomes change into procyclic trypomastigotes in the tsetse midgut, after which they move to the salivary glands and transform into epimastigotes. These epimastigotes multiply and transform into metacyclic trypomastigotes, the infective stage of trypanosomes (Figure 3) [19–21]. At this stage, the life cycle of trypanosomes is completed. The infection with trypanosomes occurs in two stages, the first stage being defined by nonspecific symptoms like fever, pruritus and lymphadenopathy. The second stage of infection involves the invasion of the central nervous system, leading to neurological and cardiological symptoms which are fatal if not treated adequately [20,22].

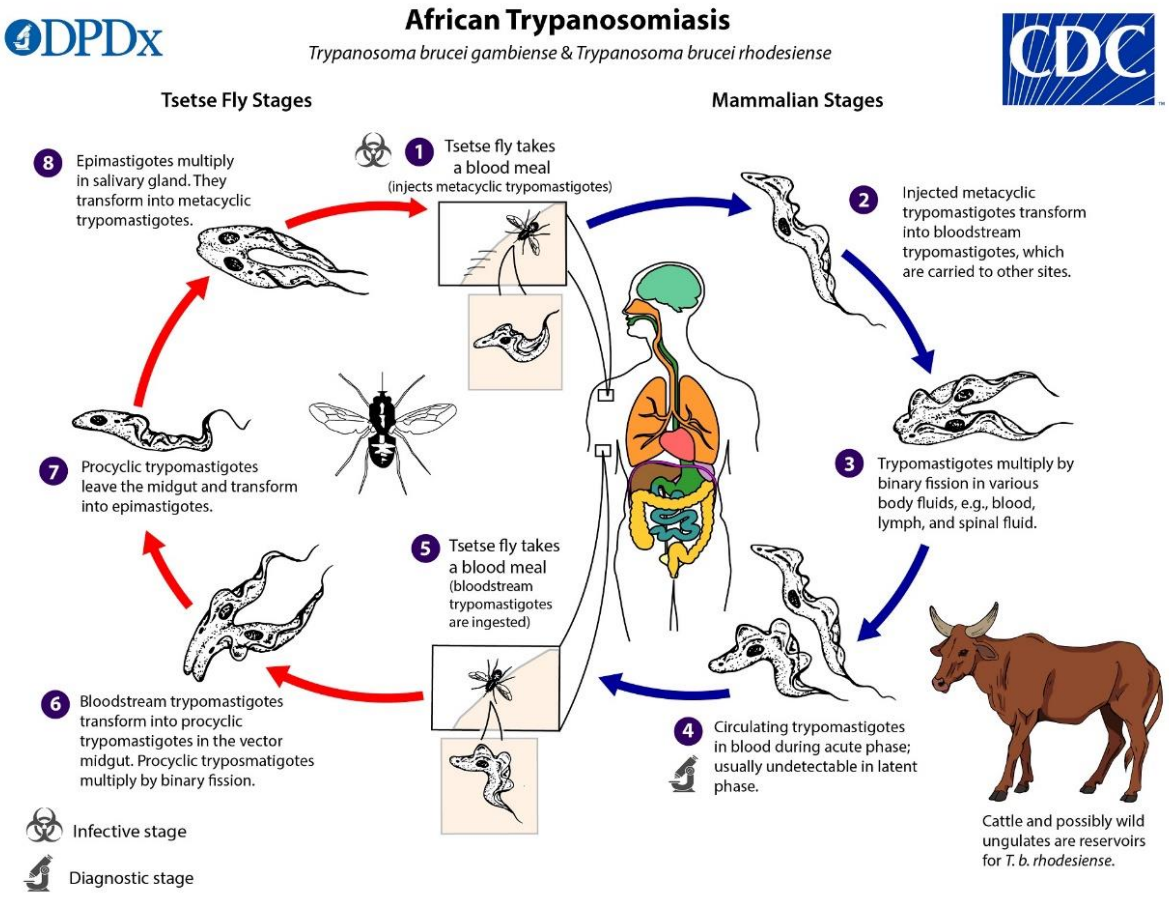

Figure 3: Life stages of trypanosomes within the host and vector [20]

1.1.2. Human African Trypanosomosis

Human African Trypanosomosis (HAT), commonly referred to as sleeping sickness, is endemic in 37 sub-Saharan African countries and can be caused by several species of *Trypanosoma* spp. parasites, among the most widely distributed are *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense* [21]. *T. b. gambiense* is responsible for the chronic form of sleeping sickness and occurs in West and Central Africa, accounting for over 92% of reported sleeping sickness cases, whereas *T. b. rhodesiense* is causing the acute form of the disease and is present in East Africa [20,22,23]. The majority of affected people live in rural regions and rely on agriculture, livestock farming and fishing and are therefore exposed to tsetse flies [22,24]. HAT is diagnosed through a combination of clinical symptom assessment and laboratory tests, such as microscopy to detect trypanosomes in body fluids and serological tests, like the card agglutination tests for trypanosomosis (CATT) or enzyme-linked immunosorbent assay (ELISA) [25,26]. Once diagnosed, the type of medical treatment is determined by the *Trypanosoma* subspecies and stage of infection. Notably, some treatments have proven to be effective against the disease, however, they can exhibit severe side effects and the potential to develop drug-resistant trypanosomes [23,27–29]. In the last decades, there have been several major outbreaks of sleeping sickness in Africa [30,31], with over 35,000 documented cases in 1998 and

estimated to be as many as 300,000 cases due to underreporting and lack of capacities for diagnosis [32]. Coordinated efforts have been undertaken since then through various control approaches, resulting in a decline of reported case numbers over time: from over 25,000 cases in the year 2000, 20,000 cases in 2003, to the first time below 10,000 cases in 2009. By 2017, less than 2000 cases per year were reported and only 977 cases in 2018, remaining below the level of 1000 cases per year since then [33,34]. Recent surveys indicate 675 reported cases of sleeping sickness in 2023, predominantly occurring in Central and West Africa [35]. This number of cases thus is at a historically low level, which can be attributed to the sustained vector control efforts. However, the estimated number of people at risk remains at 55 million, with 3 million being exposed to a moderate to high risk [36].

1.1.3. African Animal Trypanosomosis

African Animal Trypanosomosis (AAT), also called nagana, is caused by various *Trypanosoma* species: *Trypanosoma congolense, Trypanosoma vivax, Trypanosoma brucei, Trypanosoma simiae* and *Trypanosoma evansi.* They affect a multitude of domestic animals such as cattle, pigs, sheep, goats and horses, as well as wild ruminant mammals [37–39]. Common symptoms of infected animals are swollen lymph nodes, weight loss, fever and anaemia. Furthermore, degenerative and inflammatory lesions can develop in most of the animals' organs at a late stage of infection [40,41]. As with HAT, the disease results in the death of the infected animal if not treated. The detection of trypanosome infection in animals is conducted similarly to HAT and can be detected with serological tests and microscopy to confirm their presence [40,42]. AAT is commonly treated by intramuscular injection of trypanocidal drugs such as isometamidium chloride (0.25 – 1 mg / kg body weight) and diminazene aceturate (3.5 mg / kg body weight) [43], but the same concerns of developing trypanocide resistance as in HAT persists for AAT as well [28]. Notably, some wild mammal species can tolerate trypanosomes infection, which is expected to be a consequence of the coexistence with the parasite over long periods of time. Nevertheless, these animals act as reservoirs for trypanosomes and the vectors can still ingest the parasites and pass them on to other healthy animals [44]. AAT therefore represents a major threat to animal health and the livelihood of animal agriculturalists. The annual economic losses due to AAT on the African continent are estimated to be in the range of 3.8 billion USD [45,46], additionally a reduction of livestock production by 20-40% is projected in areas where animal agriculture and trypanosomosis coincide [47]. In conclusion, trypanosomosis presents a hindrance to the development of agricultural area and livestock keeping in the infested areas and therefore the economic development [48]. The severe impact of tsetse flies and the transmitted parasites is furthermore shown by the significant overlap of infested areas and the poorest countries in Africa [49].

1.2. Vector and disease control

1.2.1. Conventional vector control approaches

Despite extensive research, there is currently no effective vaccine available to protect against trypanosomes [50]. The potential resistance development of trypanosomes to certain drugs can also impair the effectiveness significantly [23], therefore, the direct control of the insect vector represents a promising strategy to suppress fly infestation and thus the disease burden of trypanosomosis [51,52]. Historically, many different vector control approaches have been tested and implemented over the years. One of the first applied methods was the clearing of vegetation and livestock to remove both the hosts and the habitat of tsetse flies [53]. Another strategy to reduce tsetse fly populations was spraying of insecticides over wide areas using aircrafts [54]. Although effective, these two strategies applied in the early days of vector control efforts are not considered sustainable as the use of the chemical insecticides raises environmental concerns due to their non-specific toxicity and the potential of resistance development [55,56]. The strategy of insecticide spraying was later refined to reduce negative effects on the environment and on non-target insects, which resulted in the development of the Sequential Aerosol Technique (SAT). This technique includes the sequential spraying of low volumes of non-residual insecticides that affect tsetse flies upon contact with spray mist [57]. It significantly reduced the environmental impacts and is considered accepted as a more environmentally-friendly strategy [51]. The effects of SAT were also monitored during field campaigns, with the conclusion that the effect on ecological indicator species was limited and that their populations quickly recovered to pre-intervention levels [58,59].

Over the years, further vector control strategies emerged: traps and targets, that serve both for sampling tsetse flies for field surveys and to reduce their numbers in the field [60–62]. Different traps have been developed to optimize sampling of the different species of tsetse flies and traps are typically combined with odour-baits to improve collection [63,64]. Furthermore, targets are a cheap and simple, yet effective way to collect tsetse flies. The principle of using targets for tsetse fly control is based on the behaviour of tsetse flies and their attraction to certain visual cues, namely black or blue screens, that have been found to be particularly attractive to tsetse flies [65]. These screens are either impregnated with insecticides [66], that induces mortality shortly after landing on the target, covered with glue (sticky targets) to capture the insect [67], or running an electric current through the target to electrocute flies upon landing [68].

Another method that is applied to animals to avoid the transmission of trypanosomes is the so-called cattle spraying or cattle dipping. This approach involves the use of insecticides, specifically deltamethrin (1%, rate of 1 ml / 10 kg body weight [69]) or cypermethrin (2.5 mg / kg of body weight [70]), applied to cattle at specific sites where tsetse flies are known to feed on the host [71,72]. This

treatment is not specific to tsetse flies but also affects ticks and flies of the species *Stomoxys* and there are reports of a significant decrease in their populations during tsetse control efforts [73]. Furthermore, the use of insecticides on cattle can lead to contamination of their excrements and subsequently kill beetles and flies associated with the animal's dung. These insects perform an important task in the natural integration of manure into the soil, which is essential for the soil cycle, furthermore the insecticide residues in the dung may have long term effects on the populations of slow-breeding beetles [74,75].

1.2.2. Sterile-Insect-Technique (SIT)

Given the growing concerns over adverse environmental effects on non-target organisms posed by conventional vector control methods, there is a necessity for more sustainable approaches. The Sterile Insect Technique (SIT) is a strategy developed with the aim to suppress and eradicate insect pest populations including tsetse flies, followed by containment and prevention of their reestablishment. The basic steps underlying SIT on tsetse flies involve the mass-rearing, sterilization of males with ionizing radiation and release of sterile males in the field to mate with wild virgin females. After mating with sterile males, wild females subsequently do not produce offspring (Figure 4) [76,77]. This is due to the sterility of the males induced by ionizing radiation, leading to chromosome fragmentation of germ cells, specifically dominant lethal mutations and translocations, that subsequently result in the interruption of mitosis and death of embryos [78,79].

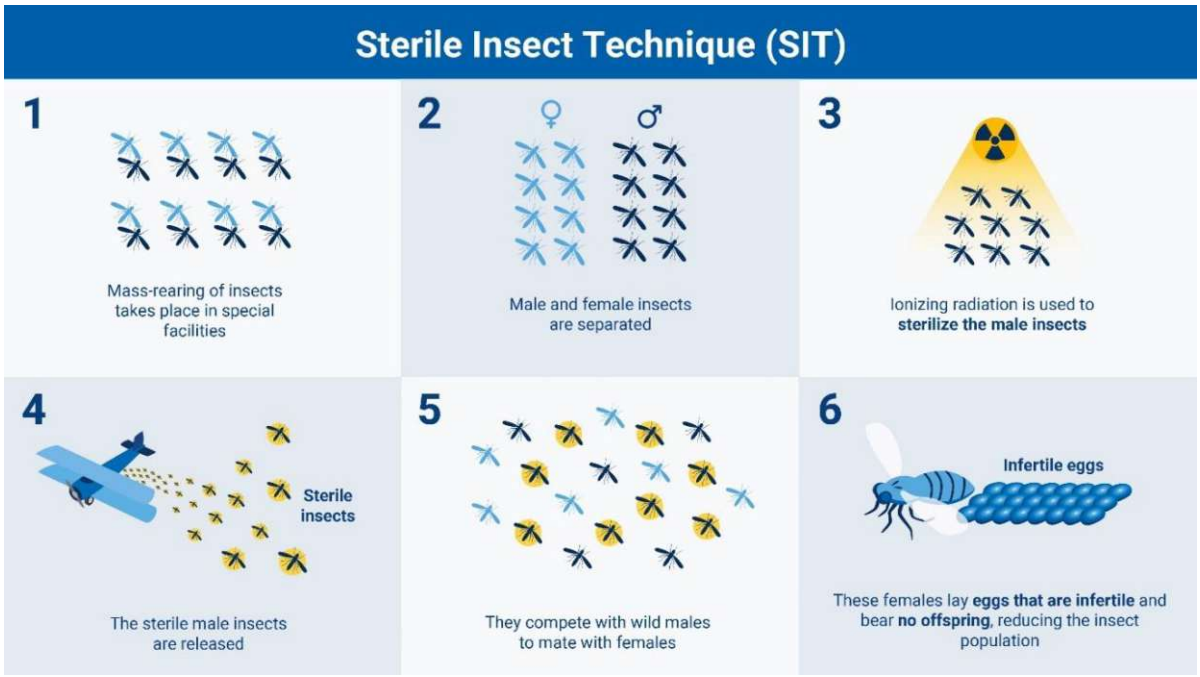

Figure 4: Schematic overview of the steps of SIT [80]

For the SIT to be efficient as possible, several key aspects must be considered before planning and implementing a campaign: First, a mass-rearing colony of the target tsetse fly species must be

established and maintained to provide a high number of insects for the following steps [81]. The second aspect is that mass-reared males must perform well and be competitive with wild males to increase the chances of successful mating with wild virgin females. Therefore, it is essential to test the performance and mating competitiveness of these sterilized males [82,83]. Additionally, irradiation protocols have to be established and optimized to achieve induced sterility levels above 95% [84]. For irradiation of males, typically gamma radiation is utilized [85]. However, it has been recently proven that X-Ray irradiation is as effective as gamma radiation, with the advantage of X-Ray irradiators being that they do not fall under the strict regulations like gamma cells [86]. Finally, strategic deployment of a large number of sterilized males across a wide area through sequential releases is imperative for the success of SIT [6,87].

Historically, SIT has been implemented for various insect species with success. The earliest historic example is the elimination of *Cochliomyia hominivorax* (new world screwworm) from North and Central America. The campaign was a long-term effort, starting in the 1950s and reaching eradication in the area by 1982. It then extended further south to Central American countries such as Mexico, Guatemala, Honduras and Costa Rica, among others. Since 2004 there is a permanent barrier in Panama and all countries north of this barrier have been declared free of the new world screwworm since then [88]. Following the success of this campaign, SIT has been effectively used in controlling populations of further insect pests, like the mediterranean fruit fly *Ceratitis capitata* in the United States and Mexico [89,90]. Populations of mediterranean fruit flies were successfully suppressed in Valencia, Spain, resulting in a decrease of insecticide use of over 90% and thus highlighting the replacement of insecticide spraying by the more sustainable SIT [91]. In the case of tsetse flies, SIT has been successfully implemented on Unguja island in Zanzibar, resulting in the eradication of *G. austeni* from the island, which has been free of this tsetse fly species since 1997 [92]. Recent efforts of an AW-IPM campaign incorporating a SIT component in the Niayes region of Senegal resulted in the eradication of *G. p. gambiensis*, leading to the elimination of AAT from the target area, enabling the sustainable development of animal agriculture in this region [93].

The presented vector control strategies may not be implemented individually but in the context of an area-wide integrated pest management (AW-IPM) approach with the option to combine different control tactics as needed to utilize the benefits of different approaches at certain densities of infestation. Some approaches like aerosol spraying are more effective at high densities of insects, whereas SIT is highly effective at moderate to low densities [77].

1.3. Population genetics as a tool to assist insect control programs

Population genetics involves the studies of genetic variability and gene flow within and between populations as well as the evolutionary forces at play [94]. This field evolved as a branch of genetics focused on the distribution and change of allele frequencies influenced by the main evolutionary processes, namely mutation, drift, migration and natural selection [95]. The foundation of population genetics is based on mathematical and statistical theory to describe populations and several key parameters have been developed over the years [94,96]. Among these parameters are Wright's F-statistics for inbreeding and differentiation [97–100], Nei's genetic distances [101], the Hardy-Weinberg-Equilibrium [102], population genetics statistics as formulated by Rousset [103] and phylogenetic analyses based on Edwards and Cavalli-Sforza's models [104]. In the framework of insect vector control, population genetics is a commonly used tool to understand the population structure and genetic variability of target field populations [105]. Specifically, it can be used to study the target populations in terms of gene frequencies, genetic drift, migration, dispersal and admixture between populations [96]. Furthermore, population genetics can help identify the degree of isolation of certain subpopulations and the origins of genetic bottlenecks and reinvasion events. In addition, population genetics data can offer estimations on effective population sizes and immigration rates from adjacent populations [88]. Insights derived from population genetics analyses are practical for the planning phase of vector control programs. For vector control, identifying isolated field populations represents an ideal situation, as these populations can be targeted by control strategies without reinvasion risks into the target area [106]. In the Niayes region of Senegal, for example, results of microsatellites, mitochondrial COI markers and morphometric analysis indicated an isolated population that was then targeted [107]. However, it is rare find fully isolated populations in the field, and in settings where no isolated population can be identified, a rolling-carpet approach may be applied. This was undertaken in the case of the new world screwworm SIT campaign in North and Central America [108].

There is a wide array of molecular genetic tools to study population genetics: Amplified fragment length polymorphism (AFLP) is a method developed in the 1990s that involves the selective PCR amplification of restriction fragments from restriction enzyme-digested genomic DNA. It can generate a large number of polymorphic markers without prior knowledge of the genome sequence [109]. Mitochondrial DNA markers are based on the Cytochrome c oxidase subunit I (COI) and are particularly useful for studying maternal inheritance patterns and to identify population structure. However, their utility is limited by the maternal inheritance, which does not entirely capture the genetic variation within populations [110] and also the lack of recombination in mtDNA can restrict the resolution of fine-scale population structure [111]. Microsatellites, also known as simple sequence repeats (SSRs) or short tandem repeats, are among the most widely used population genetics tools due to their high

polymorphism and abundance in eukaryotic species [112]. Microsatellite regions consist of repeats of a motif comprising of typically 1-6 nucleotides [113]. The primary advantage of microsatellites is their high degree of variability, which makes it possible to identify genetic differences even among closely related populations [114]. However, one disadvantage of microsatellites is that the high mutation rate, which may lead to homoplasy, where different alleles appear identical due to independent mutations resulting in the same repeat numbers, complicating the interpretation of genetic relationships, especially over long evolutionary timescales [115].

Recent advancements in next-generation-sequencing (NGS) technologies and the decreasing costs of sequencing have opened new avenues for population genetics studies. These methods allow for high-throughput and cost-effective genotyping for the identification of single nucleotide polymorphisms (SNPs) [116]. The restriction site associated DNA sequencing (RAD-seq) approach has proven particularly valuable to efficiently genotype many individuals at a large number of loci without prior knowledge of the genome sequence. The principle of RAD-seq involves the enzymatic digestion of DNA using restriction enzymes and sequencing of the fragmented DNA [117]. This method can identify thousands of SNPs across the genome which can be utilized for population genetics analysis [118]. While RAD-seq can identify a large number of SNPs, its downside is the necessity for a sufficient amount of high-quality DNA, which is not always achievable from field-collected specimens [119]. Most of the population genetics studies on tsetse flies have so far been performed using mtDNA markers and microsatellites [120–124]. However the decreasing costs of sequencing have led to recent population genetics studies using the RAD-seq approach [125,126], highlighting its increased information content.

1.4. Bacterial symbionts and viruses associated with tsetse flies

Tsetse flies harbour several bacterial endosymbionts and viruses that can significantly influence their physiology, reproductive biology and vector competence [127]. The term symbiosis was first mentioned by Anton De Bary in 1879 as “the living together of two dissimilar organisms, usually in intimate association, and usually to the benefit of at least one partner [128]”. Symbiosis is a broad field of research affecting a diverse range of species in several taxa, from bacteria, fungi, plants and animals, influencing their ecology and physiology [129]. Symbiotic relationships can be classified into mutualistic, commensal or parasitic relationships, depending on the benefits or harms experienced by the interacting species [130]. Mutualistic symbiotic interaction is defined as the relationship in which both species benefit, while in commensalism, only one species benefits and the other is not affected. A parasitic symbiotic relationship is present when one organism gains advantage at the expense of the other, typically this is the case when a parasite derives nutrients or other benefits from the host [131]. In the context of tsetse flies, the present symbionts cover the range from mutualistic to parasitic

interactions, that can exhibit significant implications on reproductive fitness in mass-rearing settings, development and modulation of the immune system, as well as vector competence [127]. The symbiotic bacteria of tsetse flies include *Wigglesworthia glossinidia*, *Sodalis glossinidius*, *Wolbachia pipientis* and the recently identified *Spiroplasma* (Figure 5). Additionally, tsetse flies also host several different viruses, such as the *Glossina pallidipes* salivary gland hypertrophy virus (*Gp*SGHV), *Glossina morsitans morsitans* Iflavirus (*Gmm*IV) and *Glossina morsitans morsitans* Negevirus (*Gmm*NegeV). These microorganisms can have a distinct role in the biology, reproductive fitness and vector competence of tsetse flies, that are subject of the following sections.

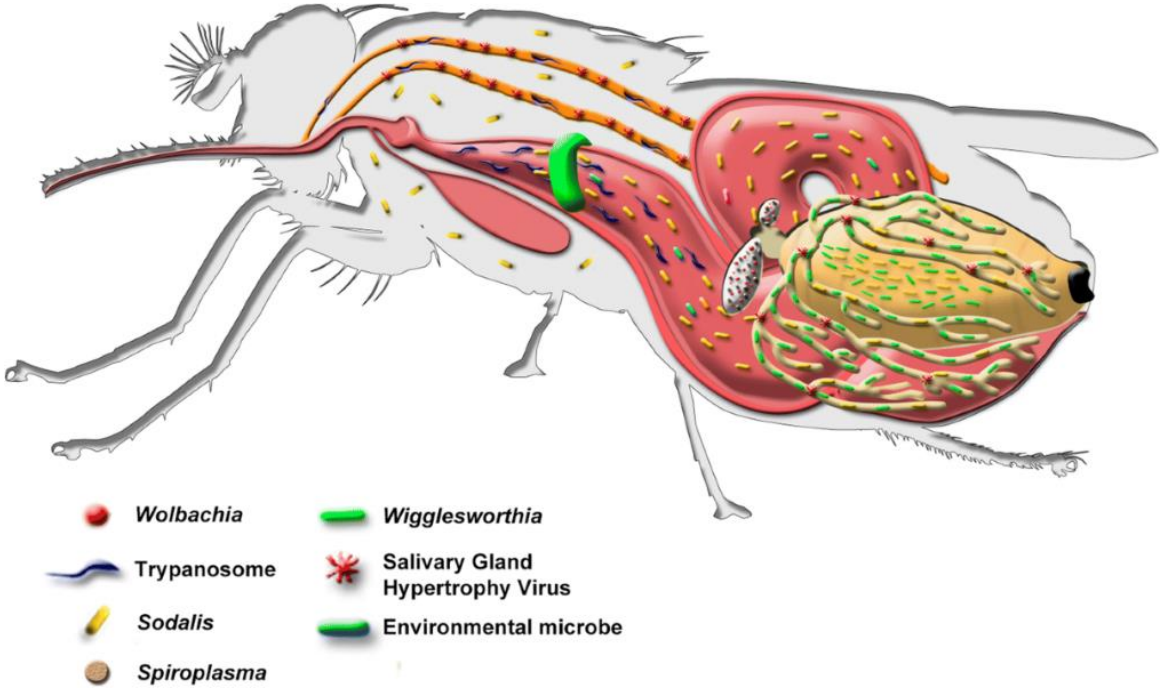

Figure 5: Localization of tsetse fly symbionts. [132] adapted from [133]

1.4.1. Wigglesworthia glossinidia

Wigglesworthia glossinidia is a gram-negative bacterium within the class of γ-Proteobacteria [127]. It is considered an obligate mutualistic symbiont, mainly residing in specialized cells named bacteriocytes within the bacteriome that is localized in the gut of tsetse flies (Figure 6) [134].

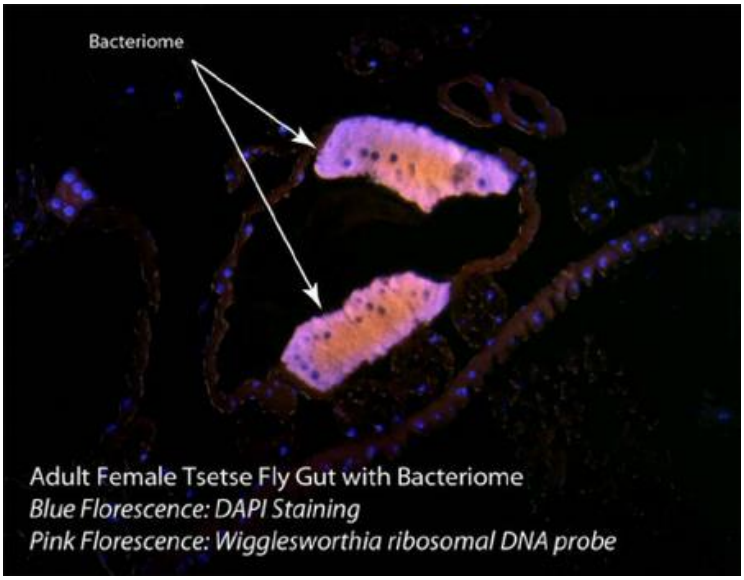

Figure 6: Localization of *Wigglesworthia* within the bacteriome in the gut *[135]*

This symbiont is present in all tsetse species and plays a crucial role for the nutritional state of tsetse flies by synthesizing essential B vitamins that are not available in the blood diet [127]. The presence of *Wigglesworthia* is vital for the successful development of larval and adult stages, impacting the reproductive success [136]. Beyond its role in reproductive fitness, *Wigglesworthia* also plays an essential part in immune system development. Tsetse flies lacking the bacterium exhibit compromised immune systems, leading to increased susceptibility to infection when challenged with *E. coli*, suggesting that this symbiont may help the fly host in maintaining an effective immune defense against pathogens [137]. Additionally, it influences the development of the gut flora, which in turn affects the vector competence for trypanosome transmission [138]. *Wigglesworthia* has a reduced genome size of approximately 700 Kb, typical for obligate mutualistic symbionts that have evolved with the host over millions of years. Despite this reduction, it retained genes for the biosynthesis of B vitamins which are a vital part of the symbiosis [134]. During the fly's reproduction, *Wigglesworthia* is transferred to the intrauterine progeny through the mother's milk gland secretions, exhibiting a strictly maternal vertical transmission [139].

1.4.2. Sodalis glossinidius

Sodalis glossinidius is a Gram-negative, rod-shaped bacterium that is part of the family of *Enterobacteriaceae* in the γ-3 subdivision of Proteobacteria [127]. It is considered a secondary symbiont of tsetse flies and is not strictly essential for fly survival or reproduction. However, the absence of *Sodalis* can decrease fly longevity and may increase susceptibility to infection with trypanosomes in some species under laboratory conditions [140,141]. There are conflicting results on the modulation of host susceptibility to trypanosome infection by *Sodalis* in wild tsetse populations depending on the *Glossina* and trypanosome species [142–144], therefore its role in vector

competence is not fully understood. *Sodalis* is localized in various tissues within its host, including the midgut, hemolymph, fat body, milk glands and reproductive organs [145]. The symbiont is transmitted maternally during intrauterine larval development, migrating into the developing larva through milk gland secretions [127]. Notably, it can also be transmitted horizontally from infected males to females during mating through seminal secretions and subsequently transmitted vertically to the progeny [146]. *Sodalis* is prevalent in all tsetse laboratory colonies, however its occurrence in wild populations is highly variable depending on the geographic region and fly species [142,144,147]. The genome size of *Sodalis* is 4.1 Mb [148] and contains several extrachromosomal plasmids [149]. Compared to *Wigglesworthia*, it shows genomic reduction to a smaller extent, suggesting that this bacterium may be in a transitional phase from free-living to a symbiotic relationship within the fly host [127]. Detailed analysis of the genome indicates that *Sodalis* has lost the arginine biosynthesis pathway, indicating that it may obtain this amino acid from the host [150]. Furthermore, it lacks the pathway for thiamine synthesis and there is concise evidence that *Sodalis* scavenges this vitamin from *Wigglesworthia* [151].

Sodalis is one of the very few insect symbionts that can be isolated and cultured *in vitro*, making it a promising candidate for paratransgenesis to control wild tsetse populations. In addition, recent advancements have made it possible to transform cultivated *Sodalis* to deliver anti-trypanosomal nanobodies to reinfected tsetse flies [152,153].

1.4.3. Wolbachia pipientis

Wolbachia is a gram-negative bacterium in the class of α-Proteobacteria that infects numerous arthropod species, making it a highly prevalent symbiont in insects [154]. This bacterium predominantly infects the ovarian tissues of female tsetse flies, from where it is maternally transmitted [145]. *Wolbachia* is present in different tsetse species from the *morsitans*, *fusca* and *palpalis* group [155,156]. The symbiotic relationship of *Wolbachia* and its hosts is generally considered parasitic, as it enhances its transmission through several mechanisms to increase female fertility [127]: parthenogenesis, feminisation, male killing and cytoplasmic incompatibility (CI) [157,158]. The most widely studied effect of *Wolbachia* in tsetse flies is cytoplasmic incompatibility, which results in embryonic mortality in offspring from mating pairs of flies with different *Wolbachia* infection status [159,160]. The role of *Wolbachia* in trypanosome infection is not fully understood - some studies suggest a negative correlation, while others did not find an impact on trypanosome establishment in the presence of the bacterium [161,162]. Given that *Wolbachia* induces a strong CI effect in *G. morsitans morsitans*, its use for suppressing field populations has been proposed. This approach would leverage CI causing embryonic lethality in crosses between released infected males with wild uninfected females, effectively reducing population sizes. Integrating *Wolbachia* into existing vector control strategies such as the SIT may increase their effectiveness [163,164].

Studies on the distribution of *Wolbachia* in different tsetse fly species reared under laboratory conditions have concluded a high prevalence in species of the *fusca* and *morsitans* groups [165,166]. In contrast to these earlier findings, it was recently concluded that *Wolbachia pipientis* is also present in *G. fuscipes fuscipes,* member of the *palpalis* group, although at very low levels. This was identified using the high sensitivity detection method blot-PCR combined with hybridizations [167]. The infection prevalence in field populations is heterogenous depending on tsetse fly species and presumably the ecological conditions in different field locations [168].

The genome size of *Wolbachia* from *G. morsitans morsitans* (*w*Gmm) is approximately 1.02 Mb and exhibits a high number of genes encoding for ankyrin domains, considered to play a vital part in interactions of host and symbiont. The genome also contains a high number of repeat sequences, transposable elements and prophages [169].

1.4.4. Spiroplasma

Spiroplasma is a bacterium within the class of Mollicutes, defined by its helical morphology (Figure 7) and lacking a cell wall [170,171]. It was relatively recently discovered in laboratory and wild populations of the tsetse fly species *Glossina fuscipes fuscipes*, *Glossina tachinoides* and *Glossina palpalis palpalis*, all belonging to the *palpalis* group [172,173].

Figure 7: Scanning electron microscopy image depicting morphology of *Spiroplasma poulsonii [174]*

In contrast to its recent discovery in tsetse flies, *Spiroplasma* has been extensively studied in other insects, revealing both beneficial and adverse effects depending on the host species: in *Drosophila melanogaster* and *Drosophila hydei*, *Spiroplasma* confers a protective effect against parasitoid wasps by preventing wasp larval development [175,176]. This protective effect is caused by ribosome-inactivating proteins (RIPs) encoded by *Spiroplasma*, which selectively interfere with the ribosomes of

the developing parasitoid larvae [177]. Similarly, this effect was also documented in *D. neotestacea,* where *Spiroplasma* protects against parasitic nematodes [178,179].

A well documented negative effect of *Spiroplasma* is male-killing in different *Drosophila* species, where it selectively kills male embryos, leading to skewed sex ratios [180,181]. The mechanism behind this phenomenon involves the production of a specific protein by *Spiroplasma*, *Spiroplasma poulsonii* androcidin (Spaid), which causes apoptosis and neural defects specifically in male embryos by targeting the dosage compensation complex on the male X chromosome, whilst leaving female embryos unaffected [180,182]. This mechanism was also observed in the butterfly *Danaus chrysippus*, resulting in female-biased populations [183]. In honeybees, *Spiroplasma apis* has been associated with significant mortality and colony decline by disrupting tissues, particularly the midgut, where initial infections are often detected [184,185]. In the leafhopper *Circulifer tenellus*, *Spiroplasma citri* is responsible for reduced longevity and fecundity through colonization of host cells and disruption of physiological processes [186,187].

In tsetse flies, *Spiroplasma* is subject of recent studies aiming to identify the effects on its host, but relatively few studies have been performed to this date. It was determined, that the *Spiroplasma* infection prevalence is higher in females and males of laboratory-reared *Gff* compared to wild populations, where overall prevalence was lower and *Spiroplasma* was only detected in females [173]. Fluorescent in-situ hybridisation revealed the localization of *Spiroplasma* in reproductive tissues of *Gff*, where it occurs sporadically in ovaries, while being present in testes at high densities, suggesting a potential for paternal transmission of the bacterium [173]. A study investigating the spatio-temporal distribution of *Spiroplasma* presented a negative correlation between *Spiroplasma* and trypanosome coinfections: out of 243 *Spiroplasma*-infected *Gff* flies collected in the field in Uganda, only 2% showed a coinfection with trypanosomes, whilst *Spiroplasma*-uninfected *Gff* showed 10% coinfection. To further address the question of potential interactions, experiments were performed with a laboratory colony of *Gff* that exhibits a varying infection prevalence with *Spiroplasma*. These flies were challenged with trypanosomes and their trypanosome and *Spiroplasma* infection status was assessed 14 days post infection. The same trend as with field flies was observed, namely a statistically significant negative correlation between occurrence of the symbiont and the parasite [188]. Despite the apparent interaction between *Spiroplasma* and trypanosomes in *Gff*, the exact mode of action of *Spiroplasma* in this host remains unknown. It remains to be clarified whether the potentially protective effect of *Spiroplasma* is caused by competition for nutritional resources within the host or through production of certain molecules, which may affect trypanosomes in a detrimental way [188].

1.4.5. *Glossina pallidipes* salivary gland hypertrophy virus (*Gp*SGHV)

Tsetse flies host different viruses such as the *Glossina pallidipes* salivary gland hypertrophy virus (*Gp*SGHV) [189]. *Gp*SGHV can significantly reduce fly survival and fecundity, posing challenges for mass-rearing efforts in support of the SIT [190]. This virus has a large circular dsDNA genome with a size of 190 Kb encoding numerous proteins involved in viral replication and transcription [191]. A large proportion of mass-reared and wild tsetse fly species is infected with *Gp*SGHV, although in most cases the infection remains asymptomatic [192]. Symptomatic infection however leads to hypertrophied salivary glands and a severe reproductive impairment in the flies [189,192]. In field populations, *Gp*SGHV is primarily transmitted vertically from infected mother to offspring [193,194]. In mass-rearing laboratory colonies, the most significant route of transmission is horizontal, occurring during feeding on artificial blood-feeding systems. This horizontal transmission is facilitated by the large number of flies feeding on the same blood batch, with infected flies releasing significant viral loads into the blood [195,196]. The expressed phenotype of *Gp*SGHV is a hindrance for mass-rearing activities, which was exemplified by the decline and eventual collapse of a mass-rearing colony due to this virus [189]. Subsequent investigations revealed that the feeding regime has a significant impact on the prevalence and phenotypic expression of *Gp*SGHV. A clean feeding method was proposed, where each cage of flies was offered a fresh batch of blood, breaking the cycle of horizontal transmission. In combination with the use of antiviral agents supplemented in the blood, this approach successfully lowered viral levels in the laboratory colonies, preventing negative impact on the colony performance [197].

1.4.6. Iflavirus and Negevirus

In addition to the *Gp*SGHV, further viruses have been recently identified in the tsetse fly species *Glossina morsitans morsitans*: *Glossina morsitans morsitans* Iflavirus (*Gmm*IV) and *Glossina morsitans morsitans* Negevirus (*Gmm*NegeV) [198]. *Gmm*IV is a positive-sense, single-stranded RNA virus belonging to the order of *Picornavirales* and has a genome size of 9685 nucleotides encoding a single polyprotein [198]. It has been detected in different tissues of *Gmm*, including the brain, fat bodies, reproductive tissues, milk glands, midgut and salivary glands (Figure 8). This widespread tissue distribution suggests vertical transmission from mother to offspring, as well as the possibility for horizontal transmission during blood feeding [198]. Iflavirus is also present in other insects, predominantly causing covert infections in different hosts [199–201], but can also lead to overt infections in honeybees, with the most prominent example being the honeybee sacbrood and deformed wing virus [202–204]. Despite the high prevalence of *Gmm*IV in a laboratory colony of *Gmm*, *Gmm*IV does not appear to cause overt symptoms in infected flies. However, further research is needed to identify the role of *Gmm*IV in fly physiology, reproductive fitness and its implications for mass-rearing settings and SIT [198,205].

*Gmm*NegeV is categorized as a positive-sense single-stranded RNA virus with a genome size of 8140 nucleotides containing two overlapping open reading frames (ORFs) [198]. ORF1 codes for a polyprotein with non-structural protein domains, while ORF2 codes for three glycosylated proteins related to the virus envelope [198]. Similar to *Gmm*IV, *Gmm*NegeV is found in various fly tissues and is colocalized with *Gmm*IV in the ovaries and somatic tissues [198]. As with *Gmm*IV, it remains unclear whether the virus infection leads to beneficial or adverse effects on tsetse fly reproductive fitness and performance in mass-rearing facilities [198].

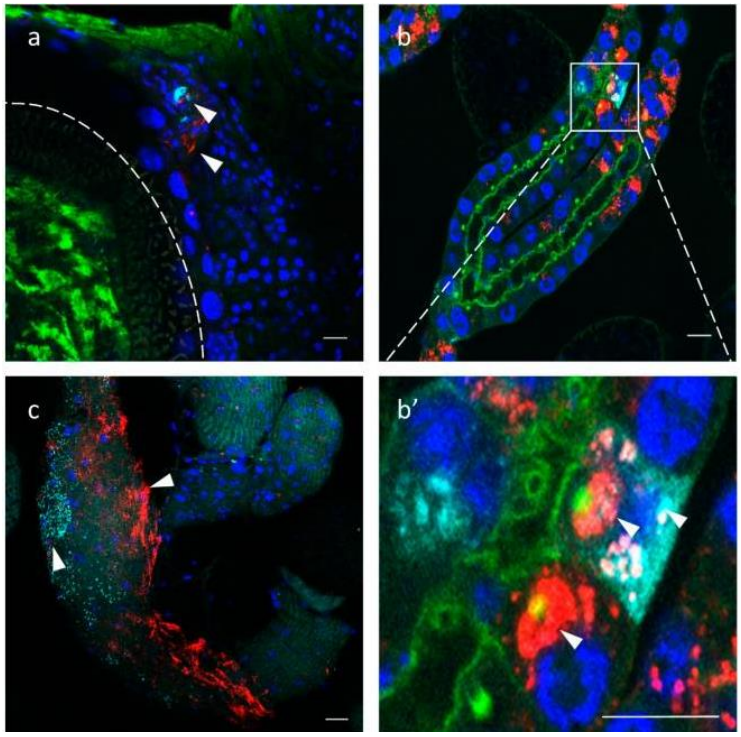

Figure 8: Fluorescent *in situ* hybridisation of Ifla- and Negevirus in gonads and milk glands of 30-day old *Glossina morsitans morsitans*. Iflavirus in red, Negevirus in cyan, *Wolbachia* and Phalloidin in green, nuclei in blue. a: ovaries. b: milk glands. b': increased magnification of milk glands. c: testes. Scale bar: 20 µm *[198]*

Objectives of the dissertation

The two main objectives of this dissertation were the development of microsatellite markers for the tsetse fly species *G. brevipalpis* for population genetics studies in support of vector control strategies and to investigate the interplay between tsetse flies and their associated symbiotic bacteria, specifically *Sodalis* and *Spiroplasma*, and the potential implications for vector competence and disease transmission. All work presented in this dissertation was performed at the Insect Pest Control Laboratory at the Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, located in Seibersdorf (Austria), under supervision of Prof. Dr. Adly M. M. Abdalla and Prof. Dr. rer. nat. Robert L. Mach from Technical University of Vienna. This dissertation is a cumulative work, comprised of six chapters.

Chapter 1 provides a framework for the work performed in this dissertation, starting with an overview of the geographic distribution, biology and systematics of tsetse flies. Subsequently, their role as vectors transmitting parasitic trypanosomes, the causative agents of African sleeping sickness, is discussed, along with the resulting medical and socioeconomic effects. Vector control strategies and the role of population genetics for intervention programmes are outlined, followed by a review about the microbial communities within tsetse flies, from bacterial symbionts to viruses and how they may affect vector competence and the transmission of disease.

Chapter 2 focuses on the development and characterization of microsatellites for the tsetse species *Glossina brevipalpis* to serve population genetics studies, in order to better understand population structure in wild populations and to provide decision-makers of vector control programmes with measures of gene flow and dispersal in the field. The study concludes a set of 9 validated microsatellites which are capable of differentiation between different laboratory and field populations, as well as between different subpopulations from South Africa and Mozambique. These findings may assist the development and execution of appropriate vector control strategies to control the trypanosomosis diseases.

Chapter 3 explores the prevalence of *Trypanosoma* and *Sodalis* in wild populations of tsetse flies, their potential interactions and implications for SIT programmes. It provides prevalence data from 10 different tsetse species across a multitude of sampling locations in 15 countries, allowing for a continent-scale perspective of infection patterns for the first time. This work concludes varying infection patterns according to tsetse fly species and geographic location. Statistically significant positive correlations between trypanosome and *Sodalis* infections were identified for *G. medicorum, G. p. gambiensis and G. pallidipes*, highlighting the significance for SIT applications.

Chapter 4 investigates the negative effects of *Spiroplasma* on the metabolic and reproductive fitness of laboratory-reared *Glossina fuscipes fuscipes.* RNA-sequencing was performed to assess sex-specific gene expression in reproductive tissues in the presence of *Spiroplasma* and functional experiments were undertaken on a heterogeneously infected *Gff* laboratory colony to identify reproductive parameters negatively affected by *Spiroplasma*. The results of this study indicate profound negative effects of *Spiroplasma* infection on the reproductive and metabolic homeostasis, as depicted by prolongation of female gonotrophic cycles due to competition for nutrients between *Spiroplasma* and the host. Furthermore, infections with *Spiroplasma* negatively impact males' sperm motility, therefore adversely affecting the competitiveness of sperm.

Chapter 5 provides a detailed report on the successful *in vitro* cultivation of *Spiroplasma* in cell-free liquid culture medium, identification of *in vitro* growth kinetics and genome sequencing of *Spiroplasma* from *in vitro* culture, laboratory and field-collected *Gff*. Subsequently, comparative genomic analysis is undertaken to identify the genomic synteny of the three genomes, their shared and unique genes, as well as genes of interest and potential genes that are involved in the symbiosis between *Spiroplasma* and *Gff*. Findings of this study deliver new insights into the genome of *Spiroplasma*, highlighting genes that may be involved in host interaction.

Chapter 6 summarizes the main findings of the dissertation and outlines future research directions. It offers a critical evaluation of the importance of the undertaken research and potential implications for vector control programmes and disease transmission. This chapter underscores the significance of the performed work and furthermore, it reflects on how such insights can help to better understand the interactions of the tsetse fly host and symbionts to refine strategies for vector control.

Chapter 2: Development and characterization of microsatellite markers for the tsetse species *Glossina brevipalpis* and preliminary population genetics analyses

Parasite **30**, 34 (2023)

https://doi.org/10.1051/parasite/2023038

PARASITE
Available online at:
www.parasite-journal.org

RESEARCH ARTICLE OPEN ACCESS

Development and characterization of microsatellite markers for the tsetse species *Glossina brevipalpis* and preliminary population genetics analyses

Fabian Gstöttenmayer[1], Percy Moyaba[1,2], Montse Rodriguez[1], Fernando C. Mulandane[3], Hermógenes N. Mucache[3], Luis Neves[3,4], Chantel De Beer[1], Sophie Ravel[5], Thierry De Meeûs[5], Robert L. Mach[6], Marc J. B. Vreysen[1], and Adly M.M. Abd-Alla[1,*]

[1] Insect Pest Control Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, Vienna International Centre, P.O. Box 100, 1400 Vienna, Austria
[2] Epidemiology, Parasites and Vectors, Agricultural Research Council-Onderstepoort Veterinary Research, 100 Soutpan Road, Private Bag X05, Onderstepoort, 0110 South Africa
[3] University Eduardo Mondlane, Centro de Biotecnologia, Av. de Moçambique Km 1.5, Maputo, Mozambique
[4] University of Pretoria, Department of Veterinary Tropical Diseases, Private Bag X04, Onderstepoort 0110, South Africa
[5] University of Montpellier, Cirad, IRD, Intertryp, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France
[6] Institute of Chemical, Environmental, and Bioscience Engineering, Vienna University of Technology, Gumpendorfer Straße 1a, 1060 Vienna, Austria

Received 21 June 2023, Accepted 23 August 2023, Published online 15 September 2023

Abstract – Tsetse flies, the vectors of African trypanosomes are of key medical and economic importance and one of the constraints for the development of Africa. Tsetse fly control is one of the most effective and sustainable strategies used for controlling the disease. Knowledge about population structure and level of gene flow between neighbouring populations of the target vector is of high importance to develop appropriate strategies for implementing effective management programmes. Microsatellites are commonly used to identify population structure and assess dispersal of the target populations and have been developed for several tsetse species but were lacking for *Glossina brevipalpis*. In this study, we screened the genome of *G. brevipalpis* to search for suitable microsatellite markers and nine were found to be efficient enough to distinguish between different tsetse populations. The availability of these novel microsatellite loci will help to better understand the population biology of *G. brevipalpis* and to assess the level of gene flow between different populations. Such information will help with the development of appropriate strategies to implement the sterile insect technique (SIT) in the framework of an area-wide integrated pest management (AW-IPM) approach to manage tsetse populations and ultimately address the trypanosomoses problem in these targeted areas.

Key words: *Glossina*, Tsetse, Microsatellites, Population genetics, South Africa, Mozambique.

Résumé – Développement et caractérisation de marqueurs microsatellites pour l'espèce de mouche tsé-tsé *Glossina brevipalpis* et analyses génétiques préliminaires des populations. Les mouches tsé-tsé, vecteurs des trypanosomes africains, sont d'une importance médicale et économique majeure et l'une des contraintes pour le développement de l'Afrique. La lutte contre la mouche tsé-tsé est l'une des stratégies les plus efficaces et durables utilisées pour contrôler la maladie. La connaissance de la structure de la population et du niveau de flux de gènes entre les populations voisines du vecteur cible est d'une grande importance pour développer des stratégies appropriées pour la mise en œuvre de programmes de gestion efficaces. Les microsatellites sont couramment utilisés pour identifier la structure de la population et évaluer la dispersion des populations cibles et ont été développés pour plusieurs espèces de glossines mais manquaient pour *Glossina brevipalpis*. Dans cette étude, nous avons criblé le génome de *G. brevipalpis* pour rechercher des marqueurs microsatellites appropriés et neuf ont été trouvés suffisamment efficaces pour faire la distinction entre différentes populations de glossines. La disponibilité de ces nouveaux locus microsatellites aidera à mieux comprendre la biologie des populations de *G. brevipalpis* et à évaluer le niveau de flux de gènes entre différentes populations. Ces informations aideront à l'élaboration de stratégies appropriées pour mettre en œuvre la technique de l'insecte stérile dans le cadre d'une approche de lutte antiparasitaire intégrée à l'échelle de la zone pour gérer les populations de glossines et, en fin de compte, résoudre le problème des trypanosomoses dans les zones concernées.

Edited by Jean-Lou Justine
*Corresponding author: a.m.m.abd-alla@iaea.org

Introduction

Tsetse flies (Diptera: Glossinidae) are the only cyclical vectors of two parasitic diseases, human African trypanosomiasis (HAT), or sleeping sickness, and African animal trypanosomosis (AAT), or nagana in livestock [6, 33, 58]. Historically, the disease has put a great burden on the development of the African continent and continues to impact food security, mixed crop livestock agriculture and human health [36]. An estimated 70 million people are at risk of contracting HAT and the economic losses due to AAT are significant [25].

Tsetse flies are distributed in sub-Saharan Africa and are part of the dipteran clade Calyptrate, in the group of blood-feeding Hippoboscidae [69], of which many species are of medical and economic importance [25]. All tsetse fly species belong to the genus *Glossina* Wiedeman, 1830 [39]. They are divided into three subgenera according to the differences of structural complexity of the genitalia, which falls in line with body hair patterns and their habitat preferences: flies of the *fusca* group inhabit the lowland rain forests and the border areas of forests and relic isolated forests, the *palpalis* group is found in lowland rain forest of the coastal region of West Africa and extends to the river systems of the humid savannah and the *morsitans* group is restricted the woodland savannah [36, 62]. *Glossina brevipalpis* belongs to the *fusca* group and is feeding on wild mammals and livestock and therefore represents an important vector of AAT [39]. The predicted geographical distribution of *G. brevipalpis* spans from Ethiopia and Somalia in eastern Africa southwards to Uganda, Kenya, Rwanda, Burundi and Tanzania [70]. Furthermore, it is found throughout southern Africa in Malawi, Zambia, Zimbabwe, Mozambique [44] and up to the southernmost populations in KwaZulu-Natal of South Africa [11].

AAT remains a major constraint for more sustainable and efficient livestock development in 36 countries infested with tsetse flies in sub-Saharan Africa. The affected area is approximately 10 million km^2 of high agricultural and livestock farming potential [24]. Studies project the annual losses in agricultural production due to AAT to be around 3 billion USD, and that the elimination of this burden by coordinated control programmes would need an investment of 12–15 billion USD in a timeframe of 10–15 years [7, 34]. The lack of vaccines against trypanosomoses, the development of resistance to the available trypanocidal drugs against AAT, and the high cost of treatment with chemotherapeutic drugs and their dangerous side effects in humans [27], make the control of tsetse flies an attractive approach for the suppression and management of trypanosomosis [62]. The evolution of tsetse control methods has changed dramatically over the years: from elimination of game and bush clearing during the early 20th century, to insecticide spraying, dipping tanks and insecticide pour-on during the 1950s and the use of insecticide-impregnated traps and tiny targets in recent years [1, 30, 32, 52, 62]. Another control tactic that has proven to be effective is the sterile insect technique (SIT) [61]. SIT, when used as part of an area-wide integrated pest management (AW-IPM) approach, can be used to suppress and even eliminate targeted tsetse fly populations [62]. The SIT requires mass-rearing of the target species in special facilities and sterilisation of males using ionising radiation, followed by systematic and sequential releases in the target area. A sufficient number of sterile males has to be released, so that they can outcompete wild males for mating with wild females. Consequently, mating of a sterile male with a virgin wild female will result in no offspring, which means that the release of massive amounts of such males can considerably reduce the size of the target population at the next generation [62]. An example for the successful implementation of the SIT was the AW-IPM programme on Unguja Island, Zanzibar from 1994 to 1997, where a population of *Glossina austeni* was eliminated and as a consequence, the transmission of trypanosomes on the island was stopped [61].

For the successful implementation of an AW-IPM, knowledge on the level of gene flow between neighbouring populations is required. A tsetse population that is isolated represents the most ideal situation to apply an eradication strategy as it will avoid remigration of flies into the area [20]. This was the case in the Niayes of Senegal, where microsatellites, mitochondrial Cytochrome c oxidase subunit 1 (COI) markers and morphometrics were used to assess population structure. It was found that the tsetse population in the Niayes of Senegal was isolated from the nearest southern populations and, hence, an eradication strategy was selected and implemented [55]. The situation becomes more complex when the targeted insect population is not isolated. Such situations require a "rolling carpet approach" [31] as was implemented during the successful SIT programme against the New World screwworm fly *Cochliomyia hominivorax* in the United States, Mexico and Central America [71].

There are currently several molecular markers available that can be used to gain insight into the population biology of wild species and to assess the level of gene flow between adjacent populations. Among these, microsatellite markers are advantageous, because they do not require sequencing nor a large amount of DNA and are relatively cheap and simple to use [56].

A study using mitochondrial DNA markers and morphometrics was carried out for *G. brevipalpis* and *G. austeni* populations of South Africa, Eswatini (formerly Swaziland) and Mozambique. Although some limitations came with the used markers, the results indicated the absence of barriers to gene flow between the populations in South Africa and Southern Mozambique [13]. The use of microsatellites might help to refine our vision of the structure of these populations and their dispersal. Microsatellite markers have been developed for tsetse species such as *Glossina fuscipes fuscipes* [2, 5, 51], *Glossina palpalis palpalis* [40], *Glossina pallidipes* [47, 48, 51], *Glossina morsitans morsitans* [3, 51] and *Glossina palpalis gambiensis* [51, 54]. No microsatellite loci have been developed specifically for *G. brevipalpis* so far, and attempts to use microsatellite markers developed for *G. pallidipes* gave only limited data for *G. brevipalpis* [48]. Therefore, the aim of this work was to develop novel microsatellite markers for *G. brevipalpis* and to test their suitability for population genetic studies.

Materials and methods

Tsetse fly samples

Glossina brevipalpis samples for this study were collected from two field locations in South Africa and three in Mozambique (Table 1) and from a laboratory colony

Table 1. Locations and geographical coordinates of tsetse fly *Glossina brevipalpis* samples collected from Mozambique and South Africa and origin of laboratory colony maintained in Onderstepoort Veterinary Research (South Africa).

Location	*N*	Label	No of traps	Latitude	Longitude
1. Hluhluwe-Imfolozi Park 1[+] (South Africa)	20	HLUE1	1	−28.09816	32.12231
2. Hluhluwe-Imfolozi Park 2[+] (South Africa)	10	HLUE2	1	−28.10031	32.12247
3. Onderstepoort Veterinary Research originated from Kibwezi Forest (Kenya)[*]	30	OVI (Colony)	0	−2.416917	37.95892
4. Reserva Especial de Maputo 1[#] (Mozambique)	9	MOZ1	1	−26.59884	32.84528
5. Reserva Especial de Maputo 2[#] (Mozambique)	7	MOZ2	1	−26.59246	32.84624
6. Reserva Especial de Maputo 3[#] (Mozambique)	4	MOZ3	3	−26.59177	32.84617

[+] Sampled between 15 October and 10 December 2018.
[#] Sampled between 5 and 7 June 2019.
[*] Approximate location.

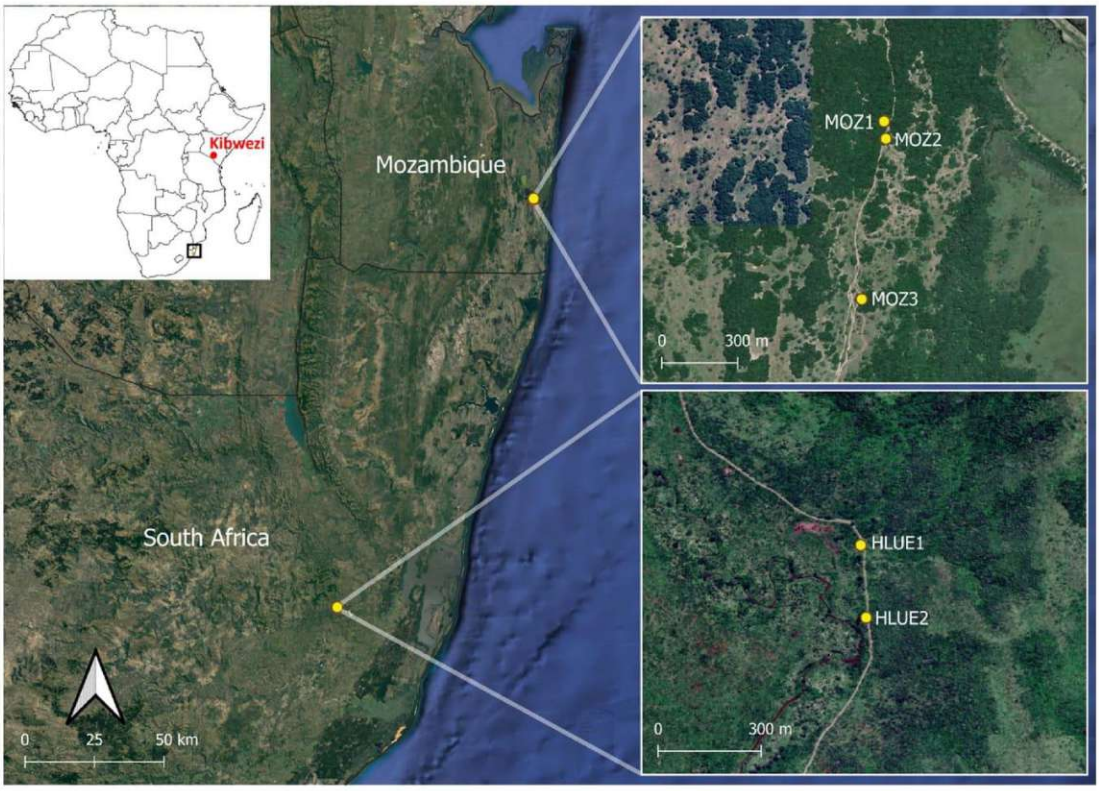

Figure 1. Locations of collected *Glossina brevipalpis* samples in Mozambique and South Africa. Kibwezi Forest (Kenya), the location of collections for Onderstepoort Veterinary Research (OVI) colony is indicated in red.

maintained in the Onderstepoort Veterinary Research (South Africa), which originated from flies caught in the Kibwezi Forest in Kenya 41 years ago. Sampling locations, Number of sampled flies, Population label, Number of traps and Geographical coordinates are indicated in Table 1. Sampling locations are furthermore shown in Figure 1.

Sampling sites in the two countries were separated by a geographical distance of 162 km. Tsetse flies were collected at Hluhluwe-Imfolozi Park (South Africa) from 15 October 2018 to 10 December 2018 and at Reserva Especial de Maputo (Mozambique) from 5 to 7 June 2019 with odour-baited H traps [37]. Assuming a two-month generation time [41], this represents three tsetse generations between the two sampling campaigns. To enhance the trapping of *G. brevipalpis*, the traps were baited with 1-octen-3-ol and 4-methylphenol at a ratio of 1:8 that were released at 4.4 mg/h and 7.6 mg/h, respectively [38]. These chemicals were dispensed from seven heat-sealed sachets (7 cm × 9 cm) made of low-density polyethylene

sleeves (wall thickness 150 microns) placed near the entrance of each trap. A 300 mL brown glass bottle that dispensed acetone through a 6 mm hole in the lid at a rate of ca. 350 mg/h was placed next to the H trap. Flies caught in the traps were guided to a non-return cage that contained a 20% ethanol solution to which an antiseptic, Savlon® (Johnson & Johnson, East London, South Africa) (0.4 mL/L) and formalin (0.4 mL/L) had been added to preserve the sampled flies as well as to prevent ant and spider predation.

Colony flies were obtained from the laboratory colony that has been maintained at the Agricultural Research Council-Onderstepoort Veterinary Research in Pretoria, South Africa for the past 21 years. This colony was originally established from seed material collected in the Kibwezi Forest in Kenya in 1982, with no new genetic material introduced to the colony since then [12]. These colony flies originating from a distant geographic location were included to test the efficiency of the newly developed microsatellites to differentiate between distant and closer field populations.

All collected flies were preserved in absolute ethanol. The samples were prepared for shipment to the Insect Pest Control Laboratory (IPCL) by replacing the ethanol with propylene glycol. Upon arrival, the propylene glycol was replaced by absolute ethanol and stored at −20 °C until DNA extraction.

Sequence analysis and microsatellite selection

Glossina brevipalpis sequence data of the following SRA files (SRR653459, SRR653479, SRR681140 and SRR1174181) available in the SRX220378: Whole Genome Sequencing of tsetse fly project at https://www.ncbi.nlm.nih.gov/sra/ were used. In addition, more sequence data of *G. brevipalpis* were provided by Otto Koekemoer, Agricultural Research Council, Onderstepoort Veterinary Research, South Africa (unpublished data) and used in the analysis. The SRA files were transformed to fastq files using the command "fastq-dump [sra file name]>output file name" in Linux Ubuntu 20.04.4 LTS. The fastq sequence data, consisting of 26209954 reads of 200 bp length were assembled with ABYSS assembler [53] with the command "abyss-pe k=64 np=16 name=sample1_2 in='1_2.fastq.gz'". The .fa files were concatenated to one file with the command "cat file1, file2, filexx>all_file". The concatenated file was screened for di- and trinucleotide motifs with a minimum of ten repeats using MSATCOMMANDER 1.08 [23] which also enables primer design with the integrated PRIMER3 software. The location of the extracted primers on the *G. brevipalpis* complete genome (https://vectorbase.org/vectorbase/app/downloads/release-59/GbrevipalpisIAEA/fasta/data/, accessed on 8 August 2022) and in coding sequence regions (CDC) was determined using Geneious Prime software, version 2022.2.2. Microsatellites with dinucleotide motifs were prioritised as they showed higher levels of polymorphism compared to trinucleotide motifs in a previous study on tsetse flies [4]. A total of 188 primer pairs were selected based on product size (180–380 bp) and the number of repeats (≥13 repeats), synthesized by Eurofins Genomics (Ebersberg, Germany), and tested for microsatellite amplification with PCR.

DNA extraction and PCR amplification

Only male tsetse fly samples were used for the microsatellite validation in this study in order to be able to exclude loci on the X chromosome. Samples were rehydrated in distilled water for 5 min after removal of the ethanol and separated into two sample types: whole body and leg samples. Total genomic DNA of *G. brevipalpis* body samples was extracted using DNeasy Blood & Tissue kit (QIAGEN Inc., Redwood City, CA, USA), following the manufacturer's instructions. DNA of three legs per fly was extracted with a Quick-DNA™ Miniprep Kit (Zymo Research, Irvine, CA, USA). The quantity and quality of extracted DNA was assessed with a Synergy H1 Hybrid Multi-Mode Reader (Agilent Technologies, Inc., Santa Clara, CA, USA).

In the first phase of the study, 188 synthesized primer pairs were tested with DNA extracted from five *G. brevipalpis* flies (whole body samples) to determine primer specificity and the amplicon profile. PCR amplification was carried out in a total reaction volume of 25 µL, with 12.5 µL QIAGEN Taq PCR 2X Master Mix (QIAGEN Inc.), 10 µL nuclease-free H_2O (Qiagen Inc., CA, USA), 0.2 µM of each primer and 1.5 µL (4 ng) of DNA. All primer sequences can be found in Supplementary Table 1. PCR conditions were 94 °C for 2 min; 35 cycles of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 1 min; ending with a final extension at 72 °C for 5 min. The PCR amplification was checked on 2% agarose E-gel™ stained with ethidium bromide (Invitrogen, ThermoFisher Scientific, Waltham, MA, USA). Out of 188 tested primers, 170 successfully amplified the expected band sizes, from which 20 were selected to be tested with 10 leg samples from each population to assess heterogeneity of fragment sizes between and within populations. A set of 12 microsatellites (Gb5, Gb28, Gb35, Gb48, Gb50, Gb66, Gb70, Gb72, Gb73, Gb92, Gb158, Gb165) exhibiting polymorphic patterns between populations was selected to be amplified for the final phase of the study, which involved amplification of the microsatellites with fluorescent dyes (6-FAM, HEX, ATTO 550 and ATTO 565). This was done by synthesizing the forward primers linked with the M13 adapter (5′-CACGACGTTGTAAAACGAC-3′). PCR was performed in a total reaction volume of 25 µL, with 12.5 µL Platinum™ II Hot Start PCR 2X Master Mix (ThermoFisher Scientific), 9.6 µL nuclease-free H_2O (QIAGEN Inc.), 0.016 µM forward primer with M13 adapter, 0.2 µM reverse primer, 0.2 µM M13 adapter labelled with fluorescent dye (6-FAM, HEX, ATTO 550 or ATTO 565) and 1.5 µL of 1:5 diluted DNA (0.75–26 ng/µL). The PCR conditions were as follows: 94 °C for 2 min; 35 cycles at 94 °C for 15 s, 58 °C for 15 s, and 68 °C for 15 s; ending with a final extension at 68 °C for 5 min. PCR products were checked on 4% agarose E-gel™ stained with ethidium bromide (Invitrogen). PCR products were then resolved on an ABI 3500XL Genetic Analyzer (Applied Biosystems, Waltham, MA, USA) with a GeneScan™ 600 LIZ™ internal size standard (ThermoFisher Scientific).

Quality assessment of *G. brevipalpis* microsatellite markers

Raw data reads were processed and allele calling performed in Genemapper version 6 (ThermoFisher Scientific). Allele calls

were transformed into a codominant matrix displaying the microsatellite loci and their respective alleles. The genetic data were formatted for Create v. 1.37 software [10], to convert the datasets into the required formats according to the software used. For Mozambique samples, individual traps were considered subpopulations, except for MOZ3, where samples from three traps, located within 50 m distance, were pooled due to the low number of samples per trap. Each of the two traps from South Africa was considered a subpopulation.

Quality of data was tested with Fstat v. 2.9.4. [28]. Presence of linkage disequilibrium (LD) between each locus pair was checked using *G*-based tests with 10,000 randomisations [29]. Furthermore *p*-values were corrected according to Benjamini and Yekutieli (BY) to assess for the false discovery rate [4] in RStudio v. 2021.09.2 [50]. *F*-statistics, namely Wright's F_{IS} for estimation of deviation from panmixia of genotypic frequencies at local scales (e.g. within subsamples), Wright's F_{ST} for estimation of subdivision, and Wright's F_{IT}, a measure of deviation from panmixia at the whole sample scale (which results from the latter two parameters: $(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST})$ [18]), were estimated using Weir and Cockerham's unbiased estimators [68]. Significant deviation from panmixia and significant subdivision were evaluated with 10,000 permutations of alleles between individuals within subsamples as well as individuals between subsamples, respectively. For the first, the statistic used was simply Weir and Cockerham's estimator of F_{IS} and for the second, the statistic used was the log-likelihood ratio *G* [29]. Confidence intervals were calculated with 5000 bootstraps over all loci.

As departures from Hardy-Weinberg-Equilibrium (HWE) may be caused by a Wahlund effect or genotyping errors such as short allele dominance (SAD), null alleles or stuttering [17], several tests were performed to investigate the influence of these scenarios on HWE. Detection of null alleles, SAD and stuttering was assessed using the strategy described in several papers [14, 16, 17, 19, 41]. The frequency of null alleles was estimated with the EM algorithm [21] using FreeNA. One-sided exact binomial tests were performed in RStudio v. 2021.09.2 [50] to test the goodness of fit of expected null homozygotes and observed missing data (putative observed null homozygotes).

Genetic differentiation

Measure of genetic differentiation was assessed with Wright's F_{ST}, corrected for the presence of null alleles with the excluding null alleles (ENA) method implemented in FreeNA [9], and labelled F_{ST_FreeNA}. The 95% confidence intervals with 5000 bootstraps over loci were also computed with FreeNA. For these computations, missing data were recoded as homozygotes for null alleles, i.e. 999,999, as recommended [9]. A standardised measure, corrected for the excess of polymorphism, was obtained with $F_{ST\_FreeNA}' = F_{ST\_FreeNA}/F_{ST\_Fmax}$, where F_{ST_max} was calculated using the software Recodedata [42]. We computed these quantities between each pair of subsamples within the two countries and computed the averages of the means across loci and of the 95% confidence intervals. We also computed $G_{ST}'' = [n(H_T - H_S)]/[(nH_T - H_S)(1 - H_S)]$ [43], where H_T and H_S are total and local genetic diversities [45] and *n* the number of subsamples. According to Wang [64], when the correlation between Nei's G_{ST} and H_S is negative, F_{ST}' (Meirmans) offers a more accurate estimate of subdivision, while G_{ST}" performs better otherwise. The correlation between G_{ST} and H_S was measured and tested with a one-sided Spearman's rank correlation test with rcmdr.

Effective population sizes (N_e)

Effective population sizes (N_e) – roughly the number of reproducing adults in a population – were estimated for wild tsetse fly samples. We utilised five different methods: linkage disequilibrium [67] corrected for missing data [49] and molecular co-ancestry [46] as computed with NeEstimator v. 2.1. [22]; the heterozygote excess method recently proposed by [15]) $N_e = -(1/2F_{IS}) - (F_{IS}/2(1 + F_{IS}))$ the one- and two-locus identity probabilities [59] with ESTIM v. 1.2. [60]; and Wang's sibship frequency method [65] with the software Colony v. 1.0.6.8 (January 5, 2022) [35]. We then computed the average N_e across methods, weighted with the number of usable values (i.e., values different from 0 or infinity).

Genetic relationships between wild populations and the laboratory colony of *G. brevipalpis*

Cavalli-Sforza and Edwards' chord distances [8] between subpopulations were computed with FreeNA with including null alleles (INA) correction for null alleles [9], which served as input for MEGA11 [57] to construct a Neighbour-Joining tree. To further visualise genetic structure, a Factorial Correspondence Analysis (FCA) was undertaken with GENETIX v. 4.0.5.2. Significance of the 10 first axes was evaluated with the broken stick method [26].

Results

Microsatellite development and validation

The MSATCOMMANDER search for di- and trinucleotides resulted in a total of 55,664 motifs and 21,827 pairs of primers. Out of them, 1244 primer pairs could be found in duplicates in the genome and therefore were excluded from the analysis. Combining the unique primer pairs with the motifs produced 20,583 primers pairs, out of them 13,420 primer pairs flanked repeats of dinucleotides and 7163 flanked trinucleotide repeats. Mapping the unique primers (no duplicates) to the *G. brevipalpis* genome indicated that all primers (41,166) mapped to the genome. When mapping the primers to the coding sequence regions (CDS), it showed that 24,609 primers matched to CDS regions, while 16,557 matched to sequences between CDS. Selecting primer pairs that produced PCR product ranging between 180 and 300 nucleotides resulted in 7328 primer pairs. Sorting the primer pairs with the number of motifs with a cut-off $\geq$ 13 repeats resulted in 253 primer pairs out of which 188 primer pairs with the highest number of motif repeats were synthesized and tested by PCR (Supplementary Table 1). Out of the 188 primer pairs tested by PCR, 170 (90.4%) showed amplification at the expected fragment size. Among all tested primer pairs, 103 (54.7%) showed monomorphic and 67 (35.6%) indicated polymorphic

Table 2. The 10 polymorphic microsatellite loci selected for *Glossina brevipalpis* including repeat motif, primer sequence, number of alleles (N_A), allele size range, heterozygosity within subsamples (H_s), total heterozygosity (H_t).

Locus name	Repeat motif	Primer sequence (5-3)	N_A	Allele size range (bp)[a]	H_s	H_t
Gb5	AG(21)	F: GTTACTAGTCACTCATAAGCACATG R: CTCATATATGCTTCTTGTTGGTCC	13	219–247	0.851	0.899
Gb28	AG(18)	F: CTACTACCAACTCCAATCAAAGAAG R: TCTACTCGCTAAGAAGTAAATCCTG	12	227–281	0.799	0.867
Gb35	AC(17)	F: TTTCTCCTTAGCCTGCATTAGATAG R: AAACACACACTAACTGAGAGAGG	10	193–239	0.673	0.783
Gb48	AC(16)	F: CTAAACCTCTAGTGAAACAGAAGC R: CTATAACATAACTCTAACGCTACGG	14	229–279	0.809	0.867
Gb66	ATC(16)	F: CAAACCTGTTGCATCTTTGTTTGG R: CCTCGTCACTACCATCGTC	14	202–259	0.861	0.893
Gb72	AC(15)	F: TTGTCACATATATGAGAACGACCTC R: GAAGTTCACTACAATGTCATCTGC	14	207–245	0.798	0.841
Gb73	AC(15)	F: TTTCTGATGTTGGAGAGTGTCTAC R: CCATTTGATTCACGATAACCAACC	11	228–256	0.764	0.849
Gb92	AG(15)	F: GAGAGCGAGAGACAACATTTC R: TTTCGTTATCTCCCTCTTTCGTTTG	13	278–324	0.755	0.802
Gb158	AC(13)	F: CACCTCAGCCTTCATAAATTAACAC R: GCTAACAAAGAGGAGTATGTAC	8	187–205	0.744	0.805
Gb165	AC(13)	F: TTTGGGAGAAACACGTACGAC R: GTTGTATACTTAAGCAACGCACAC	5	210–220	0.534	0.604

[a] Length including M13 adapter.

amplicons. From the microsatellites showing polymorphism, 12 (Gb5, Gb28, Gb35, Gb48, Gb50, Gb66, Gb70, Gb72, Gb73, Gb92, Gb158, and Gb165) were selected for population genetics analyses.

Quality assessment of microsatellite loci

Out of the 12 microsatellites, Gb50 was excluded due to the low success of fragment analysis reads. Gb70 was considered X-linked as it displayed different amplicon patterns with DNA from females comparing to males in preliminary tests (Supplementary Figure 1). Therefore, it was excluded from analysis due to X-linkage. Overall, 10 microsatellites were subjected to the quality control tests (Supplementary Table 2). Primer sequences, microsatellite motifs, number of alleles, allele size range and genetic diversities of each locus are presented in Table 2. The quality parameters, namely F_{IS}, F_{ST}, linkage disequilibrium (LD), short allele dominance (SAD), stuttering and presence of null alleles were first assessed for 10 loci on all tested samples. The *G*-based tests for LD between each pair of loci indicated that four locus pairs were in disequilibrium (Gb5xGb66, Gb28xGb72, Gb28xGb92 and Gb35x92); however, none of the pairs gave a significant *p*-value at the BY level. The global analysis over all loci and all samples indicated a significant heterozygote deficit $F_{IS} = 0.079$ in 95% CI [−0.052, 0.2101] (*p*-value = 0.0002), as indicated in Figure 2. Genetic differentiation was significant, with minor variation across loci: $F_{ST} = 0.115$ in 95% CI [0.089, 0.141] (*p*-value = 0.0001) (Supplementary Figure 2).

The correlation between F_{IS} and F_{ST} across all loci was not significant (*p*-value: 0.6336) with a trend to be negative

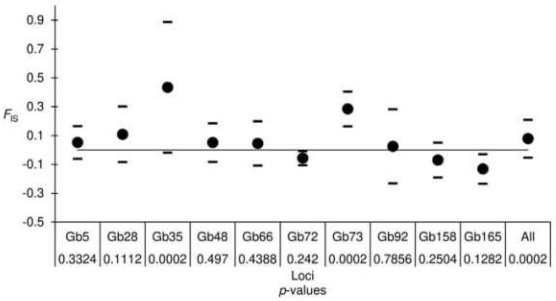

Figure 2. Average homozygosity index (F_{IS}) by locus with upper and lower limit as calculated with the formulas Li = F_{IS} − StdErr × *t* and Ls = F_{IS} + StdErr × *t* and over all loci (All) estimated over all subpopulations. The 95% confidence interval for all loci was obtained by jackknife resampling over populations for each locus and by 10,000 bootstraps over loci for the average (All). The two-sided *p*-values obtained while testing for significant deviation from panmixia and the number of observed missing data are also indicated under locus names.

($\rho = -0.115$), indicating that null alleles do not affect F_{ST} enough to result in a significant positive correlation with F_{IS} (Supplementary Figure 3). The Spearman correlation between the F_{IS} and the observed missing genotypes (blanks) for some loci (Gb5, Gb35, Gb66 and Gb73) was significantly positive (*p*-value = 0.04167) and at least 95% of the variance of F_{IS} is explained by the number of blank genotypes found across these loci (Supplementary Figure 4). Other loci displayed an excess of missing genotypes. No significant stuttering signature was

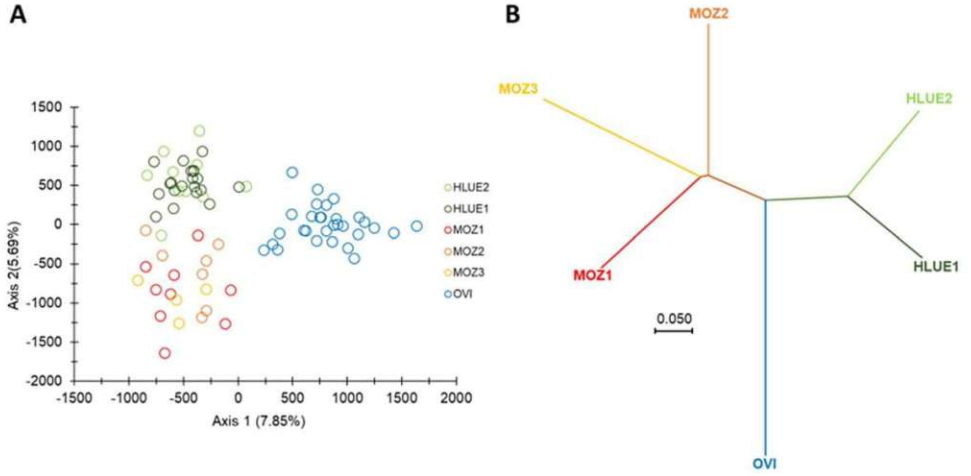

Figure 3. Population structure of *G. brevipalpis*. (A) Factorial Correspondence Analysis of six *G. brevipalpis* field and laboratory colony populations on nine selected loci as computed with GENETIX v. 4.0.5.2. For each axis, percentages of inertia are given. (B) Neighbour-joining tree constructed with MEGA11 based on Cavalli-Sforza and Edward's chord distances between individuals. Branches are labelled according to the sampling site: HLUE 1 and HLUE 2 (South Africa 1 and 2) – green; OVI (Onderstepoort Veterinary Institute) – blue and MOZ 1, 2 and 3 (Mozambique 1, 2 and 3) – red.

observed for all tested microsatellites. However, an indication of SAD was detected for the locus Gb48 with a p-value of 0.09867 and negative ρ (−0.367033) with Spearman's rank correlation test. The weighted regression correlation test was significant (p-value = 0.0456), therefore Gb48 was excluded from further analysis. In conclusion, nine microsatellites (Gb5, Gb28, Gb35, Gb66, Gb72, Gb73, Gb92, Gb158 and Gb165) were retained for further analysis after excluding Gb50 for low success of fragment analyser reads, Gb70 for X-linkage and Gb48 for SAD.

Analysing the data from all samples with the nine selected loci resulted in a total heterozygosity (H_t) of 0.816. The overall F_{IS} decreased to 0.082 (Supplementary Figure 5A). Genetic differentiation remained significant (p-value = 0.0001), with a marginal increase of F_{ST} to 0.116. Four locus pairs were in significant LD (Gb5xGb66, Gb28xGb72, Gb28xGb92 and Gb35xGb92); again none of these pairs remained significant after BY-correction. The correlation test between F_{IS} and F_{ST} remained non-significant (p-value: 0.6206) (Supplementary Figure 5B).

Genetic differences between laboratory colony and wild flies from Mozambique and South Africa

To determine the genetic relationship between the *G. brevipalpis* from a colony maintained in culture for 40 years and the field collected flies, we conducted an FCA (individual based) and NJ-tree analysis based on Cavalli-Sforza and Edwards chord distances between subsamples, with INA correction for null alleles.

The OVI laboratory colony appeared distant from the other two wild subpopulations (Figure 3). Regarding the FCA, this differentiation was the main contributor of the first axis of the FCA, while the second axis clearly separated wild subsamples from Mozambique and South Africa, though with less strength (Figure 3B). Furthermore, as seen in the NJ-tree, OVI obviously represents an outgroup.

Quality assessment of selected microsatellite loci for field samples only

In the field samples only, genetic differentiation remained significant, with F_{ST} = 0.067 in 95% CI = [0.043, 0.097] (p-value < 0.0001). There was still a significant heterozygote deficit: F_{IS} = 0.068 in 95% CI = [−0.02, 0.164] (p-value < 0.0002) (Supplementary Figure 6).

As explained above, null alleles explained most of F_{IS} variations. We always observed enough numbers of blanks as compared to the expected ones as provided by FreeNA with the EM algorithm to estimate null allele frequencies (all p-values > 0.7). The regression of F_{IS} as a function of expected null homozygotes (Nblanks – expected) provided a R^2 = 0.7251 and an intercept of $F_{IS\_no\text{-}null}$ = −0.0247 (Supplementary Figure 7), as expected in a random mating dioecious population [15]. After correction for null alleles with FreeNA and correction for polymorphism using RecodeData, the average genetic differentiation between subsample pairs across the two countries was very high: F_{ST_FreeNA}' = 0.3159 in 95% CI = [0.180, 0.457]. Nevertheless, this cannot be translated into an estimate of gene flow since we cannot separate the respective effect of temporal (three generations) and geographic (>150 km) distances.

Effective population sizes (N_e)

Results of the estimation of effective population sizes are presented in Table 3, where it is indicated that the F_{IS} based

Table 3. Average effective population size obtained for the F_{IS}-based, LD-based, Coancestry-based, the one and two locus correlation-based and the sibship frequency-based methods. For each, the minimum and maximum values are given. These were used to compute the grand average, weighted by the number of usable values (Weights) (values different from 0 or infinity).

Methods	Average N_e	N_{e_min}	N_{e_max}	Weights
FIS	10	4	17	5
LD	364	54	965	3
Coancestry	29	8	87	4
Correlations	14	14	14	1
Sibship	19	11	24	5
Grand average	76	16	192	

N_e = 10, in minmax = [4, 17] is very close to the one obtained with the intercept of the regression $F_{IS} \sim N_{blanks\text{-}expected}$ (N_e = 10.51). This brings confidence on the interpretation regarding the effect of null alleles. The weighted grand average gave effective population sizes: N_e = 76 in minmax = [16, 192].

Genetic differentiation between subsamples within South Africa and Mozambique

Within each country, the average genetic differentiation was small and not significant with F_{ST_FreeNA}' = 0.0983 in 95% CI = [−0.0160, 0.2953] (p-value = 0.3208) in South Africa and F_{ST_FreeNA}' = −0.0206 in 95% CI = [−0.1344, 0.1036] (p-value = 0.8036) on average between traps in Mozambique.

Discussion

This study aimed to develop novel microsatellite markers for *G. brevipalpis* and to evaluate the suitability and efficiency of these markers to investigate population structure and the level of gene flow between populations in the field. The search for microsatellites across the genome sequence of *G. brevipalpis* indicated that dinucleotide repeats were more common than trinucleotide repeats confirming the theory that the microsatellite abundance decreases with the increase of motif repeat number and repeat length [63]. To our knowledge, this is the first report of a set of microsatellite markers that can be used effectively for *G. brevipalpis,* allowing a broader analysis of the population genetics of this species in southeast Africa.

The selected microsatellite markers used in this study were inspected for their efficiency to explore genetic heterogeneity within and between tested populations using several quality control tests. As a result of the quality control tests, nine markers were retained, with an average genetic diversity H_t = 0.816, which is a satisfying rate for population genetics studies. Among these nine loci, four (*Gb5*, *Gb35*, *Gb73* and *Gb92*) displayed null alleles at different levels of frequency. As for the F_{ST} between the two countries, given the importance of the geographic distances between the two countries, relative to the modest number of generations separating those, it is probable that most of the genetic differentiation observed between the two countries, as measured by F_{ST_FreeNA}', was due to geography, though we cannot quantify by what proportion exactly. The same consideration applies to the FCA and NJ-tree analyses. Nevertheless, we can forecast that contemporaneous subsamples from the two zones would lead to a considerable genetic differentiation and thus that these two zones exchange very few immigrants per generation. Regarding effective population sizes, variations of N_e estimates are often important, with the LD-based estimate giving the highest values and the coancestry-based estimate providing the smallest values [15, 22, 66].

Results of the NJ-tree and FCA indicated that the OVI colony was distant from the wild populations of South Africa and Mozambique. The results agree with the results of de Beer *et al.* [13] who's morphometric and mitochondrial DNA analysis indicated that the flies from the colony could be clearly distinguished from the field samples of Eswatini, Mozambique and South Africa. However, using the selected microsatellite markers in this study, it was possible to classify the flies collected from Mozambique and South Africa into distinct groups. The difference between our results and those of de Beer *et al.* [13] can be explained by the fact that mtDNA is a conserved region and is not subject to many frequent mutations as compared to microsatellite loci.

The distinction between *G. brevipalpis* collected from different locations in Mozambique and South Africa demonstrates that microsatellite markers will provide an enhanced resolution of the genetic structure of this species and therefore allow accurate investigations of the population genetics, the immigration rate and the dispersal distances in this species. Although these results demonstrate the suitability of the selected loci to explore the genetic diversity of *G. brevipalpis* in the tested locations, the low number of tested samples, and the fact that Mozambique and South Africa populations were investigated at different times did not allow us to calculate dispersal or the immigration rates between these populations. Therefore, the genetic structure of these populations should not be considered finally established. For future studies, it is recommended to analyse more samples collected at the same time from more locations to obtain a complete and conclusive analysis of the population genetics of the flies.

Conclusions

In this study, nine selected microsatellite markers were characterised and found to be suitable for analysing the population genetics of *G. brevipalpis*. The selected microsatellite markers showed the possibility to differentiate between wild flies from different locations in Mozambique and South Africa as well as the flies from a laboratory population. Analysing more field samples collected from more locations of the same tsetse generation with these microsatellite markers will provide a better understanding of the population genetics and dynamics of *G. brevipalpis*. It will allow for precise assessment of the level of gene flow between adjacent populations that could be targeted with an area-wide integrated pest management strategy.

Acknowledgements. The authors would like to acknowledge colleagues and technicians at the Insect Pest Control Laboratory in Seibersdorf, Austria for providing guidance in the laboratory and supplying materials when needed.

Author contributions

A.M.M.A. and M.J.B.V. conceived and designed research. P.M., F.C.M., H.N.M, L.N., C.D.B., collected biological materials from the field. F.G., M.R., P.M. and A.M.M.A. collected data and conducted research. F.G., T.D.M. and A.M.M.A. analysed and interpreted data. F.G., A.M.M.A. and T.D.M. wrote the initial paper; F.G., T.D.M., M.J.B.V., S.R. and R.L.M. revised the paper; A.M.M.A. had primary responsibility for final content. All authors read and approved the final manuscript.

Funding

This study was supported by the Joint FAO/IAEA Insect Pest Control Subprogramme and the IAEA's Department of Technical Cooperation.

Data availability statement

Materials described in the paper, including all relevant raw data, are available in this link: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/SDRST2. Unpublished sequence data from Otto Koekemoer is available upon reasonable request from the corresponding author.

Ethics approval and consent to participate

Not applicable.

Conflicts of interest

The authors declare that they have no conflict of interest.

Supplementary materials

The Supplementary materials of this article are available at https://www.parasite-journal.org/10.1051/parasite/2023038/olm

Supplementary Figure 1: Determination of X-linkage for locus Gb70. PCR products were migrated on 4% agarose gel stained with ethidium bromide. The first five samples are males, the last five females. Template DNA of positive control (+) was derived from 1 female sample. M: FastRuler Low Range DNA Ladder (ThermoFisher) with band size top to bottom are 1 500, 850, 400 and 200 bp.

Supplementary Figure 2: F_{ST} values with *p*-values for subdivision of 10 loci across all samples. Dots indicate mean values, dashes show lower and upper limits calculated with the following formula: Li = F_{ST} − StdErr × *t* and Ls = F_{ST} + StdErr × *t*. Dashes for All loci are the 95% confidence intervals as obtained by jackknife resampling over populations for each locus and by 10,000 bootstraps over loci for the average.

Supplementary Figure 3: Correlation between F_{IS} and F_{ST} and between F_{IS} and the number of blanks. (A) Correlation between F_{IS} and F_{ST} for 10 loci across all samples. (B) Correlation between F_{IS} and number of blanks for all loci across all samples. (C) Correlation between F_{IS} and number of blanks for 4 selected loci (Gb66 [Nb: 1], Gb 5 [Nb: 2], Gb73 [Nb: 5] and Gb35 [Nb: 10]) across all samples.

Supplementary Figure 4: Wright's F_{IS} values *p*-values for panmixia (A) and F_{ST} values with *p*-values for subdivision (B) of nine selected loci across all samples. Dots indicate mean values, dashes show lower and upper limits calculated with the following formula: Li = F_{XY} − StdErr × *t* and Ls = F_{XY} + StdErr × *t*. Dashes for All loci are the 95% confidence intervals as obtained by jackknife resampling over populations for each locus and by 10,000 bootstraps over loci for the average.

Supplementary Figure 5: Correlation between F_{IS} and F_{ST} and between F_{IS} and the number of blanks. (A) Correlation between F_{IS} and F_{ST} for selected 9 loci across all samples. (B) Correlation between F_{IS} and number of blanks for selected nine loci across all samples. C: Correlation between F_{IS} and number of blanks for four selected loci (Gb5 [Nb: 0], Gb66 [Nb: 1], Gb73 [Nb: 3] and Gb35 [Nb:6]) across all samples.

Supplementary Figure 6: Wright's F_{IS} values *p*-values for panmixia (A) and F_{ST} values with *p*-values for subdivision (B) of nine selected loci across field samples only. Dots indicate mean values, dashes show lower and upper limits calculated with the following formula: Li = F_{XY} − StdErr × *t* and Ls = F_{XY} + StdErr × *t*. Dashes for All loci are the 95% confidence intervals as obtained by jackknife resampling over populations for each locus and by 10,000 bootstraps over loci for the average.

Supplementary Figure 7: Correlation between F_{IS} and FST (A) and between F_{IS} and number of blanks (B) for selected nine loci across field samples.

Supplementary File 1: RMarkdown for statistics in R of the manuscript "Development and characterization of microsatellite markers for the tsetse species *Glossina brevipalpis* and preliminary population genetics analyses".

Supplementary Table 1: Selected primers for *Glossina brevipalpis* microsatellite that produce PCR product length of 180–300 nucleotides with minimum 13 repeat of each motif.

Supplementary Table 2: List of genotypes (alleles) per locus for tsetse fly *Glossina brevipalpis*. Dye-labelled PCR were analysed by fragment analyser and the resulting data were read with Genemapper.

References

1. Abila PP, Kiendrebeogo M, Mutika GN, Parker AG, Robinson AS. 2003. The effect of age on the mating competitiveness of male *Glossina fuscipes fuscipes* and *G. palpalis palpalis*. Journal of Insect Science, 3, 13.
2. Abila PP, Slotman MA, Parmakelis A, Dion KB, Robinson AS, Muwanika VB, Enyaru JCK, Okedi LMA, Lokedi LM, Aksoy S, Caccone A. 2008. High levels of genetic differentiation between Ugandan *Glossina fuscipes fuscipes* populations separated by Lake Kyoga. PLoS Neglected Tropical Diseases, 2, e242.

3. Baker MD, Krafsur ES. 2001. Identification and properties of microsatellite markers in tsetse flies *Glossina morsitans sensu lato* (Diptera: Glossinidae). Molecular Ecology Notes, 1, 234–236.
4. Benjamini Y, Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. Annals of Statistics, 29, 1165–1188.
5. Brown JE, Komatsu KJ, Abila PP, Robinson AS, Okedi LMA, Dyer N, Donnelly MJ, Slotman MA, Caccone A. 2008. Polymorphic microsatellite markers for the tsetse fly *Glossina fuscipes fuscipes* (Diptera: Glossinidae), a vector of human African trypanosomiasis. Molecular Ecology Resources, 8, 1506–1508.
6. Bruce D. 1915. The Croonian Lectures on trypanosomes causing disease in man and domestic animals in Central Africa: Delivered before the Royal College of Physicians of London. British Medical Journal, 1, 1073–1078.
7. Budd LT. 1999. DFID-funded tsetse and trypanosomosis research and development since 1980. United Kingdom: Department of International Development.
8. Cavalli-Sforza LL, Edwards AW. 1967. Phylogenetic analysis. Models and estimation procedures. American Journal of Human Genetics, 19, 233–257.
9. Chapuis M-P, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution, 24, 621–631.
10. Coombs JA, Letcher BH, Nislow KH. 2008. create: a software to create input files from diploid genotypic data for 52 genetic software programs. Molecular Ecology Resources, 8, 578–580.
11. De Beer CJ, Dicko AH, Ntshangase J, Moyaba P, Taioe MO, Mulandane FC, Neves L, Mdluli S, Guerrini L, Bouyer J, Vreysen MJB, Venter GJ. 2021. A distribution model for *Glossina brevipalpis* and *Glossina austeni* in Southern Mozambique, Eswatini and South Africa for enhanced area-wide integrated pest management approaches. PLoS Neglected Tropical Diseases, 15, e0009989.
12. De Beer CJ, Venter GJ, Vreysen MJB. 2016. Improving the diet for the rearing of *Glossina brevipalpis* Newstead and *Glossina austeni* Newstead: Blood source and collection – processing – feeding procedures. PLoS One, 11, e0168799.
13. De Beer CJ, Venter GJ, Vreysen MJB, Mulandane FC, Neves L, Mdluli S, Koekemoer O. 2019. Using genetic and phenetic markers to assess population isolation within the southernmost tsetse fly belt in Africa. Onderstepoort Journal of Veterinary Research, 86, e1–e8.
14. De Meeûs T, Humair P-F, Grunau C, Delaye C, Renaud F. 2004. Non-Mendelian transmission of alleles at microsatellite loci: an example in *Ixodes ricinus*, the vector of Lyme disease. International Journal for Parasitology, 34, 943–950.
15. De Meeûs T, Noûs C. 2023. A new and almost perfectly accurate approximation of the eigenvalue effective population size of a dioecious population: comparisons with other estimates and detailed proofs. Peer Community Journal, 3, e51.
16. De Meeûs T. 2018. Revisiting FIS, FST, Wahlund effects, and null alleles. Journal of Heredity, 109, 446–456.
17. De Meeûs T, Chan CT, Ludwig JM, Tsao JI, Patel J, Bhagatwala J, Beati L. 2021. Deceptive combined effects of short allele dominance and stuttering: an example with *Ixodes scapularis*, the main vector of Lyme disease in the U.S.A. Peer Community Journal, 1, e40.
18. De Meeûs T, McCoy KD, Prugnolle F, Chevillon C, Durand P, Hurtrez-Boussès S, Renaud F. 2007. Population genetics and molecular epidemiology or how to "débusquer la bête". Infection, Genetics and Evolution, 7, 308–332.
19. De Meeûs T, Noûs C. 2022. A simple procedure to detect, test for the presence of stuttering, and cure stuttered data with spreadsheet programs. Peer Community Journal, 2, e52.
20. De Meeûs T, Ravel S, Rayaisse J-B, Courtin F, Solano P. 2012. Understanding local population genetics of tsetse: the case of an isolated population of *Glossina palpalis gambiensis* in Burkina Faso. Infection, Genetics and Evolution, 12, 1229–1234.
21. Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39, 1–22.
22. Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources, 14, 209–214.
23. Faircloth BC. 2008. msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources, 8, 92–94.
24. Feldmann U, Dyck VA, Mattioli RC, Jannin J. 2005. Potential impact of tsetse fly control involving the sterile insect technique, in Sterile insect technique. Principles and practice in area-wide integrated pest management, Dyck VA, Hendrichs J, Robinson AS, Editors. Springer: Dordrecht, The Netherlands. p. 701–723.
25. Feldmann U, Dyck V, Mattioli R, Jannin J, Vreysen M, Dyck VA, Hendrichs J, Robinson AS. 2021. Impact of tsetse fly eradication programmes using the sterile insect technique, Sterile insect technique. Principles and practice in area-wide integrated pest management, 2nd edn, Dyck VA, Hendrichs J, Robinson AS, Editors. CRC Press: Boca Raton, FL. p. 1051–1080.
26. Frontier S. 1976. Étude de la décroissance des valeurs propres dans une analyse en composantes principales: comparaison avec le modèle du bâton brisé. Journal of Experimental Marine Biology and Ecology, 25, 67–75.
27. Gooding RH, Krafsur ES. 2005. Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. Annual Review of Entomology, 50, 101–123.
28. Goudet J. 2003. FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3, https://www2.unil.ch/popgen/softwares/fstat.htm.
29. Goudet J, Raymond M, de Meeüs T, Rousset F. 1996. Testing differentiation in diploid populations. Genetics, 144, 1933–1940.
30. Green CH. 1994. Bait methods for tsetse fly control. Advances in Parasitology, 34, 229–291.
31. Hendrichs J, Vreysen M, Enkerlin W, Cayol J. 2021. Strategic options in using sterile insects for area-wide integrated pest management, in Sterile insect technique. Principles and practice in area-wide integrated pest management, 2nd edn. Dyck VA, Hendrichs J, Robinson AS, Editors. CRC Press: Boca Raton, FL. p. 841–884.
32. Holmes PH, Torr SJ. 1988. The control of animal trypanosomiasis in Africa: current methods and future trends. Outlook on Agriculture, 17, 54–60.
33. Holmes P. 2013. Tsetse-transmitted trypanosomes – Their biology, disease impact and control. Journal of Invertebrate Pathology, 112, S11–S14.
34. Hursey BS. 2001. The programme against Africa trypanosomiasis: aims, objectives and achievements. Trends in Parasitology, 17, 2–3.
35. Jones OR, Wang J. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources, 10, 551–555.
36. Jordan AM. 1986. Trypanosomiasis control and African rural development. London: Longman.

37. Kappmeier K. 2000. A newly developed odour-baited "H trap" for the live collection of *Glossina brevipalpis* and *Glossina austeni* (Diptera: Glossinidae) in South Africa. Onderstepoort Journal of Veterinary Research, 67, 15–26.
38. Kappmeier K, Nevill EM. 1999. Evaluation of a proposed odour-baited target to control the tsetse flies *Glossina brevipalpis* and *Glossina austeni* (Diptera: Glossinidae) in South Africa. Onderstepoort Journal of Veterinary Research, 66, 327–332.
39. Krafsur E. 2009. Tsetse flies: genetics, evolution, and role as vectors. Infection, Genetics and Evolution, 9, 124–141.
40. Luna C, Bonizzoni M, Cheng Q, Robinson AS, Aksoy S, Zheng L. 2001. Microsatellite polymorphism in tsetse flies (Diptera: Glossinidae). Journal of Medical Entomology, 38, 376–381.
41. Manangwa O, De Meeûs T, Grébaut P, Ségard A, Byamungu M, Ravel S. 2019. Detecting Wahlund effects together with amplification problems: cryptic species, null alleles and short allele dominance in *Glossina pallidipes* populations from Tanzania. Molecular Ecology Resources, 19, 757–772.
42. Meirmans PG. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution; International Journal of Organic Evolution, 60, 2399–2402.
43. Meirmans PG, Hedrick PW. 2011. Assessing population structure: F(ST) and related measures. Molecular Ecology Resources, 11, 5–18.
44. Moloo SK. 1993. The distribution of *Glossina* species in Africa and their natural hosts. International Journal of Tropical Insect Science, 14, 511–527.
45. Nei M, Chesser RK. 1983. Estimation of fixation indices and gene diversities. Annals of Human Genetics, 47, 253–259.
46. Nomura T. 2008. Estimation of effective number of breeders from molecular coancestry of single cohort sample. Evolutionary Applications, 1, 462–474.
47. Ouma JO, Cummings MA, Jones KC, Krafsur ES. 2003. Characterization of microsatellite markers in the tsetse fly, *Glossina pallidipes* (Diptera: Glossinidae). Molecular Ecology Notes, 3, 450–453.
48. Ouma JO, Marquez JG, Krafsur ES. 2006. New polymorphic microsatellites in *Glossina pallidipes* (Diptera: Glossinidae) and their cross-amplification in other tsetse fly taxa. Biochemical Genetics, 44, 471–477.
49. Peel D, Waples RS, Macbeth GM, Do C, Ovenden JR. 2013. Accounting for missing data in the estimation of contemporary genetic effective population size (N(e)). Molecular Ecology Resources, 13, 243–253.
50. R Studio Team. 2021. RStudio: Integrated Development for R, http://www.rstudio.com/.
51. Ravel S, Séré M, Manangwa O, Kagbadouno M, Mahamat MH, Shereni W, Okeyo WA, Argiles-Herrero R, De Meeûs T. 2020. Developing and quality testing of microsatellite loci for four species of *Glossina*. Infection, Genetics and Evolution, 85, 104515.
52. Rayaisse J-B, Courtin F, Mahamat MH, Chérif M, Yoni W, Gadjibet NMO, Peka M, Solano P, Torr SJ, Shaw APM. 2020. Delivering "tiny targets" in a remote region of southern Chad: a cost analysis of tsetse control in the Mandoul sleeping sickness focus. Parasites & Vectors, 13, 419.
53. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol İ. 2009. ABySS: a parallel assembler for short read sequence data. Genome Research, 19, 1117–1123.
54. Solano P, Duvallet G, Dumas V, Cuisance D, Cuny G. 1997. Microsatellite markers for genetic population studies in *Glossina palpalis* (Diptera: Glossinidae). Acta Tropica, 65, 175–180.
55. Solano P, Kaba D, Ravel S, Dyer NA, Sall B, Vreysen MJB, Seck MT, Darbyshir H, Gardes L, Donnelly MJ, De Meeus T, Bouyer J. 2010. Population genetics as a tool to select tsetse control strategies: suppression or eradication of *Glossina palpalis gambiensis* in the Niayes of Senegal. PLoS Neglected Tropical Diseases, 4, e692.
56. Solano P, Ravel S, de Meeûs T. 2010. How can tsetse population genetics contribute to African trypanosomiasis control? Trends in Parasitology, 26, 255–263.
57. Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution, 38, 3022–3027.
58. Taylor K, Authié EML. 2004. Pathogenesis of animal trypanosomiasis, in The trypanosomiases, 1st edn. Maudlin I, Holmes PH, Miles MA, Editors. CABI Publishing: UK. p. 331–353.
59. Vitalis R, Couvet D. 2001. Estimation of effective population size and migration rate from one- and two-locus identity measures. Genetics, 157, 911–925.
60. Vitalis R, Denis D. 2001. ESTIM 1.0: a computer program to infer population parameters from one- and two-locus gene identity probabilities. Molecular Ecology Notes, 1, 354–356.
61. Vreysen MJB, Saleh KM, Ali MY, Abdulla AM, Zhu Z-R, Juma KG, Dyck VA, Msangi AR, Mkonyi PA, Feldmann HU. 2000. *Glossina austeni* (Diptera: Glossinidae) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. Journal of Economic Entomology, 93, 123–135.
62. Vreysen MJB, Seck MT, Sall B, Bouyer J. 2013. Tsetse flies: their biology and control using area-wide integrated pest management approaches. Journal of Invertebrate Pathology, 112 (Suppl), S15–S25.
63. Wang H-L, Yang J, Boykin LM, Zhao Q-Y, Wang Y-J, Liu S-S, Wang X-W. 2014. Developing conversed microsatellite markers and their implications in evolutionary analysis of the *Bemisia tabaci* complex. Scientific Reports, 4, 6351.
64. Wang J. 2015. Does GST underestimate genetic differentiation from marker data? Molecular Ecology, 24, 3546–3558.
65. Wang J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. Molecular Ecology, 18, 2148–2164.
66. Wang J. 2016. A comparison of single-sample estimators of effective population sizes from genetic marker data. Molecular Ecology, 25, 4692–4711.
67. Waples R. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics, 7, 167–184.
68. Weir BS, Cockerham CC. 1984. Estimating *f*-statistics for the analysis of population structure. Evolution, 38, 1358–1370.
69. Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim J-W, Lambkin C, Bertone MA, Cassel BK, Bayless KM, Heimberg AM, Wheeler BM, Peterson KJ, Pape T, Sinclair BJ, Skevington JH, Blagoderov V, Caravas J, Kutty SN, Schmidt-Ott U, Kampmeier GE, Thompson FC, Grimaldi DA, Beckenbach AT, Courtney GW, Friedrich M, Meier R, Yeates DK. 2011. Episodic radiations in the fly tree of life. Proceedings of the National Academy of Sciences, 108, 5690–5695.
70. Wint W, Rogers D. 2000. Predicted distributions of tsetse in Africa. Report of the environmental research group Oxford Ltd and TALA research group. Rome: Department of Zoology, University of Oxford for the Animal Health Service of the Animal Production and Health Division of the Food and Agriculture Organisation of the United Nations.
71. Wyss JH. 2000. Screwworm eradication in the Americas. Annals of the New York Academy of Sciences, 916, 186–193.

Cite this article as: Gstöttenmayer F, Moyaba P, Rodriguez M, Mulandane FC, Mucache HN, Neves L, De Beer C, Ravel S, De Meeûs T, Mach RL, Vreysen MJB & Abd-Alla AMM. 2023. Development and characterization of microsatellite markers for the tsetse species *Glossina brevipalpis* and preliminary population genetics analyses. Parasite **30**, 34.

Chapter 3: Prevalence of *Trypanosoma* and *Sodalis* in wild populations of tsetse flies and their impact on sterile insect technique programmes for tsetse eradication

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

Prevalence of *Trypanosoma* and *Sodalis* in wild populations of tsetse flies and their impact on sterile insect technique programmes for tsetse eradication

Mouhamadou M. Dieng[1], Kiswend-sida M. Dera[1,2], Percy Moyaba[3], Gisele M. S. Ouedraogo[2], Guler Demirbas-Uzel[1], Fabian Gstöttenmayer[1], Fernando C. Mulandane[4], Luis Neves[4,5], Sihle Mdluli[6], Jean-Baptiste Rayaisse[7], Adrien M. G. Belem[8], Soumaïla Pagabeleguem[2,9], Chantel J. de Beer[1,3], Andrew G. Parker[10], Jan Van Den Abbeele[11], Robert L. Mach[12], Marc J. B. Vreysen[1] & Adly M. M. Abd-Alla[1]✉

The sterile insect technique (SIT) is an environment friendly and sustainable method to manage insect pests of economic importance through successive releases of sterile irradiated males of the targeted species to a defined area. A mating of a sterile male with a virgin wild female will result in no offspring, and ultimately lead to the suppression or eradication of the targeted population. Tsetse flies, vectors of African *Trypanosoma*, have a highly regulated and defined microbial fauna composed of three bacterial symbionts that may have a role to play in the establishment of *Trypanosoma* infections in the flies and hence, may influence the vectorial competence of the released sterile males. *Sodalis* bacteria seem to interact with *Trypanosoma* infection in tsetse flies. Field-caught tsetse flies of ten different taxa and from 15 countries were screened using PCR to detect the presence of *Sodalis* and *Trypanosoma* species and analyse their interaction. The results indicate that the prevalence of *Sodalis* and *Trypanosoma* varied with country and tsetse species. Trypanosome prevalence was higher in east, central and southern African countries than in west African countries. Tsetse fly infection rates with *Trypanosoma vivax* and *T. brucei* sspp were higher in west African countries, whereas tsetse infection with *T. congolense* and *T. simiae*, *T. simiae* (*tsavo*) and *T. godfreyi* were higher in east, central and south African countries. *Sodalis* prevalence was high in *Glossina morsitans morsitans* and *G. pallidipes* but absent in *G. tachinoides*. Double and triple infections with *Trypanosoma* taxa and coinfection of *Sodalis* and *Trypanosoma* were rarely observed but it occurs in some taxa and locations. A significant Chi square value (< 0.05) seems to suggest that *Sodalis* and *Trypanosoma* infection correlate in *G. palpalis gambiensis*, *G. pallidipes* and *G. medicorum*. *Trypanosoma* infection seemed significantly associated with an increased density of *Sodalis* in wild *G. m. morsitans* and *G. pallidipes* flies, however, there was no significant impact of *Sodalis* infection on trypanosome density.

[1]Insect Pest Control Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, 1400 Vienna, Austria. [2]Insectarium de Bobo Dioulasso-Campagne d'Eradication de la mouche tsetse et de la Trypanosomose (IBD-CETT), 01 BP 1087, Bobo Dioulasso 01, Burkina Faso. [3]Epidemiology, Vectors and Parasites, Agricultural Research Council-Onderstepoort Veterinary Research, Pretoria, South Africa. [4]University Eduardo Mondlane, Centro de Biotecnologia, Av. de Moçambique Km 1.5, Maputo, Mozambique. [5]Department of Veterinary Tropical Diseases, University of Pretoria, Private Bag X04, Onderstepoort 0110, South Africa. [6]Epidemiology Unit, Department of Veterinary Services, PO Box 4192, Manzini, Eswatini. [7]Centre International de Recherche-Développement sur l'Elevage en zone Subhumide (CIRDES), 01 BP 454, Bobo-Dioulasso 01, Burkina Faso. [8]Universite' Nazi Boni (UNB), Bobo-Dioulasso, Burkina Faso. [9]University of Dedougou, B.P. 176, Dédougou 01, Burkina Faso. [10]Roppersbergweg 15, 2381 Laab im Walde, Austria. [11]Institute of Tropical Medicine Antwerp (ITM), Antwerp, Belgium. [12]Institute of Chemical, Environmental, and Bioscience Engineering, Vienna University of Technology, Gumpendorfer Straße 1a, 1060 Vienna, Austria. ✉email: a.m.m.abd-alla@iaea.org

Abbreviations

SIT	Sterile insect techniques
qPCR	Quantitative polymerase chain reaction
BKF	Burkina Faso
ETH	Ethiopia
GHA	Ghana
GUI	Guinea
KEN	Kenya
MLI	Mali
MOZ	Mozambique
SAF	South Africa
SWA	Eswatini
ZAI	Democratic Republic of the Congo
ZAM	Zambia
ZIM	Zimbabwe
Ga	*Glossina austeni*
Gb	*G. brevipalpis*
Gff	*G. fuscipes fuscipes*
Gmm	*G. morsitans morsitans*
Gmsm	*G. m. submorsitans*
Gpg	*G. palpalis gambiensis*
Gpp	*G. p. palpalis*
Tc	*Trypanosoma congolense*
Tv	*Trypanosoma vivax*
Tz	*T. brucei* sspp.

Tsetse flies (Diptera: Glossinidae) are distributed in sub-Saharan Africa between 15° north and 26° south latitude[1]. *Glossina* spp. are the cyclic vectors[2] of unicellular protozoa of the genus *Trypanosoma* that cause African animal trypanosomosis (AAT) or nagana and human African trypanosomosis (HAT) or sleeping sickness[3,4]. Nagana in cattle is mainly caused by *T. congolense*, *T. vivax* and *T. brucei brucei*[5] and causes annual losses to agriculture estimated at $4.75 billion[6]. In addition, around 35 million doses of trypanocidal drugs are administered to livestock per year for managing AAT[7]. Human African trypanosomosis is fatal without treatment[8] and is caused by two *Trypanosoma* subspecies, i.e. *T. brucei rhodesiense* responsible for the acute form of HAT in East Africa and *T. b. gambiense* for the chronic form of HAT in western and central Africa[9]. The lack of effective vaccines and the development of resistance to the available trypanocidal drugs makes the control of AAT in the vertebrate host unsustainable[10,11]. Consequently, an effective tool to reduce *Trypanosoma* transmission would be the control of the tsetse vector. One effective method to manage populations of tsetse flies is the sterile insect technique (SIT) when used as part of an area-wide integrated pest management (AW-IPM) approach[12,13]. The SIT method relies on the mass-production and sterilization of male flies by ionizing radiation. The sterile males are released in the target area for mating with wild females and the absence of offspring will gradually reduce the density of the targeted tsetse populations[14].

The biological transmission of the *Trypanosoma* species requires the parasite to undergo a series of proliferation and differentiation steps in the tsetse alimentary tract and finally mature into an infective form in the mouthparts (*T. congolense*) or salivary glands (*T. brucei* spp.)[15]. However, tsetse flies are refractory to *Trypanosoma* infection meaning that the probability that *Trypanosoma* ingested during a blood meal complete their developmental cycle in the fly to result in a mature infection is rather low[16–18]. The endogenous bacterial microbiome seems important in providing tsetse flies the natural ability to mitigate *Trypanosoma* infections[19]. Three major endosymbiotic bacteria have been identified in tsetse flies, i.e. *Wigglesworthia glossinidia*, *Sodalis glossinidius* (hereafter mentioned as *Sodalis*) and *Wolbachia pipientis*[20]. Some studies suggested that the obligate mutualist *Wigglesworthia* must be present in the larval stage during the development of a mature tsetse fly to properly develop a well-functioning immune system contributing to a refractory phenotype against *Trypanosoma*[5,19].

Sodalis, the second mutualistic symbiont, can be found in the midgut, hemolymph, muscles, fat body, milk glands, and salivary glands of certain tsetse species and is inherited by the progeny through transovarial transmission[21]. The biological role/importance of *Sodalis* for tsetse remain unclear and needs to be clarified[22]. This symbiont might provide some benefits to the host as flies without *Sodalis* have a significantly shorter lifespan as compared with flies with it[23], however the establishment of a *Sodalis* free colony was feasible[24]. *Sodalis* also presents many ideal characteristics to be used for expressing molecular effectors in paratransgenic tsetse[25]. In addition, previous work suggested that *Sodalis* may modulate the ability of *Trypanosoma* to establish an infection in the tsetse midgut as some studies reported that the elimination of this bacterial endosymbiont results in an increased tsetse fly refractoriness to *Trypanosoma* infection[23,26,27]. Moreover, Geiger et al.,[28] suggested that specific genotypes of *Sodalis* presents in *G. p. gambiensis* from insectary colonies facilitate *Trypanosoma* infection. Soumana et al.,[29] revealed that a variation in the *Sodalis* population caused by a hosted prophage can influence the trypanosome infections. In contrast, a recent study demonstrated that the absence or presence of *S. glossinidius* in the tsetse fly does not affect the fly's susceptibility toward *Trypanosoma* infection[24]. In conclusion, from the above-described results, it is clear that our knowledge on the impact of *Sodalis* on *Trypanosoma* infection in tsetse remains limited and fragmented and is still under debate[23]. Moreover, exploring on a large scale the occurrence and possible association between *Sodalis* and *Trypanosoma* infection in wild flies is highly required. The above described potential impact of *Sodalis* to facilitate *Trypanosoma* infection in tsetse, and the fact that *Sodalis* is

Region	Country	*Sodalis* prevalence (%)*	Trypanosome prevalence (%)
	Ethiopia	94/459 (20.48)[a,b,e]	92/459 (20.04)[a,d,e]
East, central and southern Africa	Kenya	288/1008 (28.57)[a,b]	448/1008 (44.44)[a,b,e]
	Democratic R. of Congo	4/35 (11.43)[a,b,e]	1/35 (2.86)[a,e]
	Mozambique	7/100 (7.00)[a,b,e]	80/526 (15.21)[a,e]
	South Africa	9/526 (1.71)[a,c,e]	0/30 (0.00)[a,e]
	Eswatini	0/30 (0.00)[a,b,c,e]	8/100 (8.00)[a,e]
	Tanzania	227/338 (67.16)[a,d,c]	128/338 (37.87)[a,e]
	Uganda	91/210 (43.33)[d]	19/210 (9.05)[a,c,e]
	Zambia	11/210 (5.24)[a,b,e]	97/210 (46.19)[a,d,e]
	Zimbabwe	39/211(18.48)[a,b,e]	113/211 (53.55)[a,e]
	Subtotal	770/3127 (24.62)	986/3127 (31.53)
West Africa	Burkina Faso	11/2274 (0.48)[a,e]	498/2274 (21.90)[a,e]
	Ghana	0/234 (0.00)[a,e]	143/234 (61.11)[a,d]
	Guinea	90/314 (28.66)[a,e]	7/314 (2.22)[a,c]
	Mali	0/364 (0.00)[a,e]	25/364 (6.86)[a,c,e]
	Senegal	0/547 (0.00)[a,e]	78/547 (14.25)[a,e]
	Subtotal	101/3733 (2.70)	750/3733 (20.09)
	Total (average)	871/6860 (12.69)	1736/6860 (25.30)

Table 1. Global prevalence of *Sodalis* and Trypanosomes in tsetse samples analyzed per country. *Values indicated by the same lower-case letter do not differ significantly at the 5% level.

found in all laboratory-reared tsetse colonies and some wild populations[21] indicates that mitigating action, such as feeding the flies 2–3 times on blood supplemented with trypanocidal drugs before release, is required in SIT programs to minimize the risk of disease transmission by the large number of released males that harbour *Sodalis*.

Field studies in two HAT foci in Cameroon used PCR to detect *Trypanosoma* and *Sodalis* in *G. palpalis palpalis* and the results indicate that the presence of *Sodalis* favours *Trypanosoma* infections especially by *T. brucei* s.l.[30]. Furthermore, in the wildlife-livestock-human interface in the Maasai Mara National Reserve in Kenya, it was shown that *G. pallidipes* infected with *Sodalis* was associated with increased *Trypanosoma* infection rates[31]. However, other studies have found no strong association between trypanosome and *Sodalis* in some tsetse species collected in four locations in Kenya[32]. Channumsin et al.,[33] suggested that the association between *Trypanosoma* infection and the presence of *Sodalis* will vary depending on tsetse and *Trypanosoma* species. Similarly, studies carried out in the Fontem focus in Cameroon did not find a relationship between the endosymbiont and the parasite in *G. p. palpalis*[34], and no significant *Sodalis*-*Trypanosoma* infection association was found in *G. tachinoides* in two sites of the Faro and Déo Division in Adamawa region of Cameroon[35]. Likewise, no association between the presence of the parasite and *Sodalis* was found in *G. brevipalpis*, *G. m. morsitans* and *G. pallidipes* in the Luambe National Park of Zambia[36].

The overall objective of this study was to evaluate the prevalence of *Sodalis* and *Trypanosoma* in wild tsetse populations at a continental scale, i.e. Burkina Faso, Democratic Republic of Congo (DRC), Eswatini, Ethiopia, Ghana, Guinea, Kenya, Mali, Mozambique, Senegal, South Africa, Tanzania, Uganda, Zambia, and Zimbabwe and analyse these data in the context of a possible association between the occurrence of *Sodalis* and *Trypanosoma* infection in tsetse. Such information might guide the decision maker for SIT programmes to take the appropriate action, if necessary, to minimize any potential risk of increased transmission.

Results

***Trypanosoma* prevalence.** Adult tsetse flies (n = 6860) were screened for infection with *T. brucei* sspp (Tz) (*T. b. brucei, T. b. gambiense, T. b. rhodesiense)*, Tc (*T. congolense* savannah; *T. congolense* kilifi; *T. congolense* forest); Tsg (*T. simiae*; *T. simiae* tsavo; *T. godfreyi*) and Tv (*T. vivax*). The results indicate that 1736 (25.3%) adults were infected with one or more *Trypanosoma* taxa (Tables 1, 2 and 3), The *Trypanosoma* prevalence varied significantly between tsetse taxa ($X^2 = 750.18$, $df = 9$, $P << 0.001$) and between countries ($X^2 = 2038.1$, $df = 14$, $P << 0.001$). The Permanova analysis indicated as well significant differences between countries ($P = 0.009$) and taxa ($P = 0.041$) (Table 4). As all taxa were not collected from all countries, the interaction between taxa and countries was only analyzed where a taxon was collected from several countries.

Regardless of tsetse taxon, in west African countries the average *Trypanosoma* prevalence was 20% ($n = 3733$), with the highest prevalence recorded in Ghana (61%) and the lowest recorded in Guinea (2.2%). The prevalence in Burkina Faso, Mali and Senegal was 21.9, 6.9 and 14.2% respectively (Fig. 1, and Table 1). In east, central and southern African countries, the *Trypanosoma* infection prevalence was a bit higher than in west African countries with an averaged infection of 31.5% ($n = 3127$), with the highest prevalence (53.6%) in Zimbabwe and lowest prevalence (2.9%) in DRC. No *Trypanosoma* infection was detected in Eswatini (Fig. 1 and Table 1). Regardless of the country, *Trypanosoma* prevalence varied from one taxon to another, and *G. m. morsitans* showed the highest *Trypanosoma* prevalence (41%) followed by *G. pallidipes* (38.5%) and the lowest prevalence was detected

Species	*Sodalis* prevalence (%)*	Trypanosome prevalence (%)
G. austeni	5/346 (1.44)[a]	58/346 (16.76)[a]
G. brevipalpis	14/350 (4)[a]	34/350 (9.71)[a]
G. f. fuscipes	24/183 (13.11)[a,b]	31/183 (16.93)[a]
G. medicorum	8/154 (5.2)[a]	61/154 (39.6)[a,b]
G. m. morsitans	156/369 (42.27)[b]	152/369 (41.19)[a]
G. m. submorsitans	1/343 (0.29)[a]	62/343 (18.07)[a]
G. pallidipes	567/1844 (30.74)[b]	711/1844 (38.55)[a,b]
G. p. gambiensis	92/2168 (4.24)[a]	343/2168 (15.82)[a]
G. p. palpalis	4/35 (11.4)[a,b]	1/ 35 (2.8)[a,b]
G. tachinoides	0/1068 (0.0)[a]	283/1068 (26.49)[b]
Total (average)	871/6860 (12.6)	1736/6860 (25.3)

Table 2. Global prevalence of *Sodalis* and Trypanosomes in tsetse samples analyzed per tsetse species. *Values indicated by the same lower-case letter do not differ significantly at the 5% level.

Species	Country	*Sodalis* prevalence (%)*	Trypanosome prevalence (%)
G. austeni	Mozambique	0/50 (0.00)	5/50 (10.00)
	South Africa	2/226 (0.88)	49/226 (21.68)
	Eswatini	0/30 (0.00)	0/30 (0.00)
	Tanzania	3/40 (7.50)	4/40 (10.00)
G. brevipalpis	Mozambique	7/50 (14.00)[a]	3/50 (6.00)
	South Africa	7/300 (2.33)[b]	31/300 (10.33)
G. f. fuscipes	Kenya	20/89 (22.47)	21/89 (23.60)
	Uganda	4/94 (4.25)	10/ 94 (10.63)
G. medicorum	Burkina Faso	8/154 (5.20)	61/154 (39.61)
G. m. morsitans	Kenya	54/85 (63.52)[a]	2/ 85 (2.35)
	Tanzania	62/81 (76.54)[a]	43/81 (53.08)
	Zambia	8/64 (12.50)[b]	31/64 (48.43)
	Zimbabwe	32/139 (23.02)[b]	75/139 (53.95)
G. m. submorsitans	Burkina Faso	1/343 (0.30)	62/343 (18.07)
G. pallidipes	Ethiopia	94/459 (20.48)[a,b,c]	92/459 (20.04)
	Kenya	214/834 (25.65)[a,c]	425/834 (50.95)
	Tanzania	162/217 (74.65)[a,b]	81/217 (37.32)
	Uganda	87/116 (75.00)[a,b]	9/116 (7.75)
	Zimbabwe	7/72 (9.72)[a,c]	38/72 (52.77)
	Zambia	3/146 (2.05)[a,b,c]	66/146 (45.20)
G. p. palpalis	Democratic R. of Congo	4/35 (11.42)	1/35 (2.86)
G. p. gambiensis	Burkina Faso	2/943 (0.21)	235/943 (24.92)[a]
	Guinea	90/314 (28.66)	7/314 (2.22)[b]
	Mali	0/364 (0.00)	25/364 (6.87)[b,c]
	Senegal	0/547 (0.00)	78/547 (14.25)[c]
G. tachinoides	Burkina Faso	0/834 (0.00)	140/834 (16.79)[a]
	Ghana	0/234 (0.00)	143/234 (61.11)[b]
Total (average)		871/6860 (12.69)	1736/6860 (25.30)

Table 3. Global prevalence of *Sodalis* and trypanosomes in tsetse samples analyzed per country and tsetse species. *Values indicated by the same lower-case letter do not differ significantly at the 5% level.

in *G. brevipalpis* (9.71%) in east, central and southern Africa. In west Africa, *G. medicorum* showed the highest *Trypanosoma* prevalence (39.5%) and the lowest prevalence was detected in *G. p. palpalis* (2.8%) (Table 2).

Some tsetse taxa were collected from several countries as presented in Fig. 2 and Table 3. The highest *Trypanosoma* prevalence was recorded in *G. tachinoides* in Ghana (61%). This was followed by high prevalence in *G. m. morsitans* collected from Zimbabwe (53.9%), Tanzania (53%) and Zambia (48.4%). *G. pallidipes* from Zimbabwe, Kenya, Zambia and Tanzania also showed high *Trypanosoma* prevalence of 52.7%, 50.9%, 45.2% and 37.3%, respectively. The lowest *Trypanosoma* prevalence was found in *G. p. gambiensis* from Guinea (2.2%). Based on the *Trypanosoma* prevalence presented in Fig. 2 and Table 3, the tested samples can be categorized as: (i) tsetse samples with high prevalence (> 35%) detected in *G. tachinoides* from Ghana; *G. medicorum* from Burkina Faso,

Source	df	SS	MS	Pseudo-F	P (perm)	Unique perms
Countries	11	13,040	1185.4	2.6004	**0.009**	998
Species	7	7899.8	1128.5	2.4756	**0.041**	999
Residuals	5	2279.3	455.87			
Total	25	34,074				

Table 4. Permanova analysis for Countries and tsetse species for *Sodalis* and trypanosome (single and multiple) infection prevalence. Within the table, statistically significant differences ($P < 0.05$) can be seen in bold values in countries and tsetse species. Perm(s) = permutations.

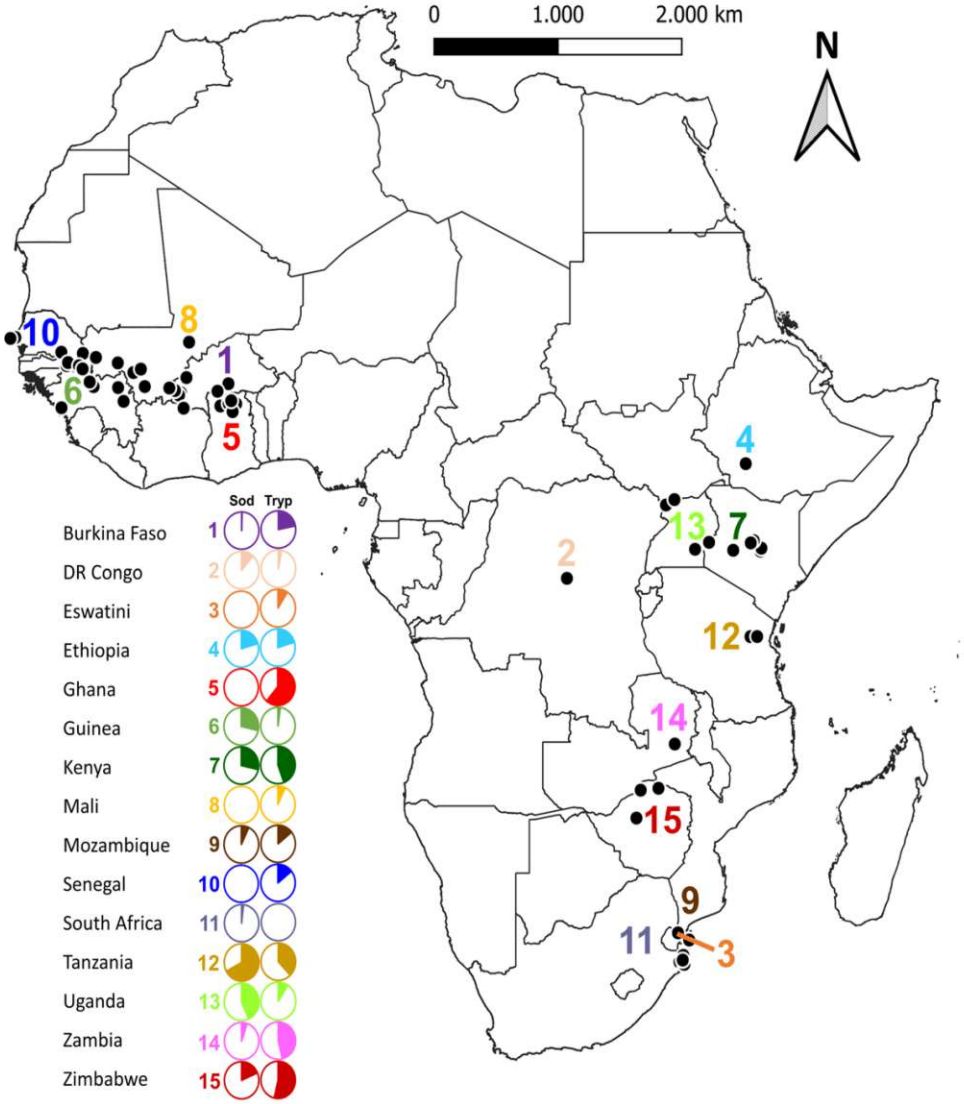

Figure 1. The geographical locations of tsetse samples in Africa. Circles indicate the total prevalence of *Sodalis* and *Trypanosoma* per country. Black dots indicate samples collection site(s) per country.

G. pallidipes from Kenya, Zambia, and Zimbabwe, *G. m. morsitans* from Tanzania, Zambia, and Zimbabwe; (ii) tsetse samples with medium prevalence (10–35%) detected in *G. austeni* from South Africa, *G. f. fuscipes* from Kenya and Uganda, *G. m. submorsitans* from Burkina Faso, *G. p. gambiensis* from Burkina Faso and Senegal and *G. tachinoides* from Burkina Faso; (iii) tsetse samples with low prevalence (< 10%) detected in the rest of the samples listed in Table 3 except the *G. austeni* collected from Eswatini. Despite the difference in *Trypanosoma*

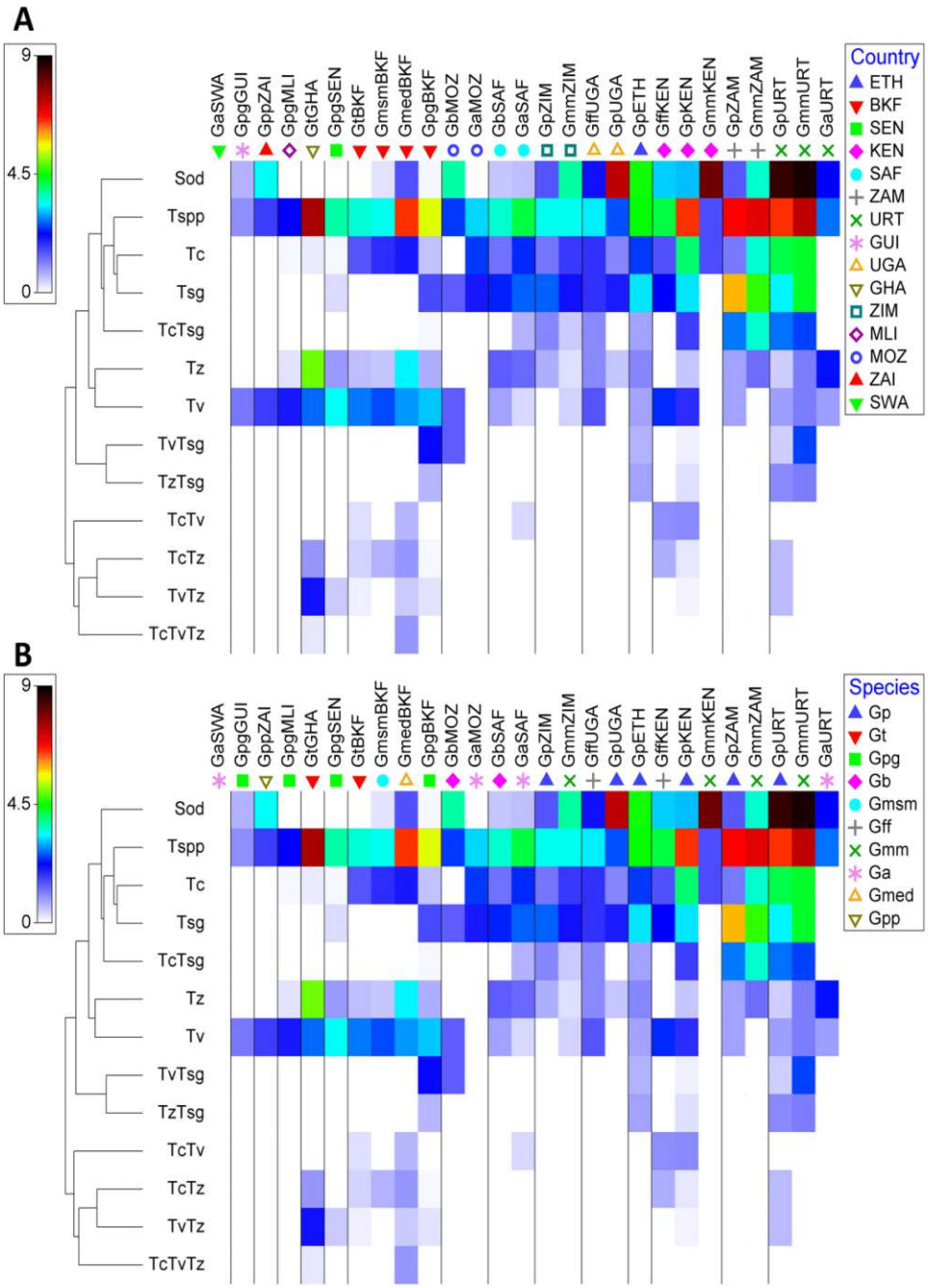

Figure 2. Prevalence of the *Sodalis* and *Trypanosoma* (single and multiple) infections per country (**A**) and tsetse species (**B**). Prevalence data were square root transformed and averaged based on country-species and the matrix display was conducted in PRIMER version 7 + software. Tree on the left of the matrix is the similarity dendrogram based on the similarity index of the square root of the prevalence values. The colour index is the square root of the prevalence values ranged 0–9 which is the square root of 0–81% prevalence. Country abbreviations follow the UNDP list of country codes https://web.archive.org/web/20060713221355/http://refgat.undp.org/genericlist.cfm?entid=82&pagenumber=1&requesttimeout=360 as follows: BKF: Burkina Faso; ETH: Ethiopia; GHA: Ghana; GUI: Guinea; KEN: Kenya; MLI: Mali; MOZ: Mozambique; SAF: South Africa; SWA: Eswatini; ZAI: Democratic Republic of the Congo; ZAM: Zambia; ZIM: Zimbabwe. Tsetse, *Sodalis* and *Trypanosoma* taxa were abbreviated as following: Ga: *Glossina austeni*; Gb: *G. brevipalpis*; Gff: *G. fuscipes fuscipes*, Gmm: *G. morsitans morsitans*; Gmsm: *G. m. submorsitans*; Gpg: *G. palpalis gambiensis*; Gpp: *G. palpalis palpalis*. Sod: *Sodalis*, Tc: *Trypanosoma. congolense* savannah; *T. congolense* kilifi; *T. congolense* forest, Tsg: *T. simiae*; *T. simiae* Tsavo; *T. godfreyi*, Tv: *T. vivax*, Tz: *T. brucei brucei, T. b. gambiense, T. b. rhodesiense.*

prevalence for each tsetse species, the differences were significant only in *G. p. gambiensis* ($X^2 = 26.71$, $df = 4$, $P < 0.001$) and *G. tachinoides*, ($X^2 = 9.38$, $df = 1, 2$, $P = 0.002$). In contrast, no significant difference was detected between countries for *G. austeni* ($X^2 = 1.47$, $df = 4$, $P = 0.688$), *G. brevipalpis* ($X^2 = 0.34$, $df = 2$, $P = 0.559$), *G. f. fuscipes* ($X^2 = 0.15$, $df = 2$, $P = 0.702$), *G. m. morsitans* ($X^2 = 1.04$, $df = 3$, $P = 0.593$) and *G. pallidipes* ($X^2 = 4.983$, df = 1,6, $P = 0.418$) (Table 3). No *Trypanosoma* infection was recorded in *G. austeni* from Eswatini. The best glm model (lowest AICc) selected for the overall *Trypanosoma* prevalence retained the countries as variables that fitted the data well (AICc = 1521.35) (Supplementary File 1).

Prevalence of different *Trypanosoma* taxa and mixed infections. The above-mentioned prevalence of *Trypanosoma* infection was comprised of several different *Trypanosoma* species and sub-species. Based on the size of the amplified fragment by PCR, the *Trypanosoma* infection was categorized into four groups: (i) the Tc group including the different forms of *T. congolense*; (ii) Tv group including *T. vivax* infections; (iii) *T. brucei* sspp (Tz) group including *T. b. brucei, T. b. gambiense, T. b. rhodesiense,* and infections; and (iv) Tsg group including the infections with *T. simiae, T. simiae* tsavo and *T. godfreyi*. The screening results revealed that tsetse flies could be infected with single or multiple (double or triple) taxa of *Trypanosoma*, and the proportion of the infections with the different *Trypanosoma* taxa and the mixed infection varied with country ($X^2 = 63.56$, $df = 14$, $P < 0.001$) and species ($X^2 = 21.86$, $df = 9$, $P < 0.001$) (Supplementary File 1).

The prevalence of the different *Trypanosoma* species with respect to the above-mentioned groups, indicate that infections with the Tsg group was the highest regardless of countries or tsetse species with an average of 7.06%. The infection rate was higher (14.13%) in east, central and southern African countries than in west Africa (1.13%). Tv infection averaged at 6.75% but with higher prevalence in west African countries (10.37%) than in east, central and southern Africa (2.43%). The prevalence of Tc infection was lower than Tv and Tsg group with an average of 4.78% with higher prevalence in central and southern Africa (8.38%) than in west Africa (1.77%). The Tz group had the lowest prevalence with an average of 2.29%. Like Tv infection, the Tz prevalence was higher in west Africa (3.16%) than central and southern Africa (1.25%).

The prevalence of infection by a single *Trypanosoma* group varied significantly from one country to another and from one tsetse species to another. For Tc, Tv, Tz and Tsg the infection prevalence varied significantly with country ($X^2 = 47.74$, $df = 14$, $P < 0.001$, $X^2 = 27.40$, $df = 14$, $P = 0.01705$, $X^2 = 106.11$, $df = 14$, $P = 0.001$ and, $X^2 = 44.74$, $df = 14$, $P = 0.001$ respectively). Regardless of tsetse species, the highest infection rate for Tc, Tv, Tz and Tsg was found in Tanzania (14.20%), Ghana (14.10%), Ghana (19.66%) and Zimbabwe (39.81%), respectively (Supplementary Table 1). Similarly, the prevalence of Tc, Tz and Tsg varied significantly with tsetse species ($X^2 = 40.364$, df = 1.9, $P << 0.001$, $X^2 = 58.253$, df = 1.9, $P << 0.001$ and $X^2 = 34.871$, df = 1.9, $P << 0.001$, respectively), however no significant difference was found in Tv prevalence between tsetse species ($X^2 = 5.475$, df = 1.9, $P = 0.07868$). Regardless of the country, the highest infection rate of Tc, Tv, Tz and Tsg was found in *G. pallidipes* (10.68%), *G. tachinoides* (12.92%), *G. medicorum* (13.64%) and *G. m. morsitans* (22.76%), respectively (Supplementary Table 2). No Tc infection was found in samples of *G. austeni* collected from Eswatini and Tanzania, *G. brevipalpis* from Mozambique, *G. p. palpalis* from DRC and *G. p. gambiensis* from Guinea. In addition, no Tv infection was detected in *G. austeni* collected from Eswatini and Mozambique, *G. m. morsitans* from Kenya and Zambia, *G. pallidipes* from Uganda and Zimbabwe. For Tz, *G. austeni* collected from Eswatini and Mozambique, *G. brevipalpis* from Mozambique, *G. f. fuscipes* from Kenya, *G. m. morsitans* from Kenya and Zambia, *G. p. palpalis* from DRC and *G. p. gambiensis* from Guinea did not show any infection (Fig. 2 and Supplementary Table 3).

Mixed infections of *Trypanosoma* groups (double or triple) are rare events with an average prevalence between 0.09 and 1.71% regardless of country or tsetse species. However, double infections seem to be more frequent in some countries than others ($X^2 = 35.01$, $df = 14$, $P = 0.001$) for Tv–Tz and in some tsetse species than others ($X^2 = 21.20$, $df = 9$, $P = 0.012$) for Tv–Tz (Supplementary File 1). The highest prevalence of the mixed infections Tv–Tz and Tc–Tz were observed in Ghana with 12.39% and 10.68%, respectively, regardless of tsetse species. Although the average Tc–Tsg prevalence was higher than that of Tv–Tz and Tc–Tz, the highest mixed infection with it was found in Zambia with 9.05%. Regardless of the country, the highest mixed infection of Tc–Tsg detected per tsetse species was ~ 5% in *G. m. morsitans* and *G. pallidipes*. The mixed infection of Tsg with either Tv or Tz or both was lower than 2% regardless of the country or tsetse species. Taking into account both the country and tsetse species, the highest mixed infection of Tc–Tsg (12.5%) was detected in *G. m. morsitans* in Zambia. However, the highest prevalence of Tc–Tz (10.68%) and Tv–Tz (12.39%) was detected in *G. tachinoides* from Ghana. Although the average prevalence of Tv–Tsg was low (0.54%), a relative high infection rate of 6.17% was found in *G. m. morsitans* from Tanzania.

A triple infection of *Trypanosoma* groups (Tc–Tv–Tz) was only detected in *G. medicorum* from Burkina Faso (1.30%) and *G. tachinoides* from Ghana (1.71%) (Fig. 2 and Supplementary Table 3, Supplementary File 1).

Prevalence of *Sodalis* infection. The prevalence of *Sodalis* infection based on the PCR results varied significantly with country ($X^2 = 108.02$, df = 1, 14, $P << 0.001$) and tsetse species ($X^2 = 69.60$, $df = 9$, $P < 0.001$). The best glm model (lowest AICc) selected for the overall *Sodalis* prevalence retained the countries, the species and their interaction (where possible) as variables that fitted the data well (AICc = 1296.12). Similar to the prevalence of *Trypanosoma*, the average *Sodalis* prevalence in east, central and southern Africa (24.6%) was higher than in west Africa (2.70%). Regardless of tsetse species, the highest prevalence of *Sodalis* infection was found in Tanzania (67.1%) followed by Uganda (43.3%), Kenya (28.5%) and Ethiopia (20.48%) (Table 1). The highest prevalence of *Sodalis* infection in west Africa was found in Guinea (28.6%). No *Sodalis* infection was found in Ghana, Mali, Senegal or Eswatini. Regardless of the country, the highest *Sodalis* prevalence per tsetse species was detected in *G. m. morsitans* (42.27%) followed by *G. pallidipes* (30.74%). No *Sodalis* infection was detected in *G. tachinoides.* The prevalence of *Sodalis* infection changed when both the countries and tsetse species are taken into

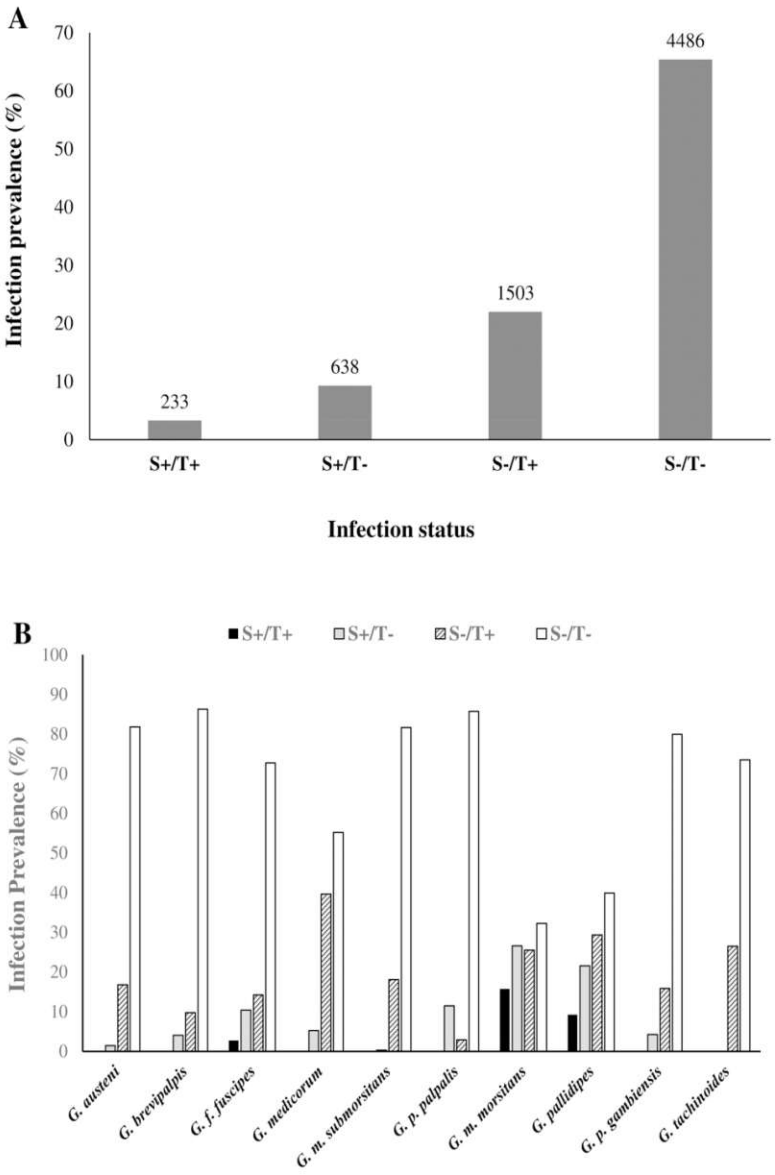

Figure 3. Prevalence of coinfection of *Sodalis* and *Trypanosoma* infection in wild tsetse populations. (**A**) Prevalence of coinfection, (**B**) prevalence of coinfection by tsetse taxa.

consideration (Table 4). Based on the *Sodalis* prevalence the tsetse samples can be categorized into four groups: (i) samples with high prevalence (> 50%) (ii) samples with medium prevalence (between < 10% and > 50%) (iii) samples with low prevalence (between > 0% and 10%) and (iv) samples with no *Sodalis* infection as shown in Fig. 2 and Table 4. The samples showing high *Sodalis* prevalence includes *G. m. morsitans* from Kenya (63.5%) and Tanzania (76.5%) and *G. pallidipes* from Tanzania (74.6%) and Uganda (75%), however the samples with no *Sodalis* infection includes *G. austeni* from Eswatini, *G. p. gambiensis* from Mali and Senegal and *G. tachinoides* from Burkina Faso and Ghana indicating that there is 95% confidence that the infection rate is less than or equal to 10%, 0.82%, 0.55%, 1.28% and 0.36%, respectively.

Interactions between *Sodalis* and *Trypanosoma* infections. *Prevalence of co-infections of Sodalis with Trypanosoma.* The screening results indicated that the single infection rate was 9.3% (n = 638) and 21.9% (n = 1503) for *Sodalis* and *Trypanosoma,* respectively, over all taxa and countries (Fig. 3A). No *Sodalis* infection was found in *G. tachinoides*, and therefore was excluded from the analysis. A Cochran–Mantel–Haenszel test for repeated tests of independence showed that infection with *Sodalis* and *Trypanosoma* did deviate from independence across all taxa ($\chi^2_{MH} = 41.73$, df = 1, $P < 0.001$) and individual Chi squared tests for independence for each

taxon showed significant deviation from independence at the Bonferroni corrected $\alpha = 0.00833$ in *G. pallidipes* ($P < 0.001$) and *G. p. gambiensis* ($P < 0.001$) (Supplementary Table 4). The prevalence of coinfection of *Sodalis* and *Trypanosoma* in wild tsetse populations varied with tsetse taxon and location. No coinfection was found in many taxa and many locations. The co-infection was found only in *G. f. fuscipes* (2.73%), *G. m. morsitans* (15.72%) and *G. pallidipes* (9.22%) in east, central and southern Africa (Fig. 3B, Table 5 and Supplementary Table 4).

Impact of co-infection on Trypanosoma and Sodalis density. Attempts were to assess the density of *Trypanosoma* and *Sodalis* under single (S^-/T^+) and (S^+/T^-) or double infection (S^+/T^+) conducted using qPCR with primers mentioned in Supplementary Table 5. The results show that *Sodalis* infections did not have a significant impact on *Trypanosoma* density ($X^2 = 0.648$, $df = 2$, $P = 0.723$), however the median value of (S^+/T^+) files were slightly (S^-/T^+) lower than (S^+/T^-) and (S^-/T^+) flies and the number of outlier samples with higher trypanosome density (S^-/T^+) flies (Fig. 4A). *Trypanosoma* infections significantly reduced the density of *Sodalis* as indicated by comparing (S^+/T^+) flies with (S^+/T^-) flies ($P = 0.014$) although the median values in (S^+/T^+) files is higher than the other samples indicating that the increased of *Sodalis* density in (S^+/T^-) might be affected with the outlier flies with high *Sodalis* density (Fig. 4B). No significant different was found in the *Trypanosoma* density determined by qPCR in the flies tested negative (S^+/T^-) or positive (S^+/T^+) and (S^-/T^+) with the standard PCR, however, *Sodalis* density showed significant difference between flies with different infection type ($X^2 = 14.54$, df = 2, $P < 0.001$) (Fig. 4B). The results showed no correlation between *Sodalis* and *Trypanosoma* density (r = 0.007, t = 0.055, $df = 69$, $P = 0.9561$) Supplementary Fig. 2, Supplementary File 1).

Discussion

The implementation of the SIT in the context of an AW-IPM strategy to eradicate tsetse flies relies on the release of sterile males in the targeted area. This was successful in eradicating a population of *G. austeni* from Unguja Island of Zanzibar[37] and significant progress has been made in the eradication programme implemented against *G. p. gambiensis* in the Niayes area of Senegal[38]. However, as both male and female tsetse flies are vectors of *Trypanosoma* species, the release of large numbers of sterile male flies bears a potential risk of temporarily increasing disease transmission during the initial release phase of an SIT programme[39]. Therefore, mitigating measures are required to reduce or eliminate this potential risk, especially in areas where sleeping sickness (HAT) is endemic. To date, to mitigate such risks, sterile males are offered two or three blood meals mixed with the trypanocidal drug isometamidium chloride, before being released which reduces the risk of *Trypanosoma* transmission significantly but does not eliminate it[40,41]. In addition, other approaches were proposed to minimize such risks such as paratransgenesis[42,43] and combining paratransgenesis with SIT[44].

The vector competence of tsetse flies for different trypanosome species is highly variable and is suggested to be affected by various factors, amongst which bacterial endosymbionts. Here, the interaction of *Sodalis glossinidius* with tsetse trypanosome infection is still under debate. Several studies reported a potential positive correlation between *Sodalis* and *Trypanosoma* infections[28,30,32,36,45–48], leading to the hypothesis that *Sodalis* might facilitate the establishment of *Trypanosoma* infections in the tsetse midgut[23,26,27]. However, other studies indicated the lack of correlation between *Sodalis* and *Trypanosoma* infection[34–36]. The presence of *Sodalis* infections in tsetse rearing colonies has been well studied and previous studies indicated that *Sodalis* is more frequently present in colonized tsetse flies than in wild tsetse populations[36,49] with a prevalence of 80 and 100% in colonized *G. m. morsitans* and *G. p. gambiensis,* respectively[49,50], which is higher than the symbiont prevalence in wild populations of these tsetse species. This seems to indicate that the rearing process of tsetse flies favours the transmission and spread of *Sodalis* infections within the colonized population. Recently, colonies of *G. pallidipes, G. p. gambiensis, G. f. fuscipes, G. m. morsitans, G. m. centralis* and *G. m. submorsitans* maintained at the FAO/IAEA Insect Pest Control Laboratory were screened for *Sodalis* infections and showed a 100% prevalence of *Sodalis;* only the *G. brevipalpis* colony had a lower prevalence of 95% (data not shown). Taken into consideration that mass-rearing conditions enhances *Sodalis* infections and that *Sodalis* infections might facilitate the establishment of a *Trypanosoma* infection in the midgut, sterile male tsetse flies that are derived from colonies might be effective vectors for different *Trypanosoma* species and, therefore, might increase the trypanosome transmission after flies being released. It is therefore important that the managers and planners of SIT programmes are aware which tsetse species show a positive correlation between *Sodalis* and *Trypanosoma* infections to be able to take the necessary mitigating actions.

Various studies have examined the prevalence of *Sodalis* and *Trypanosoma* species in wild tsetse populations[30,32,35,45,51], but our study presents for the first time the prevalence of *Sodalis* and *Trypanosoma* species on a continent-wide scale. In addition, the DNA extraction and PCR methods we have used were standardized and were all carried out in one laboratory to avoid discrepancies in the results due to different handling of tsetse samples or to different methods to discriminate trypanosome species in tsetse tissues. Our results indicate that *Sodalis* and *Trypanosoma* prevalence varied with tsetse species and geographical location (with an overall trypanosome prevalence of 23.5%), which agrees with many previous studies[52]. A high *Trypanosoma* prevalence (> 30%) was found in *G. m. morsitans* and *G. pallidipes* from central and east Africa. This finding is in agreement with previous reports on *G. m. morsitans* and *G. m. centralis* from Zambia[36,52] and *G. m. morsitans* sampled in Malawi[53]. Moreover, a high prevalence of *Trypanosoma* infection in *G. pallidipes* was also previously reported in Tanzania[54] and Kenya[33]. However, another study in northern Tanzania indicated a lower prevalence of *Trypanosoma* infection (< 10%) both in *G. m. morsitans* and *G. pallidipes*[55].

Our study showed that the prevalence of different *Trypanosoma* species and or subspecies can be different in different tsetse taxa. In *G. tachinoides* in Ghana, the *Trypanosoma vivax* (Tv) infection was high (> 10%) as well as the infections of the *T. brucei* sspp (Tz) and the *T. simiae/T. godfreyi* (Tsg) group and the mixed infections of Tv–Tsg. In contrast, the prevalence of *T. congolense* was very low. These results are in agreement with the

Glossina taxon	Country (Area, Collection Date)	N	S^+/T^+	S^+/T^-	S^-/T^+	S^-/T^-	χ^2	*P*
G. austeni	Tanzania (Jozani, 1997)	4	0	0	1	3		
G. austeni	Tanzania (Zanzibar, 1995)	6	0	1	0	5		
G. austeni	Tanzania (Uguja Island, 1995)	30	0	2	3	25		
G. austeni	South Africa (North eastern Kwazulu Natal, 1999)	39	0	2	2	35		
G. austeni	South Africa (Lower Mkhuze, 2018)	53	0	0	23	30		
G. austeni	South Africa (Saint Lucia, 2018)	57	0	0	22	35		
G. austeni	South Africa (False Bay Park, 2018)	77	0	0	2	75		
G. austeni	Mozambique (Reserva Especial de Maputo, 2019)	50	0	0	5	45		
G. austeni	Eswatini (Mlawula Nature Reserve, 2019)	30	0	0	0	30		
G. austeni	All locations	346	0	5	58	283	1.02	0.31
G. brevipalpis	South Africa (North eastern Kwazulu Natal, 1995)	50	0	0	2	48		
G. brevipalpis	South Africa (Phinda, 2018)	170	0	7	0	163		
G. brevipalpis	South Africa (Saint Lucia, 2018)	30	0	0	13	17		
G. brevipalpis	South Africa (Hluhluwe, 2018)	50	0	0	16	34		
G. brevipalpis	Mozambique (Reserva Especial de Maputo, 2019)	50	0	7	3	40		
G. brevipalpis	All locations	350	0	14	34	302	1.57	0.21
G. f. fuscipes	Uganda (Buvuma island, 1994)	94	0	4	10	80		
G. f. fuscipes	Kenya (Ikapolock, 2007) [1]	51	5	15	14	17		
G. f. fuscipes	Kenya (Obekai, 2007)	38	0	0	2	36		
G. f. fuscipes	All locations	183	5	19	26	133	0.3	0.59
G. medicorum	Burkina Faso (Comoe, 2008)	94	0	8	32	54		
G. medicorum	Burkina Faso (Folonzo, 2008)	60	0	0	29	31		
G. medicorum	All locations	154	0	8	61	85	5.53	0.02
G. m. submorsitans	Burkina Faso (Comoe, 2007)	206	0	0	20	186		
G. m. submorsitans	Burkina Faso (Folonzo, 2008)	134	0	1	42	91		
G. m. submorsitans	Burkina Faso (Sissili, 2008)	3	0	0	0	3		
G. m. submorsitans	All locations	343	0	1	62	280	0.22	0.64
G. p. palpalis	Democratic Republic of Congo (Zaire, 1995)	35	0	4	1	30		
G. m. morsitans	Tanzania (Kwekivu 2, 2005)	81	35	27	9	10		
G. m. morsitans	Zambia (Mfuwe, Eastern Zambia, 2007)	64	1	7	30	26		
G. m. morsitans	Zimbabwe (Mukondore, 2007)	13	1	2	0	10		
G. m. morsitans	Zimbabwe (M. chiuyi, 2007)	9	0	1	0	8		
G. m. morsitans	Zimbabwe (Rukomeshi, 2006)	15	0	3	0	12		
G. m. morsitans	Zimbabwe (Kemukura, NA)	18	0	4	1	13		
G. m. morsitans	Zimbabwe (Mushumbi, 2006)	6	0	0	2	4		
G. m. morsitans	Zimbabwe (Makuti, 2006)	78	19	2	52	5		
G. m. morsitans	Kenya (Kari, 2006)	85	2	52	0	31		
G. m. morsitans	All locations	369	58	98	94	119	1.8	0.18
Glossina taxon	**Country (Area, Collection Date)**	**N**	**S^+/T^+**	**S^+/T^-**	**S^-/T^+**	**S^-/T^-**	**χ^2**	***P***
G. pallidipes	Zambia (Mfuwe, Eastern Zambia, 2007)	146	2	1	64	79		
G. pallidipes	Kenya (Mwea, Katotoi, Emsos, Kari, Kiria, Koibos,Meru and Ruma national park, 2007)	834	88	126	337	283		
G. pallidipes	Ethiopia (Arba Minch, 2007)	459	15	79	77	288		
G. pallidipes	Tanzania (Kwekivu 1, 2005)	217	54	108	27	28		
G. pallidipes	Zimbabwe (Mushumbi 2006)	26	1	0	4	21		
G. pallidipes	Zimbabwe (Gokwe, 2006)	4	0	0	0	4		
G. pallidipes	Zimbabwe (Rukomeshi, 2006)	4	0	0	0	4		
G. pallidipes	Zimbabwe (Makuti, 2006)	38	6	0	27	5		
G. pallidipes	Uganda (Lira,Omogo, Budaka, Moyo, NA)	116	4	83	5	24		
G. pallidipes	All locations	1844	170	397	541	736	25.4	0
G. p. gambiensis	Burkina Faso (Lorepeni)	10	0	0	8	2		
G. p. gambiensis	Burkina Faso (Bouroum bouroum)	18	0	0	16	2		
G. p. gambiensis	Burkina Faso (Kourignon)	24	0	0	10	14		
G. p. gambiensis	Burkina Faso (Kampty)	98	0	0	85	13		
G. p. gambiensis	Burkina Faso (Ouarkoye)	5	0	0	5	0		
G. p. gambiensis	Burkina Faso (Dedougou)	57	0	0	33	24		
G. p. gambiensis	Burkina Faso (Bama)	77	0	0	0	77		
Continued								

Glossina taxon	Country (Area, Collection Date)	N	S^+/T^+	S^+/T^-	S^-/T^+	S^-/T^-	χ^2	P
G. p. gambiensis	Burkina Faso (Comoe)	123	0	0	3	120		
G. p. gambiensis	Burkina Faso (Folonzo)	212	0	2	25	185		
G. p. gambiensis	Burkina Faso (Kartasso)	136	0	0	0	136		
G. p. gambiensis	Burkina Faso (Kenedougou)	41	0	0	0	41		
G. p. gambiensis	Burkina Faso (Moussodougou)	142	0	0	49	93		
G. p. gambiensis	Guinea (Bafing)	33	0	0	1	32		
G. p. gambiensis	Guinea (Dekonkore)	16	0	0	1	15		
G. p. gambiensis	Guinea (Kangoliya	126	0	90	0	36		
G. p. gambiensis	Guinea (Kerfala	13	0	0	1	12		
G. p. gambiensis	Guinea (Kifala)	30	0	0	0	30		
G. p. gambiensis	Guinea (Lemonako)	20	0	0	0	20		
G. p. gambiensis	Guinea (Mimi)	45	0	0	1	44		
G. p. gambiensis	Guinea (Tinkisso)	31	0	0	2	29		
G. p. gambiensis	Mali	364	0	0	25	339		
G. p. gambiensis	Senegal	547	0	0	79	469		
G. p. gambiensis	All locations	2168	0	92	343	1733	18.06	0
G. tachinoides	Burkina Faso	834	0	0	140	694		
G. tachinoides	Ghana	234	0	0	143	91		
G. tachinoides	All locations	1068	0	0	283	785		

Table 5. Distribution of the association between the presence of *Trypanosoma* spp and the presence of *Sodalis* according to the tsetse species and the country.

prevalence of *T. brucei* s.l (11%) and *T. congolense* forest type (2.6%) reported in the same tsetse species in Cameroon. The same study reported a prevalence of 13.7% of *T. congolense* savannah type[35], which was not observed in our study. Our results of trypanosome infection rates in *G. tachinoides* also agree with former studies[56,57], except for T,c for which a high fly infection rate (31.8%) was previously shown[57]. The Tc infection rates in our study were high in *G. pallidipes* and *G. m. morsitans;* for the latter tsetse fly species, a study in Malawi reported a high prevalence for *T. brucei* (64.4%) but much lower for all other *Trypanosoma* infections(< 10%)[58]. The mixed infection of *Trypanosoma* species/subspecies is in agreement with previous reports[35,52,57,59].

Likewise, the prevalence of *Sodalis* infection varied significantly with tsetse taxon and location and the highest prevalence was found in *G. m. morsitans* and *G. pallidipes*. Our results agree with the high prevalence of *Sodalis* reported in *G. pallidipes* (~ 50%) in one location in Kenya regardless of the fly age[33]; however, the same study reported low *Sodalis* prevalence in another location. In another study in Kenya, Wamwiri et al.,[32] reported moderate *Sodalis* prevalence in *G. pallidipes* (16%) and low prevalence in *G. austeni* (3.7%), which is in agreement with our results. On the other hand, our results are different from the low prevalence (< 8%) found in *G. m. morsitans* and *G. pallidipes* in Zambia[36]. In another study in Zambia, *Sodalis* prevalence in *G. m. centralis*, was reported to be 15.9% with no significant difference between inter-site prevalence[52]. The prevalence of *Sodalis* in *G. brevipalpis* in our study was found to be low (< 2.3%) which contradicts with the high prevalence (93.7%) reported in this species in Zambia[36]. In the DRC, the global prevalence of *Sodalis* in *G. fuscipes quanzensis* midgut averaged 15.5%, but in certain locations the prevalence exceeded 40%[60]. In Nigeria, *Sodalis* prevalence in *G. p. palpalis* and *G. tachinoides* was 35.7%[61] which is higher that the prevalence reported in our study for both species.

Our data indicate that the *Trypanosoma* and *Sodalis* infections were very low or absent in some tsetse taxa from certain locations such as *G. austeni* in Eswatini for *Trypanosoma* and *Sodalis* infections and several species in west Africa for *Sodalis*. The lack of *Sodalis* and/or *Trypanosoma* infection in these samples might be due to (i) low number of tested samples (ii) the use of the DNA extracted from the whole body of tsetse adults (iii) the possibility of the collected samples being infected with different strains/genotypes that might not be detected with the primers used and (iv) the infection of *Sodalis* and *Trypanosoma* are under the detection limit of the used PCR. It is important to note that due to the high number of samples tested in our study, the more sensitive nested PCR to detect low infection level was excluded due to logistic reasons.

Our results indicate significant deviation from independence (correlation) of *Sodalis* and *Trypanosoma* infections in *G. medicorum, G. p. gambiensis* and *G. pallidipes*. However, the lack of detection of any tsetse adult with co-infection of *Sodalis* and *Trypanosoma* in *G. medicorum,* and *G. p. gambiensis* might indicate a negative correlation. Such negative trend might be supported by the lower density of *Sodalis* in the flies with co-infection (S^+/T^+) compared to these with *Sodalis* infection only (S^+/T^-). On other hand the lack of impact of *Sodalis* infection on *Trypanosoma* density does not support the negative trend and agreed with the results of Trappeniers et al.,[24] reported on colonized flies. This results also agreed with previous results reporting the absence of direct correlation between the presence of *Sodalis* and the acquisition of a *Trypanosoma* infection[63]. However, an inverse correlation was reported between *Sodalis* and the vector competence where the presence of *Sodalis* in both midgut and proboscis of *G. p. gambiensis* was associated with its status as a poor vector, whereas it is not found in the proboscis of *G. m. morsitans* (major vector). It is worth noting that all previous studies of *Sodalis* infection in *G. p. gambiensis* and its interaction with *Trypanosoma* infection was carried out with flies reared under laboratory conditions[28,29,64]. The correlation between *Sodalis* and *Trypanosoma* infection in *G. pallidipes* is

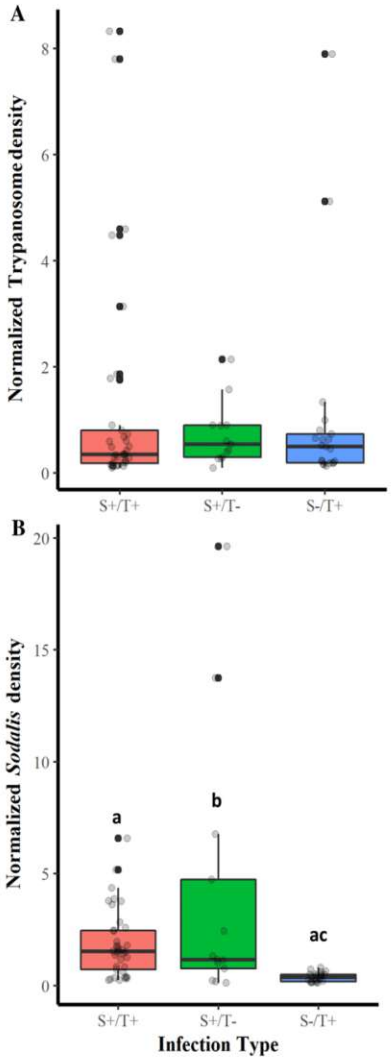

Figure 4. Impact of co-infection with *Trypanosoma* and *Sodalis* on *Trypanosoma* (**A**) and *Sodalis* (**B**) density in *Glossina pallidipes* and *G. m. morsitans*. Bars marked with the same lower-case letter do not differ significantly at the 0.05 level.

positive, evidenced with the relative high number (n = 170) of tsetse with co-infection. This positive correlation was also found in *G. pallidipes* from Kenya although with too few flies with co-infection to enable us to draw a definite conclusion[32]. Although co-infections were found in *G. m. morsitans* and *G. f. fuscipes* in some locations, the global correlation was missing. This is in agreement with the positive correlation found between *Sodalis* and *Trypanosoma* infection in *G. m. centralis* in Zambia, in which there was a 6.2 fold increase in the likelihood of a fly being infected with *Trypanosoma* if *Sodalis* was present[52]. More studies are needed to enhance the potential control interventions mediated by endosymbionts to reduce parasitic infections[61].

The results of this study clearly indicate that the interaction between *Sodalis* and *Trypanosoma* infection is complex, species-specific and requires further investigation. The prevalence results indicate that *Sodalis* and *Trypanosoma* infections are not independent in some species, such as *G. p. gambiensis* and *G. medicorum* in west Africa and *G. pallidipes* in central and east Africa, In case of a positive correlation between *Sodalis* and *Trypanosoma* infection in these species, additional measures could be suggested when implementing the SIT to reduce the *Sodalis* density in the sterile males released in the targeted area to maximize the safe implementation of the SIT. These measures might include the mixing of *Sodalis* phage(s)[29,65] with the blood meals to feed the mass-reared flies to reduce the *Sodalis* density in these flies. In addition, the blood meal offered to the males before release can be supplemented with one or more of the following antimicrobial products to reduce *Sodalis* density, i.e. streptozotocin[23], indolicidin and OaBAC 5 mini[66]. The use of the *Sodalis* phage as well as these antimicrobial agents requires further studies to (1) develop methods to isolate the phage, (2) determine the conditions (e.g. suitable concentration) for its use, and (3) determine the impact on *Sodalis* density, tsetse productivity and

survival. For *G. m morsitans* and *G. pallidipes*, our results suggest that *Sodalis* infection does not have an impact on *Trypanosoma* infection so here no additional measures need to be taken during the implementation of SIT against these species.

Conclusion

Sodalis and *Trypanosoma* infection varied with tsetse taxon and location. There is a significant positive correlation between *Sodalis* and *Trypanosoma* infection in *G. medicorum*, *G. p. gambiensis* and *G. pallidipes*; however, no significant correlation was found in other tsetse taxa and locations. The results of this study will enable the decision makers of SIT projects to better plan and take the necessary measures to fine-tune and optimize SIT efficiency and safety.

Methods

Tsetse collection and DNA extractions. Tsetse flies were collected in 1995 and between 2005 and 2018 from 95 different geographical locations in fifteen countries in east, central, southern, and western Africa (Table 6, Supplementary Table 6). The tsetse flies were collected with species-specific traps which included the biconical trap[67], the monoconical trap[68], the Vavoua trap[69], the Ngu trap[70,71], the odour-baited Epsilon trap[72], the NZI trap[73], and the odour baited H trap[74]. A total of 6860 tsetse flies, belonging to ten tsetse species, were collected for this study (Table 6). The majority of the samples were collected in Burkina Faso (2274), Kenya (1008), Senegal (547) and South Africa (526). As the distribution of most tsetse species is allopatric (only few species are sympatric), not all tsetse species were collected from each country. Following collection, fly samples were preserved in 95% ethanol or propylene glycol and shipped to the FAO/IAEA Insect Pest Control Laboratory (IPCL) in Seibersdorf, Austria and stored at − 20 °C until analysis. Total DNA was extracted from individual whole fly bodies using the DNeasy tissue kit (QIAGEN Inc., Valencia, CA) following the supplier's instructions. The DNA quality and concentration were measured by spectrophotometry (Synergy H1 Multi-Mode Reader, BioTek, Instruments, Inc., USA) and subsequently kept at 4 °C until screened for *Sodalis* and *Trypanosoma* infections. To verify the quality of the extracted DNA, a set of specific primers amplifying the *Glossina* spp. microsatellite GpCAG133 sequence (Supplementary Table 5) and only the successful samples were included in the analysis[21,75].

***Trypanosoma* prevalence and genotyping.** Polymerase chain reaction (PCR), following the method of Njiru et al.[76] that used the primers ITS1-CF and ITS1-BR (Supplementary Table 5) previously designed to amplify the internal transcribed spacer (ITS1) of the ribosomal DNA, was used to detect *Trypanosoma* infection and *Trypanosoma* species in the fly samples. The PCR was carried out in 25 µl reaction mixtures containing 22.5 µl of 1.1 × Pre-Aliquoted PCR Master Mix (0.625 units Thermoprime Plus DNA Polymerase, 75 mM Tris–HCl (pH 8.8 at 25 °C), 20 mM $(NH_4)2SO_4$, 2.0 mM $MgCl_2$, 0.01% (v/v) Tween-20 and 0.2 mM each of the dNTPs (ABgene, UK), 1 µl primers (at 200 nM final concentration of forward and reverse primer) and 1.5 µl of template DNA. PCR cycles were: 94 °C for 15 min, followed by 40 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s, and final extension 72 °C for 5 min. Interpretation of the results after resolving the amplification products in a 2% agarose gel (Fisher Biotech) stained with SafeGreen or ethidium bromide, was based on the characteristic band size of *Trypanosoma* taxa: all members of the subgenus *T. brucei* sspp (*T. b. brucei*, *T. b. gambiense*, *T. b. rhodesiense*: 480 bp); *T. congolense* savannah (700 bp); *T. congolense* Kilifi (620 bp); *T. congolense* forest (710 bp); *T. simiae* (400 bp); *T. simiae* Tsavo (370 bp); *T. godfreyi* (300 bp) and *T. vivax* (250 bp). The positive control DNA was from *T. congolense* savannah, *T. congolense* forest, *T. b. brucei*, *T. b. gambiense*, *T. b. rhodesiense*, *T. evansi*, and *T. vivax*. DNA samples validated with GpCAG133 primer amplification were screened for trypanosome infection. A tsetse sample was recorded as positive if one or more of the indicated band sizes was detected. *Trypanosoma* infection status and species were recorded for each fly.

Prevalence of *Sodalis* infection. The detection of *Sodalis* in natural tsetse samples was based on the *Sodalis fliC* (flagellin) gene which results in an amplicon length of about 508 base pairs with the *Sodalis* specific primers Sod-fliC-F and Sod-fliC-R (Supplementary Table 5)[77]. These primers were used in single pairs or in multiplex PCR with GpCAG133 primers. For all PCR reactions, 22.5 µl of 1.1 × Pre-Aliquoted PCR Master Mix (ABgene, UK) was used. In a final volume of 25 µl, 1.5 µl of template DNA plus forward and reverse primers were added to a final concentration of 0.2 mM per primer in a volume of 1 µl. Samples were considered *Sodalis*-infected if the expected symbiont PCR product amplicon was detected. Data were accepted only if the control gene GpCAG133 sequence was amplified. The PCR cycling conditions were: 95 °C for 5 min followed by 34 cycles of 95 °C for 30 s, 52.5 °C for 30 s, 72 °C for 30 s and lastly at 72 °C for 10 min; PCR products were separated by agarose (2%) gel electrophoresis and SafeGreen or ethidium bromide staining.

Analysis of the *Trypanosoma* and *Sodalis* infection in wild tsetse populations. *Co-infection of tsetse adults with Sodalis and Trypanosoma infection.* The co-infection of *Sodalis* and *Trypanosoma* infection was evaluated based on the PCR prevalence. The infection status was divided into four categories *Sodalis* positive and *Trypanosoma* positive (S^+/T^+), *Sodalis* positive and *Trypanosoma* negative (S^+/T^-), *Sodalis* negative and *Trypanosoma* positive (S^-/T^+) and *Sodalis* negative and *Trypanosoma* negative (S^-/T^-).

Analysis of the Trypanosoma and Sodalis density. Samples showing *Trypanosoma* infection (not *T. vivax*) with *Sodalis* (S^+/T^+) and samples not infected with *Trypanosoma* but infected with *Sodalis* (S^+/T^-) were evaluated with quantitative PCR (qPCR) to assess the impact of *Trypanosoma* infection (regardless the *Trypanosoma* type) on *Sodalis* density. The qPCR was performed using a CFX96 Real Time PCR Detection System (Bio-Rad). The *fliC*

Country	No. of locations	No. of collection flies with valid DNA	Collection year
Ethiopia	1	459	2007
Kenya	11	1008	2007, 2008, 2009
Uganda	5	210	2007
Tanzania	5	338	2005, 2009
Democratic R. of Congo	1	35	1995
Zambia	1	210	2007
Zimbabwe	7	211	2006
South Africa	7	526	1995, 2018, 2019
Mozambique	1	100	2019
Eswatini	1	30	2018, 2019
Burkina Faso	14	2274	2008, 2010, 2013, 2015, 2018, 2019
Ghana[a]	11	234	2008
Guinea[a]	8	314	2008, 2009
Mali[a]	10	364	2008, 2010, 2011, 2012, 2013
Senegal	12	547	2008, 2009
Total	95	6860	

Table 6. List of collections of tsetse adults with valid DNA screened for *Sodalis* and Trypanosome[a] infection in wild tsetse population in east, central, southern and west Africa. [a]Part of the trypanosome infection in west Africa was screened by Ouedraogo et al. 2018.

gene was amplified with the following primers: sodqPCR-FliCF and sodqPCR-FliCR[78] (Supplementary Table 5) to assess the density of the symbiont present within *Trypanosoma* infected and noninfected, additional criteria for the selection of the samples was the presence of the two groups (S^+/T^+) and (S^+/T^-) in a given population. Based on the preceding criteria 96 individual flies (52 and 44 flies with infection status of (S^+/T^+) and (S^+/T^-), respectively, were selected from the *G. pallidipes* and *G. m. morsitans* collected in Kenya, Tanzania and Zimbabwe. In addition, samples with (S^+/T^+) and (S^-/T^+) were used to assess the impact of *Sodalis* infection on *Trypanosoma* density. *Trypanosomatidae*18S specific primers (18S_Typ_F and18S_Typ_R) (Supplementary Table 5) were used to assess the *Trypanosoma* density in the tested samples. The DNA from all selected samples was diluted to a final concentration of 4 ng/µl and 5 µl of the diluted DNA was used for qPCR to determine *Sodalis* and *Trypanosoma* DNA density normalized to the housekeeping β-tubulin gene. The amplification mixture contained 5 µl of DNA template, 200 nM of each primer, and 7.5 µl iQ™ SYBER Green Supermix (Bio-Rad). qPCR cycling conditions for *Sodalis* were as follows: initial denaturation at 95 °C for 2 min; 39 cycles of 95 °C for 5 s, 55 °C for 30 s, one step at 95 °C for 5 s and a melting curve constructed from 65 °C to 95 °C in increments of 0.5 °C for 5 s. The same conditions were used for *Trypanosoma* except the annealing temperature was at 60 °C. The analysis of the *Sodalis, Trypanosoma* and Tubulin densities was based only on qPCR data with the expected melting curve at 81.5 °C, 85.5 °C and 86 °C, respectively.

Data analysis. The prevalence data were recorded and analyzed with the general linear model (GLM)[79]. The prevalence of *Sodalis*, *Trypanosoma* species and each *Trypanosoma* species and co-infection were tested for differences between the tsetse taxa and between countries. For each country, the prevalence was assessed again for differences between the localities where the flies were collected and between the tsetse species present in each country. In the absence of PCR detected *Sodalis* or *Trypanosoma* infection, the upper 95% confidence interval for the true rate of infection was calculated following the method of Couey and Chew[62]. *Trypanosoma* prevalence between taxa was compared between species by a pairwise comparison of proportions with a Bonferroni correction and Benjamini–Hochberg correction. The analyses were executed in R v 4.0.5[79] using RStudio V 1.4.1106[80,81] with the packages ggplot2 v3.3.2.1[82], lattice v0.20-41[83], car[84], ggthems[85] and MASS v7.3-51.6[86] except for the Chi squared tests for independence, Spearman correlation coefficient and Cochran–Mantel–Haenszel test for repeated tests of independence, which were performed using Excel 2013 The R Markdown file is available in Supplementary File 1.

To analyse the qPCR data, normalized density of *Trypanosoma* and *Sodalis* against the house keeping gene (tubulin) was extracted from the CFX Maestro software. Samples giving a valid density (not N/A) for both *Trypanosoma* and *Sodalis* were retained for further statistical analysis in R. Similarities in the structure of *Sodalis* and *Trypanosoma* (single and multiple) infection and the role of different factors such as countries and tsetse taxa, were assessed using the matrix display and metric multidimensional scaling (mMDS) plot with bootstrap averages in PRIMER version 7+. The bootstrap averages plots were displayed with a Bray and Curtis matrix based on the square-root transformation of the *Sodalis* and *Trypanosoma* (single and multiple) infection abundance data[87]. The tests were based on the multivariate null hypothesis via the use of the non-parametric statistical method PERMANOVA[88]. The Permanova test was conducted on the average of the abundance data based on the country-species after excluding the data of Eswatini (low number of tested samples).

Data availability

Materials described in the paper, including all relevant raw data, are available in this link https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/WOTAIY).

Received: 2 December 2021; Accepted: 3 February 2022
Published online: 28 February 2022

References

1. Elsen, P., Amoudi, M. A. & Leclercq, M. First record of *Glossina fuscipes fuscipes* Newstead, 1910 and *Glossina morsitans submorsitans* Newstead, 1910 in southwestern Saudi Arabia. *Ann. Soc. Belg. Med. Trop.* **70**, 281–287 (1990).
2. Leak, S. G. A. *Tsetse Biology and Ecology: Their Role in the Epidemiology and Control of Trypanosomosis*. (ILRI (aka ILCA and ILRAD), 1999).
3. Mattioli, R. C. *et al.* Tsetse and trypanosomiasis intervention policies supporting sustainable animal-agricultural development. *J. Food Agric. Environ.* **2**, 310–314 (2004).
4. Cecchi, G., Mattioli, R. C., Slingenbergh, J. & Rocque, S. D. L. Land cover and tsetse fly distributions in sub-Saharan Africa. *Med. Vet. Entomol.* **22**, 364–373 (2008).
5. Aksoy, S., Gibson, W. C. & Lehane, M. J. Interactions between tsetse and trypanosomes with implications for the control of trypanosomiasis. In *Advances in Parasitology* vol. 53 1–83 (Academic Press, 2003).
6. Roditi, I. & Lehane, M. J. Interactions between trypanosomes and tsetse flies. *Curr. Opin. Microbiol.* **11**, 345–351 (2008).
7. Geerts, S. & Holmes, P. H. *Drug management and parasite resistance in bovine trypanosomiasis in Africa*. (Food and Agriculture Organization of the United Nations, 1998).
8. Reinhardt, E. Travailler Ensemble: La Mouche Tse′ -tse′ et La Pauvrete Rurale. Accessed Nov 2021. https://www.un.org/french/pubs/chronique/2002/numero2/0202p17_la_mouche_tsetse.html. (2002).
9. Geiger, A. *et al.* Transcriptomics and proteomics in human African trypanosomiasis: Current status and perspectives. *J. Proteom.* **74**, 1625–1643 (2011).
10. Aksoy, S. & Rio, R. V. M. Interactions among multiple genomes: Tsetse, its symbionts and trypanosomes. *Insect Biochem. Mol. Biol.* **35**, 691–698 (2005).
11. Delespaux, V., Geysen, D., Van den Bossche, P. & Geerts, S. Molecular tools for the rapid detection of drug resistance in animal trypanosomes. *Trends Parasitol.* **24**, 236–242 (2008).
12. Vreysen, M. J. B. Principles of area-wide integrated tsetse fly control using the sterile insect technique. *Méd. Trop.* **61**, 397–411 (2001).
13. Schofield, C. J. & Kabayo, J. P. Trypanosomiasis vector control in Africa and Latin America. *Parasit. Vectors* **1**, 24 (2008).
14. Molyneux, D. H., Hopkins, D. R. & Zagaria, N. Disease eradication, elimination and control: The need for accurate and consistent usage. *Trends Parasitol.* **20**, 347–351 (2004).
15. Abbeele, J. V. D., Claes, Y., Bockstaele, D. V., Ray, D. L. & Coosemans, M. *Trypanosoma brucei* spp. development in the tsetse fly: Characterization of the post-mesocyclic stages in the foregut and proboscis. *Parasitology* **118**, 469–478 (1999).
16. Maudlin, I. & Welburn, S. C. Maturation of trypanosome infections in tsetse. *Exp. Parasitol.* **79**, 202–205 (1994).
17. Moloo, S. K., Asonganyi, T. & Jenni, L. Cyclical development of *Trypanosoma brucei gambiense* from cattle and goats in Glossina. *Cycl. Dev. Trypanos. Brucei Gamb. Cattle Goats Glossina* **43**, 407–408 (1986).
18. MacLeod, E. T., Darby, A. C., Maudlin, I. & Welburn, S. C. Factors affecting trypanosome maturation in tsetse flies. *PLoS One* **2**, 1–5 (2007).
19. Weiss, B. L., Wang, J., Maltz, M. A., Wu, Y. & Aksoy, S. Trypanosome infection establishment in the tsetse fly gut is influenced by microbiome-regulated host immune barriers. *PLoS Pathog.* **9**, e1003318 (2013).
20. O'Neill, S. L., Gooding, R. H. & Aksoy, S. Phylogenetically distant symbiotic microorganisms reside in Glossina midgut and ovary tissues. *Med. Vet. Entomol.* **7**, 377–383 (1993).
21. Cheng, Q. & Aksoy, S. Tissue tropism, transmission and expression of foreign genes in vivo in midgut symbionts of tsetse flies. *Insect Mol. Biol.* **8**, 125–132 (1999).
22. Wang, J., Weiss, B. L. & Aksoy, S. Tsetse fly microbiota: Form and function. *Front. Cell. Infect. Microbiol.* **3**, 69 (2013).
23. Dale, C. & Welburn, S. C. The endosymbionts of tsetse flies: Manipulating host–parasite interactions. *Int. J. Parasitol.* **31**, 628–631 (2001).
24. Trappeniers, K., Matetovici, I., Van Den Abbeele, J. & De Vooght, L. The tsetse fly displays an attenuated immune response to its secondary symbiont, *Sodalis glossinidius*. *Front. Microbiol.* **10**, 1650 (2019).
25. Medlock, J., Atkins, K. E., Thomas, D. N., Aksoy, S. & Galvani, A. P. Evaluating paratransgenesis as a potential control strategy for African trypanosomiasis. *PLoS Negl. Trop. Dis.* **7**, e2374 (2013).
26. Geiger, A., Ponton, F. & Simo, G. Adult blood-feeding tsetse flies, trypanosomes, microbiota and the fluctuating environment in sub-Saharan Africa. *ISME J.* **9**, 1496–1507 (2015).
27. Maudlin, I. & Ellis, D. S. Association between intracellular rickettsial-like infections of midgut cells and susceptibility to trypanosome infection in *Glossina* spp. *Z. Für Parasitenkd.* **71**, 683–687 (1985).
28. Geiger, A. *et al.* Vector competence of Glossina palpalis gambiensis for *Trypanosoma brucei* s.l. and genetic diversity of the symbiont *Sodalis glossinidius*. *Mol. Biol. Evol.* **24**, 102–109 (2007).
29. Hamidou Soumana, I. *et al.* The transcriptional signatures of *Sodalis glossinidius*. in the *Glossina palpalis gambiensis* flies negative for *Trypanosoma brucei gambiense* contrast with those of this symbiont in tsetse flies positive for the parasite: Possible involvement of a *Sodalis*-hosted prophage in fly *Trypanosoma* refractoriness?. *Infect. Genet. Evol.* **24**, 41–56 (2014).
30. Farikou, O. *et al.* Tripartite interactions between tsetse flies, *Sodalis glossinidius* and trypanosomes—An epidemiological approach in two historical human African trypanosomiasis foci in Cameroon. *Infect. Genet. Evol.* **10**, 115–121 (2010).
31. Makhulu, E. E. *et al.* Tsetse blood-meal sources, endosymbionts, and trypanosome infections provide insight into African trypanosomiasis transmission in the Maasai Mara National Reserve, a wildlife-human-livestock interface. *bioRxiv* https://doi.org/10.1101/2020.04.06.027367 (2020).
32. Wamwiri, F. N. *et al. Wolbachia*, *Sodalis* and trypanosome co-infections in natural populations of *Glossina austeni* and *Glossina pallidipes*. *Parasit. Vectors* **6**, 232 (2013).
33. Channumsin, M., Ciosi, M., Masiga, D., Turner, C. M. R. & Mable, B. K. *Sodalis glossinidius* presence in wild tsetse is only associated with presence of trypanosomes in complex interactions with other tsetse-specific factors. *BMC Microbiol.* **18**, 163 (2018).
34. Kanté Tagueu, S., Farikou, O., Njiokou, F. & Simo, G. Prevalence of *Sodalis glossinidius* and different trypanosome species in *Glossina palpalis palpalis* caught in the Fontem sleeping sickness focus of the southern Cameroon. *Parasite* **25**, 44 (2018).
35. Kame-Ngasse, G. I. *et al.* Prevalence of symbionts and trypanosome infections in tsetse flies of two villages of the "Faro and Déo" division of the Adamawa region of Cameroon. *BMC Microbiol.* **18**, 159 (2018).
36. Dennis, J. W. *et al. Sodalis glossinidius* prevalence and trypanosome presence in tsetse from Luambe National Park, Zambia. *Parasit. Vectors* **7**, 378 (2014).

37. Vreysen, M. J. B., Saleh, K. M., Zhu, Z.-R. & Suleiman, F. W. Responses of *Glossina austeni* to sticky panels and odours. *Med. Vet. Entomol.* **14**, 283–289 (2000).
38. Vreysen, M. J. B. *et al.* Area-wide integrated pest management of a *Glossina palpalis gambiensis* population from the Niayes area of Senegal: A review of operational research in support of an operational phased conditional approach. in. In *Area-Wide Integrated Pest Management: Development and Field Application* (eds Hendrichs, J. *et al.*) (CRC Press, 2021).
39. Van Den Abbeele, J. *et al.* Enhancing tsetse fly refractoriness to trypanosome infection—A new IAEA coordinated research project. *J. Invertebr. Pathol.* **112**, S142–S147 (2013).
40. van den Bossche, P. *et al.* Effect of isometamidium chloride treatment on susceptibility of tsetse flies (Diptera: Glossinidae) to trypanosome infections. *J. Med. Entomol.* **43**, 564–567 (2006).
41. Bouyer, J. Does isometamidium chloride treatment protect tsetse flies from trypanosome infections during SIT campaigns?. *Med. Vet. Entomol.* **22**, 140–143 (2008).
42. De Vooght, L., Caljon, G., De Ridder, K. & van den Abbeele, J. Delivery of a functional anti-trypanosome Nanobody in different tsetse fly tissues via a bacterial symbiont, *Sodalis glossinidius*. *Microb. Cell Factories* **13**, 1–10 (2014).
43. Caljon, G., De Vooght, L. & Van Den Abbeele, J. Options for the delivery of anti-pathogen molecules in arthropod vectors. *J. Invertebr. Pathol.* **112**, S75–S82 (2013).
44. Demirbas-Uzel, G. *et al.* Combining paratransgenesis with SIT: Impact of ionizing radiation on the DNA copy number of *Sodalis glossinidius* in tsetse flies. *BMC Microbiol.* **18**, 160 (2018).
45. Aksoy, E. *et al.* Analysis of multiple tsetse fly populations in Uganda reveals limited diversity and species-specific gut microbiota. *Appl. Environ. Microbiol.* **80**, 4301–4312 (2014).
46. Soumana, I. H. *et al.* The bacterial flora of tsetse fly midgut and its effect on trypanosome transmission. *J. Invertebr. Pathol.* **112**, S89–S93 (2013).
47. Maudlin, I., Welburn, S. C. & Mehlitz, D. The relationship between rickettsia-like-organisms and trypanosome infections in natural populations of tsetse in Liberia. *Trop. Med. Parasitol.* **41**, 265–267 (1990).
48. Doudoumis, V. *et al.* Tsetse-*Wolbachia* symbiosis: Comes of age and has great potential for pest and disease control. *J. Invertebr. Pathol.* **112**, S94–S103 (2013).
49. Baker, R. D., Maudlin, I., Milligan, P. J. M., Molyneux, D. H. & Welburn, S. C. The possible role of Rickettsia-like organisms in trypanosomiasis epidemiology. *Parasitology* **100**, 209–217 (1990).
50. Soumana, I. H. *et al.* Population dynamics of *Glossina palpalis gambiensis* symbionts, *Sodalis glossinidius*, and *Wigglesworthia glossinidia*, throughout host-fly development. *Infect. Genet. Evol.* **13**, 41–48 (2013).
51. Alam, U. *et al. Wolbachia* symbiont infections induce strong cytoplasmic incompatibility in the tsetse fly *Glossina morsitans*. *PLoS Pathog.* **7**, e1002415 (2011).
52. Mbewe, N. J., Mweempwa, C., Guya, S. & Wamwiri, F. N. Microbiome frequency and their association with trypanosome infection in male *Glossina morsitans centralis* of Western Zambia. *Vet. Parasitol.* **211**, 93–98 (2015).
53. Musaya, J. *et al.* Polymerase chain reaction identification of *Trypanosoma brucei* rhodesiense in wild tsetse flies from Nkhotakota Wildlife Reserve, Malawi. *Malawi Med. J. J. Med. Assoc. Malawi* **29**, 5–9 (2017).
54. Malele, I. I. *et al.* Multiple Trypanosoma infections are common amongst Glossina species in the new farming areas of Rufiji district, Tanzania. *Parasit. Vectors* **4**, 217 (2011).
55. Ngonyoka, A. *et al.* Patterns of tsetse abundance and trypanosome infection rates among habitats of surveyed villages in Maasai steppe of northern Tanzania. *Infect. Dis. Poverty* **6**, 126 (2017).
56. Lefrançois, T. *et al.* Polymerase chain reaction characterization of trypanosomes in *Glossina morsitans submorsitans* and *G. tachinoides* collected on the game ranch of Nazinga, Burkina Faso. *Acta Trop.* **72**, 65–77 (1999).
57. Djohan, V. *et al.* Detection and identification of pathogenic trypanosome species in tsetse flies along the Comoé River in Côte d'Ivoire. *Parasite Paris Fr.* **22**, 18 (2015).
58. Nayupe, S. F. *et al.* The use of molecular technology to investigate trypanosome infections in tsetse flies at Liwonde Wild Life Reserve. *Malawi Med. J. J. Med. Assoc. Malawi* **31**, 233–237 (2019).
59. Karshima, S. N., Ajogi, I. & Mohammed, G. Eco-epidemiology of porcine trypanosomosis in Karim Lamido, Nigeria: Prevalence, seasonal distribution, tsetse density and infection rates. *Parasit. Vectors* **9**, 448 (2016).
60. Simo, G. *et al.* Molecular identification of *Wolbachia* and *Sodalis glossinidius* in the midgut of *Glossina fuscipes quanzensis* from the Democratic Republic of Congo. *Parasite Paris Fr.* **26**, 5 (2019).
61. Odeniran, P. O., Macleod, E. T., Ademola, I. O. & Welburn, S. C. Endosymbionts interaction with trypanosomes in Palpalis group of Glossina captured in southwest Nigeria. *Parasitol. Int.* **70**, 64–69 (2019).
62. Couey, H. M. & Chew, V. Confidence limits and sample size in quarantine research. *J. Econ. Entomol.* **79**, 887–890 (1986).
63. Geiger, A., Ravel, S., Frutos, R. & Cuny, G. *Sodalis glossinidius* (Enterobacteriaceae) and vectorial competence of *Glossina palpalis gambiensis* and *Glossina morsitans morsitans* for *Trypanosoma congolense* savannah type. *Curr. Microbiol.* **51**, 35–40 (2005).
64. Geiger, A., Cuny, G. & Frutos, R. Two tsetse fly species, *Glossina palpalis gambiensis* and *Glossina morsitans morsitans*, carry genetically distinct populations of the secondary symbiont *Sodalis glossinidius*. *Appl. Environ. Microbiol.* **71**, 8941–8943 (2005).
65. Geiger, A. *et al.* Differential expression of midgut proteins in *Trypanosoma brucei gambiense*-stimulated versus non-stimulated *Glossina palpalis gambiensis* flies. *Front. Microbiol.* **6**, 444 (2015).
66. Haines, L. R., Hancock, R. E. W. & Pearson, T. W. Cationic antimicrobial peptide killing of African trypanosomes and *Sodalis glossinidius*, a bacterial symbiont of the insect vector of sleeping sickness. *Vector-Borne Zoonotic Dis.* **3**, 175–186 (2003).
67. Challier, A. & Laveissière, C. Un nouveau piege pour la capture des glossines (Glossina: Diptera, Muscidae): Description et essais sur le terrain. *Cah. ORSTOM Sér. Entomol. Médicale Parasitol.* **11**, 251–262 (1973).
68. Lancien, J. Description du piege monoconique utilise pour l'elimination des glossines en Republique Populaire du Congo. *Cah. ORSTOM Sér. Entomol. Médicale Parasitol.* **19**, 235–238 (1981).
69. Laveissière, C. & Grébaut, P. The trapping of tsetse flies (Diptera: Glossinidae). Improvement of a model: The Vavoua trap. *Trop. Med. Parasitol. Off. Organ Dtsch. Tropenmedizinische Ges. Dtsch. Ges. Tech. Zusammenarbeit GTZ* **41**, 185–192 (1990).
70. Brightwell, B. *et al.* A new trap to *Glossina pallidipes*. *Trop. Pest Manag.* **33**, 151–159 (1987).
71. Brightwell, R., Dransfield, R. D. & Kyorku, C. Development of a low-cost tsetse trap and odour baits for *Glossina pallidipes* and *G. longipennis* in Kenya. *Med. Vet. Entomol.* **5**, 153–164 (1991).
72. Hargrove, J. E. & Langley, P. A. Sterilizing tsetse (Diptera: Glossinidae) in the field: A successful trial. *Bull. Entomol. Res.* **80**, 397–403 (1990).
73. Mihok, S. The development of a multipurpose trap (the Nzi) for tsetse and other biting flies. *Bull. Entomol. Res.* **92**, 385–403 (2002).
74. Kappmeier, K. A newly developed odour-baited 'H trap' for the live collection of *Glossina brevipalpis* and *Glossina austeni* (Diptera: Glossinidae) in South Africa. *Onderstepoort J. Vet. Res.* **67**, 15–26 (2000).
75. Baker, M. D. & Krafsur, E. S. Identification and properties of microsatellite markers in tsetse flies *Glossina morsitans* sensu lato (Diptera: Glossinidae). *Mol. Ecol. Notes* **1**, 234–236 (2001).
76. Njiru, Z. K. *et al.* The use of ITS1 rDNA PCR in detecting pathogenic African trypanosomes. *Parasitol. Res.* **95**, 186–192 (2005).
77. Toh, H. *et al.* Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res.* **16**, 149–156 (2006).
78. Weiss, B. L., Maltz, M. & Aksoy, S. Obligate symbionts activate immune system development in the tsetse fly. *J. Immunol.* **188**, 3395–3403 (2012).

79. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2021).
80. Baier, T. & Neuwirth, E. Excel: COM: R. *Comput. Stat.* **22**, 91–108 (2007).
81. RStudio Team. *RStudio: Integrated Development Environment for R*. (RStudio, Inc., 2016).
82. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
83. Sarkar, D. *Lattice: Multivariate Data Visualization with R*. (Springer Science & Business Media, 2008).
84. Fox, J. & Weisberg, S. *An R Companion to Applied Regression*, 2nd ed. (Sage, 2019).
85. Jeffrey B. Arnold. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R package version 4.2.4. (2021).
86. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S*. (Springer, 2002).
87. Clarke, K. R. & Gorley, R. N. Getting started with PRIMER v7. Accessed Nov 2021. http://updates.primer-e.com/primer7/manuals/Getting_started_with_PRIMER_7.pdf. (2016).
88. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32–46 (2001).

Acknowledgements
The authors thank many collaborators in Africa who helped to collect tsetse fly samples.

Author contributions
A.M.M.A., R.L.M. and M.J.B.V. Conceived and designed Research; M.M.D., M.K.D., P.M., G.M.S.O., G.D.-U., F.G., F.C.M., L.N., S.M., J-P.R., A.M.G.B., S.P. and C.J.B., collected data and conducted research; A.M.M.A., M.M.D., M.K.D., J.V.D.A. and A.G.P. analyzed and interpreted data; A.M.M.A., M.M.D., M.K.D. wrote the initial paper; A.M.M.A., A.G.P., J.V.D.A., R.L.M., and M.J.B.V. revised the paper; A.M.M.A had primary responsibility for final content. All authors read and approved the final manuscript.

Funding
This study was supported by the Joint FAO/IAEA Insect Pest Control Subprogramme.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-06699-2.

Correspondence and requests for materials should be addressed to A.M.M.A.-A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Chapter 4: Infection with endosymbiotic *Spiroplasma* disrupts tsetse (*Glossina fuscipes fuscipes*) metabolic and reproductive homeostasis

PLOS PATHOGENS

RESEARCH ARTICLE

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

Jae Hak Son[1☯], Brian L. Weiss[1☯*], Daniela I. Schneider[1¤a], Kiswend-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. Abd-Alla[2], Serap Aksoy[1*]

1 Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, Connecticut, United States of America, **2** Insect Pest Control Laboratory, Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture, 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, Gulu, Uganda, **5** Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut, United States of America

☯ These authors contributed equally to this work.
¤a Current address: Moffit Cancer Center, Tampa, Florida, United States of America
¤b Current address: Department of Biology, Utah State University, Logan, Utah, United States of America
¤c Current address: Department of Entomology and Nematology, University of California, Davis, United States of America
¤d Current address: Wadsworth Centre, New York State Department of Health, Slingerlands, New York, United States of America
* brian.weiss@yale.edu (BLW); serap.aksoy@yale.edu (SA)

OPEN ACCESS

Citation: Son JH, Weiss BL, Schneider DI, Dera K-sM, Gstöttenmayer F, Opiro R, et al. (2021) Infection with endosymbiotic *Spiroplasma* disrupts tsetse (*Glossina fuscipes fuscipes*) metabolic and reproductive homeostasis. PLoS Pathog 17(9): e1009539. https://doi.org/10.1371/journal.ppat.1009539

Editor: Bruno Lemaitre, EPFL, SWITZERLAND

Received: April 6, 2021

Accepted: September 7, 2021

Published: September 16, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.ppat.1009539

Data Availability Statement: Most of the data are contained within the manuscript and/or supporting information files. RNA-seq data has been submitted to NCBI, GEO accession number

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* lab 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

GSE183197 and BioProject accession number PRJNA759598.

Funding: Funding was generously provided by the NIH/NIAID (RO1AI068932), the Ambrose Monell Foundation (monellfoundation.org), and the Li Foundation (lifoundation.org) to SA. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

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

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.

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

Infection with *Spiroplasma* can also impact host reproductive fitness and lead to reproductive abnormalities. In *Drosophila*, females infected with *Spiroplasma* are less fecund and produce fewer eggs, which may be a consequence of nutritional competition between the fly and bacterium for metabolically important free (hemolymph-borne) lipids [20]. In several insect taxa, the bacterium induces a male killing phenotype [21–23]. The potential functional role of *Spiroplasma* in tsetse reproductive physiology is currently unknown. Herein we performed a detailed investigation of the molecular and physiological associations that characterize reproductive aspects of the *Gff-Spiroplasma* symbiosis. Specifically, we performed RNA sequencing (RNA-seq) of reproductive tissues from field captured (Uganda) *Spiroplasma* infected and uninfected male and female *Gff* and report on genes and pathways that are differentially regulated in the presence of the bacterium. We also made use of a *Gff* lab line that carries a heterogenous infection with *Spiroplasma* to characterize the trans-generational transmission dynamics of this endosymbiont and to characterize *Spiroplasma*-induced metabolic and reproductive phenotypes in tsetse. Knowledge obtained from this study provides insight into the physiological mechanisms that underlie the tsetse-*Spiroplasma* symbiosis and may have translational implications with respect to controlling tsetse populations and the ability of the fly to transmit African trypanosomes.

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 in S1 Appendix).

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

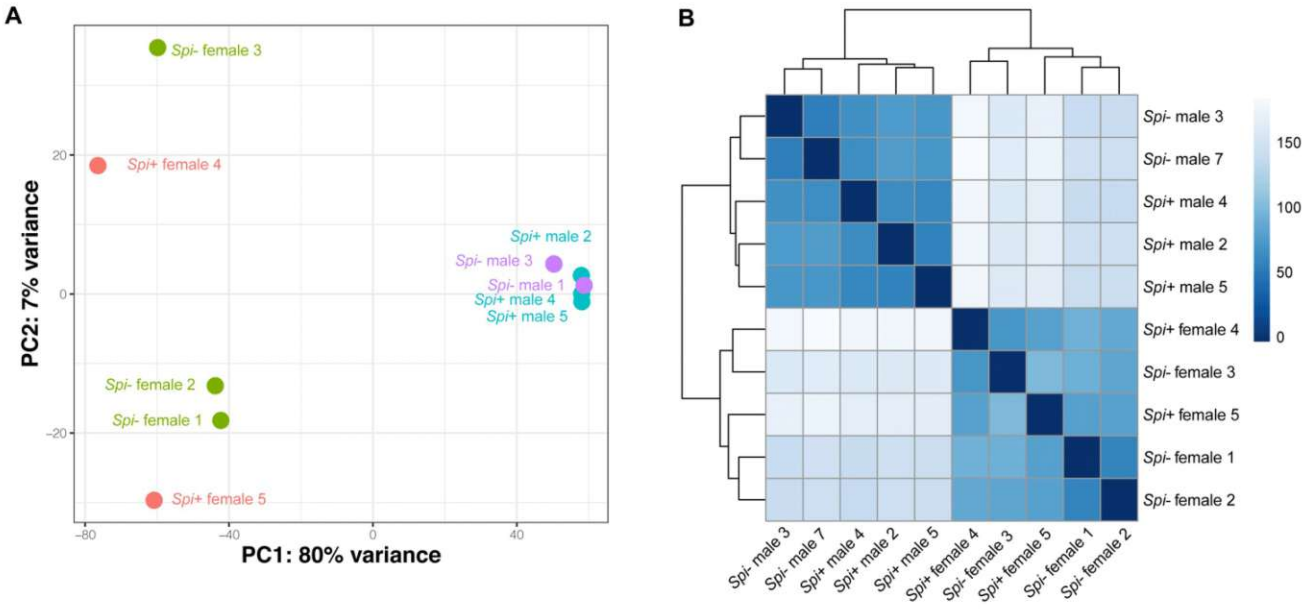

Fig 1. Principal component analysis (A) and Hierarchical clustering (B) of expression data. The Principal Component Analysis (PCA) is based on differentially expressed genes from the female and male samples, while the Hierarchical Clustering (HC) reflects all genes within the dataset. *Spiroplasma*-infection status and sex are color-coded in the PCA.

https://doi.org/10.1371/journal.ppat.1009539.g001

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 expression within the female reproductive tissues when compared to males, which is similar to the results shown in our PCA.

The impact of *Spiroplasma* on sex-biased gene expression in males and females

As our biological replicates clustered by sex, we first quantified the proportion of differentially expressed (DE) genes between females and males (sex-biased genes) using sex as a factor in the model we created in DESeq2 (see Materials and Methods for details). Of the 15,247 genes annotated in the *Gff* genome, we detected 10,540 genes expressed in our transcriptomes, with an adjusted *p*-value created by FDR correction (S1 Data) [24]. We observed that 21.7% of these genes (2,288/10,540) were preferentially expressed in females (female-biased genes; $\log_2$male/female<0 and adjusted $p<0.05$), while 20.5% (2,164/10,540) were preferentially expressed in males (male-biased genes; $\log_2$male/female>0 and adjusted $p<0.05$). Using pairwise comparisons for *Spiroplasma* infection status within each sex from the DESeq2 model, we found that a total of 194 and 299 genes were DE upon *Spiroplasma* infection in males and females, respectively. Within the male dataset, we determined that 117 (60.3%) and 77 (39.7%) of the DE genes were up- and down-regulated upon infection (adjusted $p<0.05$), respectively (Fig 2A, S2 Table in S1 Appendix and S2 Data). A similar analysis of the 299 DE genes in females showed that 61 (21.4%) and 238 (79.6%) were up- and down-regulated in the presence of *Spiroplasma*, respectively (Fig 2B and S3 Table in S1 Appendix).

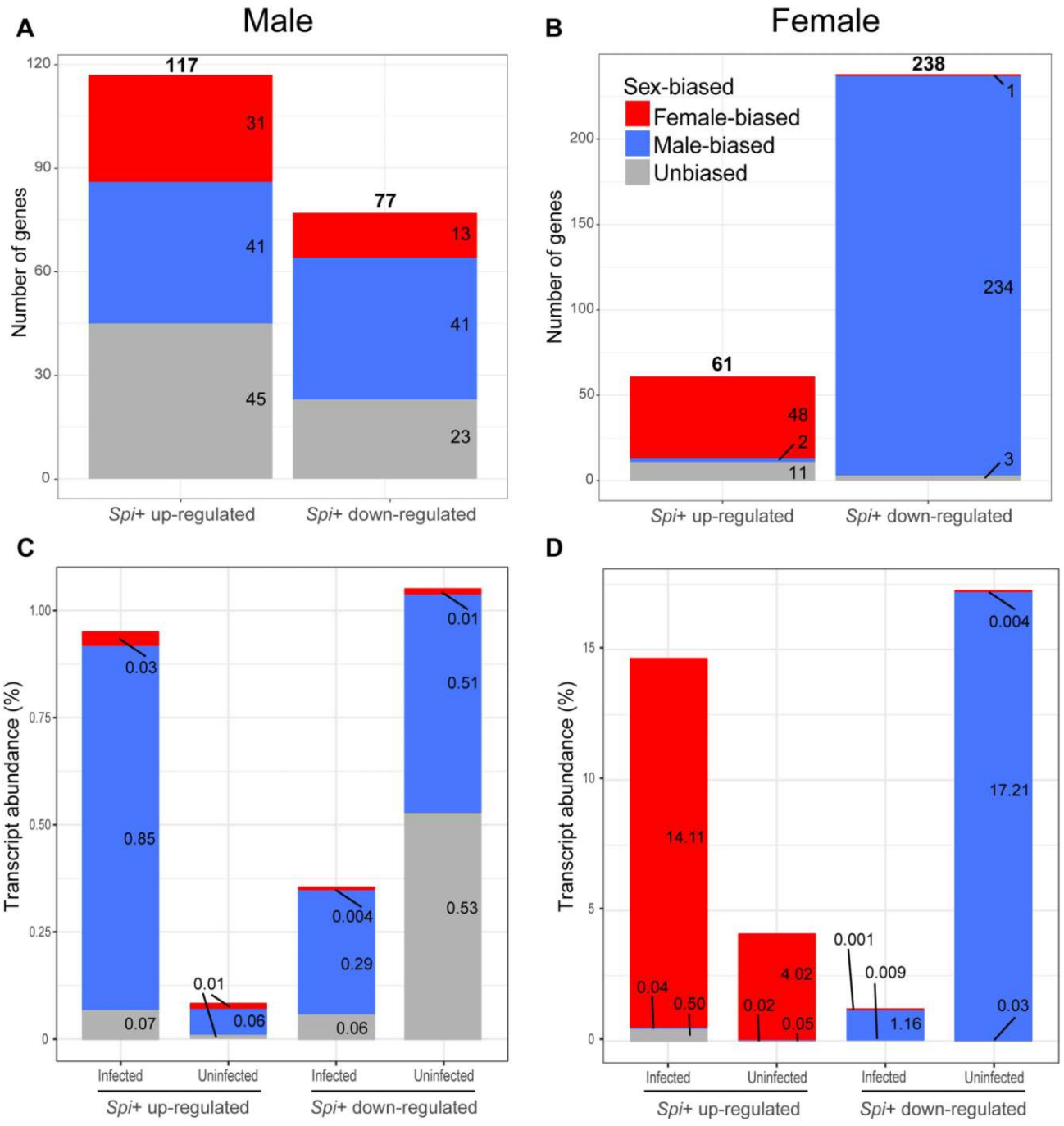

Fig 2. Enrichment of genes with sex-biased expression that are differentially expressed between Gff^{Spi+} and Gff^{Spi-} flies. (A) The number of genes with sex-biased expression that are up- and down-regulated in Gff^{Spi+} males. (B) The number of genes with sex-biased expression profiles that are up- and down-regulated in Gff^{Spi+} females. (C) Transcript abundance (%) of up- and down-regulated genes relative to total transcript abundance in Gff^{Spi+} and Gff^{Spi-} males. (D) Transcript abundance of up- and down-regulated genes relative to total transcript abundance in Gff^{Spi+} and Gff^{Spi-} females. Female-biased genes are represented as red blue and male-biased genes as blue. Genes with no sex-bias are indicated as gray.

https://doi.org/10.1371/journal.ppat.1009539.g002

Spiroplasma infections can manipulate the reproductive physiology of their host insects [25]. In *Drosophila*, infections with the *S. poulsonii* strain MSRO affects gene regulation and dosage compensation machinery in males, resulting in male-killing during embryogenesis [22,26]. Because genes displaying sex-biased expression profiles in reproductive tissues can confer sexual 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 and S2 Table in S1 Appendix). 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 and S2 Table S1 Appendix; $p<0.001$ 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 in S1 Appendix; $p<10^{-20}$ in Fisher's exact test). Conversely, only one of the 238 genes down-regulated in Gff^{Spi+} females was female-biased (Fig 2B and S3 Table in S1 Appendix), while 234 (98%) were male-biased (Fig 2B and S3 Table in S1 Appendix). The male-biased genes were thus overrepresented within the down-regulated genes (S3 Table in S1 Appendix; $p<10^{-149}$ 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 transcriptome in the infected group, up from the 4.02% transcript abundance in the uninfected group (Fig 2D). The transcript abundance of the 234 down-regulated male-biased genes in the female dataset represented 1.16% of the total transcriptome in the infected group, down from 17.21% transcript abundance in the uninfected group (Fig 2D).

In both sexes, unbiased (non-sex biased) genes were underrepresented in the DE datasets when compared to the genome-wide unbiased genes (Fig 2, S2 and S3 Tables in S1 Appendix; $p<10^{-4}$ for males and $p<10^{-9}$ for females in Fisher's exact test). Collectively, these results suggest that *Spiroplasma* infections strongly influence the expression of sex-biased genes in tsetse's reproductive tissues. Given that sexual dimorphisms are largely shaped by sex-biased gene expressions [27,28], the significant number of sex-biased genes and transcript abundance affected by *Spiroplasma* in females relative to males suggests that the endosymbiont may affect female reproductive physiology to a greater extent.

Spiroplasma infection effects on genes that encode spermatophore constituents

When tsetse copulate, the ejaculate, which is composed of sperm and secretory products (i.e., seminal fluid) derived from the male testes and accessory glands, are transferred to the female uterus where they are encapsulated into a spermatophore. The spermatophore functions as a protective container for the ejaculate and ensures that sperm and seminal fluid are delivered to the female spermathecal ducts, which in turn modifies female behavior, including potential inhibition of further matings by competing males [3–5]. Within 24 h of the commencement of copulation, sperm are transferred to the female's spermatheca and the spermatophore is discharged from the uterus [3]. We had previously identified the proteins detected in the *G.*

morsitans morsitans (Gmm) spermatophore by collecting the structures from the female spermatheca shortly after completion of copulation [2]. We first identified the homologs of these *Gmm* protein encoding genes 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 peptide *defensin* (GFUI031425), and *easter* (GFUI050711), which encodes a serine-protease required to process the extracellular Toll ligand Spätzle. We also noted that in addition to the Toll signaling pathway, several 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]. Taken together, this analysis indicates that *Spiroplasma* infection induces the upregulation of a highly restricted repertoire of immune related genes in the reproductive tract of *Gff* males. The bacterium had no impact on immune gene expression in *Gff* females.

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 (S1A Fig and S3 Data). The down-regulated genes in Gff^{Spi+} males were enriched for lipid catabolic process and peptidyl-dipeptidase activity (S1B Fig and S3 Data). The up-regulated genes in Gff^{Spi+} females were enriched for lipid metabolic process and defense response (S1C Fig and S3 Data) while the down-regulated genes in Gff^{Spi+} females were enriched for peptidase activity, proteolysis, and chitin metabolic process (S1D Fig and S3 Data).

Spiroplasma infection effects on female gene expression

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

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 *Gff* 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 3A and S4 Data). 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.

We have previously demonstrated that experimental depletion of *Wigglesworthia* density (via treatment with antibiotics) significantly impairs tsetse's reproduction [12,32–34]. With this in mind, we measured the relative densities of endosymbiotic *Wigglesworthia* and *Sodalis* in Gff^{Spi+} and Gff^{Spi-} females to determine if these symbionts are impacted by the presence of *Spiroplasma*. We determined that Gff^{Spi+} and Gff^{Spi-} females house similar densities of *Wigglesworthia* (Fig 3C, left graph) and *Sodalis* (Fig 3C, right graph). Because *Spiroplasma* infection did not impact the density of these tsetse symbionts, we next investigated whether the prolonged GC length presented by Gff^{Spi+} females reflected 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) [35], which originate in tsetse's fat body and are transferred through hemolymph to the milk gland where they are incorporated into maternal milk secretions [36,37]. To determine if *Spiroplasma* hijacks tsetse TAG, at the nutritional expense of developing intrauterine larvae, we compared TAG levels in

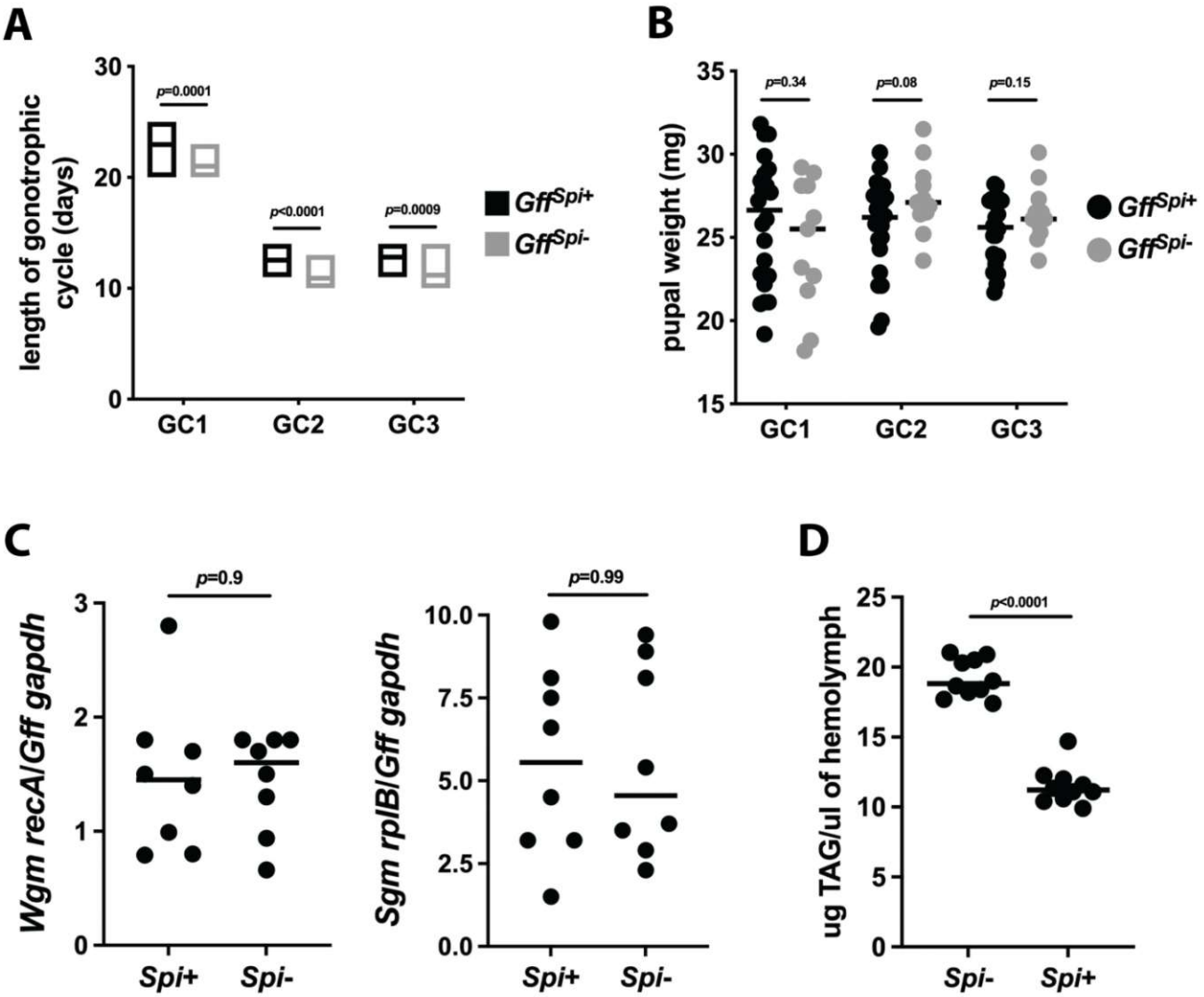

Fig 3. Impact of *Spiroplasma* infection on the reproductive and nutritional fitness of female tsetse flies. (A) Gonotrophic cycle (GC) length of offspring from Gff^{Spi+} and Gff^{Spi-} females. Age-matched, pregnant females from each group ($n = 34$ per group) were housed in individuals cages and monitored daily to observe frequency of pupal deposition. Statistical significance was determined via log-rank test. (B) The weight of pupae deposited by Gff^{Spi+} and Gff^{Spi-} females. Each dot represents an individual pupa, bars represent median values of pupae form each GC. Statistical significance was determined via multiple t-tests with correction via the Holm-Šídák method. (C) Relative densities of *Wigglesworthia* and *Sodalis* in Gff^{Spi+} and Gff^{Spi-} females. Relative *Wigglesworthia recA and Sodalis rplB* copy number was quantified using genomic DNA derived from Gff^{Spi+} or Gff^{Spi-} female midguts (including the *Wigglesworthia* harboring bacteriome organ). *Wigglesworthia recA and Sodalis rplB* was normalized relative to *Gff gapdh* copy number in each sample. Each dot represents one biological replicate (three midguts per replicate), and bars indicate median values. Statistical significance was determined via students t-test. (D) 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 (S2 Fig). Each dot represents an individual pupa, bars represent median values of pupae form each GC. Statistical significance was determined via unpaired t-test.

https://doi.org/10.1371/journal.ppat.1009539.g003

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 3D). 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

important lipids may result in the relatively long GC we observed in Spi^+ mothers because less of this nutrient is available for incorporation into milk secretions.

Impact of *Spiroplasma* infection on *Gff* male reproductive fitness

Our transcriptomic data informed us that infection with *Spiroplasma* impacts the expression of several genes that encode spermatophore-associated proteins. These proteins, many of which arise from the male accessory gland and are transferred in seminal fluid, play a prominent role in sperm fitness [2,38]. We thus compared the fitness of sperm derived from Spi^+ compared to Spi^- males to determine if infection with the bacterium impacts the fitness off *Gff* sperm. To do so we quantified the transcript abundance of *sperm-specific dynein intermediate chain* (*sdic*; VectorBase gene ID GFUI025244) in the spermathecae of 7 day old female flies two days after having mated with either Spi^+ or Spi^- males. The *Drosophila* homologue of this gene is expressed exclusively in sperm cells [39] and has been used to quantify sperm abundance and competitiveness (as a function of motility) [40–42]. We observed that *sdic* transcript abundance in the spermathecae of females that mated with Spi^- males was significantly higher than in that of their counterparts that mated with Spi^+ males (Fig 4). This conspicuous difference in the abundance of *sdic* transcripts expressed by sperm within the spermathecae of females that mated with Spi^- versus Spi^+ males was surprising considering the fact that expression of this gene in the reproductive tract of field captured Spi^- and Spi^+ males was not significantly different when measured by RNAseq (S1 Data). To validate this RNAseq data, we used RT-qPCR to measure *sdic* expression in sperm housed in the reproductive tract of 7 days old Spi^- and Spi^+ lab reared *Gff* males (these sperm were the same age as those used to measure *sdic* transcript abundance in the spermathecae of mated females). Similar to what we observed via RNAseq, *sdic* expression was not significantly different in sperm housed in the male reproductive tract (Fig 4).

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

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 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.

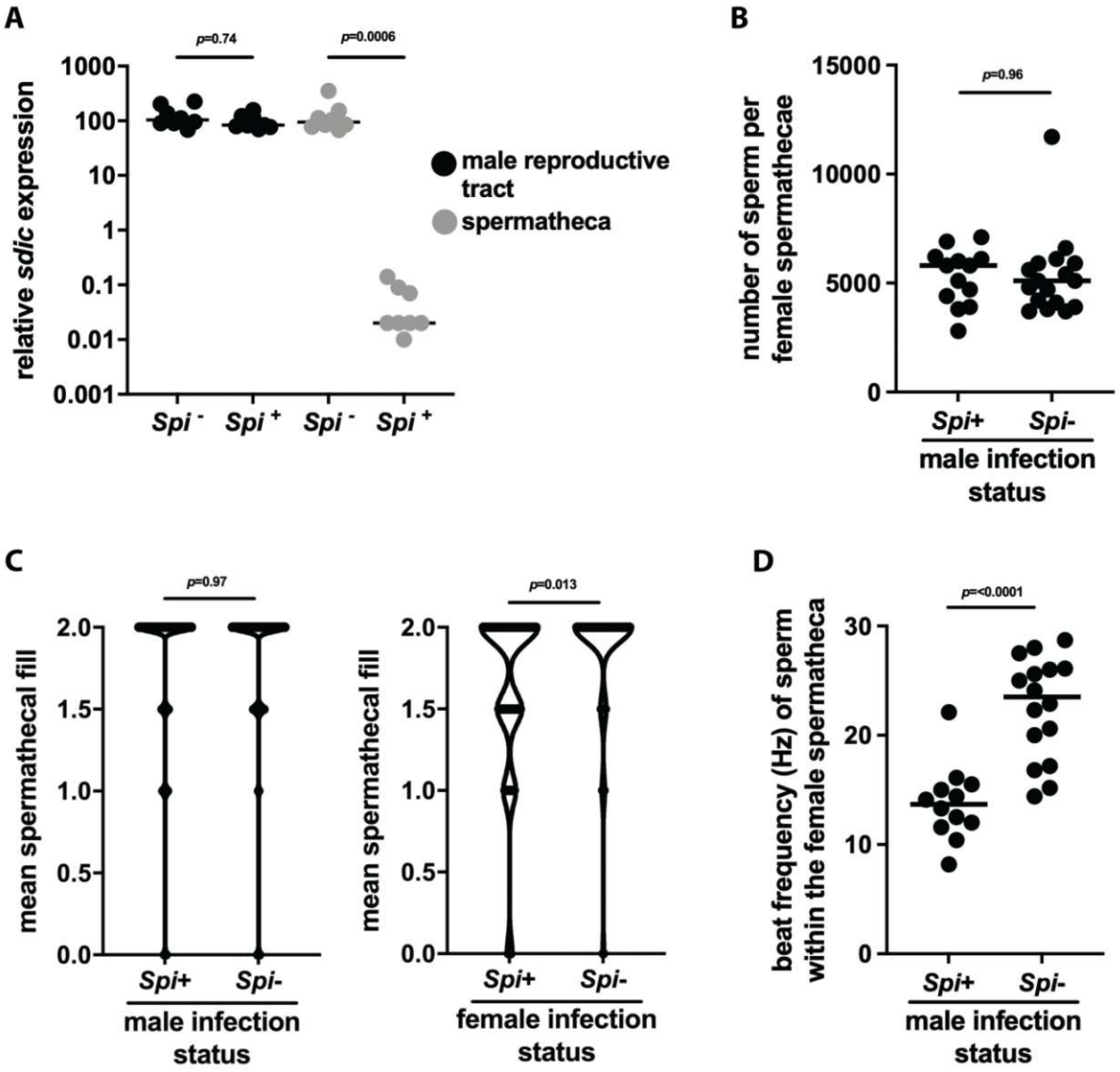

Fig 4. The impact of *Spiroplasma* infection on male *Gff* reproductive fitness. (A) Relative expression of *sperm-specific dynein intermediate chain* (*sdic*) in sperm located within the reproductive tract of Gff^{Spi+} and Gff^{Spi-} males or within the female spermatheca following copulation. *Sdic* 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 [75,76]. (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 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.

https://doi.org/10.1371/journal.ppat.1009539.g004

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%. We performed three distinct experiments as a means of quantitating the fidelity of *Spiroplasma* vertical transmission (see Materials and Methods for experimental details).

For experiment 1, we observed that 86% of adult offspring (from three GCs combined) derived from matings between Spi^+ mothers and Spi^+ fathers were Spi^+ as were 97% of adult offspring from matings between Spi^+ mothers and Spi^- fathers. Also in experiment 1 we observed no Spi^+ adult offspring when mothers lacked the bacterium, regardless of their mate's infection status (Table 1 and S5 Data). For experiment 2, 67% of pupal offspring (from three GCs combined) derived from matings between Spi^+ mothers and Spi^+ fathers were Spi^+ as were 44% of pupal offspring from matings between Spi^+ mothers and Spi^- fathers (Table 1 and S5 Data). In experiment 2, four matings between Spi^- mothers and Spi^+ fathers collectively resulted in the deposition of nine pupae, four of which carried *Spiroplasma* infections. Interestingly, in three out of four mating pairs, the second offspring (GC2) was positive for the *Spiroplasma* infection while the first offspring (GC1) was negative (Table 1 and S5 Data), indicating imperfect paternal transmission at best. The varying results we obtained between these experiments could reflect the different developmental stages of the offspring we tested (experiment 1, newly eclosed adults; experiment 2, 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 2 would persist through the lengthy pupal stage (~ 30 days) and into adulthood.

Finally, in experiment 3, we determined *Spiroplasma* infection status of mothers only. Of the Spi^+ mothers observed, 73% (11/15) and 100% (9/9) of their pupal offspring from GCs 1 and 2, respectively, were infected with *Spiroplasma* (no Spi^+ mothers in this experiment survived long enough to produce GC 3 offspring). Nine out of the 22 total progeny from Spi^- mothers were Spi^+, although transmission did not occur across sequential GCs (S5 Table in

Table 1. Vertical transmission of *Spiroplasma* across three gonotrophic cycles (GCs).

	Spiroplasma infection status			# of *Spiroplasma* infected offspring/total offspring sample/GC			
experiment #[a]	mother	father	number of mating pairs[b]	GC1	GC2	GC3	total % transmission[c]
Exp 1	+	+	7	6/7	6/7	6/7	86% (18/21)
		-	13	13/13	13/13	12/13	97% (38/39)
	-	+	9	0/9	0/9	0/9	0% (0/27)
		-	5	0/5	0/5	0/5	0% (0/15)
Exp 2	+	+	5	4/5	2/3	0/1	67% (6/9)
		-	7	6/7	3/5	2/2	79% (11/14)
	-	+	4	1/4	3/3	0/2	44% (4/9)
		-	1	0/1	0/1	0/0	0% (0/2)

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

[b]Number of mating pairs at the beginning of each experiment. Not all mothers survived for three GCs.

[c]Percentage of *Spiroplasma* infected offspring derived from each parental mating pair over the course of all three GCs combined.

https://doi.org/10.1371/journal.ppat.1009539.t001

S1 Appendix and S5 Data). While this experiment suggests possible paternal *Spiroplasma* transmission, we cannot confirm the route of infection because we did not test the *Spiroplasma* status of these fathers.

Taken together, these transmission data indicate that *Spiroplasma* is vertically transmitted via maternal lineages with high frequency, although paternal transmission may provide a less frequent route.

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

Sex-biased gene expression strongly influences sexual phenotypes in most animals [27,28], and a large proportion of sex-biased genes are expressed in reproductive tissues [43,44]. With this in mind, the differences in sex-biased gene expression we observed in between Gff^{Spi-} and Gff^{Spi+} females suggest that housing the bacterium could induce beneficial reproductive phenotypes. Our transcriptomic data from Gff^{Spi+} females showed that female-biased genes are up-regulated while male-biased genes are down-regulated relative to their uninfected (Gff^{Spi-}) counterparts. Notably, we observed that *Spiroplasma* infection induced a significant upregulation of genes that encode tsetse milk gland proteins (MGPs). These tsetse specific molecules are abundantly expressed in the accessory gland of pregnant females and are a prominent component of tsetse milk [1,45]. MGPs 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 [37]. Thus, when faced with depleted levels of circulating TAG, which comprises a substantial fraction of the lipids found in tsetse milk [35], *Spiroplasma* infected females may adaptively increase MGP production in an effort to scavenge and transport other lipids to developing intrauterine larvae. This outcome may reflect one *Spiroplasma* induced change in sex-biased gene expression that confers a fitness benefit to Gff^{Spi+} females.

The depletion of circulating TAG by hemolymph borne *Spiroplasma* likely accounts for the increase we observed in GC length in pregnant mothers that harbor the bacterium. Despite the longer larval development period we observed for progeny of Spi^+ compared to Spi^- mothers, the weight of pupal offspring deposited was similar between the two groups. Because female tsetse produce unusually few offspring (6–8) over the course of their lifespan (compared to other insects), an increase in GC length would likely result in a significant reduction in population size over time. In fact, infection with a trypanosome strain that induces a metabolically costly immune response also lengthens tsetse's GC by a duration similar to that (approximately 2 days) induced by infection with *Spiroplasma* [46]. Mathematical modelling indicates that this increase in GC length would theoretically reduce tsetse fecundity by approximately 30% over the course of a female's reproductive lifespan [46]. Thus, a moderate trypanosome infection prevalence of 26%, which is similar to the *Spiroplasma* infection prevalence in field captured *Gff* (5–34%, depending on population geographic location) [14,15], could significantly decrease fly population size [46]. The models also predict that infection prevalence above these frequencies would result in a population crash.

Our results indicate that infection with *Spiroplasma* also exerts a significant impact on reproductive processes in laboratory reared *Gff* males. We observed that sperm from both Spi^+ and Spi^- males expressed the same abundance of sperm-specific *sdic* transcripts prior to mating, but that following insemination and transfer to the female spermatheca, sperm from Spi^+ males expressed conspicuously fewer transcripts of this gene than did sperm that had originated from Spi^- males. The *Drosophila* homolog of this gene encodes a cytoplasmic dynein intermediate chain that is necessary for the proper function of the cytoplasmic dynein motor protein complex [39], which is involved in sperm motility [47]. Accordingly, one prominent phenotype we observed was that following insemination and transfer to the female spermatheca, the motility of sperm that had originated from Spi^+ males was impaired when compared to that of sperm that had originated from Spi^- males. This finding implies that *sdic* exhibits the same function in *Gff* as it does in *Drosophila*. Motility is of paramount importance to sperm competitiveness [48], and sperm with impaired motility have less fertilization success than heathy sperm [49]. Field data indicate that wild populations of *Gff* females can exhibit polyandry and store sperm from more than one mate [7]. Our data suggest that sperm transferred by Spi^+ males may be at a competitive disadvantage in comparison to sperm derived from Spi^- males because they exhibit reduced motility. It remains to be determined if, as a means of promoting polyandry, females seek additional mates if they first copulate with a Spi^+ male that transfers motility compromised sperm.

Spiroplasma mediated mechanisms that influence the regulation of *sdic* expression are currently unknown. *Spiroplasma* could mark sperm in the reproductive tract of *Gff* males [50], possibly via post-transcriptional modification of *sdic*, such that sperm motility is compromised following insemination and transfer to the female spermatheca. Alternatively, female contributions to the spermatophore could also influence *sdic* expression after sperm are packaged into the structure. In *Spiroplasma* infected females, we have observed 12 DE genes that encode products identified as constituents spermatophore. Further experiments are necessary to determine if any of these *Spiroplasma* modified female products, including two transcription factors and three serine protease inhibitors, play a role in the sperm modifications we observed here.

Infections with heritable endosymbionts can come at a significant metabolic cost to their host [51]. This outcome appears to also be the case with the *Spiroplasma* and tsetse symbiosis, and results in a significant reduction in the reproductive fitness of both female and male flies. However, from an evolutionary perspective, these microbes, including *Spiroplasma*, usually also provide an overall fitness advantage in order for infections to be maintained. In the tsetse

model system, the fly, the commensal symbiont *Sodalis*, as well as pathogenic trypanosomes, typically compete for host nutrients as they are auxotrophic for metabolically critical B vitamins [12,52–54]. Interestingly, these nutrients are present in low quantities in tsetse's vertebrate blood specific diet, and they are instead supplemented via the mutualistic endosymbiont, *Wigglesworthia* [11–13]. Despite the negative impacts of *Spiroplasma* on tsetse's fecundity, this bacterium also benefits its host by creating an environment within the fly that is hostile to parasitic trypanosomes [15]. The parasite resistance phenotype conferred by *Spiroplasma* infections could arise from induced host immune responses, or could reflect competition for critical nutrients (i.e., found in the bloodmeal and/or produced by tsetse) that both organisms require to sustain their metabolic needs [55–57]. This effect of inhibiting metabolically costly trypanosome infections could thus offset the cost of housing similarly detrimental *Spiroplasma* infections. Additionally, we demonstrate that while *Spiroplasma* can be both matrilineally and patrilineally transmitted, vertical transmission of the bacterium within our *Gff* colony does not occur at a frequency of 100%. This data corroborates observations from the field, which reveal that individual flies between and within distinct populations also present heterogeneous *Spiroplasma* infections [14,15]. In our study we observed a high frequency of maternal transmission into pupal and adult offspring as well as low frequency paternal transmission into pupal offspring. If *Spiroplasma* detected in the pupal stage survive and colonize adults, this outcome could contribute to the heterogenous infection prevalence we observe in the *Gff* lab line. The absence of steadfast vertical transmission of *Spiroplasma* by *Gff* could facilitate fly survival by reducing the population level detrimental impact (i.e., prolonged GC length and reduced sperm motility) associated with housing the bacterium. This may represent another advantageous factor that facilitates *Spiroplasma's* ability to sustain infections within the fly across multiple generations.

Gff is the prominent vector of pathogenic African trypanosomes in east and central Africa, and reducing the size of fly populations is currently the most effective method for controlling parasite transmission. One means of accomplishing this is through the use of sterile insect technique (SIT), which involves the sequential release of a large number of sterilized males into a target environment [58,59]. Because of their large number, these infertile males outcompete wild males for resident female mates, and the population size drops significantly, or the population is eliminated [60,61]. Knowledge gained from this study, in conjunction with that from previous studies, will have a significant impact on tsetse rearing efficiency for the SIT programs. Specifically, decreased fecundity can have a negative impact on the success of colony maintenance. However, enhanced resistance to trypanosomes is advantageous for SIT applications as the released males, which also feed exclusively on vertebrate blood, have the potential to serve as vectors of disease. Thus, *Spiroplasma* has the potential to reduce the impact of this shortcoming by inducing a trypanosome refractory phenotype in released males such that they would be less likely vector disease, thus increasing the overall safety and efficacy of the control program.

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% [15]. Reproductive tissues (may have been contaminated with milk gland tissue, which is composed of an extensive tubular network that is difficult to separate from female 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 collected tissues by performing PCR using primers that amplify sex specific genes (see S4 Table in S1 Appendix for gene names 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 [62,63]. 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 [64]. The function 'htseq-count' in HTSeq v0.11.2 [65] 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 [66]. We first developed a DESeq2 model to predict differential gene expression as a function of sex and *Spiroplasma*-infection status, and the interaction between sex and *Spiroplasma*-infection status. We then extracted the $\log_2$ fold changes ($\log_2$FCs) for each gene with false discovery rate (FDR) adjusted *p*-values [24]. Finally, we defined genes as DE if the $\log_2$ fold change ($\log_2$FC) was significantly different from 0, and we only considered genes with adjusted *p*-values as expressed for downstream analysis.

We next conducted a PC and HC analysis to analyze normalized expression count data from the DESeq2 program. We used a regularized log transformation of the normalized data created by "rlog" function DESeq2 [66]. For the HC by sample-to-sample distance, we used

"pheatmap" function in R. We measured the expression abundance as the sum of TPM that were calculated using the TPM Calculator [67].

Because the function of genes in *Gff are* not well annotated, we performed 'tblastx' in CLC Genomics Workbench (CLC Bio, Cambridge, MA) with *Gff* transcript to search for predicted protein sequences bearing the closest homology to those from *G. m. morsitans*, *Musca domestica*, and *Drosophila melanogaster*. Using gene-associated GO terms [31], we performed enrichment analysis of gene ontology (GO) terms with the topGO R package [68].

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

Gff pupae were obtained from Joint FAO/IAEA IPCL insectary in Seibersdorf, Austria and reared at Yale University insectary at 26˚C with 70–80% relative humidity and a 12 hr light: dark photo phase. Flies received defibrinated bovine blood every 48 hr through an artificial membrane feeding system [69].

Spiroplasma infection status of all lab-reared female, male and pupal *Gff* was determined by extracting genomic DNA (gDNA) from individual whole organisms or fly legs using a DNeasy Blood and Tissue kit according to the manufacturer's (Qiagen) protocol. To confirm that intact gDNA was successfully extracted, all samples were subjected to PCR analysis using primers that specifically amplify *Gff tubulin* or microsatellite region *GpCAG* (as controls for genomic DNA quality) and *Spiroplasma 16S rRNA*. All PCR primers used in this study are listed in S4 Table in S1 Appendix. The *Spiroplasma 16S rRNA* locus was amplified by PCR (as described in [14]) using the following parameters: initial denaturation at 94˚C for 5 min; 34 cycles of 94˚C for 45 s, 59˚C for 45 s, and 72˚C for 1 min; and a final extension at 72˚C for 10 min. *Spiroplasma* PCR reactions were carried out in a volume of 25 μl containing 1.5 μl gDNA. PCR products were analyzed on 2% agarose gels, and samples were considered infected with *Spiroplasma* if the expected PCR product of 455 bp was detected.

To confirm the sensitivity of this leg snip PCR assay, and thus eliminate the possibly of false negative outcomes, we randomly selected 10 *Spiroplasma* negative samples and two positive controls (as determined by endpoint PCR described above) and subjected them RT-qPCR using the *Spiroplasma ropB*. All RT-qPCR results were normalized to tsetse's constitutively expressed *pgrp-la* gene (S2 Fig). RT-qPCR confirmed our endpoint PCR data (RT-qPCR output shown in S6 Data).

Impact of *Spiroplasma* infection on *Gff* metabolism and reproductive fitness

Hemolymph triacylglyceride (TAG) assay: Hemolymph (3 μl/fly) was collected (as described in [70]) from two-week-old pregnant female *Gff*, centrifuged (4˚C, 3000xg for 5 minutes) to remove bacterial cells, diluted 1:10 in PBS containing 1.2 μl/ml of 0.2% phenylthiourea (to prevent hemolymph coagulation) and immediately flash frozen in liquid nitrogen. Hemolymph TAG levels were quantified colorimetrically by heating samples to 70˚C for 5 min followed by a 10 min centrifugation 16,000xg. Five μl of supernatant was added to 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 [71]. All *Gff* sample spectra data were compared to that generated from a triolein standard curve (0–50 μg, 10 μg increments; S3 Fig).

Impact on fecundity and endosymbiont density

The effect of *Spiroplasma* infection on *Gff* fecundity was measured by quantifying the length of three gonotrophic cycles (GC) and by weighing pupal offspring. To measure GC length, *Gff*

females were mated as five days old adults and thereafter maintained in individual cages. All females were monitored daily to determine when they deposited larvae, and all deposited larvae were weighed.

To confirm the sensitivity of this leg snip PCR assay, and thus eliminate the possibly of false negative outcomes, we randomly selected 10 *Spiroplasma* negative samples and two positive controls (as determined by endpoint PCR described above) and subjected them RT-qPCR using the *Spiroplasma ropB*. All RT-qPCR results were normalized to tsetse's constitutively expressed *pgrp-la* gene (S2 Fig). RT-qPCR confirmed our endpoint PCR data (RT-qPCR output shown in S6 Data).

Sperm-specific gene expression

Individual five-day old male and female flies (each fed twice) were placed together into tubular cages (height, 12.7 cm; diameter 6 cm) and allowed to mate for two days. Individual mating pairs were subsequently separated, and RNA (using Trizol reagent) was isolated from male and female reproductive tracts. RNA was then DNase treated and reverse transcribed into cDNA (using a Bio-Rad iScript cDNA synthesis kit) by priming the reaction with random hexamers. All males were then screened by PCR as described above to determine their *Spiroplasma* infection status.

Gff sperm-specific dynein intermediate chain (*sdic*; GFUI025244) transcript abundance (measured via RT-qPCR, primers used are listed in S4 Table in S1 Appendix) was used as a proxy measurement of sperm density and sperm fitness in male reproductive tracts and the spermathecae of pregnant females [39,47]. All RT-qPCR results were normalized to tsetse's constitutively expressed *pgrp-la* gene, and relative expression of *sdic* was compared between male reproductive tracts and sperm (in the female spermathecae) that originated from Spi^+ and Spi^- males. All RT-qPCR assays were carried out in duplicate, 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) (S4 Fig), and the amount of sperm transferred was then computed as the mean spermathecal filling values of the spermathecae pairs [72,73].

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-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 [74].

Vertical transmission efficiency

Three separate experiments, performed 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 copulation began, mating pairs were collected and placed into individual cages (4 cm diameter x 6 cm height) where they were kept together for 24 h. Males were subsequently removed (and conserved in absolute ethanol at -20˚C for further analysis), and mated females were pooled together in a larger cage and maintained under normal rearing conditions for 10 days. Finally, pregnant females were again separated into individual cages and maintained under normal conditions over the course of three GCs (~40 days). Individual pupa were collected from each female following each GC and conserved in ethanol at -20˚C for further analysis 24h and 72 h post deposition.

Experiment 3: This experiment was performed the same in the same manner as was experiment 2, with the exception that male flies were disposed of after mating.

Statistical analyses

All statistical analyses were carried out using GraphPad Prism (v.9), Microsoft Excel or RStudio (v.1.2.5033 and v.1.3.1073). All statistical tests used, and statistical significance between treatments, and treatments and controls, are indicated on the figures or in their corresponding legends. All sample sizes are provided in corresponding figure legends or are indicated graphically as points on dot plots. Biological replication implies distinct groups of flies were collected on different days.

Supporting information

S1 Appendix. Supporting tables S1-S5.
(DOCX)

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.
(TIF)

S2 Fig. Relative expression of *peptidoglycan recognition protein la* (*pgrp-la*) in the reproductive tract of females that mated either Gff^{Spi+} or Gff^{Spi-} males. *Pgrp-la* expression in each sample was normalized relative to geometrical mean of tsetse's constitutively expressed *gapdh and β-tubulin* genes. Each dot represents one biological replicate, and bars indicate median values. Statistical significance was determined via students t-test.
(TIF)

S3 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.
(TIF)

S4 Fig. Diagrammatic guide used to visually quantify spermathecal fill in females 24 h post-copulation. Image generated by Dr. Güler Demirbas-Uzel.
(TIF)

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

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

S3 Data. GO enrichment analysis results.
(XLSX)

S4 Data. Gonotrophic cycle (GC) length in relation to maternal *Spiroplasma* infection status.
(XLSX)

S5 Data. *Spiroplasma* vertical transmission data.
(XLSX)

S6 Data. RT-qPCR confirmation of *Spiroplasma* infection status of samples used for vertical transmission experiments. Ten *Spiroplasma* negative samples and two *Spiroplasma* positive samples (as determined by endpoint PCR described in the Materials and Methods) were randomly selected for this analysis (describe in the Materials and Methods, subsection '*Laboratory maintenance of Gff and determination of Spiroplasma infection status*').
(XLSX)

Acknowledgments

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.

Author Contributions

Conceptualization: Jae Hak Son, Brian L. Weiss, Daniela I. Schneider, Kiswend-sida M. Dera, Fabian Gstöttenmayer, Geoffrey M. Attardo, Maria Onyango, Adly M. M. Abd-Alla, Serap Aksoy.

Data curation: Jae Hak Son, Brian L. Weiss, Daniela I. Schneider, Kiswend-sida M. Dera, Fabian Gstöttenmayer, Geoffrey M. Attardo, Maria Onyango, Adly M. M. Abd-Alla, Serap Aksoy.

Formal analysis: Jae Hak Son, Brian L. Weiss, Daniela I. Schneider, Kiswend-sida M. Dera, Fabian Gstöttenmayer, Geoffrey M. Attardo, Maria Onyango, Adly M. M. Abd-Alla, Serap Aksoy.

Funding acquisition: Serap Aksoy.

Investigation: Jae Hak Son, Brian L. Weiss, Daniela I. Schneider, Kiswend-sida M. Dera, Fabian Gstöttenmayer, Robert Opiro, Richard Echodu, Norah P. Saarman, Geoffrey M. Attardo, Maria Onyango, Adly M. M. Abd-Alla, Serap Aksoy.

Methodology: Jae Hak Son, Brian L. Weiss, Daniela I. Schneider, Kiswend-sida M. Dera, Fabian Gstöttenmayer, Geoffrey M. Attardo, Adly M. M. Abd-Alla, Serap Aksoy.

Project administration: Brian L. Weiss, Serap Aksoy.

Resources: Brian L. Weiss, Robert Opiro, Richard Echodu, Norah P. Saarman, Adly M. M. Abd-Alla, Serap Aksoy.

Software: Jae Hak Son, Brian L. Weiss.

Supervision: Brian L. Weiss, Adly M. M. Abd-Alla, Serap Aksoy.

Validation: Jae Hak Son, Brian L. Weiss.

Visualization: Jae Hak Son, Brian L. Weiss.

Writing – original draft: Jae Hak Son, Brian L. Weiss, Adly M. M. Abd-Alla, Serap Aksoy.

Writing – review & editing: Jae Hak Son, Brian L. Weiss, Adly M. M. Abd-Alla, Serap Aksoy.

References

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. https://doi.org/10.1146/annurev-ento-010814-020834 PMID: 25341093
2. Scolari F, Benoit JB, Michalkova V, Aksoy E, Takac P, Abd-Alla AM, et al. The Spermatophore in *Glossina morsitans morsitans*: Insights into Male Contributions to Reproduction. *Sci Rep* 2016, 6:20334. https://doi.org/10.1038/srep20334 PMID: 26847001
3. Pollock JN. Sperm transfer by spermatophores in Glossina austeni Newstead. *Nature* 1970, 225:1063–1064. https://doi.org/10.1038/2251063b0 PMID: 16056910
4. Saunders DS, Dodd CWH. 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, et al. Interpreting Morphological Adaptations Associated with Viviparity in the Tsetse Fly *Glossina morsitans* (Westwood) by Three-

Dimensional Analysis. *Insects* 2020, 11:651. https://doi.org/10.3390/insects11100651 PMID: 32977418

7. Bonomi A, Bassetti F, Gabrieli P, Beadell J, Falchetto M, Scolari F, 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. https://doi.org/10.1371/journal.pntd.0001190 PMID: 21666797
8. Vreysen MJ, Saleh K, Mramba F, Parker A, Feldmann U, Dyck VA, et al. 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. https://doi.org/10.1371/journal.pntd.0002857 PMID: 24874883
9. Kabayo JP. Aiming to eliminate tsetse from Africa. *Trends Parasitol* 2002, 18:473–475. https://doi.org/10.1016/s1471-4922(02)02371-1 PMID: 12473355
10. Bing X, Attardo GM, Vigneron A, Aksoy E, Scolari F, Malacrida A, et al. 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, 284:20170360. https://doi.org/10.1098/rspb.2017.0360 PMID: 28659447
11. Akman L, Yamashita A, Watanabe H, Oshima K, Shiba T, Hattori M, et al. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat Genet* 2002, 32:402–407. https://doi.org/10.1038/ng986 PMID: 12219091
12. Snyder AK, Rio RV. "*Wigglesworthia morsitans*" folate (vitamin B9) biosynthesis contributes to tsetse host fitness. *Appl Environ Microbiol* 2015, 81:5375–5386. https://doi.org/10.1128/AEM.00553-15 PMID: 26025907
13. Rio RV, Symula RE, Wang J, Lohs C, Wu YN, Snyder AK, et al. Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: glossinidae) obligate symbiont *Wigglesworthia*. *mBio* 2012, 3:e00240–11. https://doi.org/10.1128/mBio.00240-11 PMID: 22334516
14. Doudoumis V, Blow F, Saridaki A, Augustinos A, Dyer NA, Goodhead I, 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. https://doi.org/10.1038/s41598-017-04740-3 PMID: 28680117
15. Schneider DI, Saarman N, Onyango MG, Hyseni C, Opiro R, Echodu R, 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. https://doi.org/10.1371/journal.pntd.0007340 PMID: 31369548
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. https://doi.org/10.1111/jeb.12260 PMID: 24118386
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. https://doi.org/10.1073/pnas.1518648113 PMID: 26712000
18. Ballinger MJ, Perlman SJ. The defensive *Spiroplasma*. *Curr Opin Insect Sci* 2019, 32:36–41. https://doi.org/10.1016/j.cois.2018.10.004 PMID: 31113629
19. Perlmutter JI, Bordenstein SR. Microorganisms in the reproductive tissues of arthropods. *Nat Rev Microbiol* 2020, 18:97–111. https://doi.org/10.1038/s41579-019-0309-z PMID: 31907461
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. https://doi.org/10.7554/eLife.02964 PMID: 25027439
21. Hurst GD, Bandi C, Sacchi L, Cochrane AG, Bertrand D, Karaca I, et al. *Adonia variegata* (Coleoptera: Coccinellidae) bears maternally inherited flavobacteria that kill males only. *Parasitology* 1999, 118 (Pt 2):125–134. https://doi.org/10.1017/s0031182098003655 PMID: 10028525
22. Harumoto T, Lemaitre B. Male-killing toxin in a bacterial symbiont of *Drosophila*. *Nature* 2018, 557:252–255. https://doi.org/10.1038/s41586-018-0086-2 PMID: 29720654
23. Harumoto T, Fukatsu T, Lemaitre B. Common and unique strategies of male killing evolved in two distinct *Drosophila* symbionts. *Proc Biol Sci* 2018, 285:20172167. https://doi.org/10.1098/rspb.2017.2167 PMID: 29563258
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. https://doi.org/10.1038/ncomms12781 PMID: 27650264

27. Ellegren H, Parsch J. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* 2007, 8:689–698. https://doi.org/10.1038/nrg2167 PMID: 17680007

28. Mank JE. The transcriptional architecture of phenotypic dimorphism. *Nat Ecol Evol* 2017, 1:6. https://doi.org/10.1038/s41559-016-0006 PMID: 28812569

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. https://doi.org/10.1111/j.1462-5822.2011.01627.x PMID: 21740495

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. https://doi.org/10.1111/imb.12340 PMID: 28833767

31. Attardo GM, Abd-Alla AMM, Acosta-Serrano A, Allen JE, Bateta R, Benoit JB, et al. Comparative genomic analysis of six *Glossina* genomes, vectors of African trypanosomes. *Genome Biol* 2019, 20:187. https://doi.org/10.1186/s13059-019-1768-2 PMID: 31477173

32. Nogge G. Sterility in tsetse flies (*Glossina morsitans* Westwood) caused by loss of symbionts. *Experientia* 1976, 32:995–996. https://doi.org/10.1007/BF01933932 PMID: 986317

33. Hill P, Saunders DS, Campbell JA. Letter: The production of "symbiont-free" *Glossina morsitans* and an associated loss of female fertility. *Trans R Soc Trop Med Hyg* 1973, 67:727–728. https://doi.org/10.1016/0035-9203(73)90051-5 PMID: 4798390

34. Michalkova V, Benoit JB, Weiss BL, Attardo GM, Aksoy S. Vitamin B6 generated by obligate symbionts is critical for maintaining proline homeostasis and fecundity in tsetse flies. *Appl Environ Microbiol* 2014, 80:5844–5853. https://doi.org/10.1128/AEM.01150-14 PMID: 25038091

35. 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. https://doi.org/10.1016/0010-406x(69)92169-0 PMID: 4310397

36. Attardo GM, Benoit JB, Michalkova V, Yang G, Roller L, Bohova J, et al. Analysis of lipolysis underlying lactation in the tsetse fly, *Glossina morsitans*. *Insect Biochem Mol Biol* 2012, 42:360–370. https://doi.org/10.1016/j.ibmb.2012.01.007 PMID: 22509523

37. 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. https://doi.org/10.1016/j.jinsphys.2011.08.009 PMID: 21875592

38. Sepil I, Hopkins BR, Dean R, Bath E, Friedman S, Swanson B, 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. https://doi.org/10.1073/pnas.2009053117 PMID: 32611817

39. 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. https://doi.org/10.1038/25126 PMID: 9859991

40. Yeh SD, Do T, Chan C, Cordova A, Carranza F, Yamamoto EA, 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. https://doi.org/10.1073/pnas.1121327109 PMID: 22308475

41. Weiss BL, Maltz MA, Vigneron A, Wu Y, Walter KS, O'Neill MB, et al. Colonization of the tsetse fly midgut with commensal *Kosakonia cowanii* Zambiae inhibits trypanosome infection establishment. *PLoS Pathog* 2019, 15:e1007470. https://doi.org/10.1371/journal.ppat.1007470 PMID: 30817773

42. Jayaswal V, Jimenez J, Magie R, Nguyen K, Clifton B, Yeh S, et al. A species-specific multigene family mediates differential sperm displacement in *Drosophila melanogaster*. *Evolution* 2018, 72:399–403. https://doi.org/10.1111/evo.13417 PMID: 29315521

43. Ingleby FC, Flis I, Morrow EH. Sex-biased gene expression and sexual conflict throughout development. *Cold Spring Harb Perspect Biol* 2014, 7:a017632. https://doi.org/10.1101/cshperspect.a017632 PMID: 25376837

44. Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, et al. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* 2003, 299:697–700. https://doi.org/10.1126/science.1079190 PMID: 12511656

45. Yang G, Attardo GM, Lohs C, Aksoy S. Molecular characterization of two novel milk proteins in the tsetse fly (*Glossina morsitans morsitans*). *Insect Mol Biol* 2010, 19:253–262. https://doi.org/10.1111/j.1365-2583.2009.00987.x PMID: 20136662

46. Hu C, Rio RV, Medlock J, Haines LR, Nayduch D, Savage AF, 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. https://doi.org/10.1371/journal.pntd.0000192 PMID: 18335067

47. 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. https://doi.org/10.4161/cib.21136 PMID: 23181161

48. Civetta A, Ranz JM. Genetic factors influencing sperm competition. *Front Genet* 2019, 10:820. https://doi.org/10.3389/fgene.2019.00820 PMID: 31572439

49. Singh SR, Singh BN, Hoenigsberg HF. Female remating, sperm competition and sexual selection in *Drosophila*. *Genet Mol Res* 2002, 1:178–215. PMID: 14963827

50. Shropshire JD, Leigh B, Bordenstein SR. Symbiont-mediated cytoplasmic incompatibility: what have we learned in 50 years? *Elife* 2020, 9:e61989. https://doi.org/10.7554/eLife.61989 PMID: 32975515

51. 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. https://doi.org/10.1111/evo.13740 PMID: 30990223

52. Rio RVM, Jozwick AKS, Savage AF, Sabet A, Vigneron A, Wu Y, et al. Mutualist-provisioned resources impact vector competency. *mBio* 2019, 10:e00018–19. https://doi.org/10.1128/mBio.00018-19 PMID: 31164458

53. 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. https://doi.org/10.1098/rspb.2010.0364 PMID: 20356887

54. Snyder AK, Rio RV. Interwoven biology of the tsetse holobiont. *J Bacteriol* 2013, 195:4322–4330. https://doi.org/10.1128/JB.00487-13 PMID: 23836873

55. McCutcheon JP, Boyd BM, Dale C. The life of an insect endosymbiont from the cradle to the grave. *Curr Biol* 2019, 29:R485–R495. https://doi.org/10.1016/j.cub.2019.03.032 PMID: 31163163

56. 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:e01006–16. https://doi.org/10.1128/mBio.01006-16 PMID: 27406568

57. Sudakaran S, Kost C, Kaltenpoth M. Symbiont acquisition and replacement as a source of ecological innovation. *Trends Microbiol* 2017, 25:375–390. https://doi.org/10.1016/j.tim.2017.02.014 PMID: 28336178

58. Percoma L, Sow A, Pagabeleguem S, Dicko AH, Serdebeogo O, Ouedraogo M, et al. Impact of an integrated control campaign on tsetse populations in Burkina Faso. *Parasit Vectors* 2018, 11:270. https://doi.org/10.1186/s13071-017-2609-3 PMID: 29703229

59. 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. https://doi.org/10.1016/j.jip.2012.07.026 PMID: 22878217

60. 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. https://doi.org/10.1016/j.cois.2015.05.011 PMID: 29588003

61. McGraw EA, O'Neill S. Beyond insecticides: new thinking on an ancient problem. *Nat Rev Microbiol* 2013, 11:181–193. https://doi.org/10.1038/nrmicro2968 PMID: 23411863

62. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 2015, 12:357–360. https://doi.org/10.1038/nmeth.3317 PMID: 25751142

63. 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. https://doi.org/10.1038/s41587-019-0201-4 PMID: 31375807

64. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, et al. The UCC genome browser database: update 2006. *Nucleic Acids Res* 2006, 34:D590–598.

65. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015, 31:166–169. https://doi.org/10.1093/bioinformatics/btu638 PMID: 25260700

66. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014, 15:550. https://doi.org/10.1186/s13059-014-0550-8 PMID: 25516281

67. 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. https://doi.org/10.1093/bioinformatics/bty896 PMID: 30379987

68. Alexa A, Rahnenfuhrer, J. topGO: Enrichrichment analysis for gene ontology. R package version 2.42.0. Edited by; 2020.

69. Moloo SK. An artificial feeding technique for *Glossina*. *Parasitology* 1971, 63:507–512. https://doi.org/10.1017/s0031182000080021 PMID: 5139030

70. Benoit JB, Vigneron A, Broderick A, Wu Y, Sun JS, Carlson JR, et al. Symbiont-induced odorant binding proteins mediate insect host hematopoiesis. *Elife* 2017, 6:e19535. https://doi.org/10.7554/eLife.19535 PMID: 28079523

71. Agbu P, Cassidy JJ, Braverman J, Jacobson A, Carthew RW. MicroRNA miR-7 Regulates Secretion of Insulin-Like Peptides. *Endocrinology* 2020, 161:bqz040. https://doi.org/10.1210/endocr/bqz040 PMID: 31875904

72. Mutika GN, Kabore I, Parker AG, Vreysen MJ. Storage of male *Glossina palpalis gambiensis* pupae at low temperature: effect on emergence, mating and survival. *Parasit Vectors* 2014, 7:465. https://doi.org/10.1186/s13071-014-0465-y PMID: 25287653

73. Mutika GN, Marin C, Parker AG, Boucias DG, Vreysen MJ, Abd-Alla AM. Impact of salivary gland hypertrophy virus infection on the mating success of male *Glossina pallidipes*: consequences for the sterile insect technique. *PLoS One* 2012, 7:e42188. https://doi.org/10.1371/journal.pone.0042188 PMID: 22912687

74. Hansen JN, Rassmann S, Jikeli JF, Wachten D. SpermQ-A simple analysis software to comprehensively study flagellar beating and sperm steering. *Cells* 2018, 8:10. https://doi.org/10.3390/cells8010010 PMID: 30587820

75. Mangiafico S. Rcompanion: functions to support extension education program evaluation. Edited by; 2020. vol R package version 2.3.26.]

76. Ogle DH, Wheeler, P., Dinno, A. FSA: fisheries stock analysis. R package version 0.8.32. Edited by; 2021.

Chapter 5: *In vitro* cultivation and genomic insights into the *Spiroplasma* symbiont of *Glossina fuscipes fuscipes*

DRAFT: *In vitro* cultivation and genomic insights into the *Spiroplasma* symbiont of *Glossina fuscipes fuscipes*

Daniel J. Bruzzese[a*], Fabian Gstöttenmayer[b*], Brian Weiss[a], Robert L. Mach[c], Serap Aksoy[a] and Adly M. M. Abdalla[b#]

*equally contributing first authors

[a] Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, Connecticut, United States of America.

[b] Insect Pest Control Laboratory, Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture, A-1400 Vienna, Austria.

[c] Institute of Chemical, Environmental, and Bioscience Engineering, Research Area Biochemical Technology, Vienna University of Technology, Gumpendorfer Straße 1a, 1060 Vienna, Austria.

[#]Corresponding Author

Adly M. M. Abdalla,

Insect Pest Control Laboratory,

Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture, International Atomic Energy Agency, P.O. Box 100, 1400 Vienna, Austria. Tel: +43 1 2600 28425, Fax: +43 1 2600 28447.

Email: A.M.M.Abd-Alla@iaea.org

Abstract

Tsetse flies (*Glossina* spp.) are cyclic vectors of African trypanosomes, unicellular parasites of the genus *Trypanosoma*, causing the diseases Human African Trypanosomosis (HAT), also called sleeping sickness in humans and African Animal Trypanosomosis (AAT) or nagana in animals. These diseases pose a severe threat to the public health and socioeconomic development of affected countries in sub-Saharan Africa. The reproductive biology and vector competence may be influenced by the complex interplay between tsetse flies and their bacterial symbionts, namely *Wigglesworthia, Sodalis, Wolbachia* and *Spiroplasma*. The latter was recently identified in wild and laboratory populations of the tsetse fly species *Glossina fuscipes fuscipes* (*Gff*) and *Glossina tachinoides* (*Gt*). *Spiroplasma* is a wall-less helical bacterium within the class of Mollicutes, known to infect various plant and arthropod species. In the tsetse fly, it was found to confer a protective effect against trypanosomes in laboratory conditions, however also negatively impacts the reproductive and metabolic fitness in the HAT vector, *Glossina fuscipes fuscipes* (*Gff*). To further study this bacterium and the potential interactions with its host and trypanosomes, we present results on the successful *in vitro* cultivation and genome sequencing of *Spiroplasma* from *Gff*, herein named *s*Gff. The cultured *s*Gff remained infectious to naïve *Gff* and resulted in reduced survival rates of injected flies. Genome sequencing and comparative genomic studies on *Spiroplasma* genomes from a colony fly, field fly and *in vitro* culture revealed structural variation between colony and cultured *Spiroplasma* as indicated by a duplicated and translocated 20 Kb phage sequence in the cultured *Spiroplasma* genome, as well as several notable differences between the *Spiroplasma* genomes of field fly and colony fly. Phylogenomic analysis placed the three genomes in a novel *Spiroplasma* clade that is closely related to the *S. poulsonii* clade.

The findings from the present study enhances the understanding of *Spiroplasma* on a genomic level and highlights the core metabolic genes and putative genes involved in the interaction with the host and trypanosomes that should be subject of future studies to further uncover the relationships between the host *Gff*, *Spiroplasma* and trypanosomes.

Introduction

Tsetse flies (*Glossina* spp.) are distributed across sub-Saharan Africa, covering an extensive range of approximately 10 million km^2 and 37 countries (Moloo, 1993). These strictly blood-feeding insects represent vectors of African trypanosomes, protozoan parasites of the genus *Trypanosoma*, which cause Human African Trypanosomosis (HAT), known as sleeping sickness in humans, and African Animal Trypanosomosis (AAT), commonly referred to as nagana in animals (Jordan, 1976; Aksoy et al., 2003). The life cycle of trypanosomes involves several stages, starting from a tsetse fly ingesting blood containing trypanosomes from an infected human or animal, development in the midgut, migration to the salivary glands and subsequent transmission to a new host during a fly's blood meal (Jordan, 1976; Centers for Disease Control and Prevention, 2019). Despite extensive work, there is currently no vaccine available to prevent these diseases (La Greca and Magez, 2011) and medication can pose severe side effects (Venturelli et al., 2022). As a result, vector control is an appropriate strategy to reduce tsetse fly populations and consequently lower the incidence of the HAT and AAT, thereby improving public health and economic development in affected countries (Allsopp, 2001; Vreysen et al., 2013). Conventional methods of tsetse fly vector control such as sequential insecticide spraying, traps, targets and treating cattle with insecticidal formulations have been effectively implemented in the framework of an area-wide integrated pest management (AW-IPM) approach, utilizing the advantages of each strategy in combination (Vreysen, 2001). A vector control strategy that is species-specific and considered sustainable is the Sterile Insect Technique (SIT), which consists of mass-rearing of the target insects, sex-separation and irradiation of males to render them sterile, before releasing them in the target area to compete with wild males to mate with wild virgin females. As a result, these wild females do not produce offspring, which leads to a decline and eventually the removal of the target population (Dyck et al., 2021). There have been several outbreaks of HAT and AAT in the past decades, but the long-lasting, sustained vector control efforts of international non-governmental organisations (NGOs) and governments have led to a considerable reduction of HAT cases, remaining at numbers below 1000 cases every year since 2018 (Franco et al., 2022; World Health Organization, 2023). Nonetheless, it is

still important to note that 3 million people are exposed to a medium to high risk of infection (Simarro et al., 2012).

The effectiveness of mass-rearing activities for SIT may furthermore be linked to symbiotic bacteria living within the tsetse fly host that can significantly influence their nutritional state, reproduction and vector competence. Tsetse flies host several different bacteria, namely *Wigglesworthia*, *Sodalis*, *Wolbachia* and *Spiroplasma,* ranging from mutualistic to parasitic relationships depending on the bacterium (Attardo et al., 2020). Among these symbionts harbored by tsetse flies, *Spiroplasma*, a wall-less bacterium within the class of Mollicutes was relatively recently identified in wild and laboratory populations of *Glossina fuscipes fuscipes* (*Gff*) (Doudoumis et al., 2017). *Spiroplasma* is known to infect various plant and insect species, exhibiting complex interactions that can result in beneficial or detrimental effects: in the case of *Drosophila hydei* and *Drosophila neotestacea*, the bacterium confers protection against parasitic wasps and nematodes (Xie et al., 2010; Ballinger et al., 2019). However, in other arthropod species, *Spiroplasma* is considered a reproductive parasite, selectively killing male embryos and therefore skewing the sex ratio of affected populations towards females (Harumoto and Lemaitre, 2018). In tsetse flies, the impact of *Spiroplasma* on *Gff* has been subject of recent studies, suggesting a protective effect against trypanosome infection (Schneider et al., 2019), however the exact mode of interaction is unknown. Infection with *Spiroplasma* has also been associated with negative effects on the reproductive fitness of *Gff* in laboratory settings, resulting in extended gonotrophic cycles and reduced lipid levels in the hemolymph of infected females. In infected males, *Spiroplasma* infection significantly reduces sperm motility and thus raises concerns about the males' competitiveness. Transcriptomic analysis of female and male *Gff* in their infected and uninfected state indicated sex-biased gene expression and differential gene expression in infected vs uninfected individuals, suggesting a list of genes that may be involved in the interaction with the host or trypanosomes (Son et al., 2021). However, there remains a gap in knowledge about further involved genes, largely due to the absence of a complete genome sequence. In the present study, we aim to provide further insights on *Spiroplasma* from *Gff* through *in vitro* cultivation of the bacterium based on available cultivation protocols for *S. poulsonii*, characterization of its growth kinetics, genome

sequencing of *Spiroplasma* from three sources (colony fly, field fly, culture) and comparative genomic analyses. These findings enhance the understanding of the interactions between *Spiroplasma* and *Gff* by providing novel genomic insights and resources, as well as enabling future studies by describing the successful *in vitro* cultivation system.

Methods

Initiation of *Spiroplasma in vitro* culture

Initiation of *Spiroplasma in vitro* culture followed the protocol of (Masson et al., 2018), optimized for the *in vitro* cultivation of *Spiroplasma poulsonii* MSRO. Briefly, Barbour-Stoenner-Kelly H (BSK-H) medium (Bio&Sell, Germany) designed for Borrelia was supplemented with rabbit serum, lipids, antibiotics and amino acids (see: Appendix 1). Culture inoculation and maintenance steps were all performed under sterile conditions.

Flies used for the initiation of the culture were derived from a high *Spiroplasma* prevalence *Glossina fuscipes fuscipes* (*Gff*) laboratory line established and described in (Dera et al., 2024). Fly hemolymph was extracted from 10 surface-sterilized females and 10 males aged 7 to 21 days by removal of one front leg and aspiration of the exposed hemolymph droplet with a 10 µl pipette tip. Precultures were initiated in three biological replicates by inoculating 3.2 ml of BSK-H medium supplemented with 5% fly extract with 8 µl of harvested haemolymph and incubated at 25°C for 14 days without agitation, utilizing the CampyGen™ Gas-Pak system (ThermoFisher Scientific) to establish microaerophilic atmospheric conditions. Microscopy was undertaken on a Leica DMi8 inverted fluorescent microscope (Leica Microsystems) with Syto9 live staining (0.025 mM) (ThermoFisher Scientific) to validate the presence and morphology of *Spiroplasma* in the precultures. Additionally, presence was confirmed by performing PCR using *Spiroplasma* 16s rRNA-specific primers (Table S1). Precultures were then pelleted by centrifugation for 40 min at 2000 relative centrifugal force (rcf), resuspended in 3.2 ml BSK-H-spiro medium (BSK-H + 7.5% fly extract, 6% rabbit serum, Penicillin, 7% arginine, 5% lipid

mix) and aliquots of 100 µl stored at -80°C. Cultures were then set up with 100 µl frozen preculture aliquots in 3.2 ml of BSK-H-spiro medium at 25°C in microaerophilic conditions without agitation for the establishment of active long-term cultures and cultures for growth curve analysis. Cultures underwent routine microscopy checks every 10-14 days to validate presence of *Spiroplasma* and to assess bacterial density before passaging. Passaging of bacteria was done every 10-14 days by diluting the culture 1:1 with freshly prepared BSK-H-spiro medium. Every fourth passage until at least passage 12, bacteria were pelleted at 2000 rcf for 40 minutes and media were completely renewed.

In vitro culture growth kinetics

Samples for the growth curve analysis were collected in threefold biological replication from day 0 to day 15 post inoculation of BSK-H-spiro media with preculture aliquots. Sample collection was performed by adding 5 µl of culture to 200 µl ddH_2O in a sterile PCR tube. Samples were stored at -20°C until processing for quantitative PCR. All samples underwent osmotic heat-shock at 95°C for 10 minutes to lyse bacteria before subjecting to quantitative PCR. Reactions were set up in three-fold technical replication, in a total volume of 15 µl with 7.5 µl iQ™ SYBR® Green Supermix (Bio-Rad, Hercules, CA, USA), 5.5 µl PCR-grade water, 0.5 µl *Spiroplasma* 16s_qPCR primer F + R (Table S1), respectively, and 1 µl heat-shocked bacterial culture. Samples were amplified alongside a standard of known concentration and copy numbers to assess primer efficiency and estimate the copy numbers of the culture aliquots. This standard was prepared by purification of *Spiroplasma* 16s qPCR amplicons, measurement of DNA concentration and serial 1:10 dilutions. Cycling conditions comprised of initial denaturation at 95°C for 2 min, followed by 40 cycles of 95°C for 5 sec and 56°C for 30 sec on a Bio-Rad CFX96 Touch Real-Time PCR Detection System (Hercules, CA, USA). Melt curves were inspected and amplicons migrated on a 2% agarose gel to verify specific amplification. Raw data quantification cycle (Cq) values were extracted from the CFX maestro software and data analysis conducted in R Studio (R Core Team, 2021a; R Studio Team, 2021). Briefly, the mean value of the three technical replicates of each sample was calculated and the expression of each sample was

normalized to its day 0 sample before calculation of fold-change ($2^{(-dCq)}$). Copy numbers of the culture samples were calculated by comparing to the standard curve of known copy numbers and culture doubling time was calculated based on the copy numbers.

Injections of *in vitro Spiroplasma* into naïve *Gff*

Teneral female and male *Gff* were collected from the laboratory colony at FAO/IAEA Insect Pest Control Laboratory (IPCL) in Seibersdorf (Austria). Sampling for a non-destructive *Spiroplasma* infection screening on 190 Gff was undertaken by excision of one of the mesothoracic legs and proceeding for DNA extraction using the ZR Quick-DNA 96 kit (Zymo Research, USA) while keeping the fly specimens in individual cages until injection. Extracted DNA was subjected to PCR amplifying tsetse fly tubulin to confirm DNA extraction and specific 16s rRNA primers for the detection of *Spiroplasma*. Primers and cycling conditions are indicated in Supplementary table 1. PCR products were resolved on 2% agarose gels stained with Ethidium Bromide. *Spiroplasma*-negative flies were retained and split into five groups (n= 18-20 each): high dose, low dose, BSK-H, PBS injected and noninjected control. *Gff* were injected using 29 gauge insulin syringes on a micrometer syringe unit injector (Burkard Scientific Ltd., UK), offered a first blood meal after 3 hrs post injections and then maintained under standard rearing conditions for 14 days. Mortality was recorded daily and 2 female and 1 male samples were collected for qPCR at designated time points 0, 1, 6, 9 and 14 days. Survival data was plotted as Kaplan-Meier survival curves in RStudio (R Core Team, 2021b; R Studio Team, 2021). Relative quantification of *Spiroplasma* titers was assessed via qPCR in threefold technical replication, normalizing to the expression of tubulin and calculation of fold-change ($2^{(-dCt)}$).

Sample collection and nanopore sequencing

Female *Gff* flies were collected from the colony located at the FAO/IAEA Insect Pest Control Laboratory (IPCL) in Seibersdorf, Austria in 2023 and from a field population located at Toloyang

village (3.304659058, 32.37430505) in 2019, a site in Northwestern Uganda known to have a high prevalence of *Spiroplasma*-infected flies (Schneider et al., 2019). Additionally, 6 ml pelleted high density *in vitro Spiroplasma* culture was collected at passage 9 in 2024. For the colony tsetse flies, unmated 10-day old females were starved 48 hours and were snap frozen and stored at -80 °C. High molecular weight (HMW) DNA was extracted from single colony flies using the Zymo Quick-DNA™ HMW MagBead Kit with wide bore pipette tips. For the field tsetse flies, females were collected from traps and were snap frozen in liquid nitrogen, transported on dry ice, and stored long-term at -80 °C. HMW DNA was extracted from individual field flies using the NEB Monarch® Genomic DNA Purification Kit, following the manufacturer's suggested modifications for insects. For the *Spiroplasma* culture, HMW DNA was extracted using the Qiagen PureGene kit following standard protocols.

All HMW DNA was checked for the presence of *Spiroplasma* PCR with primers outlined in Schneider et al (2019). For *Spiroplasma* positive samples, extraction quality was evaluated using a NanoDrop® and the Agilent TapeStation, favoring samples with NanoDrop 260/280 values closest to 1.8 and 260/230 > 2.0, and TapeStation DNA integrity values closest to 10 with the longest read lengths possible. The best *Spiroplasma* samples from the colony, the field, and the *Spiroplasma* culture were selected for sequencing. The field sample extracted with the NEB column kit had a high proportion of short DNA fragments. To mitigate this, the PacBio short read eliminator kit was used to deplete reads shorter than 10 kb that would have negatively impacted the ONT sequencing run.

ONT LSK14 libraries were prepared for each sample following standard protocols except for increasing the amount of input DNA to between 2000 ng and 3000 ng to produce more concentrated libraries, enough for multiple flow cell loads. Each library was sequenced on a separate 10.4.1 minION flow cell. To improve throughput, flow cells were washed using the ONT wash kit v4 and reloaded with fresh library 3X times over the 48 hours sequencing run.

Spiroplasma genome assembly

Generated POD5 signal data files were basecalled with Dorado v0.5.3 using the duplex pipeline. Simplex donor reads (reads that were merged into duplex reads) were dropped with Samtools v1.18, and sequencing adapters were trimmed using Dorado. The produced FASTQ files were quality filtered with nanoq v0.10.0 to have a minimum quality of q10 and minimum 1000 bp read length.

Due to the use of different extraction methods and variance between the sequencing runs, the colony fly sample with the greatest N50 read length and high sequencing depth was selected to de novo assemble the *Spiroplasma* reference assembly. Flye v2.9.3 with a minimum overlap set to 4000 bp assembled the tsetse and symbiont genomes simultaneously. BlobTools2, BLAST+ v2.15.0, and minimap2 v2.25 were used to filter the assembly and only export reads that mapped to *Spiroplasma* sequences. These *Spiroplasma* reads were then reassembled using multiple assemblers (Flye, Canuv2.2, Miniasm v0.3 and Minipolish v0.1.3, and raven v1.8.3) and varied parameters to generate a consensus assembly using the Trycycler v0.5.4 consensus tool and manual curation. The final colony assembly was polished with medaka v1.11.3.

To generate the field fly *Spiroplasma* assembly, reads were mapped to the colony consensus assembly with minimap2 and unmapped reads were dropped with Samtools. The mapped reads were assembled with flye and polished with medaka to generate a final field *Spiroplasma* assembly. To generate the culture *Spiroplasma* assembly, quality filtered reads were assembled with flye and polished with medaka. Plasmid consensus sequences were obtained by mapping all reads to the plasmids from the colony fly *Spiroplasma* assembly with minimap2 v.2.24. Additionally, 300 ng DNA was sent to Eurofins Genomics for Illumina sequencing on a NovaSeq platform to yield approximately 8.5M high quality paired-end reads of 2x150 bp to polish the assembly with short read data using Polypolish v.0.6.0. Pl

All completed circular genomes and plasmids were reoriented to start with the *dnaA* or *repA* respectively with Dnaapler v0.7.0. The Assembled genomes were evaluated for completeness with BUSCO v5.7.0 (Simão et al., 2015) and annotated with batka v1.9.2.

Comparative genomic analysis was undertaken by identifying genome synteny, number of SNPs and indels using mauve v2.4.0, minimap2 v2.26, SYRI v1.6 and PlotSR v1.4. Shared and unique genes of the three genomes and an outgroup genome (sTU-14) were visualized on a venn diagram constructed with Intervene v0.6.5.

Spiroplasma phylogenomics

We constructed phylogenies involving *Spiroplasma* reference genomes from the *Spiroplasma melliferum*, *Spiroplasma citri*, *Spiroplasma phoeniceum*, and *Spiroplasma poulsonii* clades. Bakta annotated all the reference genomes and Panaroo v1.5 (Tonkin-Hill et al., 2020) identified sets of orthologous single-copy genes. The concatenated genes were aligned with MAFFT and RAxML-NG (Kozlov et al., 2019) constructed phylogenetic trees using the GTR+G substitution model, 50 random and 50 parsimony-based starting trees, with 1000 bootstrap replicates.

Results

Successful establishment of *s*Gff cultures

Spiroplasma in vitro cultures from *Gff* hemolymph were successfully established at 25°C under microaerophilic conditions utilizing the existing cultivation protocol and the supplemented BSK-H medium for *S. poulsonii*. Microscopic examination with Syto9-stained cultures revealed the characteristic helical morphology of *Spiroplasma* and served for density assessments of cultures prior

to passaging (Fig 1A and 1B).

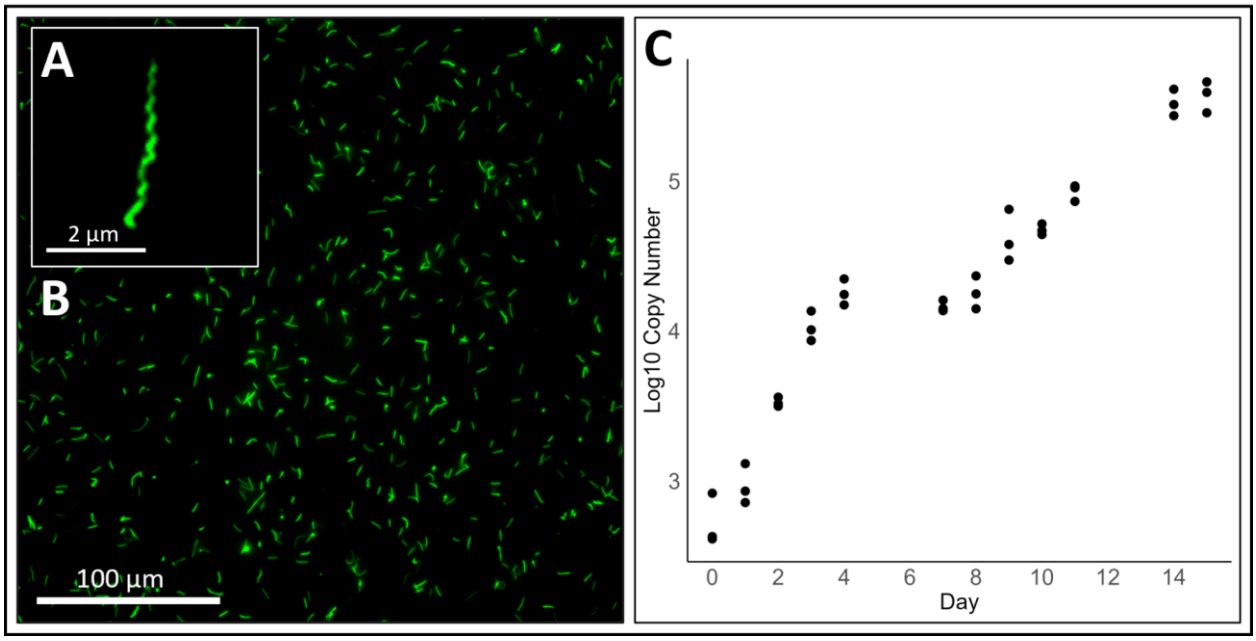

Figure 1. Successful culturing of *s*Gff.

A) Fluorescent microscopy of *in vitro Spiroplasma* stained with Syto9 and imaged with Leica DMi8 inverted microscope on FITC channel. Close-up highlighting morphology of *Spiroplasma*. B) High density *Spiroplasma* culture at passage 9. C) Growth kinetics of *Spiroplasma* in supplemented BSK-H medium at 25°C under microaerophilic conditions, as assessed with quantitative PCR in threefold technical- and biological replication.

The growth curve of cultures inoculated with frozen preculture aliquots indicated two phases of growth separated by a plateau phase. The first phase of growth was characterized by exponential growth until day 4, followed by a plateau of bacterial density until day 7. Subsequently, exponential growth resumed until reaching another plateau at day 15 (Fig 1C). The calculated doubling time of *Spiroplasma in vitro* cultures based on the growth curve was approximately 40 hours. Cultures demonstrated long-term viability in the described culture conditions (> 6 months) with routine passaging every 10-14 days. Furthermore, 1:1 diluted cultures with fresh medium reached high density within 14 days.

To test if the cultivated *Spiroplasma* remains infectious and can replicate in vivo, we injected naïve *Gff* flies with high and low doses of *s*Gff. The Kaplan-Meier survival curves demonstrate differences in the survival rates across the different experimental groups over a period of 14 days (Fig S1). The noninjected control and PBS injected group showed the highest survival rates, closely followed by the BSK-H injected group. *Gff* injected with a high dose of cultivated *Spiroplasma* exhibited a reduction in survival rate while *Gff* injected with a low dose showed intermediate survival rates, with a drop in survival on day 8. Relative quantification of *Spiroplasma* titers in injected flies shows a sex-specific and injected dose-dependent increase of *Spiroplasma* in naïve *Gff* (Fig S2 A, B).

*s*Gff complete genome assembly

We assembled three closed *s*Gff genomes taken from a whole colony fly (*s*Gff-colony), from culture (*s*Gff-culture), and from a fly specimen collected in the field (*s*Gff-field). The assemblies were approximately the same size (~1.4 -1.5 Mb) and had BUSCO scores (> 97.3%) consistent with other complete *Spiroplasma* genomes. The *s*Gff-field only had three plasmids compared to the four plasmids found in the *s*Gff-colony and *s*Gff-cult assemblies. We identified 1869 coding sequences in the genome of *s*Gff_colony, among these are core metabolic genes of *Spiroplasma* such as dnaX, ruvX, uvrX, gyrX, fruA, ftsZ, treA, atpD and recD that are essential for the base metabolism of the bacterium for processes like DNA replication and recombination, homologous recombination, sugar and carbon metabolism and energy metabolism. Furthermore, the presence of genes encoding for spiralins and ribosome-inactivating proteins (RIPs) underscores that *Spiroplasma* interacts with host cells and potentially trypanosomes. A multitude of mobile genetic elements was identified in the genomes, namely integrated phage sequences, transposable elements and ICE elements that highlight the plasticity of these genomes.

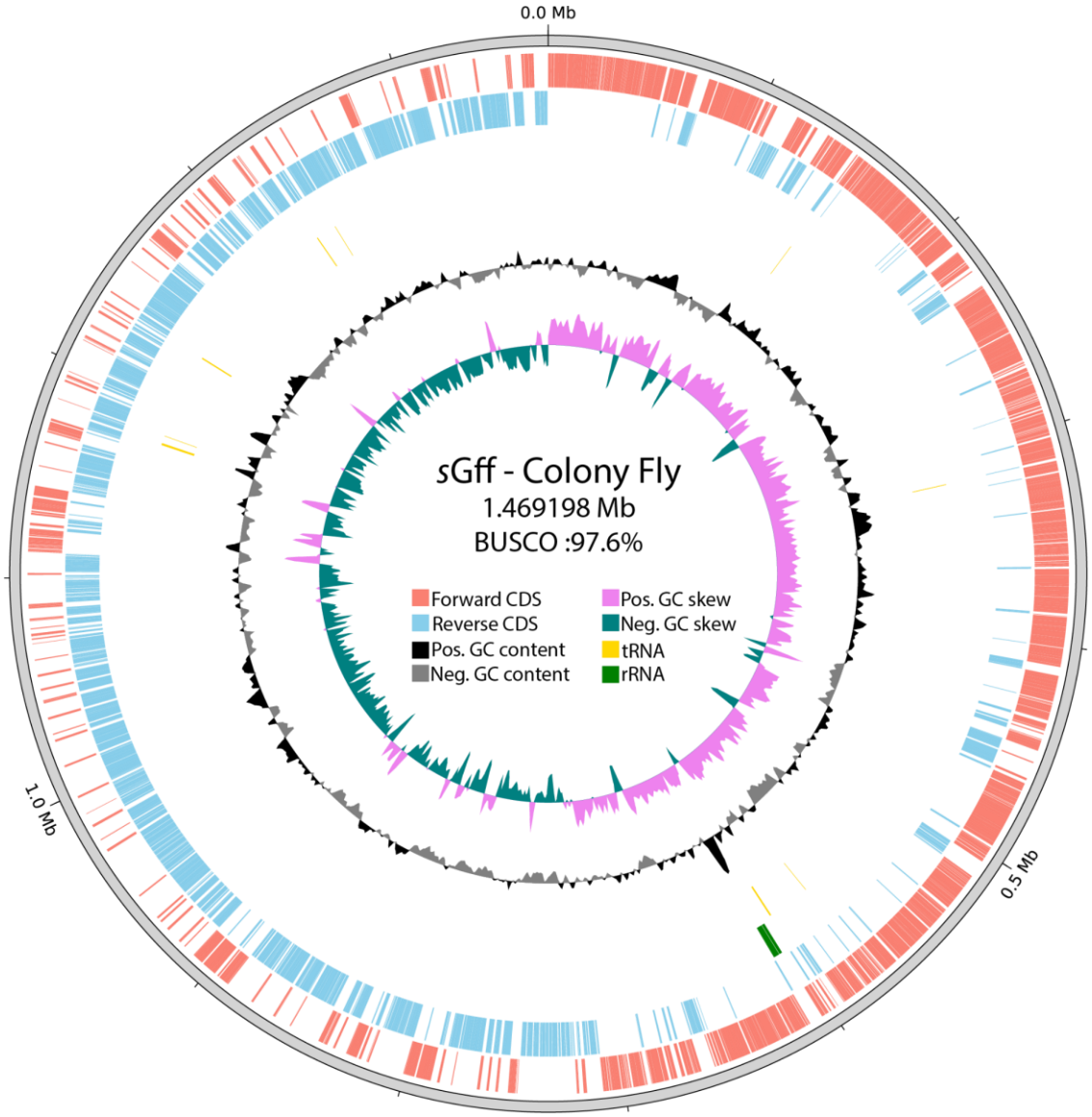

Figure 2. Circos plot for the *s*Gff-colony genome.

*s*Gff is sister to the *S. poulsonii* clade

To place the newly assembled *s*Gff genomes in phylogenetic context, we constructed a RAxML tree based on 302 single-copy orthologous loci shared among the sequenced *s*Gff and *Spiroplasma* reference genomes from multiple clades. We find that *s*Gff and two sister taxa, *s*TU-14 and *s*NBRC_100390, are sister to the *S. poulsonii* clade and constitute their own clade herein coined the *Spiroplasma glossina* clade (Fig 3). A higher resolution phylogeny consisting of *Spiroplasma* from the

poulsonii and *glossina* clades based on 459 single copy shared orthologous genes reconstructed the same relationships but also resolved the *s*Gff polytomy, finding the *s*Gff-field diverged from the *s*Gff-colony and the *s*Gff-cult.

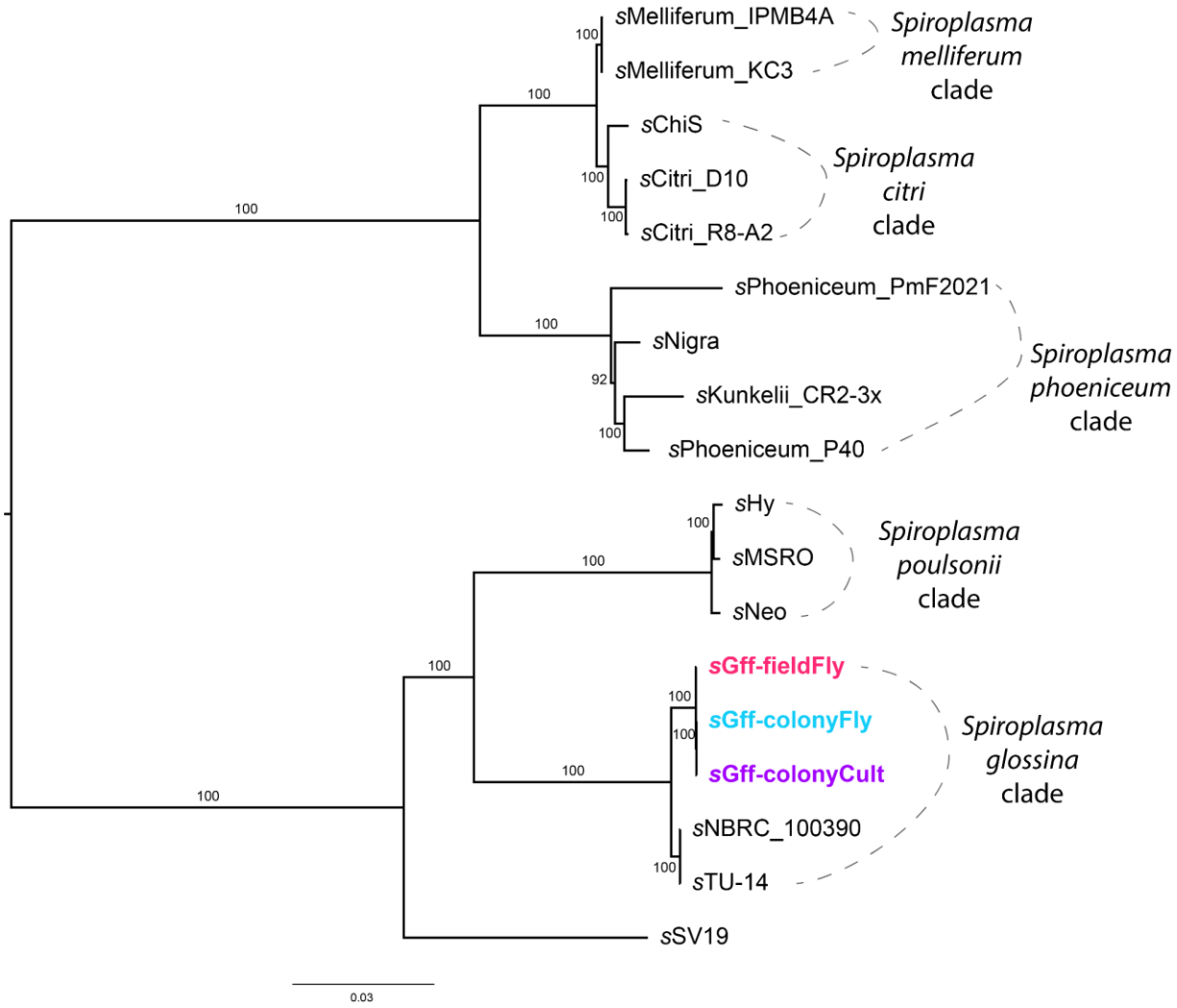

Figure 3. RAxML phylogeny of *s*Gff and representative *Spiroplasma* genomes.

*s*Gff genome comparisons

The comparative analysis between *s*Gff_colony and *s*Gff_cult revealed minimal genetic differences on the level of SNPs and indels. Specifically, there was a difference of 1 SNP, 33 insertions and 14 deletions. However, we identified a 20 Kb inserted sequence in the culture genome which was

found to be a duplicated and translocated phage sequence that is present only once in the *s*Gff_colony genome but twice in the *s*Gff_cult genome. Comparing the *s*Gff_colony and the *s*Gff_field genomes resulted in a higher level of variation, with a total of 557 SNPs, 55 insertions and 88 deletions. We then illustrated the shared and unique genes of the three genomes and an outgroup (sTU-14) with a venn diagram (Fig 4). A total of 840 genes are common across all four genomes, indicating a large core of shared genes. 495 genes are shared among the three *s*Gff genomes but are absent in the outgroup sTU-14. The *s*Gff_field genome exhibits 8 unique genes that are absent in all other genomes and the *s*Gff_colony and *s*Gff_cult genomes have 75 genes in common, that are unique to these two genomes and not found in the other genomes.

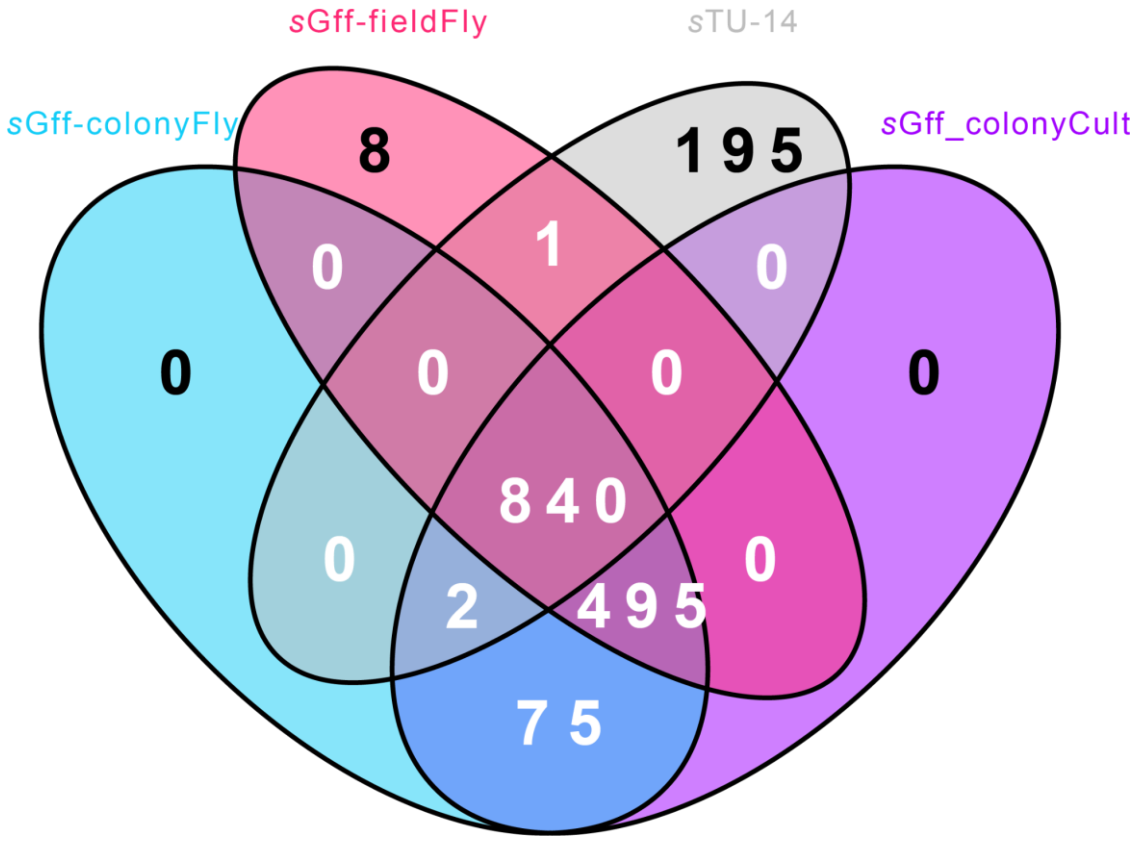

Figure 4. Shared and unique genes for sGff and an outgroup - TU-14.

Discussion

The present study provides a comprehensive genomic analysis of the *Spiroplasma* symbiont of *Glossina fuscipes fuscipes*, involving the characterization of *in vitro*-cultivated *Spiroplasma* based on published cultivation protocols of *S. poulsonii*, genome sequencing and bioinformatic analyses to place the genomes into context with other reference genomes and to identify similarities and differences between these genomes.

The establishment of *Spiroplasma in vitro* cultures demonstrates that the optimized BSK-H medium with supplementations enables the growth of *s*Gff, representing the first successful attempt to cultivate this strain of *Spiroplasma* outside of its host *Gff.* The growth kinetics as measured with quantitative PCR revealed a growth pattern in two exponential phases, separated by a plateau phase. This finding goes in line with published results on cultivation of *S. poulsonii,* the sex-distorting symbiont of *Drosophila*, where it was also found that the growth of this bacterium plateaus at specific bacterial densities and then continues to grow (Masson et al., 2020). The calculated doubling time of approximately 40 hours is similar to the doubling time of cultivated *S. poulsonii* (Masson et al., 2018). These findings on the cultivation of *s*Gff not only confirm the viability and proliferation of the *in vitro* cultures based on available protocols, but also provide a new model for further studies on the physiology and host interactions of *s*Gff, opening the way for more controlled experimental set ups to study the bacterium.

To test if the cultivated *Spiroplasma* remains infectious to its native host *Gff*, we injected uninfected *Gff* with high and low doses of *in vitro Spiroplasma* along with injection controls (BSK-H medium and PBS) and an uninjected control group. The preliminary results indicate reduced survival rates of high-dose injected *Gff* and intermediate survival rates of low-dose injected *Gff* as compared to the injection controls and the uninjected control. These findings indicate that the negative impact of *Spiroplasma* in *Gff* is dependent on the injected dose and leads to reduced survival rates. Furthermore, the preliminary results of the *Spiroplasma* quantification from injected *Gff* on days 0,1,6,9 and 14 highlights a sex-specific and dose-dependent increase of *Spiroplasma* over time, suggesting that the

injected bacteria replicate in naïve *Gff*. In addition, the observed sex-specific differences in *Spiroplasma* titers suggest potential variations in host susceptibility between female and male flies. Overall, these findings demonstrate that the cultivated bacteria remain infectious after long-term cultivation (> 6 months), enabling a multitude of potential future experiments to further study *s*Gff and its interaction with the host *Gff*, as well as studies on the host range of *Spiroplasma* by injections into novel hosts.

In this study we furthermore present high quality genome sequences of *Spiroplasma* strains from different sources (colony fly, culture, field fly), that have not been available until now. The analysis of *s*Gff genomes revealed the core set of bacterial genes that are essential for fundamental processes such as DNA replication, homologous recombination, sugar and carbohydrate metabolism. The presence of certain core genes such as dnaX, ruvX, ftsZ and fruA underscores essential metabolic pathways required for the proliferation of this bacterium. The three sequenced genomes showed a high BUSCO score of > 97.3%, similar to other published *Spiroplasma* reference genomes and indicating a high degree of genome completeness. Additionally, the identification of genes encoding spiralins and ribosome-inactivating proteins (RIPs) suggests specialized roles in the symbiotic relationship between *Spiroplasma* and its tsetse fly host *Gff*. We hypothesize that spiralins are involved in the interaction with host cells, while RIPs may play a role in the interaction with trypanosomes, in a similar manner compared to the protective effect of *Spiroplasma* in other hosts (Garcia-Arraez et al., 2019; Higareda-Alvear et al., 2020). Future studies should focus on these genes of interest that are likely to play a vital part in the interaction with the host and potentially trypanosomes. With the assembly of the complete *s*Gff genomes we provide novel genomic data that can serve as a resource for future studies through further bioinformatic analyses as well as functional studies on the specific genes of interest.

Phylogenomic analysis placed the *s*Gff genomes in close relation to the *S. poulsonii* clade, forming a distinct clade that we herein named the *Spiroplasma glossina* clade. This phylogenomic clustering enhances the understanding of evolutionary relationships of different *Spiroplasma* strains and their most closely related strains. The divergence between *s*Gff_field and *s*Gff_colony strains indicate geographical and environmental influences on *Spiroplasma* evolution, with the field strain exhibiting unique genes that are absent in the other genomes. Interestingly, the *s*Gff_cult genome

exhibited a duplicated and translocated phage sequence of 20 Kb that is found in two copies as compared to the one copy in the *s*Gff_colony and *s*Gff_field genomes. This duplication and translocation event may have occurred during the long-term *in vitro* cultivation process, but future studies should focus on the dynamics between *Spiroplasma* and its phages to further clarify this matter.

Conclusions

The presented findings provide significant advancements in our understanding of *s*Gff, particularly through the demonstration that *Spiroplasma* from *Gff* can be successfully cultivated *in vitro* outside of its host, through genome sequencing and bioinformatic analyses. We identified core metabolic genes, as well as potential symbiosis genes of *s*Gff, that should be subject of future functional studies. The phylogenomic analysis indicated that *s*Gff is closely related to strains of the *S. poulsonii* clade and may exhibit similar modes of interaction with its host. Overall, this study lays the foundation for future studies on the host-symbiont interactions and potential implications on vector competence of *Gff*.

References

Aksoy S, Gibson WC, Lehane MJ. 2003. Interactions between tsetse and trypanosomes with implications for the control of trypanosomiasis. In: Advances in Parasitology. Vol. 53. Academic Press. p. 1–83. https://doi.org/10.1016/S0065-308X(03)53002-0.

Allsopp R. 2001. Options for vector control against trypanosomiasis in Africa. Trends Parasitol. 17(1):15–19. https://doi.org/10.1016/s1471-4922(00)01828-6.

Attardo GM, Scolari F, Malacrida A. 2020. Bacterial Symbionts of Tsetse Flies: Relationships and Functional Interactions Between Tsetse Flies and Their Symbionts. In: Kloc M, editor. Symbiosis: Cellular, Molecular, Medical and Evolutionary Aspects. Vol. 69. Cham: Springer International Publishing. (Results and Problems in Cell Differentiation). p. 497–536. https://doi.org/10.1007/978-3-030-51849-3_19.

Ballinger MJ, Gawryluk RMR, Perlman SJ. 2019. Toxin and Genome Evolution in a *Drosophila* Defensive Symbiosis. Sloan D, editor. Genome Biology and Evolution. 11(1):253–262. https://doi.org/10.1093/gbe/evy272.

Centers for Disease Control and Prevention. 2019 Jun 3. Trypanosomiasis, African. Available from https://www.cdc.gov/dpdx/trypanosomiasisafrican/index.html.

Dera K-SM, Daouda DB, Benewende Aristide Kabore, et al. 2024. *Spiroplasma* infection in a colonized *Glossina fuscipes fuscipes* colony: Impact on mass rearing and the Sterile Insect Technique. Preprint.

Doudoumis V, Blow F, Saridaki A, et al. 2017. Challenging the *Wigglesworthia, Sodalis, Wolbachia* symbiosis dogma in tsetse flies: *Spiroplasma* is present in both laboratory and natural populations. Sci Rep. 7(1):4699. https://doi.org/10.1038/s41598-017-04740-3.

Dyck VA, Hendrichs J, Robinson AS, editors. 2021. Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management. 2nd ed. Boca Raton, FL: CRC Press. xvii+1200 p. https://doi.org/10.1201/9781003035572.

Franco JR, Cecchi G, Priotto G, et al. 2022. Human African trypanosomiasis cases diagnosed in non-endemic countries (2011-2020). PLoS Negl Trop Dis. 16(11):e0010885. https://doi.org/10.1371/journal.pntd.0010885.

Garcia-Arraez MG, Masson F, Escobar JCP, et al. 2019. Functional analysis of RIP toxins from the *Drosophila* endosymbiont *Spiroplasma poulsonii*. BMC Microbiology. 19(1). https://doi.org/10.1186/s12866-019-1410-1.

Harumoto T, Lemaitre B. 2018. Male-killing toxin in a bacterial symbiont of *Drosophila*. Nature. 557(7704):252–255. https://doi.org/10.1038/s41586-018-0086-2.

Higareda-Alvear VM, Mateos M, Quezada DC, et al. 2020. Transcriptomic response during wasp parasitism in the *Drosophila* - *Spiroplasma* interaction. https://doi.org/10.1101/2020.08.27.270462.

Jordan AM. 1976. Tsetse flies as vectors of trypanosomes. Vet.Parasitol. 2:143–152. https://doi.org/10.1016/0304-4017(76)90059-5.

Kozlov AM, Darriba D, Flouri T, et al. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Wren J, editor. Bioinformatics. 35(21):4453–4455. https://doi.org/10.1093/bioinformatics/btz305.

La Greca F, Magez S. 2011. Vaccination against trypanosomiasis: can it be done or is the trypanosome truly the ultimate immune destroyer and escape artist? Hum Vaccin. 7(11):1225–1233. https://doi.org/10.4161/hv.7.11.18203.

Masson F, Calderon Copete S, Schüpfer F, et al. 2018. *In Vitro* Culture of the Insect Endosymbiont *Spiroplasma poulsonii* Highlights Bacterial Genes Involved in Host-Symbiont Interaction. mBio. 9(2):e00024-18. https://doi.org/10.1128/mBio.00024-18.

Masson F, Schüpfer F, Jollivet C, et al. 2020. Transformation of the *Drosophila* Sex-Manipulative Endosymbiont *Spiroplasma poulsonii* and Persisting Hurdles for Functional Genetic Studies. Johnson KN, editor. Appl Environ Microbiol. 86(14):e00835-20. https://doi.org/10.1128/AEM.00835-20.

Moloo SK. 1993. The distribution of *Glossina* species in Africa and their natural hosts. International Journal of Tropical Insect Science. 14(4):511–527. https://doi.org/10.1017/S1742758400014211.

R Core Team. 2021a. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Available from https://www.R-project.org/.

R Core Team. 2021b. R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing.

R Studio Team. 2021. RStudio: Integrated Development for R. Available from http://www.rstudio.com/.

Schneider DI, Saarman N, Onyango MG, et al. 2019. Spatio-temporal distribution of *Spiroplasma* infections in the tsetse fly (*Glossina fuscipes fuscipes*) in northern Uganda. Caljon G, editor. PLOS Neglected Tropical Diseases. 13(8):e0007340. https://doi.org/10.1371/journal.pntd.0007340.

Simão FA, Waterhouse RM, Ioannidis P, et al. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 31(19):3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

Simarro PP, Cecchi G, Franco JR, et al. 2012. Estimating and mapping the population at risk of sleeping sickness. PLoS Negl.Trop.Dis. 6:e1859-. https://doi.org/10.1371/journal.pntd.0001859.

Son JH, Weiss BL, Schneider DI, et al. 2021. Infection with endosymbiotic *Spiroplasma* disrupts tsetse (*Glossina fuscipes fuscipes*) metabolic and reproductive homeostasis. Lemaitre B, editor. PLoS Pathog. 17(9):e1009539. https://doi.org/10.1371/journal.ppat.1009539.

Venturelli A, Tagliazucchi L, Lima C, et al. 2022. Current Treatments to Control African Trypanosomiasis and One Health Perspective. Microorganisms. 10(7):1298. https://doi.org/10.3390/microorganisms10071298.

Vreysen MJ. 2001. Principles of area-wide integrated tsetse fly control using the sterile insect technique. Med Trop (Mars). 61(4–5):397–411.

Vreysen MJB, Seck MT, Sall B, et al. 2013. Tsetse flies: their biology and control using area-wide integrated pest management approaches. J Invertebr Pathol. 112 Suppl:S15-25. https://doi.org/10.1016/j.jip.2012.07.026.

World Health Organization. 2023. Human African trypanosomiasis. Number of reported cases of human African trypanosomiasis (*T. b. gambiense*): 2023. Available from https://apps.who.int/neglected_diseases/ntddata/hat/hat.html.

Xie J, Vilchez I, Mateos M. 2010. *Spiroplasma* Bacteria Enhance Survival of *Drosophila hydei* Attacked by the Parasitic Wasp *Leptopilina heterotoma*. Raine NE, editor. PLoS ONE. 5(8):e12149. https://doi.org/10.1371/journal.pone.0012149.

Supplementary Files

Appendix 1

sGff *Spiroplasma in vitro* cultivation medium.
Adapted from Masson, F., Calderon Copete, S., Schüpfer, F., Garcia-Arraez, G., Lemaitre, B., 2018. *In Vitro* Culture of the Insect Endosymbiont *Spiroplasma poulsonii* Highlights Bacterial Genes Involved in Host-Symbiont Interaction. mBio 9, e00024-18. https://doi.org/10.1128/mBio.00024-18

Preparation of fly extract

Collect 6 g flies (*Gff*) aged 10-24 days, starved for 2 days
Crush with sterile mortar and pestle in BSK-H, 30 ml per 6 g flies
Incubate 20 min at 56 °C, then let cool to room temperature
Centrifuge 15' at 4000 rpm
Centrifuge 15' at 13200 rpm
Filter at 0.45 µM
Filter at 0.22 µM
Store at -20 °C

Components of lipid mix

Cholesterol	10 mg
Palmitic acid	5 mg
Sphingomyelin	10 mg
Ethanol absolute warmed to 30 °C	1.8 mL
1-Palmitoyl-2-oleoyl-sn-glycerol (10 mg/ml in Ethanol)	100 µl
1,2-dioleoyl-sn-glycerol (20 mg/ml in Ethanol)	100 µl
Tween 40	50 µl
Tween 80	50 µl
Oleic acid	5.6 µl

Prepare 6% BSA stock (essentially fatty acid free) and mix 19.6 mL BSA 6% + 400 µl lipids

BSK-H-spiro

BSK-H medium without L-glutamine (Bio&Sell)	314 mL
Penicillin	2 mL
Arginine	10 mL
Rabbit Serum	24 mL
Fly extract in BSK-H	30 mL
Lipids mix	20 mL

Adjust pH to 7.5 and filter at 0.22 µM. Store at 4 °C for up to 2 weeks.

Table S1. Primers

Primer Name	Primer Sequence (5'-3')	Annealing Temperature (˚C)	Amplicon Size (bp)	Reference
Spi_16srRNA_F	GGGTGAGTAACACGTATCT	55	1000	Son et al 2021
Spi_16srRNA_R	CCTTCCTCTAGCTTACACTA			
Spiroplasma 16s_qPCR_F	ATGGCCTACCAAGACAATGATAC	56	138	
Spiroplasma 16s_qPCR_R	CCTTACAACAGACCTTTACATCC			
Tsetse-tubulin_F	ACGTATTCATTTCCCTTTGG	55	350	Son et al 2021
Tsetse-tubulin_R	AATGGCTGTGGTGTTGGACAAC			
Tubulin_qPCR_F	GATGGTCAAGTGCGATCCT	56	355	
Tubulin_qPCR_R	TGAGAACTCGCCTTCTTCC			

Table S2. Collection metadata

Sample name	Sample Type	Sample locality	Collection year	HMW DNA Method
Colony *Gff*	Female whole fly	*G.f.f* colony from the FAO/IAEA Insect Pest Control Laboratory in Seibersdorf, Austria	2023	Zymo Quick-DNA™ HMW MagBead Kit
Spiroplasma culture	Pelleted passage 9 *Spiroplasma* culture	*G.f.f* colony from the FAO/IAEA Insect Pest Control Laboratory in Seibersdorf, Austria	2024	Qiagen PureGene kit
Field *Gff*	Female whole fly	Trapped *G. f. f.* from Toloyang village (3.304659058, 32.37430505)) in Northwestern Uganda	2019	NEB Monarch® Genomic DNA Purification Kit

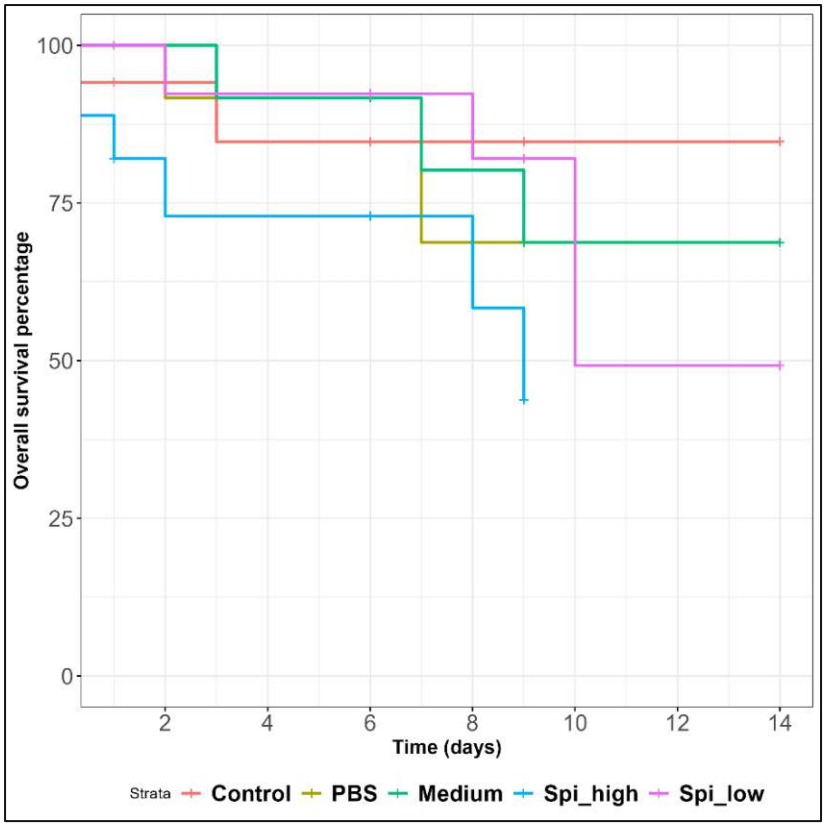

Figure S1. Preliminary results on the survival of injected Gff

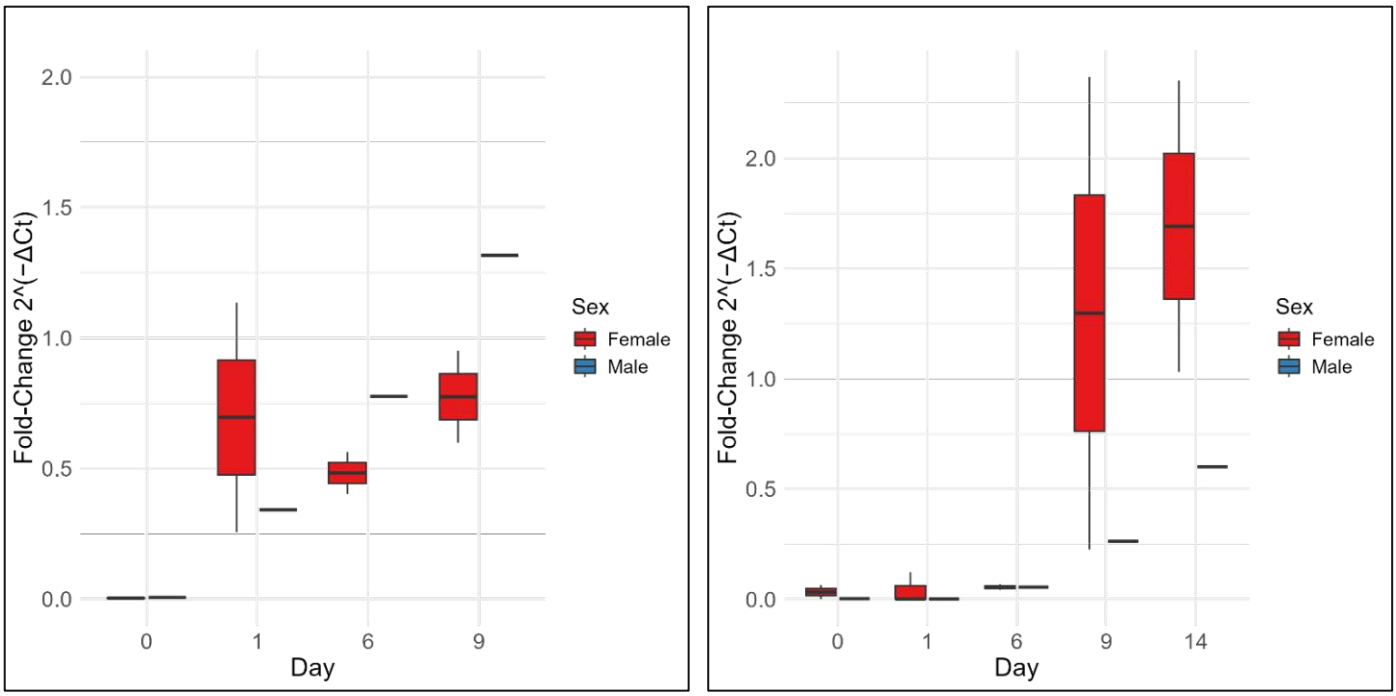

Figure S2 A, B. Preliminary results on the increase of *Spiroplasma* titers in injected, naïve *Gff*

A B

Table S3. Assembly statistics for the three *s*Gff genome assemblies.

BUSCO score is for the entomoplasmatales dataset.

Genome name	Size (Mb)	Contig #	Plasmid #	CDS #	BUSCO score
*s*Gff-colony	1.469	1	4	1837	97.6%
*s*Gff-cult	1.489	1	4	1869	97.3%
*s*Gff-field	1.404	1	3	1717	97.3%

6. General Discussion and Conclusions

This general discussion serves as a synthesis of the results from this dissertation, provides a detailed discussion and proposes future research directions based on the published findings. The performed work aimed to address two pivotal objectives in the realm of tsetse fly research, with potential implications for vector control strategies and the management of trypanosomosis. The first objective focussed on the development and characterization of microsatellite markers for the tsetse fly species *Glossina brevipalpis*, in order to facilitate larger scale population genetics studies to get a better understanding on population structure and gene flow of field populations. The second objective delved into the interactions between tsetse flies and their symbiotic bacteria, specifically *Sodalis* and *Spiroplasma* to identify the implications for vector competence, disease transmission and impacts on mass rearing settings for SIT applications.

Chapter 2 consists of the development and characterization of novel microsatellite markers for the tsetse fly species *G. brevipalpis.* Microsatellites are a widely used molecular tool in population genetics for the study of population structure, gene flow and dispersal of studied populations [206]. They are routinely applied in studies on various species of interest, including tsetse flies. Microsatellites have been developed for many tsetse fly species [120,122–124], with the exception of *G. brevipalpis*, the aim of the study therefore was to design suitable markers that are capable of differentiating between different laboratory and field populations. It mainly served as a proof of concept, as samples were selected from distinct populations to identify whether the developed microsatellites can successfully differentiate between these populations. Starting from many potential markers, nine selected microsatellites were retained after quality control and found to be suitable for discerning samples from the different locations. Screening the genome sequence of *G. brevipalpis* for microsatellite regions revealed a predominance of dinucleotide repeats over trinucleotide repeats, consistent with the documented inverse relationship between motif lengths and repeat numbers. The selected microsatellites underwent rigorous quality control tests to evaluate their effectiveness in assessing genetic heterogeneity within and between populations. Analysis of genetic distance parameters from populations sampled in neighbouring countries suggests that the observed genetic differentiation reflects geographic factors. Consistent findings from NJ-tree and FCA analyses indicated pronounced genetic differentiation between the studied regions, suggesting restricted population gene flow of the studied populations. The analyses revealed distinct genetic clusters of the investigated populations from a laboratory colony and field populations from South Africa and Mozambique. The genetic differentiation of *G. brevipalpis* from different localities underscores the enhanced resolution provided by microsatellite markers compared to mtDNA markers in terms of genetic structure, enabling more precise assessments of population structure, gene flow and dispersal capacities.

Nonetheless, the limited sample size and temporal disparities in sample collection between South Africa and Mozambique preclude definitive conclusions regarding the field situation. Therefore, it is recommended to undertake large-scale sampling surveys in these areas, including more samples from further field locations, sampled in the same timeframe. This approach will provide a more complete picture of the population structure of *G. brevipalpis* from these countries allowing for further conclusions about admixture of populations in close proximity and aiding decision-makers to develop the most suitable vector control strategies using a targeted area-wide integrated pest management approach. To further enhance the resolution of population structure in the field, it is proposed to test sequencing-based population genetics approaches such as RAD-seq [125] in future studies to obtain more data and potentially higher information content.

The prevalence of *Trypanosoma* and *Sodalis* in wild populations of tsetse flies across the African continent, their potential interactions and implications for SIT programmes were the subject of **Chapter 3.** In general, the vector competence of tsetse flies for various species of trypanosomes can be dynamic and influenced by different factors, among these can be the interactions with bacterial symbionts [127,147]. The role of *Sodalis* in facilitating trypanosome infections in tsetse flies is controversial: several studies suggested a positive correlation between *Sodalis* and trypanosomes infections [142,207,208], conversely, other studies have found no such correlation [161,209,210], therefore it is expected that the geographic distribution and ecological settings also play an important factor. The performed study on field-sampled tsetse flies highlighted considerable variability in trypanosome prevalence among different tsetse fly species and regions, with a total prevalence of 23.5%. Particularly high prevalence of trypanosomes was observed in *G. m. morsitans* and *G. pallidipes* populations from Central and East African countries, underscoring the presence of HAT and AAT in these areas. Similarly, *Sodalis* infection prevalence varied significantly across tsetse species and locations, with the average *Sodalis* prevalence being significantly higher in East, Central and Southern Africa (24.60%) compared to West Africa (2.70%). *Sodalis* was detected in most of the screened tsetse fly species but was not detected in *G. tachinoides* from Burkina Faso and Ghana and *G. p. gambiensis* from Mali, Guinea and Senegal. The performed statistical tests on coinfections with trypanosomes and *Sodalis* showed negative correlations in *G. medicorum* and *G. p. gambiensis* and a trend for positive correlation in *G. pallidipes*. These results underscore that coinfections of trypanosomes and *Sodalis* are not independent, varying depending on tsetse fly species and location. Further investigations are required to gain a more comprehensive view on the complex interactions between *Sodalis* and trypanosomes. In addition, future studies should focus on additional factors that potentially influence these interactions, such as environmental conditions and ecological habitats that different tsetse fly

species occur in. In addition, implications for sterile male releases in the context of SIT campaigns must be considered, given the statistically significant interactions.

Spiroplasma, another symbiotic bacterium was relatively recently identified in laboratory and field populations of the tsetse fly species *G. fuscipes fuscipes* and *G. tachinoides* [173]. While the negative and protective effects have been extensively studied in other insect hosts [176,177,180], little is known about the physiological impact on tsetse flies. However, there are indications that *Spiroplasma* may interact with trypanosomes and suppress its prevalence [188]. The effects of *Spiroplasma* on the reproductive and metabolic physiology of *Gff,* an important vector of sleeping sickness, were thus the subject of **Chapter 4**. The transcriptomic data presented in this study showed that infection with *Spiroplasma* significantly shifts gene expression in the reproductive tract of females and males compared to the uninfected control group. Functional assays on a heterogeneously infected laboratory colony revealed a significant increase of gonotrophic cycle length of females infected with *Spiroplasma.* Since female tsetse flies produce relatively few offspring in their lifetime compared to other insects, the increased length of the gonotrophic cycle may significantly influence population sizes and the production performance in mass-rearing facilities. This result may be explained by the also identified decreased lipid levels in infected flies, indicating that *Spiroplasma* derives lipids from the host to sustain its growth. Additionally, *Spiroplasma*-infected males exhibited significantly reduced sperm motility compared to the uninfected control group. It was observed that sperm from both infected and uninfected males had the same expression of sperm-specific *sdic* transcripts prior to mating. However, after insemination and transfer of sperm to the female spermathecae, significantly fewer *sdic* transcripts were expressed by sperm from infected males compared to their uninfected counterparts. The homologous gene of *sdic* in *Drosophila* encodes for a sperm dynein intermediate chain, necessary for the function of the cytoplasmic dynein motor protein complex [211], which plays a crucial role in sperm motility. These findings on sperm motility potentially reduce overall fecundity and raise concerns about the competitiveness of released sterile males in the field. As stated in **Chapter 1**, the use of competitive males is of utmost importance for the success of SIT and the findings from this study should be considered when planning and implementing such strategy. Moreover, the negative impact on female productivity as exemplified by prolonged gonotrophic cycles presents a negative effect on tsetse fly mass-rearing that should be considered.

In pursuit of better understanding the role of *Spiroplasma* and its interactions with the host and trypanosomes, further research was conducted which is presented in **Chapter 5**. This chapter focused on the *in vitro* cultivation and genome sequencing of *Spiroplasma* from *Glossina fuscipes fuscipes.* A *Spiroplasma in vitro* culture was established from hemolymph of infected *Gff,* making it the first successful attempt to cultivate the bacterium outside of the tsetse fly host. This achievement allowed

for detailed assessments on the *in vitro* growth kinetics, bacterial morphology and genome sequencing. Fluorescent microscopy examinations revealed the helical shape characteristic for *Spiroplasma,* as observed in other insect hosts. Growth kinetics assessment with quantitative PCR indicated growth in two phases, with a plateau phase in between exponential growth phases. The bacterial load was furthermore quantified by the calculation of copy numbers, resulting in a doubling time of approximately 40 hours under *in vitro* cultivation settings. As the genome sequence of *Spiroplasma* from *Gff* was previously unavailable, sequencing was performed utilizing the Oxford Nanopore technology. To allow for comparative genomic analyses, genome sequencing was conducted from three sample types: *Spiroplasma* from the *in vitro* culture, *in vivo Spiroplasma* from field-collected and laboratory colony *Gff*. The results of genome sequencing results revealed the complete circular genome of *Spiroplasma* with a size of 1.48 Mb along with several associated plasmids. Notably, the genome contained evidence of mobile genetic elements, including phages and transposons, which may play a role in horizontal gene transfer and genomic plasticity. Comparative genomics and synteny analysis of the three genomes showed no major differences in the genomic makeup as indicated by the difference of very few SNPs distinguishing them, except for a 20 Kb phage sequence insertion in the *Spiroplasma* genome derived from the *in vitro* culture. Phylogenetic tree reconstruction based on over 300 single-copy orthologous genes placed the three *Spiroplasma* genomes from *Gff* in context with other known *Spiroplasma* species, highlighting their evolutionary relationships. Gene annotation identified the core metabolic genes of the bacterium and potential symbiosis genes, including those coding for spiralins, adhesion-related proteins and ribosome-inactivating proteins (RIPs). These genes are likely essential for *Spiroplasma's* interaction with host cells and potentially trypanosomes. Metabolomic analyses were performed to gain insight on the metabolic pathways of *Spiroplasma*. In addition, the successful cultivation of *Spiroplasma* opens numerous avenues for future research, as the bacterium can be studied by itself outside from its host. This allows for more controlled experimental setups on the physiology of *Spiroplasma,* host interactions and impacts on vector competence through artificial infections of studied tsetse fly species.

Taken together, the work presented in this dissertation advances the understanding of population genetics and host-symbiont interactions of tsetse flies. It introduces novel microsatellite markers to study population structure and gene flow of wild *G. brevipalpis* populations, which can be of help in planning and implementing appropriate vector control strategies. Additionally, it provides detailed insights of *Sodalis* and trypanosome distribution on a continental scale, highlighting their correlations for some species and implications for SIT applications. The dissertation also describes the *in vitro* cultivation and genomic characterization of the bacterial symbiont *Spiroplasma*, revealing its impact

on host physiology and potential interactions with trypanosomes. This provides a comprehensive resource to further study of the symbiont's role in vector biology. Insights gained from this work raise considerations for mass-rearing conditions for SIT and deepen the knowledge about the interaction between tsetse flies, their heritable bacteria and trypanosomes. Overall, this dissertation contributes new knowledge to the scientific understanding of population genetics and symbiotic relationships in tsetse flies. It reveals how these interactions can inform decision-makers of vector control programs and influence disease transmission dynamics. Importantly, the findings have implications for reducing the disease burden in African countries affected by Human African Trypanosomosis (HAT) and African Animal Trypanosomosis (AAT). Enhanced vector control approaches, informed by population genetics and symbiont interactions can lead to more efficient and sustainable interventions, aiming to reduce the incidence of these diseases. In turn, this contributes to the efforts to improve public health and the livelihoods of affected communities.

Bibliography

1. Wiegmann, B.M.; Trautwein, M.D.; Winkler, I.S.; Barr, N.B.; Kim, J.-W.; Lambkin, C.; Bertone, M.A.; Cassel, B.K.; Bayless, K.M.; Heimberg, A.M.; et al. Episodic radiations in the fly tree of life. Proc. Natl. Acad. Sci. U.S.A. 2011, 108, 5690–5695, doi:10.1073/pnas.1012675108.

2. Buxton, P.A. The natural history of tsetse flies. An account of the biology of the genus *Glossina* (Diptera); London School of Hygiene and Tropical Medicine Memoir; H.K. Lewis & Co., Ltd: London, 1955.

3. Jordan, A.M. Tsetse flies as vectors of trypanosomes. Vet. Parasitol. 1976, 2, 143–152, doi:10.1016/0304-4017(76)90059-5.

4. Krafsur, E.S. Tsetse flies: Genetics, evolution, and role as vectors. Infect Genet Evol 2009, 9, 124–141, doi:10.1016/j.meegid.2008.09.010.

5. Pollock, J.N. Ecology and behaviour of tsetse; Food and Agriculture Organization (FAO): Caracalla (Rome, Italy), 1982; pp. 1–79.

6. Vreysen, M.J.B.; Seck, M.T.; 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, doi:10.1016/j.jip.2012.07.026.

7. Gooding, R.H.; Krafsur, E.S. Tsetse genetics: Contributions to biology, systematics, and control of tsetse flies. Annu Rev Entomol 2005, 50, 101–123, doi:10.1146/annurev.ento.50.071803.130443.

8. Rogers, D.J.; Randolph, S.E. Distribution and abundance of tsetse flies (*Glossina* spp.). The Journal of Animal Ecology 1986, 55, 1007, doi:10.2307/4430.

9. Moloo, S.K. The distribution of *Glossina* species in Africa and their natural hosts. International Journal of Tropical Insect Science 1993, 14, 511–527, doi:10.1017/S1742758400014211.

10. Cecchi, G.; Mattioli, R.C.; Slingenbergh, J.; de la Rocque, S. Land cover and tsetse fly distributions in sub-Saharan Africa. Med Vet Entomol 2008, 22, 364–373, doi:10.1111/j.1365-2915.2008.00747.x.

11. Distribution maps all species. ERGO: Environmental Research Group Oxford. Oxford, UK. 2024. Available online: https://ergodd.zoo.ox.ac.uk/tseweb/all_species.htm (accessed on 02 June 2024).

12. Benoit, J.B.; Attardo, G.M.; Baumann, A.A.; 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, doi:10.1146/annurev-ento-010814-020834.

13. Attardo, G.M.; Benoit, J.B.; Michalkova, V.; Yang, G.; Roller, L.; Bohova, J.; Takác, P.; Aksoy, S. Analysis of lipolysis underlying lactation in the tsetse fly, *Glossina morsitans*. Insect Biochem.Mol.Biol. 2012, 42, 360–370, doi:10.1016/j.ibmb.2012.01.007.

14. Jackson, C.H.N. The biology of tsetse flies. Biol. Rev. Camb. Philos. Soc. 1949, 24, 174–1989, doi:10.1111/j.1469-185X.1949.tb00574.x.

15. Attardo, G.M.; Guz, N.; Strickler-Dinglasan, P.; Aksoy, S. Molecular aspects of viviparous reproductive biology of the tsetse fly (*Glossina morsitans morsitans*): Regulation of yolk and milk gland protein synthesis. Journal of Insect Physiology 2006, 52, 1128–1136, doi:10.1016/j.jinsphys.2006.07.007.

16. Kabayo, J.P.; Langley, P.A. The nutritional importance of dietary blood components for reproduction in the tsetse fly, *Glossina morsitans*. Journal of Insect Physiology 1985, 31, 619–624, doi:10.1016/0022-1910(85)90060-5.

17. Entomologie médicale et vétérinaire; Duvallet, G., Fontenille, D., Robert, V., Eds.; Référence; IRD Éditions: Marseille, 2018; ISBN 978-2-7099-2377-4.

18. Aksoy, S.; Gibson, W.C.; Lehane, M.J. Interactions between tsetse and trypanosomes with implications for the control of trypanosomiasis. In Advances in Parasitology; Academic Press, 2003; Vol. 53, pp. 1–83.

19. Barrett, M.P.; Burchmore, R.J.S.; Stich, A.; Lazzari, J.O.; Frasch, A.C.; Cazzulo, J.J.; Krishna, S. The trypanosomiases. Lancet 2003, 362, 1469–1480, doi:10.1016/S0140-6736(03)14694-6.

20. Centers for Disease Control and Prevention Trypanosomiasis, African. Available online: https://www.cdc.gov/dpdx/trypanosomiasisafrican/index.html (accessed on 19 March 2024).

21. Welburn, S.C.; Maudlin, I. Tsetse-trypanosome interactions: Rites of passage. Parasitol Today 1999, 15, 399–403, doi:10.1016/s0169-4758(99)01512-4.

22. World Health Organization. Human African Trypanosomiasis (Sleeping Sickness). Available online: https://www.who.int/health-topics/human-african-trypanosomiasis#tab=tab_1 (accessed on 19 March 2024).

23. Gehrig, S.; Efferth, T. Development of drug resistance in *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense*. Treatment of human African trypanosomiasis with natural products (Review). Int J Mol Med 2008, 22, 411–419.

24. Büscher, P.; Cecchi, G.; Jamonneau, V.; Priotto, G. Human African trypanosomiasis. Lancet 2017, 390, 2397–2409, doi:10.1016/S0140-6736(17)31510-6.

25. Vincent, I.M.; Daly, R.; Courtioux, B.; Cattanach, A.M.; Biéler, S.; Ndung'u, J.M.; Bisser, S.; Barrett, M.P. Metabolomics identifies multiple candidate biomarkers to diagnose and stage human African trypanosomiasis. PLoS Negl Trop Dis 2016, 10, e0005140, doi:10.1371/journal.pntd.0005140.

26. Zillmann, U.; Albiez, E.J. The Testryp CATT (Card Agglutination Test for Trypanosomiasis): A field study on gambiense sleeping sickness in Liberia. Trop Med Parasitol 1986, 37, 390–392.

27. P De Koning, H. The drugs of sleeping sickness: Their mechanisms of action and resistance, and a brief history. Trop Med Infect Dis 2020, 5, 14, doi:10.3390/tropicalmed5010014.

28. Venturelli, A.; Tagliazucchi, L.; Lima, C.; Venuti, F.; Malpezzi, G.; Magoulas, G.E.; Santarem, N.; Calogeropoulou, T.; Cordeiro-da-Silva, A.; Costi, M.P. Current treatments to control African trypanosomiasis and one health perspective. Microorganisms 2022, 10, 1298, doi:10.3390/microorganisms10071298.

29. Kennedy, P.G. Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). Lancet Neurol 2013, 12, 186–194, doi:10.1016/S1474-4422(12)70296-X.

30. Headrick, D.R. Sleeping sickness epidemics and colonial responses in East and Central Africa, 1900-1940. PLoS Negl Trop Dis 2014, 8, e2772, doi:10.1371/journal.pntd.0002772.

31. Hide, G. History of sleeping sickness in East Africa. Clin Microbiol Rev 1999, 12, 112–125, doi:10.1128/CMR.12.1.112.

32. WHO. Control and surveillance of human African trypanosomiasis: Report of a WHO expert committee; 2013.

33. Franco, J.R.; Cecchi, G.; Priotto, G.; Paone, M.; Kadima Ebeja, A.; Simarro, P.P.; Diarra, A.; Sankara, D.; Zhao, W.; Dagne, D.A. Human African trypanosomiasis cases diagnosed in non-endemic countries (2011-2020). PLoS Negl Trop Dis 2022, 16, e0010885, doi:10.1371/journal.pntd.0010885.

34. Franco, J.R.; Cecchi, G.; Priotto, G.; Paone, M.; Diarra, A.; Grout, L.; Simarro, P.P.; Zhao, W.; Argaw, D. Monitoring the elimination of human African trypanosomiasis at continental and country level: Update to 2018. PLoS Negl Trop Dis 2020, 14, e0008261, doi:10.1371/journal.pntd.0008261.

35. World Health Organization. Human African Trypanosomiasis. Number of reported cases of human African trypanosomiasis (*T. b. gambiense*): 2023 Available online: https://apps.who.int/neglected_diseases/ntddata/hat/hat.html (accessed on 13 July 2024).

36. Simarro, P.P.; Cecchi, G.; Franco, J.R.; Paone, M.; Diarra, A.; Ruiz-Postigo, J.A.; Fevre, E.M.; Mattioli, R.C.; Jannin, J.G. Estimating and mapping the population at risk of sleeping sickness. PLoS Negl.Trop.Dis. 2012, 6, e1859-, doi:10.1371/journal.pntd.0001859.

37. Okello, I.; Mafie, E.; Eastwood, G.; Nzalawahe, J.; Mboera, L.E.G. African animal trypanosomiasis: A systematic review on prevalence, risk factors and drug resistance in sub-Saharan Africa. Journal of Medical Entomology 2022, 59, 1099–1143, doi:10.1093/jme/tjac018.

38. Hamill, L.C.; Kaare, M.T.; Welburn, S.C.; Picozzi, K. Domestic pigs as potential reservoirs of human and animal trypanosomiasis in Northern Tanzania. Parasites Vectors 2013, 6, 322, doi:10.1186/1756-3305-6-322.

39. Auty, H.; Anderson, N.E.; Picozzi, K.; Lembo, T.; Mubanga, J.; Hoare, R.; Fyumagwa, R.D.; Mable, B.; Hamill, L.; Cleaveland, S.; et al. Trypanosome diversity in wildlife species from the Serengeti and Luangwa Valley ecosystems. PLoS Negl.Trop.Dis. 2012, 6, e1828-, doi:10.1371/journal.pntd.0001828.

40. Eshetu, E.; Begejo, B. The current situation and diagnostic approach of nagana in Africa: A review. 2015.

41. Kasozi, K.I.; Namayanja, M.; Gaithuma, A.K.; Mahero, M.; Matovu, E.; Yamagishi, J.; Sugimoto, C.; MacLeod, E. Prevalence of hemoprotozoan parasites in small ruminants along a human-livestock-wildlife interface in Western Uganda. Veterinary Parasitology: Regional Studies and Reports 2019, 17, 100309, doi:10.1016/j.vprsr.2019.100309.

42. Desquesnes, M.; Gonzatti, M.; Sazmand, A.; Thévenon, S.; Bossard, G.; Boulangé, A.; Gimonneau, G.; Truc, P.; Herder, S.; Ravel, S.; et al. A review on the diagnosis of animal trypanosomoses. Parasites Vectors 2022, 15, 64, doi:10.1186/s13071-022-05190-1.

43. Giordani, F.; Morrison, L.J.; Rowan, T.G.; DE Koning, H.P.; Barrett, M.P. The animal trypanosomiases and their chemotherapy: A review. Parasitology 2016, 143, 1862–1889, doi:10.1017/S0031182016001268.

44. Kasozi, K.I.; Zirintunda, G.; Ssempijja, F.; Buyinza, B.; Alzahrani, K.J.; Matama, K.; Nakimbugwe, H.N.; Alkazmi, L.; Onanyang, D.; Bogere, P.; et al. Epidemiology of trypanosomiasis in wildlife—Implications for humans at the wildlife interface in Africa. Front. Vet. Sci. 2021, 8, 621699, doi:10.3389/fvets.2021.621699.

45. Hursey, B.S. The programme against Africa trypanosomiasis: Aims, objectives and achievements. Trends Parasitol. 2001, 17, 2–3, doi:10.1016/S1471-4922(00)01851-1.

46. Budd, L.T. DFID-funded tsetse and trypanosomosis research and development since 1980; Department of International Development: United Kingdom, 1999.

47. Swallow, B.M. Impacts of trypanosomiasis on African agriculture; PAAT: Nairobi, 1999; ISBN 92-5-104413-9.

48. Feldmann, U.; Hendrichs, J. The concept for integration of the sterile insect technique as a key component of future sub-regional, area-wide tsetse and trypanosomosis management operations.; Backhuys Publishers, 1995; pp. 193–214.

49. Feldmann, U.; Dyck, V.A.; Mattioli, R.C.; Jannin, J. Potential impact of tsetse fly control involving the sterile insect technique. In Sterile Insect Technique; Dyck, V.A., Hendrichs, J., Robinson, A.S., Eds.; Springer-Verlag: Berlin/Heidelberg, 2005; pp. 701–723 ISBN 978-1-4020-4050-4.

50. La Greca, F.; Magez, S. Vaccination against trypanosomiasis: Can it be done or is the trypanosome truly the ultimate immune destroyer and escape artist? Hum Vaccin 2011, 7, 1225–1233, doi:10.4161/hv.7.11.18203.

51. Vreysen, M.J.B.; Seck, M.T.; 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, doi:10.1016/j.jip.2012.07.026.

52. Allsopp, R. Options for vector control against trypanosomiasis in Africa. Trends Parasitol 2001, 17, 15–19, doi:10.1016/s1471-4922(00)01828-6.

53. Hocking, K.S.; Lamerton, J.F.; Lewis, E.A. Tsetse-fly control and eradication. Bull World Health Organ 1963, 28, 811–823.

54. Allsopp, R. Control of tsetse flies (Diptera: *Glossinidae*) using insecticides: A review and future prospects. Bull.Entomol.Res. 1984, 74, 1–23, doi:10.1017/S0007485300009895.

55. Grant, I.F. Insecticides for tsetse and trypanosomiasis control: Is the environmental risk acceptable? Trends Parasitol 2001, 17, 10–14, doi:10.1016/s1471-4922(00)01848-1.

56. Dransfield, R.D.; Williams, B.G.; Brightwell, R. Control of tsetse flies and trypanosomiasis: Myth or reality? Parasitology Today 1991, 7, 287–291, doi:10.1016/0169-4758(91)90099-A.

57. Adam, Y.; Cecchi, G.; Kgori, P.M.; Marcotty, T.; Mahama, C.I.; Abavana, M.; Anderson, B.; Paone, M.; Mattioli, R.; Bouyer, J. The sequential aerosol technique: A major component in an integrated strategy of intervention against riverine tsetse in Ghana. PLoS Negl Trop Dis 2013, 7, e2135, doi:10.1371/journal.pntd.0002135.

58. Ciss, M.; Bassène, M.D.; Seck, M.T.; Mbaye, A.G.; Sall, B.; Fall, A.G.; Vreysen, M.J.B.; Bouyer, J. Environmental impact of tsetse eradication in Senegal. Sci Rep 2019, 9, 20313, doi:10.1038/s41598-019-56919-5.

59. Lyczkowski, E. The effects of the 2006 spraying of deltamethrin for tsetse control on aquatic macro-invertebrates in the Kwando/Linyanti River system, Botswana 2006.

60. Malele, I.I.; Ouma, J.O.; Nyingilili, H.S.; Kitwika, W.A.; Malulu, D.J.; Magwisha, H.B.; Kweka, E.J. Comparative performance of traps in catching tsetse flies (Diptera: *Glossinidae*) in Tanzania. Onderstepoort J Vet Res 2016, 83, a1057, doi:10.4102/ojvr.v83i1.1057.

61. Asfaw, N.; Hiruy, B.; Worku, N.; Massebo, F. Evaluating the efficacy of various traps in catching tsetse flies at Nech Sar and Maze National Parks, Southwestern Ethiopia: An implication for Trypanosoma vector control. PLoS Negl Trop Dis 2022, 16, e0010999, doi:10.1371/journal.pntd.0010999.

62. Vale, G.A.; Hargrove, J.W.; Lehane, M.J.; Solano, P.; Torr, S.J. Optimal strategies for controlling riverine tsetse flies using targets: A modelling study. PLoS Negl Trop Dis 2015, 9, e0003615, doi:10.1371/journal.pntd.0003615.

63. Rayaisse, J.B.; Tirados, I.; Kaba, D.; Dewhirst, S.Y.; Logan, J.G.; Diarrassouba, A.; Salou, E.; Omolo, M.O.; Solano, P.; Lehane, M.J.; et al. Prospects for the development of odour baits to control the tsetse flies *Glossina tachinoides* and *G. palpalis* s.l. PLoS Negl Trop Dis 2010, 4, e632, doi:10.1371/journal.pntd.0000632.

64. Kappmeier, K. A newly developed odour-baited "H trap" for the live collection of *Glossina brevipalpis* and *Glossina austeni* (Diptera: Glossinidae) in South Africa. Onderstepoort J Vet Res 2000, 67, 15–26.

65. Green, C.H. The effects of odours and target colour on landing responses of *Glossina morsitans morsitans* and *G. pallidipes* (Diptera: *Glossinidae*). Bull.Entomol.Res. 1993, 83, 553–562, doi:10.1017/S0007485300039985.

66. Esterhuizen, J.; Rayaisse, J.B.; Tirados, I.; Mpiana, S.; Solano, P.; Vale, G.A.; Lehane, M.J.; Torr, S.J. Improving the cost-effectiveness of visual devices for the control of riverine tsetse flies, the major vectors of human African trypanosomiasis. PLoS Negl Trop Dis 2011, 5, e1257, doi:10.1371/journal.pntd.0001257.

67. Mbewe, N.J.; Saini, R.K.; Torto, B.; Irungu, J.; Yusuf, A.A.; Pirk, C.W.W. Sticky small target: An effective sampling tool for tsetse fly *Glossina fuscipes fuscipes* Newstead, 1910. Parasites Vectors 2018, 11, 268, doi:10.1186/s13071-018-2840-6.

68. Santer, R.D.; Vale, G.A.; Tsikire, D.; Torr, S.J. Optimising targets for tsetse control: Taking a fly's-eye-view to improve the colour of synthetic fabrics. PLoS Negl Trop Dis 2019, 13, e0007905, doi:10.1371/journal.pntd.0007905.

69. Bekele, J.; Asmare, K.; Abebe, G.; Ayelet, G.; Gelaye, E. Evaluation of deltamethrin applications in the control of tsetse and trypanosomosis in the Southern Rift Valley areas of Ethiopia. Vet.Parasitol. 2010, 168, 177–184.

70. Rothwell, J.; Hacket, K.; Friend, M.; Farnsworth, W.; Lowe, L. Efficacy of zeta-cypermethrin as pour-on or spray formulations for the control of buffalo fly (*Haematobia irritans exigua*) in cattle. Aust Veterinary J 1998, 76, 610–612, doi:10.1111/j.1751-0813.1998.tb10240.x.

71. Okello, W.O.; MacLeod, E.T.; Muhanguzi, D.; Waiswa, C.; Welburn, S.C. Controlling tsetse flies and ticks using insecticide treatment of cattle in Tororo District Uganda: Cost benefit analysis. Front Vet Sci 2021, 8, 616865, doi:10.3389/fvets.2021.616865.

72. Okello-Onen, J.; Heinonen, R.; Ssekitto, C.M.B.; Mwayi, W.T.; Kakaire, D.; Kabarema, M. Control of tsetse flies in Uganda by dipping cattle in deltamethrin. Trop.Anim.Health Prod. 1994, 26, 21–27, doi:10.1007/BF02241128.

73. Leak, S.G.A.; Mulatu, W.; Rowlands, G.J.; D'Ieteren, G.D.M. A trial of a cypermethrin "pour-on" insecticide to control *Glossina pallipes*, *G. fuscipes fuscipes* and *G. morsitans submorsitans* (Diptera:

Glossinidae) in South-West Ethiopia. Bull.Entomol.Res. 1995, 85, 241–251, doi:10.1017/S0007485300034325.

74. Vale, G.A.; Grant, I.F.; Dewhurst, C.F.; Aigreau, D. Biological and chemical assays of pyrethroids in cattle dung. Bull.Entomol.Res. 2004, 94, 273–282, doi:10.1079/BER2004300.

75. Vale, G.A.; Grant, I.F. Modelled impact of insecticide-contaminated dung on the abundance and distribution of dung fauna. Bull.Entomol.Res. 2002, 92, 251–263, doi:10.1079/BER2002162.

76. Knipling, E.F. The sterility principle of insect population control. Pesticides Abstracts and News Summaries 1964, 10, 587–603.

77. Vreysen, M.J. Principles of area-wide integrated tsetse fly control using the sterile insect technique. Med Trop (Mars) 2001, 61, 397–411.

78. Bakri, A.; Mehta, K.; Lance, D. Sterilizing insects with ionizing radiation. In Sterile Insect Technique. Principles and practice in area-wide integrated pest management; Dyck, V.A., Hendrichs, J.P., Robinson, A.S., Eds.; CRC Press: Boca Raton, 2021; pp. 355–398 ISBN 978-1-00-303557-2.

79. Robinson, A.S. Genetic basis of the sterile insect technique. In Sterile Insect Technique; CRC Press: Boca Raton, 2021; pp. 143–162 ISBN 978-1-00-303557-2.

80. IAEA Sterile insect technique and biological control: How radiation helps fight pests 2023. Available online: https://www.iaea.org/newscenter/news/sterile-insect-technique-and-biological-control-how-radiation-helps-fight-pests (accessed on 05 June 2024).

81. Mews, A.R.; Langley, P.A.; Pimley, R.W.; Flood, M.E.T. Large-scale rearing of tsetse flies (*Glossina* spp.) in the absence of a living host. Bull. Entomol. Res. 1977, 67, 119–128, doi:10.1017/S0007485300010944.

82. Vreysen, M.J.B.; Saleh, K.M.; Lancelot, R.; Bouyer, J. Factory tsetse flies must behave like wild flies: A prerequisite for the sterile insect technique. PLoS Negl Trop Dis 2011, 5, e907, doi:10.1371/journal.pntd.0000907.

83. Menjeta, M.; Seal, N.J.; Feldmann, U.; McCall, P.J. Mating competitiveness of irradiated *Glossina austeni* tsetse flies. Annals of Tropical Medicine & Parasitology 2004, 98, 539–542, doi:10.1179/000349804225003587.

84. Parker, A.; Mehta, K. Sterile insect technique: A model for dose optimization for improved sterile insect quality. Florida Entomologist 2007, 90, 88–95, doi:10.1653/0015 4040(2007)90[88:sitamf]2.0.co;2.

85. Hendrichs, J.P. Use of the sterile insect technique against key insect pests. Sustainable Development International 2000, 2, 75–79.

86. Kaboré, B.A.; Nawaj, A.; Maiga, H.; Soukia, O.; Pagabeleguem, S.; Ouédraogo/Sanon, M.S.G.; Vreysen, M.J.B.; Mach, R.L.; De Beer, C.J. X-rays are as effective as gamma-rays for the sterilization of *Glossina palpalis gambiensis* vanderplank, 1911 (Diptera: *Glossinidae*) for use in the sterile insect technique. Sci Rep 2023, 13, 17633, doi:10.1038/s41598-023-44479-8.

87. Knipling, E.F. Sterile-male method of population control. Science 1959, 130, 902–904, doi:10.1126/science.130.3380.902.

88. Sterile Insect Technique: Principles and practice in area-wide integrated pest management; Dyck, V.A., Hendrichs, J., Robinson, A.S., Eds.; 2nd ed.; CRC Press: Boca Raton, FL, 2021; ISBN 978-1-00-303557-2.

89. Villaseñor, A.; Carrillo, J.; Zavala, J.; Stewart, J.; Lira, C.; Reyes, J. Current progress in the Medfly program Mexico-Guatemala. In; Tan, K.H., Ed.; Joint Proceedings of the International Conference on Area-Wide Control of Insect Pests, May 28-June 2, 1998 and the Fifth International Symposium on Fruit Flies of Economic Importance, June 1-5, 2008, Penang, Malaysia; Penerbit Universiti Sains Malaysia: Penang, Malaysia, 1998; pp. 361–368 ISBN 983-861-195-6.

90. Hendrichs, J.; Robinson, A.S.; Cayol, J.P.; Enkerlin, W.R. Medfly areawide sterile insect technique programmes for prevention, suppression or eradication: The importance of mating behavior studies. Fla.Entomol. 2002, 85, 1–13, doi:10.1653/0015-4040(2002)085[0001:MASITP]2.0.CO;2.

91. Plá, I.; García De Oteyza, J.; Tur, C.; Martínez, M.Á.; Laurín, M.C.; Alonso, E.; Martínez, M.; Martín, Á.; Sanchis, R.; Navarro, M.C.; et al. Sterile insect technique programme against Mediterranean fruit fly in the Valencian Community (Spain). Insects 2021, 12, 415, doi:10.3390/insects12050415.

92. Vreysen, M.J.B.; Saleh, K.M.; Ali, M.Y.; Abdulla, A.M.; Zhu, Z.-R.; Juma, K.G.; Dyck, V.A.; Msangi, A.R.; Mkonyi, P.A.; Feldmann, H.U. *Glossina austeni* (Diptera: *Glossinidae*) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique. J.Econ.Entomol. 2000, 93, 123–135, doi:10.1603/0022-0493-93.1.123.

93. Seck, M.T.; Fall, A.G.; Ciss, M.; Bakhoum, M.T.; Sall, B.; Gaye, A.M.; Gimonneau, G.; Bassène, M.D.; Lancelot, R.; Vreysen, M.J.B.; et al. Animal trypanosomosis eliminated in a major livestock production region in Senegal following the eradication of a tsetse population. Parasite 2024, 31, 11, doi:10.1051/parasite/2024010.

94. Brookfield, J. Population genetics. Current Biology 1996, 6, 354–356, doi:10.1016/S0960-9822(02)00493-1.

95. Futuyma, D. Evolutionary biology; 3rd ed.; Sinauer Associates, Inc.: MA, United States, 2006; ISBN 978-0-87893-185-9.

96. Vogel, F.; Motulsky, A.G. Population genetics: Description and dynamics. In Human genetics; Springer Berlin Heidelberg: Berlin, Heidelberg, 1997; pp. 495–548 ISBN 978-3-662-03358-6.

97. Wright, S. Evolution and the genetics of populations. Volume 1. Genetics and biometric foundations.; University of Chicago Press.: Chicago, IL, USA., 1969; Vol. 1; ISBN 978-0-226-91038-3.

98. Wright, S. Evolution and the genetics of populations. Volume 2. The theory of gene frequencies.; University of Chicago Press.: Chicago, IL, USA., 1969; Vol. 2; ISBN 978-0-226-91039-0.

99. Wright, S. Evolution and the genetics of populations. Volume 3. Experimental results and evolutionary deductions.; University of Chicago Press.: Chicago, IL, USA., 1978; Vol. 3; ISBN 978-0-226-91051-2.

100. Wright, S. Evolution and the genetics of populations. Volume 4. Variability within and among natural populations.; University of Chicago Press.: Chicago, IL, USA., 1978; Vol. 4; ISBN 978-0-226-91041-3.

101. Nei, M. Genetic distance between populations. The American Naturalist 1972, 106, 283–292.

102. Hardy, G.H. Mendelian proportions in a mixed population. Science 1908, 28, 49–50, doi:10.1126/science.28.706.49.

103. Rousset, F.; Raymond, M. Statistical analyses of population genetic data: New tools, old concepts. Trends Ecol Evol 1997, 12, 313–317, doi:10.1016/S0169-5347(97)01104-X.

104. Cavalli-Sforza, L.L.; Edwards, A.W. Phylogenetic analysis. Models and estimation procedures. Am J Hum Genet 1967, 19, 233–257.

105. Stevens, L.; Dorn, P.L. Population genetics of *Triatominae*. In American Trypanosomiasis Chagas disease; Elsevier, 2017; pp. 169–196 ISBN 978-0-12-801029-7.

106. De Meeûs, T.; Ravel, S.; Rayaisse, J.-B.; Courtin, F.; Solano, P. Understanding local population genetics of tsetse: The case of an isolated population of *Glossina palpalis gambiensis* in Burkina Faso. Infect Genet Evol 2012, 12, 1229–1234, doi:10.1016/j.meegid.2012.04.005.

107. Solano, P.; Kaba, D.; Ravel, S.; Dyer, N.A.; Sall, B.; Vreysen, M.J.B.; Seck, M.T.; Darbyshir, H.; Gardes, L.; Donnelly, M.J.; et al. Population genetics as a tool to select tsetse control strategies: Suppression or eradication of *Glossina palpalis gambiensis* in the Niayes of Senegal. PLoS Negl Trop Dis 2010, 4, e692, doi:10.1371/journal.pntd.0000692.

108. Wyss, J.H. Screwworm eradication in the Americas. Annals of the New York Academy of Sciences 2000, 916, 186–193, doi:10.1111/j.1749-6632.2000.tb05289.x.

109. Guillot, G.; Santos, F. Using AFLP markers and the Geneland program for the inference of population genetic structure. Molecular Ecology Resources 2010, 10, 1082–1084, doi:10.1111/j.1755-0998.2010.02864.x.

110. Harrison, R.G. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends in Ecology & Evolution 1989, 4, 6–11, doi:10.1016/0169-5347(89)90006-2.

111. Khasdan, V.; Levin, I.; Rosner, A.; Morin, S.; Kontsedalov, S.; Maslenin, L.; Horowitz, A.R. DNA markers for identifying biotypes B and Q of *Bemisia tabaci* (Hemiptera: *Aleyrodidae*) and studying population dynamics. Bull. Entomol. Res. 2005, 95, 605–613, doi:10.1079/BER2005390.

112. Zietkiewicz, E.; Rafalski, A.; Labuda, D. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 1994, 20, 176–183, doi:10.1006/geno.1994.1151.

113. Rakoczy-Trojanowska, M.; Bolibok, H. Characteristics and a comparison of three classes of microsatellite-based markers and their application in plants. Cell Mol Biol Lett 2004, 9, 221–238.

114. Bruford, M.W.; Wayne, R.K. Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 1993, 3, 939–943, doi:10.1016/0959-437x(93)90017-j.

115. Putman, A.I.; Carbone, I. Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecol Evol 2014, 4, 4399–4428, doi:10.1002/ece3.1305.

116. Etter, P.D.; Bassham, S.; Hohenlohe, P.A.; Johnson, E.A.; Cresko, W.A. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In Molecular methods for evolutionary genetics; Orgogozo, V., Rockman, M.V., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; Vol. 772, pp. 157–178 ISBN 978-1-61779-227-4.

117. Baird, N.A.; Etter, P.D.; Atwood, T.S.; Currey, M.C.; Shiver, A.L.; Lewis, Z.A.; Selker, E.U.; Cresko, W.A.; Johnson, E.A. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 2008, 3, e3376, doi:10.1371/journal.pone.0003376.

118. Reitzel, A.M.; Herrera, S.; Layden, M.J.; Martindale, M.Q.; Shank, T.M. Going where traditional markers have not gone before: Utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. Molecular Ecology 2013, 22, 2953–2970, doi:10.1111/mec.12228.

119. Graham, C.F.; Glenn, T.C.; McArthur, A.G.; Boreham, D.R.; Kieran, T.; Lance, S.; Manzon, R.G.; Martino, J.A.; Pierson, T.; Rogers, S.M.; et al. Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADseq). Molecular Ecology Resources 2015, 15, 1304–1315, doi:10.1111/1755-0998.12404.

120. Baker, M.D.; Krafsur, E.S. Identification and properties of microsatellite markers in tsetse flies *Glossina morsitans* sensu lato (Diptera: *Glossinidae*). Mol Ecol Notes 2001, 1, 234–236, doi:10.1046/j.1471-8278.2001.00087.x.

121. Abila, P.P.; Slotman, M.A.; Parmakelis, A.; Dion, K.B.; Robinson, A.S.; Muwanika, V.B.; Enyaru, J.C.K.; Okedi, L.M.A.; Lokedi, L.M.; Aksoy, S.; et al. High levels of genetic differentiation between Ugandan *Glossina fuscipes fuscipes* populations separated by Lake Kyoga. PLoS Negl Trop Dis 2008, 2, e242, doi:10.1371/journal.pntd.0000242.

122. Brown, J.E.; Komatsu, K.J.; Abila, P.P.; Robinson, A.S.; Okedi, L.M.A.; Dyer, N.; Donnelly, M.J.; Slotman, M.A.; Caccone, A. Polymorphic microsatellite markers for the tsetse fly *Glossina fuscipes fuscipes* (Diptera: *Glossinidae*), a vector of human African trypanosomiasis. Mol Ecol Resour 2008, 8, 1506–1508, doi:10.1111/j.1755-0998.2008.02328.x.

123. Ravel, S.; Séré, M.; Manangwa, O.; Kagbadouno, M.; Mahamat, M.H.; Shereni, W.; Okeyo, W.A.; Argiles-Herrero, R.; De Meeûs, T. Developing and quality testing of microsatellite loci for four species of *Glossina*. Infect Genet Evol 2020, 85, 104515, doi:10.1016/j.meegid.2020.104515.

124. Ouma, J.O.; Cummings, M.A.; Jones, K.C.; Krafsur, E.S. Characterization of microsatellite markers in the tsetse fly, *Glossina pallidipes* (Diptera: *Glossinidae*). Mol Ecol Notes 2003, 3, 450–453, doi:10.1046/j.1471-8286.2003.00480.x.

125. Gloria-Soria, A.; Dunn, W.A.; Telleria, E.L.; Evans, B.R.; Okedi, L.; Echodu, R.; Warren, W.C.; Montague, M.J.; Aksoy, S.; Caccone, A. Patterns of genome-wide variation in *Glossina fuscipes fuscipes* tsetse flies from Uganda. G3 (Bethesda) 2016, 6, 1573–1584, doi:10.1534/g3.116.027235.

126. Saarman, N.P.; Son, J.H.; Zhao, H.; Cosme, L.V.; Kong, Y.; Li, M.; Wang, S.; Weiss, B.L.; Echodu, R.; Opiro, R.; et al. Genomic evidence of sex chromosome aneuploidy and infection-associated genotypes in the tsetse fly *Glossina fuscipes*, the major vector of African trypanosomiasis in Uganda. Infection, Genetics and Evolution 2023, 114, 105501, doi:10.1016/j.meegid.2023.105501.

127. Attardo, G.M.; Scolari, F.; Malacrida, A. Bacterial symbionts of tsetse flies: Relationships and functional interactions between tsetse flies and their symbionts. In Symbiosis: Cellular, molecular, medical and evolutionary aspects; Kloc, M., Ed.; Results and Problems in Cell Differentiation; Springer International Publishing: Cham, 2020; Vol. 69, pp. 497–536 ISBN 978-3-030-51848-6.

128. De Bary, A. Die Erscheinung der Symbiose; Verlag Trubner: Strassburg, 1879.

129. Klepzig, K.D.; Adams, A.S.; Handelsman, J.; Raffa, K.F. Symbioses: A key driver of insect physiological processes, ecological interactions, evolutionary diversification, and impacts on humans. en 2009, 38, 67–77, doi:10.1603/022.038.0109.

130. Paszkowski, U. Mutualism and parasitism: The yin and yang of plant symbioses. Current Opinion in Plant Biology 2006, 9, 364–370, doi:10.1016/j.pbi.2006.05.008.

131. Martin, B.D.; Schwab, E. Current usage of symbiosis and associated terminology. IJB 2012, 5, p32, doi:10.5539/ijb.v5n1p32.

132. Kariithi, H.M.; Meki, I.K.; Schneider, D.I.; De Vooght, L.; Khamis, F.M.; Geiger, A.; Demirbaş-Uzel, G.; Vlak, J.M.; iNCE, I.A.; Kelm, S.; et al. Enhancing vector refractoriness to trypanosome infection: Achievements, challenges and perspectives. BMC Microbiol 2018, 18, 179, doi:10.1186/s12866-018-1280-y.

133. Aksoy, S.; Caccone, A.; Galvani, A.P.; Okedi, L.M. *Glossina fuscipes* populations provide insights for human African trypanosomiasis transmission in Uganda. Trends in Parasitology 2013, 29, 394–406, doi:10.1016/j.pt.2013.06.005.

134. 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, doi:10.1038/ng986.

135. Rio, R.V.M.; Symula, R.E.; Wang, J.; Lohs, C.; Wu, Y.; Snyder, A.K.; Bjornson, R.D.; Oshima, K.; Biehl, B.S.; Perna, N.T.; et al. Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: *Glossinidae*) obligate symbiont *Wigglesworthia*. mBio 2012, 3, e00240-11, doi:10.1128/mBio.00240-11.

136. Pais, R.; Lohs, C.; Wu, Y.; Wang, J.; Aksoy, S. The obligate mutualist *Wigglesworthia glossinidia* influences reproduction, digestion, and immunity processes of its host, the tsetse fly. Appl Environ Microbiol 2008, 74, 5965–5974, doi:10.1128/AEM.00741-08.

137. Weiss, B.L.; Maltz, M.; Aksoy, S. Obligate symbionts activate immune system development in the tsetse fly. The Journal of Immunology 2012, 188, 3395–3403, doi:10.4049/jimmunol.1103691.

138. Bing, X.; Attardo, G.M.; Vigneron, A.; Aksoy, E.; Scolari, F.; Malacrida, A.; Weiss, B.L.; Aksoy, S. Unravelling the relationship between the tsetse fly and its obligate symbiont *Wigglesworthia*: Transcriptomic and metabolomic landscapes reveal highly integrated physiological networks. Proc. R. Soc. B. 2017, 284, 20170360, doi:10.1098/rspb.2017.0360.

139. Aksoy, S. *Wigglesworthia* gen. nov. and *Wigglesworthia glossinidia* sp. nov., taxa consisting of the mycetocyte-associated, primary endosymbionts of tsetse flies. International Journal of Systematic Bacteriology 1995, 45, 848–851, doi:10.1099/00207713-45-4-848.

140. Dale, C.; Welburn, S.C. The endosymbionts of tsetse flies: Manipulating host-parasite interactions. Int J Parasitol 2001, 31, 628–631, doi:10.1016/s0020-7519(01)00151-5.

141. Hamidou Soumana, I.; Loriod, B.; Tchicaya, B.; Simo, G.; Rihet, P.; Geiger, A. The transcriptional signatures of *Sodalis glossinidius* in the *Glossina palpalis gambiensis* flies negative *for Trypanosoma brucei gambiense* contrast with those of this symbiont in tsetse flies positive for the parasite: TPossible involvement of a *Sodalis*-hosted prophage in fly *Trypanosoma* refractoriness? Infection, Genetics and Evolution 2014, 24, 41–56, doi:10.1016/j.meegid.2014.03.005.

142. Farikou, O.; Njiokou, F.; Mbida Mbida, J.A.; Njitchouang, G.R.; Djeunga, H.N.; Asonganyi, T.; Simarro, P.P.; Cuny, G.; Geiger, A. Tripartite interactions between tsetse flies, *Sodalis glossinidius* and trypanosomes—An epidemiological approach in two historical human African trypanosomiasis foci in Cameroon. Infection, Genetics and Evolution 2010, 10, 115–121, doi:10.1016/j.meegid.2009.10.008.

143. Geiger, A.; Ravel, S.; Frutos, R.; Cuny, G. *Sodalis glossinidius* (*Enterobacteriaceae*) and vectorial competence of *Glossina palpalis gambiensis* and *Glossina morsitans morsitans* for *Trypanosoma congolense* savannah type. Curr.Microbiol. 2005, 51, 35–40, doi:10.1007/s00284-005-4525-6.

144. Hamidou Soumana, I.; Berthier, D.; Tchicaya, B.; Thevenon, S.; Njiokou, F.; Cuny, G.; Geiger, A. Population dynamics of *Glossina palpalis gambiensis* symbionts, *Sodalis glossinidius*, and *Wigglesworthia glossinidia*, throughout host-fly development. Infect Genet Evol 2013, 13, 41–48, doi:10.1016/j.meegid.2012.10.003.

145. Cheng, Q.; Aksoy, S. Tissue tropism, transmission and expression of foreign genes *in vivo* in midgut symbionts of tsetse flies. Insect Molecular Biology 1999, 8, 125–132, doi:10.1046/j.1365-2583.1999.810125.x.

146. De Vooght, L.; Caljon, G.; Van Hees, J.; Van Den Abbeele, J. Paternal transmission of a secondary symbiont during mating in the viviparous tsetse fly. Mol Biol Evol 2015, 32, 1977–1980, doi:10.1093/molbev/msv077.

147. Mfopit, Y.M.; Weber, J.S.; Chechet, G.D.; Ibrahim, M.A.M.; Signaboubo, D.; Achukwi, D.M.; Mamman, M.; Balogun, E.O.; Shuaibu, M.N.; Kabir, J.; et al. Molecular detection of *Sodalis glossinidius, Spiroplasma* and *Wolbachia* endosymbionts in wild population of tsetse flies collected in Cameroon, Chad and Nigeria 2023.

148. Toh, H.; Weiss, B.L.; Perkin, S.A.H.; Yamashita, A.; Oshima, K.; Hattori, M.; Aksoy, S. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. Genome Res. 2006, 16, 149–156, doi:10.1101/gr.4106106.

149. Akman, L.; Rio, R.V.M.; Beard, C.B.; Aksoy, S. Genome size determination and coding capacity of *Sodalis glossinidius*, an enteric symbiont of tsetse flies, as revealed by hybridization to *Escherichia coli* gene arrays. J Bacteriol 2001, 183, 4517–4525, doi:10.1128/JB.183.15.4517-4525.2001.

150. Belda, E.; Moya, A.; Bentley, S.; Silva, F.J. Mobile genetic element proliferation and gene inactivation impact over the genome structure and metabolic capabilities of *Sodalis glossinidius*, the secondary endosymbiont of tsetse flies. BMC Genomics 2010, 11, 449, doi:10.1186/1471-2164-11-449.

151. Snyder, A.K.; Deberry, J.W.; Runyen-Janecky, L.; Rio, R.V.M. Nutrient provisioning facilitates homeostasis between tsetse fly (Diptera: *Glossinidae*) symbionts. Proc.R.Soc.B 2010, 277, 2389–2397, doi:10.1098/rspb.2010.0364.

152. De Vooght, L.; De Ridder, K.; Hussain, S.; Stijlemans, B.; De Baetselier, P.; Caljon, G.; Van Den Abbeele, J. Targeting the tsetse-trypanosome interplay using genetically engineered *Sodalis glossinidius*. PLoS Pathog 2022, 18, e1010376, doi:10.1371/journal.ppat.1010376.

153. De Vooght, L.; Caljon, G.; De Ridder, K.; Van Den Abbeele, J. Delivery of a functional anti-trypanosome nanobody in different tsetse fly tissues via a bacterial symbiont, *Sodalis glossinidius*. Microb Cell Fact 2014, 13, 156, doi:10.1186/s12934-014-0156-6.

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

155. Doudoumis, V.; Tsiamis, G.; Wamwiri, F.; Brelsfoard, C.; Alam, U.; Aksoy, E.; Dalaperas, S.; Abd-Alla, A.; Ouma, J.; Takac, P.; et al. Detection and characterization of *Wolbachia* infections in laboratory and natural populations of different species of tsetse flies (genus *Glossina*). BMC Microbiol 2012, 12 Suppl 1, S3, doi:10.1186/1471-2180-12-S1-S3.

156. Schneider, D.I.; Garschall, K.I.; Parker, A.G.; Abd-alla, A.M.M.; Miller, W.J. Global *Wolbachia* prevalence, titer fluctuations and their potential of causing cytoplasmic incompatibilities in tsetse flies and hybrids of *Glossina morsitans* subgroup species. J Invertebr Pathol 2013, 112 Suppl, S104-115, doi:10.1016/j.jip.2012.03.024.

157. Stouthamer, R.; Breeuwer, J.A.; Hurst, G.D. *Wolbachia pipientis*: Microbial manipulator of arthropod reproduction. Annu Rev Microbiol 1999, 53, 71–102, doi:10.1146/annurev.micro.53.1.71.

158. Tram, U.; Ferree, P.M.; Sullivan, W. Identification of *Wolbachia*–host interacting factors through cytological analysis. Microbes and Infection 2003, 5, 999–1011, doi:10.1016/S1286-4579(03)00192-8.

159. Bourtzis, K. *Wolbachia*-based technologies for insect pest population control. In; Aksoy, S., Ed.; Springer Science+Business Media, LLC 233 Springer Street, New York, New York 10013, USA (www.springer.com): New York, 2008; pp. 104–113.

160. Clark, M.E.; Veneti, Z.; Bourtzis, K.; Karr, T.L. *Wolbachia* distribution and cytoplasmic incompatibility during sperm development: The cyst as the basic cellular unit of CI expression. Mech Dev 2003, 120, 185–198, doi:10.1016/S0925-4773(02)00424-0.

161. Kame-Ngasse, G.I.; Njiokou, F.; Melachio-Tanekou, T.T.; Farikou, O.; Simo, G.; Geiger, A. Prevalence of symbionts and trypanosome infections in tsetse flies of two villages of the "Faro and Déo" Division of the Adamawa Region of Cameroon. BMC Microbiol 2018, 18, 159, doi:10.1186/s12866-018-1286-5.

162. Kanté, S.T.; Melachio, T.; Ofon, E.; Njiokou, F.; Simo, G. Detection of *Wolbachia* and different trypanosome species in *Glossina palpalis palpalis* populations from three sleeping sickness foci of Southern Cameroon. Parasites Vectors 2018, 11, 630, doi:10.1186/s13071-018-3229-2.

163. Alam, U.; Medlock, J.; Brelsfoard, C.; Pais, R.; Lohs, C.; Balmand, S.; Carnogursky, J.; Heddi, A.; Takac, P.; Galvani, A.; et al. *Wolbachia* symbiont infections induce strong cytoplasmic incompatibility in the tsetse fly *Glossina morsitans*. PLoS Pathog 2011, 7, e1002415, doi:10.1371/journal.ppat.1002415.

164. Zabalou, S.; Riegler, M.; Theodorakopoulou, M.; Stauffer, C.; Savakis, C.; Bourtzis, K. *Wolbachia* -induced cytoplasmic incompatibility as a means for insect pest population control. Proc. Natl. Acad. Sci. U.S.A. 2004, 101, 15042–15045, doi:10.1073/pnas.0403853101.

165. Cheng, Q.; Ruel, T.D.; Zhou, W.; Moloo, S.K.; Majiwa, P.; O'neill, S.L.; Aksoy, S. Tissue distribution and prevalence of *Wolbachia* infections in tsetse flies, *Glossina* spp. Medical Vet Entomology 2000, 14, 44–50, doi:10.1046/j.1365-2915.2000.00202.x.

166. O'Neill, S.L.; Gooding, R.H.; Aksoy, S. Phylogenetically distant symbiotic microorganisms reside in *Glossina* midgut and ovary tissues. Medical Vet Entomology 1993, 7, 377–383, doi:10.1111/j.1365-2915.1993.tb00709.x.

167. Schneider, D.I.; Parker, A.G.; Abd-alla, A.M.; Miller, W.J. High-sensitivity detection of cryptic *Wolbachia* in the African tsetse fly (*Glossina* spp.). BMC Microbiol 2018, 18, 140, doi:10.1186/s12866-018-1291-8.

168. Mouton, L.; Henri, H.; Charif, D.; Boulétreau, M.; Vavre, F. Interaction between host genotype and environmental conditions affects bacterial density in *Wolbachia* symbiosis. Biol. Lett. 2007, 3, 210–213, doi:10.1098/rsbl.2006.0590.

169. Brelsfoard, C.; Tsiamis, G.; Falchetto, M.; Gomulski, L.M.; Telleria, E.; Alam, U.; Doudoumis, V.; Scolari, F.; Benoit, J.B.; Swain, M.; et al. Presence of extensive *Wolbachia* symbiont insertions discovered in the genome of its host *Glossina morsitans morsitans*. PLoS Negl Trop Dis 2014, 8, e2728, doi:10.1371/journal.pntd.0002728.

170. Harne, S.; Gayathri, P.; Béven, L. Exploring *Spiroplasma* biology: Opportunities and challenges. Front. Microbiol. 2020, 11, 589279, doi:10.3389/fmicb.2020.589279.

171. Whitcomb, R.F. The biology of *Spiroplasma*s. Annual Review of Entomology 1981, 26, 397–425, doi:10.1146/annurev.en.26.010181.002145.

172. Dera, K.-S.M.; Dieng, M.M.; Moyaba, P.; Ouedraogo, G.M.; Pagabeleguem, S.; Njokou, F.; Ngambia Freitas, F.S.; De Beer, C.J.; Mach, R.L.; Vreysen, M.J.; et al. Prevalence of *Spiroplasma* and interaction with wild *Glossina tachinoides* microbiota. Parasite 2023, 30, 62, doi:10.1051/parasite/2023064.

173. Doudoumis, V.; Blow, F.; Saridaki, A.; Augustinos, A.; Dyer, N.A.; Goodhead, I.; Solano, P.; Rayaisse, J.-B.; 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, doi:10.1038/s41598-017-04740-3.

174. Ramond, E.; Maclachlan, C.; Clerc-Rosset, S.; Knott, G.W.; Lemaitre, B. Cell division by longitudinal scission in the insect endosymbiont *Spiroplasma poulsonii*. mBio 2016, 7, e00881-16, doi:10.1128/mBio.00881-16.

175. Xie, J.; Butler, S.; Sanchez, G.; Mateos, M. Male killing *Spiroplasma* protects *Drosophila melanogaster* against two parasitoid wasps. Heredity 2014, 112, 399–408, doi:10.1038/hdy.2013.118.

176. Xie, J.; Vilchez, I.; Mateos, M. *Spiroplasma* bacteria enhance survival of *Drosophila hydei* attacked by the parasitic wasp *Leptopilina heterotoma*. PLoS One 2010, 5, e12149, doi:10.1371/journal.pone.0012149.

177. Higareda-Alvear, V.M.; Mateos, M.; Quezada, D.C.; Tamborindeguy, C.; Martínez-Romero, E. Transcriptomic response during wasp parasitism in the *Drosophila* - *Spiroplasma* interaction 2020. Preprint. doi:10.1101/2020.08.27.270462

178. Ballinger, M.J.; Gawryluk, R.M.R.; Perlman, S.J. Toxin and genome evolution in a *Drosophila* defensive symbiosis. Genome Biology and Evolution 2019, 11, 253–262, doi:10.1093/gbe/evy272.

179. Ballinger, M.J.; Perlman, S.J. The defensive *Spiroplasma*. Curr Opin Insect Sci 2019, 32, 36–41, doi:10.1016/j.cois.2018.10.004.

180. Harumoto, T.; Lemaitre, B. Male-killing toxin in a bacterial symbiont of *Drosophila*. Nature 2018, 557, 252–255, doi:10.1038/s41586-018-0086-2.

181. Harumoto, T.; Anbutsu, H.; Fukatsu, T. Male-killing *Spiroplasma* induces sex-specific cell death via host apoptotic pathway. PLoS Pathog 2014, 10, e1003956, doi:10.1371/journal.ppat.1003956.

182. Veneti, Z.; Bentley, J.K.; Koana, T.; Braig, H.R.; Hurst, G.D.D. A functional dosage compensation complex required for male killing in *Drosophila*. Science 2005, 307, 1461–1463, doi:10.1126/science.1107182.

183. Jiggins, F.M.; Hurst, G.D.D.; Jiggins, C.D.; V. D. Schulenburg, J.H.G.; Majerus, M.E.N. The butterfly *Danaus chrysippus* is infected by a male-killing *Spiroplasma* bacterium. Parasitology 2000, 120, 439–446, doi:10.1017/S0031182099005867.

184. Mouches, C.; Bové, J.M.; Tully, J.G.; Rose, D.L.; McCoy, R.E.; Carle-Junca, P.; Garnier, M.; Saillard, C. *Spiroplasma apis*, a new species from the honey-bee *Apis mellifera*. Ann Microbiol (Paris) 1983, 134A, 383–397.

185. Zha, G.; Yang, D.; Wang, J.; Yang, B.; Yu, H. Infection function of adhesin-like protein ALP609 from *Spiroplasma melliferum* CH-1. Curr Microbiol 2018, 75, 701–708, doi:10.1007/s00284-018-1435-y.

186. Kwon, M.-O.; Wayadande, A.C.; Fletcher, J. *Spiroplasma citri* movement into the intestines and salivary glands of its leafhopper vector, *Circulifer tenellus*. Phytopathology 1999, 89, 1144–1151, doi:10.1094/PHYTO.1999.89.12.1144.

187. Killiny, N.; Castroviejo, M.; Saillard, C. *Spiroplasma citri* spiralin acts *in vitro* as a lectin binding to glycoproteins from its insect vector *Circulifer haematoceps*. Phytopathology® 2005, 95, 541–548, doi:10.1094/PHYTO-95-0541.

188. Schneider, D.I.; Saarman, N.; Onyango, M.G.; Hyseni, C.; Opiro, R.; Echodu, R.; O'Neill, M.; Bloch, D.; Vigneron, A.; Johnson, T.J.; et al. Spatio-temporal distribution of *Spiroplasma* infections in the tsetse fly (*Glossina fuscipes fuscipes*) in northern Uganda. PLOS Neglected Tropical Diseases 2019, 13, e0007340, doi:10.1371/journal.pntd.0007340.

189. Abd-Alla, A.M.M.; Parker, A.G.; Vreysen, M.J.B.; Bergoin, M. Tsetse salivary gland hypertrophy virus: Hope or hindrance for tsetse control? PLoS Negl.Trop.Dis. 2011, 5, e1220-, doi:10.1371/journal.pntd.0001220.

190. Kariithi, H.; Van Oers, M.; Vlak, J.; Vreysen, M.; Parker, A.; Abd-Alla, A. Virology, epidemiology and pathology of *Glossina* hytrosavirus, and its control prospects in laboratory colonies of the tsetse fly, *Glossina pallidipes* (Diptera; *Glossinidae*). Insects 2013, 4, 287–319, doi:10.3390/insects4030287.

191. Abd-Alla, A.M.M.; Vlak, J.M.; Bergoin, M.; Maruniak, J.E.; Parker, A.G.; Burand, J.P.; Jehle, J.A.; Boucias, D.G. *Hytrosaviridae*: A proposal for classification and nomenclature of a new insect virus family. Arch.Virol. 2009, 154, 909–918, doi:10.1007/s00705-009-0398-5.

192. Mutika, G.N.; Marin, C.; Parker, A.G.; Boucias, D.G.; Vreysen, M.J.B.; Abd-Alla, A.M.M. Impact of salivary gland hypertrophy virus infection on the mating success of male *Glossina pallidipes*: Consequences for the sterile insect technique. PLoS ONE 2012, 7, e42188, doi:10.1371/journal.pone.0042188.

193. Jura, W.G.Z.O.; Otieno, L.H.; Chimtawi, M.M.B. Ultrastructural evidence for trans-ovum transmission of the DNA virus of tsetse, *Glossina pallidipes* (Diptera: *Glossinidae*). Curr.Microbiol. 1989, 18, 1–4, doi:10.1007/BF01568821.

194. Sang, R.C.; Jura, W.G.Z.O.; Otieno, L.H.; Ogaja, P. Ultrastructural changes in the milk gland of tsetse *Glossina morsitans centralis* (Diptera; *Glossinidae*) female infected by a DNA virus. J.Invertebr.Pathol. 1996, 68, 253–259, doi:10.1006/jipa.1996.0093.

195. Feldmann, U. Guidelines for the rearing of tsetse flies using the membrane feeding technique. In Techniques of insect rearing for the development of integrated pest and vector management strategies; Ochieng'-Odero, J.P.R., Ed.; ICIPE Science Press: Nairobi, Kenya, 1994; Vol. 2, pp. 449–471 ISBN 92 9064 065 X.

196. Abd-Alla, A.M.M.; Kariithi, H.M.; Parker, A.G.; Robinson, A.S.; Kiflom, M.; Bergoin, M.; Vreysen, M.J.B. Dynamics of the salivary gland hypertrophy virus in laboratory colonies of *Glossina pallidipes* (Diptera: *Glossinidae*). Virus Res 2010, 150, 103–110, doi:10.1016/j.virusres.2010.03.001.

197. Abd-Alla, A.M.M.; Kariithi, H.M.; Mohamed, A.H.; Lapiz, E.; Parker, A.G.; Vreysen, M.J.B. Managing hytrosavirus infections in *Glossina pallidipes colonies*: Feeding regime affects the prevalence of salivary gland hypertrophy syndrome. PLoS ONE 2013, 8, e61875, doi:10.1371/journal.pone.0061875.

198. Meki, I.K.; Huditz, H.-I.; Strunov, A.; Van Der Vlugt, R.A.A.; Kariithi, H.M.; Rezapanah, M.; Miller, W.J.; Vlak, J.M.; Van Oers, M.M.; Abd-Alla, A.M.M. Characterization and tissue tropism of newly identified Iflavirus and Negeviruses in *Glossina morsitans morsitans* tsetse flies. Viruses 2021, 13, 2472, doi:10.3390/v13122472.

199. Oers, M.M. van Genomics and biology of Iflaviruses. In; Asgari, S., Johnson, K.N., Eds.; Caister Academic Press: Norfolk, UK, 2010; pp. 231–250 ISBN 978-1-904455-71-4.

200. Murakami, R.; Suetsugu, Y.; Kobayashi, T.; Nakashima, N. The genome sequence and transmission of an Iflavirus from the brown planthopper, *Nilaparvata lugens*. Virus Research 2013, 176, 179–187, doi:10.1016/j.virusres.2013.06.005.

201. Carballo, A.; Williams, T.; Murillo, R.; Caballero, P. Iflavirus covert infection increases susceptibility to nucleopolyhedrovirus disease in *Spodoptera exigua*. Viruses 2020, 12, 509, doi:10.3390/v12050509.

202. Ryabov, E.V.; Wood, G.R.; Fannon, J.M.; Moore, J.D.; Bull, J.C.; Chandler, D.; Mead, A.; Burroughs, N.; Evans, D.J. A virulent strain of deformed wing virus (DWV) of honeybees (*Apis mellifera*) prevails after *Varroa destructor*-mediated, or *in vitro*, transmission. PLoS Pathog 2014, 10, e1004230, doi:10.1371/journal.ppat.1004230.

203. Ryabov, E.V.; Fannon, J.M.; Moore, J.D.; Wood, G.R.; Evans, D.J. The Iflaviruses sacbrood virus and deformed wing virus evoke different transcriptional responses in the honeybee which may facilitate their horizontal or vertical transmission. PeerJ 2016, 4, e1591, doi:10.7717/peerj.1591.

204. Lanzi, G.; De Miranda, J.R.; Boniotti, M.B.; Cameron, C.E.; Lavazza, A.; Capucci, L.; Camazine, S.M.; Rossi, C. Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). J Virol 2006, 80, 4998–5009, doi:10.1128/JVI.80.10.4998-5009.2006.

205. Mirieri, C.K.; Abd-Alla, A.M.M.; Ros, V.I.D.; Van Oers, M.M. Evaluating the effect of irradiation on the densities of two RNA viruses in *Glossina morsitans morsitans*. Insects 2023, 14, 397, doi:10.3390/insects14040397.

206. Selkoe, K.A.; Toonen, R.J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. Ecology Letters 2006, 9, 615–629, doi:10.1111/j.1461-0248.2006.00889.x.

207. Geiger, A.; Ravel, S.; Mateille, T.; Janelle, J.; Patrel, D.; Cuny, G.; Frutos, R. Vector competence of *Glossina palpalis gambiensis* for *Trypanosoma brucei* s.l. and genetic diversity of the symbiont *Sodalis glossinidius*. Mol.Biol.Evol. 2007, 24, 102–109, doi:10.1093/molbev/msl135.

208. Soumana, I.H.; Simo, G.; Njiokou, F.; Tchicaya, B.; Abd-Alla, A.M.M.; Cuny, G.; Geiger, A. The bacterial flora of tsetse fly midgut and its effect on trypanosome transmission. J. Invertebr. Pathol. 2013, 112, S89–S93, doi:10.1016/j.jip.2012.03.029.

209. Kanté Tagueu, S.; Farikou, O.; Njiokou, F.; Simo, G. Prevalence of *Sodalis glossinidius* and different trypanosome species in *Glossina palpalis palpalis* caught in the Fontem sleeping sickness focus of the Southern Cameroon. Parasite 2018, 25, doi:10.1051/parasite/2018044.

210. Dennis, J.W.; Durkin, S.M.; Horsley-Downie, J.E.; Hamill, L.C.; Anderson, N.E.; MacLeod, E.T. *Sodalis glossinidius* prevalence and trypanosome presence in tsetse from Luambe National Park, Zambia. Parasit.Vectors 2014, 7, 378, doi:10.1186/1756-3305-7-378.

211. Yeh, S.-D.; Do, T.; Abbassi, M.; Ranz, J.M. Functional relevance of the newly evolved sperm dynein intermediate chain multigene family in *Drosophila melanogaster* males. Communicative & Integrative Biology 2012, 5, 462–465, doi:10.4161/cib.21136.

Curriculum vitae

Name	Fabian Gstöttenmayer
Address	[redacted]
Mobile phone number	[redacted]
E-Mail-Address	[redacted]
Date of birth	[redacted]

Education

2021 – now

Doctorate program – Technical University Vienna

Natural Sciences

Doctorate thesis: Host-symbiont interactions and population genetics of the tsetse fly, *Glossina* spp.

Degree: Dr. rer. nat. to be obtained 09/2024.

Key competences: microbiology, molecular biology techniques, population genetics, DNA sequencing and bioinformatics.

2017 – 2019

Master program – FH Technikum Wien

Environmental management and Toxicology

Master thesis: Population genetics of the invasive bee *Megachile sculpturalis* via Genotyping-by-Sequencing.

Degree: MSc

2013 – 2017

Bachelor program – University of Natural Resources and Life Sciences Vienna

Environment and Bio-Resources Management

Degree: BSc

Professional experience

December 2021 – now

International Atomic Energy Agency

Insect Pest Control Laboratory, Livestock Pest Group

PhD Consultancy in population genetics of *G. brevipalpis* to assist vector control programs, interactions of tsetse flies and bacterial symbiont *Spiroplasma,* supervised by Prof Adly Abdalla.

June 2021 – November 2021

Medical University Vienna

Center for Anatomy and Cell Biology

Genome Dynamics Group

Technical Assistance in studies on *Wolbachia* in *Drosophila*; Laboratory management; supervised by Prof Wolfgang Miller and Dr Martin Kapun.

July 2020 – May 2021

International Atomic Energy Agency

Insect Pest Control Laboratory, Livestock Pest Group

Internship for development of microsatellite markers to study tsetse population genetics, host-symbiont interactions, supervised by Prof Adly Abdalla.

March 2019 – June 2019

University of Natural Resources and Life Sciences Vienna

Institute for Nature Conservation Research

Internship to perform laboratory work for master thesis, genotyping invasive bee *Megachile sculpturalis* via SSR-GBS.

July 2015, 2016

Austrian Agency for Health and Food Safety

Institute for Seeds and Seedlings, Plant Protection and Bee Health.

Internship in seed viability testing and bee health monitoring.

Skills

Languages	German, native language. English, C level plus scientific terminology. Basic knowledge in French, Spanish and Latin.
Laboratory techniques	Molecular biology techniques, microscopy, microbiology techniques, sequencing library preparation.
Data science	Data exploration, visualization and statistics in R Studio Bioinformatics: bash scripting, data mining, processing of sequencing data, genome assembly, metagenomics, transcriptomics, population genetics analyses.
Driving license	EU license class B.

Personal interests

Beekeeping	since 2010, apiary with currently 15 beehives in Vienna.
Travelling	bike tours, sailing, hiking.
Music	Guitar and bass guitar, hosting and attending concerts.

Publication record

Researchgate: Fabian Gstöttenmayer

ORCID: 0000-0003-1908-7655

2023

Gstöttenmayer, F., Moyaba, P., Rodriguez, M., Mulandane, F.C., Mucache, H.N., Neves, L., De Beer, C., Ravel, S., De Meeûs, T., Mach, R.L., Vreysen, M.J.B., Abd-Alla, A.M.M., 2023. Development and characterization of microsatellite markers for the tsetse species Glossina brevipalpis and preliminary population genetics analyses. Parasite 30, 34. https://doi.org/10.1051/parasite/2023038.

2022

Dieng, M.M., Dera, K.M., Moyaba, P., Ouedraogo, G.M.S., Demirbas-Uzel, G., **Gstöttenmayer, F.**, Mulandane, F.C., Neves, L., Mdluli, S., Rayaisse, J.-B., Belem, A.M.G., Pagabeleguem, S., De Beer, C.J., Parker, A.G., Van Den Abbeele, J., Mach, R.L., Vreysen, M.J.B., Abd-Alla, A.M.M., 2022. Prevalence of Trypanosoma and Sodalis in wild populations of tsetse flies and their impact on sterile insect technique programmes for tsetse eradication. Sci Rep 12, 3322. https://doi.org/10.1038/s41598-022-06699-2.

2021

Son, J.H., Weiss, B.L., Schneider, D.I., Dera, K.M., **Gstöttenmayer, F.**, Opiro, R., Echodu, R., Saarman, N.P., Attardo, G.M., Onyango, M., Abd-Alla, A.M.M., Aksoy, S., 2021. Infection with endosymbiotic Spiroplasma disrupts tsetse (Glossina fuscipes fuscipes) metabolic and reproductive homeostasis. PLoS Pathog 17, e1009539. https://doi.org/10.1371/journal.ppat.1009539.

Lanner, J., **Gstöttenmayer, F.**, Curto, M., Geslin, B., Huchler, K., Orr, M.C., Pachinger, B., Sedivy, C., Meimberg, H., 2021. Evidence for multiple introductions of an invasive wild bee species currently under rapid range expansion in Europe. BMC Ecol Evo 21, 17. https://doi.org/10.1186/s12862-020-01729-x.