# Prevalence and associations of *Trypanosoma spp*. and *Sodalis glossinidius* with intrinsic factors of tsetse flies

Manun Wongserepipatana

This thesis is submitted in part fulfilment of the requirements for the Degree of Doctor of Philosophy.

Institute of Biodiversity, Animal Health and Comparative Medicine
College of Medical, Veterinary and Life Sciences
University of Glasgow

August 2016

## Abstract

Trypanosomiasis has been identified as a neglected tropical disease in both humans and animals in many regions of sub-Saharan Africa. Whilst assessments of the biology of trypanosomes, vectors, vertebrate hosts and the environment have provided useful information about life cycles, transmission, and pathogenesis of the parasites that could be used for treatment and control, less information is available about the effects of interactions among multiple intrinsic factors on trypanosome presence in tsetse flies from different sites. It is known that multiple species of tsetse flies can transmit trypanosomes but differences in their vector competence has normally been studied in relation to individual factors in isolation, such as: intrinsic factors of the flies (e.g. age, sex); habitat characteristics; presence of endosymbionts (e.g. *Wigglesworthia glossinidia*, *Sodalis glossinidius*); feeding pattern; host communities that the flies feed on; and which species of trypanosomes are transmitted. The purpose of this study was to take a more integrated approach to investigate trypanosome prevalence in tsetse flies. In **chapter 2**, techniques were optimised for using the Polymerase Chain Reaction (PCR) to identify species of trypanosomes (*Trypanosoma vivax*, *T. congolense*, *T. brucei*, *T. simiae*, and *T. godfreyi*) present in four species of tsetse flies (*Glossina austeni*, *G. brevipalpis*, *G. longipennis* and *G. pallidipes*) from two regions of eastern Kenya (the Shimba Hills and Nguruman). Based on universal primers targeting the internal transcribed spacer 1 region (ITS-1), *T. vivax* was the predominant pathogenic species detected in flies, both singly and in combination with other species of trypanosomes. Using Generalised Linear Models (GLMs) and likelihood ratio tests to choose the best-fitting models, presence of *T. vivax* was significantly associated with an interaction between subpopulation (a combination between collection sites and species of *Glossina*) and sex of the flies ($\chi^2$ = 7.52, df = 21, P-value = 0.0061); prevalence in females overall was higher than in males but this was not consistent across subpopulations. Similarly, *T. congolense* was significantly associated only with subpopulation ($\chi^2$ = 18.77, df = 1, P-value = 0.0046); prevalence was higher overall in the Shimba Hills than in Nguruman but this pattern varied by species of tsetse fly. When associations were analysed in individual species of tsetse flies, there were no consistent associations between trypanosome prevalence and any single factor (site, sex, age) and different

combinations of interactions were found to be significant for each. The results thus demonstrated complex interactions between vectors and trypanosome prevalence related to both the distribution and intrinsic factors of tsetse flies. The potential influence of the presence of *S. glossinidius* on trypanosome presence in tsetse flies was studied in **chapter 3**. A high number of *Sodalis* positive flies was found in the Shimba Hills, while there were only two positive flies from Nguruman. Presence or absence of *Sodalis* was significantly associated with subpopulation while trypanosome presence showed a significant association with age ($\chi^2$ = 4.65, df = 14, P-value = 0.0310) and an interaction between subpopulation and sex ($\chi^2$ = 18.94, df = 10, P-value = 0.0043). However, the specific associations that were significant varied across species of trypanosomes, with *T. congolense* and *T. brucei* but not *T. vivax* showing significant interactions involving *Sodalis*. Although it has previously been concluded that presence of *Sodalis* increases susceptibility to trypanosomes, the results presented here suggest a more complicated relationship, which may be biased by differences in the distribution and intrinsic factors of tsetse flies, as well as which trypanosome species are considered. In **chapter 4** trypanosome status was studied in relation to blood meal sources, feeding status and feeding patterns of *G. pallidipes* (which was the predominant fly species collected for this study) as determined by sequencing the mitochondrial cytochrome B gene using DNA extracted from abdomen samples. African buffalo and African elephants were the main sources of blood meals but antelopes, warthogs, humans, giraffes and hyenas were also identified. Feeding on multiple hosts was common in flies sampled from the Shimba Hills but most flies from Nguruman had fed on single host species. Based on Multiple Correspondence Analysis (MCA), host-feeding patterns showed a correlation with site of sample collection and *Sodalis* status, while trypanosome status was correlated with sex and age of the flies, suggesting that recent host-feeding patterns from blood meal analysis cannot predict trypanosome status. In conclusion, the complexity of interactions found suggests that strategies of tsetse fly control should be specific to particular epidemic areas. Future studies should include laboratory experiments that use local colonies of tsetse flies, local strains of trypanosomes and local *S. glossinidius* under controlled environmental conditions to tease out the factors that affect vector competence and the relative influence of external environmental factors on the dynamics of these interactions.

# List of Contents

# List of Tables

# List of Figures

# Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Barbara Mable for the continuous support of my Ph.D. study and related research, for her patience, motivation and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study. I would also like to thank Prof. Mike Turner, who was my supervisor for the first year. He was always available for questions, discussion and suggestion. Importantly he gave me an opportunity to know my great supervisor. I am very grateful for the opportunity to undertake this project, which was funded by the Thai government following a requirement of Rajamangala University of Technology Tawan-ok.

Besides these, I would like to thank the rest of my thesis committee: Dr. Lisa Ranford Cartwright, Dr. Harry de Koning, Prof. David Eckersall and Dr. Jane Robinson for their insightful comments and encouragement, but also for the hard questions which incented me to widen my research from various perspectives.

My sincere thanks also goes to Dr. Marc Ciosi for providing tsetse flies samples and laboratory support and Dr. James Buckley for advisement about the R statistic programs. I thank Kathleen Scullion and Alana Hamilton from the Institute of Infection, Immunity and Inflammation, who gave access to the laboratory and research facilities and provided technical support. I am very thankful to Aileen Adam and Elizabeth Kilbride for providing laboratory facilities and very for useful suggestions. A lot of biscuits, chocolates and very strong coffees from the roof laboratory are meaningful for my Ph.D. life. Without all their precious support it would not be possible to conduct this research.

Last but not the least, a huge to thank Dr. Tuonta Chansilpa, Dean of Faculty of Veterinary Medicine, Rajamangala University of Technology Tawan-Ok for all her supporting. I would like to thanks a lot for all the support and encouragement from my best friend, Vachira Hunprasit. Your advisement relating to statistics was really helpful for my thesis. Thanks also to my lovely family for their endless support. Finally, but most importantly, I want to thank my husband, Sittiporn Channumsin, who always stands side by side with me throughout laboratory experiment, data analysis, writing this thesis and my general life.

## Author's Declaration

The work that is described in this thesis was carried out in the laboratory of the Institute of Infection, Immunity and Inflammation (III) and the Institute of Biodiversity, Animal Health and Comparative Medicine (BAHCM), University of Glasgow between October 2012 and March 2016. I hereby declared that this thesis is my own and is of my own composition. No part of this thesis has been submitted for another degree.

Manun Wongserepipatana

Glasgow, August 2016

# Abbreviations and symbols

| | |
|---|---|
| A | Ampere |
| AAT | Animal African Trypanosomiasis |
| AB | abdomen |
| AMP | Anti-microbial peptides |
| bp | base pairs |
| BLAST | Basic local alignment search tool |
| BR | the Buffalo Ridge site |
| BRGb | the *Glossina brevipalpis* from the Buffalo Ridge site |
| BRGp | the *Glossina pallidipes* from the Buffalo Ridge site |
| CAT | Canine African Trypanosomiasis |
| CDC | Central Disease Control |
| CI | cytoplasmic incompatibility |
| COI | cytochrome c oxidase I |
| cytb | cytochrome b |
| $ddH_2O$ | double distilled water |
| DNA | Deoxyribonucleic acid |
| dNTPs | Deoxynucleoside triphosphate |
| d.p.e. | day post exposure |
| EDTA | Ethylene diamine tetra-acetic acid |
| ELISA | Enzyme-linked immunosorbent assay |
| EP | glutamic acid-proline protein |
| ES | ectoperitrophic space |

| | |
|---|---|
| FAO | Food and Agriculture Organisation |
| FTA | Flinders Technology Associates |
| Ga | *Glossina austeni* |
| Gb | *Glossina brevipalpis* |
| GLMs | Generalised Linear Models |
| Gl | *Glossina longipennis* |
| Gp | *Glossina pallidipes* |
| h | hour |
| h.p.e. | hours post-eclosion |
| HAT | Human African Trypanosomiasis |
| HP | head plus proboscis |
| IAEA | International Atomic Energy Agency |
| IFAT | indirect fluorescent antibody test |
| IFN-γ | interferon gamma |
| IL | Interleukin |
| ITS | Internal transcribed spacers |
| LAMP | Loop-mediated isothermal amplification |
| m | $10^{-3}$ |
| min | minute |
| mm | millimetre |
| mtDNA | mitochrondrial DNA |
| Mu | Mukinyo site |
| MuGp | *Glossina pallidipes* from the Mukinyo site |

| | |
|---|---|
| MuGl | *Glossina longipennis* from the Mukinyo site |
| MCA | Multiple Correspondence Analysis |
| NG | the Nguruman region |
| OIE | World Organization for Animal Health |
| OR | Odds Ratio |
| PBS | Phophase buffer saline |
| PCR | Polymerase chain Reaction |
| PM | peritrophic matrix |
| ROS | reactive oxygen species |
| RLO | rickettsia-like organism |
| Sa | Sampu site |
| SaGl | *Glossina longipennis* from the Sampu site |
| sec | second |
| SH | the Shimba Hills region |
| SSU | small sub-unit |
| LSU | large subunit |
| *Taq* | *Thermus aquaticus* |
| Tb | *Trypanosoma brucei* |
| TBE | Tris-borate ethylene diamine tetra acetic acid |
| Tc | *Trypanosoma congolense* |
| Tcs | *Trypanosoma congolense* savannah |
| Tck | *Trypanosoma congolense* kilifi |
| Tcf | *Trypanosoma congolense* forest |

| | |
|---|---|
| Tg | *Trypanosoma godfreyi* |
| TH | thorax |
| TRC | Trypanosomiasis Research Centre |
| Ts | *Trypanosoma simiae* |
| TNF-α | tumor necrotic factor α |
| Tv | *Trypanosoma vivax* |
| UV | Ultra violet |
| V | voltage |
| VSG | variant surface glycoprotein |
| WHO | World Health Organization |
| Zu | Zungu Luka site |
| ZuGl | *Glossina longipennis* from the Zungu Luka site |
| ZuGp | *Glossina pallidipes* from the Zungu Luka site |
| °C | degree Celsius |
| μ | $10^{-6}$ |

# Chapter 1 Introduction

## 1.1 Biology of *Trypanosoma spp*.

### 1.1.1 Taxonomy of trypanosomes

| | |
|---|---|
| Kingdom | Protista |
| Subkingdom | Protozoa |
| Phylum | Sarcomastigophora |
| Subphylum | Mastigophora |
| Class | *Zoomastigophora* |
| Order | *Kinetoplastida* |
| Family | *Trypanosomatidae* |
| Genus | *Trypanosoma* |
| Section | Stercoraria and Salivaria |
| Subgenera | *Duttonella, Nannomonas* and *Trypanozoon* |

Trypanosomes are classified in the order *Kinetoplastida*, family *Trypanosomatidae* (Maudlin *et al.*, 2004). All trypanosomes belong to the genus *Trypanosoma* and these parasites then have been separated by modes of transmission to be two groups (Figure 1.1). The first group is Stercoraria, which is distributed in the Americas: *T. theileri, T. lewisi,* and *T. cruzi* are members of this section because they transmit via faeces of vectors to vertebrate hosts. *Trypanosoma cruzi* causes American trypanosomosis in humans or "Chagas disease", which transmits from mammals, birds, reptiles and amphibians to humans by the triatomine bug (*Rhodnius pallescens*) as the principal vector in South America (Gottdenker *et al.*, 2012). The second group is Salivarian trypanosomes, which infect vertebrates when they are bitten by insect vectors. Salivarian trypanosomes or African Trypanosomes, which I focus on in this study, have been further divided into three subgenera based on different morphological characteristics: *Duttonella*, *Nannomonas,* and *Trypanozoon*. The principal species of trypanosome in *Duttonella* is *T. vivax*, which is a typical blood-stream form with an enormous kinetoplast in the posterior part of the body. *Nannomonas*, which is characterized by a small body (8 - 24 μm length) and medium sized kinetoplast at the end of the body, has *T. congolense* as a high pathogenicity member. The last subgenus is *Trypanozoon*, which is indicated by

distributions, host and vector range and genetics. *Trypanosoma brucei brucei*, *T. brucei rhodesiense* and *T. brucei gambiense* are the main pathogenic species in the subgenus. *Trypanosoma brucei rhodesiense* and *T. brucei gambiense* are causes of African trypanosomiasis in humans. For animals, the most important pathogenic species, which are the main cause of African trypanosomiasis, are *T. congolense*, *T. vivax* and *T. brucei*, which each have a very different lifecyle.

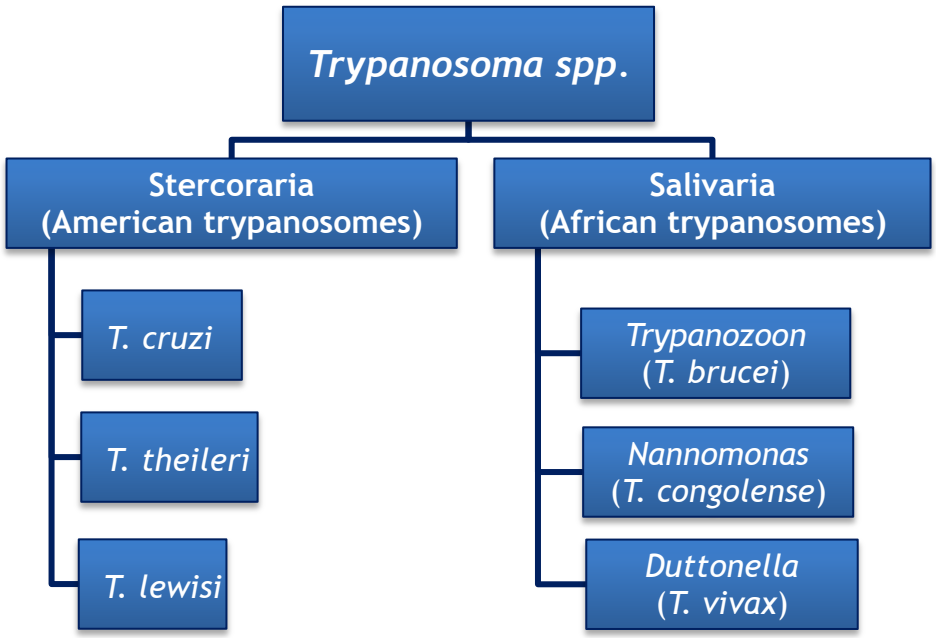



**Figure 1.1** Relationships among species of trypanosomes in the Stercoraria and Salivaria sections

Stercoraria are found in South America whereas Salivaria are found only in sub-Saharan Africa.

## 1.1.2 Life cycle of African trypanosomes

### 1.1.2.1 Life cycle of *Trypanosma brucei*

The trypanosome life cycle has been described for *T. brucei* (Figure 1.2). *Trypanosoma brucei* is an obligate extracellular blood parasitic protozoon, which has a complex life cycle in specific environments. There are six morphologically distinct stages between midgut and salivary glands of the vectors and body fluid of the vertebrate hosts: long and slender forms; stumpy forms; procyclics; mesocyclics; epimastigotes; and metacyclics (Table 1.1) (Brun *et al.*, 2009).

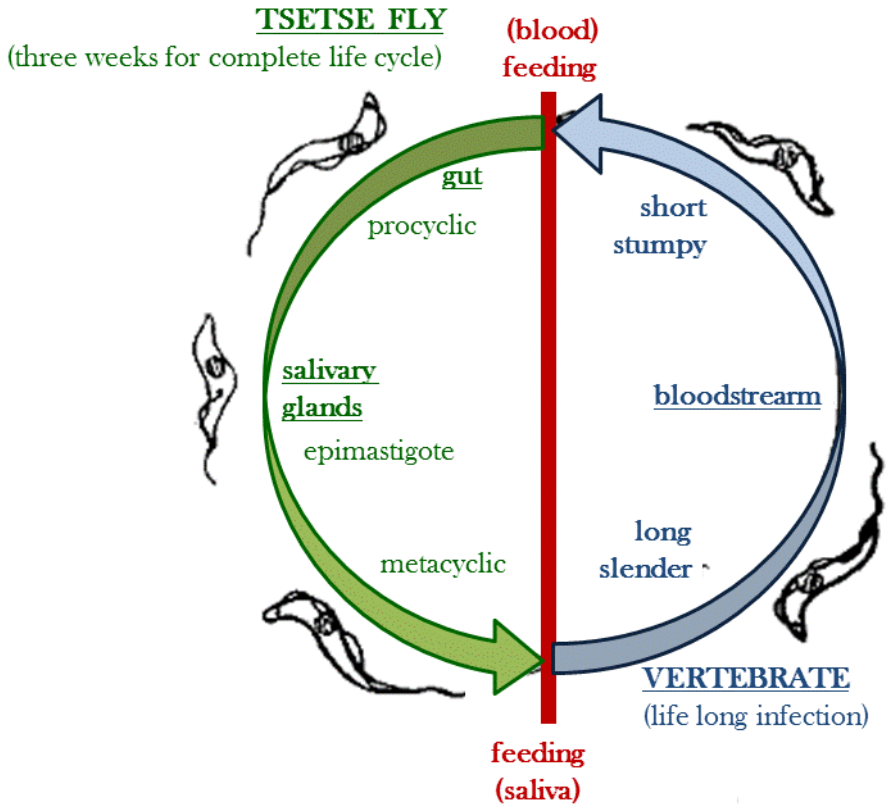
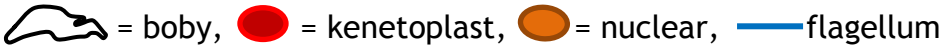



**Figure 1.2** The life cycle of *T. brucei* in vertebrate hosts and tsetse flies.

**Table 1.1** Morphological development of trypanosomes in vertebrate hosts and tsetse flies

| Stages of trypanosomes | Morphology of trypanosomes | Hosts |
|---|---|---|
| long and slender forms |  | vertebrate hosts |
| stumpy forms |  | vertebrate hosts |
| procyclics |  | tsetse flies |
| mesocyclics |  | tsetse flies |
| epimastigotes |  | tsetse flies |
| metacyclics |  | tsetse flies |



Vertebrate hosts are infected by the bite of a trypanosome-infected tsetse fly, the biological vectors. Metacyclic forms of trypomastigotes in saliva are inoculated into the hosts, where they transform to bloodstream forms that have a long and slender morphology (Vickerman, 1985). At the subdermal biting site, trypomastigotes divide by binary fission. Then they move to the blood circulation via the lymphatic system. In the bloodstream and lymphatic vessels, parasites continue to increase by binary fission and also differentiate to non-dividing short stumpy forms that are the infective stage for tsetse flies if taken up in a blood meal.

In the digestive system of tsetse flies, five of the main morphological types of trypanosomes occur: stumpy forms, procyclics, mesocyclics, epimastigotes and metacyclics (Figure 1.2). Tsetse flies ingest stumpy forms during feeding on blood meals from infected vertebrate hosts (Vickerman, 1985). In the midgut, the parasites transform to proliferative procyclic trypomastigotes within 24 hr (Peacock *et al.*, 2012a) and multiply by longitudinal binary fission at 3 days post exposure (d.p.e.) (Van Den Abbeele *et al.*, 1999). During 3 - 6 d.p.e., proliferative procyclic forms move across the peritrophic matrix (PM) and proliferate in the ectoperitrophic space (ES) (Peacock *et al.*, 2012a, Peacock *et al.*, 2012b). However, different studies have suggested different durations of trypanosome development. Parasites migrate as mesocyclics through the proventiculus, which has been reported to occur in *G. morsitans morsitans* from 6 d.p.e. (Van Den Abbeele *et al.*, 1999) to 28 d.p.e. (Peacock *et al.*, 2012b). However, in another study, Peacock *et al* (2012a) found trypanosomes in the proventiculus 7 – 10 d.p.e., with movement to the esophagous and the hypophalynx 8 – 20 d.p.e. (Peacock *et al.*, 2012a). Finally, *T. brucei* move to the salivary glands (Aksoy, 2003), where they transform into epimastigotes and use their flagella to attach themselves to the endothelum of the salivary glands. They then differentiate into non-dividing metacyclic trypomastigotes that are infective to mammalian hosts. Metacyclic trypomastigotes are detected from day 12 d.p.e. (Peacock *et al.*, 2012b). However, large variation in time taken for tsetse flies to complete their lifecycle has been reported: 12 days (Peacock *et al.*, 2012b), 16 - 80 days (Peacock *et al.*, 2012a), 16 - 35 days (Franco *et al.*, 2014) and 34 days (Bruce *et al.*, 1910). Since trypanosome infection is permanently harboured for the lifespan of tsetse flies (Franco *et al.*, 2014), this

large discrepancy in ages could affect interpretation of relative transmission risks.

#### 1.1.2.2 Life cycle of *Trypanosoma congolense*

The life cycle of *T. congolense* is similar to *T. brucei* in that all bloodstream forms of trypanosomes have the same morphology and are able to infect tsetse flies, but epimastigotes attach to the mouth parts of tsetse flies rather than in the salivary glands, where they transform into infective metacyclics in the proboscis (Aksoy, 2003, Peacock *et al.*, 2012a) (Figure 1.3). *Trypanosoma congolense* take approximately 21 days to complete their lifecycle in tsetse flies (Peacock *et al.*, 2012a), which is shorter than for *T. brucei*. *Trypansoma congolense* establish and develop in the migut of tsetse flies approximately 2 – 7 d.p.e. then move to the proventiculus at 6 – 12 d.p.e. In the foregut, trypomastigotes, epimastigotes and metacyclics are detected at 10, 15 and 21 d.p.e., respectively. In *G. m. morsitans* that fed on *T. congolense* infective blood meals for their first feed 24 – 48 hours post-eclosion (h.p.e.), the parasites were first found in proboscis parts at 13 days under controlled conditions (at 25°C and 70% relative humidity) (Peacock *et al.*, 2012a).

#### 1.1.2.3 Life cycle of *Trypanosoma vivax*

After mammalian hosts are bitten by tsetse flies, mature metacyclic forms are transferred to the animals. They then differentiate to be slender forms and stumpy forms. *Trypanosoma vivax* was detected in the blood circulation of experimentally infected goats at 5 days post inoculation (Batista *et al.*, 2011). This is different from the development of *T. brucei*, for which the developmental process is entirely restricted to the proboscis of the flies (Vickerman, 1973) (Figure 1.3); slender forms are transmitted via feeding then differentiate to procyclic, epimastigote and metacyclic forms in the mouth parts (OsóriOl *et al.*, 2008). *Glossina palpalis*, which were laboratory developed offspring of flies captured in Tanzania and fed on *T. vivax* infected goats and cattle, became infective at day 7 – 21 after exposure using microscopical examination (Bruce *et al.*, 1910).

The biological vectors of *T. brucei*, *T. congolense* and *T. vivax* are tsetse flies, in which cyclical development takes place, but tsetse flies (Taylor, 1930, Roberts

*et al*., 1989) and other blood sucking insects (Desquesnes and Dia, 2003a, Mihok *et al*., 1995, Sumba *et al*., 1998, Taylor, 1930) are able to carry these trypanosomes as mechanical vectors for approximately 30 min (Chinery, 1965). Stable flies (*Stomoxy spp*.) in the Central African Republic (D'Amico *et al*., 1996), and African tabanids (*Atylotus agrestis* and *A. fuscipes*) in Burkina Faso (Desquesnes and Dia, 2003b, Desquesnes and Dia, 2004) have been reported as mechanical vectors of *T. vivax*. Distribution of *T. vivax* by these haemotophagous biting insects clearly occurs in the apparent absence of tsetse flies (Dagnachew and Bezei, 2015); for example, *T. vivax* infection in goats in west of Santa Catarina state, Brazil (Fávero *et al*., 2016). There were also reports of *T. congolense* and *T. brucei* transmission by other mechanical vectors in Africa; *Stomoxys niger*, *Stomoxys taeniatus* (Sumba *et al*., 1998) and *Atylotus agrestis* (Desquesnes and Dia, 2003a) for *T. congolense* while *Stomoxy calcitrans* carried *T. brucei* (Taylor, 1930).

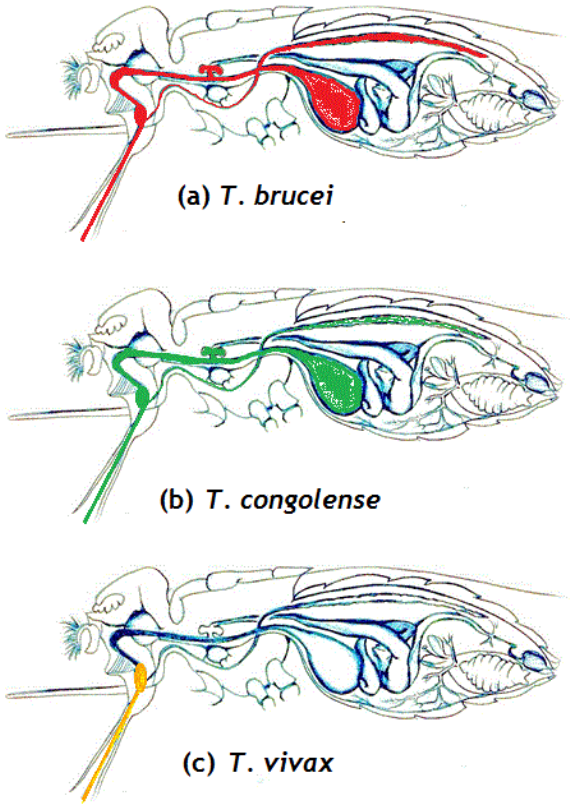



**Figure 1.3** Location of tsetse-fly tissues relating to development of *T. brucei* (a), *T. congolense* (b) and *T. vivax* (c).

Life cycles of *T. brucei* and *T. congolense* distribute in the proboscis, foregut, midgut and salivary glands while the lifecycle and development of *T. vivax* occurs only in the proboscis parts.

### 1.1.3 Importance of African trypanosomiasis

Trypanosomes are extracellular blood parasites, which are pathogens for both humans and animals. For humans in Africa, the distribution of trypanosomes is across the topical zone, usually in rural areas and far away from health services. A World Health Organization (WHO) report in 2016 declared that Human African Trypanosomiasis (HAT), also known as Sleeping Sickness, occurs in 36 countries in sub-Saharan Africa, including: Democratic Republic of the Congo, Nigeria, Cameroon, Sudan, Central African Republic, Ethiopia, Uganda, Kenya, Tanzania, Zambia, Zimbabwe and South Africa (World Health Organization, 2016b). HAT can be a cause of fatality and suffering from a variety of symptoms, for example, headache, weight loss, fatigue, intermittent fever, insomnia, tremors, abnormal movement and ultimately death (Kennedy, 2013). *Trypanosoma b. rhodesiense* causes acute HAT (the symptoms are detected within 1 – 3 weeks) (Chappuis *et al.*, 2005a) and is normally fatal within six months (Odiit *et al.*, 1997). *Trypanosoma b. gambiense* causes chronic HAT (the symptoms are detected for many months to years) (Chappuis *et al.*, 2005a, Marcello and Barry, 2007) and is normally fatal within three years (Checchi *et al.*, 2008). Among HAT cases, chronic HAT, which occurs in western and central Africa, is more commonly reported (98.0% of all trypanosome reported cases) than acute HAT (2.0% of all trypanosome reported cases), which is distributed in the eastern and southern parts of Africa (Figure 1.4) (World Health Organization, 2016b). Reported cases of both types of HAT from WHO suggest that since 1990, the number of cases was highest in 1998 (37,991 cases), with a further 300,000 estimated cases not diagnosed and treated. Since then, the number of HAT reported cases has continuously decreased and 2009 was the first time in 20 years that numbers of reported cases decreased below 10,000, after continued control efforts since 1920. This decline in numbers of cases has continued, with only 3,796 new cases reported in 2014. However, HAT is a health problem, particularly in specific regions of Africa, where the disease has been remaining (Bonnet *et al.*, 2015, Simarro *et al.*, 2010).

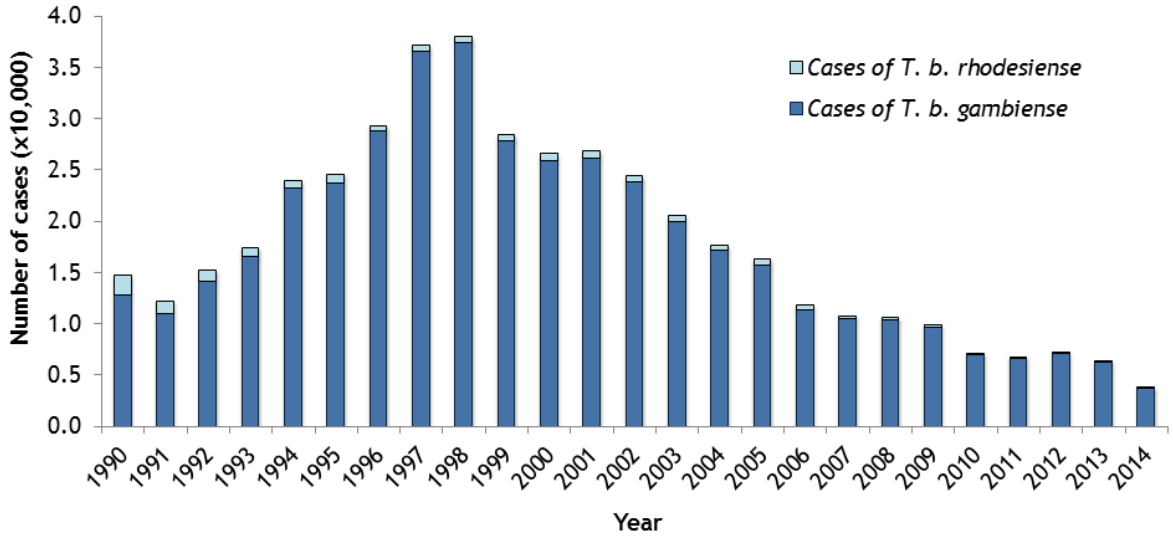



**Figure 1.4** Number of reported cases of sleeping sickness (HAT) in 1990–2014. Data derived from World Health Organization, 2016,http://apps.who.int/gho/data/node.main.A1635)

Trypanosome-infected animals can have intermittent fever, anaemia, oedema, poor condition, infertility (Wedman, 1980) and can suffer abortions (Okech *et al.*, 1996, Gutierrez *et al.*, 2005). However, wild animals have been reported as reservoir hosts of trypanosomes, for which clinical signs are undetectable (World Organisation for Animal Health, 2013). Widespread problems of Animal African Trypanosomiasis (AAT) are due to their effects on domesticated animals (Maudlin *et al.*, 2004). Many species of trypanosomes are pathogens causing a serious disease of livestock; for example, *T. congolense*, *T. vivax* and *T. b. brucei*, *T. simiae, T. godfreyi, T. suis, T. equiperdum, T. evansi*. A complex of *T. congolense*, *T. vivax* and *T. b. brucei* are the main pathogenic species collectively causing Animal African Trypanosomiasis (AAT) and "Nagana disease", which is what trypanosomiasis is referred to as in cattle. AAT has been reported in areas of Africa wherever tsetse flies exist: between latitudes 15°N and 29°S, from the southern edge of the Sahara desert to Zimbabwe, Angola and Mozambique (World Organisation for Animal Health, 2013). However, *T. vivax*, which also mechanically transmits though biting vectors, has spread into South and Central America and the Caribbean. AAT is arguably the most important disease of domesticated livestock in sub-Saharan Africa, with losses estimated at US$1.3Bn per annum (Shaw, 2004). A wide range of vertebrate animals are hosts of trypanosomes: domestic companion animals (Lisulo *et al.*, 2014, Namangala *et al.*, 2013), domestic farm animals (Ruiz *et al.*, 2015) and wildlife (Anderson *et al.*, 2011, Auty *et al.*, 2012). The main susceptible hosts of trypanosomes are

*Bovidae* and *Suidae*; however, each trypanosome species has a different host range from different regions (Table 1.2). In domestic farm animals, the disease is very important in cattle and occasionally causes serious losses in pigs, camels, goats, sheep, horses, dogs, cats, monkeys and rodents (Lisulo *et al*., 2014, Losos and Ikede, 1972).

Different trypanosome species also have been reported circulating within a wide and diverse community of wild hosts. For example, Anderson *et al* (2011) found that 13.9% of wildlife in the Luangwa Valley, Zambia were positive for trypanosomes (*T. b. gambiense*, *T. b. rhodesiense*, *T. congolense* and *T. vivax*). Infections were mainly detected in waterbucks (*Kobus ellipsiprymnus*), lions (*Panthera leo*), greater kudu (*Tragelaphus strepsiceros*) and bushbucks (*Tragelaphus scriptus*). Anderson *et al* (2011) suggested that the dominant hosts were bushbucks for *T. b. gambiense, Bovidae* for *T. congolense* and waterbucks for *T. vivax;* however, this was assessed only within a single geographic region. The particular host communities present in different geographic regions could affect the relative importance of particular host species for particular speices of trypanosomes. Wildlife hosts can also carry human infective species: *T. b. rhodesiense* was identified in African buffalo (*Syncerus caffer*) and *T. b. gambiense* in leopards (*Panthera pardus*). In addition, Auty *et al*. (2012) reported that many blood samples of wildlife from Tanzania during 2002 – 2007 and Zambia during 2005 – 2007 were trypanosome positive for a range of known species. *Trypanosoma simiae*, *T. simiae* Tsavo and *T. godfreyi* were found in warthogs; *T. brucei* were found in a zebra and a spotted hyena; lions were infected with *T. congolense* and *T. vivax* was detected in an African buffalo, a waterbuck and a giraffe. However, there were also additional genotypes of trypanosomes present that could not assigned to named species, suggesting that there could be high diversity of the parasites among wildlife hosts. These reports reflect that trypanosomiasis needs to be controlled because it is a mortal disease, which affects both humans and animals.

**Table 1.2** Review of vertebrate hosts of each Salivarian trypanosome species and subspecies in previous field studies.

| *Trypanosoma spp.* | Hosts | Study areas | References |
|---|---|---|---|
| *T. b. rhodesiense* | humans and cattle | Uganda | Welburn *et al.*, 2001 |
| | pigs | Tanzania | Hamill *et al.*, 2013 |
| *T. b. gambiense* | brush-tailed porcupines, giant rats, duiker, Mangabey monkeys, greater white-nosed monkeys, palm civets and small-spotted genets | Cameroon | Herder *et al.*, 2002 |
| | humans | Uganda and Sudan | Checchi *et al.*, 2012 |
| | humans | Zambia | Mwanakasale *et al.*, 2014 |
| | pigs | Tanzania | Hamill *et al.*, 2013 |
| *T. brucei* | brush-tailed porcupines, giant rats, sun squirrels, duikers, moustached monkeys, white-nosed monkeys, dwarf guenons, galagoes, golden pottoes, long-tailed pangolins,tree pangolins and palm civets | Cameroon | Herder *et al.*, 2002 |
| | spotted hyenas, zebras | Tanzania | Auty *et al.*, 2012 |
| *T. congolense* | brush-tailed porcupines, giant rats, duikers, moustached monkeys and small-spotted genets | Cameroon | Herder *et al.*, 2002 |
| | pigs, goats, sheep and dogs | Cameroon | Nimpaye *et al.*, 2011 |
| | lions and spotted hyenas | Tanzania | Auty *et al.*, 2012 |
| *T. vivax* | brush-tailed porcupines, giant rats, duikers, moustached monkeys, mona monkeys, greater white-nosed monkeys, dwarf guenons, galagoes, tree pangolins, palm civets, small spotted genets and monitor lizards | Cameroon | Herder *et al.*, 2002 |
| | pigs, goats, sheep and dogs | Cameroon | Nimpaye *et al.*, 2011 |
| | cape buffalo, waterbucks and giraffes | Tanzania | Auty *et al.*, 2012 |
| | pigs | Tanzania | Hamill *et al.*, 2013 |
| | cattle, camels, goats, sheep, horses and donkeys | Ethiopia | Birhanu *et al.*, 2015 |
| *T. simiae* | brush-tailed porcupines | Cameroon | Herder *et al.*, 2002 |
| | pigs and goats | Cameroon | Nimpaye *et al.*, 2011 |
| | warthogs | Zambia and Tanzania | Auty *et al.*, 2012 |
| | pigs | Tanzania | Hamill *et al.*, 2013 |
| *T. godfreyi* | warthogs | Zambia and Tanzania | Auty *et al.*, 2012 |
| | pigs | Tanzania | Hamill *et al.*, 2013 |

## 1.2 Trypanosome diagnosis

Trypanosome infection is determined by distribution areas, clinical signs and laboratory diagnosis. The available laboratory diagnosis of trypanosomes is divided into three main methods: parasitological, serological, and molecular (Wastling and Welburn, 2011), which have different advantages and are used for different objectives.

### 1.2.1 Parasitological diagnosis

Parasitological diagnosis is a direct method for simple, inexpensive and quick trypanosome identification in blood, lymph and cerebrospinal fluid (CSF) using microscopy (Moody and Chiodini, 2000). Different techniques have been applied in parasitological diagnosis, leading to different levels of sensitivity. For example, fresh blood from a finger prick, chancre fluid (secretion of ulcer lesions from tsetse-fly biting) or lymph node aspiration from humans and blook from ears or tails of animals are collected for detection of trypanosome motility for clinical investigations using the wet blood film method (Moody and Chiodini, 2000). This is the simplest method but has low sensitivity (with a detection threshold of $10^4$ trypanosomes/ml of blood sample) and needs to be screened quickly (Büscher and Lejon, 2004). Thick blood film, in which the corner of a slide is used to spread a drop of blood that is then stained with Giemsa, is applied to increase sentitivity ($5x10^3$ trypanosomes/ml of blood sample) and the samples do not need to be tested immediately (Büscher and Lejon, 2004). Centrifugation of trypanosomes in blood samples or “the Woo method” or darkground or phase contrast microscopy buffy coat techniques (Murray *et al.*, 1977) are also applied to increase sensitivity of trypanosome diagnosis ($10^3$ and $10^2$ trypanosomes/ml of blood sample, respectively) (Desquesnes and Tresse, 1996) but require specialised equipment. In addition, diagnosis of species and subspecies of trypanosomes based on these methods are limited to morphological identities, which is not possible to distinguish all species.

### 1.2.2 Serological diagnosis

Serological methods are indirect methods for trypanosomasis diagnosis from body fluids (blood, serum and cerebrospinal fluid (CSF)) based on detection of

antibodies to specific antigens or using antibodies to detect the presence of circulating antigens of parasites. Serological tests are used as tools for research, monitoring, control and surveys (Chappuis *et al.*, 2005a). Many different types of serological tests have been in used, including: haemagglutination assays (HA), immunofluorescent assays (IFA), and indirect enzyme linked immunosorbent assays (ELISA) (Ross and Novoa-Montero, 1993).

Trypanosome diagnosis based on serological methods are able to distinguish species of pathogenic trypanosomes present (Desquesnes *et al.*, 2001a) and can detect more positive samples than microscopy-based parasitological methods. For example, parasitaemic trypanosomes of N'Dama cattle in the Gambia was diagnosed at levels of 3.0% and 54.7% using the buffy coat/dark ground phase contrast technique and ELISA, respectively (Mattioli *et al.*, 2001). However, false negative and false positive trypanosome results are possibly found in the early and recovery stages of infection, respectively. Moreover, cross reaction with other parasite antigens has been reported in methods for trypanosome antigen dectection (Nantulya, 1990). Nevertheless, sensitivity of antigen tests is often low in the eary stage of each peak of healthy parasites covered by a new type of surface antigen (Uilenberg, 1998).

For HAT diagnosis, a combination of clinical signs, disease distribution (Gibson, 2003) and card agglutination tests (Testryp® CATT) are used for wide-scale screening of *T. b. gambiense* in endemic areas (Robays *et al.*, 2004), with 87–98% sensitivity and 93 – 95% specificity of the test (Jamonneau *et al.*, 2000, Truc *et al.*, 2002). However, identification of *T. b. rhodesiense* infection relies on clinical features due to the absence of available serological tests for field use (Chappuis *et al.*, 2005a)

### 1.2.3 Molecular diagnosis

The method that is currently recommended for trypanosome diagnosis in both humans and animals is molecular identification, due to the high sensitivity and sensitivity of the method (Wastling and Welburn, 2011). The principle of molecular methods is to detect DNA sequences that are specific for trypanosome subgenus, species, subspecies, type or strain. Initial tests were based on hybridisation to radioactively labeled probes designed to test for the presence of

trypanosomes but diagnostic tests based on the Polymerase Chain Reaction (PCR) allow greater sensitivity for species identification and do not require radioactive isotopes (Weiss, 1995). Specificity of the PCR tests depend on the availability of primers that are general enough to amplify all variants of a given species but are specific enough to allow distinguishing between species. For HAT a number of primers have been developed: for example, primers targeting the *T. b. gambiense* specific glycoprotein (TgsGP) (Radwanska *et al.*, 2002b) and SRA primers (Serum resistance-associated gene) for *T. b. rhodesiense* specific glycoprotein (Radwanska *et al.*, 2002a). There have also been modifications to improve PCR-based sensitivity. For example, LAMP (Loop-mediated isothermal amplification) is a method developed recently, involving autocycling strand displacement DNA synthesis using a *Bst* DNA polymerase with two sets of specially designed inner and outer primers (Notomi *et al.*, 2000). LAMP amplifies targeted DNA with high specificity, efficiency, and rapidity under isothermal conditions. Commercial kits with high specificity for detection of *T. b. rhodesiense* DNA (Namangala *et al.*, 2012) and specific glycoproteins (TgsGP) for *T. b. gambiense* (Njiru *et al.*, 2011), without cross reactivity with other species (Cunningham *et al.*, 2016). However, with its high sensitivity, DNA of dead trypanosomes in migut was detectable until 6 days post feeding (Cunningham *et al.*, 2016).

The larger number of trypanosome species associated with AAT means that many different sets of primers have been developed, both to enable amplification of all species using general primers or by the development of species-specific primers (see section 1.2.3). However, a limitation of previous studies is that primers are often designed and tested for the species found within one geographic region, without confirming that they will have the same specific in other regions, by sequencing PCR products. Moreover, false negative results might be found when primers are too specific. For example, *T. vivax* specific primers can be used to recognise specific strains; e.g. West African Tv-PCR (Masake *et al.*, 1994) vs Tv-PCR (Masake *et al.*, 1997). Moreover, PCR can be inhibited by blood contents; e.g. heme (Akane *et al.*, 1994), haemoglobin, lactoferrin (Al-Soud and Radstrom, 2001), immunoglobulin G (Al-Soud *et al.*, 2000), haematin (Opel *et al.*, 2009) and divalent ions ($Ca^{2+}$, $Na^{+}$, $K^{+}$ and $Mg^{2+}$), (Alaeddini, 2012). Nevertheless, with proper optimisation and sequencing to

confirm success of amplification, PCR remains a powerful tool for species identification in prevalence screening studies.

## 1.3 Prevalence of Animal African trypanosomes

Different techniques have been used to survey the prevalence of trypanosomes in animals, both in order to control the disease in domestic animals and to reduce risks associated with transmission to humans (Welburn *et al.*, 2001, Hamill *et al.*, 2013, Ruiz *et al.*, 2015). However, differences in detectability of trypanosomes using different techniques could bias interpretation of differences in apparent prevalence due to their different sensitivity and specificity. Although there is no comprehensive atlas map of AAT distribution in Africa as exists for HAT, the relative abundance of each trypanosome species can differ greatly between geographic regions (Table 1.3). However, this could be affected by relative detectability using various techniques.

In cattle, for example, prevalence of trypanosomes has been found to vary extensively by regions, year and seasons but different studies have also used different techniques. Using microscopy techniques, different prevalence of trypanosomes and variation of the relative prevalence of *T. congolense* and *T. vivax* was found between different studies conducted in Kenya (Mbahin *et al.*, 2013) and Ethiopia (Desta, 2014) (Table 1.3). In addition, in a single study in Ethiopia, there was relatively high consistency across two sites but prevalence increased across the four years of sampling (Rowlands *et al.*, 1993). Based on PCR screening, apparent trypanosome prevalence varied according to different primers used. Different relative prevalence of *T. congolense*, *T. vivax* and *T. brucei* in Uganda was found using general ITS primers (using both nested and non-nested approaches) compared to species-specific primers, with extensive differences between studies that were conducted in the same district (Tororo) but using different primer combinations (Cox *et al.*, 2005, Muhanguzi *et al.*, 2014). There was also extensive variation found among villages within this region (0 – 43.0%). Trypanosome prevalence of cattle from Msubugwe, Tanzania was 56.9% when a nested ITS-1 PCR technique was used but only 28.5% when diagnosed with standard PCR with species-specific primers (Adams *et al.*, 2006).

Since studies also tend to differ in whether they report overall trypanosome prevalence, the relative percentage of each species found in positive samples, or the overall prevalence of each species (Table 1.3), it is quite difficult to make direct comparisons between studies. Most studies have found relatively high prevalence of trypanosomes in cattle but this could be biased by sampling of putatively infected individuals. However, even among cattle differences in abundance have been found between seasons. Using species-specific primers in the Jos Plateau, Nigeria, Majekodunmi *et al*. (2013) found that prevalence of *T. congolense* was highest in the dry season (30.7%) while that for *T. vivax* was highest in the late wet season (29.9%). *Trypanosoma brucei* showed low prevalence all year (average 3.2%), with a peak of prevalence in the dry season (5.3%). Thus, environmental changes could influence the relative prevalence of trypanosomes but lack of consistent choice of diagnostic methods tends to obscure conclusions that can be drawn across studies.

**Table 1.3** Prevalence of *T. congolense*, *T. vivax* and *T. brucei* in cattle in Africa based on microscopy and PCR.

| *Trypanosoma spp.* | Prevalence (%) | Site of samples | Remarks | References |
|---|---|---|---|---|
| **Screening method: phase-contrast and buffy-coat technique** | | | | |
| *Trypanosoma spp* | 33.9 | Kwale, Kenya | - | Mbahin *et al.* (2013) |
| *T. congolense* | 19.0 | | | |
| *T. vivax* | 14.9 | | | |
| *Trypanosoma spp.* | 12.4 | Didessa valley, Ethiopia | - | Desta (2014) |
| *T. congolense* | 10.1 | | | |
| *T. vivax* | 1.6 | | | |
| Mixed infection | 0.7 | | | |
| *Trypanosoma spp.* | 25.0, 21.4, 38.2 and 42.8 | Ghibe, Ethiopia | collected in 1986, 1987, 1988 and 1989 | Rowlands *et al.* (1993) |
| | 23.0, 24.8, 31.4 and 32.5 | Tolley, Ethiopia | | |
| **Screening method: Polymerase chain reaction (PCR)** | | | | |
| *T. congolense* | 1.0[a], 0[b] | Soroti, Uganda | nested ITS, species-specific primers | Cox *et al.* (2005) |
| *T. vivax* | 1.0[a], 1.0[b] | | | |
| *T. brucei* | 7.0[a], 5.0[b] | | | |
| *T. congolense* | 5.0[a], 1.0[b] | Tororo, Uganda | | |
| *T. vivax* | 5.0[a], 8.0[b] | | | |
| *T. brucei* | 33.0[a], 32.0[b] | | | |
| *Trypanosoma spp* | 15.3 | Tororo, Uganda | ITS-1 CF/BR, TCS 1/2 primers | Muhanguzi *et al.* (2014) |
| *T. congolense* | 2.1 | | | |
| *T. vivax* | 13.4 | | | |
| *T. brucei* | 1.1 | | | |
| *Trypanosoma spp.* | 46.8 | Jos Plateau, Nigeria | collected in 2008, species-specific primers | Majekodunmi *et al.* (2013) |
| *T. congolense* | 27.7 | | | |
| *T. vivax* | 26.7 | | | |
| *T. brucei* | 3.2 | | | |

[a] based on nested ITS PCR, [b] based on species-specific PCR.

Although sampling is more difficult from wild animals, the same types of variation have been reported but there are also extensive differences among host species. Based on the indirect fluorescent antibody test (IFAT), trypanosome prevalence in warthogs from the Gambia was 11% (N = 62) and all were infected with *T. simiae* (Claxton *et al.*, 1992). All blood samples were from warthogs shot by local hunters at three villages where there was concern that cattle health and productivity could be at risk due to a high abundance of tsetse flies in the area, so it is possible that the IFAT underestimated prevalence. Wild

animals also might host a broader range of parasites than the tests are designed for. For example, in the Serengeti National Park, sequencing of the entire ITS-1 and ITS-2 regions of the rDNA revealed a much higher diversity of sequences than anticipated from wild hosts so species-specific tests would likely underestimate prevalence (Auty *et al.*, 2012). Among 418 wild animals from Luangwa Valley in Zambia during 2005 – 2007, although sample sizes for individual hosts was small, trypanosome prevalence and species varied greatly among wild hosts (Anderson *et al.*, 2011). These differences could reflect differences in exposure of the hosts to tsetse flies or to differences in susceptibility to the trypanosomes but potential differences in detectability of the parasites in the hosts cannot be excluded as a possibility. These types of surveys also cannot distinguish between accidental hosts and potential reservoirs of disease in domesticated animals or humans. Thus, trypanosome surveillance in the biological vectors (i.e. tsetse flies) could be important for determining the relative risk of transmission between different host species.

## 1.4 Trypanosome control and drug treatment

Three main strategies have been attempted for HAT control: the removal of susceptible populations from risk areas; mass chemoprophylaxis programmes; and active surveillance and treatment (Fèvre *et al.*, 2004, Cattand *et al.*, 2001). Trypanosomes are also controlled in livestock, sometimes through limiting movement because they are reservoirs and integral to the spread of sleeping sickness to humans and so could be responsible for new outbreaks of disease (Welburn *et al.*, 2006). Current drugs used for treatment of humans include pentamidine, suramin, melarsoprol, eflornithine and nifurtimox-eflornithine combination therapy (NECT) while drugs of choices for trypanosomiasis prevention and treatment in animals are Berenil, Samorin and Homidium (Afewerk *et al.*, 2000, Berret, 2001, Bridges *et al.*, 2007, Castro *et al.*, 2006, de Koning, 2001, Dukes, 1984, Geerts *et al.*, 2001, Matovu *et al.*, 2001, McDermott *et al.*, 2003, Priotto *et al.*, 2006, Simarro *et al.*, 2012) (Table 1.4). These drugs are not always successful for HAT and AAT treatment because of limitations from drug efficacy, resistance, and high toxicity (Simarro *et al.*, 2011). Moreover, it is difficult to apply drug or movement controls to wild animal reservoirs, which

could transmit to both livestock and humans. Thus, the safest and most effective policy is to control the vectors of trypanosomes, which in Africa are tsetse flies in the genus *Glossina*.

**Table 1.4** Summarisation of drug treatment for trypanosomasis in humans and animals

| Drugs | Advantages | Disadvantages |
|---|---|---|
| **HAT treatment** | | |
| Pentamidine | For treatment of the early stage of *T. b gambiense* infection (Damper and Patton, 1976) | Drug resistance has been reported in field isolates (Dukes, 1984, Bernhard *et al*., 2007, Bridges *et al*., 2007, Bray *et al*., 2003, de Koning, 2001) |
| Suramin | A drug of choice for the early stage of *T. b. rhodesiense* infection. | Toxic to the urinary tract by inducing a consistent reduction in the amplitude of non-voiding bladder contractions, with a tendency to reduce their frequency (Velasco *et al*., 2003)<br>Ineffective against *T. b. gambiense* infection (Pepin and Milord, 1994) |
| Melarsoprol | Recommended for use against second neurological stages of both *T. b gambiense* and *T. b rhodesiense* infection (Berret, 2001)<br>Cheaper than eflornithine | Less effective than other drugs in the past and a lot of undesirable side effects reported (Balasegaram *et al*., 2006)<br>Drug toxicity causes a dramatic reaction related to encephalopathy<br>Drug resistance (Bernhard *et al*., 2007, Bridges *et al*., 2007) |
| Eflornithine | Effective treatment for *T. b. gambiense* at both the first and the second stages (Pepin *et al*., 1987).<br>Effective treatment for the late-stage of Gambian trypanosomiasis (Balasegaram *et al*., 2006) with less toxicity or side effects when compared with melarsoprol (Chappuis *et al*., 2005b, Milord *et al*., 1992)<br>Accepted toxicity by WHO | Expensive<br>Adverse effects are alopecia, diarrhoea, leucopenia, anaemia, neurological symptoms (for example, convulsion) and death (Milord *et al*., 1992)<br>Not effective for *T.b.rhodesiense* infection |
| Nifurtimox-Eflornithine Combination Therapy (NECT) | For *T. b gambiense* treatment (Priotto *et al*., 2006)<br>The combination therapy is easier to administer than monotherapy with eflornithine, decreases toxicity, has a shorter duration for treatment, lower risk of emergence of resistance to eflornithine, and saves costs (Simarro *et al.*, 2012) | Not effective against *T. b. rhodesiense* (Vincent *et al*., 2010) |

| Drugs | Advantages | Disadvantages |
|---|---|---|
| **AAT treatment** | | |
| Diminazene Aceturate | Widely used to treat cattle trypanosomiasis (Whiteside, 1962)<br>Effective treatment in dogs, sheep, goats and caels | Drug resistance has been widely reported (Afewerk *et al*., 2000)<br>Some strains that are resistant to berenil show cross resistance to Diminazene (McDermott *et al*., 2003) |
| Isometamidium Chloride | Use for prophylactic and therapeutic purposes in drugs for animal trypanosomiasis | Toxic effects: generalized tissue damage, nervous stimulation, increased erythrocyte osmotic fragility, intravascular haemolysis and decreased erythrocytes (Ali and Hassan, 1986)<br>Wide spread resistance in *T. congolense* and *T. b. brucei* (Kaminsky and Zweygarth, 1989) |
| Homidium | Chemoprophylactic for cattle trypanosomiasis (Dolan *et al*., 1990) in areas with a high population of trypanosome-infected tsetse flies | Potential mutagenic properties and widespread resistance (Afewerk *et al*., 2000, Rowlands *et al*., 1993, Wang, 1995) |

## 1.5 Biology of *Glossina spp*.

### 1.5.1 Taxonomy of tsetse flies

| | |
|---|---|
| Kingdom: | Animalia |
| Phylum: | Arthropoda |
| Class: | Insecta |
| Subsection: | *Calyptratae* |
| Superfamily: | *Hippoboscoidae* |
| Family: | *Glossinidae* |
| Genus: | *Glossina* |

Tsetse flies (*Glossina spp*.), the biological vectors of trypanosomes (Jordan, 1976), have various sizes between 6 - 13 mm. The haematophagous insects are distinguished from other blood-sucking insects by their long proboscis extending directly forward and a special wing vein pattern called “hatchet cell”. There are 31 species and subspecies of tsetse flies, which are divided into three distinct groups by their habitats: Morsitans in savannah regions, Palpalis in high humidity areas and Fusca in evergreen forests. Species and subspecies of *Glossina* are further characterised by their geographic distribution, distinguishing

morphology, physiology, biochemistry, genetics, host range and pathogenicity (Maudlin *et al.*, 2004) (Table 1.5).

**Table 1.5** Species and subspecies of tsetse flies belonging to the Morsitans, Palpalis and Fusca groups (Pollock, 1982).

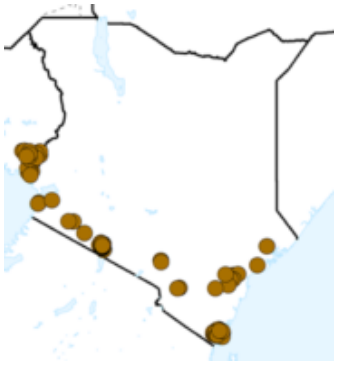
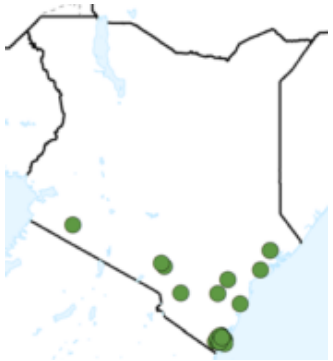
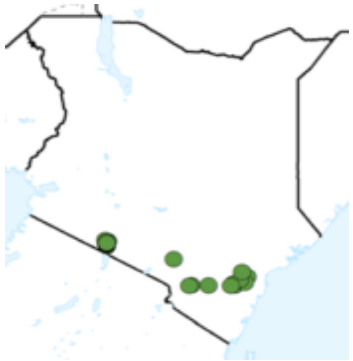
| Morsitans (Savannah) | Palpalis (Riverine) | Fusca (Forest) |
|---|---|---|
| *G. morsitans submorsitans* | *G. palpalis palpalis* | *G. nigrofusca nigrofusca* |
| *G. morsitans centralis* | *G. palpalis gambiensis* | *G. nigrofusca hopkinsi* |
| *G. morsitans morsitans* | *G. fuscipas fuscipes* | *G. fusca fusca* |
| *G. austeni* | *G. fuscipes martinii* | *G. fusca congolensis* |
| *G. pallidipes* | *G. fuscipes quanzensis* | *G. fuscipleuris* |
| *G. swynnertoni* | *G. tachinoides* | *G. haningtoni* |
| *G. longipalpis* | *G. pallicera pallicera* | *G. schwetzi* |
| | *G. pallicera newsteadi* | *G. tabaniformis* |
| | *G. caliginea* | *G. nashi* |
| | | *G. vanhoofi* |
| | | *G. medicorum* |
| | | *G. severini* |
| | | *G. brevipalpis* |
| | | *G. longipennis* |
| | | *G. frezili* |

### 1.5.2 Distribution and habitat of tsetse flies

Tsetse flies have a wide distribution, which shows a high correspondence with that of the trypanosomes they transmit to animals (Willson *et al.*, 1963). In Africa, tsetse flies are distributed from the southern edge of the Sahara desert (15°N) to Angola, Zimbabwe and Mozambique (20°S) (World Organisation for Animal Health, 2013). Habitats and distributions of each species are different because of their specific requirements; for example, temperature, humidity, cover surrounding habitats and available food. Tsetse flies do not usually move far from their preferred habitats; therefore, tsetse flies prefer to feed on locally available host animals (Weitz, 1963). In Southeast Kenya, suitable habitats are available for *G. longipennis, G. brevipalpis* and *G. pallidipes, G. austeni* but not all species are found in the same geographic locations due to their specific habitat preferences (Cecchi *et al.*, 2015) (Table 1.6).

In general, the three groups of tsetse flies inhabit specific environmental conditions (Laird, 1977). Tsetse flies in the Morsitans group settle in open areas, which are surrounded with grass, but few bushes and trees. Tsetse flies in the Palpalis group are adapted to high humidity climates, and are always found beside a river. In contrast, Fusca is a predominantly forest group, and are particularly found in moist thick evergreen forests. Temperature and humidity influence the reproductive cycle and development and so are thought to be primary factors in determining tsetse distributions (Mellanby, 1937). Temperature increases due to climate warming are thus expected to have a large impact on the population dynamics of tsetse flies (Maudlin *et al.*, 2004), geographic distribution of the diseases they transmit, and transmission dynamics (Ebikeme, 2011). Epidemics of trypanosomes are predicted to occur most between temperature ranges of 20.7 – 26.1°C (Moore *et al.*, 2012), although there is a wide range of temperatures that allow tsetse fly survival. *G. pallidipes* for example, has been found to survive both low (3.7, 8.9 and 9.6 °C) and high temperature (37.9, 36.2 and 35.6 °C) extremes, at least in short-term experiments (Terblanche *et al.*, 2008). However, more restricted temperatures could be necessary for tsetse fly development. At 24°C, tsetse flies emerge from their puparium on around day 30 (Leak, 1998) but extreme high or low temperature affect the duration of puparial development (Muzari and Hargrove, 2005). The optimum temperature for the reproductive cycle of females for laboratory colonies of *G. f. fuscipes* has been found to be 25°C (Mellanby, 1937) and a constant 30°C temperature makes the flies sterile. For *G. m. morsitans* the growth rate of oocytes, time to production of the first larvae, and development through the three stage of larvae all increase with temperature (Hargrove, 1994). The lethal temperature for the pupae depends on exposure time (Potts, 1933). The optimal temperature for incubation of the puparial stage for this subspecies is 23 °C; pupae survive for only two minutes when exposed to 55 – 57°C, half an hour for 45 – 50°C, 4 hours or more for 40°C, and failed to survive 13 days for 35°C. Temperature not only influences the distribution of tsetse flies in term of physical development but also affects the relative availability of host species, which could affect feeding behaviours and host range of tsetse flies.

**Table 1.6** General biological data of tsetse flies in Southeast Kenya: *G. austeni*, *G. pallidipes*, *G. brevipalpis* and *G. longipennis*

| Biological data | *G. austeni* | *G. pallidipes* | *G. brevipalpis* | *G. longipennis* |
|---|---|---|---|---|
| Tsetse group | Morsitans (Savannah group) | | Fusca (Morsitan group) | |
| Longevity | 54 days (Wamwiri *et al.*, 2013), 120 days (Nash *et al.*, 1968) and 151 days (Boyle, 1971) | 70 ± 27 days) (Jaenson, 1986, Wamwiri *et al.*, 2013) | Not found | |
| Habitat (Cecchi *et al.*, 2008, Ford, 1971, Jordan, 1986, Langridge *et al.*, 1963) | tall thick forest close to coast area | coastal evergreen forest and dry savannah | coastal evergreen forest (often associated with water sourses, for example, forest islands) | dry savannah more arid regions |
| Distribution map (Cecchi *et al.*, 2015) |  |  |  |  |
| Main host choices (Leak, 1998, Weitz, 1963) | feeds mainly on bushpigs (50 – 60%) and buffalo | 55 – 90% its meals from the bushbucks, warthogs, bushpigs and buffalo | 40% of its meals from the bushpigs, buffalo and bushbucks | buffalo, elephants and bushpigs |

### 1.5.3 Life cycle of tsetse flies

The life cycle of tsetse flies is different from other haematophagous insects (Leak, 1998). Females have a pair of ovaries, a uterus and the spermathecae to store spermatozoa received from copulation with males. Females restart their reproductive cycle every 9 – 10 days. The majority of female *G. m. morsitans* and *G. pallidipes* from Zimbabwe have been found to be inseminated by 4 and 7 days of age, respectively (Hargrove, 1994). After larviposition, a single ovum is fertilized with sperm from direct mating with males or a spermathecae and hatched to be the first-stage (instar) larvae in the uterus. Secretions from milk glands provide nutrients for larval development and molting, also in the uterus. Third-stage larvae are deposited on humid sandy or decaying solids in the shade. The larvae then burrow under ground and develop to a puparium stage for 30 – 40 days. After that, teneral young tsetse flies emerge from the puparium stage and start to seek hosts for feeding as soon as possible to completely develop their muscles.

Because the normal vertebrate blood diet of tsetse flies lacks some essential nutrients, their development and maintenance requires supplements obtained from endosymbionts (Douglas, 1989). These endosymbionts are necessary for complete development of the tsetse reproductive system (Nogge, 1976, Aksoy, 2000) and are vertically transmitted through maternal inheritance (Balmand *et al.*, 2013). To test whether this role of endosymbionts is essential, tsetse flies were treated with oxytetracyclin, sulphaquinoxaline and lysozyme to kill or neutralize these microbes (Nogge, 1976). This resulted in reduced growth rates and decreased numbers of eggs in the treated tsetse flies. However, the ability to reproduce a full complement of eggs without endosymbionts could be partially restored by addition of B-complex vitamins (Nogge and Gerresheim, 1982). These endosymbionts are also thought to influence the vector competence of tsetse flies for trypanosome infection. Although many bacterial flora have been investigated in tsetse flies (Geiger *et al.*, 2013), three major bacterial endosymbionts have been identified: *Wigglesworthia glossinidia* (Aksoy, 1995b)*; Wolbachia pipientis* (O'Neill *et al.*, 1993); and *Sodalis glossinidius* (Welburn *et al.*, 1993, Dale and Maudlin, 1999). *Wigglesworthia glossinidia*, a primary endosymbiont, resides intracellularly in specialized epithelial cells of the midgut. Phylogenetic comparison between tsetse flies and

*Wigglesworthia spp.* displayed completely matching relationships among the different species of tsetse and their associated endosymbionts (Aksoy, 2000). *Wigglesworthia glossinidia* enhances the immune system of mature tsetse flies (Weiss *et al.*, 2011), the fertility of females, and has been found to reduce the vector competence of flies for trypanosome infection (Pais *et al.*, 2008). *Wigglesworthia glossinidia* maternally transmits to intrauterine larvae (Rio *et al.*, 2006) and is found in ovary, egg, milk glands and spermatheca tissue (Aksoy *et al.*, 1997). *Wolbachia pipientis* is a highly diverse group of intracellular, maternally inherited endosymbiont in the Proteobacteria (Werren, 1997). It is found in reproductive tissues of both gonads and ovaries (Cheng *et al.*, 2000). *Wolbachia* has been reported to cause some reproductive abnormalities, for examples, cytoplasmic incompatibility (CI) (Stouthamer *et al.*, 1999), leading to embryonic lethality (Alam *et al.*, 2011). *Sodalis glossinidius* has been reported to be involved in regulation of iron in tsetse flies: individuals lacking *S. glossinidus* may not grow to their full size potential (Smith *et al.*, 2013). Absence of *Sodalis* also has been found to affect the reproductive capability of tsetse flies (Dale and Welburn, 2001). Some studies have suggested that the presence of the bacteria increases the biological vector competence for trypanosome infection (Welburn *et al.*, 1993, Dale and Welburn, 2001, Farikou *et al.*, 2010a, Soumana *et al.*, 2013b, Wamwiri *et al.*, 2014). However, not all studies have agreed on the nature of its potential role in mediating trypanosome infections so this remains unclear.

### 1.5.4 Age determination

Average longevity of tsetse flies is up to 4 months but differs by sex: 8 to 14 weeks for females, but only 4 to 6 weeks for males (Leak, 1998). However, lifespan also varies with tsetse species and environmental temperature of their habitat, and this can affect the relative longevity of males and females. For example, at 25°C and 60% relative humidity, male *G. m. morsitans* lived for 145.9 ± 51 (mean±SD) days, while females lived for 131.1 ± 58.3 days (Chigusa *et al.*, 1997).

In order to determine the age of wild tsetse flies,several different methods have been used: using pteridine quantification, ovary scoring and wing fray scoring (Hayes and Wall, 1999).The pteridines are pigments in the compound eyes of

*Diptera*. Pteridine accumulation is a somatic degradated product of purine metabolism. Presence and abundance of pteridine in compound eyes of tsetse flies has been studies using fluorescence spectroscopy (Harmsen, 1970, McLntyre and Gooding, 1996); levels of pteridine accumulation linearly has been found to increase with age (Lehane and Mail, 1985). A limitation of using pteridine quantification is that tsetse samples must be preserved in a dark desiccated condition, which can be difficult under field conditions. Morever, ages of field-caught flies were overestimated, which the authors suggested could be due to effects from size of tsetse flies, temperature and season (Langley *et al*., 1988). Age has also been estimated based on ovary aging by determining changes in the reproductive cycle of females. Ovaries are dissected for an examination of three points: relative sizes of the four egg follicles, number and position of follicular relics, and the contents of the uterus (Wall, 1990). Although, females are classified into eight categories based on ovary scoring, but this is not directly equivalent to the number of ovarian cycles, since categories 4 - 8 include flies older than eight (Hayes and Wall, 1999). When age determination using pteridine aging was compared with ovary dissection in *G. m. morsitans*, although a strong correlation was found for young flies, this was reduced for older flies (44 to 71 days old) (Langley *et al*., 1988). Pteridine aging and ovary scoring methods, which grade age based on biological degradation, require fresh samples to provide more accuracy than preserved samples, and ovarian scoring can only be applied to females (Njiru, 2014). Thus, wing fray scoring has been suggested for age determination of preserved tsetse flies. Wing fray scoring is a relative index that divides individuals into six age classes based on visual quantification of the degree of damage to the trailing wing edge. Wing fray was initially used to estimate the mean age of male *G. m. morsitans* (Jackson, 1946), where it was directly related to fly age. There has also been high correspondence found between wing fray and ovarian scoring methods (Woolhouse *et al*., 1993), with $R^2$ values of up to 80% (Saunders, 1962). Since the wing fray score can be applied to both males and females and can use preserved tissues, it thus provides a good relative measure of age.

### 1.5.5 *Sodalis glossinidius*

*Sodalis glossinidius* is a member of the Enterobacteriaceae family (Aksoy, 2000), which was first identified as a rickettsia-like organism (RLO) (Moloo and Shaw,

1989). It is most abundant in the midgut epithelial cells of tsetse flies and also is detectable in haemolymph, muscle, fat bodies, milk glands and salivary glands (Cheng and Aksoy, 1999). The bacteria can be transmitted both vertically to offspring via secretion from milk glands and horizontally from males to females by mating (De Vooght *et al.*, 2015).

#### 1.5.5.1 Identification of *S. glossinidius* based on PCR

Although *S. glossinidius* is able to be cultured on a semi-solid medium under micro-aerobic conditions for diagnosis (Dale and Maudlin, 1999), PCR methods have been used extensively for bacterial identification in large-scale screenings because they require less time and it is easier to identify bacteria to species if the products are sequenced. Many specific primers have been designed to identify different genes of *Sodalis* in tsetse flies; for example, pSG2 (Farikou *et al.*, 2010a), GPO1 (Dale and Maudlin, 1999), hemolysin (Hem) (Pais *et al.*, 2008). The pSG2 primers were developed to target the extrachromosomal plasmid 2 (which is abundant in this bacterial endosymbiont (Darby *et al.*, 2005)) in *G. p. palpalis*, *G. pallicera*, *G. caliginea* and *G. nigrofusca* sampled from South Cameroon (Farikou *et al.*, 2010a). Another gene on this plasmid has also been targeted (GPO1), with primers developed to detect the secondary endosymbionts in many species of laboratory-reared tsetse flies: *G. austeni, G. brevipalpis, G. f. fuscipes* and *G. tachinoides* (from the Seibersdorf Agricultural Research Laboratory, Vienna, Austria); *G. m. morsitans* (from the Tsetse Research Laboratory, Bristol University, England); and *G. palpalis* (from the University of Alberta) (O'Neill *et al.*, 1993). The nuclear hemolysin gene (which is though to help access iron from the blood meals in the tsetse gut (Toh *et al.*, 2006)) has also been used for screening, using the Hem primers developed to identify *S. glossinidius* in a *G. m. morsitans* colony maintained in the insectary at Yale University (Pais *et al.*, 2008), which were originally established from puparia from tsetse fly populations in Zimbabwe. However, direct comparisons between different sets of primers has not been made determine which are most suitable for screening and assessing genetic diversity of *S. glossinidius* in tsetse-fly samples. Moreover, as for trypanosomes, PCR-based detection could vary by tsetse tissue sampled but robust comparisons have not yet been made to determine whether conclusions about prevalence of the endosymbionts vary by the choice of tissue.

#### 1.5.5.2 Prevalence of *S. glossinidius*

Variation in *Sodalis* prevalence has been reported when tsetse flies are classified by geographic sources, species, sex and age of tsetse flies. However, different sets of PCR primers have been used in different studies and different species of tsetse are present in different regions, making it difficult to directly compare between studies. *Glossina palpalis palpalis* from Bipindi of South Cameroon had higher prevalence of *Sodalis* (64.4%) than in *G. pallicera and G. caliginea* from Campo (45.3%), based on pSG2 primers targeting a region of the plasmid DNA (Farikou *et al.*, 2010a). Using GPO1 primers (targeting a different region of the plasmids), *Sodalis* prevalence in *G. pallidipes* from Kenya (16.0%) was higher than in *G. austeni* (3.7%) (Wamwiri *et al.*, 2013). In Zambia, prevalence of *Sodalis* was the highest in *G. brevipalpis* (93.7%) using primers targeting the nuclear *GroEL* gene, followed by *G. m. morsitans* (17.5%) and *G. pallidipes* (1.4%) (Dennis *et al.*, 2014). Across the three species, prevalence in males (11.1%, 95% CI = 1.4% - 34.7%) was found to be significantly higher than in females (1.2%, 95% CI = 0.4% - 2.9%) but a difference between sexes was not found for *G. m. morsitans* on its own (P-value = 0.566; 12.0% for males and 18.8% for females). A slight decreased in the incidence of *S. glossinidius* infection with increasing fly age was revealed (Wamwiri *et al.*, 2014). In addition, *Sodalis* status showed a significantly positive correlation with trypanosome infection in *G. pallidipes* from Kenya (P-value = 0.0127) but not in *G. austeni* (P-value = 0.1554) (Wamwiri *et al.*, 2013). Although these studies on isolated factors suggest that *Sodalis* prevalence can vary by tsetse species characteristics and geographic distribution and that it could influence trypanosome prevalence, integrated studies are lacking to tease out whether *Sodalis* presence itself enhances opportunities for trypanosome infection or whether the same factors that make prevalence of *Sodalis* more likely also make trypanosome infection more likely.

### 1.5.6 The feeding process and behaviour of tsetse flies

Feeding is an important behavior in tsetse flies that shows a fixed pattern, which can be exploited in the development of control measures. Both males and females feed on vertebrate blood, with feeding divided into three steps: host

detection; landing; and feeding. Disruption of any of these steps could help to reduce trypanosome transmission.

Host detection is the first challenge and is accomplished through a combination of odour and visual cues. Tsetse flies seek hosts for approximately 25 min/day (Bursell and Taylor, 1980), flying at speeds of up to 24 km/hr (Gibson and Brady, 1988), giving a flight distance of 10 km/day (Vale *et al.*, 2014). However, this varies by habitat geometry, species of tsetse fly and sex (Vale *et al.*, 2014). Hosts can be detected by antennae at a distance of 60 – 120 m (Chapman, 1961), but this can vary with the direction of the wind (Vale, 1977). Fine-scale odor detection is also mediated through sensory pits located in the third segment of the abdomen, which contain many sensory hairs (sensillae) with 141 olfactory receptor cells (Otter and Naters, 1992). Eight odorant-binding proteins (OBPs) also play an important role in olfaction of tsetse flies by mediating interactions between odorants and odorant receptors (Liu *et al.*, 2010). There are three main types of odor attractants for tsetse flies (Torr and Vale, 2015): 1) odor associated with animal breath (e.g. acetone (Mihok *et al.*, 2007), octenol, fluid mixtures in rumens (Harraca *et al.*, 2009), and carbon dioxide (Vale and Hall, 1985); 2) odour associated with urine (e.g. phenols) (Mbahin *et al.*, 2013, Spath, 1995); and 3) odor associated with skin secretions (e.g. sebum) (Warnes, 1995). Thus, these types of chemical compounds can be used to attract tsetse flies to traps; for example, using mixtures of acetone and cattle urea (Kyorku *et al.*, 1990, Brightwell *et al.*, 1991).

Visual recognition of hosts occurs at a much closer range: about 10 m for stationary hosts but possibly a much greater distance for mobile hosts (Vale, 1977). The compound eyes compose of many units of ommatidia, which the middle ommatidium is stimulated to a different extent from each lateral eye (Turner and Invest, 1973). Tsetse flies see two points separately only if the retinal focus on both points and the space between the points involved at least three contiguous ommatidia. When two contiguous ommatidia are involved in the formation of the retinal image, the two points are seen as one larger spot. When tsetse flies find their targets, they directly fly to the hosts, expecially moving hosts (Fiske, 1920) In the second step, tsetse flies use colour as a guide to choose landing sites.(Green, 1986, Green and Flint, 1986) For example, strips of pigment in animal skins (for instance the striped pattern on zebra skins)

decreases attraction (Gibson, 1992) but black (Brady and Shereni, 1988) and blue colors (Green, 1986) are attractive, which also has been exploited in the design of traps.

In the final step, probing responses and feeding are stimulated by the heat of the hosts, which is detected by receptors on the tarsi of tsetse flies (Langley *et al.*, 1988). As the fly starts to probe, the labium moves from the palps to an angle of 90° to the skin of a host. While the labella rests on the skin, the teeth on the inner surface are everted and penetrated into lacerate capillaries, resulting in haemorrhage (Langley, 1972), which is sucked into the labrum. When the fly stops sucking, a small pool of blood forms. If blood is not found, the fly withdraws the labium partially and makes a new penetration.

Completed feeding or feeding success is defined when the abdomen of flies become fully distended with blood and there is no evidence for disturbance from the host that disrupts feeding. Based on these criteria, for *G. m.morsitans* and *G. pallidipes*, the time taken for successful feeding on domesticated ruminants was 109 ± 9 sec (Schofield and Torr, 2002). Feeding success of tsetse flies on cattle was only 15%. Young animals showed pronounced defensive behaviour and appeared to be bitten less than adults because of their small size and high activity. Feeding sucess is also related to the fat and haematin content of blood meals (Langley and Wall, 1990, Hargrove, 1976), ambient temperature, the presence of humans around the feeding areas (Hargrove, 1976) and other flies and host defenses (Schofield and Torr, 2002). Host defensive activities caused 69% disturbance of tsetse-fly feeding (Schofield and Torr, 2002), because they interrupt tsetse feeding behavior (Torr and Mangwiro, 2000) and induce host changing. The flies can be repelled by physical and/or mechanical host defensive behaviours, including tail flicking, ear flicking, skin twitching, kicking and other movements (Torr, 1994, Torr and Mangwiro, 2000). This would result in higher rates of feeding on multiple hosts, which could increase risk of transmission of trypanosomes between hosts and vectors.

Rates of feeding could also be influenced by characteristics of individual flies. Blood meals of *G. m. morsitans* are completely digested 4 to 5 days after feeding at 25°C (Langley and Stafford, 1990), but this has not been experimentally determined in other species. Host species selected for the

second blood meal of teneral male *G. palpalis gambiensis* has been found to depend on the host encountered for their first blood meal when the between-meal interval is 2 days but this preference disappears when the between-meal interval is extended to 3 days, suggesting a role of learning in host preferences (Bouyer *et al.*, 2007). The frequency of feeding has been related to the reproductive cycle of tsetse flies (Langley and Stafford, 1990). Females feed on larger blood meals than males (Langley and Stafford, 1990, Gaston and Randolph, 1993) but males feed more often. Similarly, unmated females feed on smaller meals than mated females because of their lower capacity of gut dilation and the effects of a hormone-mediated pheromone (Ejezie and Davey, 1976). Thus, intrinsic factors of flies could affect their feeding behaviours, which also might differ between species.

### 1.5.7 Hosts ranges of tsetse flies

Many species of vertebrate are fed on by tsetse flies but host preferences differ between the groups. *Suidae* and *Bovidae* are in general favorite hosts of tsetse flies in the Morsitans (e.g. *G. pallidipes* and *G. austeni*) and Fusca (e.g. *G. longipennis* and *G. brevipalpis*) groups (Table 1.7). Livestock, pigs, cattle and other small ruminants are common hosts for tsetse flies in these two groups as well as warthogs, bushbucks and wild buffalo. Tsetse flies in the Fusca group favour bushpigs, which can be found commonly in forests, woodlands and savannah habitats. However, flies from this group rarely feed on humans (Okoth *et al.*, 2007). Tsetse flies in the Palpalis group, which tend to live near rivers, generally feed on reptiles such as crocodiles and monitor lizards (Leak, 1998) but also on humans.

**Table 1.7** Host ranges of *G. pallidipes, G, austeni*, *G. brevipalpis* and *G. longipennis*.

| *Glossina spp.* | Wild animals | Domestic animals | Laboratory animals | Human |
|---|---|---|---|---|
| *G. pallidipes* | African buffalo*, bushpigs, warthogs*, other suids, giraffes, bushbucks*, dik diks, kudus, wildebeest, impalas, antelope, elands, suids, ostriches, baboons, lions, jackals, gazelles, hyenas, waterbucks, elephants, giraffes, Nile monitor lizards | cattle goats, sheep, donkeys, pigs, dogs, cats and porcupines | mice, cattle, goats and rabbits | rare |
| *G. austeni* | bushbucks*, warthogs*, buffalo*, other suids, bushpigs, kudu, antelope, reedbucks, rhinoceroses, elephants, giraffes, monitor lizards | cattle, sheep, goats, pigs, donkey and dogs | N/R | rare |
| *G. brevipalpis* | bushbucks*, buffalo*, kudus, warthogs, bushpigs*, other wild pigs, impalas, gazelles, rhinoceroses*, zebras, wildebeest, duiker, *Felidae,* elephants waterbucks, , hippopotamuses*, | donkey, cattle, sheep, goats, pigs and dogs | N/R | very rare |
| *G. longipennis* | rhinoceroses*, bushbucks*, bushpigs*, kudus, buffalo*, warthogs, red river hogs, hippopotamuses*, elephants, other wild pigs, elephants, *Felidae* (excl, lions), *Camelidae*, Rodentia reptiles and avan | cattle, sheep, goats pigs, dogs and donkey | N/R | very rare |

* The main hosts of *Glossina spp*. (Boakye *et al*., 1999, Clausen *et al*., 1998, Langridge *et al*., 1963, Leak, 1998, Muturi *et al*., 2011, Nyawira, 2009, Okoth *et al*., 2007); and N/R: no report

Although tsetse flies can feed on a wide range of hosts, there are three points that influence host choice: availability of hosts, self-defense of hosts and quality of blood meals (Glasgow and Weitz, 1956, Leak, 1998). Hosts of tsetse flies are limited by existing local variation in availability of vertebrate hosts so host choices tend to be different in rural regions and among different types of vegetation. Several studies have found that the majority of tsetse flies screened from the Nguruman region of Southern Kenya fed on warthogs (Okoth *et al*., 2007, Bett *et al*., 2008) but cattle and other *Bovidae* provided the highest proportion of blood meals in other regions of Kenya. In another study, of 13 engorged *G. pallidipes* caught in Nguruman, six fed on *Loxodonta africana* (Africa savannah elephants), five on *Phacochoerus africanus* (warthogs), one on *Syncerus caffer* (African buffalo) and one on *Papio hamadryas* (baboon) (Muturi *et al*., 2011). All of these hosts might be chosen depending on the microhabitats, the season and the landscapes of each location. For domestic livestock, bovines and camels have been found to be the major hosts of tsetse flies, but goats,

sheep, donkeys, horses, pigs (Farikou *et al.*, 2010b), and dogs and cats (Weitz, 1963) have also been reported. Thus, tsetse flies appear to be opportunistic in their feeding patterns.

Host available for tsetse feeding sometimes depends on the particular ecosystem. In the forest, seasonal changes can be a cause of animal migration and effect feeding preferences of tsetse flies (Pollock, 1982). In regions or times of year when there are fewer hosts, tsetse flies expand to less prerferred hosts. For example, rhinoceros and hippopotamus are preferred hosts of tsetse flies in the Fusca group, especially *G. longipennis* (Weitz, 1963). When rhinoceros are rare, the flies feed on other large animal populations: bushpigs; warthogs; hyenas; aardvarks; giraffes; ostriches; and humans. Thus, there could be very local patterns of host feeding, which could vary by season, habitat type or species of tsetse.

### 1.5.8 Blood meal analysis in tsetse flies

Determination of hosts of tsetse flies by their blood meal is often based on sequencing of PCR products designed to target mitochondrial genes. Using mitochondrial gene sequences for the taxonomic classification of animals has been widely used for DNA barcoding because both generic and specific primers have been designed that amplify across a range of animal species (Hebert *et al.*, 2003); thus, there are a lot of available data in Genbank to determine hosts of tsetse flies. Kocher *et al.* (1989) designed cytochrome b (cytb) L14841 and H15149 primers to amplify homologous segments of mitochondrial DNA from mammals, birds, amphibians, fishes and invertebrates (Kocher *et al.*, 1989). Ivanova *et al.* (2007) designed VF1 and used it with VR1 primers (named Fish Reverse 1 in Ward *et al.* (2005)) for identification of the cytochrome oxidase I (COI) gene in six mammal species: *Glaucomys volans* (Southern flying squirrel); *Sorex fumeus* (Smoky shrew); *Clethrionomys gapperi* (Southern red-backed vole); *Blarina brevicauda* (Northern short-tailed shrew); *Tamias striatus* (Eastern chipmunk); and *Tamiasciurus hudsonicus* American red squirrel) (Ivanova *et al.*, 2006). However, these primers have been found to also amplify DNA from mammals, reptiles and fishes (Ivanova *et al.*, 2007a) and so are useful for blood meal identification. Muturi *et al.* (2011) modified the cyt b primers of Kocher *et al.* (1989) and used these (named Cb1, Cb2) along with COI primers (VF1d_t1,

VR1d_t1, which were named VF1 and VR1 in Ivanova *et al*. (2006)) to track sources of tsetse blood meals. Using the Cb1 and Cb2 primers, African buffalo, warthogs, cattle, giraffes, spotted hyenas and baboons were reported as sources of blood meals of *G. swynertoni* from Tanzania and *G. pallidipes* from Kenya and Uganda but African buffalo, warthogs, cattle and Nile Monitor lizards were identified from the same sample set using the VF1d_t1 and VR1d_t1 primers. Thus, appropriate primers for identification sources of tsetse blood meals should also be determined by testing for reliability of amplification in local target populations.

## 1.6 *Trypanosoma spp*. infection in tsetse flies

### 1.6.1 Determinants of trypanosome infection in tsetse flies

Many factors could be related to vector competence of tsetse flies for trypanosomes. For example, Moloo *et al*. (1992) infected cattle and goats with trypanosomes, allowed different species of teneral tsetse flies (*G. pallidipes* and *G. m. centralis*) to feed on them, and then monitored infection rates in mice and goats exposed to the flies. The two species of flies showed different rates of infection, which also varied by species of trypanosomes and sex. However, both tsetse fly species had similar infection rates of *T. vivax* from a cow in Kenya and there was no significant difference of infection rates between males and females. Thus, vector competence of *G. pallidipes* was supposed to be equal to *G. m. centralis* for *T. vivax* whist *G. m. centralis* was more susceptible to *T. congolense* and *T. brucei*. However, from infected flies of both species, there was a 100% transmission rate to mice and goats. Extrinsic or environmental factors, intrinsic factors of tsetse flies and intrinsic factors of pathogenic trypanosomes all can influence trypanosome infection rates in tsetse flies. The following sections expand on what is known about each type of variable.

#### 1.6.1.1 Environmental factors

Changes of climate in Africa during the last 5 - 6 million years is thought to have influenced the prevalence of trypanosomiasis due to changes in the distribution

of host animals (Ebikeme, 2011, Gould and Higgs, 2009) and of tsetse flies (Brightwell *et al*., 1992). Temperature not only affects developmental rate and reproduction of tsetse flies but has also been found to affect the development of trypanosomes (Kinghorn and Yorke, 1912, Desowitz and Fairbairn, 1955). For example, infection rates of *T. congolense* in *G. morsitans centralis* and *G. brevipalips* incubated at 29°C in the puparial stage were higher than those that were incubated at 25°C (Ndegwa *et al*., 1992). Thus, climate change could influence the dynamics of trypanosome transmission due to direct effects on the parasites, their vectors and their final hosts.

#### 1.6.1.2 Tsetse fly factors

##### 1.6.1.2.1 Effect of tsetse species

Since there are many species of tsetse flies that can act as vectors for trypanosomes, but different species of tsetse are known to vary in their relative susceptibility to different species of trypanosome (Leak, 1998), this could have important implications for developing control measures. For example, tsetse flies in the Morsitans group, except *G. austeni*, are infected by and transmit all species of trypanosomes easily (Leak, 1998) but there is variation in relative susceptibility among tsetse species within the group, which also varies by species of trypanosomes. Although the mechanisms have not been fully established, it has been suggested that flies might have different defense mechanisms to prevent establishment of trypanosomes in different tissues, which would alter their relative sensitivity to particular species. For example, Peacock *et al*., (2012a) reported that *G. pallidipes* heavily defended against initial establishment in the midgut while *G. m. morsitans* had additional measurements to prevent trypanosome colonization of the salivary glands. In addition, the same species of tsetse flies from different sites can have different trypanosome susceptibility. For example, *G. pallidipes* from the Nguruman region of Kenya were more susceptible to *T. congolense* infection than a tsetse fly colony sampled from the Shimba Hills region (Moloo and Gooding, 2000). Tsetse flies in the Fusca group showed high infection and transmission rates of *T. congolense* and *T. vivax* compared to *T. brucei* while tsetse flies belonging to the Palpalis group have high infections and transmission rates for any species of trypanosomes (Leak, 1998). Thus, the species of tsetse flies present in a geographic region could affect the transmission dynamics to the final hosts.

Moreover, variation of trypanosome infection among individual flies could be affected by intrinsic factors of tsetse flies (sex, age, nutrition and immune status, physiological and biochemical mechanisms, and presence of endosymbionts), which also could vary by geographic region and could differ by species/strain of trypanosomes (Dyer *et al.*, 2013). Thus, it might be expected that trypanosome prevalence would depend on interactions among environmental factors, tsetse intrinsic factors, species of tsetse flies and species of trypanosomes. However, most studies so far have focused on one or a few factors rather than attempting a more holistic analysis of the combinations of these effects.

#### 1.6.1.2.2 Effects of sex

Previous studies have suggested that sex of flies can influence their susceptibility to trypanosomes but this could differ by species of both flies and trypanosomes. For example, Peakcock *et al*. (2012b) found that *T. brucei* had a higher maturation index (MI) in males, which they postulated was because parasite maturation is affected by one or more non-dosage-compensated X-linked loci in the host. Higher prevalence of trypanosomes in females could also be due to their longer lifespan compared to males (Nash, 1936). Sex can also affect relative susceptibility to trypanosomes by different age ranges of tsetse flies. According to Jackson (1946), old females are less active than younger ones, especially those in their first three weeks of their life (Jackson, 1946). Differences in activity levels could also help to explain the higher rate of infection in males than females in laboratory experiments. As reported by Moloo *et al*. (1992), male *G. m. centralis* showed higher susceptibility to *T. brucei* than females although this might differ by species; male *G. pallidipes* appeared to have slightly lower infection rates with *T. congolense* from Nigeria than females. In addition, males and females of *G. m. centralis*, *G. m. morsitans* and *G. pallidipes*, which were induced to *T. brucei* infection with D(+)-glucosamines (which blocks function of lectin, which inhibits trypanosome establishment), showed equal infection rates (Mihok *et al.*, 1992). This result was supported by trypanosome infections in *G. pallidipes* and *G. m. morsitans* from Zimbabwe, where there was not a significant difference between males and female (Woolhouse *et al.*, 1993) but female *G. m. morsitans* had additional measures for trypanosome prevention in salivary glands (Peacock *et al.*, 2012b).

#### 1.6.1.2.3 Effects of age

The age of tsetse flies can also affect their susceptibility to trypanosomes. *Glossina m. morsitans* (using a colony from the Liverpool School of Tropical Medicaine, LSTM) and *G. palpalis palpalis* (using a colony from international Atomic Energy Agency, IAEA, Entomology Labortories, Siebersdorf, Austria) were experimentally infected with laboratory strains of *T. brucei* TSW196 and *T. congolense* 1/148 BSF (Walshe *et al.*, 2011). Teneral tsetse flies, which were less than 24 hours post-eclosion (h.p.e.), were found to be twice as susceptible to trypanosome infection as tsetse flies aged 48 h.p.e. In another study, the infection rate of *T. b. rhodesiense* to teneral *G. m. m*orsitans (54.3%) was found to be higher than when the flies were eight days post-eclosion (d.p.e.) (3.4%) (Weiss *et al.*, 2013). In contrast, a survey of over 9,000 *G. pallidipes* from the Zambezi Valley, Zimbabwe (using DNA probes) found that prevalence of *Trypanosoma spp*. increased with age (Woolhouse *et al.*, 1993). However, prevalence of *T. congolense* in the same population decreased with age. Each species of tsetse has a different lifespan so classifications of old vs young could also differ by species; for example, female *G. austeni* have been observed to have a shorter lifespan than *G. pallidipes* in nature (Wamwiri *et al.*, 2013). This means that the opportunity of infection in *G. austeni* over their lifetime could be less than for *G. pallidipes*. Thus, based only on field surveys, establishing differences in susceptibility among species of tsetse species is confounded by differences in the relative risk of exposure with age.

Interactions between age and other tsetse factors could also influence infection patterns. Both age and sex influence the activities and feeding behaviours of tsetse flies. Jackson (1946) cultivated pupae and reported on the development and general activities of teneral flies. He found that they were highly active on the $2^{nd}$ – $4^{th}$ days after emergence from pupae but became inactive again after they finished their first few blood meals. When they were older, all activities decreased, but especially in females (Jackson, 1946). The feeding behaviour of females involves balancing the requirement of getting enough blood to nourish the developing larvae with allowing enough space in the abdomen for their development. (Hargrove *et al.*, 2011). Females thus might feed more frequently than males because they are forced to take smaller blood meals. Tsetse flies with a high frequency of feeding pose higher risks of transmitting trypanosome

infections but also will be at higher risk of feeding on infected animals. Moreover, age of the flies is related to the level of immunity (Kaaya and Darjt, 1988). Young tsetse flies have lower immunity than older individuals because innate immunity may be incompletely developed; for example, they might not yet have developed specific immunity to defense against trypanosome invasion. In this case, the older flies would be more resistant**.** Thus, it is important to consider both sex and age when predicting relative susceptibility of risk of exposure to trypanosomes.

#### 1.6.1.2.4 Effects of bloodmeal, nutritional status and immune status

Nutrients or contents in the blood of different animals might be associated with variation in infection rates of trypanosomes in tsetse flies. For example, *G. m. morsitans* and *G. m. centralis*, which were experimentally fed on trypanosome infected blood from goats showed a higher infection rate than from sheep, monkeys or antelope (Aksoy *et al.*, 2003). The immunological complement in blood serum might suppress infection rates (Walshe *et al.*, 2011). However, sialic acids (SAs) from host blood glycoproteins might enhance trypanosome's ability to evade immune defenses (Nagamune *et al.*, 2004). Newly emerged, previously unfed tsetse (teneral flies), show a profound susceptibility to trypanosome infection (Haines, 2013), possibly due to the nutritional stress from starvation, which influences immunity relating to trypanosome protection (Akoda *et al.*, 2009b). For example, starving 20-day-old flies increased the colonization rate of the midgut with *T. congolense* (Kubi *et al.*, 2006), and the maturation index (MI) of *T. b. brucei* was altered when the midgut infection was established prior (Akoda *et al.*, 2009a) or subsequent to starvation (Kubi *et al.*, 2006).

Trypanosomes are very effective at escaping from specific immune defense mechanism due to the ability to rapidly change the expression of their large repertoire of variant surface glycoproteins (VSGs) (Barry and Turner, 1991), which poses substantial challenges for the immune system of the flies to prevent their invasion. Starvation decreases tsetse glutamic acid–proline (EP) proteins and other defensive immune responses, leading to increased trypanosome susceptibility (Haines *et al.*, 2010). The EP protein is thought to be involved in the specific immune responsive by acting as an antagonist of trypanosome infection during initial establishment in the midgut (Haines *et al.*, 2005) because

it directly responds to changes in the VSGs of the parasites (Haines *et al.*, 2010). In addition, the innate immune system is likely important for regulating the interaction between trypanosomes and tsetse flies (Hao *et al.*, 2001). However, to develop completely it needs time and nutrients, such as lectin, which is involved in immunity (Drummond and Brown, 2013) and can be found in the haemolymph and midgut. Midgut lectin can protect the flies from trypanosome establishment through an apoptosis-like mechanism which stimulates death of trypanosomes (Murphy and Welburn, 1997). Teneral flies, which have not yet fed, lack lectin but levels increase after feeding and accumulate with age. However, lectin in haemolymph may induce establishment of trypanosomes in tsetse maturation of *T. congolense* and *T. b. rhodesiense* (Welburn and Maudlin, 1990). Glucosamine is thought to be an antagonist to lectin, but has not been demonstrated to directly affect parasite loads in the flies (Mihok *et al.*, 1992). For the acquired immune system, neutralizing antibody is the response controlling the trypanosome parasitemia (Morrison and Murray, 1985). Cytokines, such as Interleukin (IL), tumor necrotic factor α (TNF-α) and interferon γ (IFN-γ) are not directly involved (Kaushik *et al.*, 1997). Anti-microbial peptides (AMP) can also destroy trypanosomes. AMP knockdown of the transcription of attacin, which is an antibacterial protein (Wang *et al.*, 2008), leads to an increase in the trypanosome infection rate (Boulanger *et al.*, 2002). However, proper development of immunological functions does not rely only on nutritional status; symbiotic microflora can also affect immunological protection (Weiss *et al.*, 2011). For example, resistance of trypanosome challenged tsetse flies that were treated with ampicillin was lower than wild tsetse flies with the bacterial endosymbiont *Wigglesworthia* (Aksoy *et al.*, 2003, Pais *et al.*, 2008). Experimental infection of *E. coli* in tsetse flies induced immune stimulation, thereby reducing infection from trypanosome-infective feeds, and the anti-microbial peptides attacin, diptericin and defensin were detected in tissues of tsetse flies (Hao *et al.*, 2001). However, the complete dynamics of immune regulation in different species of tsetse flies has not been resolved.

#### 1.6.1.2.5 Effects of physiological and biochemical mechanisms

In addition to the immune system of tsetse flies, trypanosomes also encounter other types of barriers to differentiation, proliferation and migration in tsetse flies. The microenvironment in the midgut of tsetse flies affects trypanosome

maturation: the most suitable temperature should be a constant at 25°C (Roditi and Lehane, 2008). For example, low temperature may cause cold shock of trypanosomes in tsetses (Engstler and Boshart, 2004). Some physical and biochemical components also can provide barriers to trypanosome establishment and maturation, such as gut pH, proteases (Liniger *et al.*, 2003), the glycosaminoglycan-rich layer with chitin (peritrophic matrix; PM) (Lehane, 1997), reactive oxygen species (ROS) (Hao *et al.*, 2003) and anti-oxidant molecules (MacLeod *et al.*, 2007b). In both teneral and fed tsetse flies, the average pH is 10.6 in the proventriculus (Liniger *et al.*, 2003), 6.5 in the midgut (Wigglesworth, 1929) and 7.9 in the lumen of the hindgut (Liniger *et al.*, 2003). Thus, trypanosomes remaining in the anterior midgut will be exposed to a different environment than those in the posterior midgut. By contrast, procyclic forms of *T. brucei* are in-vitro cultured in pH 7.1 – 7.7 media, which is not the optimal pH range for growth (Brun, 1979). For the protease activity, Glossina chymotrypsin may cause trypanosome agglutination but it is inhibited by glucosamine (GlcN) and N-Acetyl glucosamine (GlcNAc) (Abubakar *et al.*, 2006). Adding anti-oxidant molecules to tsetse blood meals has been found to support trypanosome establishment and maturation (Macleod *et al.*, 2007a, MacLeod *et al.*, 2007b). The PM is a semipermeable membrane that separates the midgut epithelium from the feeding content and protects the midgut from mechanical damage, toxins and invasion of pathogens (Lehane, 1997). The PM of teneral tsetse flies is “ragged and discontinuous” and does not extend along the entirety of the midgut (Wigglesworth, 1929), enabling trypanosomes to penetrate to the ectoperitrophic space, where they multiply and continue to differentiate (Gibson and Bailey, 2003). Several of nearly 300 identified proteins contain signature Chitin Binding Domains (CBD), including novel peritrophins and peritrophin-like glycoproteins, which are essential in maintaining PM architecture and may act as trypanosome adhesins (Rose *et al.*, 2014). However, complete development of the PM is destroyed by chitinase, allowing trypanosome penetration. There has been no report that chitinase occurs naturally in tsetse flies (Roditi and Lehane, 2008) and so it was concluded that *Sodalis* in the midgut is associated with chitinase activity (Hao *et al.*, 2003, Munks *et al.*, 2005). Thus, even if tsetse flies can bypass the host immune system, there are multiple other barriers that challenge their establishment and development.

#### 1.6.1.2.6 Effects of *Sodalis glossinidius* on tsetse flies

The relationship between *S. glossinidus* and trypanosome infection in tsetse flies remains unclear. Some experiments have supported the hypothesis that this bacterial endosymbiont supports trypanosome establishment in tsetse flies while others have suggested that *Sodalis* induces immnunological defence against trypanosome infection in tsetse flies, and some experiments have found that there is no direct relationship.

Susceptibility of *T. congolense* infection in tsetse flies has been associated with the presence of *S. glossinidius* in the epithelial cells of the midgut (Maudlin and Ellis, 1985) and tsetse flies harboring the endosymbionts as their normal flora have been shown to have increased susceptibility to trypanosomes compared to flies lacking *Sodalis* (Welburn *et al.*, 1993, Welburn and Maudlin, 1999). The increasing density of *S. glossinidius* in the midgut may induce the flies to become infected by trypanosomes more easily (Cheng and Aksoy, 1999) by favouring their establishment via a complex biochemical mechanism (Dale and Welburn, 2001). For example, chitinase from *S. glossinidius* can enhance trypanosome establishment in the midgut of tsetse flies by allowing permeation of the PM (Welburn *et al.*, 1993). In field surveys using molecular techniques, a positive relationship has sometimes been found between the incidence of *S. glossinidius* and the presence of trypanosomes in tsetse flies. Prevalence of trypanosome infections in Bipindi and Campo of South Cameroon were around 75%, which was similar to the rate of *S. glossinidius* harboring in tsetse flies in these regions (Farikou *et al.*, 2010a). Tsetse flies with the endosymbionts have also been found to have a three times greater risk of trypanosome infection than those without (Soumana *et al.*, 2013b). Therefore, while *Sodalis* is not essential for trypanosome midgut infection, it has been suggested that these endosymbionts play a role in enhancing the establishment of trypanosome infections in the midgut of tsetse flies (Dyer *et al.*, 2013, Welburn *et al.*, 1993, Welburn and Maudlin, 1991). However, other studies have not found an exclusive relationship between the presence of *S. glossinidius* and the susceptibility of tsetse flies to *T. congolense* (Geiger *et al.*, 2005b). Moreover, it has been reported that the endosymbiont themselves induce immune function in the flies, which could decrease the levels of trypanosome infection (Hao *et al.*, 2001). Diptericin is an antimicrobial immune effector, which is expressed by *Sodalis* in

the gut and hemolymph of the flies (Lehane *et al.*, 2004). Furthermore, 27 proteins from *Sodalis* were suggested to have a close association with the tsetse PM (Rose *et al.*, 2014). When tsetse flies were treated with streptozotocin for selective elimination of *Sodalis* (which did not kill the obligate tsetse symbiont *W. glossinidia*), there was a 40% reduction in susceptibility to trypanosome infection in the midgut of progeny of treated flies (Dale and Welburn, 2001). The prevalence of *T. b. rhodesiense* infection in *G. m. m*orsitans that were cured of *Sodalis* by ampicillin treatment (5.3%) was higher than for wild tsetse flies that tested positive for *Sodalis* (3.4%) (Weiss *et al.*, 2013), which they concluded was due to a higher immune response in the latter. Although, *S. glossinidius* occurred in both *G. pallidipes* and *G. m. centralis*, *T. congolense* infection rates in the midgut of *G. pallidipes* were lower than in *G. m. centralis* (Shaw and Moloo, 1991), suggesting that it is not just the presence of *Sodalis* that is important for the infection. Therefore, *Sodalis* might be essential for trypanosome defense in some conditions. However, some authors have also concluded that there is no association between *Sodalis* and trypanosome infection in tsetse flies (Moloo and Shaw, 1989, Shaw and Moloo, 1991, Geiger *et al.*, 2005b, Dennis *et al.*, 2014). Therefore the role that Sodalis plays in vector competence remains to be completely elucidated.

To understand more about these complex relationships, the genetics of *S. glossinidius* have also been studied to determine whether it might be particular strains of the endosymbiont that influence trypanosome susceptibility rather than simply presence or absence. The population genetic structure of *Sodalis glossinidius* from different areas may be influenced both by geographic barriers and by specificity to different hosts (Farikou *et al.*, 2011b). Results from Geiger *et al.* (2007) showed that *T. b. gambiense* and *T. b. brucei* infection in *G. p. gambiensis* and *G. m. morsitans* was linked to specific genotypes of *S. glossinidius* based on amplified fragment length polymorphism analysis (AFLP) suggesting that vector competence of tsetse flies might be related to genotypes of *S. glossinidius* rather than to mere presence/absence of the symbiont (Farikou *et al.*, 2010a). Together, this suggests that different strains of the endosymbiont might have different effects on trypanosome susceptibility in the flies.

*Sodalis* may not be presented in all species of *Glossina* and each *Glassin*a species may carry different bacterial loads (Aksoy *et al.*, 2014). For example, prevalence

of *S. glossinidius* in *G. pallidipes* from the Shimba Hills and Arabuko-Sokoke National Reserve of Kenya was 16.0% compared to 3.7% in *G. austeni* from the same sites of collection (Wamwiri *et al.*, 2013). *Glossina austeni* and *G. brevipalpis* both showed relatively low bacterial loads when analysed based on five intensity levels of PCR amplification (negative or four levels of positives) (Cheng and Aksoy, 1999) and microscopic methods (Moloo and Shaw, 1989). In addition, the age of tsetse flies may be associated with prevalence of *Sodalis*, with a higher proportion of older individuals testing positive for *Sodalis* when screened with PCR methods than young flies (Soumana *et al.*, 2013a). Thus, interactions between *Sodalis*, tsetse flies and trypanosomes may depend on multiple other factors that could obscure the real relationship between trypanosome susceptibility and the endosymbionts.

#### 1.6.1.3 Trypanosome factors

Sites of establishment, maturation, and the life cycle of different species/strains of trypanosome in tsetse flies are factors that can affect their relative prevalence. *Trypanosoma vivax* develops exclusively in the mouthparts with its short life cyle so it has low opportunity to be exposed to anti-trypanosomal factors, such as attacin (Hu and Aksoy, 2006), lectin (Aksoy *et al.*, 2003) and proteinase (Roditi and Lehane, 2008) that are produced in the midgut. This is different from the pathogenesis of *T. brucei* and *T. congolense*, which initially establish in the midgut (Roditi and Lehane, 2008). Thus, *T. vivax* tends to show the highest infection rate in tsetse flies. This supports conclusions from a previous study of Moloo *et al.* (1992) conducted in Tanzania, Kenya and Nigeria, which showed that *G. pallidipes* had a higher infection rate with *T. vivax* than with *T. congolense* or *T. brucei*. Although non-dosage-compensated X-linked loci controls the maturation index (MI) of *T. brucei* in males (Gooding and McIntyre, 1998), metacyclogenesis of *T. congolense* in the proboscis of *G. morsitans* is independent (Dale *et al.*, 1995). For *T. brucei*, the MI is higher for faster maturing strains (Dale *et al.*, 1995), suggesting a limited time window for successful maturation after midgut infection. In addition, virulent strains of trypanosomes in animals might cause high infection rates in tsetse flies. For example, *T. congolense* strains with high virulence in mice showed also higher infection rates in the midguts of tsetse flies compared to strains with moderate or low virulence (Masumu *et al.*, 2006).Thus, it is important to consider the

species of both tsetse flies, and trypanosomes present when investigating the transmission dynamics of the parasites.

Although survival and maturation of trypanosomes in tsetse flies are known to depend on many factors (Dyer *et al.*, 2013), what has been lacking are studies that consider the full complexity of the biology of the vector-parasite-blood meal system. A recent study has demonstrated why it could be misleading to study individual factors in isolation. Using multiple regressions, the main biological and ecological determinants of trypanosome prevalence in *G. p. palpalis, G. tachinoides* and *G. morsitans* from northern Nigeria were found to be the fly sex, with more trypanosomes found in females than males, and site, with *T. congolense* subspecies being more abundant in Yankari than in Wuya (Isaac *et al.*, 2016). However, there was a significant interaction between sex and site, but only for some species of trypanosomes and for only some species of flies. Thus, some of the discrepancies between previous studies on factors affecting trypanosome prevalence could be because interactions that could affect interpretation of the results have not been considered. There are also other factors thought to be associated with trypanosome infection in tsetse flies that could show complex interactions; for instance, age, *Sodalis* status and composition of host communities. Considering potential interactions among all of these factors will provide more realistic biological information on trypanosome infection in tsetse flies and benefit understanding of the epidemiology of trypanosomes.

### 1.6.2 Trypanosome identification based on PCR method

Microscopy has been used to determine the species of trypanosome present based on the sites of trypanosome establishment and development by dissection of the flies and screening each organ separately (Lloyd and Johnson, 1924, Enyaru *et al.*, 2010, Desta, 2014). Infection by trypanosomes in the *Dutonella* subgenera (mainly *T. vivax*) is determined when parasites are identified only in the proboscis parts of tsetse flies, whereas *T. congolense* and other trypanosomes in the *Nannomonas* subgenera occur both in the proboscis and midgut parts, and *T. brucei* and other *Trypanozoon* invade the salivary glands as well (Njiru *et al.*, 2004). This method is cheap and does not require specialised materials but it is laborious work that needs a trained specialist, allows only

subgenera of trypanosomes to be identified and has low sensitivity, particularly in early stages of infection when parasite numbers are low. It is also difficult to detect mixed infection of trypanosome species because it is difficult to differentiate each species of trypanosomes in trypomastigote stages.

The polymerase chain reaction (PCR) facilitates the identification of trypanosome infections in tsetse flies because it is more sensitive than microscopy and can detect very low numbers of parasites, and it is more reliable for identifying individual subspecies of trypanosomes and mixed infections (Masiga *et al.*, 1996, Morlais *et al.*, 1998, Ouma *et al.*, 2000, Njiru *et al.*, 2004, Alibu *et al.*, 2015). Many PCR techniques have been applied for trypanosomes identification in vertebrate hosts and vectors. For example, species-specific primers have been developed to distinguish the various *T. congolense* clades (Moser *et al.*, 1989, Masiga *et al.*, 1992), *T. brucei* sub-species (Moser *et al.*, 1989), *T. simiae* (Masiga *et al.*, 1992), and *T. vivax* (Masiga *et al.*, 1992) and general primers targeting the ribosomal DNA array (often the internal transcribed spacer regions, or ITS) have been used to identify particular species of trypanosomes based on different sizes of amplification products (Njiru *et al.*, 2005). However, different approaches could have different sensitivity of detection.

Effort has been taken to improve the utility of PCR-based methods by designing increasing numbers of primer combinations to target particular species. Mouser *et al* (1989) designed TCN-1 and TCN-2 primers for *T. congolense* savannah and TBR-1 and TBR-2 for *T. brucei* identification, which they were concluded 100 times better at diagnosing infections than DNA radioactive probes (Moser *et al.*, 1989) based on experimental infections of outbred laboratory mice with lab-raised positive controls for *T. congolense*, *T. b. gambiense*, and *T. b. brucei*. Specific primers targeting *T. simiae* (TSM1 and TSM2), *T. congolense* savannah (TCS1 and TCS2), *T. congolese* forest (TCF1 and TCF2), *T. congolese* kilifi (TCK1and TCK2), and *T. vivax* West Africa (TVW1 and TVW2) were designed to enable identification of a broader range of species and subspecies (Masiga *et al.*, 1992). The reliability of the various primer combinations were tested using experimental infections of lab-reared colonies of *G. m. morsitans* (from the Tsetse Research Laboratory, TRL, Bristol), and they were found to be sensitive not only to particular species but could also detect different developmental

stages of trypanosomes and mixed infections. The authors recommended that the methodology is suitable for large-scale study of the epidemiology of trypanosome infections. However, these primers were all tested using laboratory strains of tsetse flies and parasites, without extensive investigation of whether sequence variation among wild populations might affect the sensitivity and reliability of the primers.

Using general primers to amplify all species of parasites present in a set of samples or to resolve mixed infections in single samples has also been a popular approach but relies on lack of variation within trypanosome species in the size of the amplification products produced. The internal transcribed spacers (ITS), noncoding regions situated between the small subunit (SSU) and large subunit (LSU) rRNA genes (Figure 1.5), have been the most widely used for identification of a wide range of parasite species because of their rapid rate of evolution, particularly through extensive insertions and deletions that result in characteristic variation in length among species. For trypanosomes, these regions have been demonstrated to vary in sizes among species and occasionally subspecies (Cox *et al.*, 2005) so they have been widely used for large-scale screening of trypanosome infections (Morlais *et al.*, 1998). McLaughlin *et al.* (1996) originally developed primers (named Kin) to identify the ITS-1 region of kinetoplastid species in general and showed that while there was large variation among species of trypanosomes, apicomplexan species, or bacterial or mammalian amplification products were not detected (Desquesnes *et al.*, 2001b). Cox *et al.* (2005) subsequently developed general primers that spanned the entire rDNA array and used nested PCR to allow detection of very low parasite loads. However, screening results from the ITS nested PCR are not always well correlated with those based on species-specific primers (Ahmed *et al.*, 2013). As might be expected the nested PCR produced larger numbers of positive diagnoses than specific primers for identification of some species (e.g. *T. vivax* and the three clades of *T. congolense*). However, it produced fewer positives than the specific primers for *Trypanozoon* species or than multiplex PCR for identification of *T. brucei* subspecies. Thus, these original nested primers might not result in equally efficient amplification of all trypanosome species. Other sets of primers have also been developed in an attempt to reduce such discrepancies. For example, a wide range of trypanosome species and

mixed infections were detected in the midguts of wild tsetse flies from Tanzania using ITS-1 nested PCR designed by Adams *et al*. (2006). Thus, choice of PCR primer could affect interpretations of trypanosome prevalence and infection rates in both vertebrate hosts and vectors.

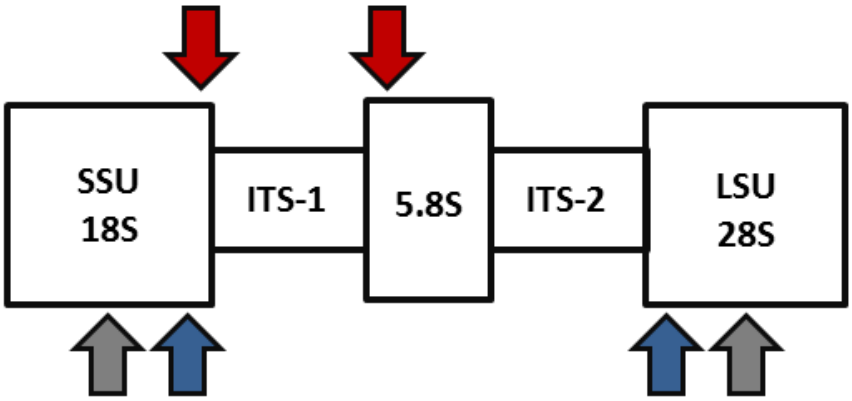



**Figure 1.5** The structure of the ribosomal RNA array spanning the internal transcribed spacer regions (ITS-1 and ITS-2).

Large boxes represent the conserved coding regions (SSU, small sub-unit; and LSU, large subunit), while the small boxes represent the spacer regions. The target sites of the ITS-1 CF and ITS-1 BR primers, developed by Njiru *et al*. (2005), are indicated by red arrows. A set of nested primers, designed by Cox *et al*. (2005), are represented by grey arrows for the first-round (outer) PCR primers (ITS-1 and ITS2) and blue arrows for the 2nd round (inner) PCR primers (ITS3 and ITS4).

Due to the increased risks of contamination using nested approaches (because smaller amounts of DNA can be amplified), non-nested approaches have also been developed. Nijuru *et al*. (2005) used well-characterised reference DNA of trypanosomes (the *Nannomonas* and *Duttonella* groups obtained from W. C. Gibson, University of Bristol; the *T. lewisi* and *T. theileri* from the KETRI cryobank; and trypanosome DNA from trapped *G pallidipes* and *G. longipennis* from Kenya) as standards to directly compare the reliability of non-nested and nested general primers compared to species-specific primers. The non-nested ITS-1 CF and ITS-1 BR primers (developed by Davila, unpublished) shower higher diagnostic sensitivity than the nested Kin primers, with the sensitivity being close to that based on species-specific primers: 95% for *Trypanosoon*; 90.0% for *T. congolense* and 77.4% for *T. vivax*. They thus recommended the use of the ITS-1 CF and ITS-1 BR primer set for routine trypanosome diagnosis. However,

using both types of ITS-1 primers (i.e. nested and non-nested) for trypanosome investigation in blood samples of cattle from Kenya showed significantly higher positive results than specific primers ($\chi^2$ = 10.7, df = 1, P-value = 0.001) (Thumbi *et al.*, 2008). Trypanosome prevalence was 10.7% (2.9% *T. congolense* savannah, 1.0% *T. congolense* kilifi, 3.9% *T. vivax* and 2.9% *T. brucei*) using four specific primers (TCN-1 and TCN-2, TCK1 and TCK2, TVW1 and TVW2 and TBR-1 and TBR-2) and no mixed trypanosome species infections were detected. In contrast, for the ITS-1 CF and ITS-1 BR primers, trypanosome prevalence was 26.2% (4.9% *T. congolense*, 17.5% *T. vivax* and 3.9% *T. brucei*), which was similar to the results from nested ITS-1 PCR (28.1%; 6.7% *T. congolense*, 23.3% *T. vivax* and 1.9% *T. brucei*). Mixed infections of trypanosome species were also detected 1.9% and 3.8% using the general primers. Thus, the most efficient approach might be to use the general single PCR primers for initial trypanosome screening and then following up with the specific primers to confirm the infection in this investigation. One limitation to these studies is that individual species of trypanosomes have not been confirmed by sequencing, which is crticial for assessing the reliability of primers when used for natural field surveys.

In addition to choice of reliable primers, the most appropriate tissue for trypanosome screening in tsetse flies is still a question when investigating a wide range of trypanosome species, due to the differences in where the various species reside. The midgut is generally dissected to investigate *T. congolense* and *T. brucei* because they establish and develop in this organ while the salivary gland is appropriate only for *T. brucei* screening. While the abdomen is easier to dissect and so using only this tissue might reduce risks of contamination due to more handling needed to isolate the salivary glands. Using only this tissue might lead to underestimation of *T. vivax* infection since its life cycle occurs in the proboscis parts. Thus, trypanosome investigation of whole flies has also been used for trypanosome identification (Salekwa *et al.*, 2014). Nevertheless, some compounds found in blood meals can inhibit amplification during the PCR cycle (Alaeddini, 2012). Head plus proboscis parts could thus be a good choice for trypanosome identification in tsetse flies because there are thought to be fewer inhibitory compounds than in the gastrointestinal tract, all of the main trypanosome species reside in those parts, and this is the site of transmission of all Salivarian trypanosomes through biting.

### 1.6.3 Trypanosome prevalence in tsetse flies

Prevalence of trypanosomes is known to vary extensively for different species of tsetse flies sampled from different regions (Table 1.8) and for different species of trypanosomes (Table 1.9). However, methods of trypanosome identification can influence interpretation of the presence of the parasites. For example, Wamwiri *et al*. (2013) reported low prevalence of trypanosomes in *G. austeni* and *G. pallidipes* using microscopic methods, but the prevalence was found to be more than two times higher when screened with PCR (Table 1.8). *Trypanosoma vivax* is the predominant species found in most studies (Table 1.9). Moreover, trypanosome prevalence can be different between the sexes of flies and sometimes has been found to increase with age but this can again vary by the species of trypanosome, the species of fly, or the geographic region sampled. For example, trypanosome prevalence of males (13.9%; 11.6% for *T. vivax*; and 2.3% for *T. congolense* forest) was slightly higher than for females (11.6%; 10.3% for *T. vivax*; and 1.3% for *T.* congolense forest) in Mtito Andei Division, Makueni County, Kenya (Nthiwa *et al*., 2015). On the other hand, trypanosome prevalence in female *G. pallidipes* (31.9%) and *G. austeni* (24.0%) from Ethiopia was higher than for males (18.4% and 21.43%, respectively) (Bitew *et al*., 2011). Prevalence of trypanosomes in the midgut of *G. pallidipes* caught from the Luangwa Valley, Zambia, ranged from 0% in wing fray category 1 to 12.6% in category 6 for *T. vivax* and 7.9% in category 6 for *T. congolense* using DNA probes (Woolhouse *et al*., 1993). Thus, different species of tsetse flies with different sexes and age from different geographic regions could carry different relative rates of different trypanosome species to transmit to their hosts.

**Table 1.8** Prevalence of *Trypanosoma spp*. in different species of tsetse flies in different geographic regions in Africa based on microscopy and PCR-based methods (2013-2015).

| Tsetse flies | Trypanosome prevalence | Sites of samples | Identification methods | References |
|---|---|---|---|---|
| *Glossina spp.*<br>*G. pallidipes*<br>*G. morsitans*<br>*G. fuscipes*<br>*G. tachinoides* | 6.0%<br>5.9<br>10.6<br>1.0<br>3.4 | western Ethiopia (longitudinal survey) | Microscopy | (Desta *et al.*, 2013) |
| *G. austeni* &<br>*G. pallidipes* | 6.0%<br>4.9% | Kenya and South Africa | Microscopy | (Wamwiri *et al.*, 2013) |
| *G. austeni* | 13%<br>12.7%<br>0%<br>8% | ASNR[a]<br>SHNR[b]<br>LSLSA[c]<br>Norhtern, South Africa | | |
| *G. pallidiles* | 9.6% | SHNR | | |
| *G. austeni* &<br>*G. pallidipes* | 13%<br>10% | Kenya and South Africa | ITS-1 CF and BR primers (Njiru *et al.*, 2005) | |
| *G. brevipalpis*<br>*G. m. morsitans*<br>*G. pallidipes* | 56.4%<br>83.9%<br>29.1% | Luambe National Park, Zambia. | ITS-1 CF and BR primers (Njiru *et al.*, 2005) | (Dennis *et al.*, 2014) |
| *G. morsitans* &<br>*G.pallidipes* &<br>*G. swynertonni* | 3.0%* | Simanjiro district, Northern Tanzania | ITS-1 CF and BR primers (Njiru *et al.*, 2005) | (Salekwa *et al.*, 2014) |
| *G. pallidipes*<br>*G. longipennis* | 5.8%*<br>23.1%* | Mtito Andei Division, Makueni County, Kenya. | Nested ITS-1 PCR (Cox *et al.*, 2005) | (Nthiwa *et al.*, 2015) |
| *G. p. palpalis* | 25.5%<br>19.3%<br>20.8%<br>41.7%<br>54.3%<br>0% | Cameroon (overall)<br>Eloa<br>Bouraka<br>Gueboba<br>Guefigue<br>Guientsing | ITS-1 primers (Desquesnes *et al.*, 2001b) | (Simo *et al.*, 2015). |

[a] ASNR: the Arabuko-Sokoke National Reserve, Kenya; [b] SHNR: the Shimba Hills National Reserve, Kenya; [c] LSLSA: South of Lake St. Lucia, South Africa

**Table 1.9** Prevalence of *T. congolense*, *T. vivax*, *T. brucei*, *T. simiae*, *T. godfreyi* and mixed infections in tsetse flies from different geographic regions in Africa based on microscropy and PCR-based methods (2011-2015).

| Tsetse flies | Trypanosome prevalence | Sites of samples | Identification methods | References |
|---|---|---|---|---|
| *G. pallidipes* | 7.5% (Tc)<br>16.5% (Tv)<br>0.5% (Tb) | the Gojeb valley of Ghibe Omo river system, Southwest of Ethiopia | Microscopy | (Bitew *et al.*, 2011) |
| *G. f. fuscipes* | 8.2% (Tc)<br>14.7% (Tv)<br>0% (Tb) | | | |
| *G. austeni* | 8.0% (Tc)<br>6.1% (Tc)<br>0% (Tc)<br>4.0% (Tc) | ASNR[a]<br>SHNR[b]<br>LSLSA[c]<br>Northern, South Africa | ITS-1 CF and BR primers ITS-1 CF and BR primers (Njiru *et al.*, 2005) | (Wamwiri *et al.*, 2013) |
| | 3.0% (Tv)<br>5.6% (Tv)<br>0% (Tv)<br>4.0% (Tv) | ASNR[a]<br>SHNR[b]<br>LSLSA[c]<br>Northern, South Africa | | |
| | 1.0% (Ts)<br>1.0% (Ts)<br>0% (Ts)<br>0% (Ts) | ASNR[a]<br>SHNR[b]<br>LSLSA[c]<br>Northern, South Africa | | |
| *G. pallidipes* | 3.6% (Tc) | SHNR[b] | | |
| | 3.6% (Tv) | | | |
| | 2.3% (Ts) | | | |
| *G. brevipalpis* | 14.5% (Tv)<br>10.9% (Tc)<br>0% (Tg)<br>18.2% (Ts) | Luambe National Park, Zambia. | ITS-1 CF and BR primers (Njiru *et al.*, 2005) | (Dennis *et al.*, 2014) |
| *G. m. morsitans* | 32.9%(Tv)<br>14.6% (Tc)<br>2.2% (Tg)<br>2.2% (Ts) | | | |
| *G. pallidipes* | 7.2% (Tv)<br>6.0% (Tc)<br>9.5% (Tg)<br>3.1% (Ts) | | | |
| *G. morsitans* | 4.3% (Tv)<br>0.95% (Tv+Tb)<br>0% (Tv+Tc) | Simanjiro district, Northern Tanzania | ITS-1 CF and BR primers (Njiru *et al.*, 2005) | (Salekwa *et al.*, 2014) |
| *G. pallidipes* | 0% (Tv)<br>0% (Tv+Tb)<br>0% (Tv+Tc) | | | |
| *G. swynertonni* | 2.9% (Tv)<br>0.27% (Tv+Tb)<br>0.13% (Tv+Tc) | | | |
| *G. pallidipes* | 3.9% (Tv)<br>1.9% (Tc) | Mtito Andei Division, Makueni County, Kenya. | Nestsed ITS-1 PCR (Cox *et al.*, 2005) | (Nthiwa *et al.*, 2015) |
| *G. longipennis* | 23.1% (Tv)<br>0% (Tc) | | | |

[a] ASNR: the Arabuko-Sokoke National Reserve, Kenya; [b] SHNR: the Shimba Hills National Reserve, Kenya; [c] LSLSA: South of Lake St. Lucia, South Africa.

Tc: *T. congolense*; Tb: *T. brucei*; Tv: *T. vivax*, Ts: *T.* simiae; and Tg: *T. godfreyi*.

### 1.6.4 Tsetse-fly control

Since there are limited vaccines or drugs that have been developed to target trypanosomes, disease management mainly relies on vector control (Aksoy, 2003). Methods for vector control are classified into three groups: environmental management, insecticide, and modification of flies (Table 1.10). Destructions of tsetse habitat and elimination of tsetse reservoir hosts have been discarded for ecological and environmental concerns so most current vector control interventions involve insecticides.

**Table 1.10** Methods of tsetse fly control (Rogers *et al.*, 1994, World Health Organization, 2016)

| Methods | Advantage | Disadvantage |
|---|---|---|
| Surround management | | |
| bush clearing | disturb life cycle | ecological and environmental concerns |
| elimination of wild animals | limit hosts | ecological and environmental concerns |
| Insecticide application | | |
| ground spraying | very effective with wide space control | environmental concerns |
| insecticide-treated animals | effective for farm animals | targets only fed flies |
| insecticide-treated targets (a combination of traps and artificial baits) | effective, low cost benefit ratio | different type of targets have different species of tsetse flies |
| sequential aerosol spraying technique (SAT) | very effective with wide area of tsetse control | high cost and environmental concerns |
| Fly Modification | | |
| sterile insect technique (SIT) | wide area of tsetse control | high cost |

Tsetse flies are susceptile to many insecticides (Hadaway, 1972) but only some chemicals with specific properites have been used so far; for example, dichlorodiphenyltrichloroethane (DDT), endosulfan, melathion, carbaryl, pyrethrin, permethrin, cypermethrin, lambda-cyhalothrin and delta-methrin (Allsopp, 1984, Grant, 2001, Makumi *et al.*, 2000). There are four main strategies for applying insecticides: ground spay, pouring on the back or biting sites of animals, using insecticide-treated baited targets and the sequential

aerosol spraying technique (SAT) (World Health Organization, 2016b). Extensive insecticide ground spraying, which was used to control tsetse in Zimbabwe and Nigeriaj, is seldom used today because of concerns about residual insecticides and limitations on operational demands that the African governments can support (World Health Organization, 2016a). Specifically targeting cattle for insecticide treatment can be effective because it saves funds and minimizes the distribution of pesticide in the environment. Cost-effectiveness could be further achieved by restricting treatment to only the body regions where most tsetse flies feed on, for instance applying insecticide to the belly and legs (Torr *et al.*, 2007). However, the limitation is the potential for re-invasion if the area is not isolated and treated animals used as baits need space and cost for maintainance. This method is appropriate for settled areas, but the benefit is limited if there are tsetse in nearby unsettled areas that can invade. In this case, it is necessary to deploy targets for a long distance to trap tsetse flies from the invasion sources, but the costs can be very high if tsetse can invade from long distances.

Targets with insecticides in general have been used for increasing the benefit-cost ratio of vector control. While the simplicity of the traps or targets lends them to be used by local communities, sometimes these are applied on such a small scale that control efforts are bound to be frustrated by re-invasion. There are many factors that influence the effectiveness of traps: e.g. color, size and shape (Tirados *et al.*, 2011). For example, the relative propensities of tsetse to divert to variously coloured targets were found to decrease in the order: black or blue, red, and yellow. For different shapes, it decreased in the order: circle, square, and horizontal or vertical oblong (Torr, 1989). There are many kinds of traps that have been specifically modified to collect tsetse flies but the effectiveness of the design depends upon the species and of tsetse flies in different sites (Malele *et al.*, 2016). For example, biconical traps are effective for *G. palpalis* and *G. fuscipes* (Esterhuizen *et al.*, 2011); Ngu and Nzi for *G. morsitans a*nd *G. pallidipes* in easthern Africa (Belete *et al.*, 2004, Mihok, 2002); Epsilon for *G. morsitans a*nd *G. pallidipes* in southern Africa (Mhindurwa, 1994); H-traps for *G. brevipalis* (Kappmeier, 2000); and Ngu and Epsilon for *G. longipennis* in Kenya and Somalia (Kyorku *et al.*, 1990, Mihok, 2002). As described in section 1.5.6, attractant odours applied with traps are artificial

solutions or secretions from tsetse hosts. For the best-cost benefit ratio, 2 - 4 targets/km$^2$ has been suggested but this could vary by the type of trap, the species of flies or the overall density of flies in a region. SAT has been undertaken in many countries under the Pan African Tsetse and Trypanosomiasis Eradication Campaine (PATTEC) projects; for example, in Angola, Zambia, Ghana and Burkina Faso. However, this method failed to eliminate tsetse flies, and so did not alleviate problems with health and animal production, including fostering security of food (Adam *et al.*, 2013).

The Sterile Insect Technique (SIT) is another approach to reduce tsetse populations by exposing males to short bursts of gamma radiation from a cobalt-60 source to induce sterility and then releasing males into wild populations to compete with wild males. This technique is based on the assumption that females mate only once in their lifetime, thus any mating with a sterile male will prevent females from giving birth to any offspring. SIT requires rearing a large number of laboratory-bred male tsetse flies but has been found to be effective in some cases. Average emergence rate of sterile male *G. p. gambiensis* was more than 69%, with median survival of 6 days, and an average flight ability of more than 35% after pupae were exposed to chilling (4 – 6°C) (Seck *et al.*, 2015). The technique has been effectively used for eradication of *G. austeni* in the Unguja island in Zanzibar (Vreysen, 2001) but effective suppression using conventional methods is a pre-requisite before this environmentally safe species-specific technique can be used to eradicate residual populations. Ecological information is also needed or else re-invasion is bound to occur (Vreysen, 2001). However, the cost of SIT is exorbitant and the mass rearing of the flies is a major problem. The results of using this technique in areas with multiple species of tsetse flies also remain doubtful. Advances in molecular technologies and the availability of genomic information have been applied to develop new control strategies directly aimed at the fly or its ability to transmit parasites. For example, SIT has been combined with adding bacterial endosymbionts that carry salivary gland hypertrophy virus, which moderates the fecundity of the infected flies (Abd-Alla *et al.*, 2013). Another approach has been to use genetic modification to insert a recombinant gene into *S. glossinidius* (rec*Sodalis*) that is then introduced to tsetse flies, to block trypanosome development (Aksoy, 2003).

Since re-invasion of cleared areas from adjacent uncontrolled areas is a major problem due to the mobility of the flies and absence of natural barriers, surveillance and mornitoring of tsetse flies and other mechanical blood sucking vectors should be carefully and continuously managed. More effective strategies for eradication and control of tsetse flies can be developed from better knowledge of interaction among trypanosomes, vectors and vertebrate hosts.

## 1.7 Aims of this thesis

Variation of trypanosome prevalence in tsetse flies is known to be related to many factors, including: geographic region, species of both tsetse flies and trypanosomes, sex, age, and *Sodalis* status (Table 1.11). However, most previous studies have investigated these factors in isolation or considering only a few interactions, which could result in misleading interpretations if associations are in different directions for different combinations of factors (Figure 1.6). For example, it has been predicted that prevalence should increase with age (e.g. Woolhouse and Hargrove, 1998) but the strength of the association might differ between males and females because females are predicted to have higher overall prevalence (e.g. Nash, 1936) (Figure 1.6a). Associations with age also might differ by geographic location (Figure 1.6b), species of tsetse flies (Figure 1.6c), species of trypanosomes (Figure 1.6d), or combinations of factors (Figure 1.6e). There also could be differences in prevalence due to differences in the prevalence of endosymbionts or the host communities available in different regions. The aim of my thesis was to take an integrated approach to investigating the intrinsic factors (species, sex, age and *Sodalis* status) that affect prevalence of different species of trypanosomes in different species of tsetse flies sampled from different geographic regions, and to determine the source of their blood meals. In this investigation, I focused on four species of tsetse flies that occur in eastern Kenya (*G. austeni*, *G. brevipalpis*, *G. longipennis* and *G. pallidipes*) and conducted my study using samples collected from two sites in each of two geographic regions (Buffalo Ridge and Zungu Luka in the Shimba Hills region; and Mukinyo and Sampu in Nguruman), where tsetse flies and trypanosomes associated with AAT are known to be abundant.

**Table 1.11** Some factors affecting trypanosome prevalence

| Factors | Relationship to trypanosome infection in tsetse flies |
|---|---|
| Site | Variation in the distributions and communities of trypanosome species, tsetse flies, and potential hosts present (Ebikeme, 2011, Gould and Higgs, 2009). |
| Trypanosome species | Variation in the virulence of available trypanosome species and strains (Masumu *et al*., 2006) and their life cycles. |
| Tsetse species | Variation in trypanosome susceptibility of tsetse fly species and strains (Peacock *et al*., 2012a) |
| Age | Persistence of trypanosome infection in old flies (Franco *et al*., 2014).<br>Higher suscepeptibility of teneral flies than after feeding (Weiss *et al*., 2013).<br>Theoretical model predicted an increase in trypanosome prevalence with age (Woolhouse and Hargrove, 1998) |
| Sex | Females have longer longevity than males so could accumulate more trypanosomes (Nash, 1936).<br>Feeding volume and frequency could differ between the sexes because of differences in reproductive requirements and behaviour (Langley and Stafford, 1990, Gaston and Randolph, 1993) (Ejezie and Davey, 1976, Adlington *et al.*, 1996). |
| *Sodalis* status | Chitinase from *Sodalis* supports trypanosome penetration through the midgut membrane (Welburn *et al*., 1993, Welburn and Maudlin, 1999).<br>Presence of *Sodalis* and the other bacterial endosymbionts is related to immune status by enhancing of trypanosome refractoriness (Weiss *et al*., 2013). |

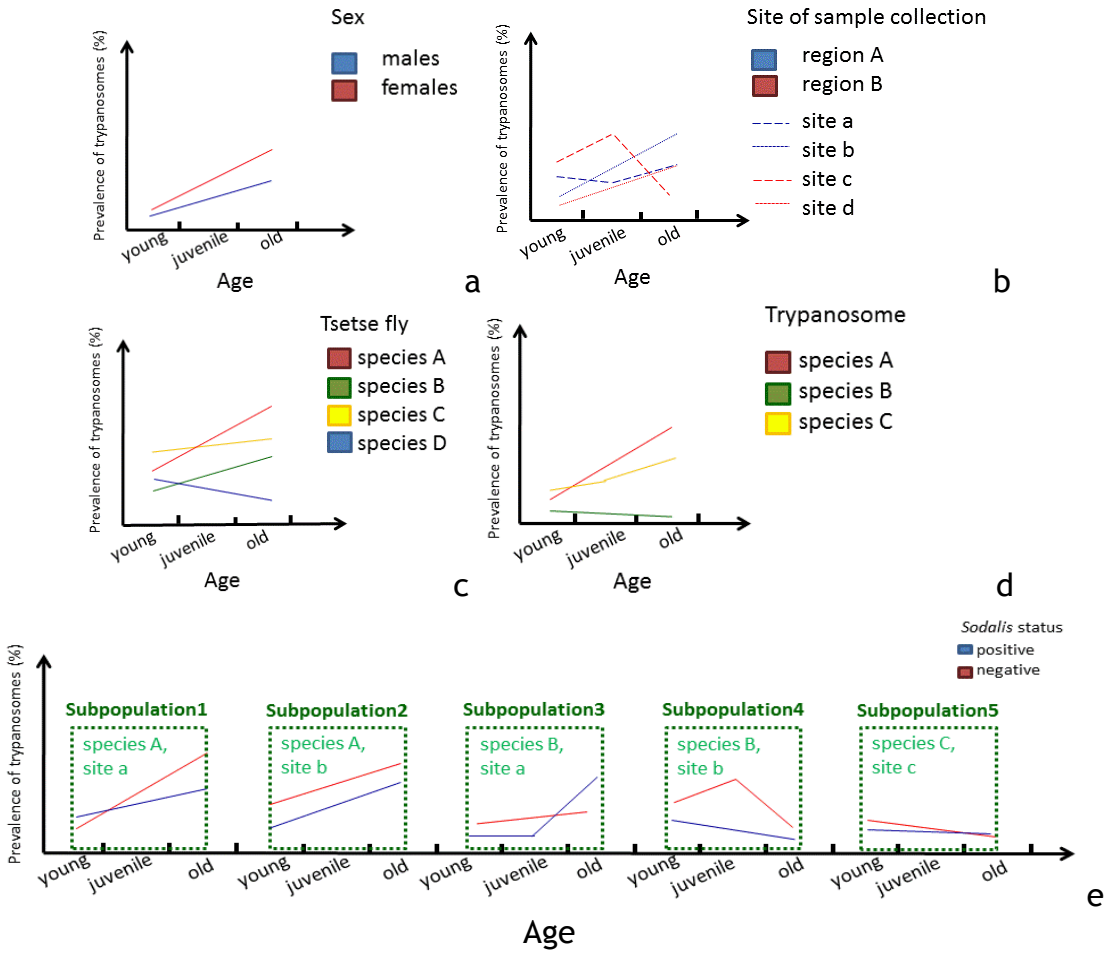



**Figure 1.6** Hypotheses of trypanosome prevalence considering interactions between multiple factors.
Variation in the potential association between trypanosome species with multivariable factors: (a) age and sex; (b) age and site; (c) age and species of tsetse flies; (d) age and species of trypanosomes; and (e) age, site, species of tsetse flies and *Sodalis* status.

There were four principle objectives to my thesis:

1. To optimise PCR methods for identification of trypanosomes, *S. glossinidius* and blood meal analysis in particular tsetse fly samples.
2. To determine which tsetse-specific factors (geographic location, species, age, sex, presence of *Sodalis*) or interactions between factors affect the relative prevalence of various species of trypanosomes in tsetse flies.
3. To determine whether the presence of *Sodalis* in different species or populations of tsetse flies affects their relative susceptibility to trypanosomes.
4. To characterise the feeding patterns of tsetse flies and determine whether this influenced prevalence of trypanosomes.

## 1.8 Chapter objectives

Many studies have used molecular techniques to amplify ITS genes for trypanosome screening in tsetse flies but effects of species, sex and age of tsetse flies from different sites and for different species of trypanosomes are still ambiguous. This is partly because different PCR techniques have been used in different studies and partly because few studies have simultaneously considered multiple factors that could affect trypanosome prevalence in the vectors. Therefore, in **chapter 2,** I investigated the distribution of tsetse fly species in the sampling region, their age and sex, and optimized methods for trypanosome identification in wild tsetse flies. I tested which tissues were most appropriate for identification of the trypanosome species present in the regions (head plus proboscis or abdomen) and tested various combinations of primers for trypanosome identification (ITS-1 universal primers and species specific primers). I then assessed the prevalence of the three main animal trypanosome species (*T. congolense, T. brucei* and *T. vivax*) in each species of tsetse fly (G. *austeni, G. brevipalpis, G. pallidipes* and *G. longipennis*) in relation to the sampling site, sex and age of the flies.

Variation in the prevalence of *S. glossinidius* (the secondary endosymbiont of tsetse flies) in different sites of sample collection and species of tsetse flies based on screening using different primers has been reported and it remains unclear whether or not the presence of the endosymbiont enhances trypanosome harboring in tsetse midguts. This is partly because many studies have focused on single sites rather than making comparisons between sites and none have taken into consideration how interactions between tsetse-specific factors might influence the relationship between the presence of *Sodalis* and relative susceptibility to trypanosomes. In **chapter 3**, I tested the reliability of *Sodalis* screening in different tissues of tsetse flies using different primers (pSG2, Hem and GPO1) and then tested the association of *Sodalis* status with multiple variables: subpopulation of the flies (a combination of species and sampling site); sex; age; and trypanosome status. I then tested whether the presence of *Sodali*s was related to the prevalence of trypanosomes.

Wild animals are reservoir hosts of trypanosomes, but they are difficult to investigate because it is difficult to sample them. Host feeding of tsetse flies

could indicate vertebrate hosts that are at risk of trypanosome infection and provide information about the distribution of trypanosomes. Thus, in **chapter 4** blood meal analysis was applied to determine sources of tsetse meals. This chapter focused on *G. pallidipes* because it was found at a sufficiently high frequency in three of the four sites sampled (Buffalo Ridge, Zungu Luka and Mukinyo) to enable assessment of the relationship between blood meal host and other factors, without confounding effects of the species of fly but enabling comparisons between populations. The following research questions were addressed in this chapter: 1) what animals were fed on by *G. pallidipes* at each of the study sites; 2) did tsetse flies tend to feed on single or multiple hosts; and 3) were host-feeding patterns related to trypanosome status? In addition, the most appropriate methods for determination of host-feeding patterns based on amplification and sequencing of mitochondrial genes from blood meals were assessed.

Lastly, in **chapter 5**, I discuss general issues associated with application of PCR methods in epidemiological studies of trypanosomiasis. I also discuss issues associated with data analysis when considering complex dynamics between hosts, parasites, and vectors. My results revealed complicated associations of trypanosome infection in tsetse flies with multiple tsetse intrinsic factors, geographic location, species and *Sodalis* status that made interpretation of Generalised Linear Models difficult. Finally, I make recommendations for how my results might inform the development of trypanosome control practices based on targeting of tsetse flies.

# Chapter 2 Tsetse fly distribution and trypanosome prevalence in tsetse flies

## 2.1 Abstract

Tsetse flies (*Glossina spp.*) are the known biological vectors for trypanosomes but there are multiple species of both the parasites and the vectors. The purpose of this study was to investigate what factors affect the relative susceptibility of the various species of tsetse to the trypanosome species present in eastern Kenya. In this survey, 1090 tsetse flies from the Shimba Hills (Buffalo Ridge and Zungu Luka) and Nguruman (Mukinyo and Sampu) regions were sampled using NG2G traps: *G. austeni, G. brevipalpis, G. longipennis* and *G. pallidipes*. Head plus proboscis parts of the tsetse flies were first screened using general primers targeting the first internal transcribed spacer region (ITS-1) of the ribosomal DNA array, designed to detect any species of trypanosome present and allowing species identification based on size of amplified products. Generalised Linear Models (GLMs) were conducted separately for each species of tsetse fly and trypanosome species to test for the effects of sex, age and site (for *G. pallidipes* and *G. longipennis*, which were the only species found at multiple sites) on trypanosome presence. The results indicated that presence of: 1) *T. congolense* in *G. brevipalpis* was significantly associated with an interaction between sex and age of the flies ($\chi^2$ = 4.12, df = 3, P-value = 0.0424); 2) *T. brucei* in *G. longipennis* was significantly related to sex ($\chi^2$ = 5.51, df = 1, P-value = 0.0190); and 3) *T. vivax* in *G. pallidipes* was associated with interactions between both site and sex ($\chi^2$ = 7.52, df = 21, P-value = 0.0061) and age and sex ($\chi^2$= 12.96, df = 6, P-value = 0.0015). Overall, these results emphasise the complexity of interactions between trypanosomes and their tsetse fly vectors and suggest that the method used for screening is important for trypanosome investigation. Appropriate screening tissues, primers and PCR conditions for trypanosome screening in tsetse flies could be dependent on geographic regions and biology of each trypanosome and tsetse fly species and so these should be newly optimized for each study region and set of species compared.

## 2.2 Introduction

Trypanosomes (*Trypanosoma spp.*) are vector-borne blood parasites that infect both animals and humans in poor rural areas. Human African Trypanosomiasis (HAT) or Sleeping sickness is a fatal infectious disease, which was first reported in 1896, and the incidence of the disease has been reported continuously since that time (Steverding, 2008, Simarro *et al.*, 2011, Franco *et al.*, 2014). The main pathogenic species of the parasites causing HAT are *T. b. rhodesiense* and *T. b. gambiense*. *Trypanosoma vivax*, *T. congolense* and *T. b. brucei* are the three major pathogenic species of Animal African Trypanosomiasis (AAT) or Nagana. These diseases affect public health and production of farm animals (Pathak, 2009, Swallow, 1999) and lead to economic loss (Swallow, 1999, Willson *et al.*, 1963, Smith *et al.*, 1998). Wildlife, such as bushbucks, waterbucks, lions, leopards, wildebeest, impalas, warthogs, hippopotamuses, pukus, and buffalo (Anderson *et al.*, 2011) are important reservoir hosts that carry a wide range of trypanosomes species (Auty *et al.*, 2012). Infected animals do not often show any clinical signs (World Organisation for Animal Health, 2013) and it is more difficult to collect blood samples from wild animals for screening of their trypanosome infection statuses than for livestock or people. Thus, the presence of trypanosomes has been screened mostly in patients, domestic animals and insect vectors (*Glossina spp.* or tsetse flies) to control and attempt to eradicate the diseases.

Microscope-based techniques are the classical method used for trypanosome identification. In tsetse flies, proboscis, salivary glands and midgut are dissected and require expertise for taxonomic assignment and identification of trypanosomes in those three organs, which tend to hold different groups of trypanosomes (Lloyd and Johnson, 1924, Enyaru *et al.*, 2010). *Dutonella* infection (mainly *T. vivax*) is diagnosed when infection of tsetse flies is only in the proboscis parts, *Nannomonas* infection (mainly *T. congolense*) when tsetse flies are trypanosome positive in both proboscis and midgut parts and *Trypanozoon* infection (mainly *T. brucei*), when all three regions of the gastrointestinal tract are trypanosome positive (Njiru *et al.*, 2004).

Subsequently, molecular identification methods based on the polymerase chain reaction (PCR) have been applied widely instead of dissection and microscope-

based techniques due to their higher sensitivity (Ouma *et al.*, 2000, Alibu *et al.*, 2015). In addition, the PCR method is relatively simple, quick and cheap for large sample sizes. Specific primer sets have been developed for identification of each *Trypanosome spp.:* TCN-1 and TCN-2 primers for *T. congolense* (Moser *et al.*, 1989), TCS1 and TCS2 for *T. congolense* savannah, TCK1 and TCK2 for *T. congolense* kilifi, TCF1 and TCF2 for *T. congolense* forest, TSM1 and TSM2 for *T. simiae,* TVW1 and TVW2 for *T. vivax* West Africa (Masiga *et al.*, 1992), and TBR-1 and TBR-2 for *T. brucei* (Moser *et al.*, 1989). In order to use those primers sets, many PCR reactions are necessary to assess trypanosome infections. Thus, Davila (unpublished) developed universal CF and BR primers for identification of mixed species of trypanosome presence in a single reaction by amplification of the first internal transcribed spacer region (ITS-1) of all pathogenic species of trypanosome (Njiru *et al.*, 2005). Screening results using DNA from well-characterised laboratory reference strains of trypanosomes indicated that these new general primers showed higher diagnostic sensitivity than Kin primers, another general set of ITS primers previously designed by McLaughlin *et al.* (1996). For blood samples of cattle from Kenya, trypanosome screening based on these primers was also more sensitive for *T. congolense*, *T. vivax* and *T. brucei* identification than using their relevant species-specific primers. However, this study did not sequence amplification products from wild-caught flies or hosts to determine whether the band sizes assigned to represent the various species of trypanosomes based on laboratory strains were also accurate for predicting the species of wild trypanosomes.

There also could be potential biases in amplification of different species of trypanosomes based on their biology (e.g. which tissues each species resides in within tsetse flies). For classical microscopic studies, midguts (Alibu *et al.*, 2015, Simo *et al.*, 2011, Farikou *et al.*, 2010a), salivary glands or mouth parts (Malele *et al.*, 2003, Masiga *et al.*, 1992, Lehane *et al.*, 2000) of tsetse flies are typically dissected for trypanosome screening, each of which are small and could be easily contaminated with other trypanosome positive samples. Dissection of tsetse flies into 3 parts (head plus proboscis (HP); thorax (TH); and abdomen (AB), which includes both the midgut and the salivary glands) is more practical and poses a lower risk for contamination than the dissection of the internal organs. Using the abdomen for *T. congolense* and *T. brucei* screening could be

more sensitive than using head plus proboscis parts but could misdiagnose *T. vivax*, which establishes and develops mostly in the mouthparts. Moreover, substances contained in blood meals in the abdomen parts could decrease efficiency of a PCR reaction (Akane *et al.*, 1994, Al-Soud *et al.*, 2000, Al-Soud and Radstrom, 2001, Opel *et al.*, 2009, Alaeddini, 2012), and so there is a risk of amplifying false positives from trypanosomes in blood meals (i.e. trypanosomes that are present but do no establish in the tsetse flies). The mouth parts are at the upper end of the gastrointestinal tract, so fast clearance of trypanosomes that will not establish should result in fewer false positives. Although it is not a main site for *T. brucei* and *T. congolense* development, the infective stage of all trypanosomes have to pass through the mouth parts to vertebrate hosts while tsetse flies are feeding. Thus, this body part could be suitable for screening of which *Trypanosoma spp.* are present in a particular geographic region but might not be as informative to establish actual prevalence of blood parasite infections in tsetse flies because false negatives might be expected due to low density of parasites.

Previously, prevalence of trypanosomes has been reported to vary by species of tsetse flies and geographic regions. In Zambia, *G. m. morsitan* (83.9%) had the highest trypanosomes prevalence base on the ITS-1 CF and BR primers when compared with *G. pallidipes* (29.1%) and *G. brevipalpis* (56.4%) (Dennis *et al.*, 2014). In Tanzania, prevalence of *T. congolense* savannah in *G. swynnertoni* and *G. m. morsitans* from the Sangaiwe region was 7.0% lower than *G. brevipalpis* and *G. pallidipes* from the Pangani region, while prevalence of *T. godfreyi* in the Sangaiwe region was approximately two times higher than in the Pangani region (Malele *et al.*, 2003). In addition, mixed infections of *Trypanosoma spp.* (double and triple infections) were reported with lower percentage than single infections in *G. pallidipes*. Which species of trypanosomes is most prevalent also has been found to vary by geographic region. In Tanzania, *T. simiae* Tsavo had the highest prevalence in the Sangaiwe and Pangani regions (Malele *et al.*, 2003) while *T. vivax* infection was highest in Simanjiro (Salekwa *et al.*, 2014). In Zambia, *T. brucei* was the predominant pathogenic trypanosome species (Dennis *et al.*, 2014) while *T. congolense* was more prevalent in Kenya (Woolhouse and Hargrove, 1998). Furthermore, prevalence of trypanosome infections in tsetse flies has been found to vary with sex of tsetse flies (Peacock *et al.*, 2012b), with

one study showing 2.5 times higher infection in female *G. pallidipes* than males. Prevalence of trypanosomes has also been related to age of tsetse flies (Woolhouse and Hargrove, 1998); for example, infection with *Trypanosoma spp*. in *G. pallidipes* from Zimbabwe was found to increase with age (Woolhouse *et al*., 1994, Woolhouse *et al*., 1993). However, this may differ by species of trypanosome: in *T. congolense* infection decreased with age and no relation was found for *T. vivax* (Woolhouse *et al*., 1993). Together, these studies suggest that there is extensive variation in tsetse-related factors that affect relative susceptibility to trypanosomes. However, most studies have investigated these factors in isolation. To achieve a better understanding of disease distribution and associations with their biological hosts, what is needed is an investigation of potential interactions between multiple intrinsic factors of tsetse flies in different sites that could influence trypanosome infection by different pathogenic species.

In the southeast of Kenya, there is an abundance of different types of habitats suitable for tsetse flies (Figure 2.1). Individual environments of each geographic region might affect behaviour and development of each community of tsetse flies. Moreover, they could feed on a variety of available trypanosome-susceptible hosts. Studying trypanosome abundance in regions that differ in habitats and hosts thus could make a good model for investigating the factors that influence trypanosome prevalence in tsetse flies and associations among those factors.

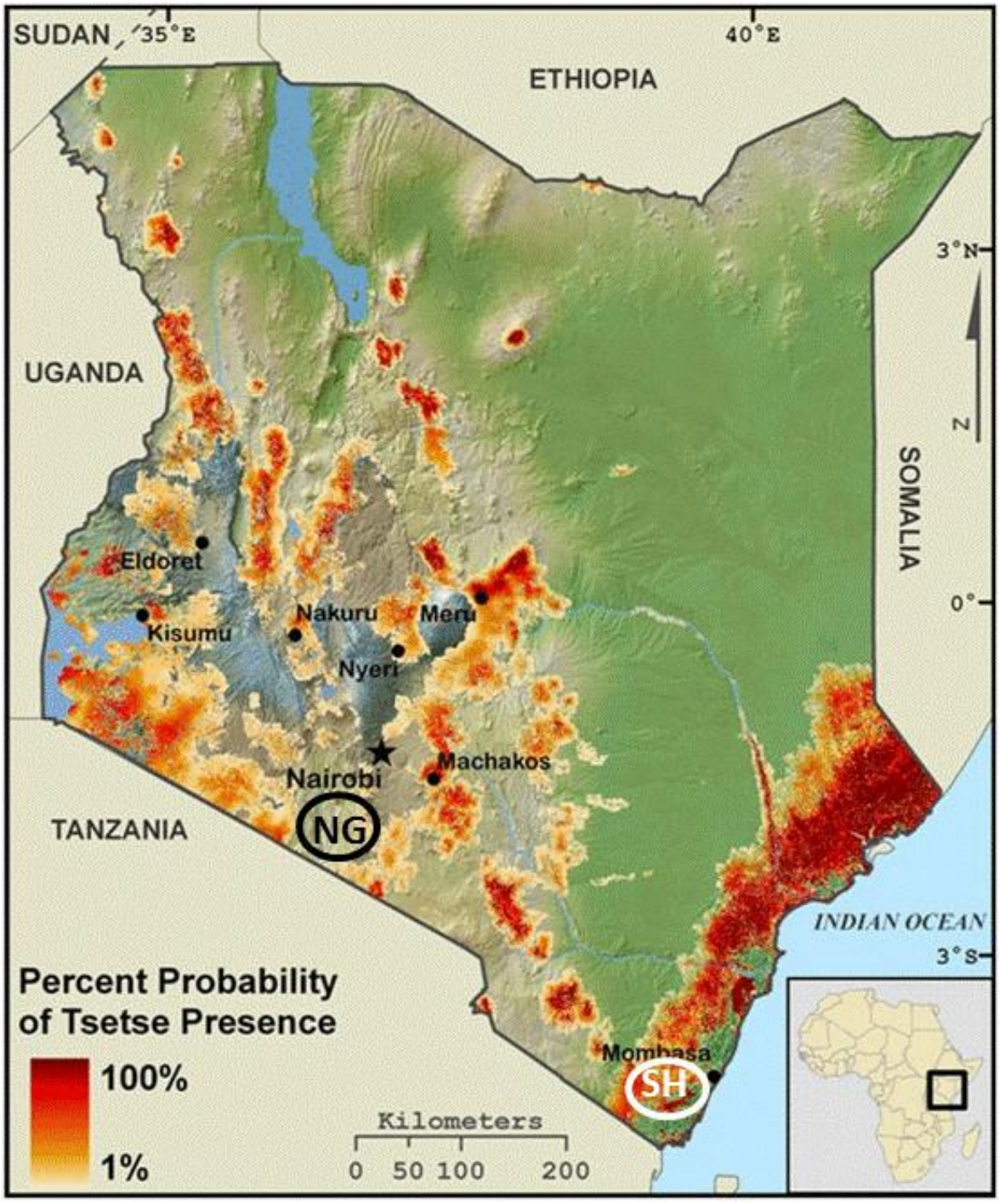



**Figure 2.1** Map of tsetse fly distribution in Kenya.

Probability of tsetse flies, climate and land cover data are represented across the country. NG: Nguruman; SH: Shimba Hills. (Joseph Messina, 2011. http://publications.nigms.nih.gov/insidelifescience/sleepsick_mystery.html. accessed date: 17 Sep 2015.)

There were three main aims of this chapter. The first was to characterize the age, sex and distribution of the species of tsetse flies found in the Shimba Hills and Nguruman regions of Kenya. The second purpose was to optimise PCR conditions for trypanosome identification in tsetse flies and choose which tissues produce the most reliable and repeatable results. The third purpose was to determine prevalence and associations of *Trypanosoma spp*. presence (*T. congolense* savannah and kilifi types, *T. vivax*, *T. brucei, T. simiae and T. godfreyi*) in relation to geographic distribution and intrinsic factors of tsetse flies. Results based on consideration of each of these factors separately were compared with generalised linear models approach to investigate whether interactions among factors might lead to different conclusions about what drives relative susceptibility to trypanosomes in tsetse flies. The main hypotheses of this study were set as following. Firstly, that there is variation among sites in tsetse fly abundance by species, sex and age, which might depend on factors such as the lifespan or reproductive outputs of each species under different ecological conditions. Secondly, that the particular PCR conditions used for trypanosome screening could affect interpretation of results due to genetic variation in strains of trypanosome present at each site (e.g. Auty *et al*. 2012) or differences in infection rates among different tissues. Thirdly, that interactions among intrinsic factors of different species of tsetse flies from different sites might significantly influence the prevalence of different trypanosome species in different directions. For example, since females have a longer lifespan than males and tend to feed more frequently, it has been suggested thay they will accumulate more trypanosomes than males (Nash, 1936) so sites with older females might be predicted to show the highest prevalence. However, this relationship could be different for different species of tsetse flies, different species of trypanosomes or different geographic regions.

## 2.3 Materials and methods

### 2.3.1 Sampling sites

Tsetse flies were collected from two sites from each of two main regions (the Shimba Hills and Nguruman). The Shimba Hills region (Figure 2.2a) is within the National Reserve Region, a large green forest full of many species of wild animals. The sites sampled in that region, Buffalo Ridge and Zungu Luka (approximately 20 km apart), have different types of vegetation and levels of human activity. Buffalo Ridge is surrounding with a thicket forest, where many tourists visit all year, while Zungu Luka has a woodland type of vegetation, close to an inhabited rural area. In contrast, the Nguruman region (Figure 2.2b) contains lowland woodland patches surrounded by open savannah, habitats which have been found to host a large number of *G. pallidipes* and *G. longipennis* (Brightwell *et al.*, 1997). The two sampled sites, Mukinyo and Sampu, are around 6 km apart and are generally similar in habitat to each other.

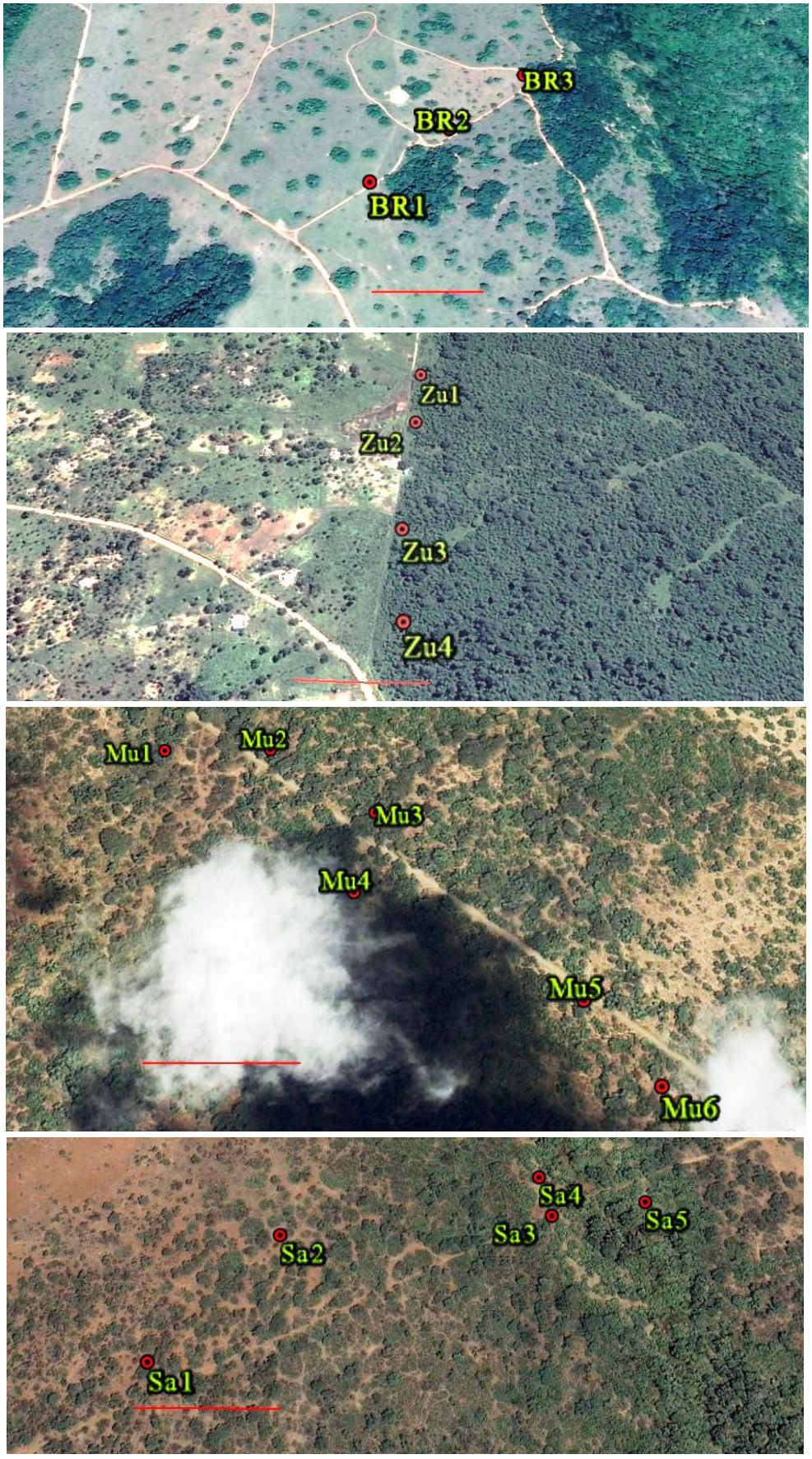



**Figure 2.2** The geographic regions of tsetse fly collection in Kenya with traps indicated.
Aerial photographs of the trap locations in: (a) Buffalo Ridge; (b) Zungu Luka; (c) Mukinyo; and (d) Sampu. Red lines are scale bars for 200 m.

### 2.3.2 Sampling method

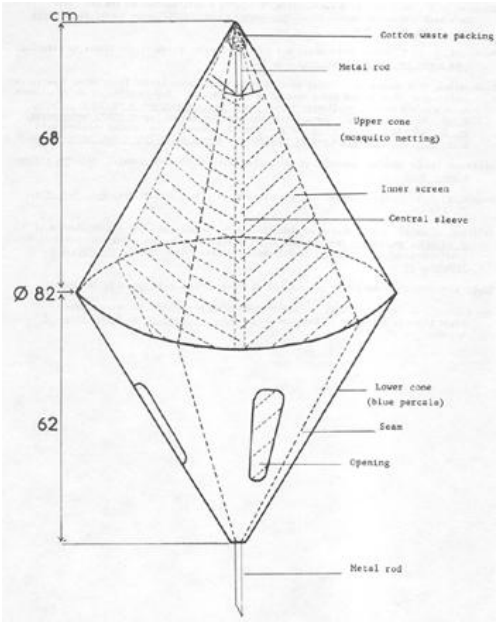
Biconical traps were widely used for tsetse control and sampling during the 1970s - 1980s. The original trap was composed of two cones joined at the base (Figure 2.3). The upper cone consisted of four pieces of nylon net and the lower cone was made from four panels of blue cloths, with an opening through the trap (Murray *et al.*, 1983a, Murray *et al.*, 1983b). Several traps have been modified from the primitive traps for greater efficiency with cheaper designs; for example, monoconical "Vavoua", pyramidal, monoconical Mérot, F3 Box, NG2B and NG2G (Leak, 1998). In this study, Marc Ciosi and his colleagues randomly collected tsetse flies from their habitats using NG2G traps. These are modified NG2B tsetse traps to which one metre of blue cloth is added on the right side (Figure 2.4) to control *G.pallidipes* and *G.longipennis* (Brightwell *et al.*, 1991). Cages were made of light transmitting netting, which was designed by raising the trap about 15 – 20 cm and were baited with acetone and cow urine. Acetone is assumed as an "odour in host breath" that is highly effective for tsetse attraction (Vale *et al.*, 1988). The urine was collected from local cows, which were older than three weeks, and prepared for a standard odour at ambient temperature (approximately 25 - 30°C). The urine was dispensed from a closed plastic container with a 2 x 4 cm slot cut at the side below the rim to release phenols at a rate of 1000 mg/h. At the same time, acetone was dispensed at 150 mg/h from a 200 ml bottle with a 0.2 cm diameter hole in the lid. Both dispensers were placed 30 cm behind the traps. At the top of the trap, a net cage for fly collection was opened through a 95% ethanol plastic bottle, which was attached at a 45° angle (Figure 2.4b).

NG2G traps were placed in the shadows of trees, which is a resting site for tsetse flies, to decrease the attractant evaporation rate. The distance between traps at each site was approximately 200 - 500 m (Figure 2.2) Sites were sampled in 2012, between 15 June and 15 July for Buffalo Ridge and Zungu Luka 10 - 14 August for Mukinyo and Sampu. This time of year was selected for sampling beause Kenya is hot and dry during January to March and then relatively cool from June to August (Mbahin *et al.*, 2013), and is humid from June to October (Moggridge, 1949). The long rain is normally between the middle of March and June while the short rain occurs between October and December (Mbahin *et al.*, 2013). A previous longitudinal survey also found a peak in prevalence of

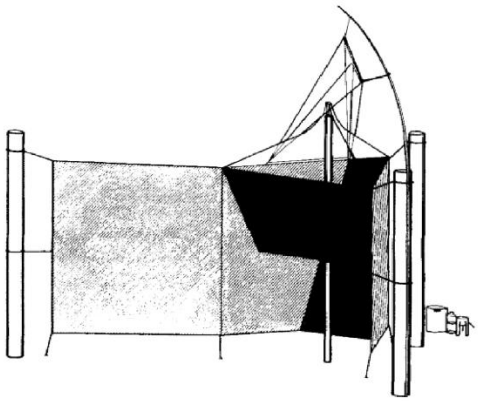
trypanosomes in June - August (Woolhouse *et al.*, 1993, Woolhouse *et al.*, 1994). In addition a higher prevalence of trypanosomes in cattle has been found in the wet season than in the dry season (Majekodunmi *et al.*, 2013). Trapped flies were collected every day at 17.00 pm, which has been found to be the highest interaction period between tsetse flies and hosts (Okoth *et al.*, 2007). Whole flies were preserved in 95% ethanol and stored at -80°C and provided to me for data collection.

**Table 2.1** Locations of traps for sample collection, indicating latitude and longitude of individual traps.

| Sites | Names of traps | Latitude | Longitude |
|---|---|---|---|
| **Buffalo Ridge** | BR1 | 4°14'33.03" S | 39°26'12.73" E |
| | BR2 | 4°14'28.61" S | 39°26'18.04" E |
| | BR3 | 4°14'23.78" S | 39°26'23.55" E |
| **Zungu Luka** | Zu1 | 4°20'0.54" S | 39°15'53.51" E |
| | Zu2 | 4°20'4.44" S | 39°15'53.10" E |
| | Zu3 | 4°20'11.95" S | 39°15'52.19" E |
| | Zu4 | 4°20'17.60" S | 39°15'52.15" E |
| **Mukinyo** | Mu1 | 1°50'2.77" S | 36° 4'54.21" E |
| | Mu2 | 1°50'2.78" S | 36° 4'59.35" E |
| | Mu3 | 1°50'6.45" S | 36° 5'4.46" E |
| | Mu4 | 1°50'10.59" S | 36° 5'3.49" E |
| | Mu5 | 1°50'15.91" S | 36° 5'13.29" E |
| | Mu6 | 1°50'19.62" S | 36° 5'16.06" E |
| **Sampu** | Sa1 | 1°53'29.95" S | 36° 4'7.52" E |
| | Sa2 | 1°53'23.87" S | 36° 4'13.20" E |
| | Sa3 | 1°53'23.11" S | 36° 4'26.25" E |
| | Sa4 | 1°53'21.05" S | 36° 4'25.76" E |
| | Sa5 | 1°53'22.48" S | 36° 4'30.85" E |




a



Brightwell, R., R.D. Dransfield, and C. Kyorku, *Development of a low-cost tsetse trap and odour baits for Glossina pallidipes and G. longipennis in Kenya*. Medical and Veterinary Entomology, 1991. **5**(153-164).

b

**Figure 2.3** Models of a biconical trap and a NG2G trap.

(a) An original biconical trap, which was designed by Challier and Laveissiere (Kaba *et al*., 2014); (b) A NG2G trap, which was modified from a biconical trap by Brightwell *et al*. (1991).

a

b

**Figure 2.4** Modified NG2G trap baited with acetone and cow urine for tsetse fly collection (created by Ciosi, 2012).

(a) A NG2G trap, which was modified from biconical traps by adding a meter of blue cloth on the right side for tsetse fly collection in this current study; (b) A net cage with a 95% ethanol plastic bottle at the top of the NG2G trap.

## 2.3.3 Intrinsic data collection

### 2.3.3.1 Tsetse fly characterisation

### 2.3.3.2 Species of tsetse flies determination

*Glossina austeni*, *G. brevipalpis*, *G. longipennis* and *G. pallidipes* were identified by their distribution and specific morphology (Pollock, 1982). The four species of tsetse flies can be distinguished based on differences in body size and colour of their abdomen parts (Figure 2.5). *G. brevipalpis* (10.2 - 13.0 mm) has an oval-shaped abdomen with a dark brown color while *G. longipennis* (9.5 - 11.0 mm) has a longer oval shaped yellow abdomen with short dark marks at both lateral sides. *G. austeni* (7.5 - 8.5 mm), the smallest species, has a column-shaped abdomen with a dark mark across the posterior part. For *G. pallidipes* (8.5 - 11.0 mm), the color marks on the abdomen part show dark lines across the whole abdomen, except the last section that attaches to the thorax.

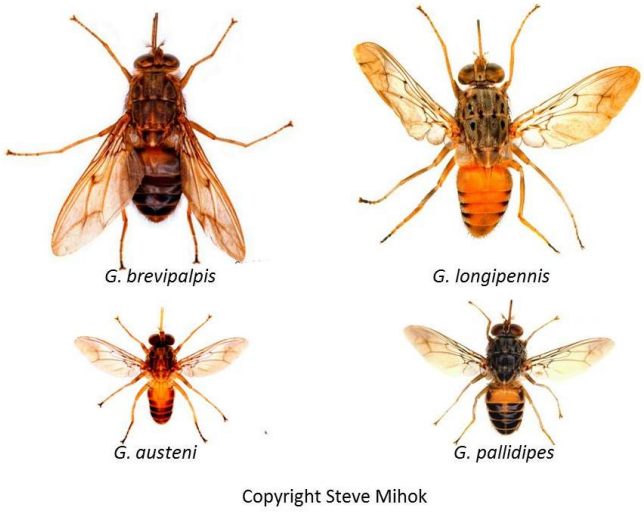



**Figure 2.5** Morphology of the four tsetse species used in this study: *G. brevipalpis*; *G. longipennis*; *G. austeni*; and *G. pallidipes* (photos are used by permission from Steve Mihok).

### 2.3.3.3 Sex determination

The sex of all tsetse samples was physically determined by an examination of the external sex organs on the ventral part of the abdomen under a stereoscope (10x magnification) (Figure 2.6). In males, a large rounded hypopygium is detectable clearly. In front of the hypopygium, there is a plate with a number of

hector plates. For females, external sex organs are restricted to only a small genital plate.

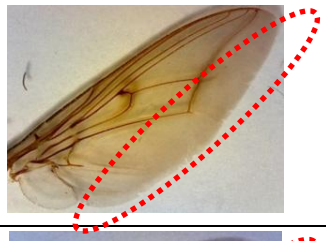
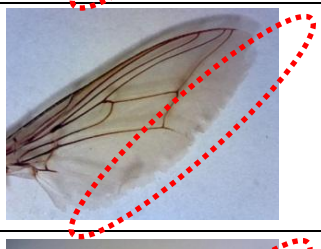
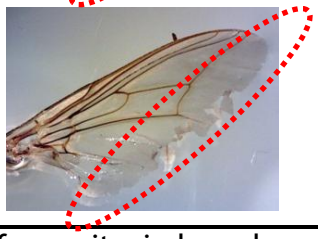

Male Female

**Figure 2.6** Sex identification of *G. spp.* using external genitalia of the ventral abdomen.

Male flies have a hypopygium (a) and hector plate (b), while female flies have a small genital plate (c) (Leak, 1998).

#### 2.3.3.4 Age determination

Age of tsetse flies was determined based on the wing fray score described by Jackson (1946) because it allows aging of both males and females, it can be conducted on preserved samples and it has been demonstrated to be correlated with age determined by other methods, such as ovary scores (Wamwiri *et al.*, 2013, World Health Organization, 2016a, Woolhouse *et al.*, 1993). This classifies the amount of damage at the rim of each wing into six levels according to the degree of wing degradation (Figure 2.7). Wings that are perfect, with no detectable damage, are classified as scale 1. For scale 2, some damage is present. For scale 3, some damage is noticed but confined to the proximal part of the wing. Individuals with definite damage on both proximal and distal parts of the wing are classified as scale 4. For scale 5, there is a saw-edged appearance to the wings. For scale 6, the wing is extensively damaged or has severe ruptures. The scores for the two wings of each individual were averaged and the age of tsetse flies was then classified into three levels: “young” for wing fray score = 1.0 - 2.5, “juvenile” for wing fray = 3.0 - 4.0 and “old” for wing fray = 4.5 - 6.0. Predictions of absolute ages based on these scores has been predicted only for some tsetse species (Jackson, 1946), so here it will be used as a relative measure to enable comparisons between species at different sites.

| Score | Estimated age[a] | Picture | Description |
|---|---|---|---|
| 1 | 1–8 days |  | Perfect |
| 2 | 8–19 days |  | Slight damage |
| 3 | 20–30 days |  | Definite damage |
| 4 | 30–40 days |  | Definite damage both proximal and distal parts |
| 5 | 40–50 days |  | Saw-edged appearance |
| 6 | 50–63 days |  | Rounded indentations |

[a] estimated ages and wing fray criteria based on *G. mositans* taken from (Jackson, 1946, Wamwiri *et al.*, 2013)

**Figure 2.7** The six levels of wing fray score for age estimation (described by Jackson, 1946).

Both wings of each fly were scored: 1 for a perfect wing; 2 for little damage at the anterior part of a wing notch; 3 for damage with a higher degree along the anterior part than the previous score; 4 for little damage before and beyond the notch; 5 for saw-edged damage; and 6 for severe damage along the wing rim that results in rounded indentations.

## 2.3.4 *Trypanosoma spp*. presence in tsetse flies

### 2.3.4.1 DNA extractions

Trapped tsetse flies were dissected into three parts: head and proboscis (HP); thorax (TH); and abdomen (AB). To minimize cross contamination between samples, scalpel blades and petri dishes used for dissections were changed for each individual. Forceps were also cleaned with 10% bleach, water and double distilled water. Head plus proboscis parts were frozen using Liquid Nitrogen and crushed with a micropestle (Starlab.Co.UK). Abdomen parts were cut into small pieces with scalpel blades at room temperature. DNA extractions were conducted separately for head plus proboscis and abdomen parts using DNeasy® Tissue & blood kits, following the manufacturer's instruction (Qiagen Inc, Paisley, UK). Both tissue parts were digested with proteinase K at 56ºC shaking incubator until solutions were clear. The elution volume of extracted DNA was 50 µl from the head plus proboscis parts and 100 µl from the abdomen parts. The concentrations of all DNA samples were then measured by using a Nanodrop 1000™ Spectrophotometer (Thermo Scientific) at 260 nm. All extracted DNA was preserved at -20°C.

### 2.3.4.2 PCR-based screening of *Trypanosoma spp*. presence

In order to determine the trypanosome status of tsetse flies, screening was performed using the general primers developed by Davila (unpublished) (Njiru *et al*., 2005) that amplify the internal transcribed spacer 1 region of the ribosomal DNA of all known trypanosome species (CF and BR primers). The PCR products of each species have been described as having their own specific sizes, as follows: 697 bp for *T. congolense* savannah (Tcs); 600 bp for *T. congolense* kilifi (Tck); 476 bp for *T. brucei* (Tb); 397 bp for *T. simiae* (Ts); 273 for *T. godfreyi* (Tg); and 250 bp for *T. vivax* (Tv). Thus, these primers have conventionally been used for trypanosome species identification (Marc *et al*., 2011, Njiru *et al*., 2005).

**Table 2.2** Primers sets of *Glossina spp*. and *Trypanosoma spp*. used for screening.
ITS-1 primers were used to identify both *Glossina spp*. and *Trypanosoma spp*., based on size of PCR products. *T. congolense* (savannah, forest and kilifi types) and *T. brucei* were also determined using TCS, TCF, TCK and TBR primers, respectively.

| Primers | Sequences | Length (bp) | Species identification | Reference |
|---|---|---|---|---|
| Glossina ITS-1 | ITS-1_for: GTG-ATC-CAC-CGC-TTA-GAG-TGA | varied[1] | *G. brevipalpis*, *G. pallidipes*, *G. austeni* and *G. longipennis* | Dyer *et al*. (2008) |
| | ITS-1_rev: GCA-AAA-GTT-GAC-CGA-ACT-TGA | | | |
| Trypanosome ITS-1 | ITS-1_CF: CCG-GAA-GTT CAC-CGA-TAT-TG | varied [2] | *T. congolense*, *T. brucei*, *T. vivax*, *T. simiae* and *T. godfreyi* | Njiru *et al*. (2005) |
| | ITS-1_BR: TTG-CTG-CGT-TCT-TCA-ACG-AA | | | |
| TCS | TCS1: CGA-GAA-CGG-GCA-CTT-TGC-GA | 319 | *T. congolense* savannah type | Masiga *et al*. (1992) |
| | TCS2: GGA-CAA-ACA-AAT-CCC-GCA-CA | | | |
| TCF | TCF1: GGA-CAC-GCC-AGA-AGG-TAC-TT | 350 | *T. congolense* forest type | Masiga *et al*.(1992) |
| | TCF2: GTT-CTC-GCA-CCA-AAT-CCA-AC | | | |
| TCK | TCK1: GTG-CCC-AAA-TTT-GAA-GTG-AT | 294 | *T. congolense* kilifi type | Masiga *et al*. (1992) |
| | TCK2: ATC-CAA-AAT-CGT-GCA-CCT-CG | | | |
| TBR | TBR-1: GAA-TAT-TAA-ACA-ATG-CGC-AG | 173 | *T. brucei* | Moser *et al*. (1989) |
| | TBR-2: CCA-TTT-ATT-AGC-TTT-GTT-GC | | | |

[1] ITS-1 PCR product sizes were different in each species of tsetse flies: *G. brevipalpis* is 778 bp; *G. pallidipes* is 919 bp; *G. austeni* is 633 bp; and *G. longipennis* is approximately 1,250 bp (in my study).

[2] ITS-1 PCR product sizes were different in each species and type of trypanosomes: *T. congolense* forest is 714 bp; *T. congolense* savannah is 697 bp; *T. congolense* kilifi is 600 bp; *T. brucei* is 476; *T. simiae* is 397 bp; *T. godfreyi* is 273 bp and *T. vivax* is 250 bp.

#### 2.3.4.3 Optimisation of PCR screening method for *Trypanosoma spp.* in tsetse flies

In order to increase sensitivity of PCR screening methods and assess the most reliable tissue type for identifying tsetse flies, amplification of ITS-1 from DNA extracted from the head plus proboscis was compared with that from the abdomen, using 50 samples of *G. pallidipes* from Buffalo Ridge (the Shimba Hills). PCR was carried out in 10 µl reaction mixtures containing 1 µ of 10X Custom PCR Master Mix – No *Taq* (45 mM Tris-HCl, pH 8.8 at 25°C, 11 mM $(NH_4)_2SO_4$, 4.5 mM $MgCl_2$, 0.113 mg/ml BSA and 4.4 mM EDTA and 1.0 mM each of dATP, dCTP, dGTP and dTTP (Thermo Scientific), 10 µM ITS-1 primers (Njiru *et al.*, 2005), DNA template (20-200 ng) and 1 unit of *Taq* DNA polymerase (Thermo Scientific). Primers were tested with two PCR conditions for ITS-1 amplification: the conditions of Njiru *et al.* (2005), which were designed for trypanosome identification in blood samples of cattle from Kenya (Njiru *et al.*, 2005) and those of Isaac *et al.* (2013), which were optimised for trypanosome screening in tsetse flies from Nigeria (Isaac *et al.*, 2013) (Table 2.3). Screening results from the two parts of tsetse flies were then compared, using the two sets of amplification conditions.

**Table 2.3** Comparison of PCR conditions for trypanosome screening with ITS-1 primers.

| PCR steps | Isaac *et al.*, 2013 | | Njiru *et al.*, 2005 | |
|---|---|---|---|---|
| | Temperature(°C) | Time (sec) | Temperature(°C) | Time (sec) |
| Preheating | 95 | 2 min | 94 | 5 min |
| Denaturation | 95 | 50 | 94 | 40 |
| Extension | 60 | 50 | 58 | 40 |
| Annealing | 72 | 60 | 72 | 90 |
| Cycle (rounds) | 30 | | 35 | |
| The last extension | 72 | 5 min | 72 | 5 min |

A comparison was also made between different types of PCR reagents: *Taq* DNA Polymerase from Invitrogen (catalogue number 18038-034), Dream*Taq* Green

PCR Master Mix from Thermo Scientific (catalogue number K1082) and *Taq* DNA Polymerase from Thermo Scientific (using the buffers that come with the enzyme) were compared with respect to amplification of ITS-1 fragments of positive controls (*T. congolense* savannah, *T. brucei* and *T. vivax*). They all used the same PCR conditions that were found most appropriate for ITS-1 identification from the previous experiment.

PCR amplicons were visualised by gel electrophoresis in 1.5% UltraPure™ Agarose gels (Invitrogen Inc, Carlsbad, CA., U.S.A.) with 2% Ethidium Bromide (Invitrogen Inc, Carlsbad, CA., U.S.A.) in 0.5X TBE buffer, containing 45 mM Tris Base (Fisher Chemical), 44.5 mM Boric acid (Fisher Chemical) and 0.05 mM EDTA (SIGMA-ALDRICH®), pH 8.0. The conditions of gel electrophoresis were 100 V and 300 A for 80 min. The results were then visualised using a gel documentation system (UVIpro Plainum, UVITEC, Cambridge, UK).

The PCR conditions that were found to produce the most consistent amplifications were used for trypanosome screening in all sampled populations of tsetse flies. In order to confirm that size-based species identification would be reliable, PCR products that, based on their sizes, were suspected to be *T. congolense* kilifi (approximately 600 bp), *T. congolense* savannah (approximately 700 bp), *T. brucei* (approximately 500 bp) and *T. vivax* (approximately 250 bp), were purified with QIAquick Gel Extraction Kits (Qiagen Inc, Paisley, UK) and cloned using TOPO®-TA Cloning Kits (Invitrogen Inc, Carlsbad, CA., U.S.A.). DNA from six plasmids of each clone were extracted with QIAprep Spin Miniprep Kits (Qiagen Inc, Paisley, UK) and sent to the DNA Sequencing and Services, University of Dundee for sequencing. Chromatographs were manually corrected and compared using the Sequencher software program, version 5.3 (Gene Codes Corporation, Ann Arbor, MI USA). The identity of the sequences was then determined by using the Basic Local Alignment Search Tool (BLAST) to confirm match to available trypanosome species in GenBank.

#### 2.3.4.4 Identification of *Trypanosoma spp*. based on specific primers

In order to confirm results from the ITS-1 primer screening, samples were also screened using species-specific primers. All DNA samples from head plus proboscis were screened with specific primers for *T. congolense* savannah, using

primers developed by Masiga *et al*. (1992) that amplify a highly conserved 319 bp fragment of nuclear DNA (TCS1 and TCS1 primers) (Masiga *et al*., 1992) (Table 2.2). Since *T. congolense* forest amplifies a similar sized product (710 bp) as *T. congolense* savannah using the general ITS-1 primers, subspecies-specific primers for forest (TCF1 and TCF2 primers developed by Masiga *et al*. (1992) (Table 2.2), were used to test randomly selected 700 bp ITS-1 positive products to determine if they any were from the forest type (N = 27). In addition, the subspecies-specific primers TCK1 and TCK2 (Masiga *et al*., 1992) were used to test randomly selected samples (N = 47) to determine if these primers were more sensitive than identification based on the 600 bp band amplified using the ITS-1 primer (Table 2.2). These primers amplify 350 and 294 bp fragments, respectively, at the location of nuclear satellite repeat monomers. The PCR mixture for the *T. congolense* primers contained 10 µl, which was composed of 5 µl Dream *Taq* Green PCR master Mix (2X) (Thermo Scientific), 10 µM primers and DNA template (20 - 200 ng). PCR conditions were as follows: preheating at 95°C for 3 min, followed by 30 cycles of denaturation at 95°C for 30 sec, annealing at 58°C for 30 sec, and extension at 72°C for 30 sec and a final elongation step at 72°C for 5 min.

For *T. brucei*, all samples (N = 75) that presented an approximately 500 bp fragment using the general ITS-1 primers were screened with a specific primer pair (TBR1 and TBR2 primers) (Table 2.2), which were developed by Moser *et al*. (1989). The primers anneal to a region of highly conserved nuclear DNA and amplify a variable region of repeats, with a 173 bp PCR product. PCR conditions for the TBR primers were as described for TCS but with an annealing temperature of 55°C (Masiga *et al*., 1992) and used Dream *Taq* Green PCR master Mix (2X) (Thermo Scientific).

In order to distinguish samples that showed lack of amplification due to absence of trypanosomes from those due to poor DNA quality, all trypanosome negative samples were screened with ITS primers designed for *Glossina spp*. (GlossinaITS-1_for and GlossinaITS-1_rev) (Table 2.2) (Dyer *et al*., 2008) that should amplify in all samples. Samples without amplification using these primers were excluded from assessment of variation in subsequent analyses. PCR reaction mixtures were as described for trypanosome amplification; the PCR conditions of Dyer *et al*. (2008) were used for amplification. Representative PCR products from each

tsetse species (seven samples: two *G. pallidipes* from Buffalo Ridge; three *G. pallidipes* from Mukinyo; one *G. longipennis* from Mukinyo; and one *G. longipennis* from Sampu) were cleaned with QIAquick PCR Purification Kits (Maryland, U.S.A.) and sent for sequencing at the DNA Sequencing and Services, University of Dundee. Chromatographs were manually corrected using the Sequencher software programme and BLAST was used to confirm match to available *Glossina* species in GenBank.

### 2.3.5 Associations of *Trypanosoma spp*. with intrinsic factors in tsetse flies

To visualise differences in trypanosome prevalence in relation to the various factors considered, 95% confidence interval (95% CI) were computed and added to summarising histograms.

$$95\%\ \mathrm{CI} = \mathrm{p} \pm 1.96\sqrt{\frac{p(1-p)}{N}}$$

| | | |
|---|---|---|
| 95% CI | = | 95% Confident interval value |
| P | = | prevalence of trypanosomes |
| N | = | total number of tsetse-fly samples |

Generalised Linear Models (GLMs) (Dobson, 1990), as implemented in the glm2 package (version 1.1.2) of the R statistical software programme (version 3.1.2), were used to test for associations of *Trypanosoma spp*. (using ITS-1 screening results) with species, collection site, sex and age of tsetse flies, as well as their interactions, as explanatory variables. Due to different species compositions in each population and region, tsetse species and collection site were collapsed into a single variable "subpopulation" (Table 2.6). All variables were treated as fixed categorical effects (Table 2.4): 1) subpopulation (seven groups); 2) sex (male and female); and 3) age (young, juvenile and old). Separate GLMs were run for each of the three most common species of trypanosomes (Models 1.1-1.3): *T. congolense*, *T. brucei* and *T. vivax*.

**Table 2.4** Variables and types for statistical analysis in GLM analyses to determine the association of trypanosome status (*T. congolense, T. brucei* and *T. vivax*) with sex and age of tsetse flies from each subpopulation (Models 1.1-1.3).

| No | Factors | Type of data | Compositions |
|---|---|---|---|
| **GLM Model 1** | | | |
| Explanatory variables | | | |
| 1. | subpopulation | categorical | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl[a] |
| 2. | sex | categorical | male and female |
| 3. | age | categorical | young, juvenile and old |
| 4 | subpopulation ∗ sex | interaction between two categorical factors | |
| 5 | subpopulation ∗ age | interaction between two categorical factors | |
| 6 | sex ∗ age | interaction between two categorical factors | |
| 7 | subpopulation ∗ sex ∗ age | interactions among three categorical factors | |
| Response variable | | | |
| Model 1.1 | *T. congolense* status | Binary | positive (1) and negative (0) |
| Model 1.2 | *T. brucei* status | Binary | positive (1) and negative (0) |
| Model 1.3 | *T. vivax* status | Binary | positive (1) and negative (0) |

[a] BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu**.**

Since some sites only contained a single species of *Glossina*, and some species were only found at a single site, separate analyses were also conducted for each species of *Glossina* to avoid confounding effects (Table 2.5). For *G. austeni* (Model 2) and *G. brevipalpis* (Model 3), associations of the presence or absence of the three trypanosome species were tested for sex and age while for *G. longipennis* (Model 4) and *G. pallidipes* (Model 5) the effects of site were also considered, since these species were found at more than one site.

All GLM Models were analysed in the binomial family of the R glm2 package, since response variables were classified as presence (0) or absence (1) of the parasites. In order to fit the best models, the variables from the full models (all explanatory variables and their interactions) were manually removed using a backward elimination technique. Variables that did not significantly (alpha =

0.05) improve the fit of the model based on Likelihood Ratio Tests (LRTs) were excluded. In order to identify relationships within variables and across variables that were significantly different from one another, TukeyHSD (Tukey's Honestly Significant Difference tests) version 3.1.2 of the R program was applied to the best fitting models for post hoc comparisons. Probabilities of trypanosome positive status for the best-fitting model for each response variable were predicted using the predict() function of R and plotted for predicted directions of each association using Microsoft Excel 2000.

**Table 2.5** Variables and types for statistical analysis in GLM analyses to determine association of trypanosome presence or absence with sex, age and sampling site of each *Glossina spp*. separately (Models 2 - 5).

| No | Factors | Type of data | Compositions |
|---|---|---|---|
| **Model 2 : Association of trypanosomes status with intrinsic factor of *G. austeni*** | | | |
| Explanatory variables | | | |
| 1 | sex | categorical | male and female |
| 2 | age | | young, juvenile and old |
| 3 | sex * age | interaction between two categorical factors | |
| Response variables | | | |
| Model 2.1 | *T. congolense* status | binary | negative (0) and positive (1) |
| Model 2.2 | *T. brucei* status | | |
| Model 2.3 | *T. vivax* status | | |
| **Model 3: Association of trypanosomes status with intrinsic factor of *G. brevipalpis*** | | | |
| Explanatory variables | | | |
| 1 | sex | categorical | male and female |
| 2 | age | | young, juvenile and old |
| 3 | sex * age | interaction between two categorical factors | |
| Response variables | | | |
| Model 3.1 | *T. congolense* status | binary | negative (0) and positive (1) |
| Model 3.2 | *T. brucei* status | | |
| Model 3.3 | *T. vivax* status | | |

| No | Factors | Type of data | Compositions |
|---|---|---|---|
| **Model 4: Association of trypanosomes status with intrinsic factor of *G. longipennis*** | | | |
| Explanatory variables | | | |
| 1 | site | categorical | Mukinyo and Sampu |
| 2 | sex | | male and female |
| 3 | age | | young, juvenile and old |
| 4 | site * sex | interaction between two categorical factors | |
| 5 | site * age | interaction between two categorical factors | |
| 6 | sex * age | interaction between two categorical factors | |
| 7 | site * sex * age | interaction between three categorical factors | |
| Response variables | | | |
| Model 4.1 | *T. congolense* status | binary | negative (0) and positive (1) |
| Model 4.2 | *T. brucei* status | | |
| Model 4.3 | *T. vivax* status | | |
| **Model 5: Association of trypanosomes status with intrinsic factor of *G. pallidipes*** | | | |
| Explanatory variable | | | |
| 1 | site | categorical | Buffalo Ridge and Mukinyo |
| 2 | sex | | male and female |
| 3 | age | | young, juvenile and old |
| 4 | site * sex | interaction between two categorical factors | |
| 5 | site * age | interaction between two categorical factors | |
| 6 | sex * age | interaction between two categorical factors | |
| 7 | site * sex * age | interaction between three categorical factors | |
| Response variables | | | |
| Model 5.1 | *T. congolense* status | binary | negative (0) and positive (1) |
| Model 5.2 | *T. brucei* status | | |
| Model 5.3 | *T. vivax* status | | |

# 2.4 Results

## 2.4.1 Tsetse fly characterisation

### 2.4.1.1 Distribution of tsetse by species

Among the four sites (Figure 2.8 and Table 2.6), more tsetse samples were collected from the Shimba Hills region than from Nguruman, despite the larger number of traps deployed in the latter (Figure 2.2 and Table 2.6). All tsetse samples were identified as one of four species of *Glossina* (Figure 2.9) but their relative abundance varied by site. *Glossina pallidipes* was at the highest frequency overall, but this also varied across sites (Table 2.6): this species was found in all populations except Sampu, and was at the highest abundance at Zungu Luka. Although the overall number of *G. austeni* sampled was higher than for *G. brevipalpis* or *G. longipennis*, this species was only found at Zungu Luka. *Glossina brevipalpis* was also only found at a single site (Buffalo Ridge) but *G. longipennis* was found at both sites within the Nguruman region and was the only species found from Sampu.

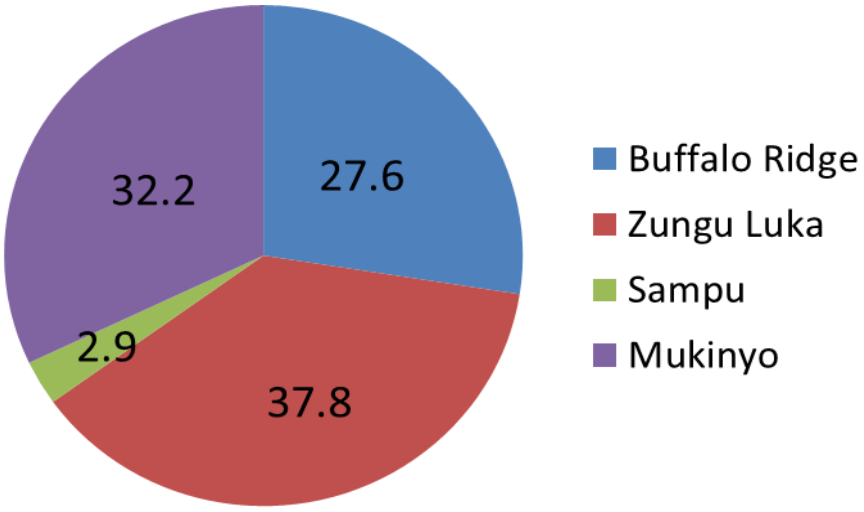



**Figure 2.8** The percentages of tsetse fly samples collected from four different sites: Bufffalo Ridge (N = 295), Zungu Luka (N = 412), Sampu (N = 325) and Mukinyo (N = 58).

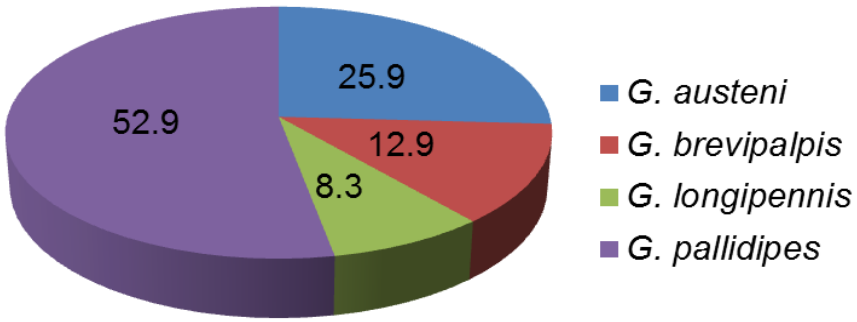



**Figure 2.9** The percentages of tsetse species among the samples collected: *G. pallidipes* (N = 577), *G. austeni* (N = 282), *G. brevipalpis* (N = 282) and *G. longipennis* (N = 90).

**Table 2.6** The numbers and percentage of tsetse fly samples trapped during June – August, 2012 in seven subpopulations, defined by the combination of site collection and *Glossina spp*.

| Duration of collection | Sites (Numbers of traps) | *Glossina spp.* | Subpopulations | Number of tsetse flies/trap | Number of samples | Percentage of samples |
|---|---|---|---|---|---|---|
| June - July 2012 | Buffalo Ridge (3 traps) | *G. pallidipes* | BRGp | 51.3 | 154 | 14.1 |
| | | *G. brevipalpis* | BRGb | 47.0 | 141 | 12.9 |
| | Zungu Luka (4 traps) | *G. pallidipes* | ZuGp | 32.5 | 130 | 11.9 |
| | | G. *austeni* | ZuGa | 70.5 | 282 | 25.9 |
| August, 2012 | Mukinyo (6 traps) | *G. pallidipes* | MuGp | 48.8 | 293 | 26.9 |
| | | *G. longipennis* | MuGl | 5.3 | 32 | 2.9 |
| | Sampu (5 traps) | *G. longipennis* | SaGl | 11.6 | 58 | 5.3 |

### 2.4.1.2 Distribution of tsetse flies by sex

Females were the dominant samples trapped in this study, except for *G. longipennis*. Overall, more female than male *G. pallidipes*, *G. brevipalpis* and *G. austeni* were sampled when compared within the same subpopulations (Figure 2.10). In contrast, for *G. longipennis*, almost twice as many males as females were found; this was at least partly due to a male bias at one site (Sampu). The highest percentage (proportion of all samples collected) of females (22.9%) was found for *G. austeni*, while the percentage of males in this species was the

lowest (2.9%); since this species was only found at Zungu Luka, whether this was due to site or tsetse species cannot be distinguished.

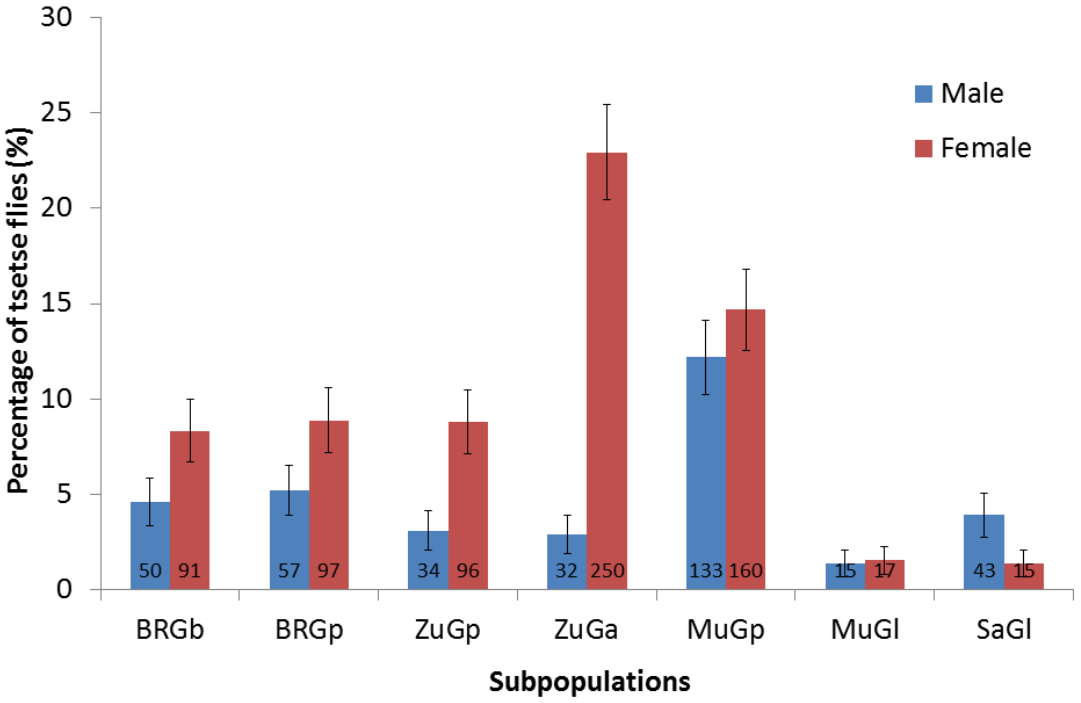



**Figure 2.10** Percentage of the total flies collected (N = 1090) for each sex in the seven subpopulations: BRGb, BRGp, ZuGa, ZuGp, MuGp, MuGl and SaGl.

Numbers of male and female flies in each subpopulation (BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu.) are indicated at the base of each bar. 95% Confidence interval (CI) bars are shown.

### 2.4.1.3 Distribution of tsetse flies by age

Although there was some variation among subpopulations, the age distribution was generally similar, with many fewer old flies then young and mid-aged flies (Figure 2.11). The age compositions of the trap-caught samples in the BRGb, ZuGa, BRGp, ZuGp and MuGp subpopulations were similar in that more than 50% were classified in a young group and had more juvenile than old flies. However, for G. *longipennis* from both Mukinyo and Sampu there were not much difference in the relative prevalence of age group.

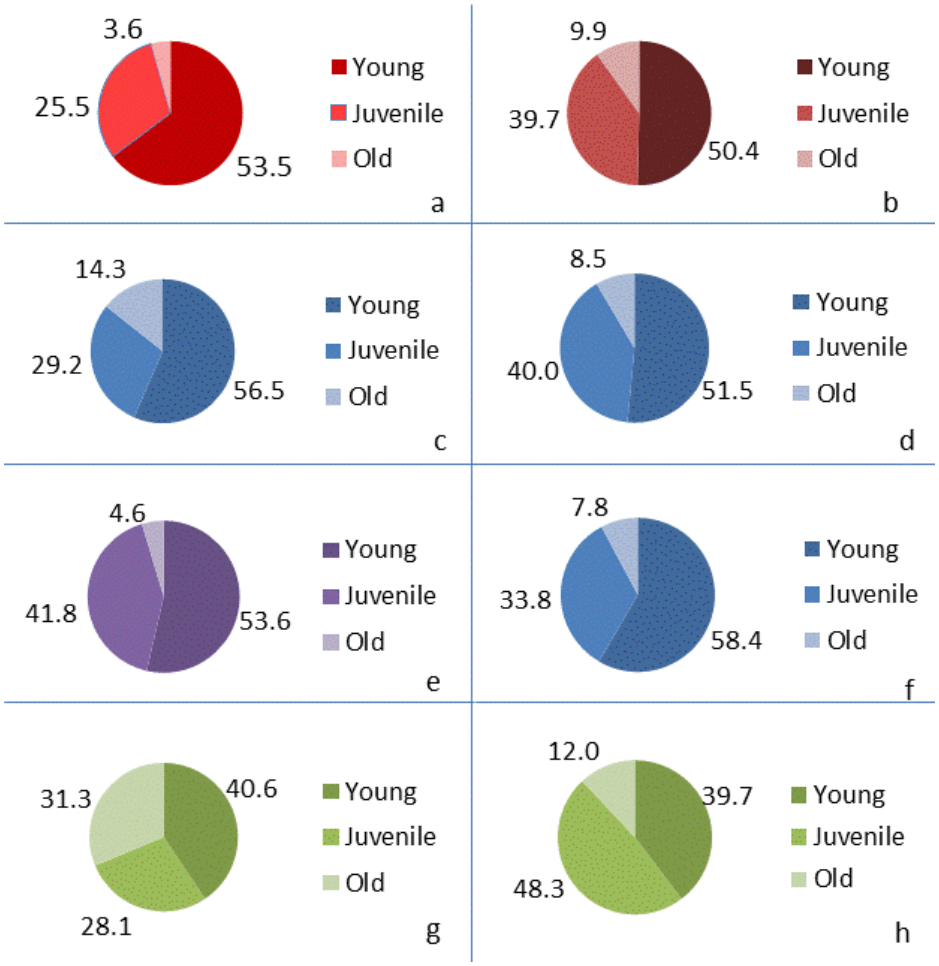



**Figure 2.11** Pie charts representing the percentage of individuals classified as young, juvenile and old using each wing fray score.

The proportions are shown for: (a) all flies; (b) BRGb, which is *G. brevipalpis* from Buffalo Ridge; (c) BRGp, which is is *G. pallidipes* from Buffalo Ridge; (d) ZuGa, which is *G. austeni* from Zungu Luka; (e) ZuGp, which is *G. pallidipes* from Zungu Luka; (f) MuGp, which is *G. pallidipes* from Mukinyo; (g) MuGl, which is *G. longipennis* from Mukinyo; and (h) SaGl, which is *G. longipennis* from Sampu

## 2.4.2 *Trypanosoma spp*. screening results using ITS-1 and specific primers

### 2.4.2.1 Comparison of amplification conditions and body parts for *Trypanosoma spp*. screening using general ITS-1 primers

A pilot study clearly indicated that the PCR conditions recommended by Njiru *et al*. (2005) (Figure 2.12, lanes 10 - 18) produced stronger amplification products from the same DNA samples (head plus proboscis of *G. pallidipes*) and was more sensitive (i.e. products were amplified from more individuals) than those of Isaac *et al*. (2013) (Figure 2.12, lanes 1 - 9).

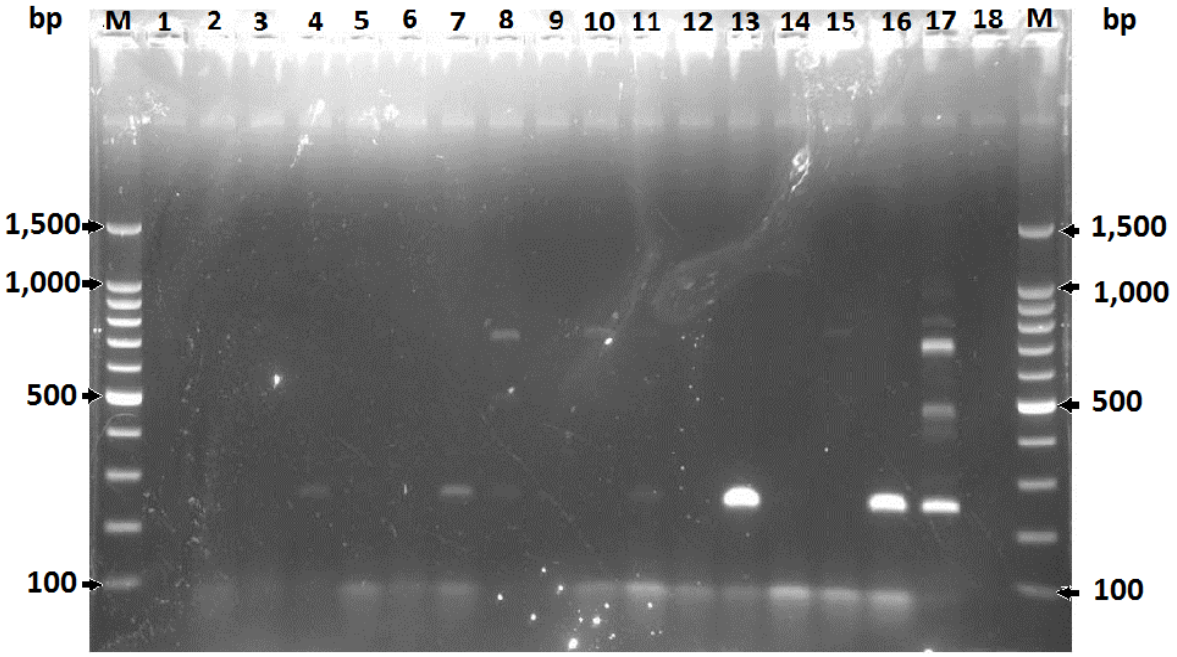



**Figure 2.12** Comparison of PCR results from head plus proboscis of the same tsetse flies between two sets of conditions.

Lanes 1 - 9: conditions of Isaac *et al*. (2013). Lane 10 - 18: conditions of Nijiru *et al*. (2005). M: DNA ladder marker 100 bp. (Promaga). Lanes 1 – 7 and 10 - 16: BRGp1HP-BRGp7HP. Lanes 8, 17: positive control mixture (*T. congolense*, *T. brucei* and *T. vivax*). Lanes 9 and 18: negative control ($ddH_2O$).

Using the Isaac *et al*. (2013) conditions, only 5/50 head plus proboscis (HP) samples and two abdomen (AB) samples tested positive (Table 2.7a) so it was not possible to quantify relative detection in the two body parts but variation among species was suggested. Mixed infections (presence of more than one species detected by PCR in a single fly) of *T. congolense* and *T. vivax* were found in both tissues but no mixed infections were found involving *T. brucei* in either. *Trypanosoma vivax* was identified in three head plus proboscis samples, and one abdomen; *T. congolense* in three head plus proboscis and one abdomen; and *T. brucei* was identified in one abdomen but no head plus proboscis.

The Njiru *et al*. (2005) PCR conditions showed much higher sensitivity than those of Isaac *et al*. (2013), allowing assessment of both the usefulness of different body parts and variation among species. There was not a high correspondence between resuls from different body parts; 10 trypanosome positive samples for head plus proboscis were negative for abdomens and eight positive samples for abdomens were negative for head plus proboscis (Table 2.7a). Mixed infections

were found in both tissues, including with *T. brucei* in abdomens. In terms of variation among species, 12 head plus proboscis parts were *T. vivax* positive, in contrast to eight abdomens. For *T. congolense*, nine samples from head plus proboscis and eight from abdomens amplified a band of the predicted size. For *T. brucei*, there were no positive results from head plus proboscis but three weakly positive bands were obtained from abdomens. These results imply that while the Njiru *et al*. (2005) conditions are more sensitive for detection in both tissues, there is not one tissue that can be concluded to provide the most reliable results. I thus chose to use head plus proboscis parts in case blood meals or other factors might inhibit amplification in abdomens and because it is the route where the infective stage of trypanosomes passes through.

Comparisons among PCR reagents also suggested variation that could affect interpretation of the presence of trypansosomes in tsetse flies. ITS-1 fragments of all three positive controls tested (*T. congolense* savannah, *T. brucei* and *T. vivax*) were amplified using *Taq* DNA Polymerase from Thermo Scientific while only one (*T. congolenses* savannah) was detectable using *Taq* DNA Polymerase from Invitrogen, and there was no ITS-1 amplification for any of the species using Dream*Taq* Green PCR Master Mix from Thermo Scientific (Figure 2.13).

**Table 2.7** Comparison of pilot ITS-1 screening from head plus proboscis and abdomen DNA preparations for trypanosome detection in 50 *G. pallidipes* sampled from the Shimba Hills.

**Table 2.7a** Trypanosome screening using the PCR conditions of Isaac.

| PCR condition of Isaac | | Head plus proboscis | | |
|---|---|---|---|---|
| | | Positive | Negative | Total |
| Abdomen | Positive | 1 | 1 | 2 |
| | Negative | 4 | 44 | 48 |
| | Total | 5 | 45 | 50 |

**Table 2.7b** Trypanosome screening using the PCR conditions of Njiru.

| PCR condition of Njiru | | Head plus proboscis | | |
|---|---|---|---|---|
| | | Positive | Negative | Total |
| Abdomen | Positive | 9 | 8 | 17 |
| | Negative | 10 | 23 | 33 |
| | Total | 19 | 31 | 50 |

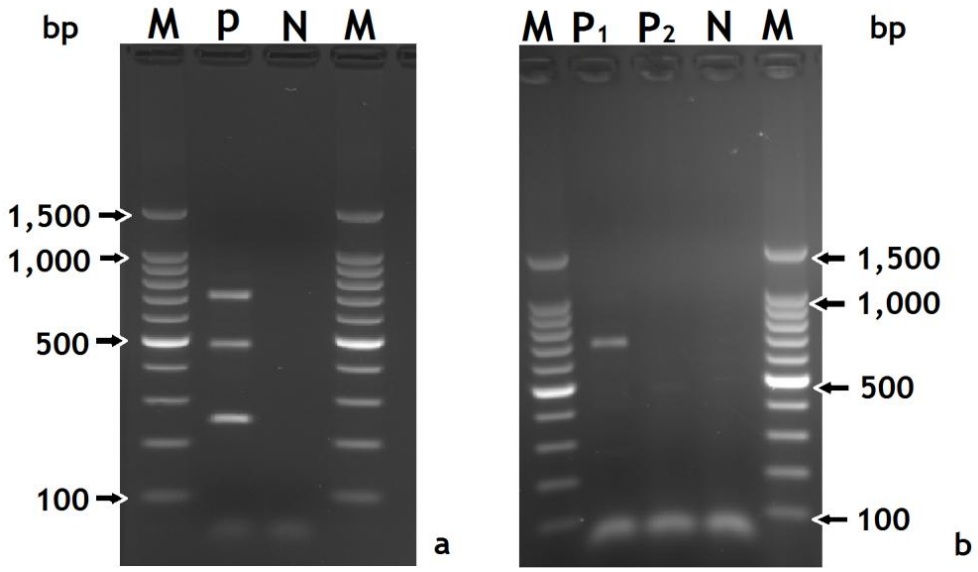



**Figure 2.13** ITS-1 amplification of *T. congolense* savannah, *T. brucei* and *T. vivax* positive controls using different *Taq* DNA Polymerase in the same PCR conditions.

(a) Amplification using *Taq* DNA Polymerase from Thermo Scientific. Lane P: amplification of *T. congolense* (approximately 700 bp), *T. brucei* (approximately 500 bp) and *T. vivax* (approximately 250 bp) positive control mixture. Lane N: negative control ($ddH_2O$). (b) Amplification using *Taq* DNA Polymerase from Invitrogen and Dream*Taq* Green from Thermo Scientific. Lane $P_1$: amplification of *T. congolense* (approximately 700 bp) positive control using *Taq* DNA Polymerase from Invitrogen. Lane $P_2$: no amplification of trypanosome positive controls using Dream*Taq* Green from Thermo Scientific. M: DNA ladder marker 100 bp. (Promega).

BLAST analysis of sequences from six plasmids of each of the purified bands predicted to correspond to the four species (approximately 600 bp for *T. congolense* kilifi, 700 bp for *T. congolense* savannah, 500 bp for *T. brucei* and 250 bp for *T. vivax*, confirmed that the PCR products from the ITS-1 primers were gene fragments belonging to *Trypanosoma spp.* (Table 2.8). The 597 bp and 688 bp PCR products were confirmed to correspond to *T. congolens*e kilifi and *T. congolens*e savannah, respectively, with high percentages of query cover and sequence identity. The 251 bp PCR products showed only 66% query cover with ITS-1 fragments of *T. vivax* in BLAST, but with more than 90% identity. The 500 bp bands were more problematic; plasmid sequencing results of ITS-1 gene fragments that were suspected to be *T. brucei* could only identify the amplified fragments as *Trypanosoma spp.,* but with low sequence cover. The identity of these fragments was closest to *T. congolense*, *T. brucei* and *T. simiae,* respectively.

**Table 2.8** Plasmid sequencing results of *Trypanosoma spp.* from ITS-1 cloning.

Shown are the clone name (indicating the sample site and species of tsetse, the trypanosome species predicted, the sequencing primer, and the clone number), the size of the fragment sequenced, a description of the closest match found in BLAST, the proportion of the sequence that matched a sequence in BLAST (% query cover), the similarity of the sequence to that in BLAST (% Identity) and the GenBank accession number.

| Plasmid names | Sizes (bp) | Description | Query cover (%) | Identity (%) | Accession |
|---|---|---|---|---|---|
| *T. congolense* kilifi | | | | | |
| SaGl18_Tck_M13F_1<br>SaGl18_Tck_M13F_2<br>SaGl18_Tck_M13F_3<br>SaGl18_Tck_M13F_4<br>SaGl18_Tck_M13F_5<br>SaGl18_Tck_M13F_6 | 597 | *T. congolense* kilifi | 100 | 94 | U22317.1 |
| *T. congolense* savannah | | | | | |
| MuGp9_Tcs_M13F_1<br>MuGp9_Tcs_M13F_3 | 688 | *T. congolense* | 100<br>95<br>100 | 97<br>97<br>94 | U22315.1<br>JN673389.1<br>FJ712718.1 |
| MuGp9_Tcs_M13F_2<br>MuGp9_Tcs_M13F_4<br>MuGp9_Tcs_M13F_5<br>MuGp9_Tcs_M13F_6 | 688 | *T. congolense* | 100<br>95<br>100 | 97<br>97<br>94 | U22315.1<br>JN673389.1<br>FJ712718.1 |
| *T. vivax* | | | | | |
| ZuGp33_Tv_M13F 10 | 251 | *T. vivax* | 66 | 98<br>98<br>97 | KC196699.1<br>EU482078.1<br>EU482082.1 |
| ZuGp33_Tv_M13F 7<br>ZuGp33_Tv_M13F 8<br>ZuGp33_Tv_M13F 9<br>ZuGp33_Tv_M13F 11<br>ZuGp33_Tv_M13F 12 | 251 | *T. vivax* | 66 | 96<br>96<br>96 | KC196699.1<br>EU482078.1<br>EU482082.1 |
| *T. brucei* | | | | | |
| ZuGp96_Tb_M13F_1 | 453 | *T. congolense* | 32 | 100 | JX910374.1 |
| ZuGp96_Tb_M13F_4<br>ZuGp96_Tb_M13F_5 | | *T. brucei* | 29 | 99 | JX910373.1 |
| ZuGp96_Tb_M13F_6 | | *T. simiae* | 19 | 100 | AB625446.1 |
| ZuGp96_Tb_M13F_2<br>ZuGp96_Tb_M13F_3 | 445 | *T. congolense* | 33 | 99 | JX910374.1 |
| | | *T. brucei* | 30 | 98 | JX910373.1 |
| | | *T. simiae* | 10 | 99 | AB625446.1 |

#### 2.4.2.2 Comparison of *Trypanosome spp*. screening results using ITS-1 primers with specific primers

In total, 1090 samples were screened using the ITS-1 general primers, with varying percentages of trypanosome species present in different sites and tsetse species (Appendix A.9). Screening using species (or subspecies) specific primers showed varying levels of correspondence with predictions based on size bands of ITS-1 products. The TCS primers yielded PCR products of the expected size (319 bp) for *T. congolense* savannah (Figure 2.14) but there were also some unexpected PCR bands (500 and 700 bp), especially in the *G. pallidipes* and *G. austeni* populations. Screening results of *T. congolense* savannah using the ITS-1 and the TCS primers in general were not well correlated (Figure 2.15), with much higher detection sensitivity using the TCS primers. Overall (N = 1090), there were 52.7% positive *T. congolense* savannah flies using the TCS primers, compared to only 11.6% using the ITS-1 primers. There were only 9.4% samples that were both ITS-1 and TCS positive while 45.2% were both negative. There were 2.1% of ITS-1 positive flies that were TCS negative whereas 43.2% of TCS positive flies were ITS-1 negative. Among the twenty-seven samples tested, no positive amplifications were identified using the TCF specific primers. For TCK primers, 294 bp PCR products were amplified, with weakly amplified 500 bp and 700 bp unexpected bands in some samples (Figure 2.16). Of the 47 flies screened, the ITS-1 screening results of *T. congolense* kilifi generally were consistent with the results from the specific primers (Figure 2.17), with the majority of samples showing similar amplification in each (only four ITS positive samples were TCK negative and only a single TCK positive sample was ITS negative). According to the screening results of the TCS and TCK primers, there also was no bias in consistency of amplification with ITS-1 associated with particular subpopulations.

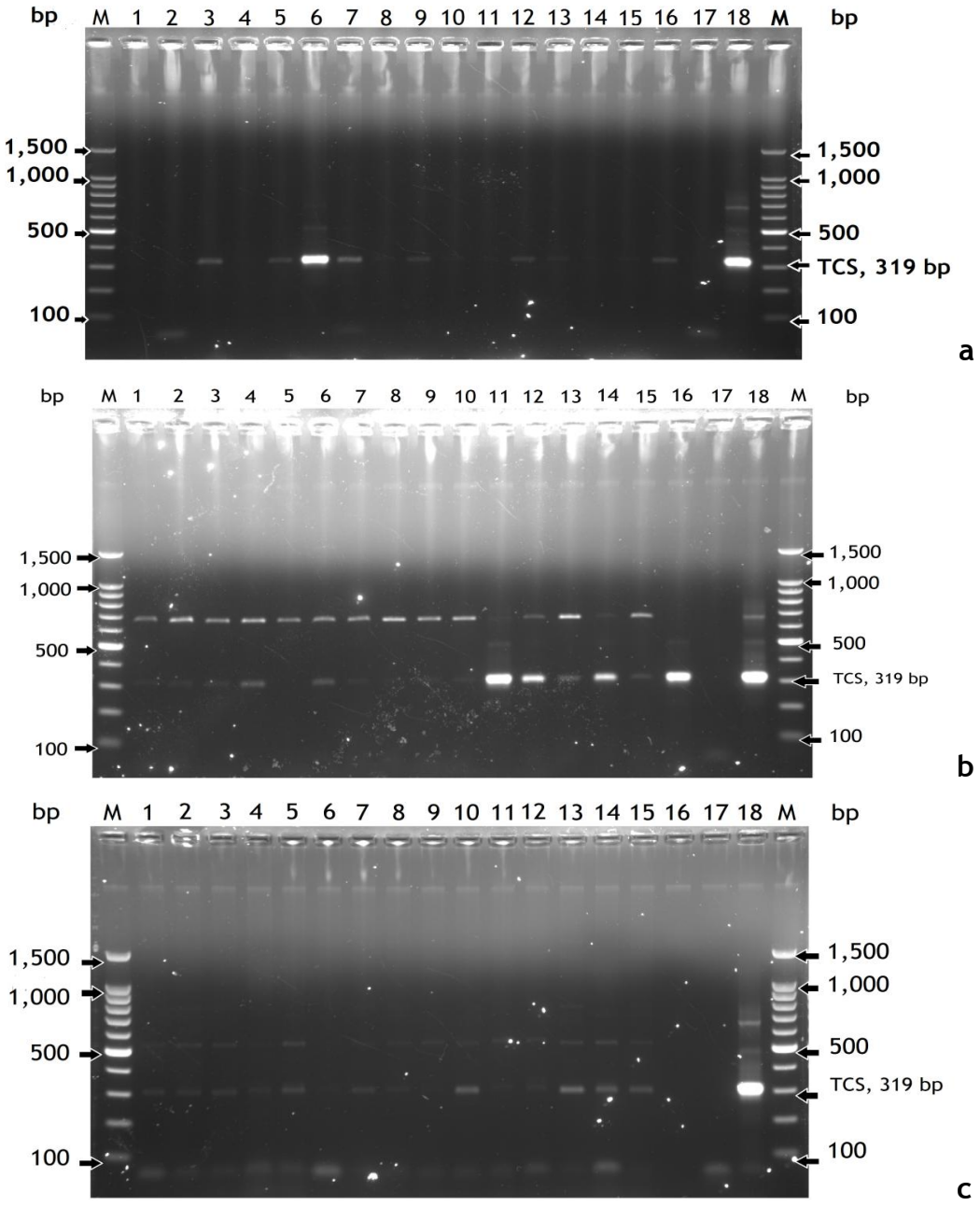



**Figure 2.14** PCR products of *T. congolense* savannah (319 bp) using TCS specific primers in tsetse flies.

(a) Screening results of 16 flies in the BRGb subpopulation. Lanes 3, 5 - 16: *T. congolense* savannah positive *G. brevipalpis*. Lanes 1, 2 and 4: lack of *T. congolense* savannah amplification. Lane 6: unexpected sized PCR products (500 and 700 bp). Lane 17: negative control (nuclease free water). Lane 18: *T. congolense* savannah positive control. (b) Screening results of 16 flies in the BRGp subpopulation. Lanes 1 - 4 and 6 - 16: *T. congolense* savannah positive *G. pallidipes*. Lane 5: lack of *T. congolense* savannah amplification. Lane 17: negative control (nuclease free water). Lane 18: *T. congolense* savannah positive control. Lanes 1 - 15: unexpected sized PCR products (700 bp). (c) Screening results of 16 flies in the ZuGa subpopulation. Lanes 1 - 5, 7 - 8 and 10 - 15: TCS positive *G. austeni*. Lanes 6, 9 and 16: lack of *T. congolense* savannah amplification. Lane 17: negative control (nuclease free water). Lane 18: *T. congolense* savannah positive control. Extra PCR bands were found in all positive samples. M: DNA ladder marker 100 bp. (Promega).

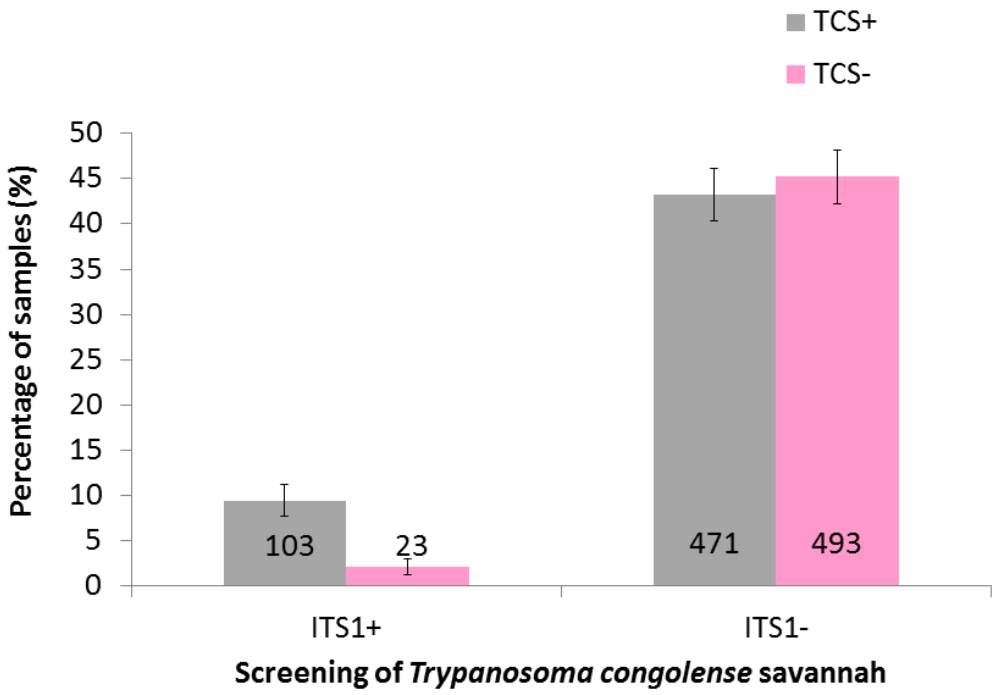



**Figure 2.15** Comparison of *T. congolense* savannah screening results from all tsetse flies sampled (N = 1090), using ITS-1 and TCS-specific primers.

Numbers of tsetse flies from the screening results are represented at the base of each bar. 95% CI bars are shown.

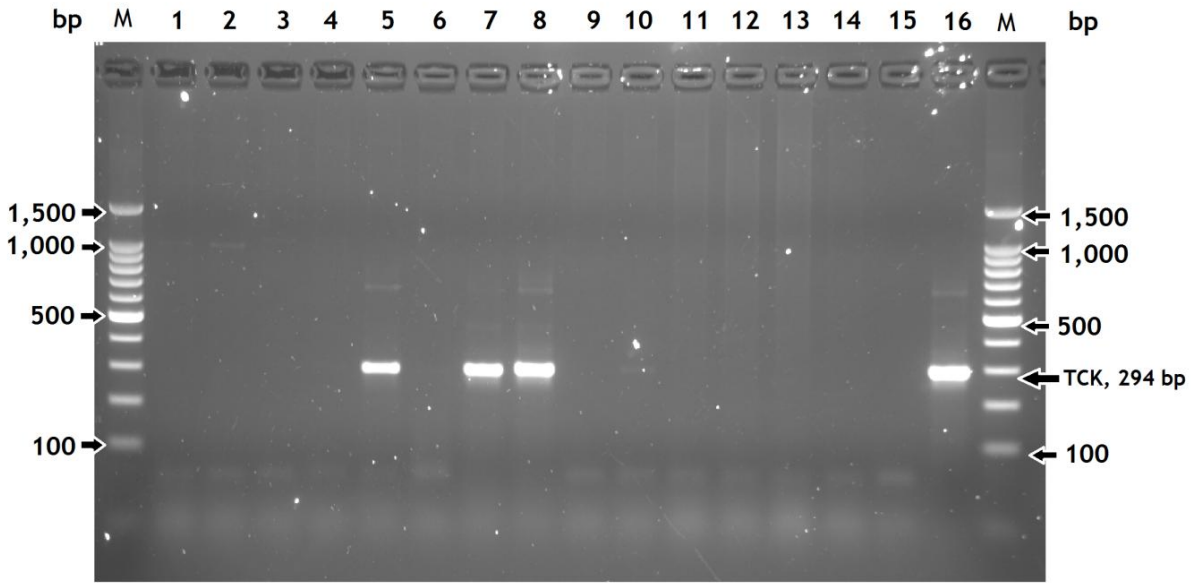



**Figure 2.16** PCR products of *T. congolense* kilifi (294 bp) amplified using TCK-specific primers in tsetse flies.

M: DNA ladder marker 100 bp. (Promega). Lanes 5 and 7 - 8: *T. congolense* kilifi positive. Lanes 1-4, 6, and 9-14: TCK negative. Lane 15; negative control (nuclease free water). Lane 16: *T. congolense* kilifi (in *G. pallidipe*) positive controls. Unexpected PCR products (around 500 and 700 bp) were visualised in all positive samples, including the positive control.

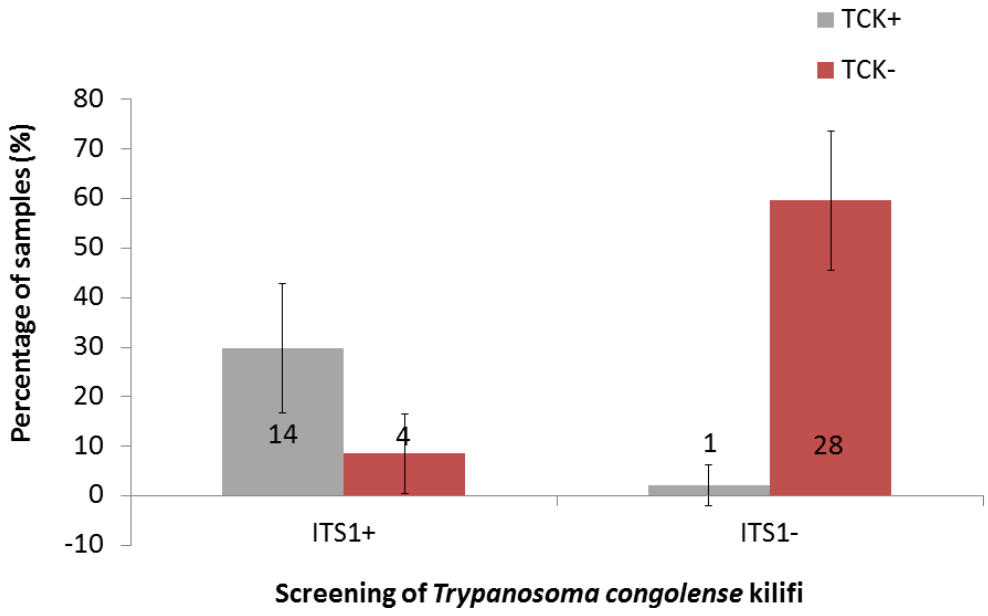



**Figure 2.17** Comparison of *T. congolense* kilifi screening results from 47 flies using ITS-1 and TCK-specific primers.

Numbers of tsetse flies from the screening results are represented at the base of each bar. 95% CI bars are shown.

Amplification of *T. brucei* fragments using the TBR-specific primers was visualised at 173 bp, with two unexpected bands (approximately 350 and 525 bp) (Figure 2.18). Only 53.3% of ITS-1 positive flies were also TBR positive. Matches of the screening results between the primers revealed a good agreement for samples from Zungu Luka but very poor for samples from the Buffalo Ridge, Mukinyo and Sampu (Table 2.9). The TBR primers seem to be sensitive for *T. brucei* screening in Zungu Luka, but they were not sensitive for *T. brucei* strain in Mukinyo and Sampu.

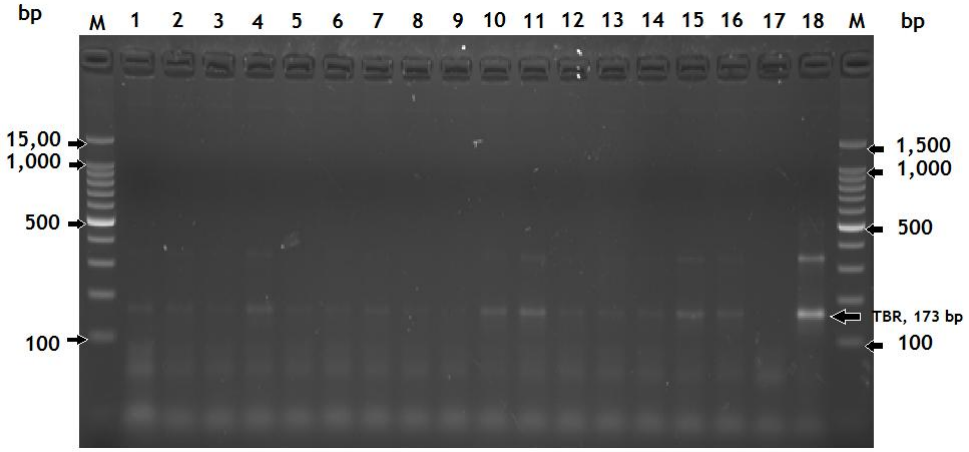



**Figure 2.18** PCR products of 173 bp fragments, which were amplified with TBR specific primers for *T. brucei*.
M: DNA ladder marker 100 bp. (Promega). Lanes 1 - 16: 500 bp ITS-1 positive samples. Lane 17; negative control (nuclease free water). Lane 18: *T. brucei* (in *G. pallidipe*) positive controls. Unexpected PCR products (around 350 and 525 bp) were visualised in positive samples, including the positive control.

**Table 2.9** Comparison of *T. brucei* identification in tsetse flies from seven subpopulations using ITS-1 and TBR specific primers.

| Source of samples | Tsetse species | Number of ITS-1 positive samples | Number of TBR positive samples |
|---|---|---|---|
| Zungu Luka | *G. austeni* | 19 | 19 |
| | *G. pallidipes* | 11 | 11 |
| Buffalo Ridge | *G. brevipalpis* | 14 | 6 |
| | *G. pallidipes* | 4 | 2 |
| Mukinyo | *G. pallidipes* | 21 | 0 |
| | *G. longipennis* | 1 | 0 |
| Sampu | *G. longipennis* | 5 | 0 |

## 2.4.3 Association of *Trypanosoma spp*. with intrinsic factors of tsetse flies

### 2.4.3.1 *Trypanosoma spp*. prevalence and distributions

All trypanosome negative samples (631 flies) were positive when they were screened with *Glossin*a ITS-1 primers, suggesting that lack of amplification during the trypanosome screening was not due to poor quality DNA. Thus, all samples were included in subsequent analyses.

Since there were differences in correspondence between each of the species-specific primers and the general ITS-1 primers that could suggest differences in relative sensitivity, I took a conservative approach and used the general primers for all of the remaining comparisons. Moreover, using a single PCR for identification of all species reduces potential for error. From the overall trapped samples (1090 flies), there were 459 trypanosome positive flies based on the general ITS-1 primers (42.1%) (Appendix A.9). Among these infected flies, there were 362 with single trypanosome species present (33.2%) (Figure 2.19) and the remaining samples showed mixed species present (8.9%): 83, 10 and 4 flies showed amplification of two, three and four trypanosome species, respectively. These will be referred to as "mixed infections" but it should be noted that detection in the PCR does not necessarily imply that the trypanosomes have established in the fly. Across all trypanosome-infected flies (Figure 2.20; Appendix A.9), *T. vivax* showed the highest positive frequency across populations, followed by *T. congolense* and *T. brucei*; *T. simiae* and *T. godfreyi* were much rarer.

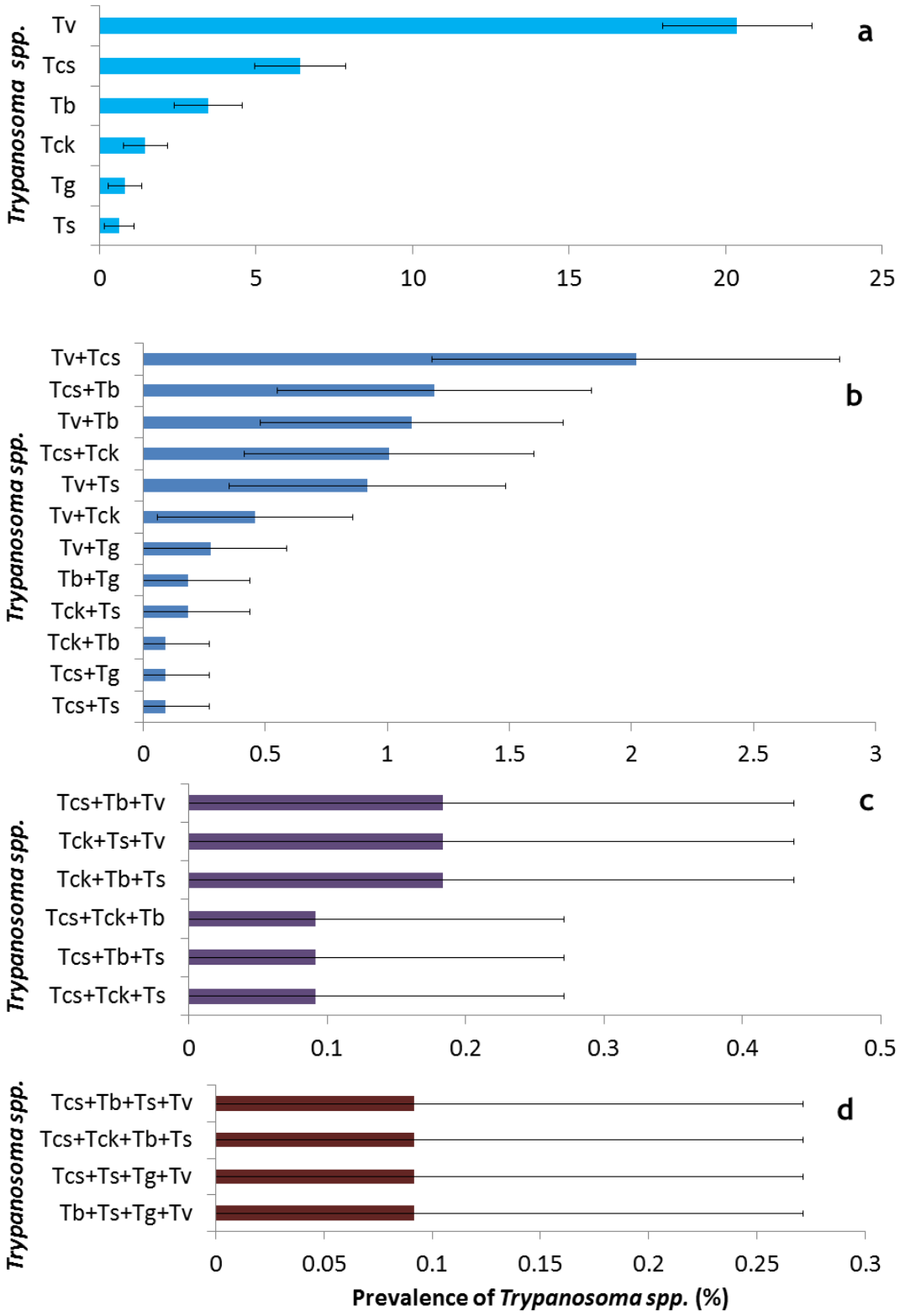



**Figure 2.19** Prevalence of single and mixed species of trypanosomes in tsetse flies identificated based on ITS-1 screening.

(a) Single species of trypanosome amplified (362 flies). (b) Two species of trypanosome amplified (83 flies). (c) Three species of trypanosomes amplified (10 flies). (d) Four species of trypanosome amplified (four flies). Numbers of tsetse flies from the screening results are represented on each bar. 95% CI bars are shown. Tcs: *T. congolense* savannah; Tck: *T. congolesne* kilifi; Tb: *T. brucei*; Tv: *T. vivax*; Ts: *T. simiae*; and Tg: *T. godfreyi*.

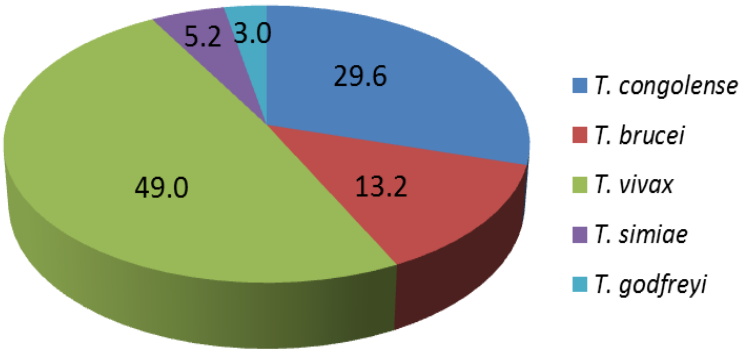
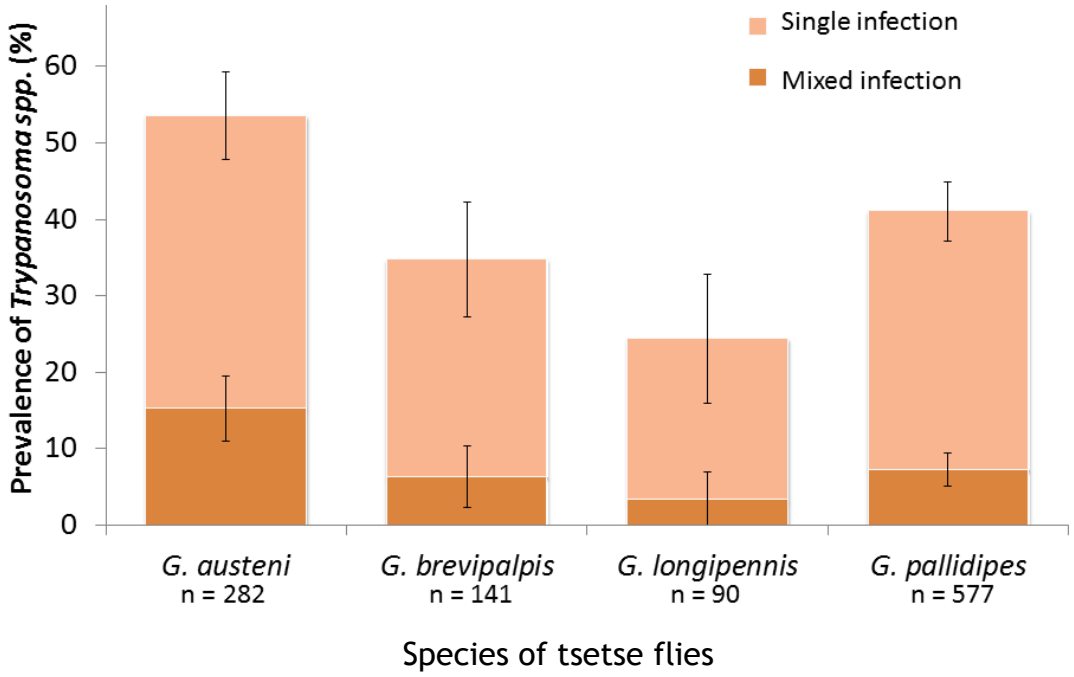



**Figure 2.20** The percentage of different trypanosome species found in all trypanosome positive results (N = 574*) based on ITS-1 trypanosome screening.

*Trypanosoma vivax* (95% CI = 44.9 - 53.0%) was the predominant species of trypanosome followed by *T. congolense* (95% CI = 25.9 - 33.4%), *T. brucei* (95% CI = 10.5 - 16.0%), *T simiae* (95% CI = 3.4 - 7.0%) and *T. godfieyi* (95% CI = 1.6 - 4.3%) * There were mixed trypanosome species present in 115 tsetse flies.

### 2.4.3.2 *Trypanosoma spp*. prevalence by *Glossina species*

Across the four *Glossina* species, *G. austeni* had the highest prevalence of trypanosomes (53.6%, with 15.3% mixed infections), compared to *G. pallidipes* (41.1%, with 7.3% mixed infections), *G. brevipalpis* (34.8%, with 6.4% mixed infections) and *G. longipennis* (24.4%, with 3.3% mixed infections) (Figure 2.21a). *Trypanosoma vivax* was the predominant pathogenic parasite in every species of tsetse fly, except for *G. brevipalpis* (Figure 2.21b). *Trypanosoma congolense* was the most prevalent in *G. brevipalpis,* but the savannah subspecies was found more frequently than the kilifi type; *T. vivax* was the next most common trypanosome species found. Although *T. simiae* was relatively rare, the pathogen was found at the highest abundance in *G. austeni*. For *G. longipennis,* equal prevalence of *T. congolense* savannah and *T. brucei* was found but no *T. godfreyi* were detected.



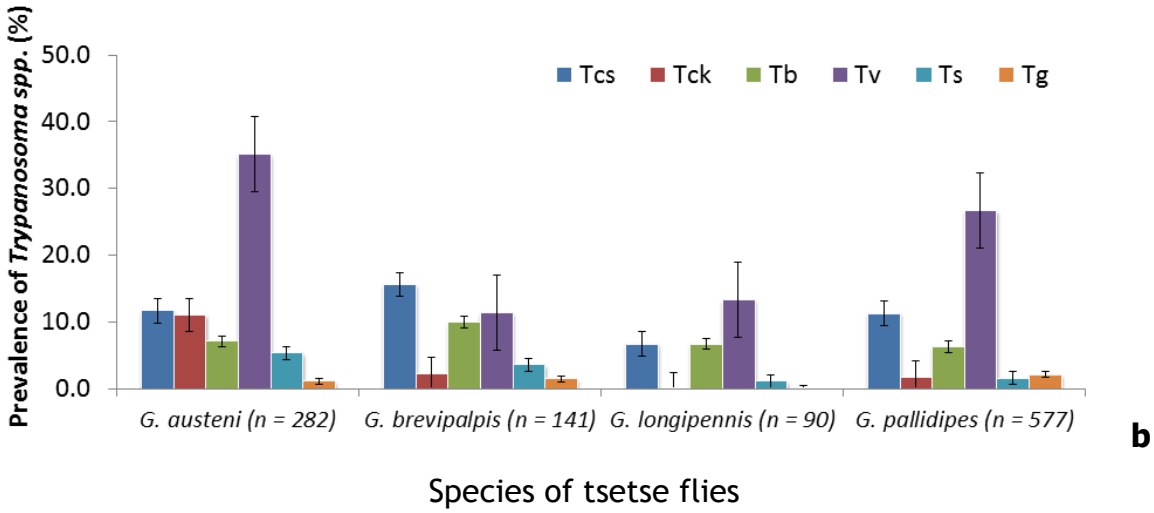




**Figure 2.21** Trypanosome prevalence among all tsetse flies sampled (N = 1090) in *G. pallidipes*, *G. austeni*, *G. brevipalpis* and *G. longipennis*.

Prevalence of: (a) single and mixed trypanosome species infections; and (b) each trypanosome species in individual *Glossina species*. Total tsetse fly numbers for each tsetse species are shown below the graphs. 95% CI bars are indicated. Tcs: *T. congolense* savannah; Tck: *T. congolense* kilifi; Tb: *T. brucei*; Tv: *T. vivax*; Ts: *T. simiae*; and Tg: *T.godfreyi*.

#### 2.4.3.3 *Trypanosoma spp*. prevalence by sampling site

Tsetse flies from the four different sites sampled showed different trypanosome prevalences (19.0 - 55.6%), with 5.2 - 12.9% mixed trypanosome species infections (Figure 2.22a) and different species compositions (Figure 2.22b). Tsetse flies from Zungu Luka showed the highest overall trypanosome prevalence (55.6%, with 12.9% mixed infections), compared to Mukinyo (36.0%, with 6.5% mixed), Buffalo Ridge (34.6%, with 6.8% mixed) and Sampu (19.0%, with 5.2% mixed). *Trypanosoma. vivax* was the predominant species of trypanosome, followed by *T. congolense*, *T. brucei* , *T. simiae* and *T. godfreyi* in tsetse flies from Buffalo Ridge, Zungu Luka and Mukinyo. However, at Sampu, *T. vivax* and *T. brucei* were found in equal proportions, followed by *T. congolense* savannah and *T. simiae*; *T. congolense* kilifi and *T. godfreyi* were not found at this site.

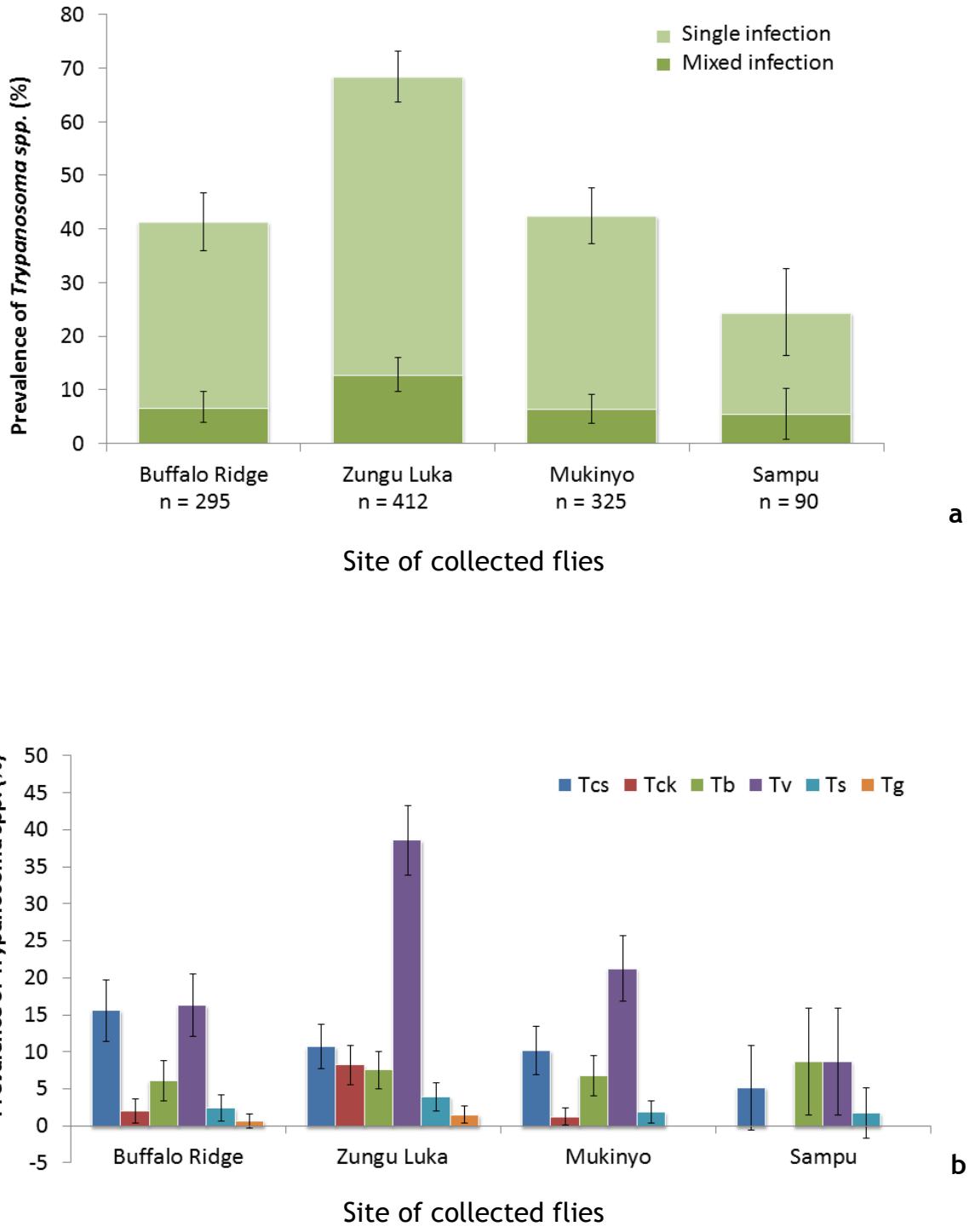



**Figure 2.22** Trypanosome prevalence of all tsetse flies sampled (N = 1090) from Buffalo Ridge, Zungu Luka, Mukinyo and Sampu.

Prevalence of: (a) single and mixed infections with trypanosomes; (b) trypanosome species present at each site. Total tsetse fly numbers for each site are shown below the graphs. 95% CI bars are indicated. Tcs: *T. congolense* savannah; Tck: *T. congolense* kilifi; Tb: *T. brucei*; Tv: *T. vivax*; Ts: *T. simiae*; and Tg: *T. godfreyi*.

#### 2.4.3.4 *Trypanosoma spp*. prevalence by subpopulations

In order to consider possible interactions between tsetse species and geographic locations, prevalence in the seven subpopulations was also considered, with extensive variation in trypanosome prevalence among site-species combinations (19.0 - 60.0%) (Figure 2.23a). In general, different species of flies from the same regions tended to show similar results in terms of overall prevalence of trypanosomes. Tsetse flies in the ZuGp subpopulation had the highest prevalence, compared to the ZuGa, MuGp, MuGl, BRGp, BRGb, and SaGl subpopulations, suggesting that site was a stronger factor than species of tsetse within these populations. Mixed trypanosome species infections were found in all subpopulations, except for MuGl. The highest prevalence of mixed infections was found in the ZuGa subpopulation. This was in contrast to *G. pallidipes* from the same sample site (ZuGp), which showed a lower number of mixed infections than any of the other populations.

In terms of specific trypanosome species (Figure 2.23b), *T. vivax* showed the highest prevalence in the ZuGp subpopulation, followed by ZuGa, MuGl, MuGp, BRGp, BRGb, and SaGl subpopulations. However, in the BRGb subpopulation, the prevalence of *T. vivax* was lower than that of *T. congolense* savannah type, and in SaGl, it was found at the same rate as *T. brucei*. Although the majority of all subpopulations had higher prevalence of *T. brucei* than *T. congolense* kilifi, it was the opposite for the ZuGa subpopulation. There were no *T. congolense* kilifi found in the MuGl or the SaGl subpopulations and they were rare in MuGp, suggesting that this type of *T. congolense* is not widely found in the Nguruman region. In contrast, the savannah type was found in all subpopulations, as similar to *T. brucei* and *T. vivax*.

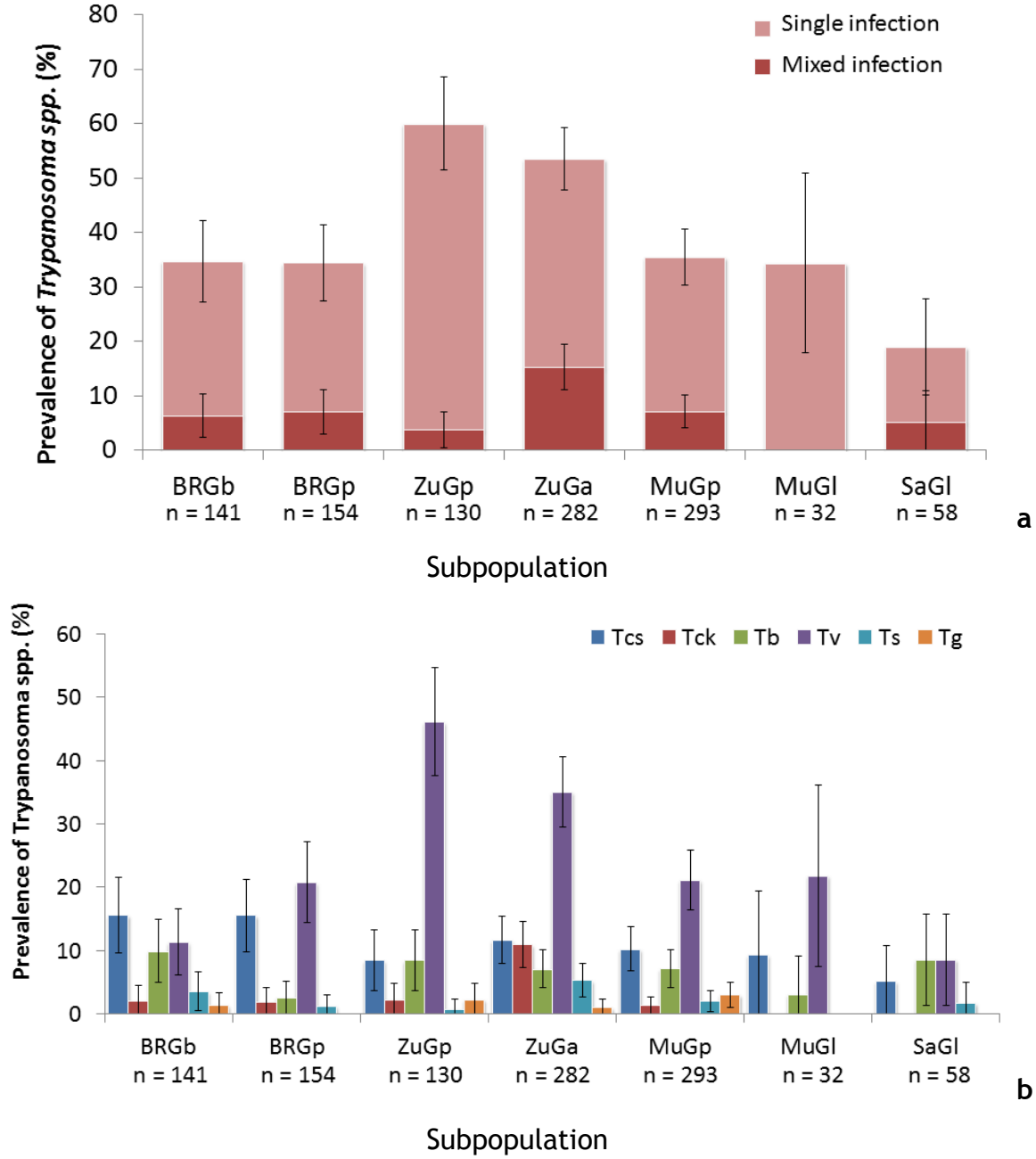



**Figure 2.23** *Trypanosoma spp*. prevalence of all sampled tsetse flies (N = 1090) in the BRGb, BRGp, ZuGp, ZuGa, MuGp, MuGl and SaGl subpopulations.

Prevalence of: (a) single and mixed infections with trypanosomes; (b) trypanosome species present in each subpopulation. Total tsetse fly numbers for each subpopulation are shown below the graphs. 95% CI bars are indicated. Tcs: *T. congolense* savannah; Tck: *T. congolense* kilifi; Tb: *T. brucei*; Tv: *T. vivax*; Ts: *T. simiae*; and Tg: *T. godfreyi*.

### 2.4.3.5 *Trypanosoma spp*. prevalence by sex

Trypanosome prevalence varied between males and females in different supopulations (Figure 2.24), but overall, it was higher in female than in male flies (Figure 2.25a). The proportion of single versus mixed trypanosome infection, however, was similar between sexes (Figure 2.25a). *Trypanosoma vivax* was the predominant pathogen in both male and female flies, followed by *T. congolense* (savannah and kilifi type), *T. simiae* and *T. godfreyi*, respectively. Prevalence of *T. vivax, T. congolense savannah, T. congolense kilifi* and *T. simiae* in males was slightly lower than in females, but prevalence of *T. brucei* and *T. godfreyi* was slightly higher in males than in females.

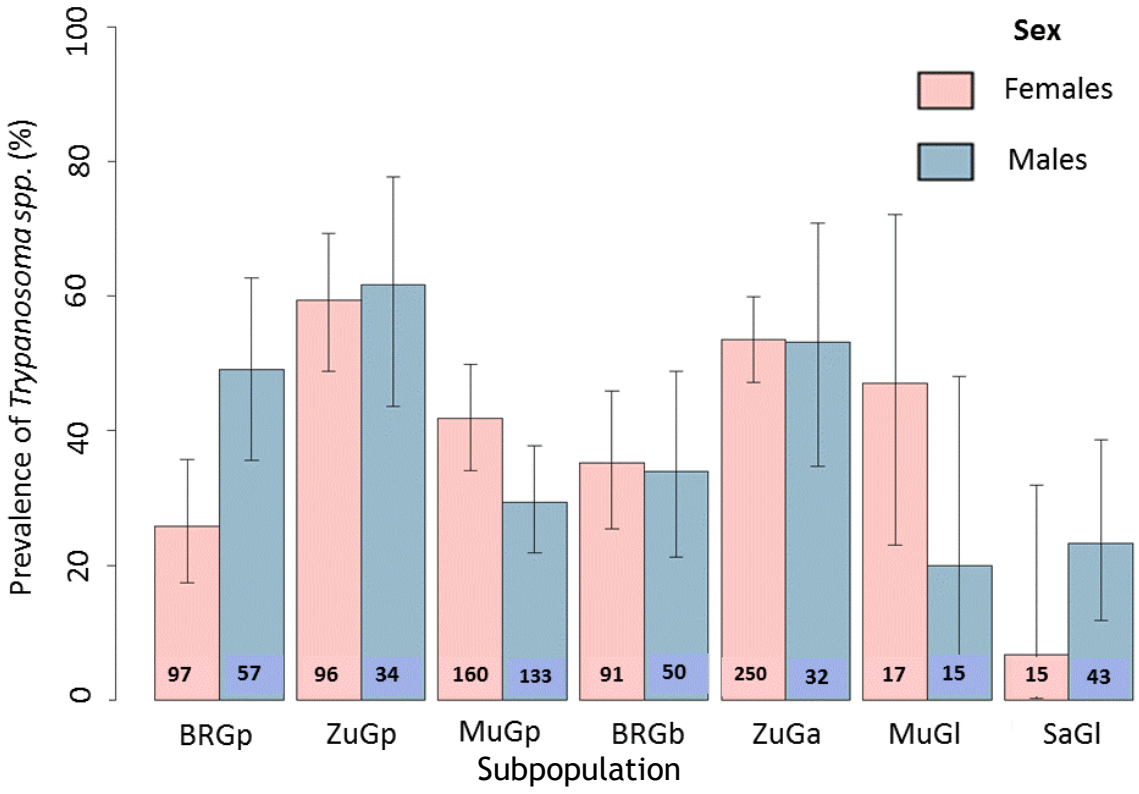



**Figure 2.24** Histogram of trypanosome prevalence in male and female tsetse samples from each of the subpopulations based on PCR screening with universal ITS-1 primers.

Numbers of male and female flies in each subpopulation are represented at the base of each bar. 95% Confidence interval (CI) bars are shown.

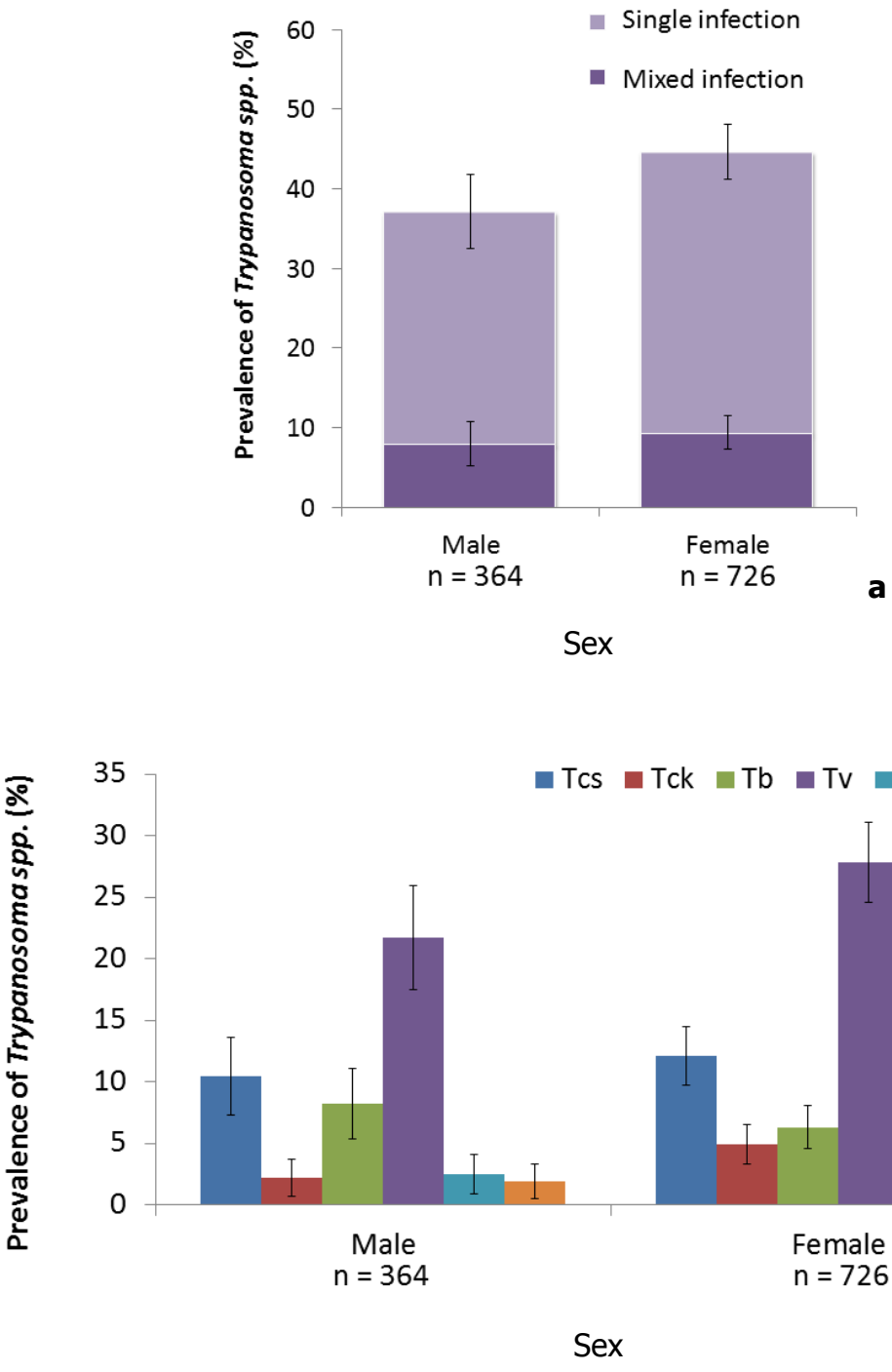



**Figure 2.25** Prevalence of trypanosomes in tsetse flies by sex using ITS-1 primers.

Prevalence of: (a) single and mixed infections with trypanosomes; (b) trypanosome species present in males and females. Total tsetse fly numbers for each sex are shown below the graphs. 95% CI bars are indicated. Tcs: *T. congolense* savannah; Tck: *T. congolense* kilifi; Tb: *T. brucei*; Tv: *T. vivax*; Ts: *T. simiae*; and Tg: *T. godfreyi*.

#### 2.4.3.6 *Trypanosoma spp*. prevalence by age

Based on wing fray scores of all sampled flies, trypanosome prevalence tended to increase with age but this was not found consistently across sites and species of flies (Figure 2.26a). Increasing prevalence with age was found for *G. austeni* and *G. pallidipes* but *G. longipennis* did not show a change in prevalence with age and *G. brevipalpis* showed decreasing prevalence with age (Figure 2.26c). The dynamics also varied by sex. For females, the highest infections were found for individuals with a wing fray score of 2 - 3 and infections declined at wing fray score 5 - 6 (Figure 2.29). In contrast, young males showed low rates of infection until wing fray score 3.5 when they showed similar rates to females; they also showed a peak in infection at wing fray score 5 - 6. There was no consistent age-related pattern for mixed infections (Figure 2.27). In most species of trypanosomes, although the same trend of increasing prevalence with age was observed, differences between age classes tended to be small and no increase was observed for *T. congolense* savannah (Figure 2.28).

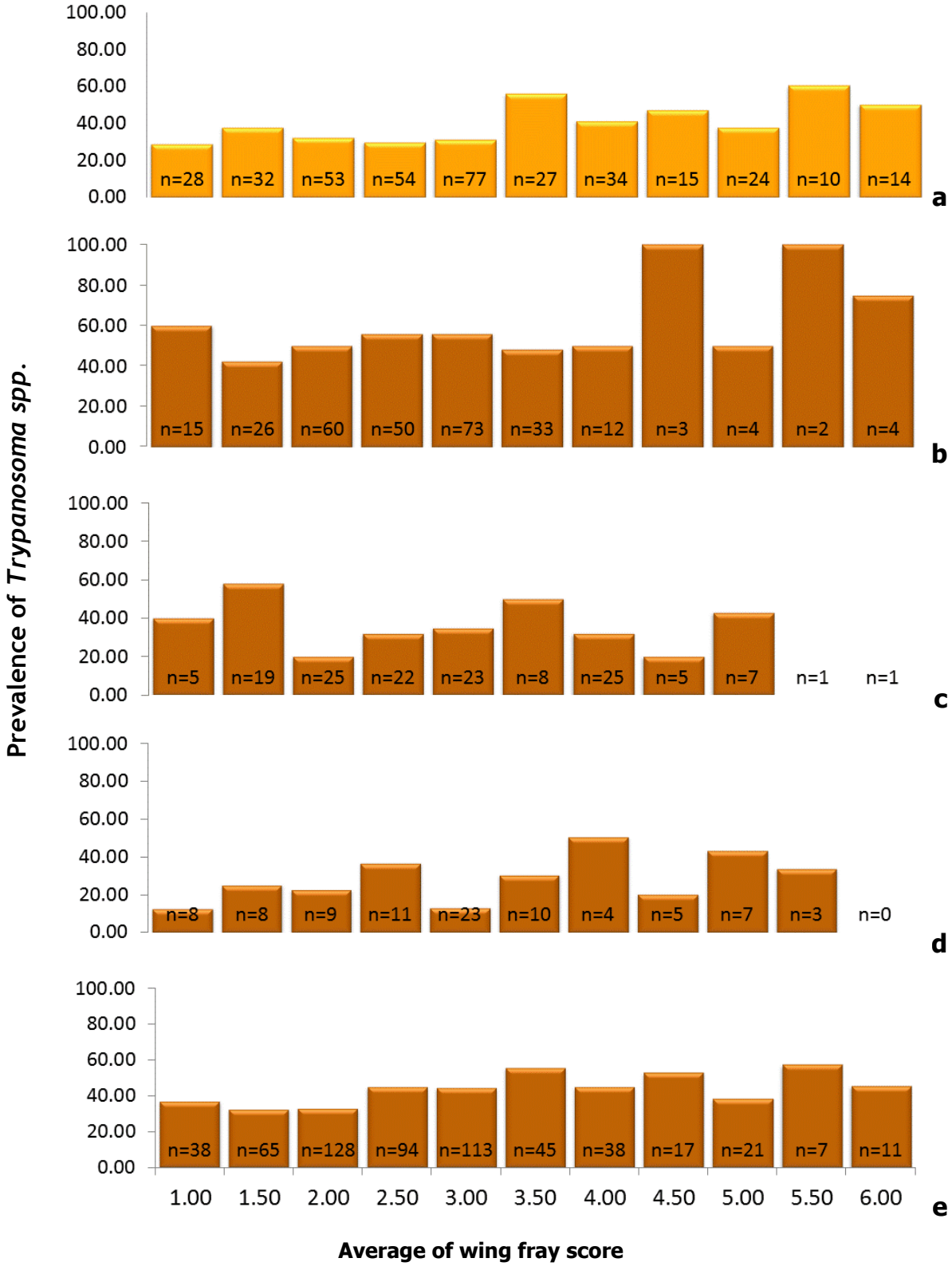



**Figure 2.26** Distribution of trypanosome prevalence by age using wing fray scores.

Prevalence of trypanosomes by age in: (a) All sampled flies; (b) *G. austeni*; (c) *G. brevipalpis*; (d) *G. longipennis*; (e) Prevalence in *G. pallidipes*. Numbers of tsetse flies from the screening results are represented at the base of each bar.

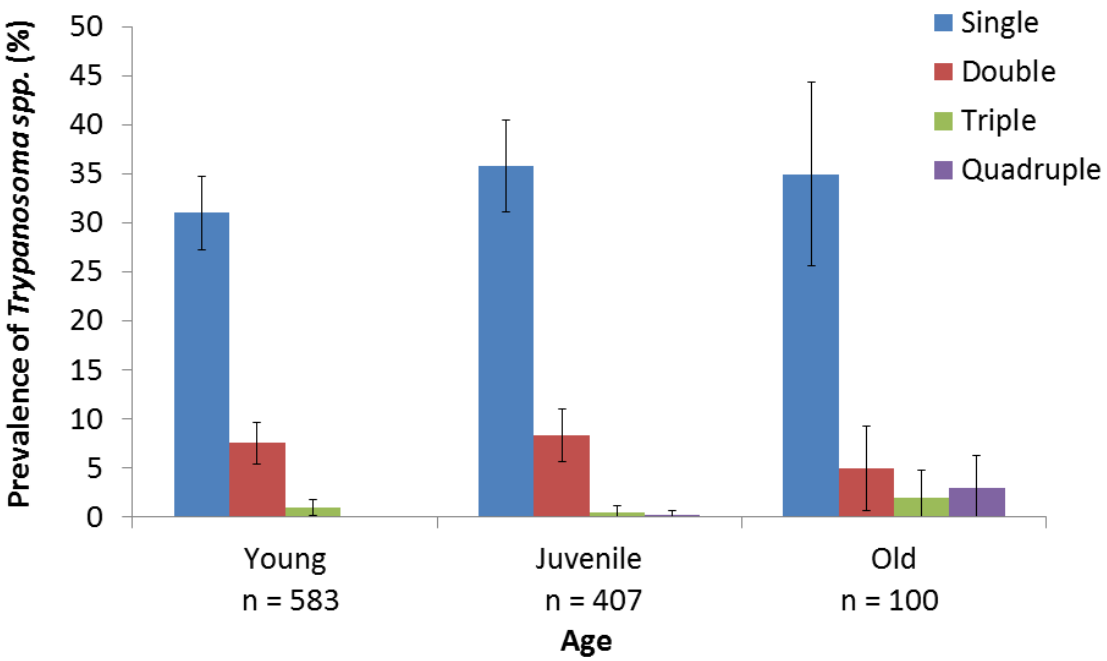



**Figure 2.27** Type of trypanosome infection (single, through quadruple) in relation to age of tsetse flies (young, juvenile and old).

Total tsetse fly numbers for each age are shown below the graphs. Dark blue bars indicate single infections; red bars indicate double infections; green bars indicate triple infections; purple bars indicate quadruple infections. 95% CI bars are indicated.

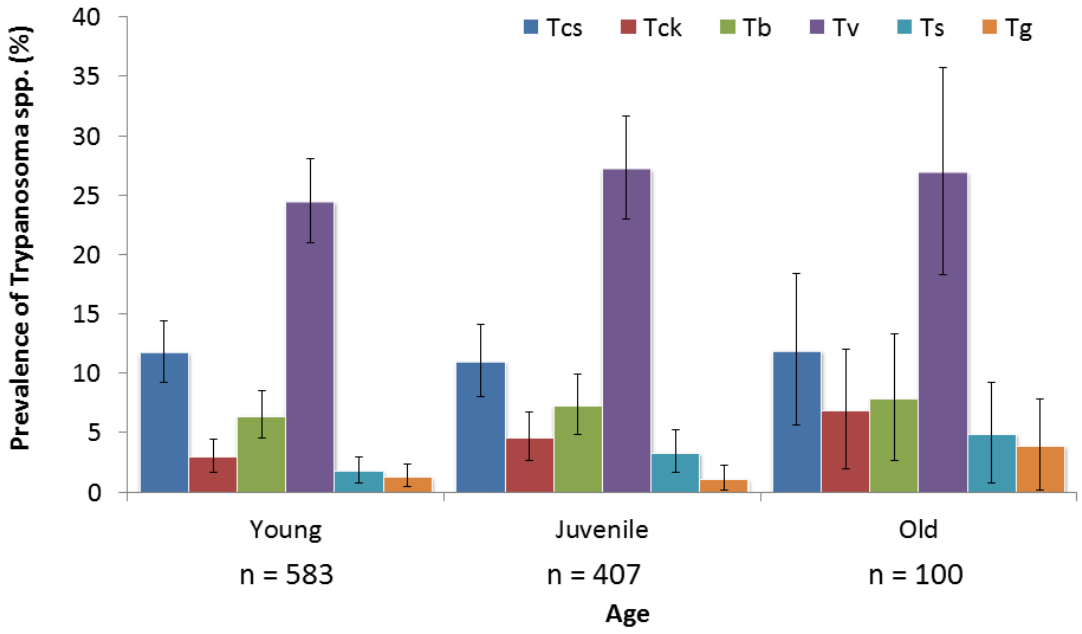



**Figure 2.28** Prevalence of individual species of trypanosomes in relation to age of tsetse flies.

Total tsetse fly numbers for each age are shown below the graphs. 95% CI bars are indicated. Tcs: *T. congolense* savannah; Tck: *T. congolesne* kilifi; Tb: *T. brucei*; Tv: *T. vivax*; Ts: *T. simiae*; and Tg: *T. godfreyi*.

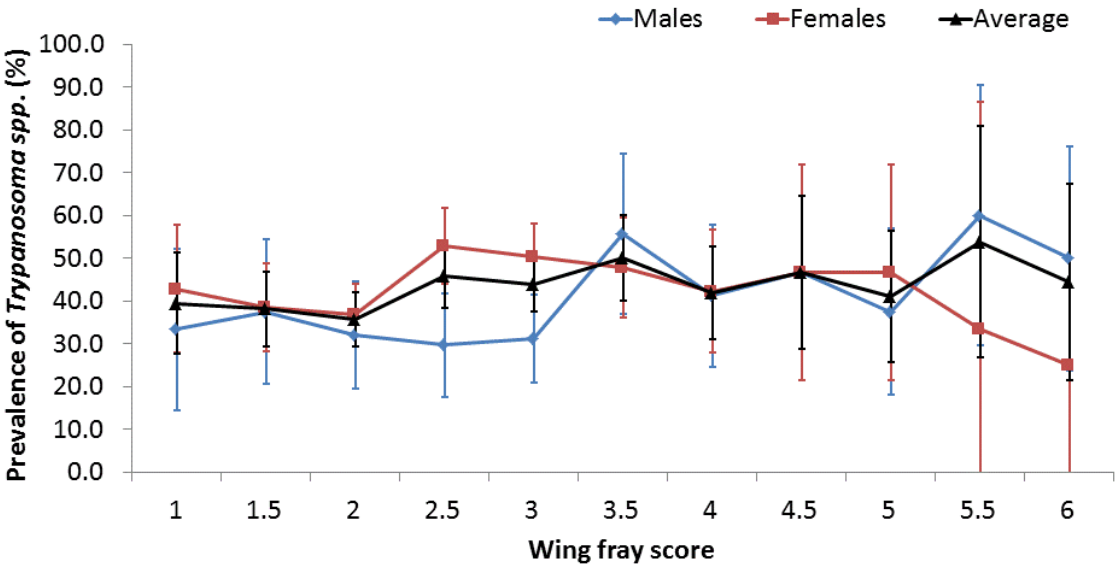



**Figure 2.29** Trypanosome prevalence in males and females in relation to age based on six levels of wing fray score.

Average prevalences in males and females showed slightly different patterns in relation to wing fray scores, tending towards lower prevalence with age and higher overall prevalence in females than in males. 95% CI bars are showed.

Plotting prevalence of trypanosomes by subpopulation, sex and age clearly demonstrated why it is important to consider interactions among multiple factors, as the overall trends of higher infections in older flies and in females were not consistent across subpopulations (Figure 2.30a).

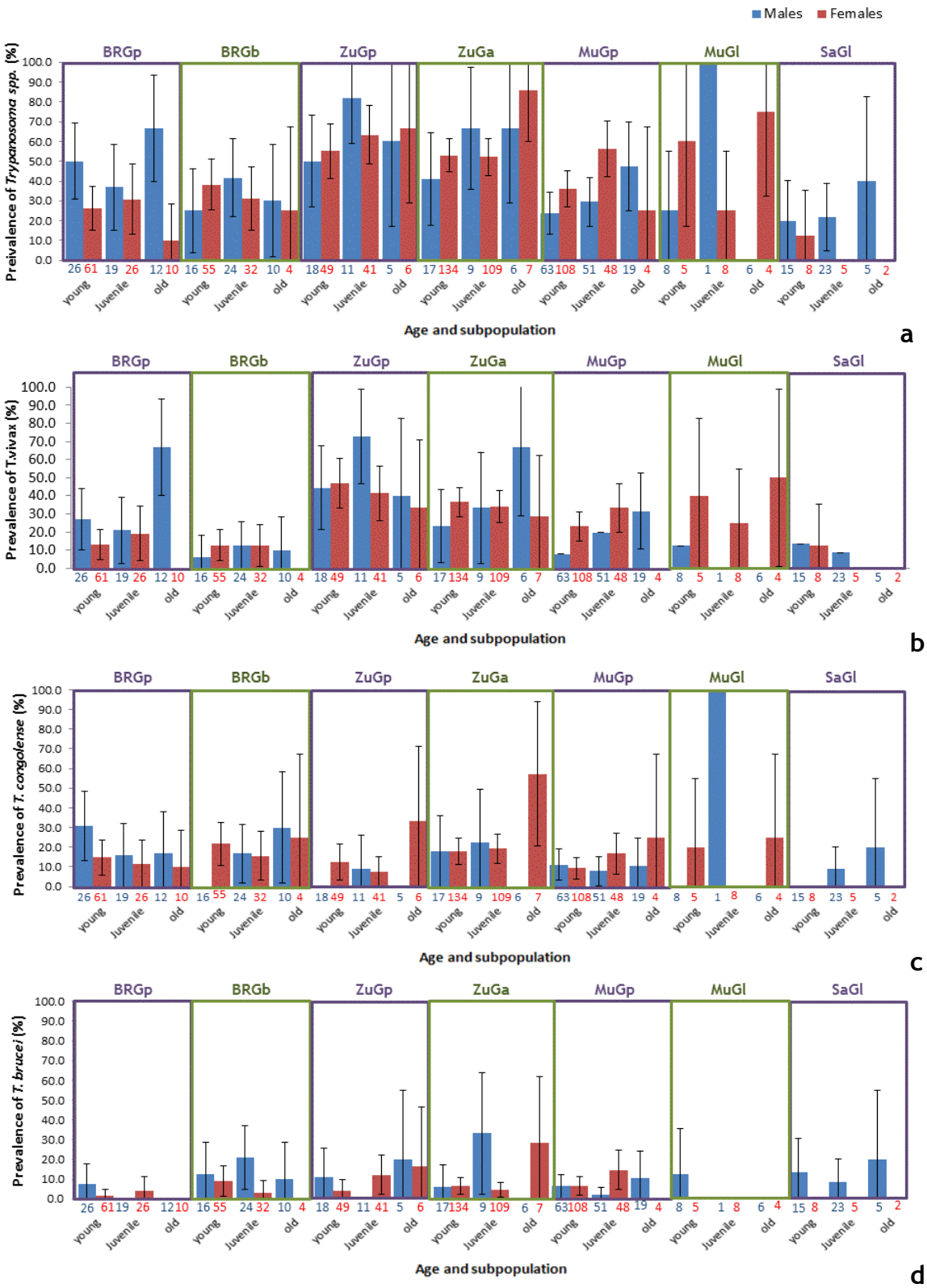



**Figure 2.30** Prevalence of trypanosomes in male and female tsetse flies with different ages in seven subpopulations

Prevalence of trypanosomes in young, juvenile and old *G. pallidipes* from Buffalo Ridge (BRGp): *G. brevipalpis* from Buffalo Ridge (BRGb); *G. pallidipes* from Zungu Luka (ZuGp); *G. austeni* from Zungu Luka (ZuGa); *G. pallidipes* from Mukinyo (MuGp); *G. longipennis* from Mukinyo (MuGl); and *G. longipennis* from Sampu (SaGl): (a)*Trypanosoma spp.* (b) *T. vivax*. (c) *T. congolense*. (d) *T. brucei*. 95% CI bars are showed. Numbers of total tsetse flies of each cathegory are shown at the base of each bar.

### 2.4.4 GLM analysis of association of *Trypanosoma spp*. with intrinsic factors of tsetse flies

Statistical analyses based on GLMs followed by post hoc tests confirmed the pattern-based results but clearly demonstrated differences between species of trypanosomes. For each species of trypanosome, the full model included subpopulation, age, sex, and their interactions; P-values are reported from likelihood ratio tests comparing this model to the best-fitting model. Presence of *T. congolense* (Model 1.1; Appendix A.10) was strongly associated with particular subpopulations ($\chi^2$ = 18.77, df = 1, P-value = 0.0046): probabilities of *T. congolense* positive tsetse flies from the Shimba Hill tended to be higher than tsetste flies from Nguruman. Among all four species of tsetse flies, *G. austeni* had the highest probability of *T. congolense* presence, followed by *G. brevipalpis, G. pallidipes* and *G. longipennis*. Although combinations between site and species affected variation of the probabilities for each subpopulation (Figure 2.31a), the subpopulation factor was not significant in post hoc tests (Appendix A.10), suggesting only weak differentiation (Table 2.10). There was no significant association of *T. brucei* presence (Model 1.2; Appendix A.11) with subpopulation, sex or age. For *T. vivax* (Model 1.3), presence was significantly associated with an interaction between subpopulation and sex ($\chi^2$ = 7.52, df = 21, P-value = 0.0061). Different predicted directions and values (Figure 2.31b) were driven by the significant differences of *T. vivax* prevalence in the BRGb, BRGp, MuGp and SaGl compared to the ZuGa, ZuGp subpopulations and 19 pairs of significant interactions between subpopulation and sex factors, in different directions (Appendix A.12).

**Table 2.10** Associations of *T. congolense*, *T. brucei* and *T. vivax* with intrinsic factors of tsetse flies based on GLMs.

For each primary model, the explanatory variables considered are shown, along with a description of the response variables. For each submodel (i.e. for each species of trypanosome considered as the response variable), indicated are the explanatory variables found in the best-fitting models, along with the degrees of freedom (df), log likelihood value (-logLik) Akaike's information criterion (AIC). See Appendix A.10 – Appendix A.12) for details of model selection based on likelihood ratio tests.

| Response variables | Explanatory variables of best fitting models | Full model | | | Best Fitting model | | |
|---|---|---|---|---|---|---|---|
| | | df | -logLik | AIC | df | -logLik | AIC |
| **Model 1** | Explanatory variables: subpopulation, sex and age<br>Response variables: *Trypanosoma spp*. status in all flies sampled (N = 1090) | | | | | | |
| 1.1 *T. congolense* | subpopulation | 28 | 424.8 | 905.6 | 7 | 436.4 | 886.7 |
| 1.2 *T. brucei* | none | 28 | 260.7 | 577.3 | 1 | 275.7 | 553.4 |
| 1.3 *T. vivax* | subpopulation * sex | 28 | 567.3 | 1190.5 | 14 | 577.1 | 1182.2 |

A * B means an interaction between A and B; red explanatory variables indicate significant differences also found in post hoc tests.

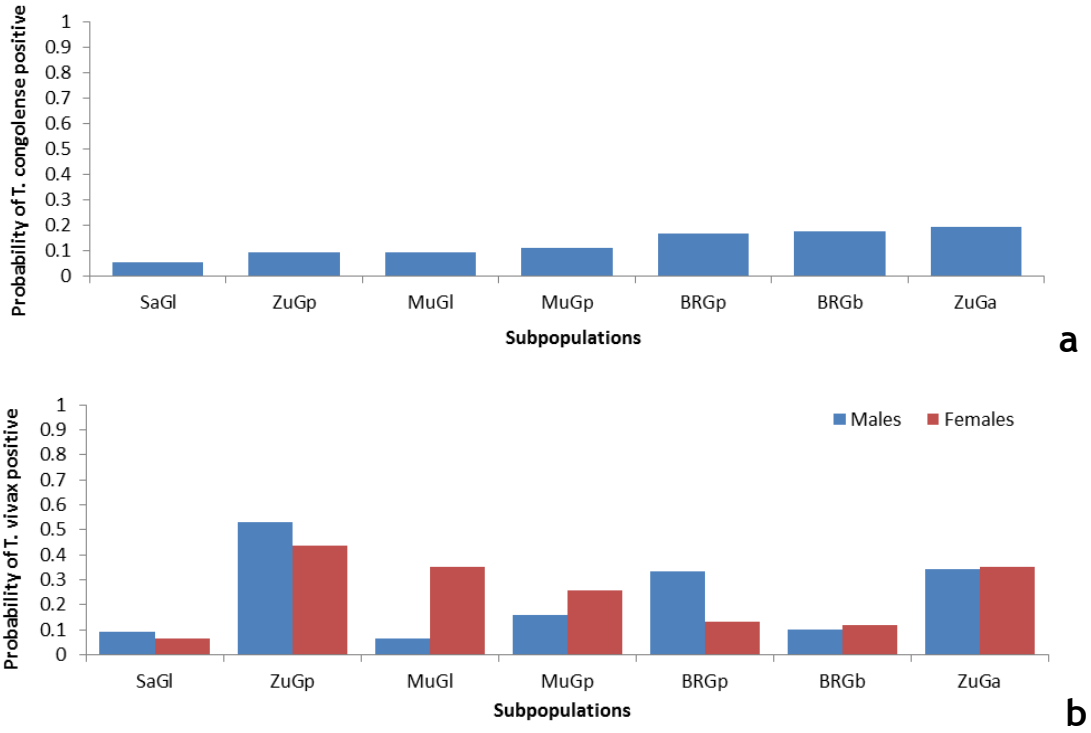



**Figure 2.31** Probabilities of trypanosome positive status in all tsetse flies based on the best fitting models.
Predicted probability of trypanosome status based on the factors that significantly explained variation for each species of trypanosome: (a) *T. congolense* with a subpopulation factor (b) *T. vivax* with an interaction between subpopulation and age. BRGp is *G. pallidipes* from Buffalo Ridge; BRGb is *G. brevipalpis* from Buffalo Ridge; ZuGp is *G. pallidipes* from Zungu Luka; ZuGa is *G. austeni* from Zungu Luka; MuGp is *G. pallidipes* from Mukinyo; MuGl is *G. longipennis* from Mukinyo; and SaGl is *G. longipennis* from Sampu.

When individual species of tsetse flies were analysed separately to more specifically test the effects of site (since only some species were found at multiple sites), there was additional variation in which factors explained variation in prevalence among the most abundant species of trypanosomes (*T. congolense*, *T. brucei* and *T. vivax*) (Table 2.11). For *G. austeni*, neither sex nor age was significantly associated with presence for any of the trypanosome species (Models 2.1 - 2.3; Appendices A.13 - A.15). For *G. brevipalpis*, although infection by *T. brucei* (Model 3.2; Appendix A.16) and *T. vivax* (Model 3.3; Appendix A.17) were not significantly associated with sex or age, there was a significant interaction between sex and age for presence of *T. congolense* ($\chi^2$ = 4.12, df = 3, P-value = 0.0424) (Model 3.1; Appendix A.18). Probabilities of detecting *T. congolense* in males tended to increase with age while for females it tended to decrease (Figure 2.32a). However, no significant differences were found in post hoc tests. For *G. longipennis,* presence of *T. brucei* (Model 4.2; Appendix A.19) was significantly associated with only sex ($\chi^2$ = 5.51, df = 1, P-value = 0.0190); males showed a higher probability of *T. brucei* presence than females (Figure 2.32b), but no significant difference was found between males and females in post hoc tests. Presence of *T. congolense* (Model 4.1; Appendix A.20) and *T. vivax* (Model 4.3; Appendix A.21) were not significantly associated with site, sex or age. For *G. pallidipes*, there were no significant factors related to *T. brucei* (Model 5.2; Appendix A.22) or *T. congolense* presence (Model 5.1; Appendix A.23) but *T. vivax* infection (Model 5.3; Appendix A.24) was associated with two interactions: between site and sex ($\chi^2$ = 7.52, df = 21, P-value = 0.0061); and between sex and age ($\chi^2$ = 12.96, df = 6, P-value = 0.0015). Males from Buffalo Ridge and Mukinyo showed an increasing probability of *T. vivax* presence with age whereas there was a slight decrease in females. However, a different range of probabilities were predicted for males and females (Figure 2.32c-1 and c-3). The associations of *T. vivax* with age in male and female *G. pallidipes* showed opposite directions (Figure 2.32c-2). From post hoc analysis of the *T. vivax* infection in *G. pallidipes*, all significant comparisons involving site included Zunga Luka, which showed a higher prevalence of *T. vivax* compared to the other sites in all cases (Appendix A.24). The significant interaction between sex and age was driven by significantly higher infections in old males compared to old females and in old males compared to young males.

**Table 2.11** Associations of *T. congolense*, *T. brucei* and *T. vivax* with intrinsic factors of *G. austeni*, *G. brevipalpis*, *G. longipennis* and *G. pallidipes* based on GLMs.

| Response variables | Explanatory variables of best fitting models | Full model | | | Best Fitting model | | |
|---|---|---|---|---|---|---|---|
| | | df | -logLik | AIC | df | -logLik | AIC |
| **Model 2** Explanatory variables: sex and age<br>Response variables: *Trypanosoma spp.* status in *G. austeni* (N = 282) | | | | | | | |
| 2.1 *T. congolense* | none | 4 | 136.2 | 280.5 | 1 | 137.7 | 277.4 |
| 2.2 *T. brucei* | none | 4 | 71.3 | 150.7 | 1 | 72.2 | 146.4 |
| 2.3 *T. vivax* | none | 4 | 181.0 | 370.0 | 1 | 182.8 | 367.5 |
| **Model 3** Explanatory variables: sex and age<br>Response variables: *Trypanosoma spp.* status in *G. brevipalpis* (N = 141) | | | | | | | |
| 3.1 *T. congolense* | sex * age | 4 | 62.9 | 133.8 | 4 | 62.9 | 133.8 |
| 3.2 *T. brucei* | none | 4 | 43.2 | 94.5 | 1 | 45.6 | 93.2 |
| 3.3 *T. vivax* | none | 4 | 49.6 | 107.2 | 1 | 49.9 | 101.8 |
| **Model 4** Explanatory variables: site, sex and age<br>Response variables: *Trypanosoma spp.* status in *G. longipennis* (N = 90) | | | | | | | |
| 4.1 *T. congolense* | none | 8 | 19.4 | 54.9 | 1 | 22.0 | 46.1 |
| 4.2 *T. brucei* | sex | 8 | 18.5 | 52.9 | 2 | 19.3 | 42.6 |
| 4.3 *T. vivax* | none | 8 | 29.9 | 75.9 | 1 | 35.3 | 72.7 |
| **Model 5** Explanatory variables: site, sex and age<br>Response variables: *Trypanosoma spp.* status in *G. pallidipes* (N = 577) | | | | | | | |
| 5.1 *T. congolense* | none | 12 | 206.2 | 436.4 | 1 | 213.2 | 428.5 |
| 5.2 *T. brucei* | none | 12 | 127.6 | 279.3 | 1 | 134.7 | 271.5 |
| 5.3 *T. vivax* | site * sex and sex * age | 12 | 306.7 | 637.5 | 8 | 307.9 | 631.8 |

A * B means an interaction between A and B; red explanatory variables indicate significant differences also found in post hoc tests.

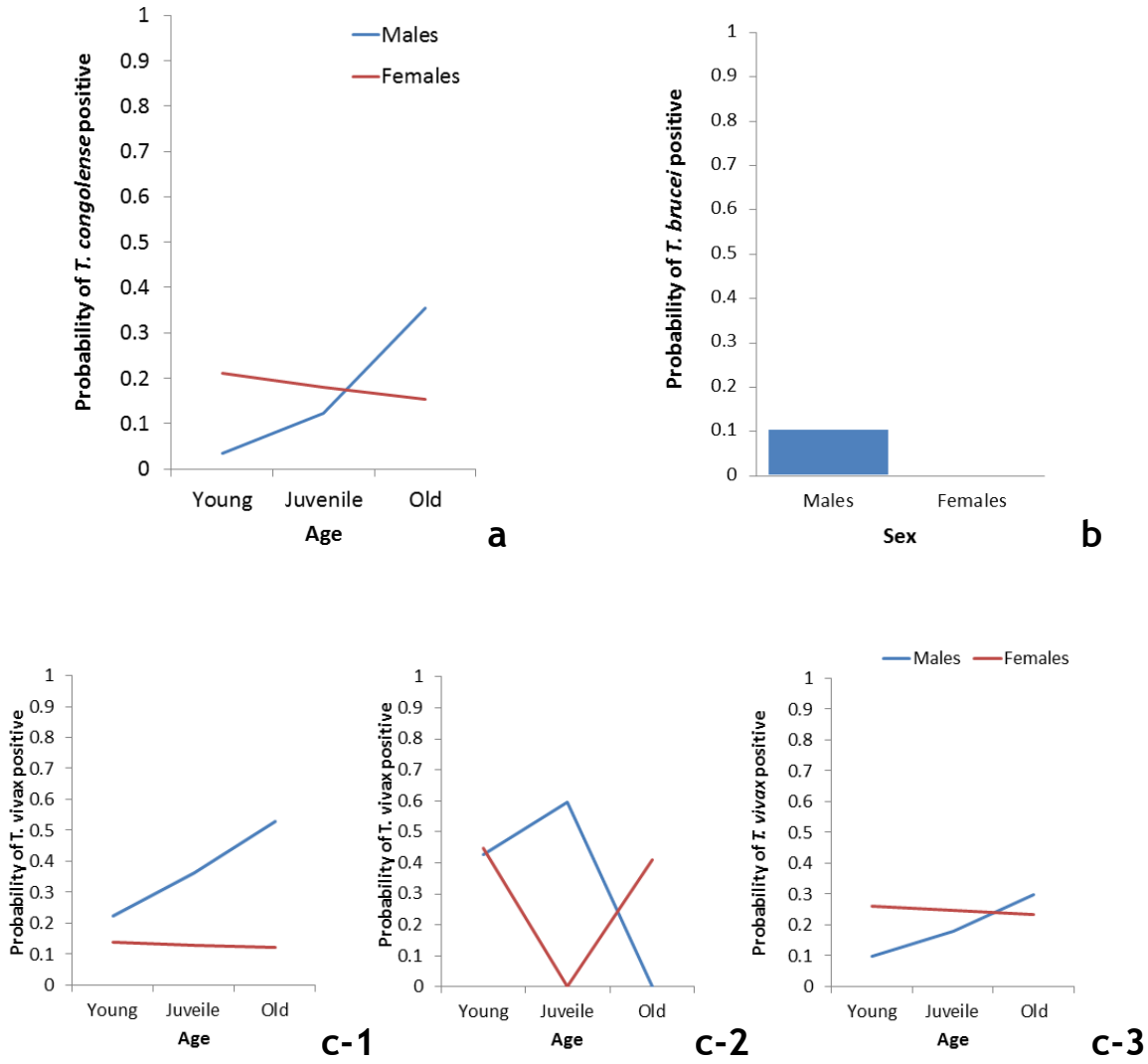



**Figure 2.32** Probabilities of trypanosome positive status in each *Glossina spp.* based on the best-fitting models
Predicted probability of trypanosome status in the final models for: (a) *T. congolense* in *G. brevipalpis;* (b) *T. brucei in G. longipennis*; and (c) *T. vivax* in *G. pallidipes* from (c-1) Buffalo Ridge, (c-2) Zungu Luka, and (c-3) Mukinyo).

## 2.5 Discussion

Overall, my results suggest extensive differences in the prevalence of different species of trypanosomes in different species of tsetse flies sampled from two different habitat regions within southeastern Kenya. Although focusing on individual species of trypanosomes in individual species of tsetse flies reduced the power of statistical tests for associations with particular characteristics of the tsetse flies, sex, age and sampling site all influenced these associations. The strength of these associations varied by species of tsetse but most strongly by sampling location.

### 2.5.1 Tsetse fly distribution and characteristics

There was extensive variation among sampling sites in the presence and relative abundance of particular species of tsetse fly. *Glossina pallidipes* was found in both geographic regions sampled (the Shimba Hills and Nguruman) and was found at three of the four sites sampled, but all other species were only found within a single region (*G. longipennis* was found at two sites but both were within the Nguruman region). This finding is consistent with previous reports that *G. pallidipes* and *G. longipennis* are distributed in areas around Nguruman while *G. pallidipes*, *G. austeni* and *G. brevipalpis* are found in areas around the Shimba Hills (Mbahin *et al.*, 2013, Cecchi *et al.*, 2015). This emphasizes that differences between tsetse species in ecological preferences or requirements is an important factor for understanding variation of tsetse fly distribution (Leak, 1998, Pollock, 1982).

Variation in the sex and age of trapped tsetse flies was found across the sites, which should reflect differences in relative abundance. All testes flies were trapped in the rainy season (June – August), when tsetse flies are most reproductively active (Mamoudou *et al.*, 2008), so emerging flies (young or juvenile tsetse flies) could have been expected to be present at high numbers. There could be a bias in trapping efficiency due to sex because the traps were modified to include blue cloth, which is predicted to increase attractiveness to resting females (Brightwell *et al.*, 1991). I could not rule out the possibility that abundance of males is underestimated in the data. This underestimation would be consistent throughout the data set and thus not cause any systematic bias. It means abundance of males is not able to be estimated accurately, but still robustly detect changes in the sex ratio between sites and species. Interestingly, the tsetse species for which males were at higher abundance (*G. longipennis*) also showed lower overall numbers of flies, which could reflect lower reproductive output due to lower numbers of females.

Temperature has been found to be a critical variable not only for the presence of tsetse flies but also for the reliability of sampling using traps (Pollock, 1982). The range of best suitable temperature for tsetse flies is around 25 °C – 26ºC while lethal temperatures are around 40°C (Nash, 1936). Too high or low temperatures influence, reproduction, puparial development (Harley, 1968) and

host seeking activity (Pollock, 1982). Previous reports have found that the number of trapped flies increases when the temperature is lower than a threshold of 34°C compared to higher temperatures (Leak, 1998). Different ranges of minimum and maximum critical temperatures have also been documented for different populations of tsetse flies. In Kenya, the critical temperatures have been found to be approximately 20 - 21°C and 43.9- 45.0°C (Terblanche *et al.*, 2008). Average temperatures of the Shimba Hills region were suitable for tsetse-fly activities while the number of tsetse flies from the Nguruman regions might be reduced due to lower temperatures at the time of sampling (Table 2.12).

Relative humidity, which is important for the development of immature flies, could also explain some of the differences in distribution or sampling. Pollock (1982) noted that dry conditions possibly limit distribution of some tsetse flies because they require a suitably humid soil in their puparial stage, while they are developing under the ground. Relative humidity also influences activity of tsetse flies (Leak, 1998), which could affect the number of collected flies, rather than their actual relative abundance: activity levels of tsetse flies under dry conditions is higher than in wet air (Bursell, 1957). However, since more flies were collected in the Shimba Hills despite less arid conditions (Table 2.12) and lower numbers of traps, this potential bias was not reflected in the sample sizes from each region. Moreover, previous studies have found that seasonal change does not influence activity and biases in the number of collected flies (Brightwell *et al.*, 1992).

The sites sampled in this study differed in some important factors that might affect tsetse fly presence due to differences in biology rather than sampling effects (Table 2.12). *Glossina austeni, G. brevipalpis, G.longipennis* and *G.pallidipes*, have been found in Kenya (Clausen *et al.*, 1998) and showed variation of density in different tempterature and vegetation type (Rogers and Randolph, 1993). *Glossina pallidipes* was determined as the predominant species of tsetse flies among the sampling sites. This species is generally able to adapt itself to inhabit both evergreen or savannah forest (Laird, 1977) and is widely found in eastern and southern Africa (Ford, 1971). Nevertheless, *G. pallidipes* was collected from Mukinyo, but not Sampu, which are around 6 km apart and have broadly similar environments. *Glossina brevipalpis* (classified in the Fusca

group) prefers to inhabit tropical rain forest (Leak, 1998), where there is high humidity similar to the thicket vegetation type found in Buffalo Ridge (Table 2.12). However, fewer *G. brevipalpis* were trapped than *G. pallidipes* at this site. *Glossina longipennis* is a special species of tsetse flies in the Fusca group in that it is distributed in dry regions (Laird, 1977), which could explain why it was only found in the Nguruman region, which has lower humidity compared to the Shimba Hills. *Glossina austeni* (classified in the Morsitans group) was found only at Zungu Luka, which is a woodland area appropriate for *G. austeni* living (Laird, 1977).

The two sampling regions are also characterized by differences in potential host species communities, which could influence the relative abundance of particular species of tsetse flies. There are more wildlife species present in the Shimba Hills than Nguruman, which could mean that there are more overall hosts available to support larger communities of tsetse flies. Zungu Luka is woodlanded grassland and close to a rural area with large numbers of livestock (Appendix A.2 – Appendix A.8) and this was the only site where *G. austeni* were trapped. Generally, *G. austeni* is thought to feed on bushbucks, warthogs, bushpigs and buffalo, which are available in many countries in Africa (Clausen *et al.*, 1998) but it is possible that they prefer livestock. Testing of this hypothesis would require detailed blood meal analysis and host feeding preference trials. Many lifestock (cattle, buffalo, cattle, sheep, goat, pig and poultry) were estimated in both collected sites in Nguruman (Appendix A.2 – Appendix A.8), but the number of trapped tsetse flies was lower than in the Shimba Hills. Ecology of the different geographic regions might be a domiant factor of tsetse-fly distribution. In addition, decreased frequency of feeding has been found to cause increased mortality of lava in the uterus of tsetse flies (Leak, 1998), so lack of preferred hosts might affect the relative recruitment of different tsetse species. In conclusion, available hosts were possibly associated with species distribution of tsetse flies and the relative population sizes of each species of tsetse found.

**Table 2.12** Factors differing between the Shimba Hills and Nguruman that might affect distribution of tsetse flies.

| Factors | The Shimba Hills | | Nguruman[a] | |
|---|---|---|---|---|
| | Buffalo Ridge | Zungu Luka | Mukinyo | Sampu |
| **Type of vegetation** | Thicket natural forest[b] | Woodlanded grassland[b] | Acacia woodland, scattered bushes and open grasslands (Tarimo-Nesbitt *et al.*, 1999), surrounded with open savannah (Brightwell *et al.*, 1997). | |
| **Temperature (average)[c]** | 22 - 30°C (26°C) | | 15 - 27°C (20°C) | |
| **Relative humidity[c]** | 54 - 94% | | 41 - 88% | |
| **Available hosts** | cattle, goats, ducks, chickens (Njenga *et al.*, 2011) and wildlife[d] | | cattle, buffalo, goats,sheep, pigs,pourtry (Food Agriculture Organization) and wildlife[e] | |

[a] Nguruman had different species of tsetse flies between the two sites; Glossina pallidipes was collected from only Mukinyo.

[b] The Shimba Hills is tropical evergreen seasonal lowland rain forest (Höft and Höft, 1995). Typing the vegetation was accomplished by comparing of satellite pictures of sampling sites (Figure 2.2) from Ciosi with the diagrams that were sketched by Leak (1998) (Appendix A.1) and description of vegetation from (Sutton *et al.*, 2002).

[c] estimated from temperature and humidity in July (for the Shimba Hills) and August (for Nguruman) in 2012 from https://www.wunderground.com/DailyHistory.html. The average temperature is indicated in parentheses.

[d] Including elephants, antelope, African buffalo, warthogs, giraffes, hyenas, bush pigs, leopards, genets, waterbuck, bushbuck, colobus monkeys, duikers, galagos, monkeys, serval cats, civet cats, monkeys and sand shrews (Mbahin *et al.*, 2013) and some other hosts were determined from G. pallidipes using blood meal analysis in chapter 4.

[e] African buffalo, warthogs, giraffes, Kirk's dik-diks, bushbuck, waterbuck, hartebeest, wildebeest, impala, Grant's gazelles, and zebra (Brightwell *et al.*, 1992) and some other hosts were determined from G. pallidipes using blood meal analysis in chapter 4.

### 2.5.2 PCR-based *Trypanosoma spp*. identification

Two issues concerning the use of a PCR-based detection methodology as employed in this thesis need to be addressed. The first is that PCR does not detect trypanosomes, it detects trypanosome DNA. Thus, it is possible that DNA from a dead trypanosome that has not been degraded at death could be amplified. This could result in false positives, leading to higher than "true" prevalence values. I have reduced the risk of this problem by analysing head plus proboscis samples, which contain the foreguts. On ingestion of an infected blood meal, trypanosomes are thought to move through the foregut rapidly into the midgut (Chinery, 1965). The classical alternative to PCR detection is traditional microscopy but this is likely to lead to false negatives and lower than true prevalence (Wamwiri *et al.*, 2013), and identification of different trypanosome species can be problematic. The second issue is that presence of trypanosomes in head plus proboscis samples does not indicate whether they are present and could be mechanically transmitted or whether full development of the life cycle has occurred. Therefore, what I have assessed is presence or absence of the trypanosomes, rather than true rates of infection but it is informative about differences in the species of trypanosomes taken up by different species of tsetse.

Comparison of different PCR cycling conditions with different tsetse tissues suggested that amplification of trypanosomes is very sensitive to the particular methods used and which body parts are used for amplification. The PCR conditions described by Njiru *et al*. (2005) to amplify the ITS-1 fragment of trypanosomes from blood samples of cattle in Kenya, tsetse flies and DNA references were more sensitive for trypanosome screening in tsetse flies from Kenya than the conditions of Isaac *et al*. (2013), which were specifically developed to identify trypanosomes from tsetse flies in Nigeria. There were more PCR cycles, a lower annealing temperature and a longer elongation time in the former, which could have resulted in higher sensitivity for detection of low levels of trypanosome infection. More amplification products were found using head plus proboscis compared to abdomens. While this might reflect differences in the relative abundance of trypanosomes present and include transient parasites that would not become established and develop in the flies (Leak, 1998), there also could be some methodological biases. Moreover, some

inhibitors, such as heme (Akane *et al.*, 1994), haemoglobin, lactoferrin (Al-Soud and Radstrom, 2001) and immunoglobulin G (Al-Soud *et al.*, 2000) present in blood meals in the abdomen of tsetse flies could reduce the potential for amplification of trypanosome positive samples. Thus, DNA in the tissues used for screening and substances that could affect PCR amplification should be considered before developing or modifying screening methods. According to the pilot experiment, the method developed for trypanosome detection in Nigeria (Isaac *et al.*, 2013) was not effective for trypanosome screening in tsetse flies from Kenya. This could imply that a suitable method of trypanosome screening in one geographic location might not be appropriate for others. I would suggest that all screening methods should be evaluated for the most reliable screening tissue and PCR conditions at a local scale.

Using *Taq* DNA Polymerase from Thermo Scientific Invitrogen and Dream*Taq* Green PCR Master Mix from Thermo Scientific showed different results for amplification of the ITS-1 fragments, especially for *T. vivax* (Panton, 2015). The different PCR mixes differed greatly in chemical composition (Table 2.13), with the Thermo Scientific Taq DNA polymerase being the most complex and showing the most reliable amplification of PCR prodcuts. $(NH_4)_2SO_4$, BSA and EDTA could be important factors that increase the efficiency of DNA amplification in the PCR reaction. In addition, KCl, Tween® 20, Nonidet® P40 and stabilizers, which are added to the storage buffer for *Taq* from Thermo Scientific, preserve enzyme function. For amplification of small amounts of DNA where inhibitors may be present, the more complex reaction mixes might therefore have substantial benefits.

**Table 2.13** Chemical components of storage buffer and reaction buffer.

The concentration of reagents is indicated. No indicates that that component was not present. *Taq* DNA Polymerase (Thermo Scientific, catalog number AB-0192/A) with 10x Custom PCR Master mix (Thermo Scientific, catalog number SM-0005); *Taq* DNA Polymerase (Invitrogen, catalog number 18038-034); and Dream*Taq* Green PCR Master Mix (Thermo Scientific, catalog number K1082)

| Chemical components | *Taq* DNA Polymerase (Thermo Scientific) | *Taq* DNA Polymerase (Invitrogen) | Dream*Taq* Green PCR Master Mix (Thermo Scientific) |
|---|---|---|---|
| Storage buffer of Taq DNA polymerase | | | |
| KCl | 100 mM | no | no |
| Tween® 20 | 0.5 % | no | no |
| Nonidet® P40 | 0.5% | no | no |
| Stabilizers | Yes | no | no |
| Reaction buffer | | | |
| dNTP mixture | 1.0 mM | 0.2 mM | 0.4 mM |
| $MgCl_2$ | 4.5 mM | 1.5 mM | 4.0 mM |
| Tris-HCl | 45 mM, pH 8.8 | 200 mM, pH 8.4 | no |
| $(NH_4)_2SO_4$ | 11 mM | no | no |
| BSA | 0.113 mg/ml | no | no |
| EDTA | 4.4 µM | no | no |
| KCl | No | 500 mM | no |
| Dream Taq buffer | No | no | yes |

no = not included in the product information.

Using sizes of PCR products from the ITS-1 primers to determine species of trypanosome infection was not well correlated with results from the TCS and TCK specific primers and showed high correspondence with TBR primers in the ZuGp subpopulation but not in the other subpopulations. Since not all products were sequenced, it is difficult to assess whether species-specific or general primers are more trustworthy based only on these comparisons. Ciosi *et al.* (unpublished; Appendix A.8) experimentally infected *G. pallidipes* individuals with *T. congolense* savannah and monitored the first date and how long trypanosomes were detected in head plus proboscis (HP) and abdomen (AB) samples for 9 days post infection. At day 0, detection of parasites was very low (less than 40%) in HP samples using the ITS-1 CF and BR primers but 100% infection was indicated using the TCS primers. No parasites were detected using the ITS-1 primers two days after infection but TCS indicated apparent infection for 8 days. The TCS primers also showed higher sensitivity than ITS-1 primers for abdomen tissues but infection rates started off at 100% and declined only

gradually using both primer sets. An important difference was that infection levels using TCS primers remained relatively high (60%) at day 11, when ITS-1 primers indicated that the parasites had been cleared from abdomens. This experiment suggests that TCS could be too sensitive because it may detect very small numbers of parasites, which could represent residual parasites from the experimental exposure rather than parasites that have become established. The relatively high proportion of ITS-1 negative samples that were TCS positive in my study could thus represent transient parasites obtained from blood meals that would not give rise to infections that could be transmitted to other hosts. Thus, the ITS-1 results, while more conservative, could be more informative about relative vector competence for this subspecies of trypanosome.

Although the ITS-1 and TBR primers were designed for *T. brucei*, sizes of PCR products from the two primers could not identify the subspecies of *T. brucei*, and the TBR primers appeared to have low sensitivity in some sampling areas, based on comparison with the ITS-1 primers (Table 2.9): TBR positive amplifications were higher in the coastal region (the Shimba Hills) (Figure 2.1) but this pattern was not found for ITS-1. The most appropriate PCR primers for *T. brucei* screening could be different in each site of sampling if different strains of trypanosomes are present at different sites. Sequencing results based on ITS-1 confirmed that trypanosomes were amplified in the reaction products sequenced but the identity of the parasites was not confirmed because there is not enough sequence polymorphism to distinguish between the known subspecies. There have been no reports of *T. b. gambiense* or *T. b. rhodesiense* in the sites that tsetse flies were sampled for this study so *T. brucei brucei* is what was expected but this could be tested using more specific primers, such as for amplification of *T. brucei gambiense* specific glycoprotein (TgsGP) (Radwanska *et al.*, 2002b) and multiplex PCR for identification of *T. b. brucei* and *T. b. rhodesiense* (Picozzi *et al.*, 2008). Alternatively, sequencing a longer region of the ribosomal RNA array could be informative to confirm subspecies identification and determine whether the differences found among populations were due to the presence of different genotypes of *T. brucei* or amplification of another species using the ITS-1 primers. For example, screening using nested PCR conditions for primers developed by Cox *et al*. (2005) to amplify a product starting in a conserved region of the 18S gene and extending to the 28S gene, which spans both the ITS-

1 and ITS-2 regions, revealed a higher than expected diversity of trypanosomes in wildlife hosts in the Serengeti ecosystem in Tanzania and the Luangwa Valley in Zambia (Auty *et al.*, 2012). Unfortunately, not enough studies have yet used these primers to provide an adequate database for reliable identification of new species (Elizabeth Panton, unpublished honour's thesis) but I recommend that future studies focused on documenting the diversity of trypanosomes present within a geographic region amplify and sequence this more variable region of the rDNA array.

In my study, tissues for screening, primers, PCR mixtures and PCR cycles all influenced the screening results for the presence of trypanosomes. It is important to choose a tissue that has the highest opportunity to find an infective stage of *Trypanosoma spp*. and the least inhibitors of PCR amplification but this was not completely clear from my results. Both tissue types resulted in positive amplifications using ITS-1 primers but not always from the same individual flies. Given the variation in amplification among sites in some cases, primers should be tested and products sequenced in the local geographic regions and for the particular species of tsetse fly targeted in a particular study. The amplicons should also have an appropriate size. Too short PCR products are easy to misdiagnosis with primer-dimers and can be difficult to sequence. Too long PCR products can also be difficult to sequence and cloning may be required to sequence through long products or when mixed infections are present, as was widely apparent in my study. This approach could be beneficial in a survey of trypanosomes in wild hosts but only one sample of each trypanosome species was sequenced from each variant so more extensive screening of clones might be necessary to distinguish new species or genotypes from cloning errors. Although I attempted to sequence products from both species-specific and ITS-1 primers, weak amplification and poor sequencing results prohibited drawing conclusions about whether species-specific or general primers produced more reliable results. Since both species-specific and ITS-1 primers amplify regions with tandem repeats that are present in many copies, the quality of their PCR product sequencing was poor. I recommend that PCR products of trypanosome-specific primers (from both expected and unexpected bands) also should be sequenced to confirm the amplification results.

### 2.5.3 Association of *Trypanosoma spp*. presence with intrinsic factors in tsetse flies

Trypanosome prevalence was much higher than previous reports (Nthiwa *et al*., 2015, Wamwiri *et al*., 2013), possibly due to several factors. First, the tissue samples I used (head plus proboscis part) are expected to have lower levels of PCR inhibitory factors such as heme (Akane *et al*., 1994) and haemoglobin (Al-Soud and Radstrom, 2001) than using whole flies for trypanosome screening. Second, some other studies investigated trypanosomes from abdomen parts so *T. vivax* and some *T. congolense* infected flies could have been misdiagnosed because developments of the two pathogenic species occure in the proboscis part. Third, the samples were trapped in a time period when trypanosomes in animals has been found to be higher than at other times of the year (Majekodunmi *et al*., 2013); thus the opportunity of trypanosome exposure might be high. Prevalence of single trypanosome infections was higher than mixed species infections, which were found with a relatively low rate but increased with age of tsetse flies. Tsetse vectors are unable to clear infections in their lifetime, which would mean that mixed infections should be highest in older flies (Figure 2.33b). However, my results still show that single infections are much more common than mixed infections. *Glossina m. morsitans* experimentally infected with *T. congolense* showed higher mortality rates and reduced longevity compared with uninfected control flies (Nitcheman, 1987). Trypanosome infection in salivary glands has been found to significantly reduce tsetse survival of *G. m. morsitans*, however, midgut infection had little or no effect (Maudlin *et al*., 1998). A relationship between *T. congolense* infection and mortality of *G. pallidipes* has also been suggested based on age of tsetse flies (Woolhouse and Hargrove, 1998, Woolhouse *et al*., 1993). Thus, lower mixed infections would be found in old flies because highly infected flies would have died. Consistent with this, Ciosi *et al*. (unpublished) found that 3/50 tsetse flies died after a few days of *T. congolense* savannah inoculation. Thus, in my study, the biological vectors could have died from high infections or mixed species infections or could have been too weak to seek hosts so they were not sampled by the trapping method. This could have led to the apparently lower rate of mixed infection in old flies than juvenile flies and fewer old flies collected than young flies. However, this assumption should be tested by experimental infection of tsetse flies with specific species of trypanosome and determination of mortality rates in relation

to trypanosome loads in single and mixed infections. A few tsetse flies in this study were positive for quadruple infection with trypanosomes, which was also reported in a previous investigation of *G. pallidipes* and *G. longipennis* from Kenya (Njiru *et al.*, 2004). However, investigation of trypanosome infection in *G. brevipalpis*, *G. pallidipes*, *G. swynnertoni*, and *G. m. morsitans* from Tanzania showed only double and triple mixed trypanosome infections (Adams *et al.*, 2006, Malele *et al.*, 2003). Motloang *et al.* (2012) reported that the rate of *T. congolense* infection in *G. austeni* was higher than *G. brevipalpis* when they fed on infected hosts (Motloang *et al.*, 2012). In addition, they found that cattle (100%) were infected with *T. congolense* when infected *G. austeni* fed on them, but no infection in cattle, which had been fed on by infected *G. brevipalpis*. Tsetse flies of each geographic region thus could have variation in susceptibility levels to trypanosomes, or the strains of trypanosomes in different geographic regions could have variation in virulence or competitiveness with other species.

The species of trypanosomes, species of tsetse flies and available hosts in each site could affect relative risk of infection by flies. Firstly, different pathways of trypanosome development in each parasitic species could be associated with differences in infection rate of the parasites (Peacock *et al.*, 2012b, Roditi and Lehane, 2008). *Trypanosoma vivax* develops exclusively in mouth parts (Osóriol *et al.*, 2008), while establishment and maturation of *T. congolense* and *T. brucei* occur in other parts of the upper gastrointestinal tract (Peacock *et al.*, 2012a). *Trypanosoma congolense* and *T. brucei* initially establish for infection in the midgut with subsequent maturation of *T. congolense* occurring in mouth parts and *T. brucei* in salivary glands (Peacock *et al.*, 2012a). Secondly, tsetse flies that are biological vectors in the *Morsitans* and Fusca groups are good biological hosts for *T. vivax* and *T. congolense* but they are not for *T. brucei* (Leak, 1998). Thirdly, many species of wildlife and farm animals are susceptible to *T. vivax, T. congolense* and *T. brucei* infection while hosts of *T. simiae* and *T. godfreyi* are limited to *Suidae*. The GLM analyses supported the conclusion that subpopulation (classified by species of tsetse flies and sampling site) was significantly associated with presence of *T. congolense*, but this varied by sex for *T. vivax* and none of the factors considered were significantly associated with presence of *T. brucei*. The absence of associations in the latter could be due to its low overall prevalence, reducing the power to detect associations. At one site

(Zungu Luka), where more than one species of tsetse was present, prevalence of trypanosomes was higher in *G. pallidipes* than *G. austeni*. Since the lifespan of *G. pallidipes* females has been found to be longer than that of *G. austeni* (Wamwiri *et al.*, 2013) and a long lifespan of tsetse flies could increase the risk of trypanosome exposure and infection, this could contribute to the observed differences. Other differences among sites, such as relative prevalence of the secondary endosymbiont that has been associated with trypanosome prevalence (Wamwiri *et al.*, 2013, Farikou *et al.*, 2010a, Geiger *et al.*, 2007) and host blood meal feeding practices, would also be interesting to consider (see chapters 3 and 4, respectively).

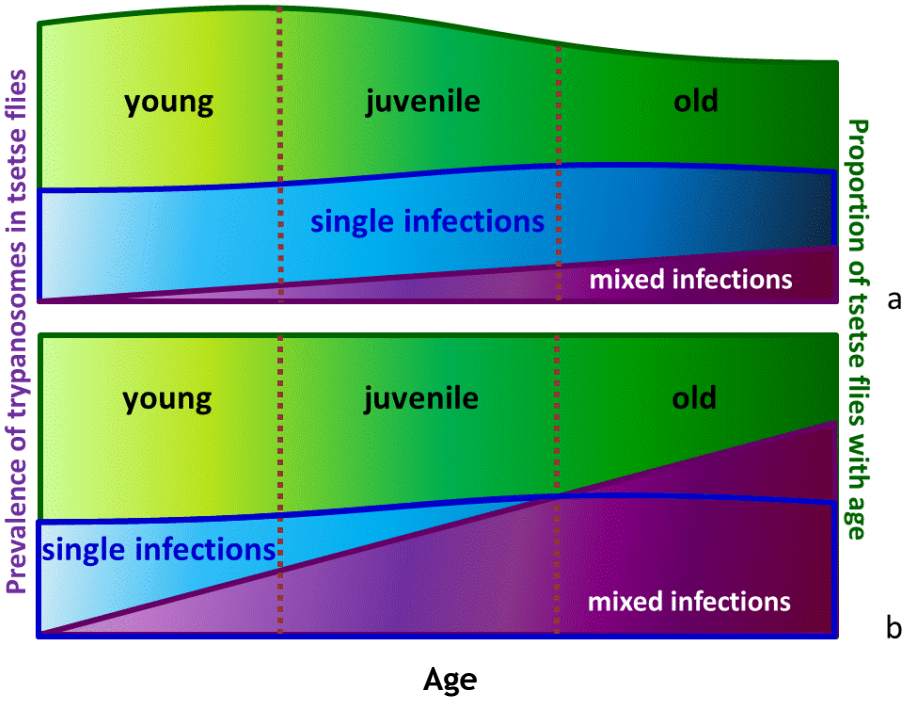



**Figure 2.33** Cartoon demonstration of trends of single and mixed trypanosome infections with age of tsetse flies.

(a) Representation of observed patterns in this study. (b) Predictions based on the theory that pathogens do not affect the health of the vector and vectors cannot clear infections. Green areas indicate the proportion of flies that are not infected with trypanosomes in young, juvenile and old flies; blue areas indicate the proportion of flies with single trypanosome infections; and purple areas indicate the proportion with mixed trypanosome infections.

The highest trypanosome prevalence was found in *G. austeni*, followed by *G. pallidipes*, *G. brevipalpis* and *G. longipennis* but this could have been influenced by sex, since there was a stronger bias towards females in the ZuGa

subpopulation compared to the others. Prevalence of females in the ZuGa subpopulation was the highest when compared with prevalence of males and females for the other subpopulations (Appendix A.9). Although this subpopulation showed a "typical" age distribution centred around young and juveniles, this also might have influenced infection status because of the female bias; females showed the highest levels of infection within this age class and decreasing infection with age, whereas males showed increasing infection rates with age. Since *G. austeni* was only found at Zungu Luka, it is not possible to separate which factor (site, species of tsetse or their age or sex) is most important for trypanosome susceptibility but it emphasises how misleading interpretation of results could be if considered in isolation. Differences among sites within species could also affect interpretation of results. Although *G. austeni* on average showed higher prevalence of trypanosomes than *G. pallidipes*, this was driven by low prevalence in this species at two sites (Buffalo Ridge and Mukinyo): at Zungu Luka *G. pallidipes* actually showed higher prevalence than *G. austeni*, particularly for *T. vivax*. Similarly, *G. brevipalpis* was the only species for which a higher infection was found for *T. congolense* savannah (TCS) than *T. vivax*, but it was found at only a single site (Buffalo Ridge), where relatively high infection with TCS was also found in the other species present (*G. pallidipes*). In general, across all subpopulations, the apparent effects of site and tsetse species were confounded by differences in sex and age distributions and potential sampling biases related to the species of tsetse found at each site (Figure 2.10). However, these results are consistent with a previous report that tsetse flies in the *Morsitans* group (*G. austeni* and *G. pallidipes*, which were the majority population in this study) are better hosts for trypanosomes than the Fusca group (*G. brevipalpis* and *G. longipennis*) (Leak, 1998). Another study reported that *G. austeni* were more susceptible to trypanosome maturation than *G. brevipalpis* (Motloang *et al.*, 2012). Consistent with this, this investigation found that mixed trypanosome infections were higher in *G. austeni* from Zungu Luka than *G. pallidipes* from the same site. Zungu Luka had both wildlife and domestic animals present so it is possible that *G. austeni* also feed on domestic animals. A previous study by Clausen *et al.* (1998) found that *G. austeni* from Muhaka Forest on the coast of Kenya fed on hares, dik-diks, bushpigs, warthogs, duiker, colobus monkeys, cattle, goats, humans and monitor lizards (Clausen *et al.*, 1998). However, blood meal analysis

should be further studied to determine whether host-feeding behaviour influences single or multi-species infection with trypanosomes.

The GLM analyses were used in attempt to unravel which factors or interactions between factors were most important in explaining variation in infection of tsetse flies with particular species of trypanosomes but results suggests that there might be too many factors interacting to resolve all significant relationships. A recent study was conducted in Nigeria comparing the influences of site, species and sex of three species of flies (*G. palpalis palpalis*, *G. tachinoides* and *G. morsitans submorsitans*) on the presence of four species of trypanosomes (*T. vivax*, *T. congolense*, *T. simae* and *T. brucei*) detected using the same ITS-1 primers that I used (Isaac *et al.*, 2016). They found that there was variation in the presence of different species of trypanosomes in different species of flies but significant associations were found only for fly sex (across all species of trypanosomes prevalence was higher in females than in males) and collection site (one site showed a higher prevalence of *T. congolense* than the other sites sampled). However, there appeared to be overall less complexity than in the region of Kenya that I studied because there were no significant interactions detected in relation to site, sex or species of flies. In my study, "subpopulation" appeared to be the explanatory variable that was most often associated with variation in trypanosome prevalence, but often in interactions with sex or age and the strength of associations varied by species of trypanosome. Presence of *T. vivax* was significantly associated with an interaction between subpopulation and sex, with many significant differences among pairwise comparisons defined in post hoc tests (Appendix A.12). In contrast, while infection of *T. congolense* was significantly associated with subpopulation, no significant differences were found in post-hoc comparisons (Appendix A.10). For individual comparisons by tsetse species, *T. vivax* infection of *G. pallidipes* was significantly associated with interactions between site and sex and sex and age, and post hoc tests found significant differences within and between the factors (Appendix A.24). For associations of *T. congolense* infection in *G. brevipalpis* (Appendix A.18) and *T. brucei* infection in *G. longipennis* (Appendix A.19), post hoc tests did not identify significant differences despite overall significance of some factors. These results suggest that despite the relatively large sample size in my study (particularly for *T. vivax*, which was the

most abundant species found), there was still not enough power to resolve the complex relationships among individual species of trypanosomes. Thus, while it would be worthwhile to investigate specific interactions between tsetse factors and a range of trypanosome species, this might require more targeted sampling approaches to increase sample sizes.

## 2.6 Conclusions

Overall, my results emphasise the challenges of resolving the epidemiology of complex vector-parasite-host systems in natural ecosystems. *Glossina pallidipes*, *G. austeni*, *G. brevipalpis* and *G. lon*gipennis were found to show variation in their distributions in the Shimba Hills and Nguruman regions of south eastern Kenya and showed varying levels of infection with different species of trypanosomes. The PCR screening method was a sensitive way to determine infection status of tsetse flies but different conclusions would have been reached for some species of trypanosome using different tsetse tissues and using general ITS-1 and species-specific primer sets. So, choice of approach might require optimisation specific for particular combinations of vectors and parasites in different regions. Previous studies have tended to investigate individual attributes of tsetse flies that could influence relative susceptibility to trypanosomes but my research clearly indicates the risk of drawing misleading conclusions if considering factors in isolation. Tsetse sample site, species, age, sex and their interactions were all found to influence trypanosome prevalence using GLM analyses so considering single factors would have lead to false predictions about direction of relationships because it would have averaged over important differences. Although the distribution of tsetse flies did not make it possible to separate the effects of geographic location from attributes of individual species, these results emphasise that tsetse flies should not be treated as a homogeneous vector group and risk factors should be conducted separately for each species of tsetse in each particular site from which they are sampled. Although I attempted to use a GLM approach to investigate factors associated with particular species of trypanosomes, biases in the distribution of trypanosome species across sites and vectors also confounded results and lead to an even more complex picture of associations. Nevertheless, they further emphasise the need for more integrated studies of factors when considering risks posed by particular disease agents.

# Chapter 3 Prevalence and association of *Sodalis glossinidius* and *Trypanosoma spp*. in tsetse flies

## 3.1 Abstract

Tsetse flies (*Glossina spp.*) are biological vectors of trypanosomiasis in humans and animals. Prevalence of *Sodalis glossinidi*us, a common secondary endosymbiont, has been reported with varied rates but whether there is an association between presence of the bacteria and trypanosome infection in tsetse flies remains uncertain. The two objectives of this study were to: 1) investigate what factors affect the presence or absence of *S. glossinidius* in *G. austeni, G. brevipalpis*, *G. longipennis* and *G. pallidipes*; and 2) to test whether presence of any *Trypanosoma spp*. found in the area (*T. congolense, T. brucei* and *T. vivax*) is associated with presence of the endosymbionts, subpopulation (a combination between sites and *Glosina spp.*), sex or age of the host flies. Among 1090 tsetse flies collected from two sites in the Shimba Hills (Buffalo Ridge and Zungu Luka) and Nguruman (Mukinyo and Sampu) regions of eastern Kenya, prevalence of *S. glossinidius* was 34.0% based on PCR screening, which was tested using primers targeting different gene regions and confirmed by sequencing. There was high variation in prevalence of the endosymbiont in relation to site of sampling (0.0% – 83.4%) and species of tsetse flies (0.0% – 97.9%). A significant association was found between *S. glossinidius* and trypanosome presence using $\chi^2$ tests ($\chi^2$ = 75.0, P-value = 0.0001) but no association was found within regions. Generalised Linear Models (GLMs) incorporating variation due to geographic location, and intrinsic tsetse factors (species, sex and age) revealed more complex patterns of association: *Sodalis* influenced trypanosome presence but only in interactions with other intrinsic factors and only in some species of trypanosomes. The strongest association was found for *T. congolense* and no association was found for *T. vivax*. Complicated interactions of this type can be visualised more clearly using Multiple Correspondence Analysis (MCA), which also suggested only a weak association between trypanosomes and *Sodalis* presence. I suggest that previous conclusions about the presence of endosymbionts increasing vector competence in tsetse flies may have been confounded by other factors, such as community composition of the tsetse flies and trypanosomes found in different regions.

## 3.2 Introduction

*Trypanosoma congolense*, *T. vivax* and *T. b. brucei* are the three main pathogens of Animal African Trypanosomiasis (AAT), which causes dramatic loss of farm animal production. These diseases affecting public health and animal husbandry lead to economic loss in both endemic and epidemic areas (Angara *et al.*, 2014, Sahaw, 2009, Wilson *et al.*, 1963, Ministry of Livestock Development (Kenya) , 2011). Trypanosomes biologically transmit to humans and animals via saliva of tsetse flies, blood-sucking vectors of trypanosomiasis. Both males and females of the insects have diets limited to vertebrate blood but other supplemental nutrients and substances that are lacking or cannot be synthesized based on this normal diet must be obtained from endosymbionts for maintenance and development (Douglas, 1989, Koch, 1967, Geiger *et al.*, 2011). Many endosymbionts have been reported (Aksoy, 2000, Lindh and Lehane, 2011) in various tissues of tsetse flies but *Wigglesworthia glossinidia, Sodalis glossinidius* and *Wolbachia spp.* are the three major bacterial species harboured by tsetse flies (O'Neill *et al.*, 1993).

The three species have similar routes of transmission but are thought to play different roles in the tsetse flies. The primary endosymbiont, *W. glossinidia*, resides in the midgut, milk glands, and fat bodies of larvae (Balmand *et al.*, 2013). These bacteria infect the next generation of tsetse flies by transovarial transmission (Attardo *et al.*, 2008). *Wolbachia* are detectable in both tsetse gonads and ovaries (Cheng and Aksoy, 1999) and infect oocytes and embryos prior to developing intrauterine (Balmand *et al.*, 2013). *Sodalis glossinidius* spreads in the midgut (Aksoy, 1995a, Dale and Maudlin, 1999, Cheng and Aksoy, 1999), haemolymph, muscle, fat bodies, salivary glands (Cheng and Aksoy, 1999), milk glands, reproductive system and first stage of larvae (Balmand *et al.*, 2013). Routes of *S. glossinidius* infection are transovarial transmission via haemolymph (Cheng and Aksoy, 1999), vertical transmission to intrauterine larvae via milk-gland secretions and horizontal transmission during mating (De Vooght *et al.*, 2015).

*Wigglesworthia glossinidia* provides nutrition for larvae developing in the uterus (Pais *et al.*, 2008, Nogge and Gerresheim, 1982) and has been implicated in immune regulation (Weiss *et al.*, 2011), while *Wolbachia spp.* infection affects

fertilization (Alam *et al.*, 2011). The functional role of *S. glossinidius* in tsetse flies has not been clearly defined (Wang *et al.*, 2013b) but presence of *S. glossinidius* has been related to vector competence of trypanosomes, possibly related to lectin-inhibitory activity (Welburn and Maudlin, 1999). It has been suggested that chitinase from *S. glossinidius* breaks down chitin and produces N-acetyl-D-glucosamine (Welburn *et al.*, 1993), which inhibits lectin function. Without lectin function, trypanosomes more easily penetrate into the midgut of tsetse flies (Welburn and Maudlin, 1999). *Glossina morsitans morsitans* that were treated with the antibiotic Steptozocin to eradicate *S. glossinidius,* showed decreased reproductive capacity of tsetse flies, decreased longevity and increased trypanosome refractoriness compared to untreated flies (Dale and Welburn, 2001, Wang *et al.*, 2013a). The ability of *G. palpalis gambiensis* from Burkina Faso to be infected by experimental inoculation with *T. b. brucei* and *T. b. gambiense* was statistically linked to existence of *S. glossinidius* based on PCR methods targeting the 16S rDNA gene (Geiger *et al.*, 2007). This suggests that different *Sodalis* genotypes might be associated with differing capacities for facilitation of trypanosome establishment. Moreover, susceptibility of tsetse flies to trypanosome infection might increase in response to a greater density of the symbiont in the fly gut (Cheng and Aksoy, 1999). A significant association between presence of *S. glossinidius* (using GPO1 primers; described by O'Neill *et al.* (1993)) and trypanosomes (using general ITS-1 CF and BR primers; described by Davila (unpublished) (Njiru *et al.*, 2005)) in *G. pallidipes* from Kenyan coastal forests was also reported (Wamwiri *et al.*, 2013). Similarly, a statistically significant correlation was found for prevalence of *S. glossinidius* and trypanosomes in tsetse flies in Campo and Bipindi based on pSG2 primers for the secondary endosymbionts and trypanosome species-specific primers described by Masiga *et al.* (1992), Moser *et al.* (1989) and Farikou *et al.* (2010a). In contrast, *Sodalis* has been reported to stimulate immune function of tsetse flies, which could decrease the levels of trypanosome infection (Dale *et al.*, 2001, Hao *et al.*, 2001, Lehane *et al.*, 2004, Rose *et al.*, 2014, Weiss *et al.*, 2013). Flies without *Sodalis* showed a reduction in trypanosome infection compared to those that had the endosymbiont (Dale and Welburn, 2001, Weiss *et al.*, 2013). However, Dennis *et al.* (2014) reported that there was no association between *S. glossinidius* and *Trypanosoma spp.* infection in *Glossina brevipalpis*, G. *pallidipes* and *G. morsitans* from Zambia. Inconsistences have also been found

based on experimental infections. Teneral (0 - 2 days old) male *G. pallidipes* from the Trypanosomiasis Research center (TRC), Kenya were inoculated with *T. b. rhodesiense* (KETRI2537; isolated from a human host, Busoga, 1972), *T. b. brucei* (KETRI3386; isolated from *G. pallidipes*, Kibwezi, 1979) and *T. congolense* (EATRO993; isolated from *G. pallidipes*, South Nyanza, 1962) (Wamwiri *et al.*, 2014). Using GPO1 primers for *Sodalis* and trypanosome species-specific primers significant associations were found between *Sodalis* status and *T. b. rhodesiense* and *T. congolense* presence with but not for *T. b. brucei*. Inconsistency between results could be because individual studies were conducted using different tsetse fly species from different regions. Effects of geographic locations, species of tsetse flies and intrinsic factors of tsetse flies (such as species, age and sex) could contribute to the conflicting results. What has been lacking is a detailed study that combines traits of the flies with prevalence of both the endosymbiont and various species of trypanosomes across different geographic regions.

The methods used to detect the presence of endosymbionts could also affect interpretation of associations with trypanosome susceptibility. Although isolation and culture are basic methods for bacterial identification, molecular diagnosis based on the polymerase chain reaction (PCR) is a practical method for screening of large numbers of samples in a short time. However, the potential sensitivity of PCR screening could be affected by diversity of *S. glossinidius* communities in different regions and sources of blood feeding. Thus, the most appropriate tissue sampling and primers used for *S. glossinidius* identification should be considered when designing a reliable diagnostic test for their presence. *Sodalis glossinidius* reside in various tissues of tsetse flies, and PCR inhibitors such as haematin, melanin, collagen (Opel *et al.*, 2009) in tsetse flies are able to interrupt PCR amplification. Thus, results of *S. glossinidius* screening in head plus proboscis could be different from those based on abdomen parts.

A number of primer sets have been developed to identify *S. glossinidius,* each developed for different target populations and for different regions. For example, pSG2 primers were developed by Farikou *et al.* (2010) for *S. glossinidius* identification in *G. palpalis palpalis*, *G. pallicera*, *G. caliginea* and *G. nigrofusca* from Cameroon and target a gene on the extrachromosomal plasmid 2, which is abundant in this bacteria endosymbiont (Darby *et al.*, 2005). GPO1 primers also target a gene on this plasmid and have been applied to detect

secondary endosymbionts in many species of laboratory-reared tsetse flies: *G. austeni, G. brevipalpis, G. fuscipes fuscipes* and *G. tachinoides* from the Seibersdorf Agricultural Research Laboratory, Vienna, Austria; *G. m. morsitans* from the Tsetse Research Laboratory, Bristol University, England; and *G. palpalis* from the University of Alberta (O'Neill *et al.*, 1993). *Sodalis* exochitinase (Sexo) primers were developed for amplification of nuclear exochitinase gene fragments in *G. palpalis gambiense* from Centre de coopération internationale en recherche agronomique pour le développement (CIRAD, Baillarguet, France) (Soumana *et al.*, 2013a). Hem primers have been developed to target a gene encoding the hemolysin protein that bacteria use to digest the cell membrane of erythrocytes and were designed to investigate *S. glossinidius* in *G. m. morsitans* from Zimbabwe (Pais *et al.*, 2008). Prevalence in *G. palpalis palpalis* from Bipindi of South Cameroon (64.4%) was higher than in *G. pallicera and G. caliginea* from Campo (45.3%) based on pSG2 primers (Farikou *et al.*, 2010a). Based on another nuclear gene (GroEL), a survey in Zambia indicated that prevalence of *S. glossinidius* was highest in *G. brevipalpis* (93.7%), followed by *G. m. morsitans* (17.5%) and *G. pallidipes* (1.4%) (Dennis *et al.*, 2014) while prevalence in *G. pallidipes* (16.0%) was higher than in *G. austeni* (3.7%) from Kenya using the plasmid GPO1 primers (Wamwiri *et al.*, 2013). Thus, the particular primer sets used could affect differences in conclusions about *Sodalis* prevalence in different geographic regions and in different species of tsetse flies. Since *Sodalis* can be found in haemolymph and muscle (Cheng and Aksoy, 1999) it should be possible to use head plus proboscis or the abdomen parts for diagnosis of prevalence but whether results based on the two body parts would give concordant results has not been tested. No studies have yet directly assessed whether interpretation of prevalence of the endosymbionts varies depending on the particular set of primers used, in a particular region or applied to a particular body part.

Because the routes of *S. glossinidius* infection are through both vertical and horizontal infection by mating (De Vooght *et al.*, 2015), sex and age of the tsetse flies also could affect prevalence. For example, a slight decrease of the *Sodalis* incidence using GPO1 primers was found with increasing fly age (Wamwiri *et al.*, 2014), as well as a significant difference between male and female *G. pallidipes* but not *G. m. morsitans* (Dennis *et al.*, 2014). Male *G. pallidipes*

(11.1%) showed a higher *Sodalis* prevalence than females (1.2%) but no difference was found between male (12.0%) and female *G. m. morsitans* (18.8%). In addition, infection with trypanosomes could affect the status of *S. glosssinidius* in tsetse flies because *T. congolense, T. brucei* and *S. glossinidius* establish in the midgut. Nevertheless, intrinsic tsetse factors associated with *Sodalis* status have not been investigated to tease out whether *Sodalis* presence itself enhances opportunities for trypanosome infection or whether the same factors that make prevalence of *Sodalis* more likely also make trypanosome infection more likely.

The overall aim of this chapter was to assess whether *S. glossinidius* is one of the factors that drive the relative susceptibility of tsetse flies to trypanosomes. I focused on two specific objectives. The first was to determine prevalence of *S. glossinidius* in *G. austeni, G. brevipalpis G. longipennis* and *G. pallidipes* from two sites each in the Shimba Hills and Nguruman regions of eastern Kenya, which are surrounded with different vegetation types, humidity and available species of tsetse flies, in order to test whether intrinsic factors (subpopulation, sex and age) of tsetse flies and overall *Trypanosoma spp*. infection status (i.e. infection with any species) are associated with presence of the endosymbionts. The second was to analyse associations of the three main animal pathogenic trypanosome species (*T. congolesne*, *T. brucei* and *T. vivax*) with *S. glossinidius* infection, tsetse fly intrinsic factors and collecting sites. In addition, to assess whether interpretation of associations could be biased by differences between methods in detection of *Sodalis*, comparisons were made between PCR-based screening based on head plus proboscis and abdomen tissues and between three different primer sets. The main hypothesis being tested in this study was that the apparent association between presence of *Sodalis* and trypanosome prevalence in tsetse flies could be confounded by site, species of fly, sex and age, as well as species of trypanosomes so it is important to account for these factors when interpreting the role that the endosymbiont plays in explaining differences in the association.

# 3.3 Materials and methods

## 3.3.1 Optimisation of PCR screening methods for *Sodalis glossinidius* in tsetse flies

### 3.3.1.1 Comparision of screening results based on different primers sets: pSG2; Hem; and GPO1 primers

All samples were initially screened with pSG2 primers (Farikou *et al*., 2010a). An initial test of 50 samples comparing amplfication in head plus proboscis and abdomen parts revealed consistent results based on the two tissues. Head plus proboscis parts were thus screened for the remaining samples to allow comparison of the presence of trypanosomes amplified from the same tissues (see Chapter 2). To determine whether other sets of primers might lead to different conclusions about presence of *Sodalis*, two additional primer sets were tested, using both head plus proboscis and abdomen parts: 20 samples that were pSG2 positive and 20 that were pSG2 negative were used and results from the three sets of primers compared. A positive control of *S. glossinidius* isolated from *G. pallidipes* from the International Atomic Energy Agency (IAEA) was also included. The pSG2 primers (Farikou *et al*., 2010a) and the GPO1 primers (O'Neill *et al*., 1993), amplify DNA fragments of extrachromosomal plasmid 2 (Darby *et al*., 2005), with expected fragments of 120 bp and 1,200 bp, respectively. The Hem primers were developed from sequences of *S. glossinidius* (GenBank accession no. AP008232) (Pais *et al*., 2008) to target 650 bp of a gene encoding the hemolysin protein (Smith *et al*., 2013).

For the pSG2 primers, PCR reaction mixtures contained 1 µl of 10X PCR Master Mix buffer (45 mM Tris-HCl, pH 8.8 at 25°C; 11 mM $(NH_4)_2SO_4$; 4.5 mM $MgCl_2$; 0.113 mg/ml BSA; 4.4 µM EDTA; and 1.0 mM each of dATP, dCTP, dGTP and dTTP (Thermo Scientific), 10 µM of each primer (Eurofins MWG Operon), 20-200 ng DNA template and 1 unit of *Taq* DNA polymerase (Thermo Scientific). GPO1 and Hem products were each amplified in 10 µl using 5 µl of Dream *Taq* Green PCR master Mix (2X) (Thermo Scientific), 10 µM of forward and reverse primers, and DNA template. PCR cycles for the four primers are described in Table 3.2. All PCR amplicons were analysed by gel electrophoresis in 1.5% UltraPure™ Agarose gels (Invitrogen, Carlsbad, CA.) with 2% Ethidium Bromide (Invitrogen, Carlsbad, CA, U.S.A.), in 1X TBE buffer (108 g of Tris Base, 55 g of Boric acid and 40 ml of

0.5 M EDTA, pH 8.0, with $ddH_2O$ added to 10 liters). The conditions of gel electrophoresis were 100 V., 300 A., 50 min. PCR products of positive samples (one sample from each primer set) and the positive control were cleaned using PCR purification using ExoSAP-IT PCR Clean-up Kits (GE Healthcare) and sent to the University of Dundee sequencing service for direct sequencing. The results were analysed using BLAST to confirm that PCR products amplified with pSG2, Hem and GPO1 primers were *S. glossinidius* genes. All chromatograms were visualised using the Geneious® 7.0.3 (Biomatters Ltd., Auckland, NZ)

To further test the consistency among primers, fragment amplification using each primer set (head and proboscis parts only) was compared across all tsetse samples (N = 1090 samples), in relation to subpopulations (i.e. the combination of site and tsetse species).

**Table 3.1** Sequences of pSG2, GPO1 and Hem primers used for *Sodalis glossinidius* screening.

Sizes of the expected amplification products and the sources of the primers are indicated.

| Name | Sequences | Size (bp) | References |
|---|---|---|---|
| pSG2 | pSG2F: TGA-AGT-TGG-GAA-TGT-CG<br>pSG2R : AGT-TGT-AGC-ACA-GCG-TGT-A | 120 | Farikou *et al.* (2010) |
| GPO1 | GPO1F: TGA-GAG-GTT-CGT-CAA-TGA<br>GPO1R: ACG-CTG-CGT-GAC-CAT-TC | 1,200 | O'Neill *et al.* (1993) |
| Hem | HemF: ATG-GGA-AAC-AAA-CCA-TTA-GCC-A<br>HemR: TCA-AGT-GAC-AAA-CAG-ATA-AAT-C | 650 | Pais *et al.* (2008) |

**Table 3.2** PCR cycles of *Sodalis glossinidius* screening using pSG2, GPO1 and Hem primer sets.

PCR conditions of pSG2, GPO1 and Hem primer sets followed Farikou *et al.* (2010), O'Neill *et al.*, (1993) and Pais *et al.* (2008), respectively.

| PCR steps | PCR conditions | | | | | |
|---|---|---|---|---|---|---|
| | pSG2 Farikou *et al.* (2010a) | | GPO1 O'Neill *et al.* (1993) | | Hem Pais *et al.* (2008) | |
| | Temp (°C) | Time (sec) | Temp (°C) | Time (sec) | Temp (°C) | Time (sec) |
| **Pre-heating** | 94 | 5 min | 94 | 5 min | 94 | 2 min |
| **Denature** | 94 | 30 | 94 | 60 | 94 | 30 |
| **Annealing** | 50 | 30 | 55 | 60 | 54 | 40 |
| **Elongation** | 72 | 60 | 72 | 60 | 72 | 60 |
| **Number of cycles** | 40 cycles | | 35 cycles | | 30 cycles | |
| **Final elongation** | 72 | 10 min | 72 | 10 min | 72 | 7 min |

### 3.3.1.2 Sequence comparison and identification of genetic variation

To test whether inconsistencies in amplification among primer sets were due to genetic variation in the endosymbionts present, I cloned and sequenced pSG2 PCR products from 10 individual flies that were Hem and GPO1 negative. Positive pSG2 fragments from DNA samples of four *G. pallidipes* (two samples from Zunku Luka and two samples from Mukinyo), four *G. austeni* (all samples from Zunku Luka), and two *G. longipennis* (from Mukinyo and Sampu) were purified and cloned using TOPO®-TA Cloning Kits (Invitrogen, Carlsbad, CA, U.S.A). After purifying six plasmids of each PCR product with QIAprep Spin Miniprep Kits (Qiagen Inc, Paisley, UK), all extracted plasmids were sent for sequencing using universal vector primers (M13F and M13R). Chromatographs were visualised and manually manipulated using the Sequencher software programme (version 4.5; Gene Codes, Inc. Michigan). Identity of *S. glossinidius* was determined using BLAST.

To find out the most appropriate primer set for determination of genetic variation of *S. glossinidius* from the seven subpopulations, 20 positive PCR fragments of *G. pallidipes*, *G. austeni* and *G. brevipalpis* from both Hem and

GPO1 primer sets were extracted using QIAquick Gel Extraction Kits (Qiagen Inc, Paisley, UK) and sent for direct sequencing at the DNA Sequencing and Services, University of Dundee. The Hem fragments were amplified from seven samples from ZuGp, five samples from ZuGa, three samples from BRGb and five samples from BRGp while the GPO1 fragments were prepared from six samples from ZuGp, six samples from ZuGa, three samples from BRGb and five samples from BRGp. All chromatographic sequencing results were visualiseded and corrected using the Sequencher software programme. In order to visualise variation among sequences, 18 GPO1 fragments (four *G. pallidipes* from Buffalo Ridge, six *G. pallidipes* from Zungu Luka, six *G. austeni* from Zungu Luka and two *G. brevipalpis* from Buffalo Ridge) were used to build a neighbour-joining tree using the Molecular Evolutionary Genetics Analysis (MEGA) software, version 6.06 (Tamura *et al.*, 2013). Not enough variation was found using Hem fragments to perform similar analyses.

### 3.3.2 Prevalence of *Sodalis glossinidius* in relation to intrinsic factors of tsetse flies and *Trypanosoma spp.* infection

Presence of *S. glossinidius* based on the most appropriate PCR screening method identified was first qualitatively compared with respect to the sex, age, species of tsetse flies, and sampling site data taken from chapter 2. To test whether there was an overall association between presence or absence of *S. glossinidius* and *Trypanosoma spp.* $\chi^2$ tests (with a significance threshold of 0.05) were initially applied to all samples and then within the two regions separately (the Shimba Hills and Nguruman).

### 3.3.3 Generalised Linear Models of *Sodalis glossinidius* and trypanosome status

Generalised Linear Models (GLMs) using the glm2 package of The R programme were applied to test the affects of these variables on *S. glossinidius* status (presence or absence) in tsetse flies, as well as all pairwise interactions, using the Binomial family, as described in chapter 2. Four models were used to investigate these relationships (Table 3.3). Firstly, GLM Model 1 tested for an association of *S. glossinidius* status ("0" or "1" for presence or absense) with presence of any trypanosome species and the tsetse intrinsic factors (from the same sets of samples and methods described in the chater 2). In the remaining

three models, associations of *S. glossinidius* status with intrinsic factors and trypanosome status were assessed separately for each of the most three abundant species of trypanosomes (*T. congolense*, *T. brucei* and *T. vivax*). In order to fit the best models, the variables from the full models (all explanatory variables and their interactions) were sequentially removed using the "step()" function and a backward elimination technique manually. In order to identify relationships within variables and across variables that were significantly different from one another, the TukeyHSD command was applied to the best fitting models for post hoc comparisons. The predict() function in R was used to predict the probability of *Sodalis* positive status for the best-fitting model for each response variable.

**Table 3.3** Variables for statistical analysis in a GLM model to determine associations of *S. glossinidius* with intrinsic factors and *Trypanosoma spp.* status, using *Sodalis* presence (1) or absence (0) as a binary response variable.

The same approach was used to compare the influence of: Model 1) overall trypanosomes; Model 2) *T. congolense;* Model 3*) T. brucei;* and Model 4) *T. vivax* Shown are the response variable, explanatory variables tested, and the classification (type of data) and levels (composition) of each variable considered.

| Intrinsic factors | Type of data | Composition |
|---|---|---|
| Explanatory variables | | |
| subpopulation | categorical | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl[a] |
| sex | categorical | male and female |
| age | categorical | young, juvenile and old |
| trypanosome status | binary | positive (1) and negative (0) |
| Interactions considered | | |
| subpopulation ∗ sex | | |
| subpopulation ∗ age | | |
| subpopulation ∗ trypanosome status | | |
| sex ∗ age | | |
| sex ∗ trypanosome status | | |
| age ∗ trypanosome status | | |
| subpopulation ∗ sex ∗ age | | |
| subpopulation ∗ sex ∗ trypanosome status | | |
| sex ∗ age ∗ trypanosome status | | |
| age ∗ trypanosome status ∗ subpopulation | | |
| subpopulation ∗ sex ∗ age ∗ trypanosome status | | |

[a] BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu.

Additional models were constructed (Models 5 – 8), also using the Binomial family of GLMs, but with trypanosome infection status as the response variable and *Sodalis* as a binary variable with values of 0 and 1 (Table 3.4). Model selection was conducted as for the comparisons using *Sodalis* as the response variable.

**Table 3.4** Variables for statistical analysis in a GLM model to determine associations of *Trypanosoma spp*. with intrinsic tsetse factors and *S. glossinidius* status of tsetse flies, using *Trypanosoma* presence (1) or absence (0).

The same approach was used considering the response variable as: Model 5) overall trypanosomes; Model 6) *T. congolense;* Model 7*) T. brucei;* and Model 8) *T. vivax*.

| Intrinsic factors | Type of data | Composition |
|---|---|---|
| Explanatory variables | | |
| subpopulation | categorical | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl[a] |
| sex | categorical | male and female |
| age | categorical | young, juvenile and old |
| *Sodalis* status | binary | positive (1) and negative (0) |
| Interactions considered | | |
| subpopulation ∗ sex | | |
| subpopulation ∗ age | | |
| subpopulation ∗ *Sodalis* status | | |
| sex ∗ age | | |
| sex ∗ *Sodalis* status | | |
| age ∗ *Sodalis* status | | |
| subpopulation ∗ sex ∗ age | | |
| subpopulation ∗ sex * *Sodalis* status | | |
| sex ∗ age ∗ *Sodalis* status | | |
| age ∗ *Sodalis* status ∗ subpopulation | | |
| subpopulation ∗ sex ∗ age ∗ *Sodalis* status | | |

[a] BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu.

### 3.3.4 Multiple correspondence analysis among status of *Sodalis* and trypanosome and intrinsic factors of tsetse flies

Principle correspondence analysis (PCA) was also applied to graphically visualise associations among trypanosomes, tsetse intrinsic factors and *Sodalis* in all collected flies (N = 1090), using Multiple Correspondence Analysis (MCA) as implemented in the FactoMineR package (version 1.30) and the "ggplot2()" function. MCA 1 analyses included five categorical variables, comprised of presence or absence of *S. glossinidius*, *Trypanosoma spp.*, as well as subpopulation, sex and age of tsetse flies (Table 3.5). MCA 2 analyses included seven categorical variables, comprised of presence or absence of *S. glossinidius*, *T. congolense*, *T. brucei*, and *T. vivax*, as well as subpopulation, sex and age of tsetse flies. The parameter $Eta^2$ is the correlation ratio and indicates the proportion of the total sum of squares that is explained by the predictor in each dimension. High $eta^2$ of each factor means strong correlations of each other so they were classified into the same cluster of correlation.

**Table 3.5** Variables used in Multiple correspondence analyses to determine associations of trypanosome and *Sodalis* status with intrinsic factors of tsetse flies.

| No | Intrinsic factors | Composition |
|---|---|---|
| **MCA graph 1** | | |
| 1. | subpopulation | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl[a] |
| 2. | sex | male and female |
| 3. | age | young, juvenile and old |
| 4. | *Sodalis* status | positive and negative |
| 5. | trypanosome status | positive and negative |
| **MCA grah 2** | | |
| 1. | subpopulation | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl[a] |
| 2. | sex | male and female |
| 3. | age | young, juvenile and old |
| 4. | *Sodalis* status | positive and negative |
| 5. | *T. congolense* status | positive and negative |
| 6. | *T. brucei* status | positive and negative |
| 7. | *T. vivax* status | positive and negative |

[a] BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu.

# 3.4 Results

## 3.4.1 Optimisation of PCR screening methods for *Sodalis glossinidius* in tsetse flies

### 3.4.1.1 Comparison of screening results based on different primers sets

*Sodalis* positive amplifications were obtained from head plus proboscis and abdomen parts using pSG2, GPO1 and Hem primers with 120, 1,200 and 650 bp of PCR product sizes, respectively for the 20 samples initially screened. Comparisons were strongly correlated (Table 3.6) between tissues but a few abdomen samples showed different results when screened with GPO1 (two samples). All *S. glossinidius* negative flies showed negative results when screened from both tissues using the three primer sets. Since both tissue types showed similar amplification, head plus proboscis was chosen as the appropriate tissue type to determine infection of *S. glossinidius* in tsetse flies because this enabled direct comparison with status of *Trypanosoma spp*. infection, which was based on head plus proboscis.

**Table 3.6** Comparison of trypanosome results in head plus proboscis and abdomen parts of tsetse flies using three sets of PCR primers: pSG2, Hem and GPO1 conditions.

| **Screening results of pSG2 primers** | | Head and proboscis | |
|---|---|---|---|
| | | positive | Negative |
| abdomen | positive | 20 | 0 |
| | negative | 0 | 20 |
| **Screening results of GPO1 primers** | | Head and proboscis | |
| | | positive | Negative |
| abdomen | positive | 18 | 0 |
| | negative | 2 | 20 |
| **Screening results of Hem primers** | | Head and proboscis | |
| | | positive | negative |
| abdomen | positive | 20 | 0 |
| | negative | 0 | 20 |

From 1090 head plus proboscis samples, there were 66.8% pSG2 positives but only 34.0% and 26.6% positives when screened with Hem and GPO1 primers, respectively. A substantial proportion of samples that were positive with pSG2 were negative for Hem (34.9%) (Figure 3.1a) or GPO1 (41.8%) (Figure 3.1b) but there was higher consistency in apparent prevalence between Hem and GPO1 (Figure 3.1c). However, 7.4% of samples also differed between these two primer sets. Comparing all primer sets, there were only 56.1% of samples that showed consistent prevalence of *S. glossinidius*: 25.0% were all positive and 31.1% were all negative. Importantly, 34.9% of pSG2 positive samples were negative for both Hem and GPO1 (Figure 3.1d).

There were also differences in the relative consistency of the primers when compared across subpopulations (Figure 3.2 – Figure **3.8**). There was a much higher discrepancy between results based on the three primer sets in flies sampled from Zunga Luga than for the other sites. The major difference in *Sodalis* screening results was higher numbers of pSG2 positive samples compared to the other primers, especially for *G. austeni* from Zungu Luka (Figure 3.5) and *G. pallidipes* from Zungu Luka and Mukinyo (Figure 3.4 and Figure **3.6**). However, *Sodalis* screening results based on Hem and GPO1 showed high consistency with one another, and showed much higher prevalence of *Sodalis* in tsetse flies from the Shimba Hills than Nguruman.

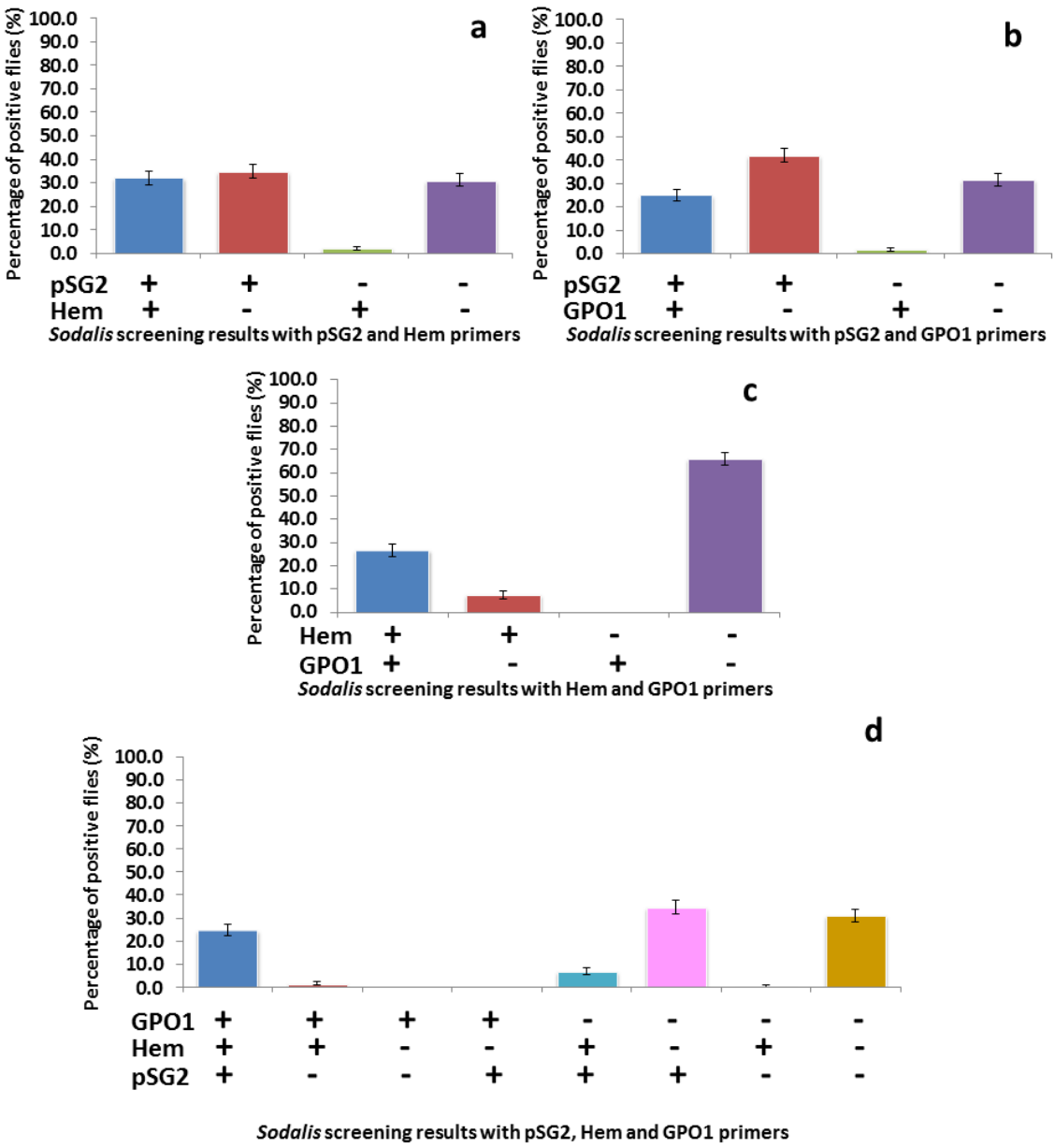



**Figure 3.1** Comparison of 1090 screening results for pSG2, Hem and GPO1 primers, across populations and species.
Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

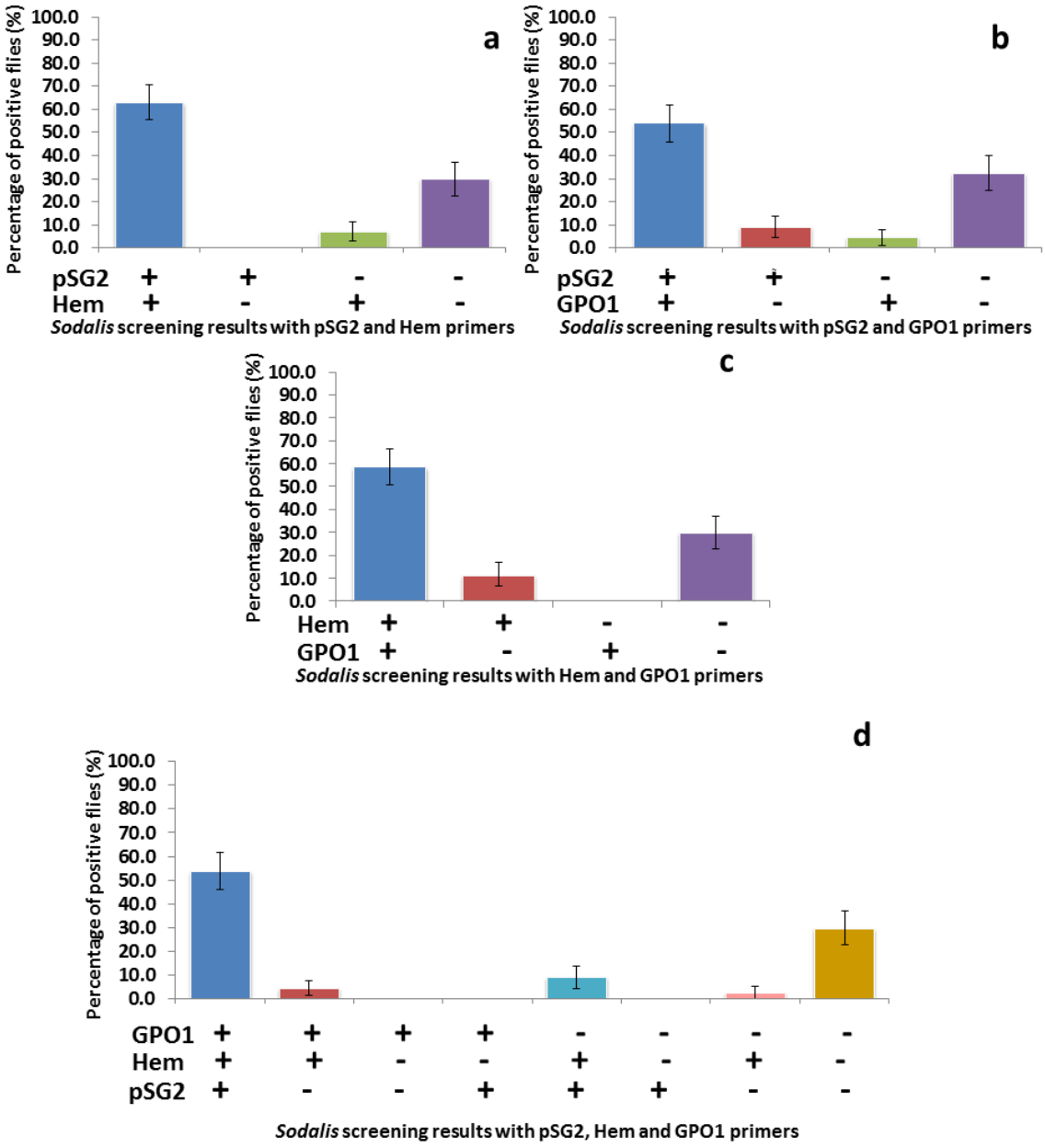



**Figure 3.2** Comparisons of 154 *G. pallidipes* screening results from pSG2, Hem and GPO1 primers in Buffalo Ridge (BRGp).

Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

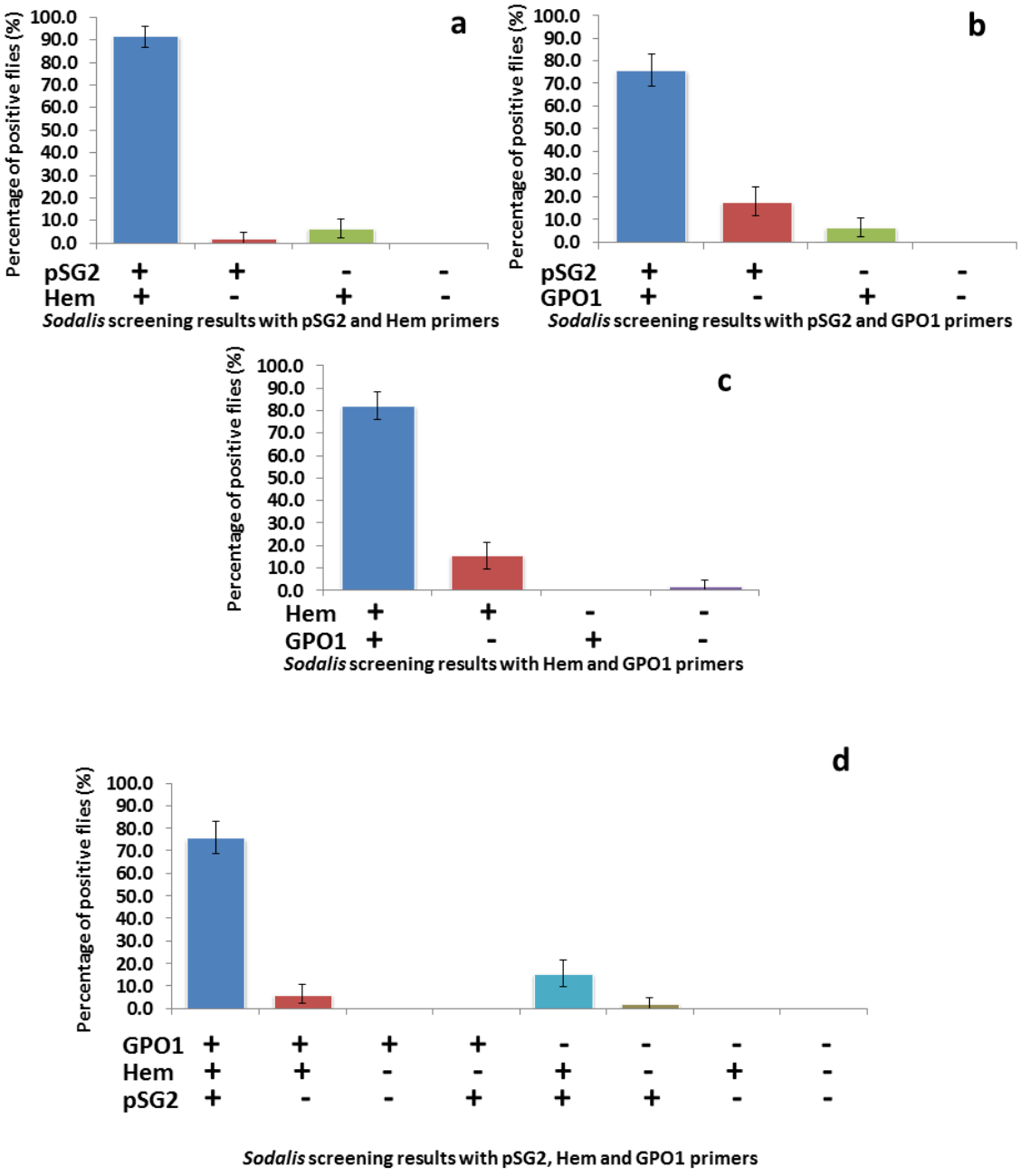



**Figure 3.3** Comparisons of 141 *G. brevipalpis* screening results from pSG2, Hem and GPO1 primers in Buffalo Ridge (BRGb).

Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

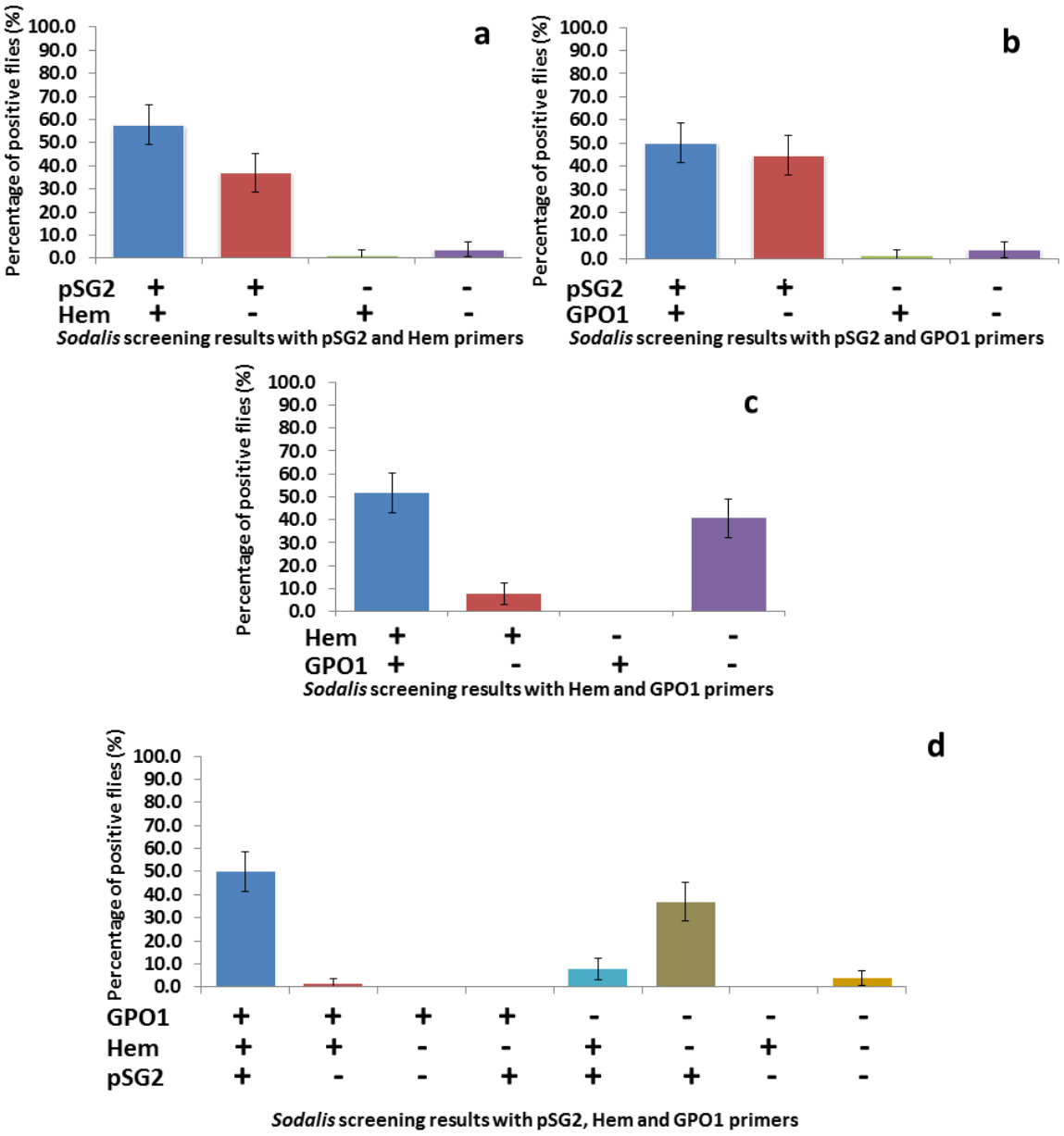



**Figure 3.4** Comparisons of 130 *G. pallidipes* screening results from pSG2, Hem and GPO1 primers in Zungu Luka (ZuGp).

Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

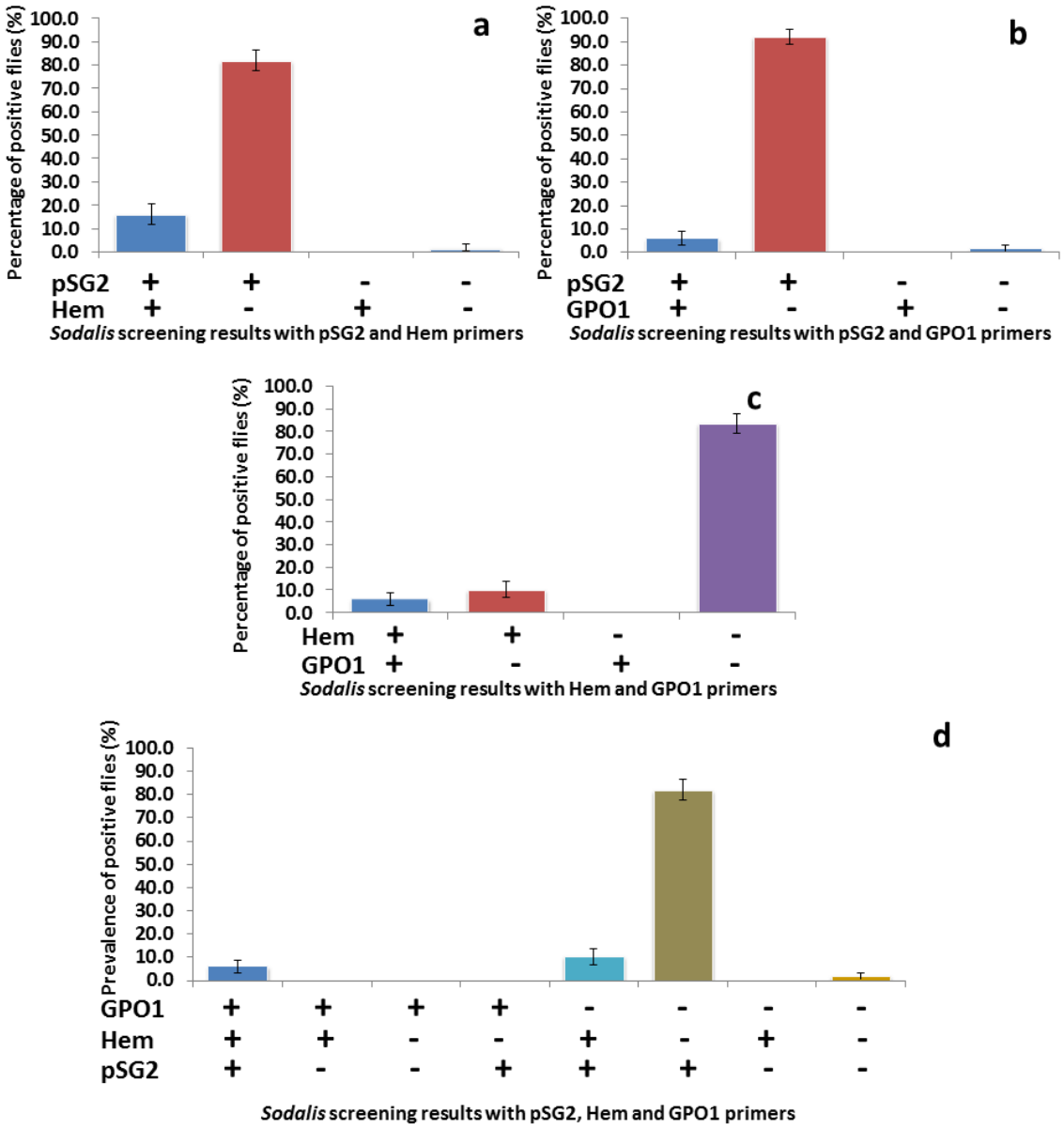



**Figure 3.5** Comparisons of 282 *G. austeni* screening results from pSG2, Hem and GPO1 primers in Zungu Luka (ZuGa).

Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

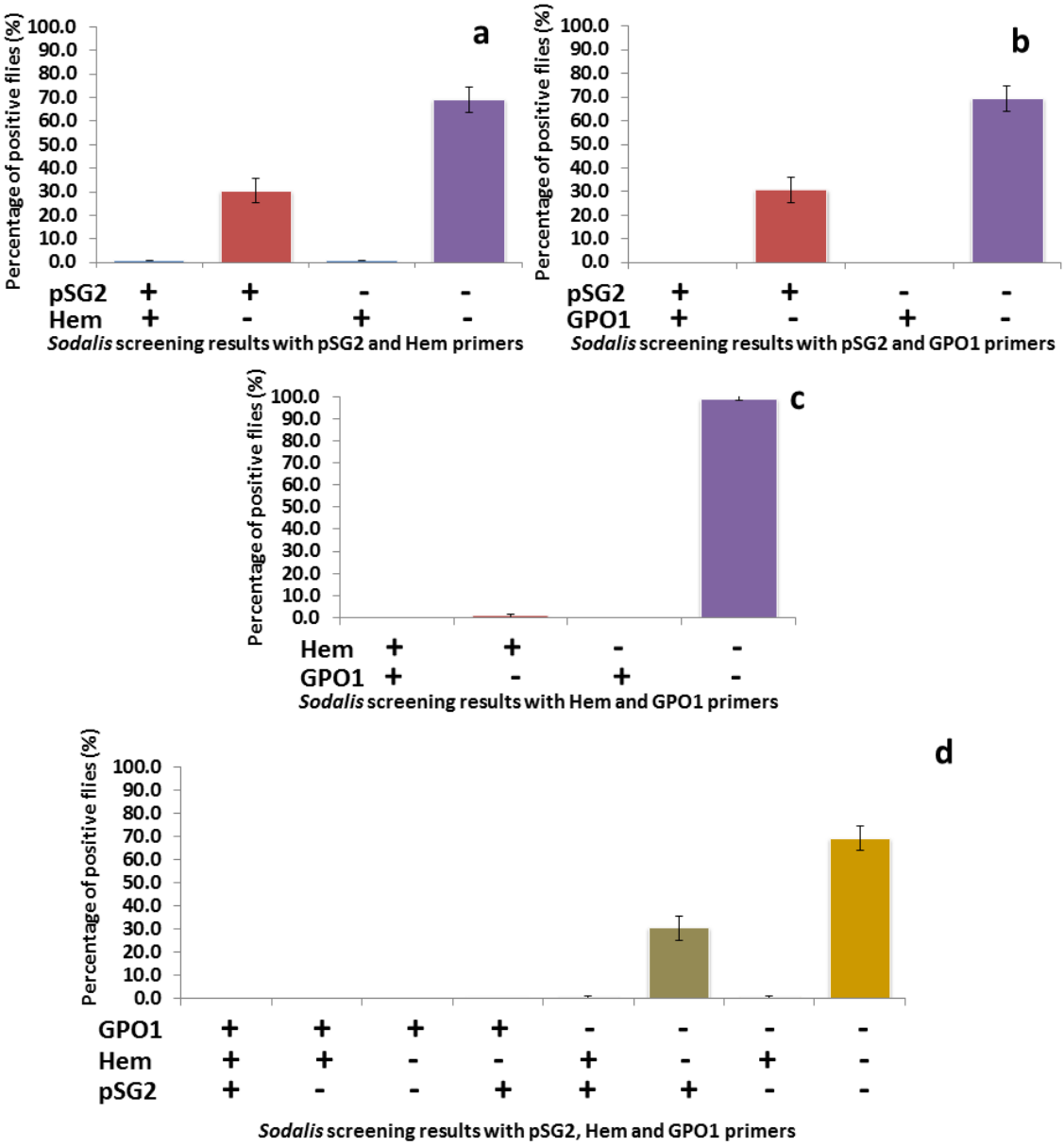



**Figure 3.6** Comparisons of 293 *G. pallidipes* screening results from pSG2, Hem and GPO1 primers in Mukinyo (MuGp).
Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

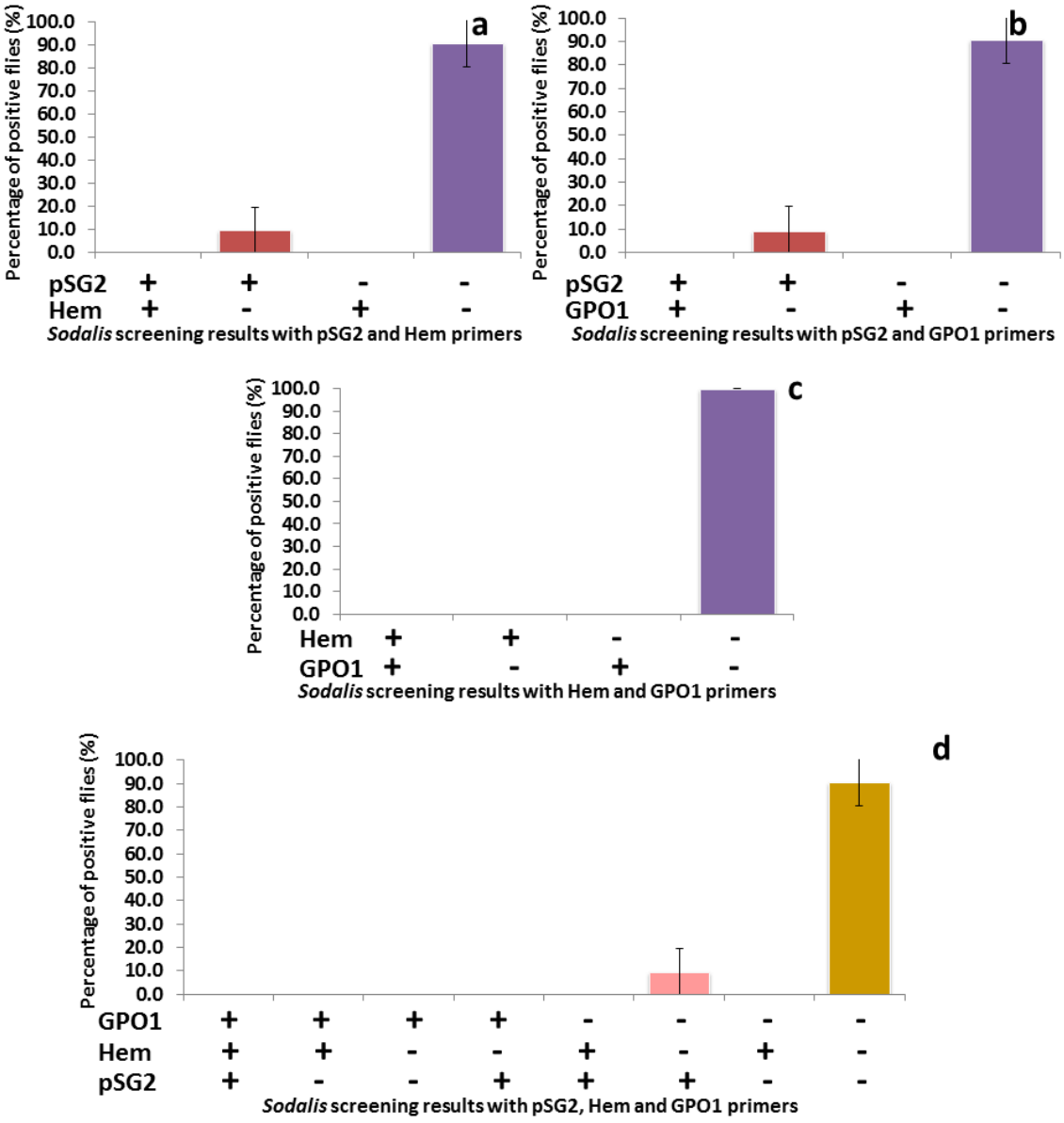



**Figure 3.7** Comparisons of 32 *G. longipennis* screening results from pSG2, Hem and GPO1 primers in Mukinyo (MuGl).
Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

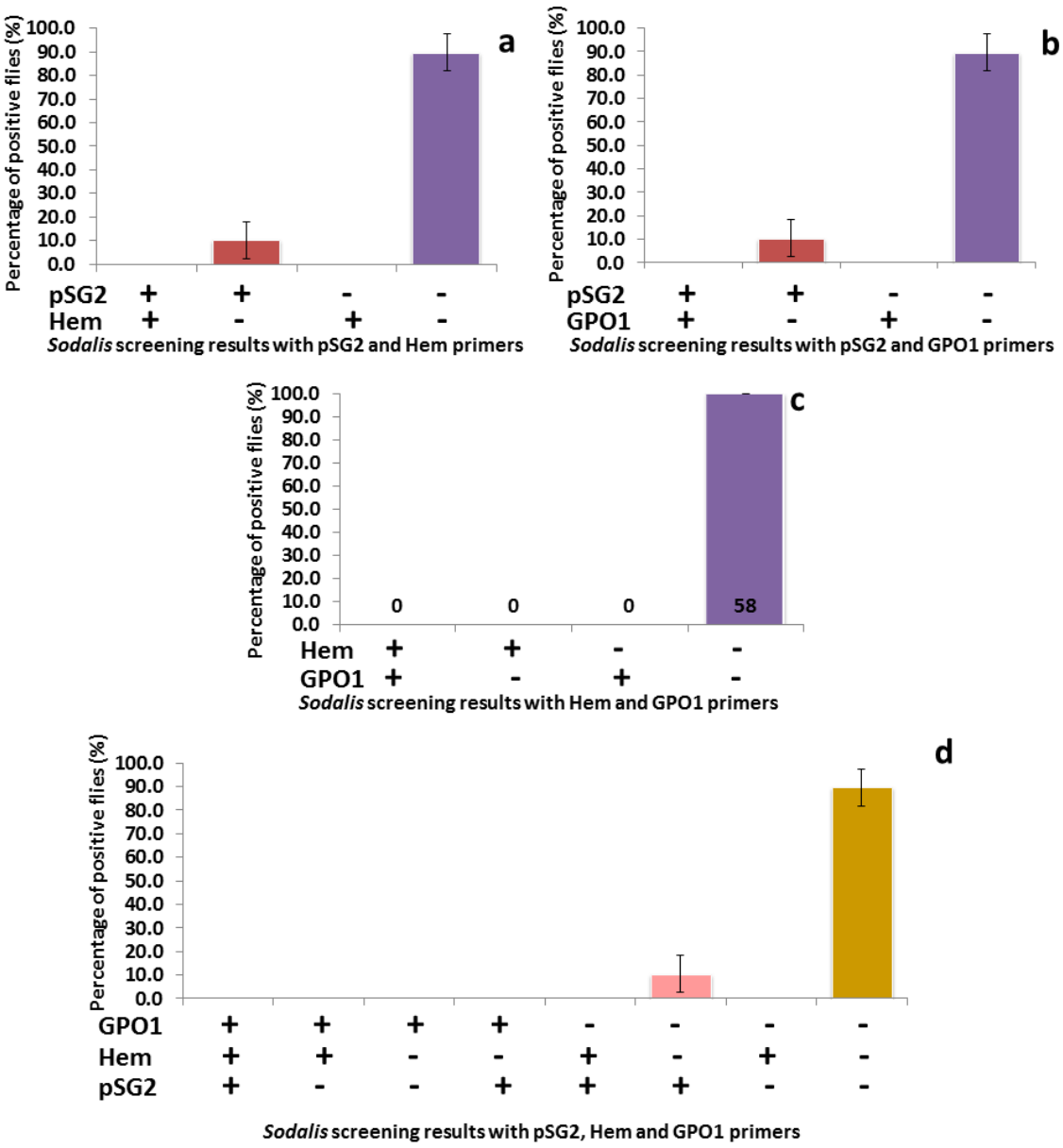



**Figure 3.8** Comparisons of 58 *G. longipennis* screening results from pSG2, Hem and GPO1 primers in Sampu (SaGl).
Percentage of positive flies found using: (a) pSG2 and Hem primers; (b) pSG2 and GPO1 primers; (c) Hem and GPO1 primers; (d) pSG2, Hem and GPO1 primers. 95% confident intervals (CI) are indicated.

### 3.4.1.2 Sequence comparison and identification of genetic variation

Although *S. glossinidius* was confirmed by sequencing of cloned amplicons, this was not the case for all amplification products for pSG2. Among the 10 pSG2 positive flies that were Hem and GPO1 negative that were cloned (Table 3.7), all clones of four samples in ZuGa and one each in ZuGp and MuGl (all showing amplification of 120 bp bands), were confirmed to be *S. glossinidius* (Table 3.7; Appendix B.1). The most similar sequences revealed through BLAST were: *S. glossinidius* str. 'morsitans' plasmid pSG2, complete sequence (GenBank accession number AP008234.1) (Toh *et al.*, 2006); *S. glossinidius* pSG2 plasmid from *G. austeni* (GenBank accession number AJ868436.1) (Darby *et al.*, 2005); and *S. glossinidius* pSG2 plasmid from *G. palpalis palpalis* (GenBank accession number AJ868435.1) (Darby *et al.*, 2005). There were also some sequences from other samples that did not match with any sequences in GenBank and had different lengths of fragments (Table 3.7; Appendix B.1). None of the 130 bp bands were confirmed to be *S. glossinidius* but this was also true for one of the 120 bp bands (the more typical size) from the SaGl sample. Small differences in size of some amplified bands (9/278) were also detected from ZuGa using 3% agarose gels, which also could represent nonspecific amplification.

**Table 3.7** The numbers of pSG2 clones for which *S glossinidius* was confirmed.

Shown are the name of the samples (indicating the site of tsese flies collection and species of tsetse, and the tissue parts of tsetse fly used), the total number of clones that were sequenced, the number of clones for which *S. glossinidius* was confirmed using BLAST and the sequence length of the cloned products (size of inserted fragments).

| Sample | Total number of clones | Sequencing results | |
|---|---|---|---|
| | | Number of *S. glossinidius* positive clones | Size of inserted fragments (bp) |
| ZuGp11HP | 10 | 2 | 120,130* |
| ZuGp125HP | 6 | 0 | 130* |
| MuGp159HP | 6 | 0 | 130* |
| MuGp250HP | 6 | 0 | 130* |
| MuGl25HP | 6 | 3 | 115*, 120, 130*,132* |
| SaGl24HP | 6 | 0 | 120* |
| ZuGa25HP | 6 | 6 | 120 |
| ZuGa43HP | 6 | 6 | 120 |
| ZuGa88HP | 6 | 6 | 120 |
| ZuGa94HP | 6 | 6 | 120 |

* sizes of unidentified PCR products

Sequencing confirmed the presence of *S. glossinidius* from all amplification products using GPO1 (Appendix B.2). Good quality sequences were obtained from all 20 samples, which all were most similar to *S. glossinidius* samples in Genbank: *S. glossinidius* str. 'morsitans' plasmid pSG2, complete sequence (GenBank accession number AP008234.1) (Toh *et al.*, 2006); *S. glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA (GenBank accession number LN887947.1) (Isaac *et al.*, 2013); *S. glossinidius* pSG2 plasmid from *G. austini* (GenBank accession number AJ868436.1) (Darby *et al.*, 2005); and *S. glossinidius* pSG2 plasmid from *G. palpalis palpalis* (GenBank accession number AJ868435.1) (Darby *et al.*, 2005). There was extensive sequence variation found among the 18 consensus fragments that were analysed (34 bp differences among samples). A phylogenetic tree (Figure 3.9) suggested division of the sequences into three groups, consisting of: 1) nine *G. pallidipes*, two *G. austeni* and one *G. brevipalpis*; 2) four *G. austeni* and one *G. pallidipes*; and 3) one sequence from a single *G. brevipalpis*, which showed the highest divergence from other groups. These results reflect genetic variation of *S. glossinidius* among the sites of sample collection and the three species of *Glossina*. There was also variation in the prevalence of trypanosomes in each group (Table 3.8).

For Hem fragments, all amplifications of PCR products were found to be most similar to the genome sequence from *S. glossinidius* str 'morsitans' (GenBank accession number AP008232.1) (Toh *et al.*, 2006) but there was a single bp difference between sequences obtained from four Hem fragments of *G. pallidipes* and four Hem fragments of *G. austeni*. Given the lack of confirmation that all amplified products using pSG2 were *S. glossinidius*, the relatively high agreement between GPO1 and Hem, and the chance that plasmids could not always be present, the nuclear primer set (Hem) was selected as the most appropriate for *S. glossinidius* screening for my tests of assocation.

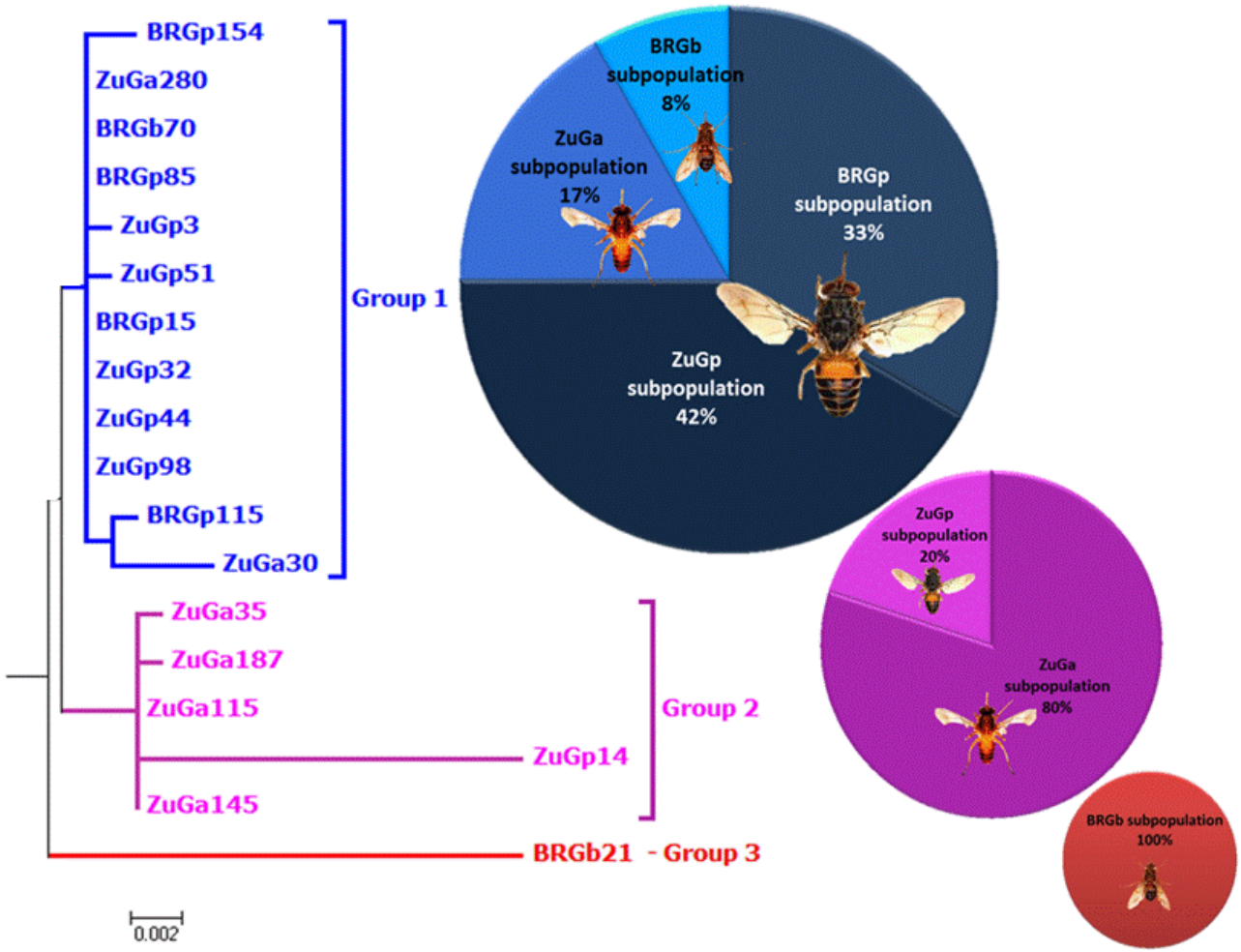



**Figure 3.9** Neighbor-joining tree representation of the GPO1 sequences belonging to *S. glossinidius*

Variants are classified into three groups. Pie charts represent the proportion of tsetse flies from subpopulation that fell into each group. Group 1 consists of *G. brevipalpis*, *G. pallidipes and G. austeni* from Buffalo Ridge and Zungu Luka. Group 2 is composed of *G. austeni* and *G. pallidipes* from Zungu Luka. Group 3 is represented by a single sample of *G. brevipalpis* from Buffalo Ridge.

**Table 3.8** Number of trypanosome positive flies (based on ITS-1 primer identification) found in the three groups of *S. glossinidius* predicted based on GPO1 sequencing

Shown are the number of each species of trypanosome found in each group (in both single and mixed infections) and the total number of tsetse flies that tested positive for trypanosomes.

| Trypanosome screening results | Group1 | Group2 | Group3 |
|---|---|---|---|
| **Single infection** | | | |
| *T. vivax* | 3 | 1 | 0 |
| *T. brucei* | 1 | 0 | 0 |
| *T. congolense* savannah | 2 | 0 | 1 |
| **Mixed infection** | | | |
| *T. congolense* savannah and *T. congolense* kilifi | 1 | 1 | 0 |
| *T. congolense* savannah, *T. congolense* kilifi and *T. simiae* | 0 | 1 | 0 |
| **Total positive** | **7** | **3** | **1** |
| **Total** | **12** | **5** | **1** |

### 3.4.2 Prevalence of *Sodalis glossinidius* in relation to intrinsic factors of tsetse flies and *Trypanosoma spp*. infection

Using the Hem primer screening conditions, overall prevalence of *S. glossinidius* from all 1090 tsetse samples was 34.0%. This varied (Appendix B.3) in relation to tsetse sex, age, species and sampling site. Prevalence of *S. glossinidius* in females was slightly higher than for males (Figure 3.10a). In relation to age, young flies had the lowest *S. glossinidius* prevalence but there was not much variation in relation to age (Figure 3.10b). Nearly all *G. brevipalpis* showed the presence of *S. glossinidius*, in contrast to *G. pallidipes* and *G. austeni* which showed very low prevalence (Figure 3.10c). None of the *G. longipennis* samples were positive for *S. glossinidius*. *Sodalis glossinidius* prevalence in tsetse fly samples from the Shimba Hills was substantially higher than for Nguruman. In the Shimba Hills, most of the tsetse flies from Buffalo Ridge were *S. glossinidius* positive, but this was not the case for tsetse flies from Zungu Luka (Figure 3.10d). For Nguruman, only two *Sodalis* positive flies were identified from Mukinyo and no infected flies were identified from Sampu.

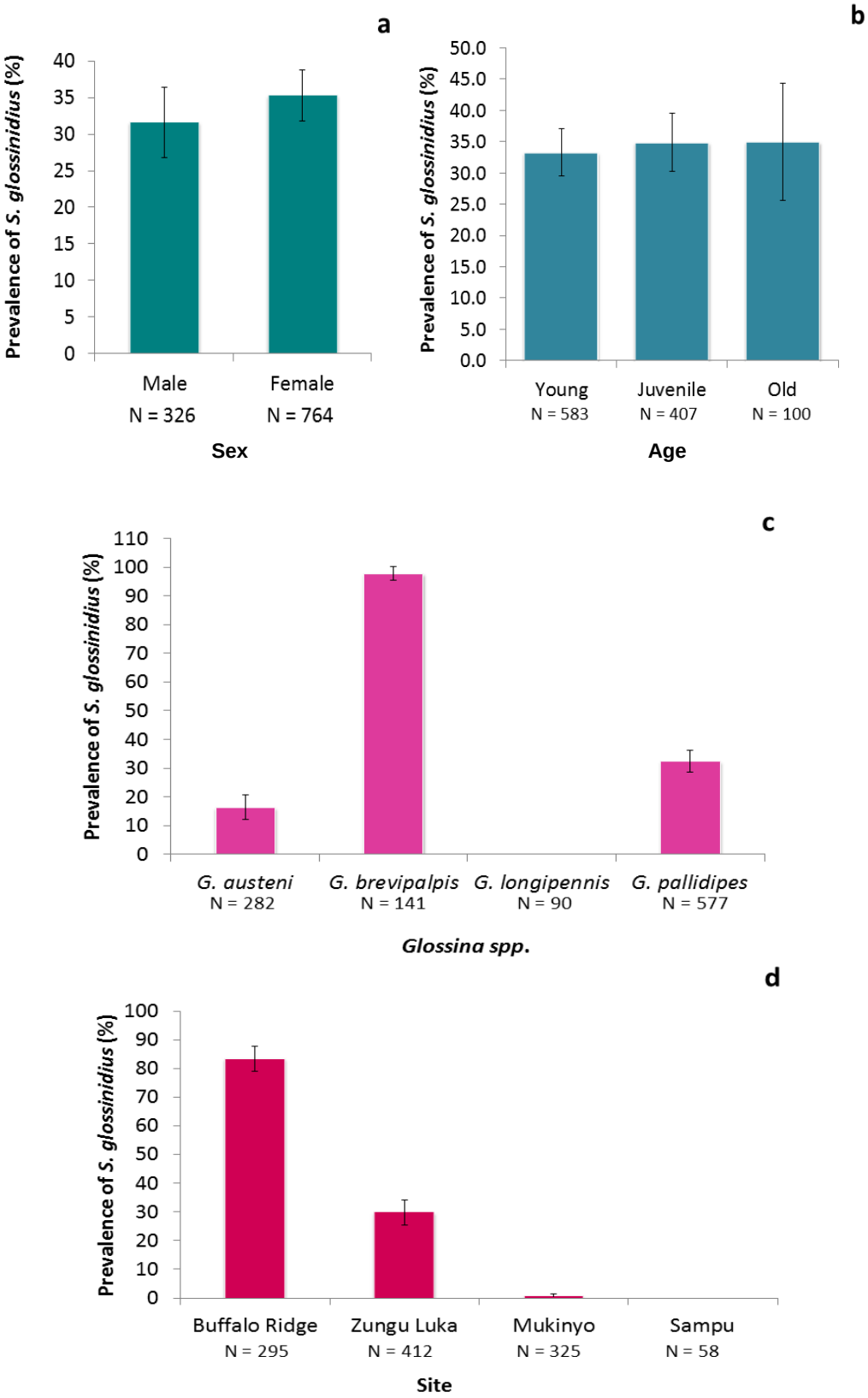



**Figure 3.10** Prevalence of *Sodalis glossinidius* in tsetse flies based on Hem primer screening of 1090 tsetse flies.
Histograms indicate prevalence of the endoparasite by: (a) sex; (b) age; (c) species of tsetse fly (*G. austeni*, *G. brevipalpis*, *G. longipennis* and *G. pallidipes*); and (d) site (Buffalo Ridge, Zungu Luka, Mukinyo and Sampu). 95% confident intervals (CI) are indicated. N is the number of tsetse flies screened in each population.

Across all samples, significant associations were found between *S. glossinidius* harbouring and infection with *Trypanosoma spp*. ($\chi^2$= 75.0, df = 1, P-value < 0.0001) based on $\chi^2$ tests (Figure 3.11a). However, this appears to be driven by the samples that were negative for trypanosomes (Figure 3.11); among samples that were positive for trypanosomes, there was an equal chance of being infected or not with *Sodalis*. The pattern also varied by geographic region. More specifically, samples from the Shimba Hills, which showed high prevalence of *S. glossinidius* infection, showed no significant association with trypanosomes ($\chi^2$= 2.7, df = 1, P-value = 0.1032) (Figure 3.11b). Samples from Nguruman, which showed low prevalence of *S. glossinidius* infection, also showed no significant association ($\chi^2$= 2.9, df = 1, P-value = 0.0889): only two samples tested positive for *Sodalis* and a much higher proportion of samples were negative for trypanosomes compared to the Shimba Hills. So, the apparent patterns again appear to be driven by the negative, rather than the positive samples.

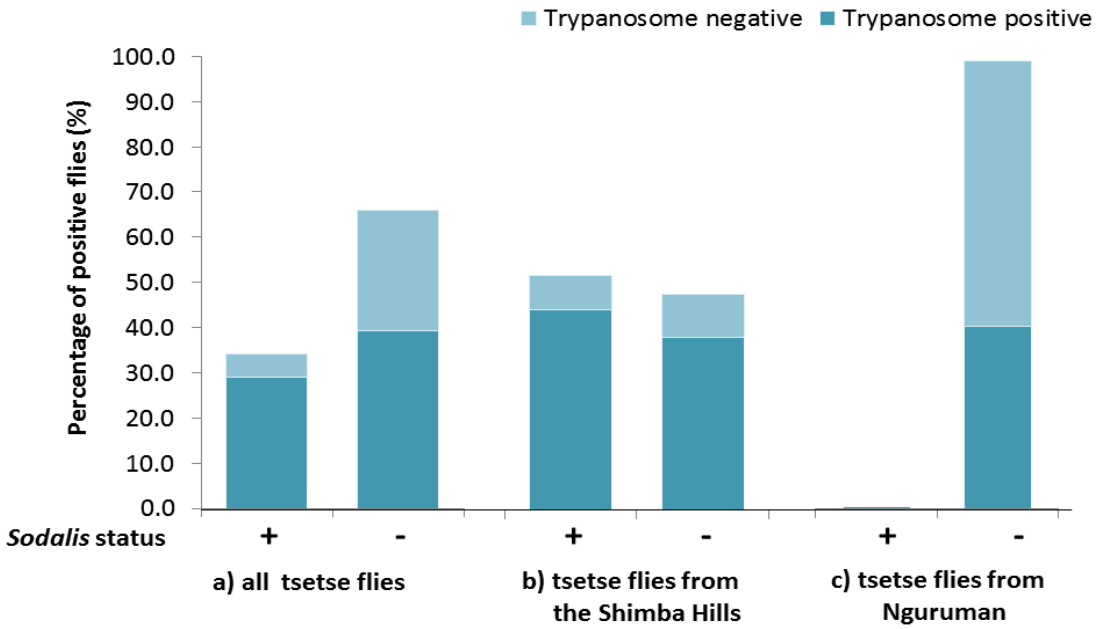



**Figure 3.11** Comparison of *Trypanosoma spp*. and *Sodalis glossinidius* screening results in tsetse flies tested using $\chi^2$ contingency tests.

Histograms indicate the percentage of *Sodalis* positive flies for: (a) all tsetse flies (N = 1090 samples); (b) tsetse flies from the Shimba Hills (N = 707 samples); and (c) tsetse flies from Nguruman (N = 383 samples)

### 3.4.3 Generalised Linear Models of *Sodalis glossinidius* and trypanosome status

Using likelihood ratio tests for model selection in GLM analyses, a significant association between *S. glossinidius* presence or absence and tsetse factors was found only for subpopulation ($\chi^2$ = 730.6, df = 1, P-value < 0.0001) in GLM Model 1 (Table 3.9; Appendix B.4). The predicted direction of *Sodalis* status was higher infection in the Shimba Hills (Figure 3.12), compared to the other site and species combinations. Post hoc tests (Appendix B.4) showed that this was driven predominantly by higher prevalence in the Buffalo Ridge (10/11 comparisons significant, with the nonsignfiicant comparison being with the ZuGp subpopulation) and the Zunga Luka regions (7/8 comparisons, excluding comparisons with Buffalo Ridge). When each species of trypanosome was included as an explanatory variable in separate models, subpopulation was again the only significant factor for both *T. brucei* ($\chi^2$= 730.6, df = 1, P-value < 0.0001) (Model 3; Appendix B.6) and *T. vivax* status ($\chi^2$ = 730.6, df = 1, P-value < 0.0001) (Model 4; Appendix B.7). A significant three-way interaction was found involving sex, age and *T. congolense* status ($\chi^2$ = 6.28, df = 48, P-value = 0.0122) in Model 2 (Appendix B.5). Since only two flies were positive for Sodalis in Nguruman, predicted values could only be fit for somples from the Shimba Hills. *Trypanosoma congolense* positive males of all four subpopulations showed the highest probabilities to be *Sodalis* positive, regardless of age, while predicted values in females decreased with age and old females showed a much lower probability to be *Sodalis* positive than males (Figure 3.13). However this was driven by only a single significant difference (Appendix B.5): old females that were *Sodalis* positive showed lower infections than juvenile males that were *Sodalis* positive.

**Table 3.9** Summary of the best-fitting models for each species of trypanosome when *S. glossinidius* was considered as the response variable (GLM Models 1 – 4).

In addition to the trypanosomes, the full models in each case included subpopulation, sex and age of the tsetse flies. Shown are the explanatory variables and interactions remaining in the best fitting models, along with the degrees of freedom (df), log likelihood (-logLik), Akaike information criterion (AIC), Red indicates interactions that showed significant differences among levels in the post hoc tests. Please see Appendix B.4 – Appendix B.7 for full details of the model selection.

| Model Number | Explanatory variables of best fitting models | Full model | | | Best fitting model | | |
|---|---|---|---|---|---|---|---|
| | | df | - logLik | AIC | df | - logLik | AIC |
| **Model 1** (*Trypanosoma spp.*) | subpopulation | 55 | 316.6 | 743.2 | 7 | 333.7 | 681.4 |
| **Model 2** (*T. congolense*) | subpopulation + sex * age * *T. congolense* | 52 | 313.3 | 730.7 | 14 | 325.5 | 679.0 |
| **Model 3** (*T. brucei*) | subpopulation | 50 | 316.1 | 732.1 | 7 | 333.7 | 681.4 |
| **Model 4** (*T. vivax*) | subpopulation | 54 | 319.1 | 746.3 | 7 | 333.7 | 681.4 |

A * B means interaction between A and B.

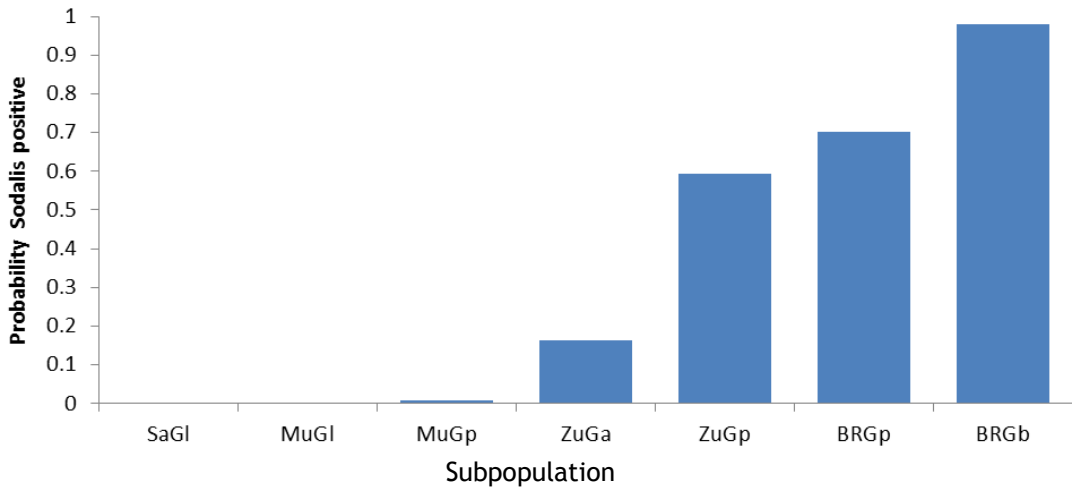



**Figure 3.12** Probabilities of *S. glossinidius* positive status in relation to subpopulation (the only significant factor in the best-fitting model) for Model 1.

Since there also was no influence of trypanosomes in Models 3 or 4 these would reduce to the same as Model 1. BRGp: *G. pallidipes* from Buffalo Ridge; ZuGp: *G. pallidipes* form Zungu Luka; MuGp: *G. pallidipes* form Mukinyo; BRGb: *G. brevipalpis* from Buffalo Ridge; ZuGa: *G. austeni* from Zungu Luka; MuGl: *G. longipennis* from Mukinyo; and SaGl; *G. longipennis* from Sampu.

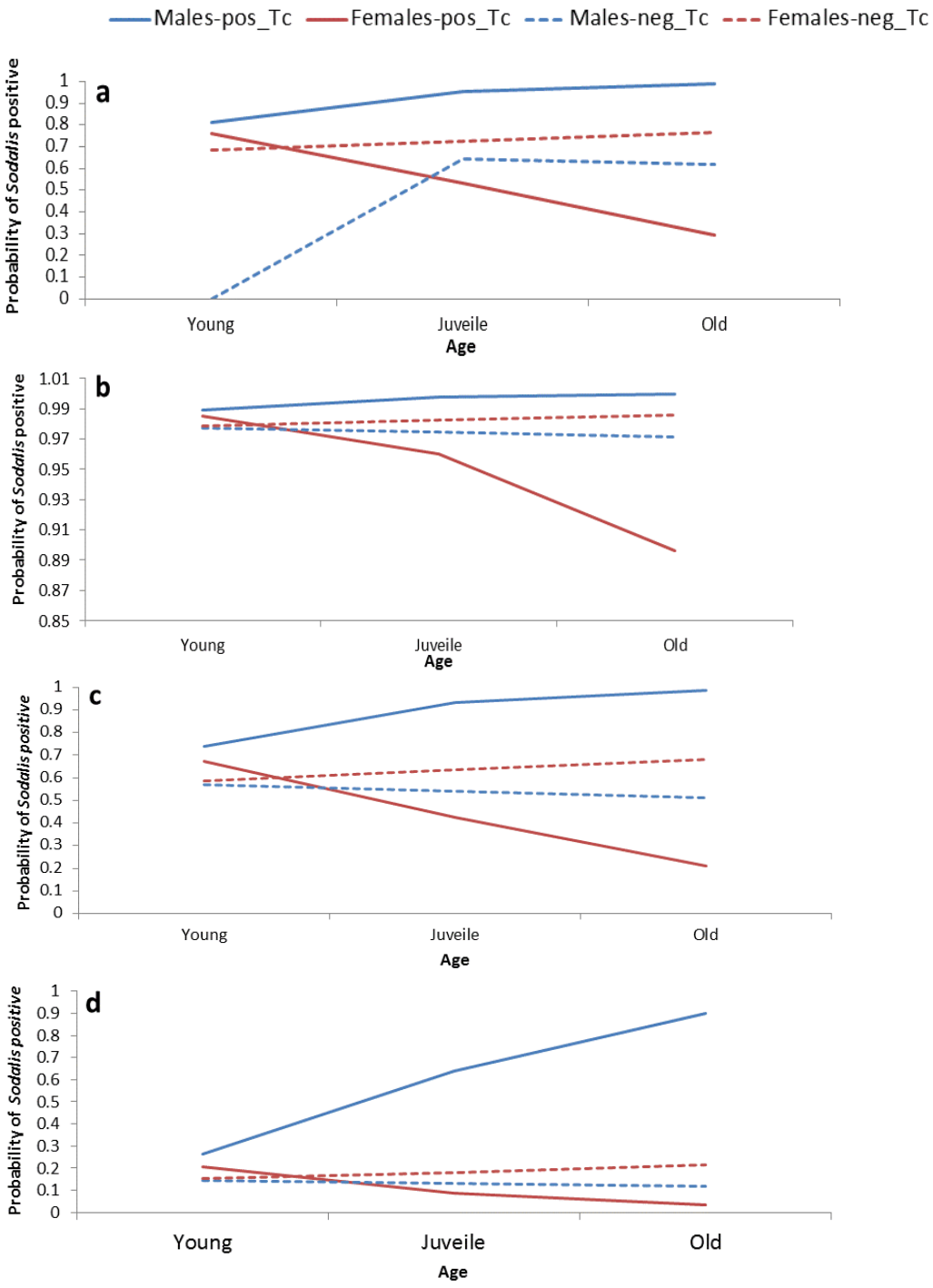



**Figure 3.13** Probabilities of *S. glossinidius* positive status in tsetse samples when analysed with *T. congolense* status (Tc) as an explanatory variable (Model 2)

Predicted values are shown for the four subpopulations from the Shimba Hills (because only two individuals from Nguruman were *Sodalis* positive: (a) *G. pallidipes* from Buffalo Ridge; (b) *G. brevipalpis* from Buffalo Ridge; (c) *G. pallidipes* from Zungu Luka; and (d) *G. austeni* from Zungu Luka.

Using all trypanosomes as the response variable, GLM Model 5 showed that age ($\chi^2$ = 4.65, df = 14, P-value = 0.0310) and a two-way interaction between subpopulation and sex ($\chi^2$= 18.94, df = 10, P-value = 0.0043) were significantly associated with presence of trypanosomes in tsetse flies (Table 3.10; Appendix B.8), with no effect of *Sodalis* status. It was clear that probabilities of trypanosomes presence in tsetse flies in all subpopulations increased when age increased, and some but not all subpopulations showed higher numbers of trypanosome present in females than in males (Figure 3.14). Males of the BRGp, ZuGp and SaGl subpopulations were predicted to show higher trypanosome risk than females. When considered separately, *T. vivax* status was significantly associated with significant interactions between subpopulation and sex ($\chi^2$= 18.90, df = 11, P-value = 0.0043) and between sex and age ($\chi^2$= 7.52, df = 22, P-value = 0.0059), without involving *Sodalis* status (Model 8; Appendix B.9), have decreasing probability of *T. vivax* presence in females (Figure 3.15). None of the pairwise comparisons using post hoc tests (Appendix B.9) were significant for the sex by age interaction, but multiple significant differences were found in relation to sex and subpopulation, all involving flies sampled from the Zungu Luka region. When *T. congolense* (Model 6; Appendix B.10) was analysed on its own as the response variable, a three-way interaction was found but this time involving *Sodalis*: an interaction among sex, age and *Sodalis* status, ($\chi^2$ = 6.84, df = 41, P-value = 0.0089). The probability of *T. congolense* presence was predicted to increase for *Sodalis* positive males with increasing age, but to decrease for *Sodalis* positive females (Figure 3.16). On the other hand, absence of *Sodalis* in females tended to increase probabilities of *T. congolense* presence with age, but decreased for *Sodalis* negative males. However for *T. brucei* (Model 7; Appendix B.11), the results were even more complicated, with subpopulation, sex, age, *Sodalis* status factors and three different three-way interactions involving *Sodalis* status: a three-way interaction among sex, age and *Sodalis* status ($\chi^2$= 7.94, df = 41, P-value = 0.0048); a three-way interaction among subpopulation, age and *Sodalis* status ($\chi^2$= 13.84, df = 39, P-value = 0.0031); and a three-way interaction among subpopulation, sex and *Sodalis* status ($\chi^2$= 9.27,df = 40, P-value = 0.0097). Although it is not possible to interpret the biology of such complex interactions, predicted values could be used to demonstrate the overall directions of associations (Figure 3.17). For example, both species of tsetse flies (*G. pallidipes* and *G. brevipalpis*) from Buffalo Ridge tended to show decreased

probabilities of *T. brucei* presence with age but this varied by sex and *Sodalis* status: males showed a higher probability of *T. brucei* presence than females and lower values in older flies, but an association with age was only found for females that lacked *Sodalis*. In contrast, even though *G. pallidipes* was also found at Zungu Luka, different patterns of relationships were found among sex, age and presence of Sodalis in relation to probabilities of presence of *T. brucei*.

**Table 3.10** Summary of the best-fitting models when presence or absence of trypanosomes was considered as the response variable (GLM models 5 – 8).

In addition to *Sodalis* status, the full models in each case included subpopulation, sex and age of the tsetse flies. See Table 3.8 and Appendix B.8 - Appendix B.11 for details. Red indicates interactions that showed significant differences among levels in the post hoc tests. Red indicates interactions that showed significant differences among levels in the post hoc tests.

| Model Number | Explanatory variables of best fitting models | Full models | | | Best Fitting models | | |
|---|---|---|---|---|---|---|---|
| | | df | - logLik | AIC | df | - logLik | AIC |
| **Model 5** (*Trypanosoma spp.*) | age + subpopulation * sex | 44 | 588.0 | 1464 | 15 | 701.3 | 1432.5 |
| **Model 6** *(T. congolense)* | sex * *Sodalis* status * age + subpopulation | 44 | 412.7 | 913.4 | 8 | 434.7 | 885.35 |
| **Model 7** *(T. brucei)* | subpopulation * sex * age + subpopulation * sex * *Sodalis* status + subpopulation * age * *Sodalis* status and sex * age * *Sodalis* status | 44 | 250.4 | 588.8 | 42 | 250.4 | 584.8 |
| **Model 8** (*T. vivax)* | subpopulation + sex * age | 44 | 562.1 | 1212.2 | 16 | 572.1 | 1176.3 |

A * B means interaction between A and B.

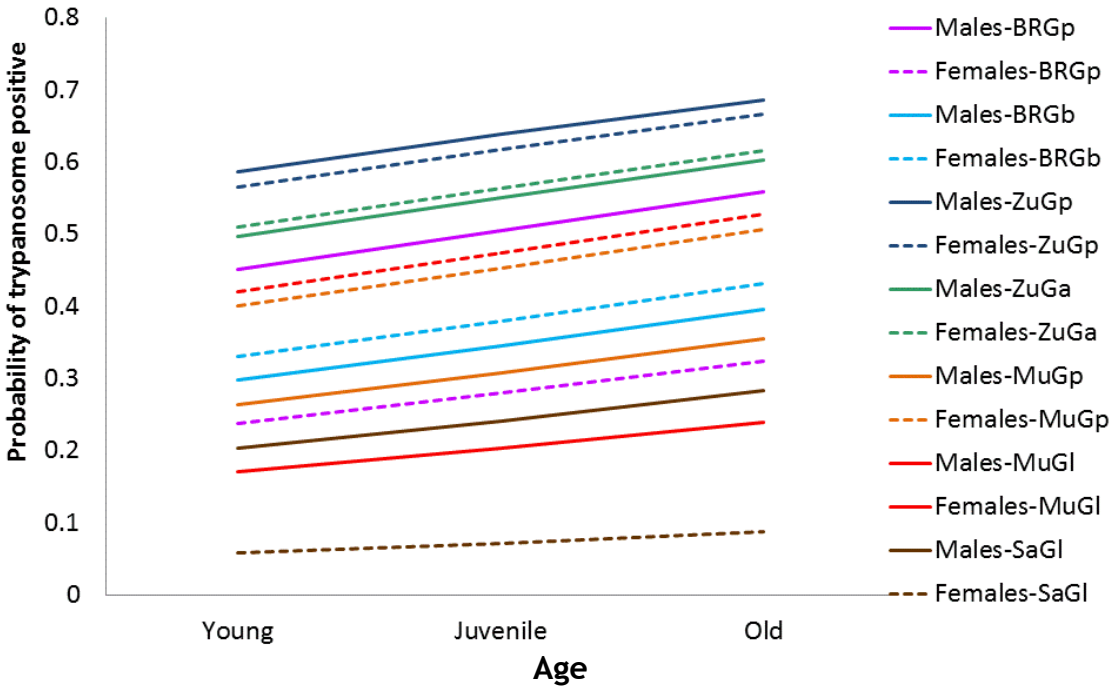



**Figure 3.14** Probabilities of trypanosome presence in tsetse samples from the best-fitting model (Model 5).

Predicted values are shown for males and females in the seven subpopulations. Note that while all lines are roughly parallel with respect to age, whether the probability is higher in males or in females differs by subpopulation. BRGp: *G. pallidipes* from Buffalo Ridge; ZuGp: *G. pallidipes* form Zungu Luka; MuGp: *G. pallidipes* form Mukinyo; BRGb: *G. brevipalpis* from Buffalo Ridge; ZuGa: *G. austeni* from Zungu Luka; MuGl: *G. longipennis* from Mukinyo; and SaGl; *G. longipennis* from Sampu.

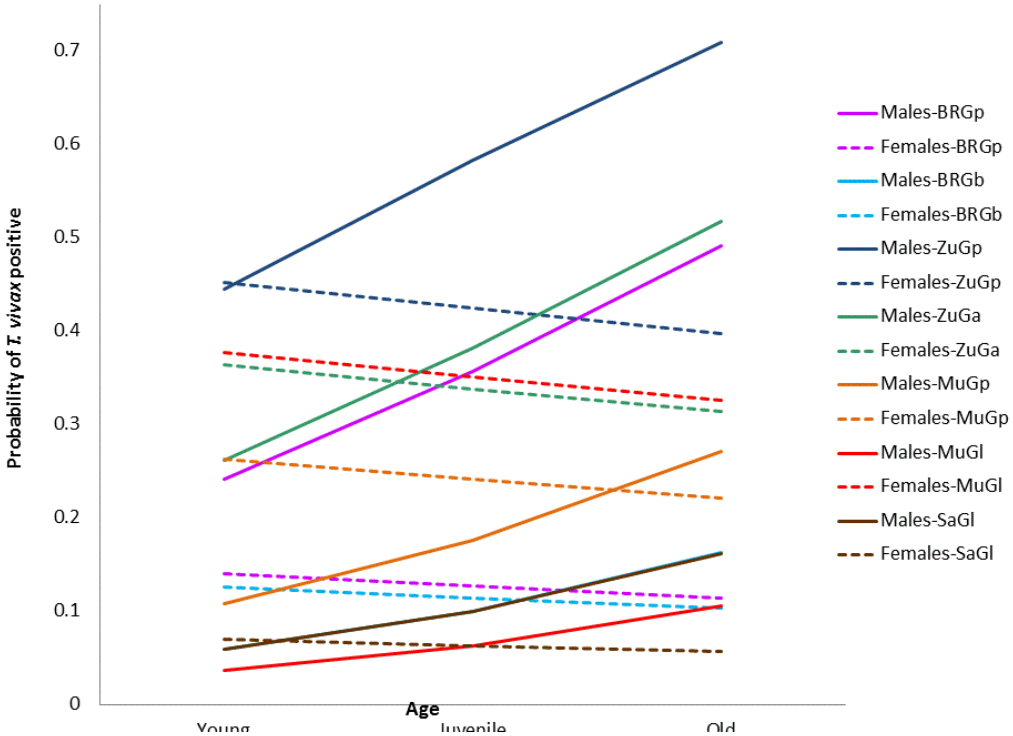



**Figure 3.15** Probabilities of *T. vivax* presence in tsetse samples from the best-fitting model (Model 8).

Predicted values are shown for males and females in the seven subpopulations. BRGp: *G. pallidipes* from Buffalo Ridge; ZuGp: *G. pallidipes* form Zungu Luka; MuGp: *G. pallidipes* form Mukinyo; BRGb: *G. brevipalpis* from Buffalo Ridge; ZuGa: *G. austeni* from Zungu Luka; MuGl: *G. longipennis* from Mukinyo; and SaGl; *G. longipennis* from Sampu.

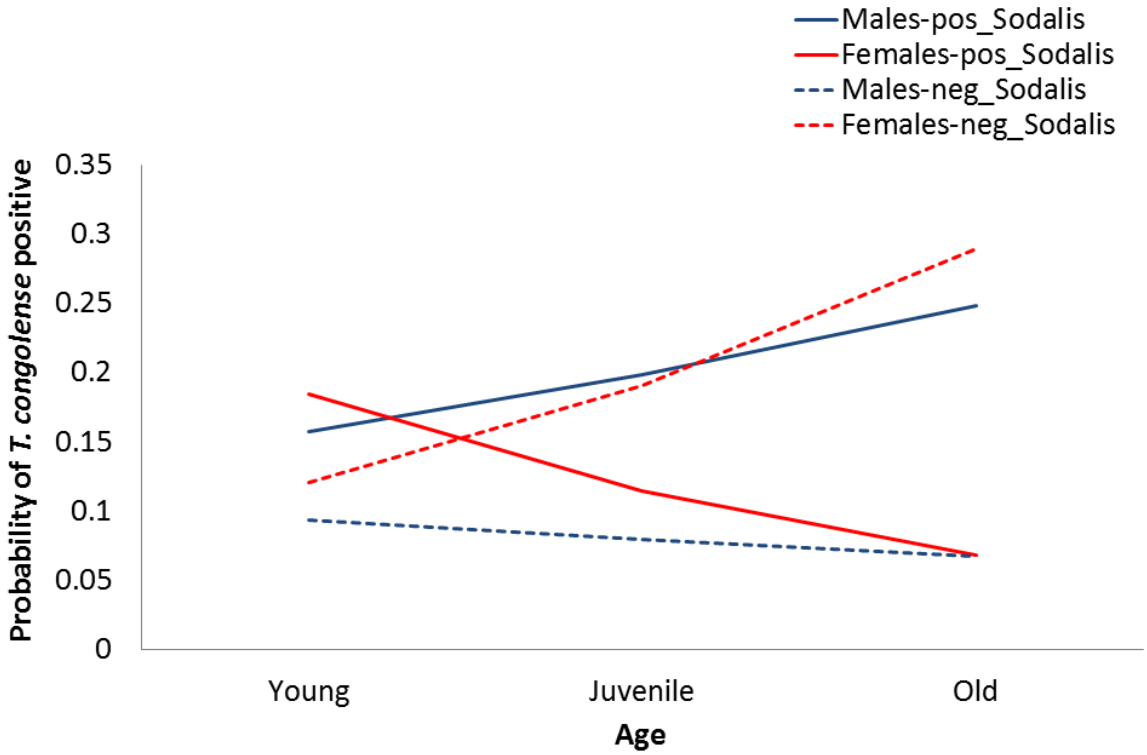



**Figure 3.16** Probability of *T. congolense* presence in tsetse samples from the best-fitting model (Model 6).

Predicted values are shown for males and females that are positive or negative for *Sodalis*, averaged across subpopulations.

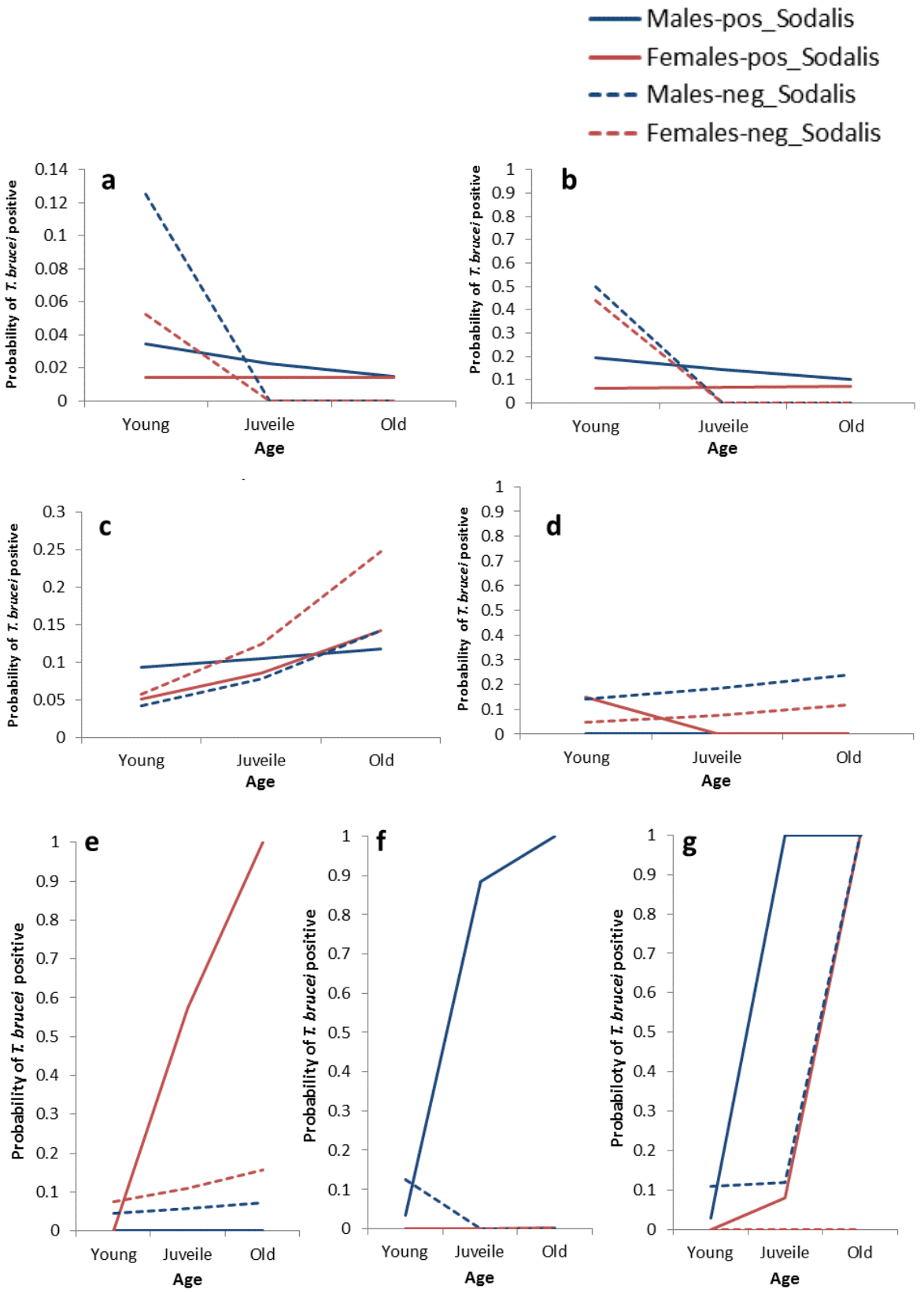



**Figure 3.17** Probability of *T. brucei* presence in tsetse samples from the best-fitting model (Model 7).

Predicted values are shown for males and females in the seven subpopulations: (a) *G. pallidipes* from Buffalo Ridge. (b) *G. brevipalpis* from Buffalo Ridge. (c) *G. pallidipes* from Zungu Luka. (d) *G. austeni* from Zungu Luka. (e) *G. pallidipes* from Mukinyo. (f) *G. longipennis* from Mukinyo. (g) *G. longipennis* from Sampu.

### 3.4.4 Multiple correspondence analysis among status of *Sodalis* and trypanosome and intrinsic factors of tsetse flies

Given the complexity of interactions found in the GLM, two-dimensional MCA 1 analyses were used to more clearly demonstrate relationships among subpopulation, sex, age, trypanosome and *Sodalis* status (Figure 3.18 -Figure 3.19). A strong correlation between *Sodalis* status and subpopulation was represented in the first principle component (dimension 1; Table 3.11; Figure 3.18). The majority of tsetse flies from Nguruman (MuGp, MuGl and SaGl) and the ZuGa subpopulation were *Sodalis* negative (Cluster 1) when screened with the Hem primers while tsetse flies from the Shimba Hills (BRGp, BRGb and ZuGp) tended to have more individuals that were *Sodalis* positive (Cluster 2). A correlation between subpopulation and sex, which was weakly related to trypanosome status, was found in the second principle component (dimension 2; Table 3.11; Figure 3.18). Males, especially the old flies in five of the subpopulations (BRGp, BRGb, MuGp, MuGl and SaGl), tended to be trypanosome negative (Cluster 3). On the other hand, there was a tendency for females in ZuGp and ZuGa to be trypanosome positive (Cluster 4). The third principle component showed a correlation among trypanosome status, subpopulation and age (dimension 3; Table 3.11; Figure 3.19). Young flies in BRGp, MuGp and BRGb tended to be trypanosome negative (Cluster 1) while many tsetse flies in ZuGp were trypanosome positive (Cluster 2). Overall, these results confirmed the conclusions based on the GLM analysis that *Sodalis* was not a strong driving factor in the prevalence of trypanosomes when considered across species but that subpopulation (and so tsetse intrinsic factors) did explain a substantial amount of the variation in *Sodalis* prevalence.

Similarly, the complicated associations of each trypanosome species with subpopulation, sex, age and *Sodalis* status in tsetse flies were confirmed in the multivariate analyses (Figure 3.20). Presence of *Sodalis* was still most strongly influenced by subpopulation (Table 3.12), driving the differentiation along dimension 1 (Clusters 1 and 2). Subpopulation and sex were most separated along dimension 2, with a weak correlation with *T. vivax* presence (Table 3.12; Figure 3.20). Correlation of *T. vivax* infection with subpopulation, sex and age (Cluster 3) was distinct from *T. brucei* and *T. congolense* (Cluster 4) in the MCA graph 2.

**Table 3.11** Adjusted eta-squared for the combination of variables in dimensions 1 - 3 in MCA 1.

| Variables | Dimension 1 | Dimension 2 | Dimension 3 |
|---|---|---|---|
| *Sodalis*_status | 0.863 | 0.009 | 0.000 |
| subpopulation | 0.870 | 0.586 | 0.398 |
| sex | 0.020 | 0.537 | 0.025 |
| age | 0.000 | 0.210 | 0.385 |
| Trypanosome_status | 0.003 | 0.158 | 0.351 |

**Table 3.12** Adjusted eta-squared for the combination of variables in dimensions 1 - 3 in MCA 2.

| Variables | Dimension 1 | Dimension 2 | Dimension 3 |
|---|---|---|---|
| *Sodalis*_status | 0.860 | 0.005 | 0.005 |
| subpopulation | 0.867 | 0.585 | 0.423 |
| sex | 0.018 | 0.543 | 0.018 |
| age | 0.000 | 0.203 | 0.058 |
| *T. congolense* | 0.017 | 0.009 | 0.255 |
| *T. brucei* | 0.000 | 0.003 | 0.154 |
| *T. vivax* | 0.001 | 0.160 | 0.307 |

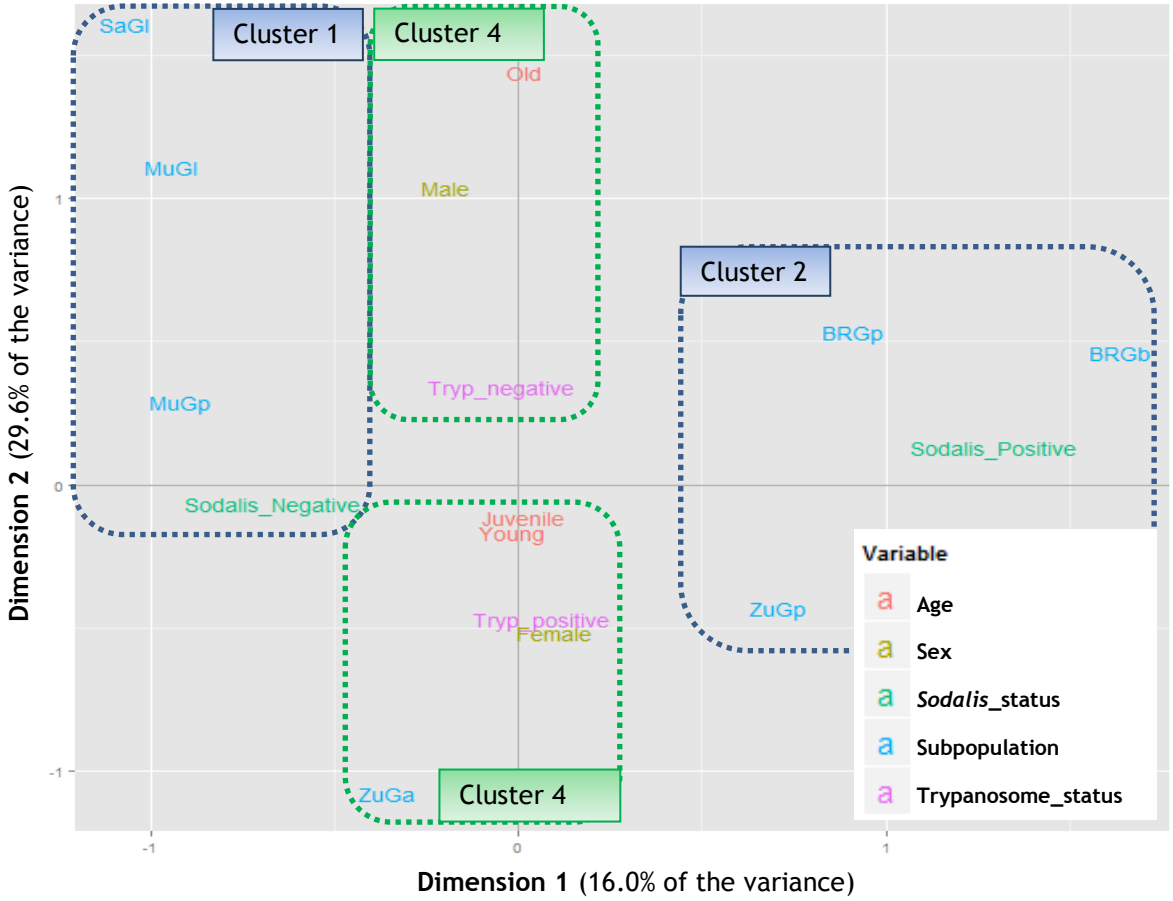



**Figure 3.18** Dimension 1 and 2 of the MCA 1 for explaining relationships between the trypanosome status of tsetse flies and their intrinsic factors, as implemented in the R package FactoMineR statistical software programme.

Sex (male and female), age (young, juvenile and old) and Sodalis status (negative and positive) of the seven subpopulations (BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu) were plotted for describing associations with the trypanosome status of tsetse flies (negative and positive).

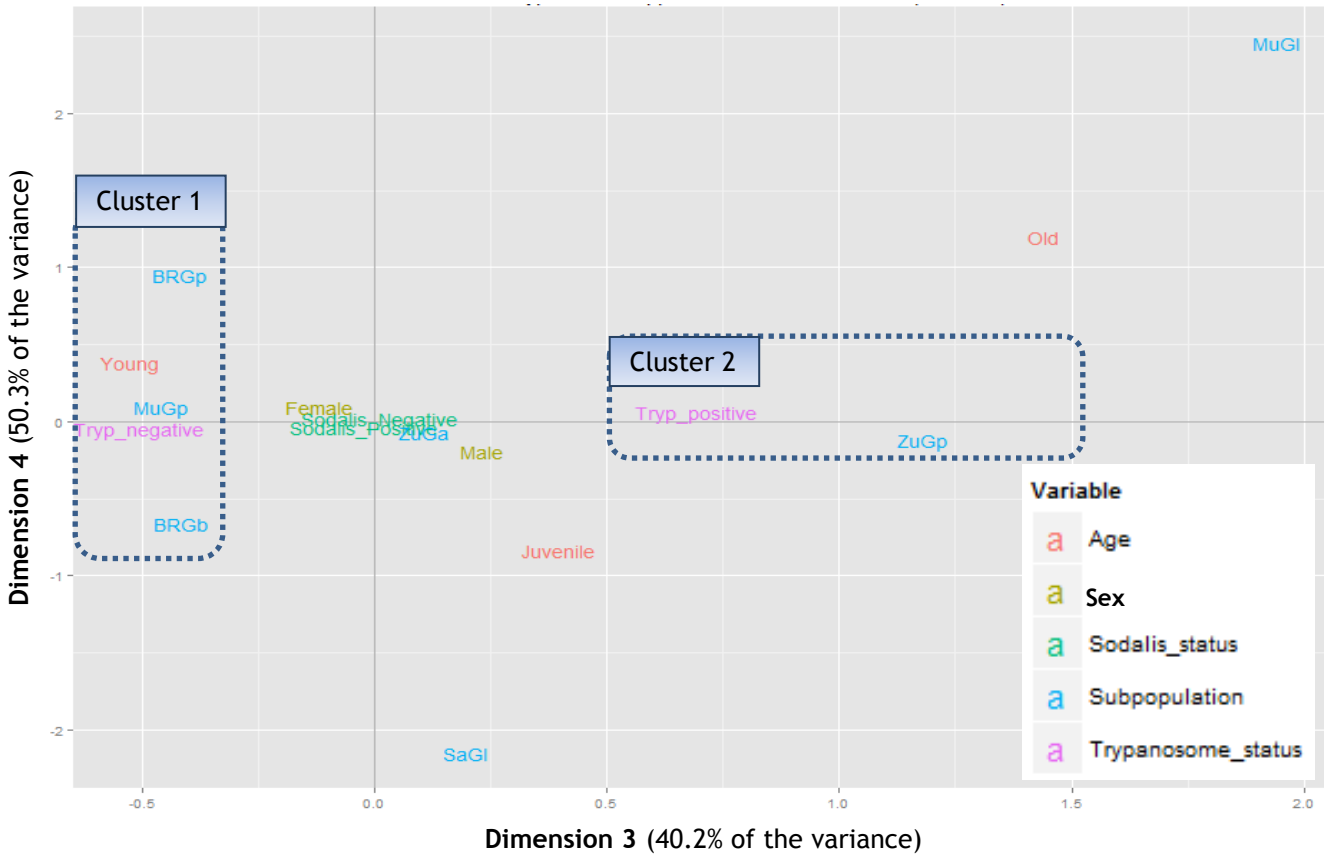



**Figure 3.19** Dimension 3 and 4 of the MCA 1 for explaining relationships between the trypanosome status of tsetse flies and their intrinsic factors.

Sex (male and female), age (young, juvenile and old) and *Sodalis* status (negative and positive) of the seven subpopulations (BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu) were plotted for describing associations with trypanosome status of tsetse flies (negative and positive).

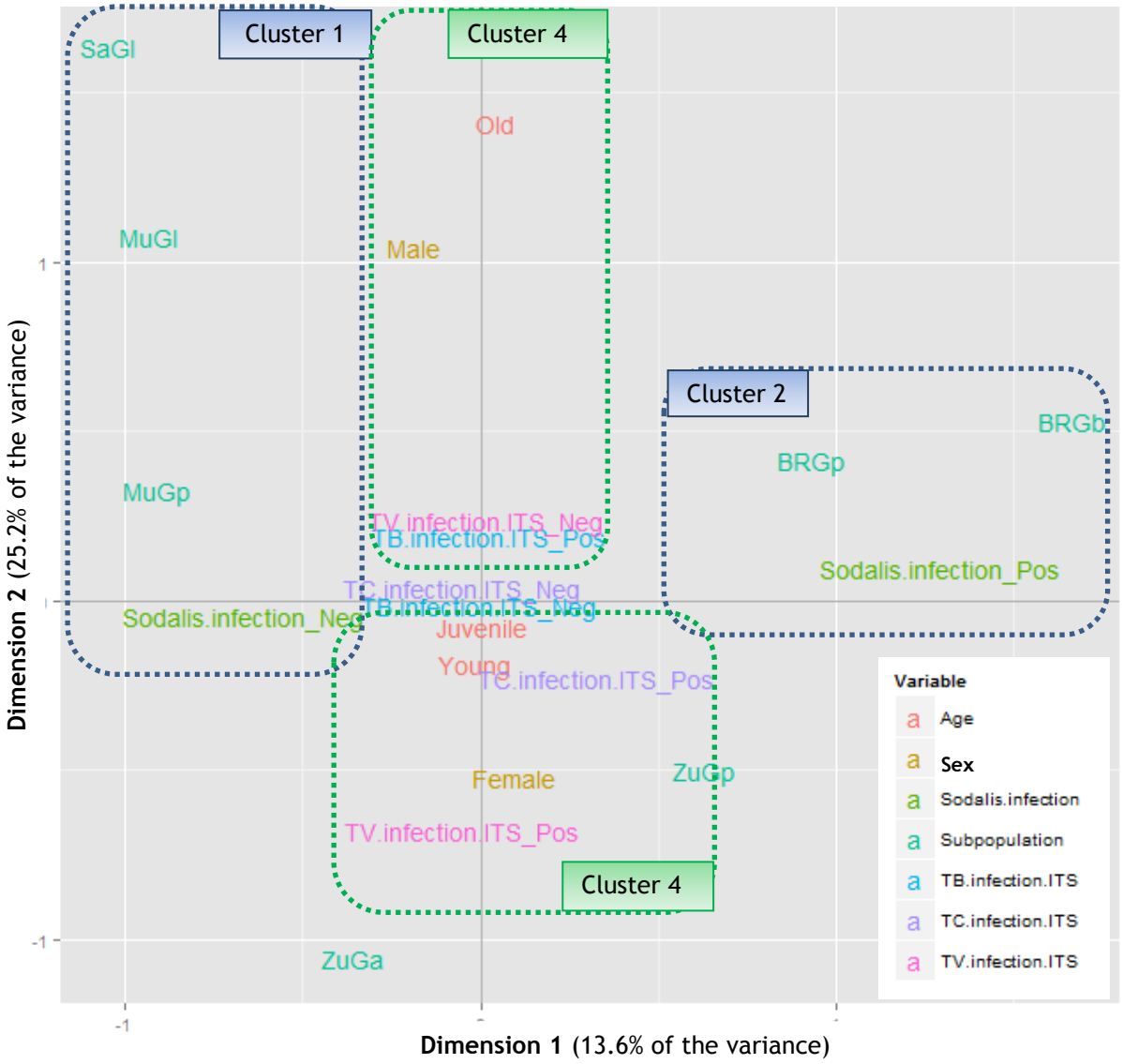



**Figure 3.20** Associations of dimension 1 and 2 of MCA 2 for for explaining relationships among infections of *Trypanosoma spp.*, *S. glossinidius* and instrinsic factors of tsetse flies.

Infections of the three main pathogenic *Trypanosoma spp*. (*T. congolense* (TC): *T. brucei* (TB):*T. vivax* (TV)) and infection of *S. glossinidius* of 1090 tsetse flies were analysed with the intrinsic tsetse variables (subpopulation, sex and age). (BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu)

## 3.5 Discussion

### 3.5.1 Optimisation of PCR screening methods for *Sodalis glossinidius* in tsetse flies

Using all three sets of screening primers, *S. glossinidius* was detected in both abdomen and head plus proboscis parts, which was consistent with previous reports that *S. glossinidius* is found in many tissues, such as the midgut (Aksoy, 1995a, Dale and Maudlin, 1999, Cheng and Aksoy, 1999), haemolymph, muscle, fat bodies, salivary glands (Cheng and Aksoy, 1999), milk glands (Attardo *et al.*, 2008), and first stage of larvae in the uterus (Balmand *et al.*, 2013). However, this study showed that primer selection and PCR sequencing were critical for determination of *S. glossinidius* prevalence. Although the GPO1 and Hem primers were developed for amplification of different DNA regions of *S. glossinidius* in tsetse flies from different geographic locations, the screening results were more correlated with one another than the screening results of the pSG2 primers, which amplified a fragment of the same plasmid of *S. glossinidius* as the GPO1 primers. The 3.3% difference in apparent prevalence based on GPO1 and Hem screening could be due to variation in the priming sites among different strains of *S. glossinidius* but is also within the range of expected PCR errors if there is only a low quantity of the bacteria present. The pSG2 primers were very sensitive but a number of PCR bands whose identity could not be confirmed by sequencing were also detected, which were of similar size as the positive bands confirmed to be *S. glossinidius*. Previous studies using the pSG2 primers for investigation of *Sodalis* in tsetse flies, did not consider sequencing of PCR products (Farikou *et al.*, 2010a) but this is necessary to resolve the source of the differences in conclusions that might be reached using different regions of DNA. The small size of the PCR products for pSG2 made it difficult to sequence but it also limits identification of genetically distinct types. The larger discrepancies between the pSG2 and the other primers are indicative of either variation among the *S. glossinidius* strains present in different regions or nonspecific amplification of another bacterial species. Diversity of bacterial communities in *G. palpalis palpalis, G. pallicera, G. nigrofusca* and *G. caliginea* from Cameroon based on PCR (Geiger *et al.*, 2011), *G. fuscipes fuscipes* from Kenya (Lindh and Lehane, 2011) and *G. f. fuscipes, G. m. morsitans,* and *G. pallidipes* from Uganda (Aksoy *et al.*, 2014) based on qualitative PCR (qPCR) have been

reported. Moreover, genetic variation of *S. glossinidius* has been found samples taken from *G. palpalis gambiensis, G. m. morsitans* (Geiger *et al.*, 2005a), and *G. p. palpalis* (Farikou *et al.*, 2011b). There could thus be a higher risk of false positives than for the other primers, which were confirmed to amplify only *S. glossinidius*.

GPO1 produced the largest fragment and has the potential to reveal the most differences among genotypes. Sequence variation was revealed among subpopulations of tsetse flies, which agreed with previous studies that genetic diversity of *S. glossinidius* depended on species of tsetse hosts (Geiger *et al.*, 2005a, Farikou *et al.*, 2011b). The 18 GPO1 sequences in my analysis of genetic variation could be divided into three different groups, which tended to be associated with particular species of tsetse flies. Group 1 were predominantly sequences of *Sodalis* obtained from *G. pallidipes*, whereas group 2 were mostly from *G. austeni*. Although sample sizes were small, infection patterns with trypanosomes also tended to vary between groups (Table 3.8). Previous studies have also found an association between genetic variation in *Sodalis* and species of tsetse flies. For example, using amplified fragment length polymorphism (AFLP) analysis, the presence of specific genotypes of *S. glossinidius* was linked to the presence of *T. b. gambiense* and *T. b. brucei* infection in wild *G. palpalis palpalis* from Burkina Faso that show variation in vector competence (Geiger *et al.*, 2007). It is thus possible that particular strains of *Sodalis* or specific genotypes might have different effects on trypanosome infection. As even the most variable gene I compared (GPO1) did not fully resolve relationships among sequences, comparison of whole genomes could be a worthwhile approach for assessing the extent of genetic variation in *S. glossinidius* and whether this affects vector competence.

Also lacking in previous studies has been consistent use of a single primer set to draw conclusions about *S. glossinidius* presence; my results clearly demonstrate that differences in sensitivities of various primers could be responsible for differing conclusions among studies. GPO1 is located on extrachromosomal DNA (pSG2 plasmid), so there is a risk that lack of amplification could be due to missing plasmids. This plasmid has been assumed to occur in all wild *Sodalis* positive tsetse flies but although this conclusion was only based on a survey of nine flies (Darby *et al.*, 2005). In the same study, none of the plasmids

associated with *Sodalis* (pSG1, pSG2 pSG3 and pSG4) were detected in *S. glossinidius* negative flies but pSG3 was present in only 7/9 *S. glossinidius* positive flies. There is thus a risk when using primers that target plasmids that they might not be present in all bacterial cells. I thus recommend the use of the Hem primers for screening of *S. glossinius* because they provide reliable amplification of a nuclear gene region. However, the GPO1 primers since the screening results were very similar to Hem, the small risk of false negatives using plasmid-based primers might be compensated for by the higher amount of sequence variation detected in GPO1 if genetic variation among *Sodalis* is the key question to be addressed (rather than just presence or absence). More sequencing is necessary to determine the extent of such variation in relation to tsetse species, geographic location and trypanosome asssociations but my results suggest that this could be worthwhile to investigate. Moreover, different host feeding could affect this variation (see chapter 4).

### 3.5.2 Prevalence of *Sodalis glossinidius* in relation to intrinsic factors of tsetse flies and *Trypanosoma spp*. infection

Although *Sodalis* is involved in the development (Smith *et al.*, 2013) and immune defence mechanism of tsetse flies (Dale and Welburn, 2001, Hao *et al.*, 2001, Lehane *et al.*, 2004, Rose *et al.*, 2014), almost all flies from the Nguruman regions were *Sodalis* negative. This suggests that *Sodalis* is not an essential endosymbiont and its distribution might be related to the local environment of each region, such as type of blood meal hosts and ecological conditions. Prevalence of *S. glossinidius* was not different between males and females; however, it slightly increased with age of tsetse flies. This is consistent with a report that *S. glossinidius* is horizontally transmited via mating (De Vooght *et al.*, 2015) and permanently infects tsetse flies, so it should accumulate in older flies (Rita *et al.*, 2006). Nevertheless, Warmwiri *et al.* (2014) found that the number of *S. glossinidius* positive *G. pallidipes* slightly decreased with increasing age when midguts of flies were screened using GPO1 primers. It is possible that the bacterium was present at too low a density to amplify, rather than complete absence. Density of *S. glossinidius* is associated with that of other endosymbionts in tsetse flies, such as *Halomonas, Pseudomonas, Stenotrophomonas, Pseudomondaceae, Rhodopsedomonas, Wigglesworthia, Shewanella* and *Enterobacteriae* (Aksoy *et al.*, 2014). It is possible that

competition with other bacterial endosymbionts in limited areas of the tsetse body could lead to decrease of *S. glossinidius* density, which would reduce detectability using the PCR screening method.

Nevertheless, based on standard PCR, *S. glossinidius* prevalence showed a strong pattern of variation in relation to sites of sampling and species of tsetse flies, as has been found for other symbionts, such as *Wigglesworthia spp*. (Aksoy, 2000) and *Wolbachia spp*. (Cheng *et al*., 2000). Environmental circumstance, such as humidity and available host communities, could influence *S. glossinidius* distribution. Tsetse flies from Buffalo Ridge and Zungu Luka in the Shimba Hills, which is close to coastal areas, had consistently higher prevalence of *S. glossinidius* than Mukinyo and Sampu in Nguruman. Buffalo Ridge is classified as a thicket type of vegetation while Zungu Luka is a woodland site and Mukinyo and Sampu are both savannah. Different vegetation types of each site of sampling could influence the *S. glossinidius* distribution by providing different host communities, which will be investigated in chapter 4. For example, the growth rate of *S. glossinidius* could possibly have varied by the relative availability of nutrients in host blood from different types of wildlife, as suggested by a laboratory experiment that found that *S. glossinidius* required 20% faetal bovine serum (FBS) and 10% packed horse blood for cultivation (Matthew *et al*., 2005). Since I did not specifically examine differences in environmental parameters between sites, I cannot resolve the reasons for the much lower prevalence of *Sodalis* in Nguruman compared to the Shimba Hills but it would be interesting to more explicitly test the role that differences in which hosts are fed on might play in determining colonisation of the flies by the endosymbionts.

There could also be differences in the relative ability of different species of tsetse flies to support *Sodalis*. Consistent with my results, previous studies have also found variation in prevalence of *S. glossinidius* among *G. brevipalpis*, *G. pallidipes*, *G. austeni* and *G. longipennis* (see references in Table 3.13). In my study, almost all *G. brevipalpis* were infected with *S. glossinidius* but there was more variation among other species. Across studies for *G. pallidipes, S. glossinidius* shows a wide range of prevalence based on different primer combinations, geographic regions and years of sampling (Table 3.13). A moderate rate of *S. glossinidius* prevalence in *G. pallidipes* was detected in my

investigation based on Hem primers but another study in the Shimba Hills Kenya found much lower presence based on GPO1 primers (Wamwiri *et al*., 2013). Moreover, although low prevalence was found using primers targeting the Gro*EL* gene in Zambia (Dennis *et al*., 2014) and in Zimbabwe (Matthew, 2007), high prevalence was revealed in Tanzania using the same gene (Matthew, 2007). For *G. austeni*, a previous study in Kenya using the GPO1 primers (Wamwiri *et al*., 2013) also focused on the Buffalo Ridge region as well as Arabuko-Sokoke but found much lower prevalence than in my current study. I did not detect *Sodalis* positive *G. longipennis* and there has been no previous study of *S. glossinidius* prevalence in this fly species. Together, these results suggest that *G. brevipalpis* is a highly suitable host for the secondary endosymbionts, whereas *G. longipennis* might not support their growth. However, experimental infection of *S. glossinidius* in different species of tsetse flies in the same surrounding environment should be conducted to test whether the apparent differences were due to differences in host suitability or factors related to differences in geographic distributions of the host species. In order to differentiate sensitivities of the primers used to detect *S. glossinidius* in each population, the primers should be applied in *Sodalis* positive samples from different populations.

**Table 3.13** Comparision of *S. glossinidius* prevalence in *G. brevipalpis*, *G. pallidipes*, *G. austeni* and *G. longipennis* from different geographic regions.

| Prevalence (%) | Year of sampling | Primer Set | Sample sites | Reference |
|---|---|---|---|---|
| ***G. brevipalpis*** | | | | |
| 97.9 | 2012 | Hem | Kenya | this study |
| 93.7 | UI | Gro*EL* | Zambia | Dennis *et al.*,(2014) |
| 100 | UI | | Tanzania | Matthew (2007)[a] |
| ***G. pallidipes*** | | | | |
| 32.4 | 2012 | Hem | Kenya | this study |
| 15.9 | 2009 - 2011 | GPO1 | Kenya | Wamwiri *et al.*, (2013) |
| 1.4 | 2009 - 2011 | | Zambia | Dennis *et al.*, (2014) |
| 83.3 | UI | Gro*EL* | Tanzania | Matthew (2007)[a] |
| 17.6 | UI | | Zimbabwe | Matthew (2007)[a] |
| ***G. austeni*** | | | | |
| 16.3 | 2012 | Hem | Kenya | this study |
| 2.0 | 2009 - 2011 | GPO1 | Kenya[c] | Wamwiri *et al.*, (2013) |
| 4.6 | 2009 - 2011 | | Kenya[d] | Wamwiri *et al.*, (2013) |
| ***G. longipennis*** | | | | |
| 0.0 | 2012 | Hem | Nguruman, Kenya | this study |

[a] Dennis *et al.*, (2014) discussed S. glossinidius prevalence in tsetse flies and refers to a thesis by Matthew (2007), but the thesis cannot be accessed directly: [b] S. *glossinidius* prevalence in the Shimba Hills and in Nguruman were 59.2% and 0.7%, respectively: [c] from the Shimba Hills: [d] from Arabuko-Sokoke: UI (unidentified) means year of sample collection was not indicated in the reference.

Using $\chi^2$ tests, infection of trypanosomes showed a significant association with harbouring of *S. glossinidius* in tsetse flies when considered across all populations ($\chi^2$= 75.0, df = 1, P-value < 0.0001). The apparent significance is driven by a large proportion of samples that were both trypanosome negative and *Sodalis* negative; however, there was also a relatively large proportion that were trypanosome positive and *Sodalis* negative. Importantly, in the Shimba Hills, although many samples were both trypanosome positive and *Sodalis* positive, there was about an equal proportion of trypanosome positive flies that were *Sodalis* negative; similarly, there was no difference in the proportion of trypanosome negative flies that were *Sodalis* negative or positive. Thus, caution must be employed when interpreting broad-scale associations based on the $\chi^2$ test, particularly when other factors may be important in explaining the

associations. For example, among 450 tsetse flies (including *G. palpalis palpalis*, *G. pallicera*, *G. caliginea* and *G. nigrofusca*) from Bipindi and Campo, infection by trypanosomes was found to be significantly associated with *Sodalis* harbouring (Table 3.14) (Farikou *et al.*, 2010a). Prevalence of both *Sodalis* (64.4%) and trypanosomes (53.3%) was higher in Bipindi than in Campo (43.3% and 32.4%, respectively) and among trypanosome negative flies in the former there was no association with *Sodalis* so the pattern was mainly driven by the Campo population. There were more flies in both populations that were positive for both *Sodalis* and trypanosomes than *Sodalis* negative with trypanosomes but there was not a comparison between species of tsetse flies. In addition, the *Sodalis* screening results in that study were based on the pSG2 primers and no attempt was made to confirm species identification of *Sodalis* based on sequencing. Thus, simple association tests might not reveal the full complexity of interactions between the endosymbionts, the flies and the parasites.

**Table 3.14** Screening results of *Sodalis glossinidius* and *Trypanosoma spp.* in tsetse flies based on pSG2 and trypanosome species-specific primers (Farikou *et al.*, 2010a).

| Sites | Total | S+ | S+I- | S-I- | I+ | S+I+ | S-I+ |
|---|---|---|---|---|---|---|---|
| Bipindi | 225 | 145 | 53 | 52 | 120 | 92 | 28 |
| Campo | 225 | 102 | 49 | 103 | 73 | 53 | 20 |
| Total | 450 | 247 | 102 | 155 | 193 | 145 | 48 |

"I" refers to trypanosome screening results; "S" refers to *Sodalis* screening results.

### 3.5.3 Generalised Linear Models of *Sodalis glossinidius* and trypanosome status

A combination of collecting sites and tsetse fly species (subpopulation) was strongly associated with *S. glossinidius* presence, as demonstrated by both GLM Model 1 and the graphical MCA. However, *T. congolense* was the only pathogenic trypanosome species whose presence was significantly associated with *S. glossinidius* infection when the latter was considered as the response variable ($\chi^2$= 6.28, df = 48, P-value = 0.0122). Location of colonisation and development could be related to this association because both *T. congolense* and *S. glossinidius* colonise the midgut of tsetse flies (Peacock *et al.*, 2012a). However,

it is interesting that *T. brucei*, which also harbours in the midgut (Peacock *et al.*, 2012a), was not significantly associated with the presence of *Sodalis* in tsetse flies (see Model 3; Appendix B.6). *Trypanosoma congolense* mostly moves to the proboscis for development of the infective stage while that of *T. brucei* develops in salivary glands (Peacock *et al.*, 2012a). Thus remaining *T. congolense* in the midgut might be related to the presence of *Sodalis*. However, prevalence of *T. brucei* (13.2%) might have been too low to accurately define the association. Although *Sodalis* is also able to horizontaly transmit from males to females (De Vooght *et al.*, 2015), the prevalence of *T. congolense* in females tended to show negative correlations with increasing age. More samples are required for a better prediction from the models, especially *Sodalis* positive samples from the Nguruman region. However, it would be interesting to further study about the possible effects of sex on *Sodalis* status.

I also attempted to use "levels" of *S. glossinidius* present based on intensity of bands in standard PCR as a response variable in GLM analyses (Appendix B.12). The association lead to complicated conclusions of interactions among tsetse factors (subpopulation and sex) and trypanosome levels. This might indicate that levels of infection are what is affected by host factors; alternatively, it could be due to noise due to differences in DNA quality among samples, rather than real differences in endosymbiont density among samples. More sensitive methods, such as qPCR, could be useful for not only detecting low-density endosymbionts but also to enable quantification of levels of infections (Aksoy *et al.*, 2014). Aksoy *et al.* (2014) successfully used qPCR for *S. glossinidius* detection in the midgut of *G. fuscipes fuscipes*, *G. m. morsitans* and *G. pallidipes* and found a wide range of *S. glossinidius* densities (Aksoy *et al.*, 2014). Thus, investigation of *S. glossinidius* using qPCR should be studied to provide a more reliable assessment of the relative importance of tsetse factors to the relationship between *Sodalis* presence in infection with trypanosomes.

GLMs were also conducted for presence of any trypanosome species as the response variable to determine whether there would be more power to detect the associations than when based on single species. However, there was still no evidence for an association between trypanosome presence and *Sodalis* presence and so the best-fitting model was the same as reported in chapter 2 with sex, subpopulation and age all influencing the association. *T. vivax* showed the

highest prevalence in all trapped tsetse flies (49.0%), but showed the least complicated associations (GLM Model 8). However, the addition of *Sodalis* to the best-fitting models for *T. congolense* and *T. brucei* resulted in complex patterns of interactions. In old flies, presence of *Sodalis* in females and absence in males was predicted to be associated with a decrease in *T. congolense* prevalence; however, there was no difference between males and females for young flies (Figure 3.16). In addition, the absence of *Sodalis* appeared to increase susceptibility of *T. brucei* in tsetse flies from Buffalo Ridge but the pattern was less clear for tsetse flies from Zungu Luka (Figure 3.17). This suggests that the role of *Sodalis* in affecting immune mechanisms might be different in males and females and in different species of flies from different geographic regions. *Sodalis* appeared to enhance trypanosome harbouring in males but trypanosome defence in females, but the mechanisms remain unclear. *Sodalis* could thus play a role in trypanosome prevalence but could differ by species, consistent with biological predictions based on the site of harbouring (Peacock *et al.*, 2012a). However, *T. brucei* showed the lowest prevalence (13.2%) but the best-fitting model was the most complicated (GLM Model 7), which could suggest that the apparent pattern of interactions is driven by low power to resolve the most influential explanatory variables when there are many potential factors that explain relationships. Thus, while a GLM approach is more appropriate than $\chi^2$ tests for interpreting relationships among multiple variables, it still can be difficult to draw firm conclusions when these variables interact in complicated ways.

### 3.5.4 Multiple correspondence analysis of *Sodalis*, trypanosomes and intrinsic factors of tsetse flies

Since it is not possible to biologically interpret such complex interactions in GLM models, multivariate analyses, such as the MCA used here, provide an appropriate means of visualising correlations among factors. Graphical MCA supported conclusions from the GLM analyses that presence of *S. glossinidius* was distinctly correlated to only particular subpopulations of tsetse flies while trypanosome infection was correlated more with sex, age and subpopulation than *Sodalis* status. The association of *T. brucei*¸with sex and age was different than for *T. vivax* and *T. congolense*. *Trypanosoma brucei* positive flies tended to be found in old and male flies whereas the other two were more common in

young female flies. The environment of salivary glands belonging to female *G. m. morsitans* and *G. pallidipes* is more inhospitable for *T. brucei* than males and could because of higher effective immune response in the former (Peacock *et al.*, 2012b). Although these patterns were quite weak and no strong predictions were suggested about the association of each species of trypanosomes with tsetse factors (including *Sodalis* status), I suggest that multivariate analyses such as this could have more potential for resolving complex associations and for designing more targeted studies to identify risk factors at a more local scale.

## 3.6 Conclusions

Investigation of *S. glossinidius* prevalence based on the PCR method, combined with sequencing, was sufficiently sensitive to detect and determine variation in bacterial genotypes among species of tsetse flies from different sampling sites but the choice of primers is critical to interpretations. My research clearly shows that prevalence of *S. glossinidius* is significantly associated with subpopulation which is consistent with previous studies that both species of tsetse flies and sites of collection influence the presence of *S. glossinidius*. Athough association between trypanosome status and *Sodalis* status was not clear, *T. congolense* and *T. brucei* were significantly associated with presence of *Sodalis* in some analyses, but only in complex interactions with other tsetse-specific factors. These results imply that there are many factors that affect the presence or absence of *S. glossinidius* and that each species of trypanosome could show different directions of relationships. It is thus important not to study potential associations in isolation; combined with chapter 2, my results emphasise that the factors that determine whether tsetse flies are infected with trypanosomes is complex and simple associations might not be informative enough to inform management programmes. In order to control trypanosomiasis relating to *Sodalis*, it could be effective for *T. congolense* and *T. brucei* control but there might not be a benefit for *T. vivax*. Better understanding of host-feeding patterns of tsetse flies would also be interesting to assess in relation to risk of trypanosome transmission between hosts (see chapter 4). The next chapter will focus on *G. pallidipes* as a model for understanding the association of trypanosome infection status with feeding patterns and other tsetse specific factors.

# Chapter 4 Tsetse host identification and association of tsetse hosts to trypanosome and *Sodalis* status in *Glossina pallidipes*

## 4.1 Abstract

*Glossina pallidipes* is a common biological vector for African trypanosomiasis in southeastern Kenya. Both male and female flies transmit and/or acquire trypanosomes while feeding on domestic animals and wildlife. With different status of the reproductive cycle and sexes, host-feeding patterns could be associated with trypanosome infection in the flies. The objectives of this study were to investigate patterns of host feeding in *G. pallidipes* from Kenya and determine whether host-feeding status (unfed, single or multiple hosts) was associated with differences in trypanosome prevalence among populations. Sources of blood meals of tsetse flies were identified by sequencing the mitochondrial cytochrome B (cytb) gene from gut DNA samples of 577 *G. pallidipes* from three geographic locations (Buffalo Ridge, Zungu Luka and Mukinyo). Cytochrome B fragments from vertebrate hosts were detected in 74.7% of collected samples but it was only possible to identify the host species using direct sequencing for 28.6% of samples, due to the presence of multiple trypanosome species within individual flies and variable sequence quality. In Mukinyo most flies fed on single hosts but in the other populations cloning confirmed multiple host feeding in a subset of samples that appeared heterozygous in direct sequences. Blood meal analysis results were in agreement with previously published reports that *G. pallidipes* fed predominantly on African buffalo and African elephants, along with a range of other less common hosts, including humans. There was no association between host-feeding patterns and trypanosome presence but there were differences in the relative prevalence of particular species of trypanosomes among sites that differed in the host communities that the flies fed on. Multiple Correspondence Analysis suggested complex relationships among feeding patterns, trypanosome presence, tsetse-specific factors and geographic location. I would suggest that environment and host availability may affect host feeding of tsetse flies. Details about tsetse hosts are thus an important key to understanding the epidemiology of trypanosomes in order to control the disease in particular areas.

## 4.2 Introduction

Trypanosomes biologically transmit to animals and humans via tsetse flies (*Glossina spp.*). The blood parasites possibly harbor and conceal themselves in tsetse flies and wild animals, which are reservoir hosts and do not show clinical signs of the infection (World Organisation for Animal Health, 2013). Complex interactions among trypanosomes, vertebrate hosts and tsetse flies likely increase the difficulty of controlling and eradicating trypanosomes but few studies have attempted to combine investigation of host-feeding patterns in relation to trypanosome infection with intrinsic factors of tsetse flies distributed in different regions. Since both sexs of teneral tsetse flies feed on blood meals, they are both at risk of exposure to trypanosomes from infected vertebrate hosts after they emerge from their puparium (Jackson, 1946). The feeding-interval time of tsetse flies is approximately two, although that of males has been found to be longer than that for females (Turner, 1987) and times can extend to three days in mature tsetse flies seeking their preferred hosts (Bouyer *et al.*, 2007). Both tsetse flies and *Trypanosoma spp.* are catholic feeders on a wide variety of vertebrate host species (Weitz, 1963), including: dogs (Eloy and Lucheis, 2009, Lisulo *et al.*, 2014); cat (Clausen *et al.*, 1998, Maudlin *et al.*, 2004); horses (Maudlin *et al.*, 2004); chickens (Clausen *et al.*, 1998); ruminants (Clausen *et al.*, 1998, Maudlin *et al.*, 2004, Nyawira, 2009, Turner, 1987); antelopes (Turner, 1987); elephants (Muturi *et al.*, 2011, Nyawira, 2009); warthogs (Nyawira, 2009); and reptiles (Turner, 1987). Host feeding of tsetse flies is thus imortant for investigating the distribution and epidemiology of trypanosomiasis.

Vertebrate hosts of tsetse flies have been identified based on blood meal analysis to determine animal species that are at risk of trypanosome infection via these vectors. Serological methods have been applied to determine sources of fresh blood feeding; for example, using precipitin and haemagglutination (Weitz, 1963); and enzyme-linked immunosorbent assays (ELISA) (Clausen *et al.*, 1998). These techniques are, however, time-consuming, and have low sensitivity (Maleki-Ravasan *et al.*, 2009). As an alternative, PCR-based techniques have been used to amplify target mitochondrial genes of fed hosts from blood meal contents in guts of tsetse flies, followed by DNA sequencing to confirm host species. Mitochondrial genes, which are inherited from mother to offspring (Dyer *et al.*, 2008), have been widely used for the taxonomic classification of animals

based on DNA barcoding (Lah *et al.*, 2012, Pradhan *et al.*, 2015) and so make useful markers for host determination in tsetse flies (Njiokou *et al.*, 2004, Nyawira, 2009, Farikou *et al.*, 2010b, Muturi *et al.*, 2011). Blood meal analysis has been based most often on the mitochondrial genes cytochrome b (cytb) and cytochrome c oxygenase1 (CO1) genes (Kocher *et al.*, 1989) because of the availability of a large database of reference sequences in GenBank. However, the specific primers used to target these regions have varied in whether they target any vertebrates, mammals more specifically or individual host species. For example, Muturi *et al.* (2011) surveyed blood meals of *G. swynertoni* from Tanzania and *G. pallidipes* from Kenya and Uganda using primers targeting the cytB (Cb1 and Cb2) and COI (VF1d_t1, VR1d_t1) genes. African elephants, African buffalo, warthogs, lizards, giraffes, spotted hyenas and baboons were identified as sources of blood contained in tsetse abdomens using the cytB primers whereas the COI primers also identified Nile Monitor lizards in blood contents but were not able to identify the more common mammalian hosts. Thus, the choice of primers could be important to develop surveys for estimating how tsetse-feeding pattern affects transmission of trypanosome to and from vertebrate hosts.

Moreover, the blood meal might only be detected for a short amount of time due both to gastric clearing and to degradation of host DNA over time whereas development of infective stages of trypanosomes could be much longer. For example, in *G. palpalis*, *T. vivax*, *T. congolence* and *T. brucei* become infective 7, 13 and 12 days post exposure, repectively (Bruce *et al.*, 1910). However, these timings have not been established for *G. pallidipes* so it is not possible to predict the duration of blood meal detection in relation to the rate of trypanosome development in this species. Thus, PCR-based blood meal analysis can neither indicate that trypanosome infection causes changes in feeding pattern of flies or which hosts have contributed to observed trypanosome infections. Nevertheless, geographic regions where flies have the opportunity to feed on multiple host species within a short amount of time, for example, could increase risk of transmitting trypanosomes between vertebrate hosts than areas where flies feed on single hosts in a similar time period. Thus, blood meal analysis could be informative about the relative risks of trypanosome transmission in different environments.

However, to fully understand the risk dynamics, it is also important to know whether different parasites are carried by different host species, which can also be determined using molecular techniques. For example, using the Internal Transcribed Spacer (ITS) primer sets designed by Cox *et al*. (2005) and specific PCR for *T. brucri* detection, Anderson *et al*. (2011) found high host specificity of particular species of trypanosomes in blood samples of wild vertebrate hosts sampled from Zambia: bushbucks for *T. b. gambiense*; and bovines for *T. congolense*. Moreover, Auty *et al*. (2012) used nested PCR (based on both the ITS-1 and ITS-2 regions) to investigate *Trypanosoma spp*. in blood samples from wildlife sampled from Tanzania and Zambia. Although they identified the full range of named species expected to be found in wild animals in these regions (*T. vivax*, *T. simiae* Tsavo and *T. godfreyi*, *T. brucei* and *T. congolense*) and they identified specific associations with particular host species, they also found new sequences that did not correspond to previously recognized species of trypanosome. They concluded that there was a higher diversity of trypanosome sequences from wild vertebrate hosts than had previously been anticipated, but the implications for transmission dynamics remain unclear because it is not known how this relates to the diversity of parasites carried by tsetse flies. It is known that there can be differences among species of tsetse flies in the relative prevalence of different species of trypanosomes (see chapters 2 and 3), but different species of tsetse fly are also characterized by different sets of preferred hosts (Leak, 1998). Not only will this affect which species of trypanosomes are transmitted but the quality of the blood meal could also differ between hosts (Geigy *et al*., 1971). Thus, disease control measures might need to be specifically targeted to the particular communities of tsetse flies, vertebrate hosts and trypanosome species present in a particular region.

In this chapter I focused on *G. pallidipes,* for three reasons. Firstly, *G. pallidipes* was sampled in larger numbers than the other species of tsetse collected and so could provide enough blood meal results for robust statistical analysis. Secondly, *G. pallidipes* was collected from three sites (Buffalo Ridge, Zungu Luka and Mukinyo) so sample sources could be considered as one of the explanatory variables for association analysis. Thirdly, blood meal analysis of these tsetse flies species has been the most commonly studied so provides most potential to compare with other studies. This species has been demonstrated to feed on a

wide range of hosts, including bovines (Muturi *et al.*, 2011, Clausen *et al.*, 1998, Turner, 1987, England and Baldry, 1972, Okoth *et al.*, 2007), suvids (Bett *et al.*, 2008, Okoth *et al.*, 2007), elephants (Muturi *et al.*, 2011, Nyawira, 2009), antelopes (Allsopp *et al.*, 1972) and cattle (Muturi *et al.*, 2011). It thus provides a good model to test whether tsetse-specific factors (e.g. sex and age) in a single vector species from different sites influence feeding patterns that could affect the dynamics of trypanosome distribution.

The purpose of this study was to identify feeding statuses (fed or unfed), feeding patterns (unfed, single or multiple hosts) and hosts of *G. pallidipes* using blood meal analysis based on PCR and sequencing techniques in order to determine whether adding host-feeding patterns affected interpretation of the association between trypanosome infection and tsetse-specific factors (age, sex, habitat, presence of endosymbionts). The ultimate aim is to determine whether more knowledge about these associations would be useful for a better understanding of trypanosomiasis distribution and control. My hypothesis was that feeding patterns demonstrated by recent blood meals (unfed, single or multiple hosts) of *G. pallidipes* would be associated with trypanosome status: specifically, that tsetse-fly populations with the opportunity to feed on multiple host species in a limited time would show higher prevalence of trypanosomes due to higher risk of exposure from infected hosts. I also hypothesised that feeding patterns might vary by age, sex, or *Sodalis* status of the flies and so could also influence risk of trypanosome infection.

## 4.3 Materials and methods

### 4.3.1 Optimisation of blood meal analysis based on PCR

#### 4.3.1.1 Choosing primers for blood meal analysis

To determine which primers were most appropriate for identification of blood meal sources, 12 extracted DNA samples of *G. pallidipes* from Buffalo Ridge (two samples), Zungu Luka (two samples) and Mukinyo (eight samples) (from chapter 2) were initially screened using Cb1, Cb2 and VF1d_t1, VR1d_t1 primers (Table 4.1). For this optimisation step, samples from the other testse species were also

tested: *G. austeni* (four samples), *G. brevipalpis* (12 samples) and *G. longipennis* (12 samples) to find out the appropriate methods for blood meal analysis in tsetse flies. The Cb primers, developed by Kocher *et al.*, (1989), target a 359 bp fragment of the mitochondrial gene cytochrome B (cytb) gene in mammals. The VF1d-t1 and VR1d-t1 primers described by Ivanava *et al.*, (2006), were designed for detection of a 648 bp fragment of the mitochondrial cytochrome C oxygenase 1 (COI) gene in mammals, reptiles and fish (Ivanova *et al.*, 2007b). PCR cycling for both primer sets were carried out in 25 µl reaction mixtures containing: 1X PCR buffer minus Mg; 0.2 mM dNTP mixture; 1.5 mM $MgCl_2$ (Thermo Scientific); 0.5 µM of each primer; 1 unit of *Taq* DNA polymerase (Invitrogen Inc, Carlsbad, CA., U.S.A.); and 2 µl DNA template. PCR conditions of as described by Muturi *et al*. (2011): samples were pre-heated at 94°C for 5 min, denatured at 94°C for 30 sec, annealed at 55°C for 45 sec, then extended at 72°C for 30 sec, with 35 cycles of the amplification. Following this, there was a final extension at 72°C for 10 min. PCR products were visualised using 1.5% UltraPure™ Agarose gels (Invitrogen, Paisley) with 2% Ethidium Bromide (Invitrogen, Paisley) in 1X TBE buffer (108 g of Tris Base, 55 g of Boric acid and 40 ml of 0.5 M EDTA, with $ddH_2O$ added to 10 litters). The conditions of gel electrophoresis were 100 V., 300 A., 50 min. The results were visualised and analysed by a gel documentation system (UVIpro Plainum, UVITEC, Cambridge, UK or GeneDoc, BioRad Inc, UK).

**Table 4.1** Sequences of Cb1, Cb2 primers and VF1d_t1, VR1d_t1 primers used for blood meal analysis. Sizes of PCR products and references are indicated

| Primers | Sequences | size (bp) | Reference |
|---|---|---|---|
| Cb1 | CCATCCAACATCTCAGCATGATGAAA | 359 | Kocher *et al.* (1989) |
| Cb2 | GCCCCTCAGAATGATATTTGTCCTCA | | |
| VF1d_t1* | **TGTAAAACGACGGCCAG**TTCTCAACCAACCACAAAGACATTGG | 648 | Ivanova *et al.* (2006) |
| VR1d_t1* | **CAGGAAACAGCTATGAC**TAGACTTCTGGGTGGCCAAAGAATCA | | Ward *et al.* (2005) |

* includes M13 tag (M13 taq are indicated in boldface)

#### 4.3.1.2 Selection of sequencing methods for host determination

A pilot experiment was conducted to also determine the most reliable method of sample preparation for sequencing. Three extracted DNA samples of *G. pallidipes* collected from Buffalo Ridge were amplified using the mammalian specific Cb1, Cb2 primers (Muturi *et al.*, 2011) and used to test three sequencing methods: 1) direct sequencing of unpurified PCR products; 2) direct sequencing after PCR product cleaning using either QIAquick PCR Purification Kits (Qiagen Inc, Paisley, UK) or ExoSAP-IT PCR Clean-up Kits (GE Healthcare); 3) sequencing of cloned products. For the latter, PCR products were inserted into plasmid vectors and transformed into *E. coli* competent cells using StrataClone PCR Cloning Kits (Agilent Technologies UK Limited), with three single colonies of each sample sub-cultured in LB broth and incubated at 37°C overnight. Then, the plasmid DNA was extracted using QIAprep Spin Miniprep Kits (Qiagen Inc, Paisley, UK). All samples were sent for sequencing at the DNA Sequencing and Services, University of Dundee. Sequencing results of the three unpurified PCR products, the six purified PCR products (three for each purification kit) and the nine purified plasmids were visualized using Geneious® 7.0.3 (Biomatters Ltd., Auckland, NZ) and analysed for host determination using BLAST.

### 4.3.2 Blood meal analysis

All abdominal DNA samples of G. *pallidipes* (N = 577 samples) from chapter 2 were screened using the cytb primers to identify feeding status of tsetse flies. *Glossina pallidipes* (i.e. Cb positive were classified as “fed” and Cb negative as “unfed”), with sequencing used to identify host species fed on. Yield of all PCR products was estimated by comparing fragments to a 100 bp DNA ladder (Promega, Corporation, Madison, U.S.A.). PCR products that amplified at a high concentration (at least 20 ng/15 µl) were cleaned using ExoSAP-IT PCR Clean-up Kits (GE Healthcare). In cases where yield of PCR products was lower than 20 ng/15 µl, multiple PCR products were concentrated and QIAquick PCR Purification Kits (Qiagen Inc, Paisley, UK) were applied to extract the PCR products from agarose gels. All purified samples were sent for DNA sequencing in both forward and reverse directions, using the Sequencing Service at the University of Dundee.

Chromatogram sequences were corrected and aligned using the Sequencher software program, version 5.3 (Gene Codes Corporation, Ann Arbor, MI USA). BLAST was used to identify the closest matching sequences in the GenBank database to determine host identity of each consensus sequence. Chromatographs with only single peaks based on direct sequences were classified as "single host feeding" and could be used to confidently identify host species. Sequences that showed multiple clearly heterozygous positions were classified as "multiple host feeding" but it was normally not possible to predict host species accurately due to the difficulty of resolving the phase of mutations. However, it was sometimes possible to identify polymorphisms within a single host species based on the heterozygous patterns (i.e. based on only a few mutations). Chromatographs that did not show clear homozygous or heterozygous peaks for at least 100 bp of the sequence were removed from the results.

Three sets of samples were rechecked to confirm host-feeding patterns by cloning using TOPO®-TA Cloning Kits (Invitrogen, UK), with six plasmids of each sample sent for sequencing, after purifying using QIAprep Spin Miniprep Kits (Qiagen Inc, Paisley, UK). The first set were samples whose chromatographs showed double or triple peaks at single positions in the direct sequences, to confirm that heterozygous peaks were due to multiple feeding hosts rather than poor quality sequences. The second set was samples identified as single host feeding flies, to confirm that those flies had fed only on single hosts. The final set was DNA extracted from my buccal cells to differentiate between my DNA and human DNA samples from Africa. BLAST was used to determine host identity from all these sequencing results.

### 4.3.3 Prevalence of *Trypanosoma spp*. and associations of hosts with trypanosome status of *G. pallidipes*

Data about *Trypanosoma spp*. screening in head plus proboscis of *G. pallidipes* using ITS-1 CF and ITS-1 BR primers from chapter 2 were used to quantify prevalence of trypanosomes in the set of samples used for blood meal analysis. $\chi^2$ contingency analyses were initially used to specifically test for an association between trypanosome status with feeding status (fed or unfed) and with feeding patterns (single or multiple feeding) of *G. pallidipes*. These tests were

conducted both across all sites and within each of the three sampling sites using $\chi^2$ tests.

Principle correspondence analysis (PCA) was then applied to visualise graphic correlations of trypanosome status with intrinsic factors, *Sodalis* status and feeding patterns in the *G. pallidipes* samples for which feeding patterns could be assigned (N = 361) using Multiple Correspondence Analysis (MCA), as implemented in the FactoMineR package (version 1.30) and the "ggplot2()" function in R. Analysis of MCA graph 1 included six categorical variables, comprised of presence or absence of *S. glossinidius* and *Trypanosoma spp.*, as well as site, sex, age and feeding pattern of *G. pallidipes* (unfed, single and multiple feeding).

## 4.4 Results

### 4.4.1 Optimisation of blood meal analysis based on PCR

#### 4.4.1.1 Choosing primers for host determination

The strength of amplification and numbers of positive samples tended to be greater using the Cb1, Cb2 primers (cytb) than the VF1d_t1, VR1d_t1 (COI) primer sets in the same samples (Figure 4.1), despite the fact that the latter were more conservative as they were designed to target all vertebrates. Among 12 DNA samples of *G. pallidipes*, cytb fragments were amplified in nine individuals tested, whereas none were amplified using the COI primers from the same sets of DNA samples (Table 4.2). A number of DNA samples belonging to *G. austeni*, *G. brevipalpis* and *G. longipennis* showed amplification of COI but this was consistently lower than for cytb. Overall, there were 24/40 cytb positive samples that were COI negative. Therefore, it was decided to use only cytochrome B for the blood meal analysis, to avoid a high number of false positives.

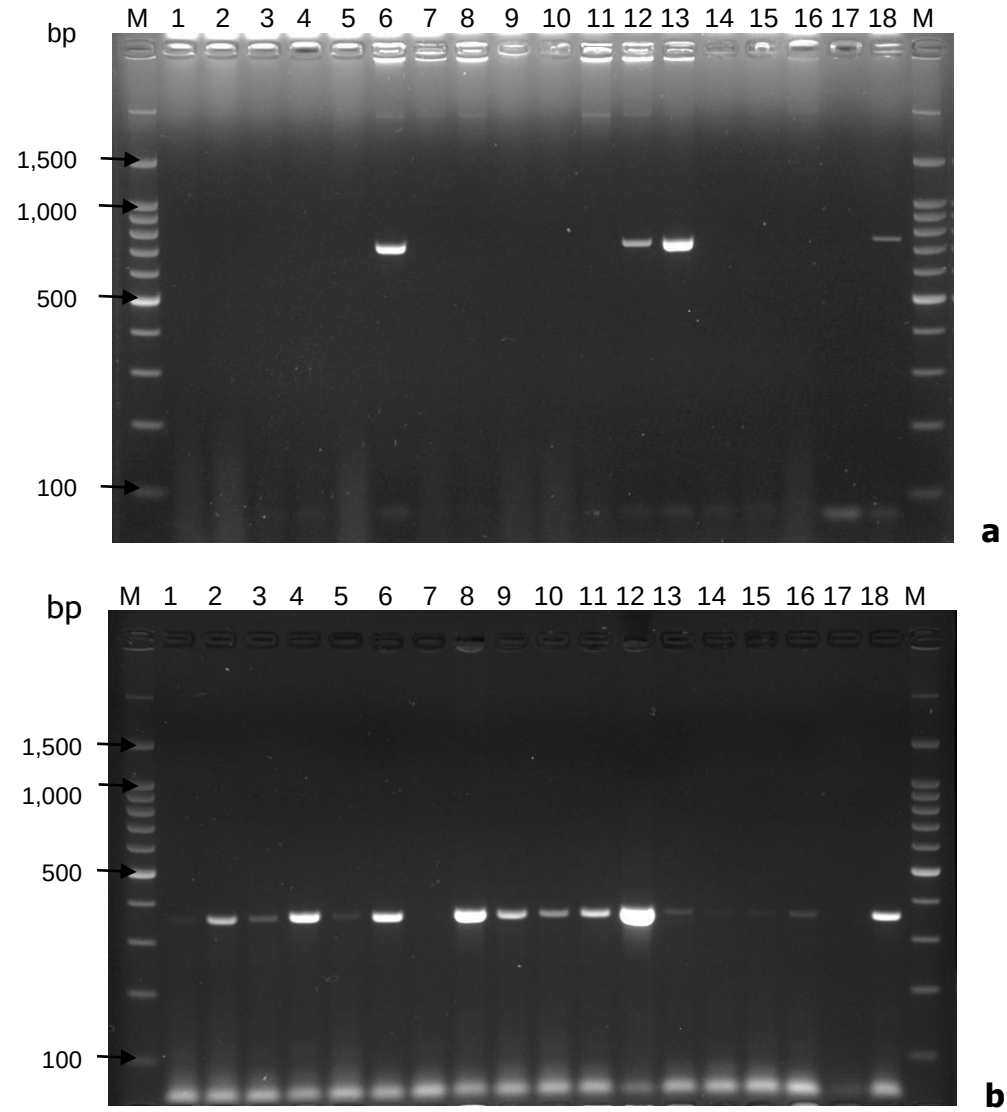



**Figure 4.1** Comparison of agarose gel electrophoresis from COI- and cytb-based blood meal identification from abdomen samples (AB) of *Glossina spp*. (Gp = *G. pallidipes;* Gb = *G. brevipalpis*; and Gl = *G. longipennis*) from three populations (BR = Buffalo Ridge; Mu = Mukinyo; Sa = Sampu).

The gels indicate PCR products using: (a) VF1d_t1, VR1d_t1 primers; and (b) Cb1, Cb2 primers. Lane 1: BRGb 30AB. Lane 2: BRGb 33AB. Lane 3: BRGb 60AB. Lane 4: MuGp 9AB. Lane 5: MuGp 13AB. Lane 6: MuGl 1AB. Lane 7: SaGl 24AB. Lane 8: SaGl34AB. Lane 9: MuGp 109AB. Lane 10: MuGp 110AB. Lane 11: MuGl 2AB. Lane 12: MuGl 6AB. Lane 13: SaGl 7AB. Lane 14: MuGp 275AB. Lane 15: MuGp 283AB; 16: MuGl 25AB. Lane 17: Negative control (nuclease free water. Lane 18: positive control (BRGp 4AB). Lane M: molecular standard 100 bp DNA ladder (Promega, Corporation, Madison, U.S.A.).

**Table 4.2** Correspondence between mitochondrial DNA fragments amplified from *G. pallidipes*, *G. austeni*, *G. brevipalpis* and *G. longipennis* using two sets of primers; Cb1, Cb2 (cytb) and VF1d_t1, VR1d_t1 (COI).

| Primers | | Cb1, Cb2 | | Total |
|---|---|---|---|---|
| | | Positive | negative | |
| ***G. pallidipes*** | | | | |
| VF1d_t1, VR1d_t1 | positive | **0** | **0** | 0 |
| | negative | **9** | **3** | 12 |
| Total | | 9 | 3 | 12 |
| ***G. austeni*** | | | | |
| VF1d_t1, VR1d_t1 | positive | **1** | **0** | 1 |
| | negative | **1** | **2** | 3 |
| **Total** | | 2 | 2 | 4 |
| ***G. brevipalpis*** | | | | |
| VF1d_t1, VR1d_t1 | positive | **3** | **0** | 3 |
| | negative | **9** | **0** | 9 |
| Total | | 12 | 0 | 12 |
| ***G. longipennis*** | | | | |
| VF1d_t1, VR1d_t1 | positive | **4** | **0** | 4 |
| | negative | **5** | **3** | 8 |
| Total | | 9 | 3 | 12 |

#### 4.4.1.2 Selection of sequencing methods for host determination

Chromatographs of sequences from the unpurified PCR products were not as clean as when compared with those of the purified PCR products. Both purification methods produced readable sequences and there was not much difference between them (Figure 4.2). The purified plasmids showed the highest consistency and longest readable sequences of the chromatographs when compared with those of the unpurified PCR products and the purified PCR products, but the differences were not substantial compared to the purified products (Figure 4.2). For all sequences, *Syncerous caffer* (African buffalo) was the only host that was found among the three samples (Table 4.3). According to the results, I concluded that direct sequencing using purified DNA for sequencing is feasible for blood meal analysis rather than more expensive and time-consuming cloning.

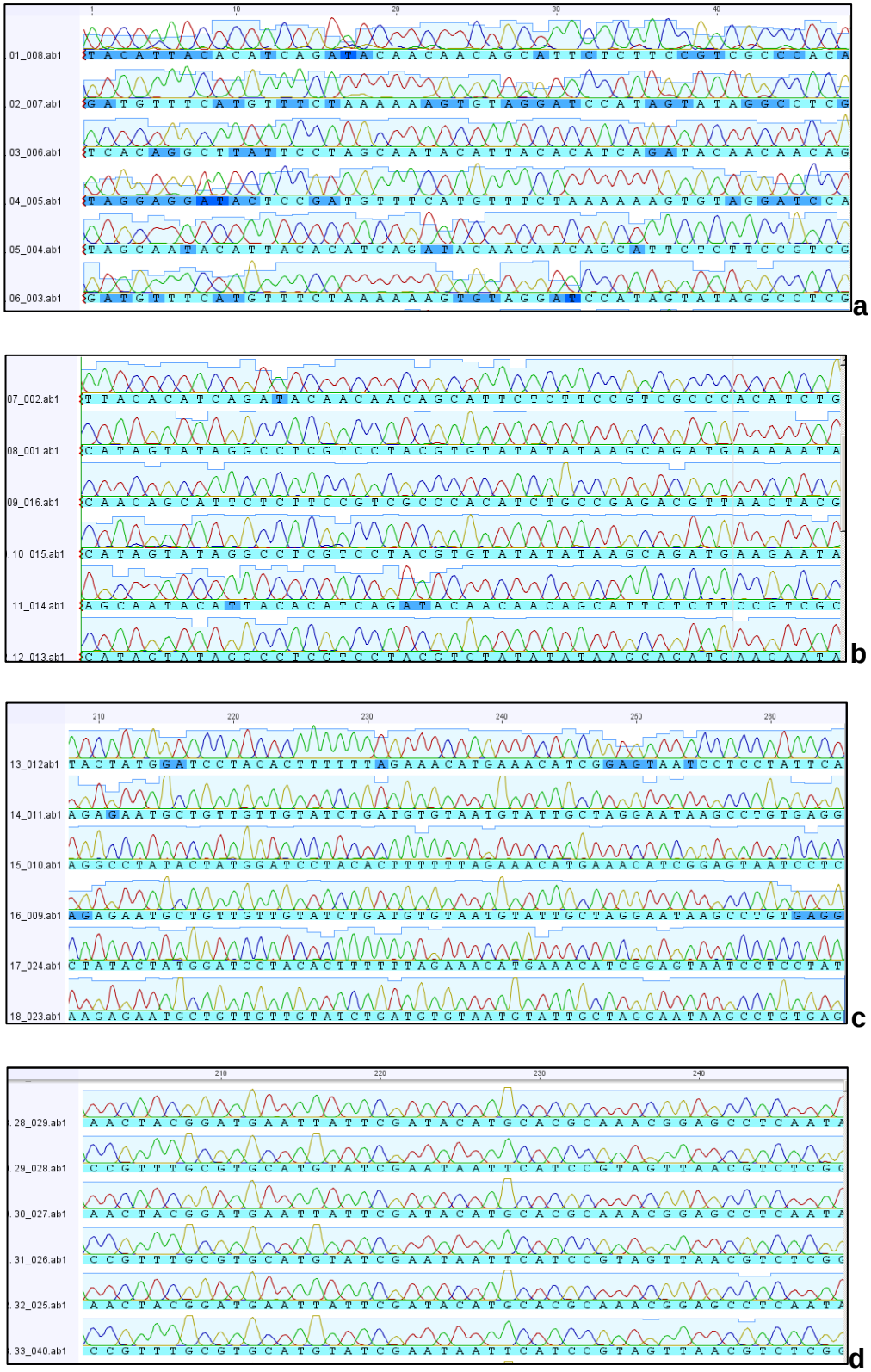



**Figure 4.2** Chromatographs of the sequencing results.

(a) Unpurified PCR products; (b) PCR products purified using QIAquick PCR Purification Kits (Qiagen Inc, Paisley, UK); (c) PCR products purified using ExoSAP-IT PCR Clean-up Kits (GE Healthcare); (d) Purified plasmids.

**Table 4.3** Results of blood meal analysis using direct sequences from PCR products of cytb primers from three *G. pallidipes* samples from Buffalo Ridge, with host identity confirmed by BLAST.
Shown are the sample name (indicating the site of samples, species of tsetse and the tissue part of tsetse flies), the method of DNA preparation, a description of the closest match found in BLAST, the proportion of the sequence that matched a sequence in BLAST (% query cover), the similarity of the sequence to that in BLAST (% Identity) and the GenBank accession number.

| Samples | DNA preparation | Description | Query cover (%) | Identification (%) | Accession |
|---|---|---|---|---|---|
| **BRGp4AB** | Unpurified | *Syncerus caffer* isolate 3479 mitochondrion, complete genome | 100 | 100 | JQ235539.1 |
| **BRGp7AB** | Unpurified | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 100 | 100 | JQ235544.1 |
| **BRGp14AB** | Unpurified | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 100 | 100 | JQ235544.1 |
| **BRGp4AB** | QIA Quick | *Syncerus caffer* isolate 3479 mitochondrion, complete genome | 100 | 100 | JQ235539.1 |
| **BRGp7AB** | QIA Quick | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 100 | 100 | JQ235544.1 |
| **BRGp14AB** | QIA Quick | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 100 | 100 | JQ235544.1 |
| **BRGp4AB** | EXO SAP | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 96 | 100 | JQ235544.1 |
| **BRGp7AB** | EXO SAP | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 100 | 100 | JQ235544.1 |
| **BRGp14AB** | EXO SAP | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 100 | 100 | JQ235544.1 |
| **BRGp4AB.1** | Cloning | *Syncerus caffer* isolate 3479 mitochondrion, complete genome | 82 | 99 | JQ235539.1 |
| **BRGp4AB. 2** | Cloning | *Syncerus caffer* isolate 3479 mitochondrion, complete genome | 82 | 99 | JQ235539.1 |
| **BRGp4AB.3** | Cloning | *Syncerus caffer* isolate 3479 mitochondrion, complete genome | 81 | 99 | JQ235539.1 |
| **BRGp7AB. 1** | Cloning | *Syncerus caffer* isolate 9084 mitochondrion, complete genome | 84 | 99 | JQ235526.1 |
| **BRGp7AB. 2** | Cloning | *Syncerus caffer* isolate 9084 mitochondrion, complete genome | 79 | 99 | JQ235526.1 |
| **BRGp7AB. 3** | Cloning | *Syncerus caffer* isolate 9084 mitochondrion, complete genome | 81 | 99 | JQ235526.1 |
| **BRGp14AB.1** | Cloning | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 82 | 99 | JQ235544.1 |
| **BRGp14AB.2** | Cloning | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 82 | 99 | JQ235544.1 |
| **BRGp 14AB3** | Cloning | *Syncerus caffer* isolate 671 mitochondrion, complete genome | 82 | 99 | JQ235544.1 |

### 4.4.2 Blood meal analysis

From 577 *G. pallidipes*, 431 tsetse flies showed amplification products following screening with Cb1, Cb2 primers, which were used to classify feeding status (Table 4.4). Among the fed flies, chromatograms of 216 samples were not sufficient quality to determine the source of the blood meals. There were 50 samples that showed clear heterozygous peaks that could be confidently assigned as flies having fed on multiple hosts rather than poor sequence quality but the proportion of samples differed by site. The two sites from the Shimba Hills (Buffalo Ridge and Zunga Luka) showed a higher proportion of flies that appear to have fed on multiple hosts than the site from Nguruman (Mukinyo): 13.0% from Buffalo Ridge (95% CI = 7.6 - 18.3%); 10.0% from Zungu Luka (95% CI = 6.2 - 13.0%); and 0.7% from Mukinyo (95% CI = 0 - 1.6%). There were 165 samples that produced readable sequences for which blood meals could be classified as single hosts in the recent blood meals based on the absence of heterozygous bands and for which host identity could be confirmed using BLAST. Flies from Mukinyo (45.1%, 95% CI = 39.4 - 50.7%) had the highest rate of feeding on single hosts, followed by Buffalo Ridge (17.5%, 95% CI = 11.5 - 23.5%) and Zungu Luka (2.0%, 95% CI = 0.4 - 3.7%). The remainder of the samples in each population either showed no evidence of a recent blood meal ("unfed") or did not have clean enough sequence patterns to be assigned confidently ("unidentified"; Table 4.4; Figure 4.3 - Figure 4.6 -4.6). Across populations, the predominant hosts of tsetse flies were African buffalo (*S. caffer*; 110 samples), followed by African elephants (*Loxodonta africana* and *L. cyclotis*; 32 samples), antelopes (*Tragelaphus scriptus*; seven samples), warthogs (*Phacochoerus africanus*; six samples), humans (*Homo sapiens*; five samples), giraffes (*Giraffa Camelopardalis*; four samples) and spotted hyenas (*Crocuta crocuta*; one sample) (Figure 4.3; Appendix C.1).

**Table 4.4** Summary of blood meal analysis results based on direct sequencing.

Cytb negative samples were classified as "unfed flies". Single host feeding refers to cases where the cytb sequence had only single chromatograph peaks. Multiple host feeding were samples for which cytb was amplified but the sequences showed multiple chromatograph peaks (i.e. they were clearly heterozygotes). Unidentifed tsetse flies were characterised based on strongly amplified cytb PCR that could not be confirmed by sequencing due to failed reactions or poor quality.

| Sites | Number of *Glossina pallidipes* | | | |
|---|---|---|---|---|
| | Unfed | Patterns of feeding | | |
| | | Single host | Multiple host | unidentified |
| Buffalo Ridge (N = 154) | 41 | 27 | 20 | 66 |
| Zungu Luka(N = 130) | 32 | 6 | 28 | 64 |
| Mukinyo(N = 293) | 73 | 132 | 2 | 86 |
| **Total** | **146** | **165** | **50** | **216** |

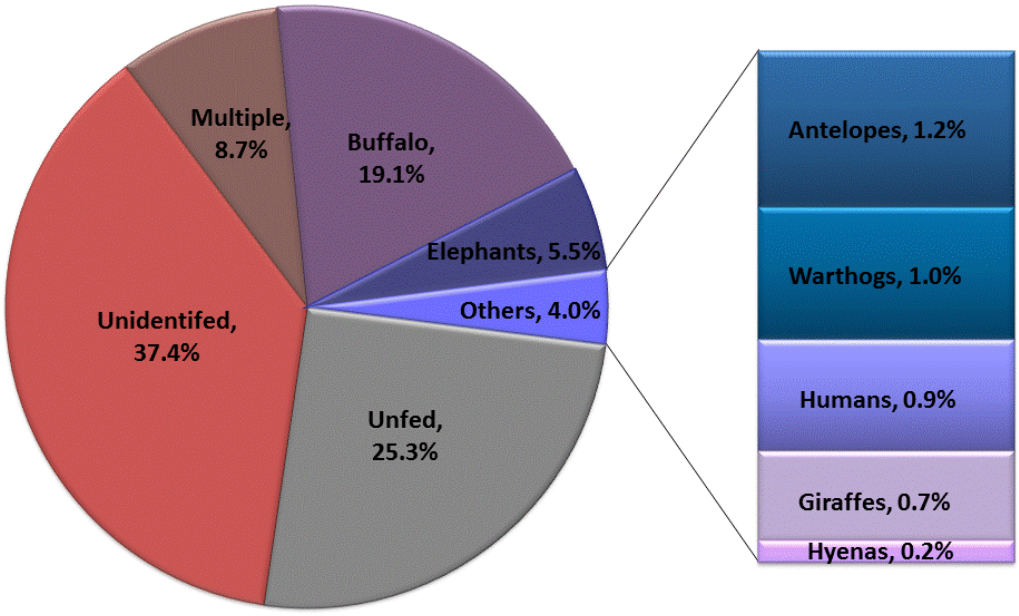



**Figure 4.3** Results of blood meal analysis from all *G. pallidipes* (N = 577) analysed.
Feeding status of all flies was classified using PCR screening with cytb primers. Multiple host feeding and host identification were characterised based on BLAST analysis of DNA sequences. Unidentified refers to cases where a clear amplification product was obtained to suggest feeding but the sequence quality was not sufficient to resolve whether single or multiple hosts were used. 95% $CI_{unfed}$ = 21.8 - 28.9, 95% $CI_{single\ feeding}$ = 24.9 - 32.3%), 95% $CI_{multiple\ feeding}$ = 6.4 - 11.0%.

For *G. pallidipes* from Buffalo Ridge, although 113/154 fed, only 47 could be clearly assigned to host-feeding patterns because 66 produced DNA sequences that were not of sufficient quality to interpret (unidentified) (Table 4.4; Figure 4.4). Among the fed flies, there was a similar proportion that showed multiple as single host feeding and most of the latter were from African buffalo (19 samples) but there was also single feeding on African elephants, antelopes, warthogs and humans. In Zungu Luka a similar proportion of flies fed as in Buffalo Ridge but there was a much higher proportion of multiple host feeding (Table 4.4; Figure 4.5). Single feeding was on four antelopes, one elephant and one human. A different pattern was found in Mukinyo, where there was a much higher proportion that fed on single hosts (Table 4.4; Figure 4.6). The hosts were mostly African buffalo (91 flies) and African elephants (29 flies) but a wide range of other hosts were also identified (antelopes, humans, warthogs, giraffes, spotted hyena).

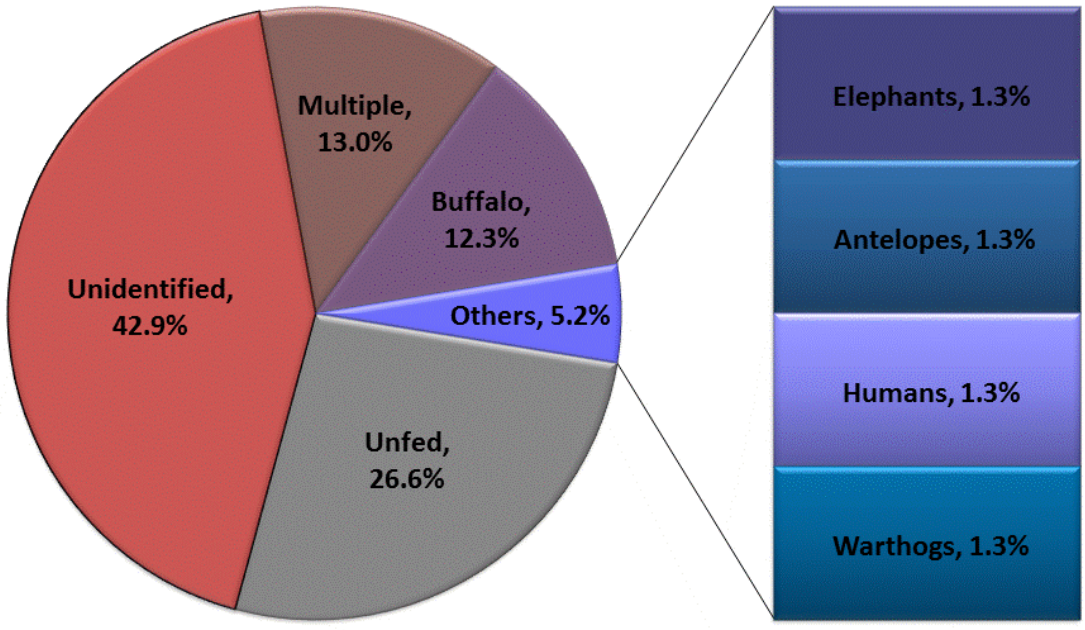



**Figure 4.4** Results of blood meal analysis of *G. pallidipes* from Buffalo Ridge (N = 154), as described in figure 4.3.
95% $CI_{unfed}$ = 19.6 - 33.6%, 95% $CI_{single\ feeding}$ = 11.5 - 23.5%, 95% $CI_{multiple\ feeding}$ = 7.7 - 18.3%.

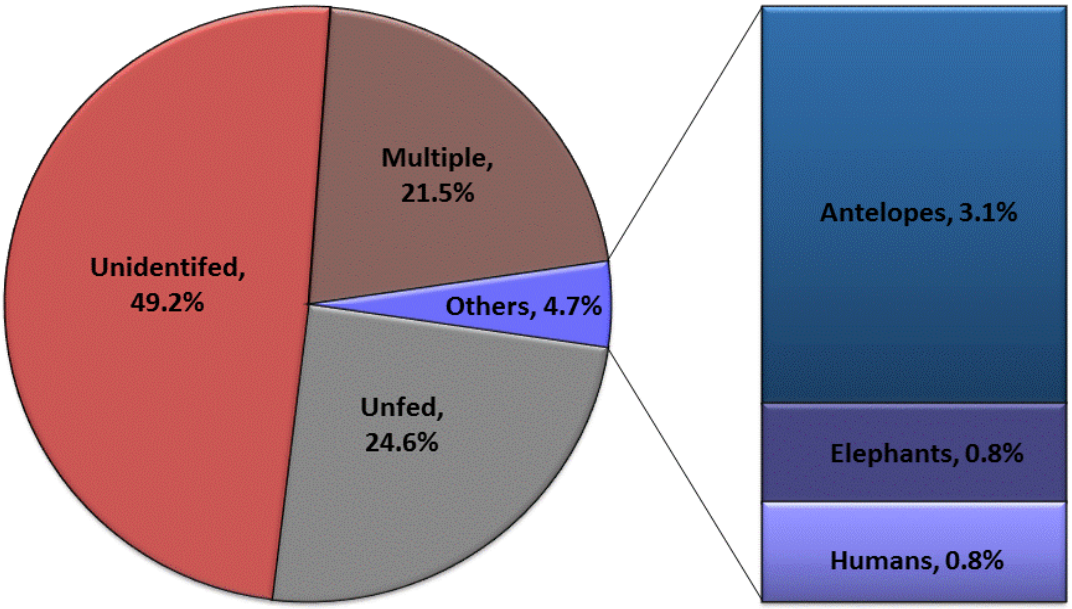



**Figure 4.5** Results of blood meal analysis of all *G. pallidipes* from Zungu Luka (N = 130), as described in figure 4.3.

95% $CI_{unfed}$ = 7.4 - 15.5%, 95% $CI_{single\ feeding}$ = 0.4 - 3.7%, 95% $CI_{multiple\ feeding}$ = 6.2 - 12.9%.

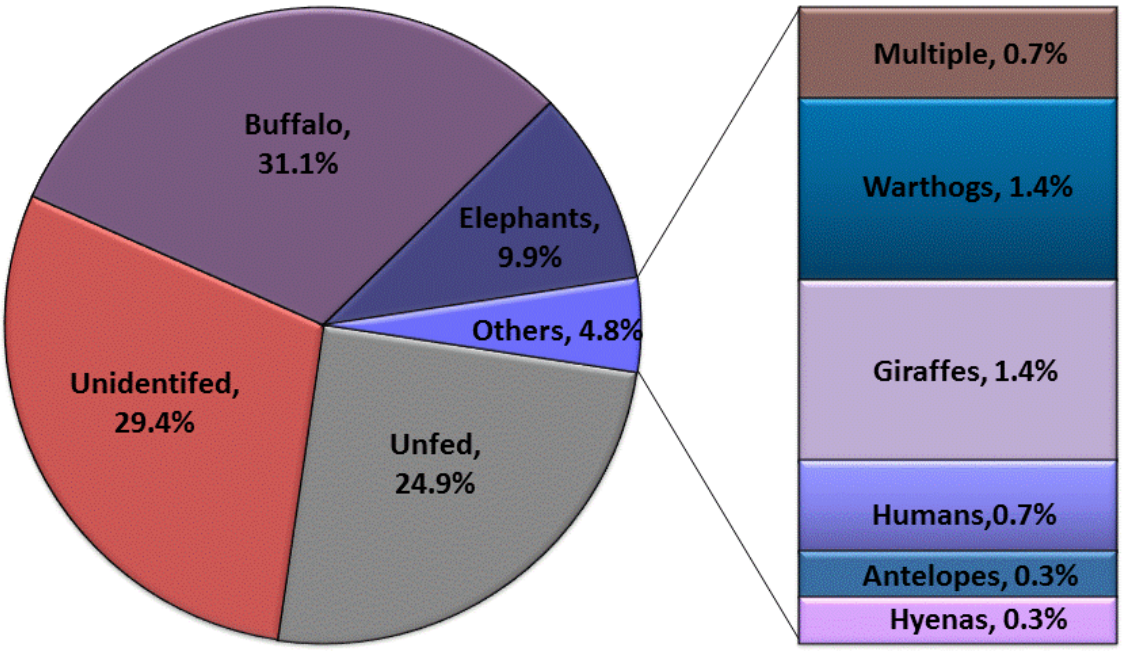



**Figure 4.6** Results of blood meal analysis of *G. pallidipes* from Mukinyo (N = 293), as described in figure 4.3.

95% $CI_{unfed}$ = 20.0 - 29.9%, 95% $CI_{single\ feeding}$ = 39.4.5 - 50.7%), 95% $CI_{multiple\ feeding}$ = 0 - 1.6%.

Plasmid sequencing confirmed multiple host identity in the 10 samples that were cloned because the direct sequences had shown heterozygous peaks (Table 4.5; Appendix C.2). Sources of tsetse blood meals from the Shimba Hills (Buffalo Ridge and Zungu Luka) included various combinations of humans, domestic animals, pests (mouse) and wild animals. For Mukinyo, sources of blood meals predominantly involved African buffalo and African elephants. Sequencing also confirmed single host feeding in the four samples included (Table 4.5) that appeared to be homozygous based on direct sequencing (warthogs and antelopes).

Based on BLAST analysis, sequences of my buccal samples were identical to cytb sequences from humans from Taiwan while all sequences obtained from human hosts in *G. pallidipes* were similar to cytb sequences belonging to people from Jamaica, Angola and Zambia. This confirmed that cytb PCR products were amplified from the abdomen contents of the tsetse samples, rather than from contamination during DNA extractions or PCR.

**Table 4.5** Host identification based on plasmid sequencing of the cytb gene.

Shown are the clone name (indicating the site and species of tsetse and the tissue part of tsetse flies), and the hosts identified, with numbers of clones with that sequence shown in parentheses. Individuals predicted to show single host feeding from the direct sequencing are indicated in boldface.

| Samples | Host determination (Numbers of plasmids) |
|---|---|
| **Buffalo Ridge (BRGp)** | |
| BRGp34AB | humans (4), goats (1), antelopes (1) |
| BRGp108AB | mouse (2), goats (2), humans (1), African buffalo (1) |
| BRGp109AB | African elephants (3) humans (1), mouse (1), goats (1) |
| BRGp126AB[a] | humans (2), goats (2), antelopes (1) |
| BRGp143AB | humans (3), goats (2), mouse (1) |
| **BRGp33AB** | **warthogs (6)** |
| **BRGp52AB** | **antelopes (6)** |
| **Zungu Luka (ZuGp)** | |
| ZuGp2AB | humans (2), goat (2) and mouse (1), warthogs (1) |
| ZuGp96AB | antelopes (4) and humans (2) |
| ZuGp102AB | chickens (3), humans (2) and insects (1) |
| **ZuGp104AB** | **warthogs (6)** |
| **ZuGp54AB** | **antelopes (6)** |
| **Mukinyo (MuGp)** | |
| MuGp9AB | African elephants (4), African buffalo (2) |
| MuGp287AB[a] | African elephant (2), African buffalo (1) |
| **Control** | |
| Buccal | human from Taiwan (6) |

[a] the sequencing failed for one plasmid from Buffalo Ridge and three plasmids from Mukinyo.

### 4.4.3 Prevalence of *Trypanosoma spp*. and associations of hosts with trypanosome status of *G. pallidipes*

Patterns of host feeding did not strongly correlate with infection status with trypanosomes (Table 4.6). For both trypanosome positive and trypanosome negative flies, there were more fed than unfed flies and more had fed on single than multiple hosts. A contingency $\chi^2$ analysis found no significant association between host-feeding status (fed or unfed) and trypanosome infection status but a weak association with feeding pattern (single vs multiple). No significant association was found when "unfed" was added to the single vs multiple comparison).

**Table 4.6** Association between trypanosome status with host-feeding status of *G. pallidipes* based on contingency $\chi^2$ tests.

The observed and expected values are shown, along with the $\chi^2$ test statistic and the P-value.

| Feeding status | Trypanosome status | | | | Correlation analysis | |
|---|---|---|---|---|---|---|
| | Observed | | Expected | | $\chi^2$ | P-value |
| | Positive | Negative | Positive | Negative | | |
| Fed | 176 | 255 | 177 | 254 | 0.04 | 0.841 |
| Unfed | 61 | 85 | 60 | 86 | | |
| Single | 63 | 102 | 69 | 96 | 3.95 | 0.047 |
| Multiple | 27 | 23 | 21 | 29 | | |
| Unfed | 61 | 85 | 61 | 85 | 3.95 | 0.127 |
| Single | 63 | 102 | 69 | 96 | | |
| Multiple | 27 | 23 | 21 | 29 | | |

However, *G. pallidipes* that fed on different hosts showed differences in trypanosome prevalence (Figure 4.7; Appendix C.3). Although the predominant single hosts of *G. pallidipes* were African buffalo, the percentage of trypanosome positive flies that fed on elephants (53.1%) was higher than for African buffalo (33.6%). The proportion of human-fed flies that were infected with trypanosomes was high (80.0%) when compared with warthogs (16.7%), giraffes (25.0%) antelopes (42.9%) and hyenas (0.0%).

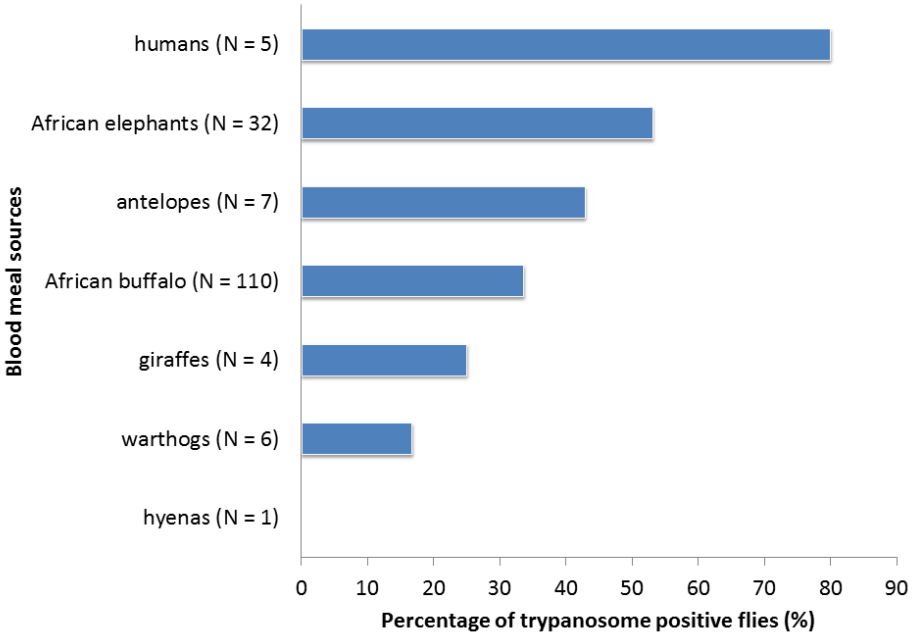



**Figure 4.7** Percentage of *G. pallidipes* that fed on different hosts that tested positive for trypanosomes.
Total number of tsetse flies that fed on each host are indicated in parentheses.

Among tsetse flies whose feeding source could be determined, the majority of the flies that were trypanosome positive showed the presence of single species of trypanosome and had fed recently on African buffalo. Since single species infections were highest in Muinkyo, it provided the most information about the relative prevalence of different species of trypanosomes that could have been obtained from different host species (Table 4.7). *Trypanosoma vivax* was the predominant species identified in all populations of *G. pallidipes* and occurred in flies that had fed recently on African buffalo, African elephants, antelopes, giraffes and humans. *Trypanosoma congolense* was found in Mukinyo and Buffalo Ridge but not Zunga Luka. All species of trypanosomes (including *T. godfreyi* and *T. simaie*) were found in flies that had fed on African Buffalo or African elephants. A high proportion of flies that fed on humans were trypanosome positive (9/11): five had *T. vivax* only; two had *T. brucei only*; one had *T. congolense* only and one showed a mixed infection with *T. vivax* and *T. brucei*. Of these flies, five fed only on humans (Table 4.7) and six had also fed on other hosts (up to 3 other additional hosts for the same flies; Table 4.8), including goats (6), mouse (4), chickens (1), antelopes (3), warthogs (1) and African buffalo (1). Only a few of the flies for which blood meal analyses showed single species of trypanosomes present were positive for *S. glossinidius*, but there did not appear to be an association with host species.

**Table 4.7** Single and mixed species of trypanosomes detected in *G. pallidipes* in relation to single blood meal hosts and *Sodalis* status.

Shown are the host determined, site of sample collection (the number of each host species identified at each site is indicated in parentheses after the name), the species of trypanosomes identified, the number of samples, and the status of *S. glossinidius*.

| Host species | Sites ( Number. of total infected flies) | *Trypanosoma spp.* | Number of samples | *Sodalis* status |
|---|---|---|---|---|
| Single trypanosome species detected | | | | |
| African buffalo | Buffalo Ridge (19) | Tv | 1 | 1 positive |
| | | Tc | 2 | 1 positive |
| | Mukinyo (91) | Tv | 18 | negative |
| | | Tc | 6 | negative |
| | | Tb | 1 | negative |
| | | Tg | 1 | negative |
| African elephants | Mukinyo (29) | Tv | 9 | negative |
| | | Tc | 2 | negative |
| | | Tb | 2 | negative |
| Antelopes | Zungu Luka (4) | Tv | 3 | 2 positive |
| Giraffes | Mukinyo (4) | Tv | 1 | negative |
| Humans | Buffalo Ridge (2) | Tv | 1 | negative |
| | Zungu Luka (1) | Tv | 1 | positive |
| | Mukinyo (2) | Tv | 1 | negative |
| | | Tb | 1 | negative |
| Warthogs | Buffalo Ridge (2) | Tc | 1 | 1 positive |
| Multiple trypanosome species detected | | | | |
| African buffalo | Buffalo Ridge (19) | Tv/Tc | 1 | negative |
| | Mukinyo (91) | Tv/Tc | 4 | negative |
| | | Tc/Tb | 1 | negative |
| | | Tb/Tv | 1 | negative |
| | | Tv/Tc/Tb | 1 | negative |
| African elephants | Mukinyo (29) | Tv/Tc | 2 | negative |
| | | Tv/Tb | 1 | negative |
| | | Tv/Tb/Ts/Tg | 1 | negative |

Tv = *T. vivax*; Tc = *T. congolesnse*; Tb = *T. brucei*; Tg = *T. godfreyi*; Ts = *T. simiae*.

Based on the samples that were cloned (which included only individuals that were positive for single species of trypanosomes), multiple feeding flies (N = 10) showed the presence of *T. vivax* (N = 4), *T. congolense* (N = 2) or *T. brucei* (N = 1) (Table 4.8). All *T. vivax* positive flies were collected from the Shimba Hills (Buffalo Ridge and Zungu Luka) and their predominant hosts were humans and domestic animals. The *T. congolense* positive tsetse fly from Buffalo Ridge fed only on humans and domestic animal while the one from Mukinto fed only on wildlife. The *T. br*ucei tsetse fly from Zungu Luka fed on antelope and humans. Only a single fly from each site was trypanosome negative but the Shimba Hills samples included both domestic (including humans) and wild hosts whereas the Mukinyo samples had fed only on wild hosts. Among the multiple feeding flies that were cloned, only two tested positive for *S. glossinidius* (Table 4.8). Both were from Zungu Luka and fed mostly on humans and domesticated animals (but one also had fed on a warthog). One tested positive for *T. vivax* but one was trypanosome negative.

Based on MCA 1 analyses, strong correlations among site of *G. pallidipes* collection, *Sodalis* status and host-feeding pattern were resolved into two clusters for dimension 1 (Table 4.9; Figure 4.8). Feeding on single hosts was related to the absence of *Sodalis* and the Mukinyo site (Cluster 1). On the other hand, *Sodalis* positive flies were associated with multiple host feeding and the Shimba Hills region (Buffalo Ridge and Zungu Luka) (Cluster 2). Two clusters of correlations among site, trypanosome status, age and sex were found in dimension 2 (Table 4.9; Figure 4.8). Trypanosome positive flies showed an association with juvenile and old males (Cluster 3), while trypanosome negative flies tended to be found in young females (Cluster 4).

**Table 4.8** Summary of trypanosome and *Sodalis* status, for the samples that were cloned for PCR products from the cytb gene and that showed evidence of having fed on multiple hosts.

Shown are the clone name (indicating the site and species of tsetse and the tissue part of tsetse flies), presence or absence of each species of trypanosome, the *Sodalis* status, and presence or absence of hosts identified. All samples showed presence of only single species of trypanosomes.

| Sample | Status | | Hosts Identified | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Trypanosomes | *Sodalis* | Humans | Goats | Mouse | Chickens | African elephants | African buffalo | Antelopes | Warthogs |
| ZuGp102AB | *T. vivax* | Positive | + | - | - | + | - | - | - | - |
| BRGp108AB | | Negative | + | + | + | - | - | + | - | - |
| BRGp109AB | | | + | + | + | - | + | - | - | - |
| BRGp126AB | | | + | + | - | - | - | - | + | - |
| BRGp143AB | *T. congolense* | Negative | + | + | + | - | - | - | - | - |
| MuGp9AB | | | - | - | - | - | + | + | - | - |
| ZuGp96AB | *T. brucei* | Negative | + | - | - | - | - | - | + | - |
| ZuGp2AB | Negative | Positive | + | + | + | - | - | - | - | + |
| BRGp34AB | | Negative | + | + | - | - | - | - | + | - |
| MuGp287AB | | | - | - | - | - | + | + | - | - |

**Table 4.9** Adjusted eta squared of the combination of variables in dimensions 1 - 3 in MCA 1.

| Variables | Dimension 1 | Dimension 2 | Dimension 3 |
|---|---|---|---|
| *Sodalis*_status | 0.661 | 0.078 | 0.012 |
| site | 0.828 | 0.475 | 0.192 |
| sex | 0.016 | 0.200 | 0.436 |
| age | 0.009 | 0.297 | 0.214 |
| Trypanosome_status | 0.057 | 0.335 | 0.120 |
| Feeding_pattern | 0.517 | 0.045 | 0.174 |

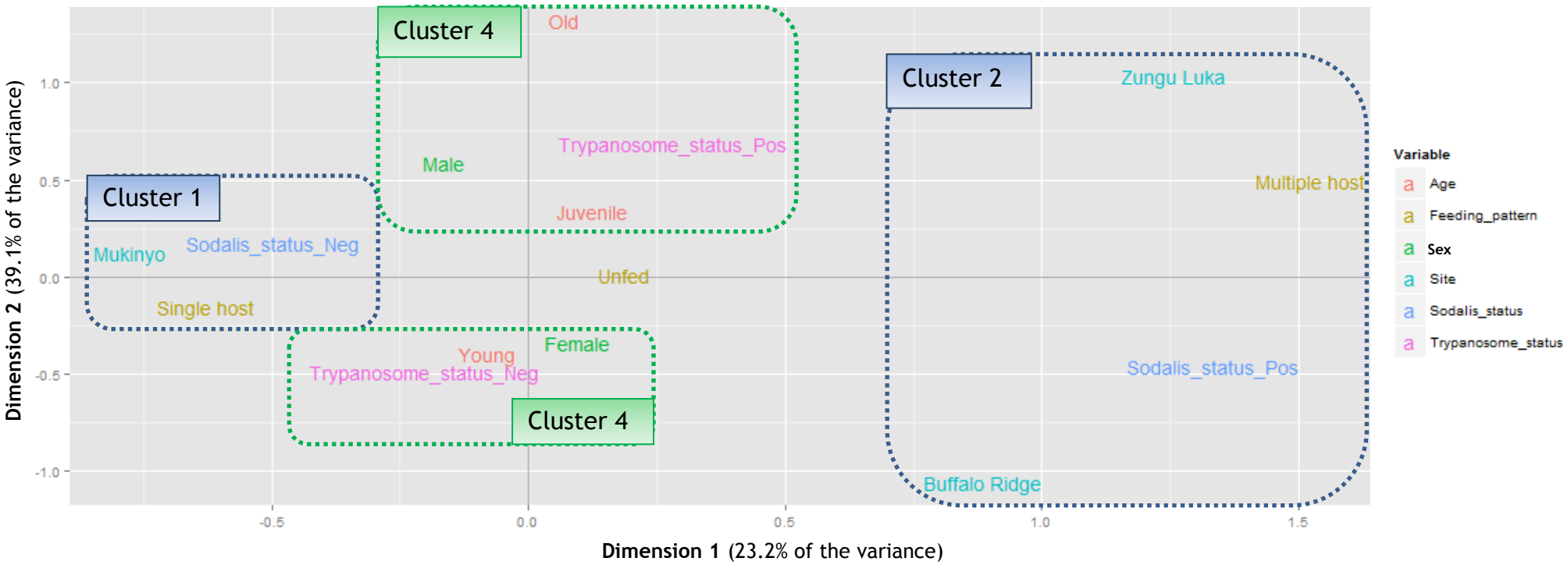



**Figure 4.8** Multiple Correspondence Analysis (MCA 1), showing associations of dimension 1 and 2 in relation to *Trypanosoma spp*. status, site (Buffalo Ridge, Zungu Luka and Mukinyo) and intrinsic factors of *G. pallidipes* (sex, age, *Sodalis* status; and host-feeding pattern) (N = 364).

## 4.5 Discussion

Based on PCR screening, a large proportion of the *G. pallidipes* sampled were classified in the fed status but the proportion of single and multiple host feeding flies varied by geographic region. The majority of single host feeding flies fed on African buffalo and African elephants but a wide range of other hosts was identified, including humans. Flies from Nguruman tended to feed on single hosts and there was no evidence for feeding on domestic animals (Table 4.7; Figure 4.6), in contrast to the Shimba Hills, where feeding on multiple host species was common and there was evidence for feeding on domesticated as well as wild animals, but never on only a single species (Table 4.8; Figure 4.4, Figure 4.5). Humans were fed on in both regions and showed high prevalence of trypanosomes; they were the only species that showed the presence of all three species of trypanosomes among cloned samples (Table 4.8). Patterns of feeding were not strongly associated with trypanosome status and MCA suggested that feeding patterns were associated more with geographic location (confounded with *Sodalis* status) than with intrinsic tsetse factors or with trypanosome status. Together, these results suggest that differences in host communities in different regions could influence risk of transmission between vectors and hosts in complex ways and so control programmes should be designed on a local scale, which considers all interacting components of the system.

### 4.5.1 Optimisation of blood meal analysis based on PCR

As for the other diagnostics discussed in chapters 2 and 3, my results demonstrate the importance of testing primers to be used for blood meal analysis. Despite the fact that the COI primers (VF1d_t1, VR1d_t1) were designed to be degenerate enough to amplify all vertebrates, the cytb primers used (Cb1, Cb2) provided more positive samples and a higher yield of PCR products. This supports a previous report that low numbers of samples amplified using the VF1d_t1, VR1d_t1 primers in *G. pallidipes* from Nguruman and Busia, Kenya: only 12/20 PCR products of flies that had fed (based on visual inspection) were amplified using this primer set (Nyawira, 2009). Moreover, a wider range of animal species was detected among blood meals of *G. pallidipes* from Nguruman using the cytb primers (Cb1, Cb2) than using the VF1d_t1, VR1d_t1 primers (Muturi *et al.*, 2011). They found that African buffalo, giraffes, spotted hyenas

and baboons were determined from blood meals using cytb, while the COI primers detected only African elephants and warthogs from the same set of samples. The sequences obtained using the cytb primers are divergent enough to distinguish among mammalian hosts of tsetse flies (Kocher *et al.*, 1989) but other types of hosts, such as monitor lizards that have been reported in blood meals of tsetse flies using the COI primers (Nyawira, 2009, Muturi *et al.*, 2011), would not be detected. In my survey, only a single sequence from a cytb amplicon was similar to an insect sequence available in Genbank (low percent query cover and low percent identity to *Drosophila permilis*) and mammalian host DNA was also detected in this sample, so these primers do not appear to present a high risk of amplifying tsetse DNA rather than host DNA. A study of *G. swynnertoni* from Tanzania using the COI primers reported only amplification of the fly DNA (Nyawira, 2009, Muturi *et al.*, 2011) but this could have been because the flies had not fed. In my study, since only a single primer combination was used for each gene, my results cannot tell which gene region would be optimal for resolving between closely related hosts. Therefore, I would recommend that a combination of the two primer sets might be the most thorough approach for identifying blood meal sources.

Based on the cytb gene survey, a relatively large number of collected tsetse flies were classified as having fed on multiple hosts but this was likely under estimated. Host determination from sequencing chromatographs that had double or triple peaks at single positions of the direct sequences could be used to indicate multiple feeding but was often ambiguous because it was difficult to distinguish poor sequencing quality from heterozygous sequences in some cases. Moreover, it was not possible to resolve the phase of polymorphisms to identify which hosts had fed on, based on direct sequencing. Cloning and plasmid sequencing methods reliably confirmed multiple feeding in the subset of samples that were checked. Only six clones were sequenced per individual so it was not possible to tell relative quantity of amplification products from each host but it was useful for predicting dominant hosts within geographic regions. In addition, more detailed plasmid sequencing, combined with experiments and host field surveys, could provide information on whether the hosts identified were those that had most recently been fed on, the preferred hosts, reflected relative availability of hosts in the environment, or indicated hosts with low defence

mechanisms. For example, *G. pallidipes* from Zungu Luka fed multiply on humans and domesticated animals (e.g. chickens and goats) (Table 4.8), but whether this is because these were the most abundant hosts available at the time, the preferred host, or the blood meals that lasted the longest is not clear. However, this was based on a few samples and there was also a high proportion of samples whose hosts could not be resolved (Figure 4.5).

It is also not possible to tell the rate of feeding on different hosts based on my blood meal analysis. A recent study found evidence that blood meals could be used to identify hosts in sand flies that had fed on humans and chickens up to 48 h post feeding (Sant'Anna *et al.*, 2008) and 72 h after *Culex pipiens* had fed on humans (Kent and Norris, 2005). However, there has been no report about the time duration that sources of blood meals are detectable in tsetse flies after feeding. The duration of detection will also depend on the rate of clearing (complete digestion) of the blood meal, which has been found to be 4 - 5 day after blood meal for *G. morsitans morsitans* (Langley and Stafford, 1990). Theoretically, the number of clones could be used to predict which host was last fed on, but this would also depend on the rate of feeding of the fly (e.g. if they were interrupted and switched hosts very rapidly, more than one blood meal might have a similar DNA concentration) and lack of bias in PCR amplifications. Thus, studying the maximum times after feeding on different host species that blood meal sources can be detected in tsetse flies using PCR methods should be further investigated, combined with detailed documentation of the relative density of available hosts in the particular environments where the flies have been feeding. In addition, using real time PCR might provide a more precise estimate of the interval of time between each meal and DNA concentration left in the abdomen parts.

### 4.5.2 Blood meal analysis

Across populations, African buffalo were the main hosts of *G. pallidipes* found in my study, which supports previous reports from the International Atomic Energy Agency (IAEA, Vienna) that ruminants are attractive to adult *G. pallidipes*, *G. fuscipes* and *G. brevipalpis* (Harraca *et al.*, 2009). Selective feeding of savannah tsetse flies might depend on odour attraction to large-size animals (such as cattle, buffaloas described by Torr and Vale (2015). However, Clausen *et al*

(1998) reported variation in dominant hosts in different areas, which might relate to overlapping of habitat and activities between tsetse flies and hosts. This has also been found in other studies in eastern Africa (Table 4.10). For example, approximately half of the blood meal sources of *G. pallidipes* from Kenya and Ethiopia belonged to ruminants when blood contents were screened with an enzyme-linked immunosorbent assay (ELISA) (Clausen *et al.*, 1998) (Table 4.10). In the Busia region of Kenya, three different studies found that bovines (mostly cattle) were the primary hosts (Okoth *et al.*, 2007, Nyawira, 2009, Muturi *et al.*, 2011). However, in other regions differences in dominant hosts have been found between studies conducted in different years and by different groups of researchers (Table 4.10). For example, in Nguruman, warthogs have been found in consistently high proportions across four different studies (Bett *et al.*, 2008, Okoth *et al.*, 2007, Nyawira, 2009, Muturi *et al.*, 2011) but the relative proportion of ruminants has varied. In my study, at the Mukinyo site within the Ngruman region, warthogs were found relatively rarely whereas buffalo were identified in nearly a third of the flies surveyed (Table 4.10). No domestic cattle were detected in my study. Whether this is due to differences in methodology for detecting hosts, differences in the season that the studies were conducted (which could affect the relative availability of migratory buffalo or nomadic cattle), differences in the vegetation type or cover at particular sites that flies were sampled within the main geographic region, or particular environmental conditions in different years is not clear. It is possible that ruminants are the preferred hosts but when they are not readily available, tsetse flies move to feed on other available hosts such as African elephants, warthogs and antelopes. What is intriguing from my study in Mukinyo is that there was a very low proportion of flies that fed on multiple hosts and a predominance of buffalo among the single host blood meals that could be resolved. Although the blood meal analysis cannot predict host preferences, it is possible that buffalo are abundant hosts that are easy to feed on and so flies could learn to return to the same host species (Bouyer *et al.*, 2007), especially females, which are more motile than males (Vale *et al.*, 2014). However, study of feeding preference should be further investigated by giving host choices for tsetse rather than assuming that a high frequency of host feeding reflects preferences rather than relative host abundance.

There are several reasons that could explain the detection of multiple hosts in a blood meal. If feeding on an initial host is interrupted or too low quality ("unsuccessful"), flies might switch hosts to obtain a sufficient fat-haematin level for development of larvae in the uterus of females (Randolph and Rogers, 1978) and reproductive success in males (Yuval, 2006). In Zimbabwe, feeding success of tsetse flies on cattle was reported to be lower, feeding was interrupted earlier, and the time taken to complete a meal was lower than for stable flies but unsuccessful feeding was attributed more to host defense (69%) and than to competition with other biting insects (31%) (Schofield and Torr, 2002). I found nearly 30% of trapped *G. pallidipes* had fed only on single hosts, but this was due mostly to flies sampled from Mukinyo, where domesticated animals were not detected in the recent blood meals. This could suggest a difference in the rate of "successful" feeding in different regions, which could be related to the types of host present. It is possible that wildlife react to tsetse flies with less defensive activities than domesticated animals or that fewer other flies interrupt feeding of the tsetse flies on large wildlife. Humans could be inappropriate hosts because they camouflage their odours, apply chemical repellents, and react strongly to tsetse bites, which could result in unsuccessful feeding (Baylis, 1996, Baylis and Nambiro, 1993, Hargrove, 1976) that would lead to host switching. I found evidence for up to four different hosts detected within the time window of blood meal detection, but this was only found in some of the sites sampled (Table 4.5). Whether the differences in rates of multiple feeding is due to interruptions leading to more frequent feeding in flies sampled in some regions, or due to the presence of a wider range of hosts in a limited geographic area cannot be distinguished based only on the blood meals because it is possible that flies also feed multiple times on the same hosts in some reason. Feeding behaviours of individual flies also might differ depending on age and sex, which also could contribute to differences among populations. For example, the frequency of feeding in females has been found to increase when there are larvae in the uterus (Langley and Stafford, 1990), but this also varies by species of tsetse flies. In a laboratory experiment, *G. pallidipes* females have been found to require more blood meals than *G. m. morsitans* to achieve the same reproductive performance (Langley and Stafford, 1990). So, it is possible that *G. pallidipes* also feed more frequently in the wild.

**Table 4.10** Summary of blood meal sources from G. pallidipes in previously published studies.

Shown are the source of tsetse fly samples (including the year of collection in parentheses), the host determined from blood meal analysis, the proportion of each host found and the reference source of data.

| Site of collection (year) | Blood meal sources | Proportion (%) | References |
|---|---|---|---|
| Kenya and Ethiopia (1983 - 1993) | ruminants* | 52.2 | Clausen *et al.* (1998) |
| | suvid** | 36.2 | |
| | mammals*** | 8.2 | |
| | Primates | 2.3 | |
| | monitor lizards | 0.6 | |
| | Avian | 0.5 | |
| Nguruman, Kenya (2004 - 2005) | bovines (did not indicate species) | 16.8 | Bett *et al.* (2008) |
| | Antelopes | 9.8 | |
| | Warthogs | 28.3 | |
| | Elephants | 13.3 | |
| | Zebras | 11.5 | |
| | Giraffes | 6.2 | |
| | Goats | 6.2 | |
| | Ostriches | 5.3 | |
| | Sheep | 2.0 | |
| | Humans | 0.9 | |
| Nkineji, Kenya (2004 - 2005) | bovines (did not indicate species) | 53.9 | |
| | Sheep | 15.4 | |
| | Antelopes | 15.2 | |
| | Goats | 11.5 | |
| | Baboons | 3.8 | |
| Busia, Kenya | bovines (did not indicate species) | 58.0 | Okoth *et al.* (2007) |
| | Warthogs | 14.8 | |
| | Humans | 4.9 | |
| | Antelopes | 2.5 | |
| Nguruman, Kenya | Warthogs | 35.0 | |
| | Bovines (did not indicate species) | 16.7 | |
| | Antelopes | 13.3 | |
| | Giraffes | 11.7 | |
| | African buffalo | 6.7 | |
| | Zebras | 1.7 | |
| | Ostriches | 1.7 | |
| | Lions | 1.7 | |
| | Goats | 1.7 | |
| Nguruman, Kenya | African elephants | 46.2 | Nyawira, (2009) |
| | Warthogs | 38.5 | |
| | African buffalo | 7.7 | |
| | Baboons | 7.7 | |
| Busia, Kenya | Cattle | 100.0 | |
| Nguruman, Kenya (2008 - 2009) | African elephants | 46.2 | Muturi *et al.* (2011) |
| | Warthogs | 38.5 | |
| | African buffalo | 0.8 | |
| | Baboons | 0.8 | |
| Busia, Kenya (2008 - 2009) | Cattle | 100 | |

*The majority of ruminants were buffalo (265), bushbuck (201), cattle (78), non-specific ruminants (44), wild ruminants (39), duikers (8), waterbucks (7), and small ruminants (3), (N =645).

** The majority of suids were bushpigs (179), warthogs (135), non-specific *Suidae* (113), and wild *Suidae* (22), (N = 449).

*** The majority of mammals were hippopotamus (50), elephants (17), *Equidae* (11), *Camelidae* (8), *Rodentia* (5), *Felidae* (excl. lion) (4), lions (3), and *Canidae* (3), (N = 101).

Determination of feeding on multiple hosts in some of the other studies cited was limited by antibodies for ELISA methods or the PCR-based methods used, so that comparisions between relative rates of single and multiple-host feeding could not be made. However, my study suggests that this could be important for determining relative risk of trypanosome transmission among vectors and hosts. Among the three sites of sample collection, there was not very much difference in the proportion of tsetse flies that did not feed (24.6 - 26-6% in Figure 4.4 - Figure 4.6) but there were variation in the feeding patterns of flies that did feed. The majority of chromatogram profiles from the Shimba Hills indicated multiple-host feeding whereas the majority from Nguruman (Mukinyo) showed single host feeding. There are several reasons that could explain the results. Firstly, the location of fly trapping in the Shimba Hills could have had lower density of the preferred hosts. Thus, tsetse flies had to seek and feed on other available animals. The plasmid sequencing of DNA samples from Zungu Luka, which is close to a human settlement, showed that *G. pallidipes* fed on goats, chickens, mouse, warthogs, antelopes and humans but did not feed on cattle, African buffalo or African elephants (Table 4.5). Although it has been predicted that there are many lifestock (cattle, buffalo, cattle, sheep, goat, pig and poultry) in the Mukinyo region (Appendix A.2 – Appendix A.8) and the Shimba Hills region is approximately less than one km away from predicted available animal husbandry, domesticated animals were not detected from blood meals of single feeding flies from any sites. *Glossina pallidipes* might prefer feeding on wildlife more than available domestic animals, or the density and location of domesticated animals was lower than the predicted maps in tsetse-fly habitats within each site (Appendix A.2 – Appendix A.8). Unfortunately, distributions of wildlife in these geographic regions, which would be very useful for this discussion, have not been quantified. Secondly, tsetse flies from the Shimba Hills could have had more choices for their meals than tsetse flies from Mukinyo because the Shimba Hills is a National Reserve area so a wide range of wild animals are available (Appendix C.4). Thirdly, the Shimba Hills is a tourist attraction, where many people visit. However, relatively few tsetse flies fed singly on humans; only five humans were identified among the flies that had fed on single hosts (N = 33). This could be an accidental human feeding of those flies. In contrast, the cloning revealed that humans were involved in mixed feeding in all eight of the flies that had fed on multiple hosts in the Shimba Hills

(but not in the two from Mukinyo), so it is possible that the presence of humans influences the feeding behaviour of tsetse flies. Hargrove (1976) found that presence of humans not only repelled tsetse flies but also inhibited the landing response to approach other potential hosts nearby. This might causes flies to search more broadly for more appropriate hosts. I could not find previous studies that have assessed the relationships between host availability, host preferences and patterns of multiple feeding in tsetse flies, but it is known that host selection of tsetse flies depends on confounding factors, for example, available host community, attractive properties of hosts (size, colour, odour, movement) (Onyiah, 1980, Torr, 1989, Willemse and Takken, 1994), host defenses, and foraging experience (Bouyer *et al.*, 2007). For mosquitos, nutritional value of blood, geographic regions, seasons, and ecology of microhabitats can be causes of spatial and temporal variaition in the behaviour of host selection (Lyimo and Ferguson, 2009). In order to define host preference in field studies, the "feeding index" is often used, which is the ratio between observed relative abundance of two host species (Kay *et al.*, 1979). On the other hand, sophisticated experiments have been conducted to examine how relative availability of preferred hosts influence mosquito behaviours, using semi-field systems that allow manipulation of host choices under controlled but environmentally relevant conditions (Lyimo *et al.*, 2013). It would be valuable to extend these types of studies to tsetse flies in order to test hypotheses about what drives their host-feeding patterns.

### 4.5.3 Prevalence of *Trypanosoma spp*. and associations of hosts with trypanosome status of *G. pallidipes*

There was not a strong association between prevalence of trypanosomes and host-feeding patterns in the tsetse flies. The reason for initially considering fed vs unfed separately was that the fed class included a large proportion of flies for which it could not be determined whether they had fed single or multiply due to poor sequence quality. This substantially reduced the sample size. However, the lack of an association is consistent with the MCA, which suggested a stronger correlation among feeding pattern, site and *Sodalis* status than trypanosome status, sex and age of tsetse flies. Adding feeding status to the already complicated GLM models described in chapter 3 resulted in multiple three-way interactions (data not shown), again emphasising the complexity of associations.

Leak (1998) suggested that trypanosome infection might influence feeding success and feeding behaviour of the flies. For example, high numbers of *T. congolense*, which attach to the cuticle of the proboscis, could interrupt feeding and result in more frequent probing. Thus, an association between the frequency of feeding and trypanosome infection status should be further studied in laboratory experiments to indicate whether "trypanosome infection causes a feeding pattern change" or "difference in feeding pattern promotes trypanosome infection".

Bouyer *et al.*, (2007) suggested that repeated feeding on the same host species was likely to increase risk of trypanosome transmission within species, but to decrease risk between species. In my study, for flies that had fed on multiple hosts and were positive for trypanosomes, four were identified as *T. vivax* (which is not pathogenic to humans), two were *T. congolense* (which is not known to be pathogenic to humans) and one was *T. brucei* (which has some subspecies that are known to be zoonotic (Maudlin *et al.*, 2004). *Trypanosoma brucei* has not previously been reported in my study area but if tsetse flies tend to feed on multiple hosts when they only have small domesticated species available, this could increase risks to humans in regions where the human associated parasites are common.

African buffalo were the most frequent hosts detected, but a larger proportion of flies that had fed on African elephants were trypanosome positive (Figure 4.7). It has been reported that African buffalo are efficient at limiting trypanosome parasitaemia and show few or no signs of *T. congolense* infection (Grootenhuis *et al.*, 1990). *Trypanosoma congolense* and *T. vivax* resistance in African buffalo has also been suggested (Olubayo *et al.*, 1996). It is possible that African buffalo are more resistant to trypanosome invasion than African elephants. However, trypanosome infection experiments in African elephants should be further studied to test this. There also could be a dilution effect of common hosts (for example, African buffalo, African elephants), when there is potentally changing of host structure and so higher diversity of potential hosts species, as suggested in a previous report of *T. cruzi* (which causes Chagas disease) (Gottdenker *et al.*, 2012). That study also found that blood meal species diversity, host population and habitat type were important determinants

of vector infection. So, the relationship between which hosts are fed on and risk of trypanosome transmission are not straight forward.

## 4.6 Conclusions

Identification of the hosts that *G. pallidipes* fed on based on direct PCR sequencing revealed evidence for both use of a wide range of hosts and multiple feeding bouts by individual flies. However, the direct sequencing method could not determine blood meal sources of tsetse flies that fed on multiple hosts. I would suggest that cloning and plasmid sequencing should be used to confirm multiple feeding flies and determine the dominant hosts. Based on the single host sequences, the flies in my study fed predominantly on wildlife, and most frequently on African buffalo. *Trypanosoma vivax* was the main pathogenic species found in tsetse flies; however, one tsetse fly that fed on humans was infected with *T. brucei* suggesting that more detailed investigations are warranted for the relative risk to humans posed by the evidence for feeding on multiple hosts in my study. In order to assess HAT, I would recommend targeting subspecies of *T. brucei* (*T. b. gambiense, T. b. rhodesiense* and *T. b, brucei*) and host determination in different species of tsetse flies from these areas. Although I did not find evidence for extensive feeding on cattle in my study, ruminants (both wild and domesticated) in particular areas should be monitored for trypanosomiasis as a routine investigation, especially in areas in which *G. pallidipes* lives. In addition, since many species of animals were identified, associations of trypanosome status within each host species should be further studied.

# Chapter 5 General discussion

African trypanosomiasis, a neglected tropical disease in humans, domestic animals and wild animals cyclical transmits via tsetse flies. Wide variation in trypanosome infection rates have been reported in tsetse flies based both on field and laboratory studies (Moloo *et al.*, 1987, Mihok *et al.*, 1992, Mekata *et al.*, 2008, Dennis *et al.*, 2014, Desta, 2014, Salekwa *et al.*, 2014, Nthiwa *et al.*, 2015). Thus, vector competence, which refers to the ability of trypanosomes to infect tsetse flies and transmission between other hosts, has been studied in relation to different individual tsetse-specific factors, geographic location and the presence of endosymbionts using different samples and methods (Moloo *et al.*, 1992, Geiger *et al.*, 2005b, Geiger *et al.*, 2007, Akoda, 2009, Motloang *et al.*, 2012), as reviewed previously. Howerever, interactions among all these factors have not been investigated previously. The main hypothesis of my thesis was that studying these factors in isolation might be misleading because of the expectation that at least some of these factors would not operate independently. My main aims were thus to take an integrated approach to investigating associations among variables using Generalised Linear Models (GLMs) and multivariate analyses. This could be informative for predicting better understanding biology of trypanosome infection in tsetse flies. My study results showed clearly that the prevalence of different species of trypanosomes was significantly associated with different interactions among the intrinsic factors of different tsese species in different grographic regions, which could be partly due to differences in the communities of hosts available. Results of specific associations with intrinsic of tsetse flies in each subpopulation suggested a complicated bilological combination among vectors, trypanosomes, hosts and geographic regions. My results also suggest that previous conclusions from other studies about the role that endosymbionts play in enhancing trypanosome susceptibility might have been confounded by not considering these other types of interactions: I found no clear association with *Sodalis* as an independent variable but it sometimes was involved in interactions with other factors. The implications are that interventions would need to be targeted to consider particular communities of hosts, vectors and parasites in particular geographic regions.

## 5.1 Molecular application for trypanosome epidemiology

Molecular analyses have been applied in many studies of epidemiology (Foxman and Riley, 2001, Lymbery and Thompson, 2012) because they are appropriate for a large number of samples, and they have high sensitivity and high sensitivity. In my study, trypanosomes were identified based on PCR screening, confirmed by sequencing and comparison with published databases. Similar approaches were used to identify *S. glossinidius* presence and host-feeding patterns of tsetse flies. My results were variable depending on which primers were used; since a number of primers have been designed for identification and genetic diversity of both trypanosomes (Adams *et al.*, 2010, Auty *et al.*, 2012, Botero *et al.*, 2013) and *S. glossinidius* (Aksoy *et al.*, 1997, Farikou *et al.*, 2011a, Farikou *et al.*, 2011b), appropriate tissue samples, primers and PCR conditions should be tested for specific species of trypanosomes and *S. glossinidius* identification in tsetse flies from each region.

The tissue type screened could affect interpretation of prevalence of pathogens in vectors. The midgut and salivary glands of tsetse flies are located in the abdomen parts, where biological PCR inhibitors exist; for example, complex polysaccharides, proteinase, urea and blood contents (Alaeddini, 2012). Using head plus proboscis parts avoids these types of substances, while still allowing detection of trypanosomes and *S. glossinidius,* which are also found in these tissues. *Trypanosoma congolense, T. vivax* and *T. brucei* are found in mouth parts (Lloyd and Johnson, 1924), the route of transmission to the final hosts, while *S. glossinidius* is found in the haemolymph (Cheng and Aksoy, 1999) and so can be found all over the body. I found more positive trypanosome samples using head plus proboscis than abdomen parts (see chapter 2) but no differences in results of *Sodalis* screening between the two tissues (see chapter 3). The head plus proboscis is easier to prepare and there is less risk of contamination than for more delicate internal organ dissection so use of this single tissue might be most informative about relationships between trypanosome prevalence and endosymbiont presence.

Choice of PCR primers and reaction conditions could also seriously affect interpretation of infection patterns. Prevalence of trypanosomes (see chapter 2)

and *S. glossinidius* (see chapter 3) were found with different rates using different available primers and PCR conditions. The different screening results emphasised that choice of identification methods is a critical issue for epidemiological studies. In this study, I concluded that the ITS-1 CR and BR universal primers and PCR cycles described by Njiru *et al*. (2005) and the nuclear Hem primers and PCR conditions as described by Pais *et al*. (2008) were suitable for screening trypanosomes and *S. glossinidius* in tsetse fly samples from Kenya. However, both of these primers amplify relatively small fragments of DNA, which limits the potential to assess genetic variation of trypanosomes and *S. glossinidius* in tsetse flies. The GPO1 primers amplified a longer fragment that demonstrated sequence variation among the *Sodalis* isolated from tsetse flies but use of primers targeting extrachromosomal DNA risks underestimation of presence of the endosymbionts if they lack the plamids. Although the other set of plasmid primers (pSG2) resulted in higher predicted prevalence of *Sodalis* than the nuclear primer (suggesting that absence of the plasmid is not of concern), the sequences for some of the *G. pallidipes* and *G. longipennis* samples could not be identified using BLAST, which showed that using pSG2 primers for *Sodalis* screening in the samples risked over diagnosis. However, the small size of the amplified product limits usefulness for identifying species based on the sequences. For trypanosomes, I attempted to obtain longer sequences to assess genetic variability in more detail by using primers targeting both the ITS-1 and ITS-2 regions for which Auty *et al*. (2012) found extensive sequence polymorphism. Unfortunately, not enough studies have used these primers to allow identification of the trypanosome species by comparison with the Genbank database to make this a reliable diagnostic approach at this time. However, discrepancies between predicted trypanosome species based on fragment sizes of the ITS-1 region and the sequences that I obtained, emphasises the importance of confirming identification based on sequencing. Therefore, I would recommend that DNA sequencing is a necessary method to confirm amplification for biological identification based on PCR techniques.

For blood meal analysis, mitochondrial cytochrome B (cytb) and cytochrome C oxygenase1 (COI) genes have mainly been used to identify sources of blood content in guts of blood sucking insects. In my study, feeding patterns of *G. pallidipes* were identified using the Cb1, Cb2 primers targeting the cytb gene

(Kocher *et al.*, 1989) because they provided more positive amplification products than using the VF1d_t1, VR1d_t1 primers targeting the CO1 gene (Ivanova *et al.*, 2006, Ward *et al.*, 2005). These primers can determine a wide range of mammals (Nyawira, 2009, Muturi *et al.*, 2011) and so they are appropriate for the study areas, which were anticipated to include a wide variety of potential hosts. I also was able to detect multiple feeding on different hosts based on heterozygosity of the chromatographs, including feeding on different individuals within species (i.e. multiple genotypes of elephants and buffalo were both detected). However, there could be a trade-off in primer selection between accurate targeting of particular hosts and detecting the full range of hosts. For example, a limitation of the cytb primers that I used is that they would not detect other vertebrate hosts, such as reptiles, which tsetse flies are also known to feed on (Muturi *et al.*, 2011, Nyawira, 2009). Alternative approaches have been suggested to target particular host species rather than attempting to examine the full range of hosts. For example, multiplex PCR using four forward primers designed to detect humans, dogs, cattle and pigs with a universal reverse primer detected dual and multiple feeding of *G. palpalis* and *G. tachinoides* from Nigeria on these hosts (Karshima *et al.*, 2016) but the paper did not describe confirmation of host species using sequencing. Once again, my results indicate the usefulness of sequence-based approaches to species identification, rather than relying only on presence or absence of amplification products of the expected size.

## 5.2 Tsetse fly distribution and association with trypanosome infection

Variation in both species of tsetse flies (Cecchi *et al.*, 2015) and prevalence of trypanosomes (Simo *et al.*, 2012, Salekwa *et al.*, 2014, Duguma *et al.*, 2015) in different regions has been reported previously. There has also been investigation of differences in trypanosomes infection related to sex and age of tsetse flies but the associations remain unclear (Jackson, 1946, Moloo *et al.*, 1992, Woolhouse *et al.*, 1993, Peacock *et al.*, 2012b, Walshe *et al.*, 2011). An important contribution of my study is simultaneously considering multiple explanatory variables and investigating whether interactions between factors

could affect interpretation of apparent patterns. Moreover, there have not been previous reports of the use of multivariate approaches to investigating these complex interactions.

The composition of the tsetse fly communities varied extensively by geographic region, which prevented separation of geographic effects from species of tsetse flies present. Since age distributions could also vary by species and sex, there was also a risk of nonindependence of these factors. From 1090 collected tsetse flies, *G. austeni, G. brevipalpis* and *G. pallidipes* were found in the Shimba Hills (Buffalo Ridge and Zungu Luka) while *G. pallidipes* and *G. longipennis* were collected from Nguruman (Mukinyo and Sampu) (see chapter 2). *Glossina pallidipes* was the predominant species found in both regions and although females were mostly more abundant than males and young and juvenile flies were more common than old flies for each tsetse fly species, the distributions did vary by species (Figure 2.10 and Figure 2.11). Finding of different tsetse fly communities and densities is consistent with previous publications (Mbahin *et al.*, 2013, Cecchi *et al.*, 2015), including finding *G. pallidipes* as the predominant species in Kenya (Nthiwa *et al.*, 2015). Tsetse flies are thought to be at their highest densities in humid seasons (Moggridge, 1949), corresponding to observations that after a couple months of rains, a high rate of trypanosome infections occur in cattle (Tarimo-Nesbitt *et al.*, 1999). This implies that distribution of tsetse flies is related to environmental circumstance and location, which influences trypanosome distribution. My results suggest that multivariate approaches throughout the year to test associations between the environmental determinants of tsetse fly distributions (as well as age, sex and *Sodalis* distributions) and trypanosome infections could be more appropriate than analyses that assume independence of factors.

My results also suggest that it might be important to design studies to test associations separately by species of trypanosomes. *Trypanosoma vivax* was the main pathogenic trypanosome species found in most of the populations of tsetse flies collected (see Chapter 2). This supports results from a previous study conducted in the Mtito Andei Division (approximately 290 km south east of Nairobi, Kenya), which found that *T. vivax* was the main species occurring in *G. pallidipes* and *G. longipennis*, based on nested PCR (Nthiwa *et al.*, 2015). *Trypanosoma vivax* was also the most common pathogen found in *G.*

*swynertonni*, *G. m. morsitans* and *G. pallidipes* sampled from the Simanjiro district of Tanzania (Salekwa *et al.*, 2014). In a laboratory study, *G. pallidipes* was found to be infected with *T. vivax* at a higher rate than *T. congolense* and *T. b. brucei* (Moloo *et al.*, 1992). Varied experimental *T. vivax* infection rates also have been reported for different species of tsetse flies: *G. m. centralis* from Tanzania; *G. austeni* from Zanzibar; *G. p. palpalis* from Nigeria; *G. p. gambiensis* from Burkina Faso; *G. f. fuscipes* from Central African Republic; *G. tachinoides* from Chad; and *G. brevipalpis* from Kenya (Moloo *et al.*, 1987). In my study, *G. austeni* showed the highest bias towards *T. vivax* infection whereas *T. vivax* was not the most dominant species found for *G. brevipalpis* (Figure 2.21). This could suggest that the high overall prevalence of *T. vivax* is related to higher susceptibility in *G. pallidipes* and *G. austeni*, which were the common species found in my study. Previous studies have also suggested that there could be differences related to different strains of *T. vivax* (Moloo *et al.* 1987) so it could be important to study the dynamics of relationships between tsetse species and trypanosomes using locally collected strains of the pathogens found in each geographic region.

Different tsetse fly communities also showed different patterns of prevalence in relation to sites, species, sex and age for different species of trypanosomes. Mixed infections increased with age, which could imply that trypanosomes establish permanent infections in tsetse flies, as suggested in a previous study (Soumana *et al.*, 2014). Based on my GLM analyses (see chapter 2), sex was associated with the presence of trypanosomes, but often in combination with age and not for all species of trypanosomes or for all species of tsetse flies. Thus, previous simple conclusions about higher prevalence in females (Isaac *et al.*, 2016) might have been misleading because of differences in sex bias in different regions or interactions with other factors such as age, species of tsetse or species of trypanosomes. Similarly, conclusions about the association of susceptibility to trypanosomes with age of flies might be too simplistic. Differences in virulence of each trypanosome species and defence mechanisms of each tsetse species could possibly be important for driving this variation in associations. So, experiments to identify particular risk factors for trypanosome infection might have to be set up with more complex factorial designs than have been attempted previously.

## 5.3 Prevalence and association of *S. glossinidius* status

My study did not support previous predictions that the endosymbiont *S. glossinidius* promotes trypanosome infections in tsetse flies. *Sodalis glossinidius* produces chitinase, which breaks down chitin and produces N-acetyl-D-glucosamine (Welburn *et al.*, 1993), which is thought to lead to increased trypanosome susceptibility in tsetse flies (Welburn and Maudlin, 1999). Presence of the secondary endosymbiont has thus been suggested to favour trypanosome infection in tsetse flies (Farikou *et al.*, 2010a). In field studies, although a significant association was reported between trypanosomes and *Sodalis* infection in *G. pallidipes* from Kenyan coastal forests (trapped in 2009–2011), there was no significant association found for *G. austeni* from South Africa (trapped in 2008) (Wamwiri *et al.*, 2013). Dennis *et al.* (2014) reported that there was no association between presence of *Sodalis* and trypanosome in *G. brevipalpis, G. m. morsitans* and *G. pallidipes* sampled from Luambe National Park, Zambia. My results suggest that this discrepancy among studies could be because potentially confounding factors were not considered.

In my study, if I had only compared the distribution of *Sodalis* and trypanosomes without considering species of tsetse fly, age or sex distribtutions, geographic location, or species of trypanosome, I would have concluded that there was at least a weak association between the endosymbiont and the parasites. Prevalence of *Sodalis* was high in the Shimba Hills but very low in the Nguruman region (Figure 3.10d), where there was also low prevalence of trypanosomes (Figure 2.22a). However, within the Shimba Hills region, there was no association, with equal proportions of trypanosome-infected flies testing postive or negative for *Sodalis*. The apparent overall association was thus driven by the high numbers of negatives in Nguruman. Since subpopulation was the main factor relating to the presence of *Sodalis* in tsetse flies (Table 3.9), it is difficult to tell whether differences in prevalence are due to the community of tsetse species present or environmental factors that reduce both endosymbiont and parasite invasion.

In addition, different species of parasites might have different relationships with the endosymbiont. There was evidence from the GLMs that the presence of *T.*

*congolense* and *T. brucei* but not *T. vivax* were influenced by *Sodalis* status in tsetse flies but only in complex interactions with tsetse-specific factors or geographic location and the effect was strongest in *T. congolense* (Table 3.9). The role of *Sodalis* thus might differ by species of trypanosomes but not considering interactions with tsetse factors or sampling site might lead to misleading conclusions. An experiment using a laboratory strain of *G. pallidipes* (from the Trypanosomiasis Research Centre, TRC), found that the prevalence of *Sodalis* in tsetse flies that were trypanosome positive after experimental inoculations varied by species of trypanosome: 38.9% for *T. b. rhodesiensei;* 36.4% for *T. b. brucei*; and 32.7% for *T. congolense* (Wamwiri *et al.*, 2014). It is also possible that there are specific interactions between strains of *S. glossinidius* and species of tsetse flies that could favour *Trypanosoma spp.* infections, as suggested in a previous publication (Farikou *et al.*, 2011b). They found genetic variation among *S. glossinidius*, which was also suggested in my study for the GPO1 gene (Figure 3.9). For all trypanosome species, my study suggested some apparent trends in *Sodalis* status with age or sex but the GLMs indicated that it is not appropriate to consider these factors in isolation. The multivariate analyses also indicated that *Sodalis* status is resolved on a different dimension than trypanosome status and is associated with variation among subpopulations and not tsetse-specific factors (Figure 3.2). Therefore, interpretation of the influence of *Sodalis* may be confounded by multiple other factors that affect trypanosome prevalence in tsetse flies.

There is also a possibility that particular types of blood meals influence the apparent relationship between *Sodalis* and trypanosome status. For the blood meal analysis, I focused only on *G. pallidipes* to enable comparision across populations but without the confounding effects of tsetse species. Although there was no association between *Sodalis* and trypanosome status, the MCA demonstrated a strong correlation between *Sodalis* status and host-feeding pattern (Figure 4.8). However, this was again confounded by geographic region, since there was also a strong correlation between these variables and sample site: tsetse flies from Nguruman were associated with single host feeding and absence of *Sodalis*. It would be intriguing to speculate that this could be due to differences in the quality of blood: perhaps not enough nutrition is obtained from single hosts to enable proliferation of the endoparasite. Moreover,

different blood parameters have been reported in antelopes, buffalo and giraffes from Kenya (Drevemo *et al.*, 1974), which could affect the presence of *Sodalis* in tsetse flies in geographic regions that differ in host availability. Teasing out factors that directly affect *Sodalis* prevalence in the MCA might allow more accurate interpretation of relationships with trypanosome prevalence in different geographic regions or species of tsetse.

## 5.4 Association of host feeding of *G. pallidipes* with trypanosome status

Tsetse flies are cyclical insect vectors for trypanosomes in Africa. Host and feeding patterns of tsetse flies might reflect trypanosome distribution and so might be informative about the relative risk of wildlife to be reservoirs of trypanosomiasis in humans and domesticated animals. However, hosts of tsetse flies vary by species of flies, host availability and geographic location (Clausen *et al.*, 1998) and so simple relationships are again not expected. In my study, African buffalo were found to be the most common wildlife source of blood meals for the *G. pallidipes* screened, followed by African elephants, antelopes, warthogs, giraffes and hyenas (Figure 4.3). Although there was some feeding on humans and domesticated animals (primarily goats and chickens) this was at much lower proportions than for the wildlife and domesticated animals tended only to be involved in mixed feeding (Table 4.5). The proportion of single and multiple-feeding flies of each site also were different. The percentage of tsetse flies from Mukinyo that fed on single hosts was higher than multiple hosts (Figure 4.6Figure 4.6) while tsetse flies from Zungu Luka fed on multiple hosts more than single hosts (Figure 4.5). Proportions of single and multiple-host feeding flies from Buffalo Ridge were similar to one another but approximately 40% of the sampled flies had fed but the source of blood meals could not be confirmed (Figure 4.4). Results of host determination in this investigation were consistent with previous studies in Kenya finding that ruminants were the main hosts for *G. pallidipes* from Nguruman and Nkineji (Bett *et al.*, 2008), and Busia (Okoth *et al.*, 2007, Nyawira, 2009, Muturi *et al.*, 2011). African elephants, which were the second predominant hosts for *G. pallidipes* in my study, were occasionally major hosts in Nguruman (Nyawira, 2009, Muturi *et al.*, 2011). Bushpigs and warthogs

have been reported as predominant hosts for *G. pallidipes* from the Shimba Hills, Mwalewa, Matuga, Diani and Muhaka (Snow *et al.*, 1988) and Nguruman (Okoth *et al.*, 2007) regions. Altitude was found to be a significant risk factor for trypanosomiasis whilst host migration also influenced risks for animal trypanosomiasis using logistic regression analysis (Majekodunmi *et al.*, 2013). Thus, host feeding could relate to abundance of available animals in each region and animal migration in different seasons.

Alhtough the host species could be identified only for a subset of the trypanosome positive *G. pallidipes* screened, *T. vivax* was the predominant pathogenic species found.Thus, while my results suggest that which hosts are fed on could influence trypanosome prevalence, this could be biased by the dominance of *T. vivax* in the geographic regions surveyed. Nevertheless, prevalence of trypanosome in tsetse flies that fed on African buffalo was proportionately lower than for African elephants. I also found a relatively high proportion of mixed species of trypanosomes in tsetse flies that fed on African buffalo, African elephants or domestic animals. Animals that have low self-protection mechanisms and settle around the habitat of tsetse flies could possibly be fed on more than others. My analyses likely under-estimated mixed feeding because I took a conservative approach; poor resolution of sequences was interpreted as multiple feeding only when clear heterozygous peaks were resolved but amplification of multiple species could also reduce overall resolution of sequences. There is also a possibility that not all host species amplified equally well using the primer set used; for example, some sequences that were not included in the final analyses showed similarity to goats in BLAST analyses but the sequence quality was not sufficient to resolve whether they were from single or multiple-host blood meals. Nevertheless, my results indicate that wildlife can support multiple species of trypanosomes that can be transmitted to tsetse fly. Risks associated with wild animal movements in areas where animals are domesticated should thus be studied in more detail, including comparison of whether the same strains of the parasites are shared among vectors, wild animals and livestock.

## 5.5 Limitations

There are four primary issues that limited my study: 1) differences in communities of tsetse flies sampled from different geographic regions; 2) sample size and preservation of samples; 3) PCR methods; and 4) available data on habitat charateristics and density of possible host species for discussion.

Firstly, tsetse flies sampled showed variation in species, sex and age across sites (Appendix B.3), making it difficult to interpret the relative associations of trypanosomes with intrinsic factors. I would suggest that samples should be trapped every month through the years or at least from representative seasons. The numbers of traps would depend on the target number of trapped flies/day to gain the required appropriate number suggested by power analyses. Temperature and relative humidity should also be measured besides each trap to allow an assessment of whether it is microclimate or other features of the habitat that drive any differences observed among sites or times of year.

Secondly, numbers of samples were very small in some populations and for some combinations of factors (Appendix B.3). Although associations involving overall species of trypanosome (N = 1090) and *T. vivax* on its own were analysed using a relatively large sample size (N = 281), there were some combinations of variables that did not occur sufficiently frequently enough to have much power to test significance. For example, old flies that were both *Sodalis* and *T. brucei* positive were rare, especially among samples from Nguruman. While this is potentially informative about differences in composition of vector–symbiont-parasite community in different regions, more targeted sampling would need to be used to specifically test associations when sample sizes are this small. In order to increase the potential of the models to predict associations, more targeted sampling of particular communities might be necessary. All collected samples were preserved in 99% ethanol to protect degradation of DNA by driving out water from the tissue, cells and blood contents in the gastrointestinal tract of tsetse flies. Thus, age estimation of females had to be based on the wing fray score, rather than a more commonly used estimate based on ovary scores (Saunders, 1962). The ethanol preservation also prevented classification of flies as having recently fed or not so this had to be inferred only based on PCR. In addition, the dark brown colour solutions in the lysis step with proteinase K

during DNA extraction could easily be confused between blood meals and tsetse colour pigment. Thus, fresh samples are possibly more practicable for determination of feeding status than preserved samples. I would suggest that the feeding status (fed and unfed) and wing fray score of trapped flies should be determined every day by dissection using microscopy. Blood meals from the dissection could then be collected using whatman filter paper.

Thirdly, biological identification based on PCR methods is unable to differentiate between infection and exposure status. For example, flies that have recently taken up parasites that do not subsequently establish or transmit should not be considered "infected" but PCR might still detect low-level presence of the parasites. In my study I attempted to reduce the risk of this problem by analysing head plus proboscis parts instead of whole flies or abdomen parts (see chapter 2). Although this still does not allow distinction of whether parasites have established in the flies, it does include parasites that could be mechanically transmitted rather than just those taken up in the blood meals but would have no chance of transmission. To test for false positives, head plus proboscis parts could be screened using both the general ITS-1 prrimers for trypanosome diagnosis and the cytb primers for vertebrates, to distinguish possibly parasites taken up by recent feeding. For my blood meal analysis, direct sequencing of PCR products was only informative for single host feeding. PCR amplification that provided low DNA yield for sequence analysis and direct sequences that corresponded to two or more hosts (with obvious double peaks) could not be used reliably for host identification. Cloning and plasmid sequencing is suggested to identify the predominant hosts of tsetse flies that feed on multiple hosts. In addition, the maximum periods over which it is possible to track blood meal sources needs to be further investigated to distinguish whether there are remnants of old feeding or whether only the most recent meals can be identified.

Finally, density of wildlife abundance and livestock in the study areas were not available to access, thus limiting an informed discussion of the correlation between the available animals with trypanosome prevalence and feeding patterns or predicting whether relative proportion of hosts found reflected abundance available or tsetse host preferences. Abundant wild animals and livestock should be estimated from direct observation within 10 km (the flight

distance/day that has been predicted for individual flies (Vale *et al.*, 2014)) around the local areas in the period of sample collection.

## 5.6 Broader implications of this work

Epidemiological studies of trypanosomiasis require trypanosome investigation in vectors or hosts, as well as planning and implementing control (Desta, 2014). Use of insecticides for the control of insect vectors was reviewed and it was concluded that low dosage aerosols from fixed-wing aircraft have the greatest potential for rapid and effective control over large areas, with minimal risk of long-term environmental contamination (Allsopp, 1984). An integrated insecticidal approach combining the attributes of various chemicals and possibly non-chemical control methods also have been used for tsetse flies control; for example, ground spraying, aerosol spraying, odour-baited traps, insecticide-treated targets application with animals and sterile male techniques (Kabayo, 2002). However, refinements to improve efficiency, consider environmental effects and develop sustainable control methods are needed. My study found that there are multiple levels of interactions that could affect the dynamics of trypanosome transmission between tsetse flies and animal hosts, suggesting that control measures should be developed to specifically target particular communities of vectors, parasites and hosts at small geographical scales.

From my investigation, age and an interaction between subpopulation and sex were significantly associated with status of *Trypanosoma spp*. but different species of tsetse flies showed different factors significantly affecting trypanosomes infections (Table 3.10). When intrinsic factors and *Sodalis* status were analysed, complicated associations of each trypanosome species were detected. These types of interactions could be a reason for the difficulty of trypanosome control and eradication. Different associations were detected in particular sites of tsetse fly collection and for different species of tsetse flies, which further suggests that strategies for disease control should be developed under specific localised conditions. The complexity of tsetse behaviours in various habitats and interactions with available hosts are such that the creation and maintenance of fly-free zones is a complex task (Shaw *et al.*, 2013). I

recommend that control of *Glossina pallidipes*, which was the most common species found in my study, should be based both on environmental manipulation and by strategies to prevent it from feeding on its main hosts. My blood meal analysis showed that approximately 75% of tsetse flies showed recent feeding activity; thus, trap setting near stalls and treating livestock with insecticides on legs and belly (Torr *et al.*, 2007) would be practical and beneficial for control of tsetse, but also of other blood-sucking insects and ticks. Using repellents is not always successful for biting protection (because it might just shift feeding to another nearby host) and it sometimes causes multiple feeding (Leak, 1998), which is supported by my results suggesting that blood meals involving humans often also included other hosts. Other broad and more sustainable measures should be applied in conjunction with this low-technology approach for open areas. For example, using a combination between Nzi traps baited with acetone, urine or ocetone (Mihok *et al.*, 2007) and SIT might benefit local *G. pallidipes* control. The baits might also be adapted to specifically target the particular species of tsetse flies in each region (as described in section 1.6.4) for better efficiency and improving cost-benefit ratios. If there is an influence of *Sodalis* on prevalence of particular strains of trypanosomes, antibiotic treatment could help to reduce infection levels (e.g. of *T. congolense* and *T. brucei*) in tsetse flies in high-density *Sodalis* areas. In the future, SIT could also be combined with recombinant *Sodalis* approaches. For example, mechanisms are being developed to use recombinant *Sodalis* for activation of the tsetse immune system to block trypanosome development in tsetse hosts (Aksoy, 2003) or to express antibodies against the variant surface glycoprotein (VSG) of trypanosomes (De Vooght *et al.*, 2014, De Vooght *et al.*, 2012).

The blood meal analysis results are useful for indicating vertebrate hosts that could be at risk of trypanosome infection via tsetse fly biting in particular regions. In order to control trypanosomiasis in Buffalo Ridge and Zungu Luka, many species of domestic animals should be screened because tsetse flies in these regions fed on a wide range of domestic and wild animals (including chickens, mouse, goats and other ruminants). For Mukinyo, trypanosome control should focus more on wild animals, especially African buffalo and African elephants, since no flies were found to feed on livestock. The relative risks of

exposure to humans and domesticated animals thus should be considered in the design and implementation of control measures.

## 5.7 Future work

My work has revealed many uncontrolled factors that could influence trypanosome infection in tsetse flies under natural conditions. In chapters 2 and 3, site and subpopulation significantly correlated with trypanosome infection but it is not clear whether this is due to the influence of environmental conditions, refractoriness of local colonies of tsetse flies, or virulence of local strains of trypanosomes. Moreover, there are some conditions that cannot be controlled in field studies; for instance, numbers of tsetse flies of the different species, sex and age, food conditions, trypanosome density and *Sodalis* density. Therefore, more laboratory experiments are required to explicitly test factors that influence infection by trypanosomes in tsetse flies, but using wild strains of tsetse flies and their geographically relevant trypanosomes rather than laboratory strains, as has been done in most previous studies. Trypanosomes sampled from hosts from the same geographic region as the target tsetse flies could be cultured using laboratory mice for use in these experiments. Power analysis could be used to establish the number of tsetse flies that would be required for robust tests of all hypotheses, prior to sampling. After establishing cultures of trypanosomes and collecting samples of tsetse from the wild, there could be two phases of the experiment. The first would be to characterise growth rates, longevity, mortality rate and sex biases using tsetse-fly cultures without trypanososome infection under laboratory conditions, initiated using the field-collected flies. Wing-fray scores of both males and females (using flies anaesthetised using carbon dioxide (Nilson *et al.*, 2006) could then be directly related to age of flies, to determine the most reliable classification of flies into age groups. The experiment also would test the parameters separately for each species of tsetse-fly to consider the relationship between actual age and wing fray scores. The second phase would be challenge experiments conducted using flies of the different species raised in the laboratory that are separated into age and sex groups. The number of tsetse samples would be determined from power analysis and mortality rate of tsetse flies in the first phase. Flies would be

exposed to trypanosomes by feeding them on the mice infected to culture the local strains and then would be fed on uninfected blood sources for three weeks. They would then be starved for a further week to be sure that there would be no blood content in the abdomen parts. They would then be screened for trypanosome infection; one month is the duration that all species of trypanosome completely develop to be an infective stage (Aksoy *et al.*, 2001) and so should assess differences in transmittable trypanosome numbers rather than just differences in uptake from the blood. Screening results of head plus proboscis parts and abdomen parts of tsetse flies would be compared to test the hypothesis that trypanosome screening from abdomen parts should have a lower sensitivity than that from head plus proboscis part due to PCR inhibitors in blood content. The data would then be analysed using GLMs to test for associations of each field trypanosome strain with species of the fly, age and sex. The same approach could be used to test specific hypotheses. For example, as discussed in chapter 2, mixed infections of trypanosome species were a suspected cause of high mortality of infected tsetse flies; thus, it would be interesting to inoculate various species of trypanosomes into male and female tsetse flies of different age classes to test this assumption.

Similarly, in order to test specific associations related to *Sodalis*, experiments could be conducted to compare infection levels using different species of flies and different species trypanosomes, with and without *Sodalis*. *Sodalis glossinidius* isolated from local tsetse flies could be cultured and separated into two stocks: one used for an antibiotic sensitivity test to choose a drug that will ensure all of the bacteria are killed and one used to inoculate flies. The flies could first be treated with the selected antibiotic, to ensure that they are *Sodalis* free prior to the experimental infection. Experimentally infected and noninfected flies could then be compared for infectivity of different species of trypanosomes. Experimental infection of local strains of trypanosome species and *S. glossinidius* in local colonies of tsetse flies under environmental control should be tested to compare infection rates of each tsetse fly group.

For identification of feeding patterns, it would be informative to experimentally test the maximum time of cytb detection from the abdomen part of tsetse flies with different volumes of blood meals, in relation to particular activities of tsetse flies (e.g. feeding on the same or different species of hosts in a short time

frame; relative activity levels) and the surrounding environmental conditions (e.g. temperature and relative humidity). These results would enhance interpretation of blood meal sources from wild flies, as described in chapter 4. Flies would be raised under standard laboratory conditions (e.g. 25 °C, 60% relative humidity, 12 hr light - 12 hr dark), which could also be altered to compare the effects environmental conditions on blood meal detection.Anaesthetised flies would be weighed before and after blood meals to determine the volume taken up; flies would then be randomly selected for destructive sampled at varying times after feeding to allow screening of DNA extracted from the abdomens using mtDNA primers. The experiment would require housing of individual flies (to allow relating decay in blood meal detection to the volume taken up) but could also consider differences in sex, age, relative activity levels of the flies after feeding, or differences in detection between blood meals with multiple feeding on the same or different hosts within a short time frame. Differences in emptying times of the foregut between species of flies or species of trypanosomes could also be explored. However, each factor considered would increase the scale of the experiment, since each fly could only be tested once, so large numbers of individual flies would have to be reared to allow a complete assessment of time. Nevertheless, my results clearly indicate that more details are required about the sensitivity of PCR-based detection that is informed by the species, sex and age of tsetse flies rather than generalising across these factors.

# Appendices

## Appendix A Supplementary Information for chapter 2

### A.1. Drawings of vegetation types, taken from Leak (1998).


**Woodland**, *Brachystegia, Terminalia* spp.

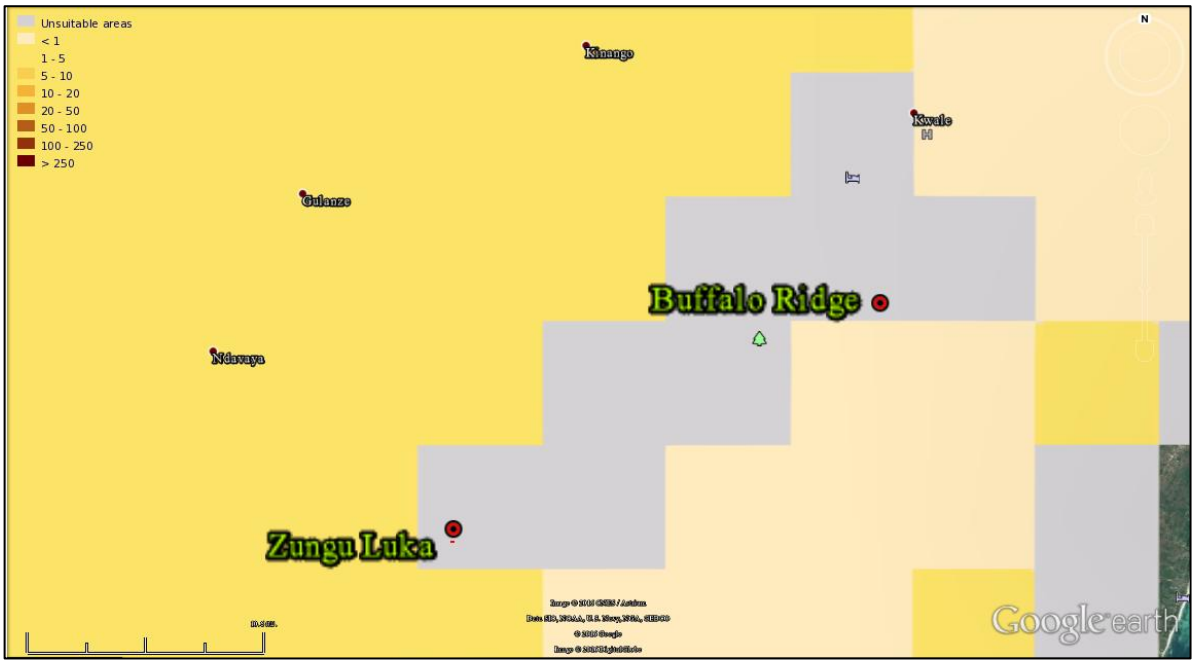

**Thicket**, *Commiphora, Combretum, Acacia, Teclea, Maba.*

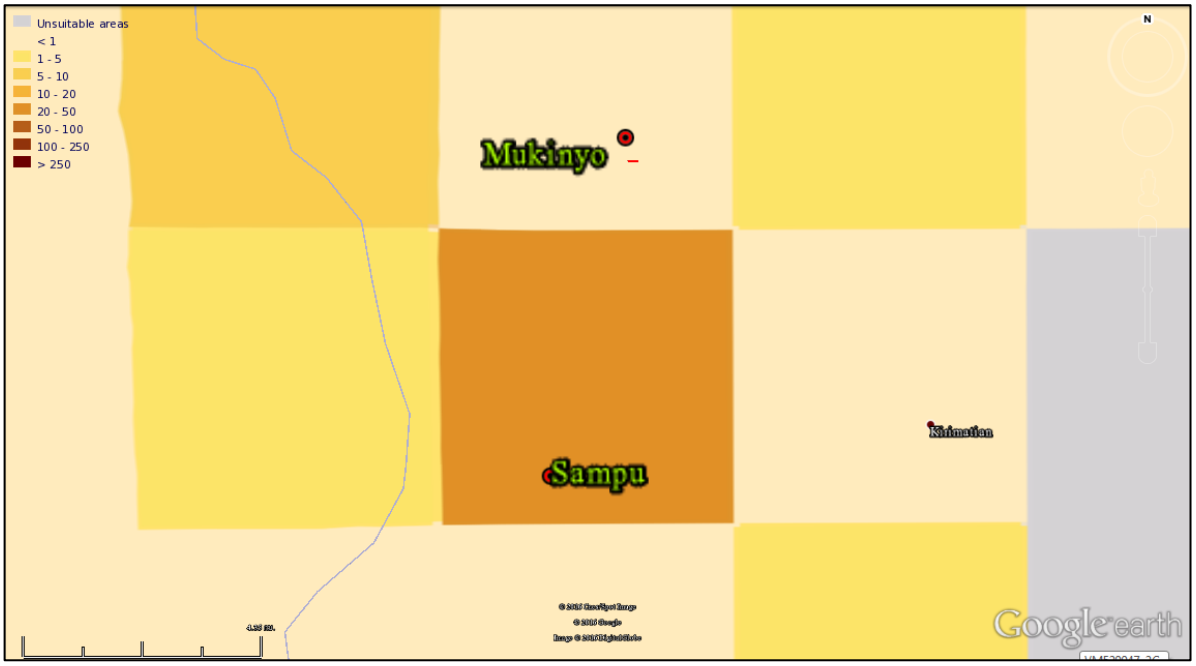

**Savanna**, *Combretum, Acacia, Borassus.*


**Acacia Savanna**, *Acacia* spp.

## A.2. Estimated densities of cattle around Buffalo Ridge, Zungu Luka, Mukinyo and Sampu

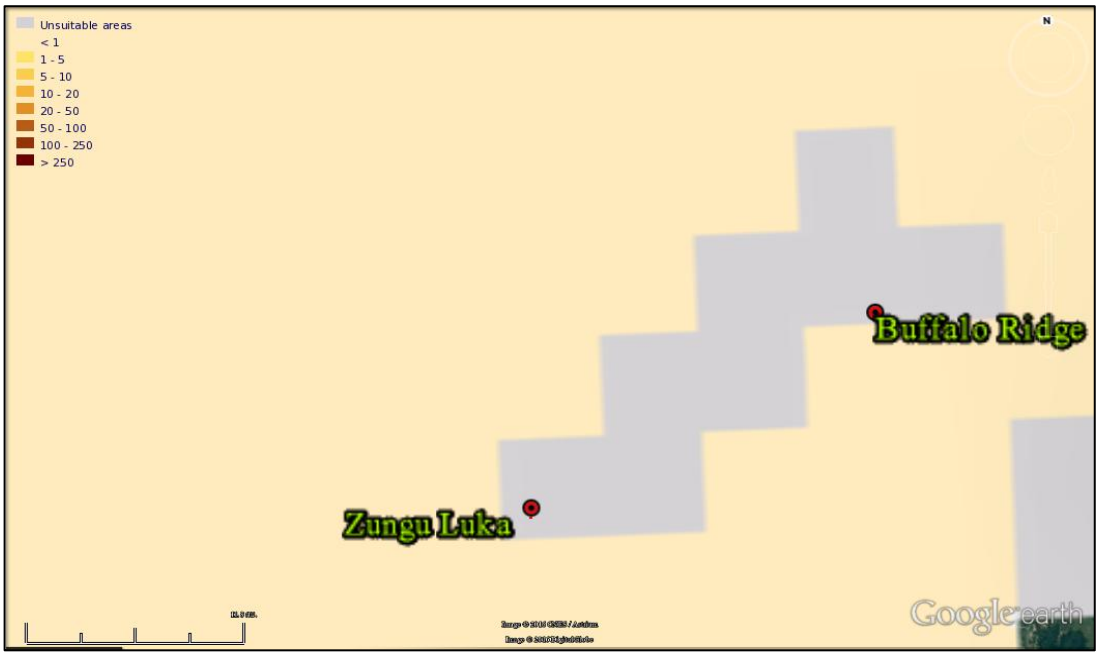
Predicted density of animals in 2005, mapped to Google earth based on the Food Agriculture Organisation web-based database of global cattle density (http://www.fao.org/home/en/).

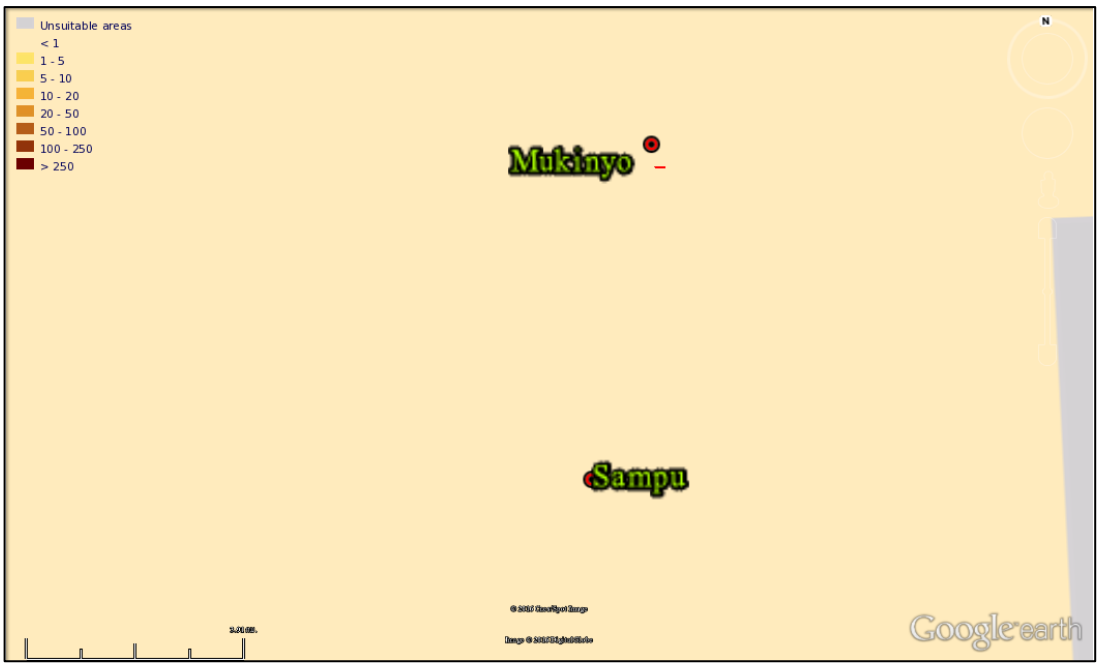


## A.3. Estimated densities of buffalo around Buffalo Ridge, Zungu Luka, Mukinyo and Sampu

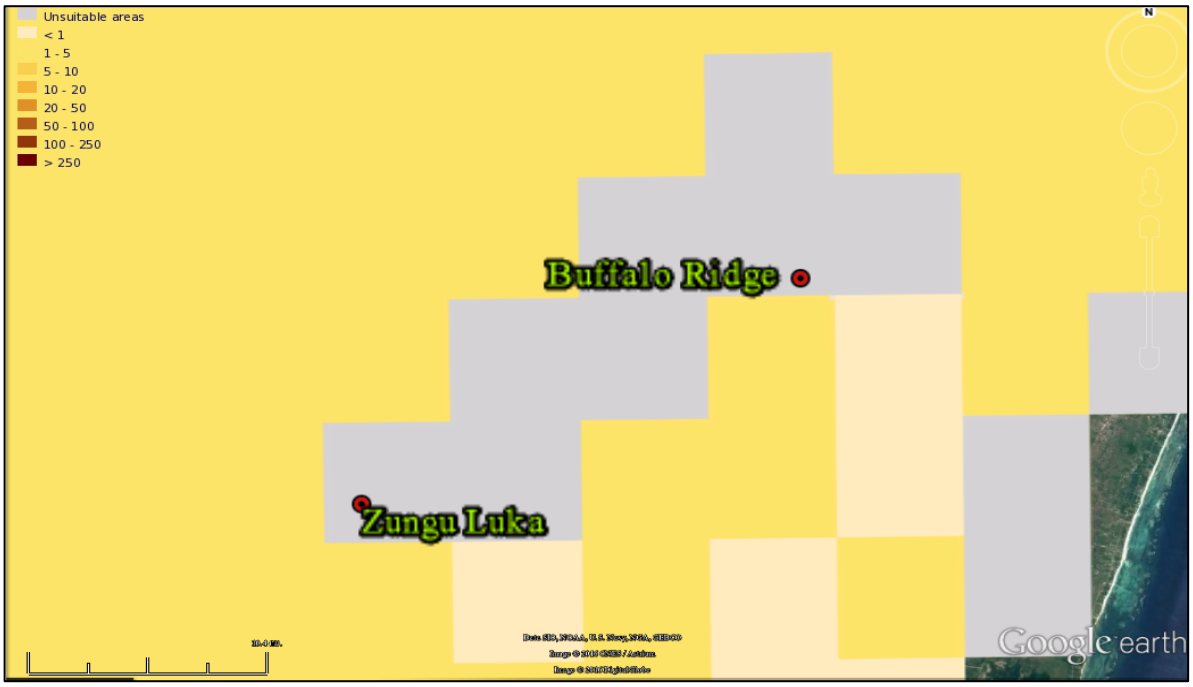
Predicted density of animals in 2005, mapped to Google earth based on the Food Agriculture Organisation web-based database of global cattle density (http://www.fao.org/home/en/).

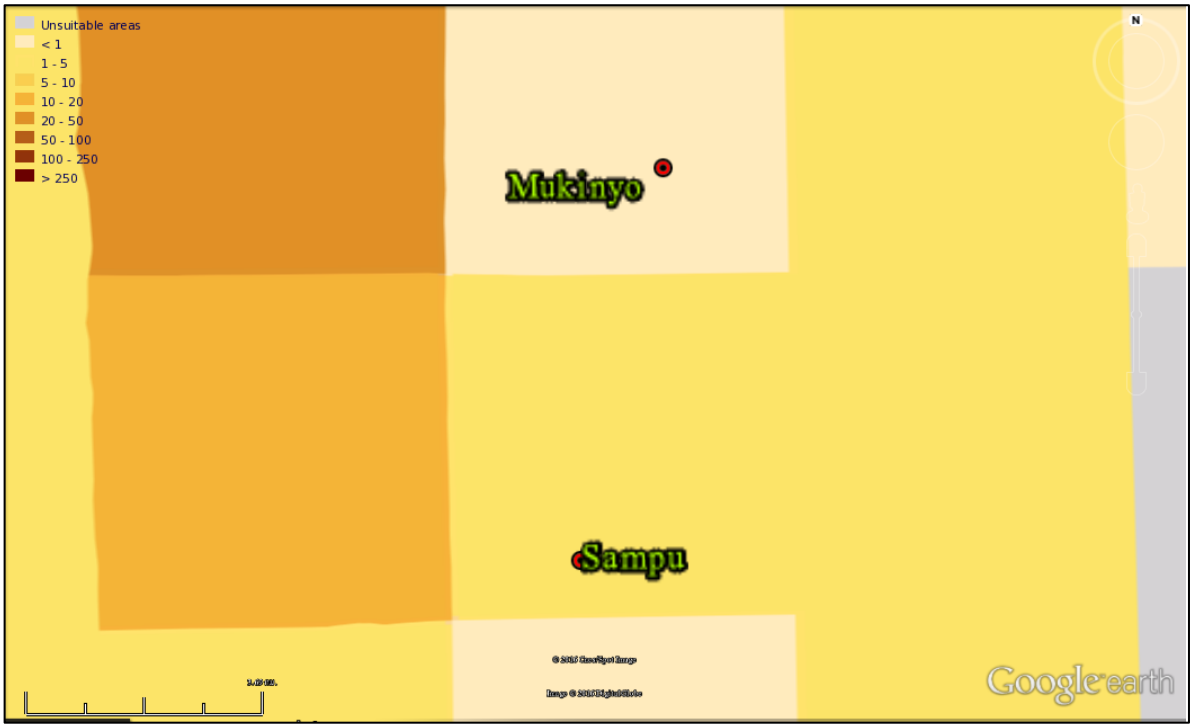


## A.4. Estimated densities of goats around Buffalo Ridge, Zungu Luka, Mukinyo and Sampu

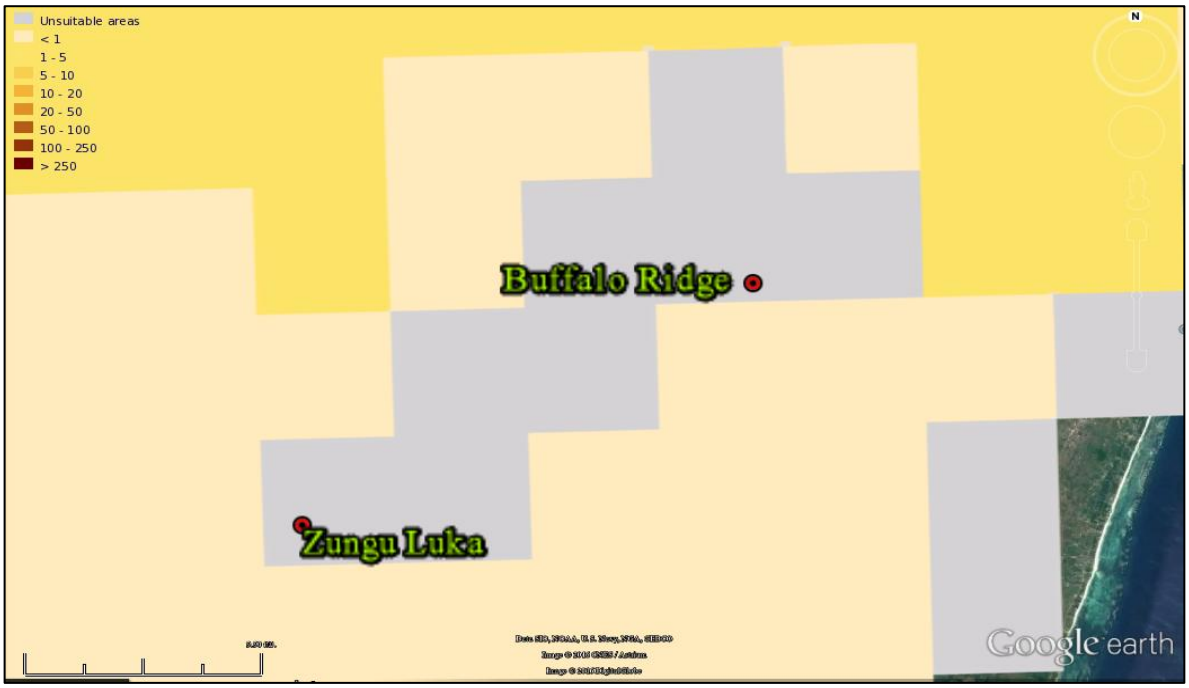
Predicted density of animals in 2005, mapped to Google earth based on the Food Agriculture Organisation web-based database of global cattle density (http://www.fao.org/home/en/).

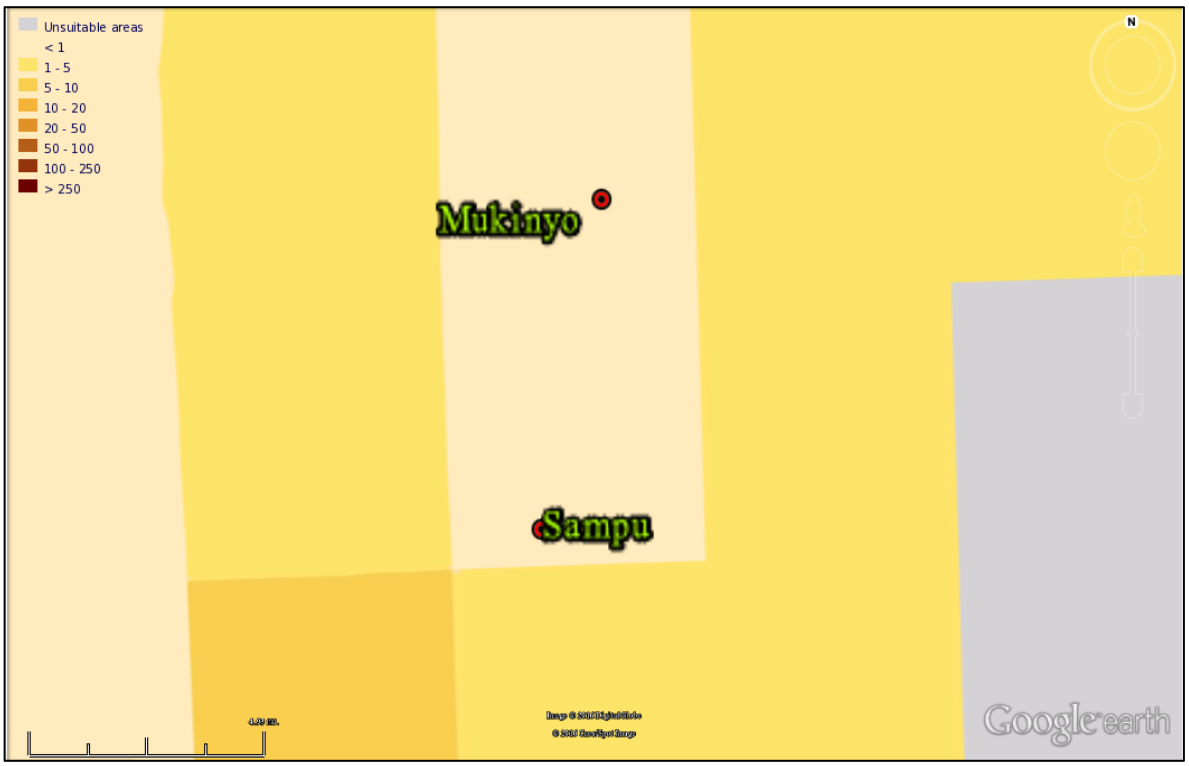


## A.5. Estimated densities of sheep around Buffalo Ridge, Zungu Luka, Mukinyo and Sampu

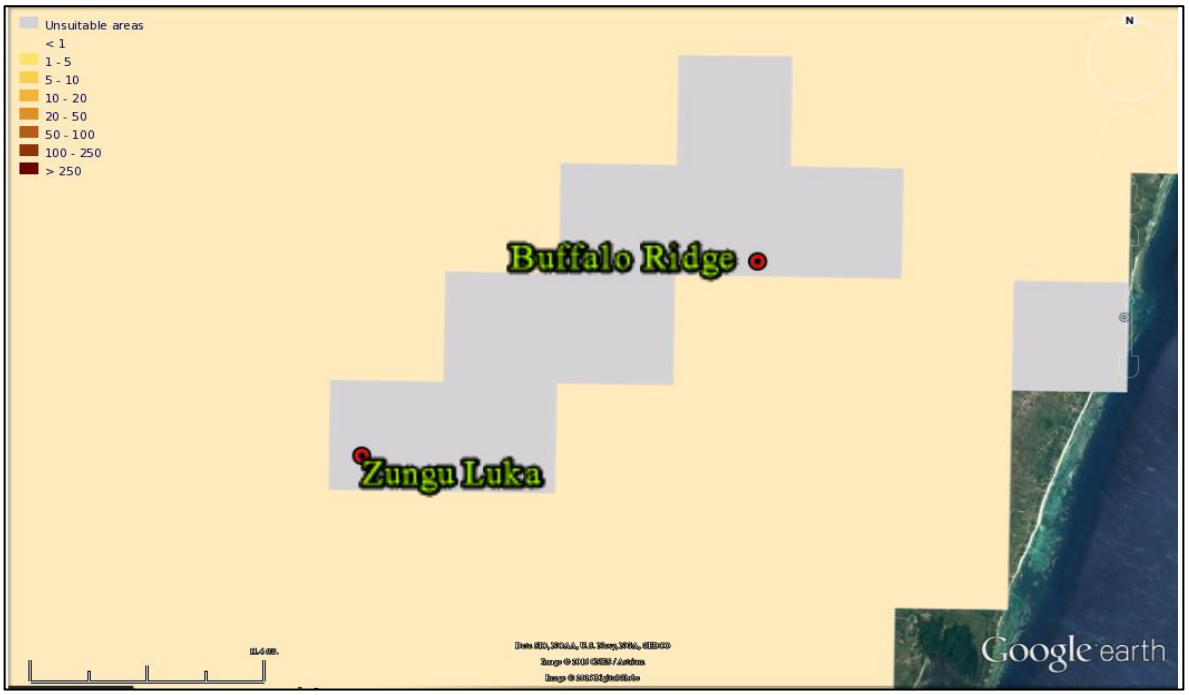
Predicted density of animals in 2005, mapped to Google earth based on the Food Agriculture Organisation web-based database of global cattle density (http://www.fao.org/home/en/).

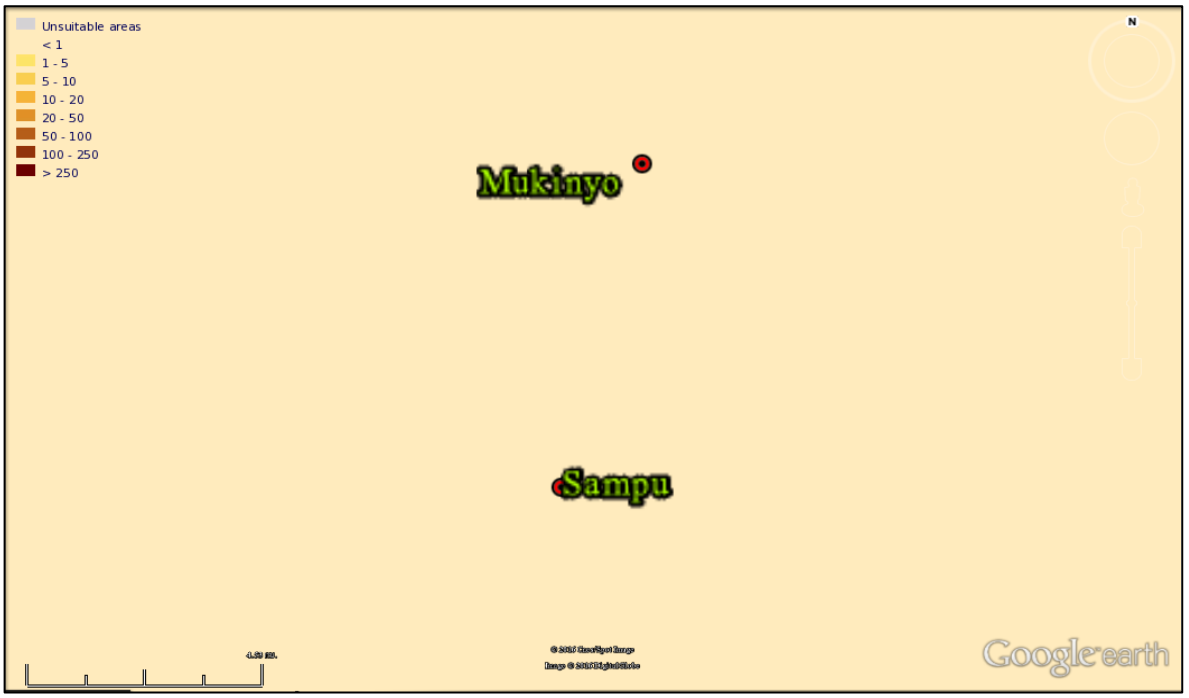


## A.6. Estimated densities of pigs around Buffalo Ridge, Zungu Luka, Mukinyo and Sampu

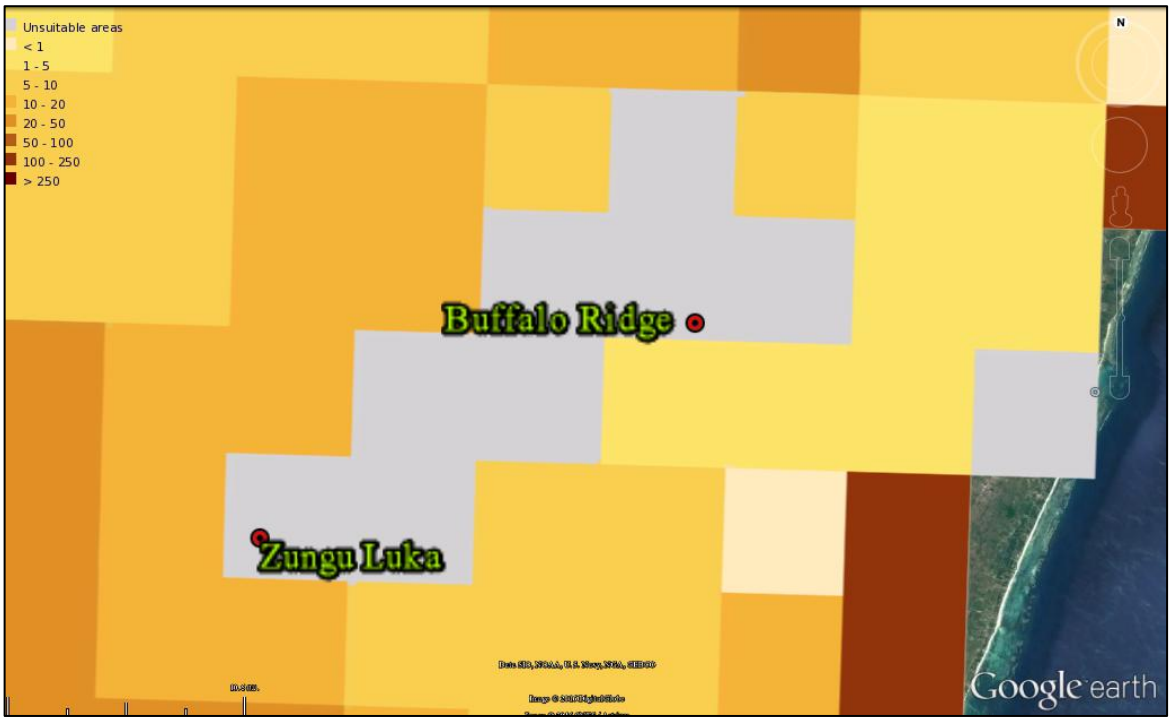
Predicted density of animals in 2005, mapped to Google earth based on the Food Agriculture Organisation web-based database of global cattle density (http://www.fao.org/home/en/).

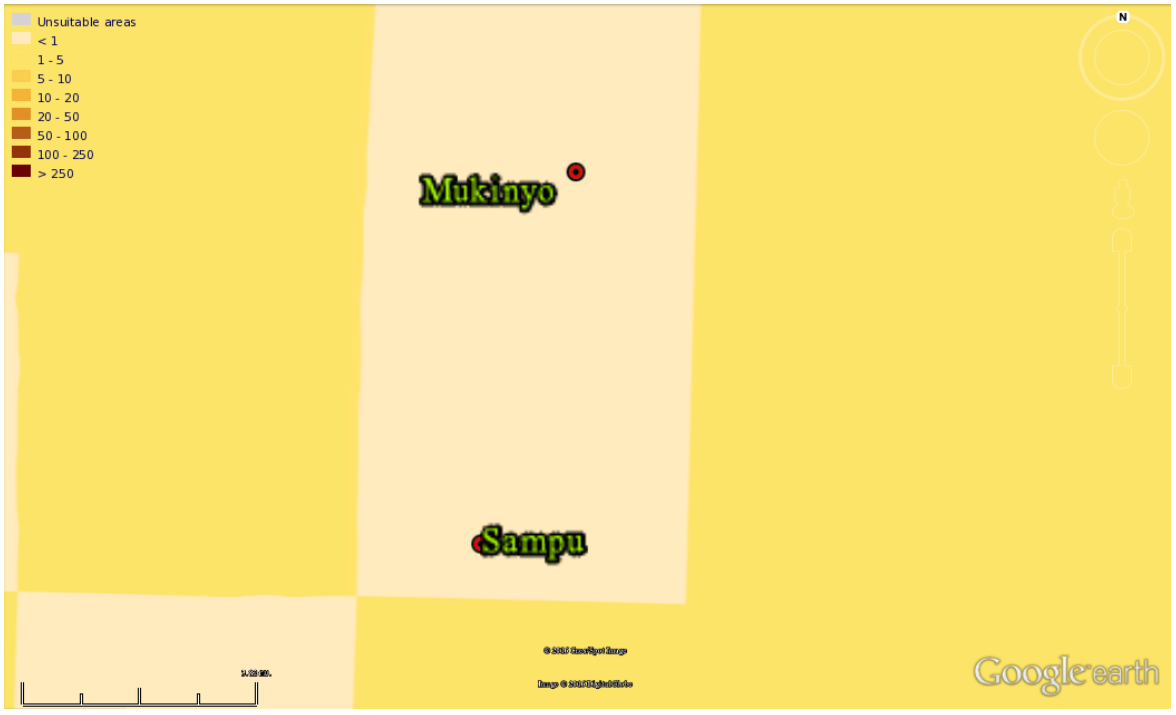


## A.7. Estimated densities of poultry around Buffalo Ridge, Zungu Luka, Mukinyo and Sampu

Predicted density of animals in 2005, mapped to Google earth based on the Food Agriculture Organisation web-based database of global cattle density (http://www.fao.org/home/en/).



## A.8. Experimental infection of *T. congolense* savannah in *G. pallidipes*.

This study was conducted by Marc Ciosi, with PCR screening conducted by Manun Wongserepipatana and Alana Hamilton.

*Trypanosoma congolense* savannah strain KETRI2885 (IL1180) was inoculated into *G. pallidipes*. Then, head plus proboscis and abdomen of 10 tsetse flies were dissected for DNA extraction at each time point from samples collected at days 0 (6 h after the inoculation), 2, 5 and 8. On day 11 of this experiment, there were only 7 tsetse flies left so all were sampled. Screening results from general ITS-1 and subspecies-specific TCS primers were analysed. The graph of trypanosome identification (Figure A.2) suggests that it was better to use mouth parts and ITS-1 primers rather than abdomen parts and TCS primers for trypanosome screening because the higher sensitivity risked confusing established infections with transient parasites remaining from the experimental exposure.

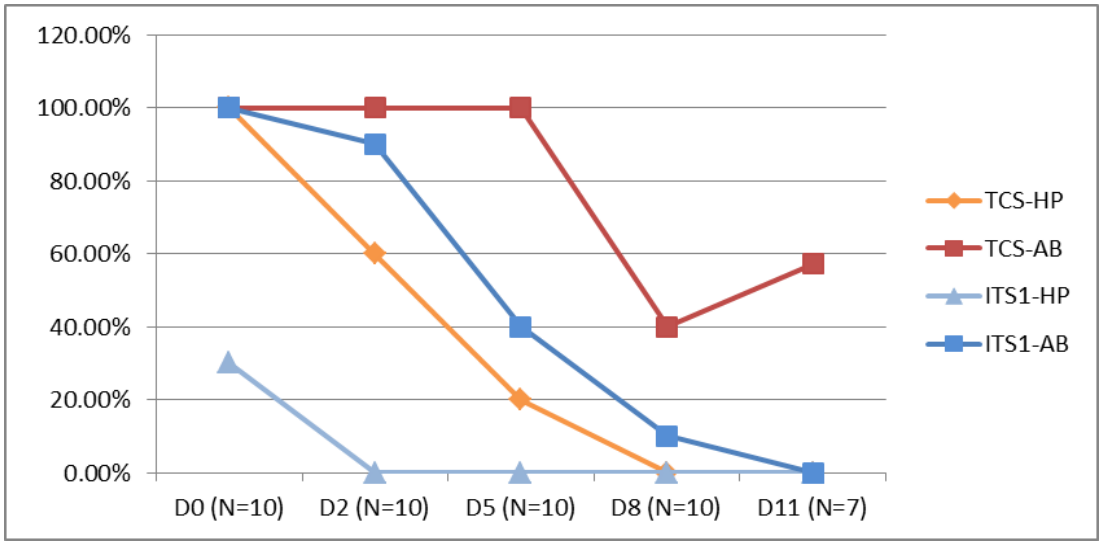



**Figure A.2** Percentage of *T. congolense* savannah in *G. pallidipes* using ITS-1 specific primers.

D0 is the day of trypanosome inoculation. D2, D5, D8 and D11 are timepoints after the inoculation. Numbers of tsetse flies are represented below the graph. (Ciosi, unpublished)

## A.9. Percentage amplification of PCR products corresponding to five species of trypanosomes based on ITS-1 screening.

Head plus proboscis of *G. austeni*, *G. brevipalpis*, *G. pallidipes* and *G. longipennis* from Zungu Luka, Buffalo Rodge, Mukinyo and Sampu were screened for *T. congolense* savannah (Tcs), *T. congolense* kilifi (Tck), *T. brucei* (Tb) *T. vivax* (Tv), *T. simiae* (Ts) and *T. godfreyi* (*Tg*).

| *Trypanosoma spp.* | Prevalence of trypanosome infection in *Glossina spp.* | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *G. austeni* | | | *G. brevipalpis* | | | *G. longipennis* | | | | | | *G. pallidipes* | | | | | | | |
| | Zungu Luka | | | Buffalo Ridge | | | Mukinyo | | | Sampu | | | Buffalo Ridge | | | Zungu Luka | | | Mukinyo | |
| | Male | Female | Total (N=282) | Male | Female | Total (N=141) | Male | Female | Total (N=32) | Male | Female | Total (N=58) | Male | Female | Total (N=154) | Male | Female | Total (N=130) | Male | Female | Total (N=293) |
| **Tcs** | 1.06 | 10.64 | 11.70 | 3.55 | 12.05 | 15.60 | 0.35 | 6.25 | 9.38 | 5.17 | 0.00 | 5.17 | 7.79 | 7.79 | 15.58 | 0.77 | 7.69 | 8.46 | 4.50 | 5.73 | 10.23 |
| **Tck** | 1.42 | 9.57 | 10.99 | 1.42 | 0.71 | 2.13 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.65 | 1.30 | 1.95 | 0.00 | 2.31 | 2.31 | 0.35 | 1.02 | 1.37 |
| **Tb** | 1.42 | 5.67 | 7.09 | 5.67 | 4.26 | 9.93 | 0.35 | 0.00 | 0.35 | 8.62 | 0.00 | 8.62 | 1.30 | 1.30 | 2.60 | 2.31 | 6.15 | 8.46 | 2.48 | 4.79 | 7.17 |
| **Tv** | 3.90 | 31.21 | 35.11 | 3.55 | 7.80 | 11.35 | 3.13 | 18.75 | 21.88 | 17.24 | 1.72 | 8.62 | 12.34 | 8.44 | 20.78 | 13.85 | 32.30 | 46.15 | 7.17 | 13.99 | 21.16 |
| **Ts** | 1.06 | 4.26 | 5.32 | 2.13 | 1.42 | 3.55 | 0.00 | 0.00 | 0.00 | 0.00 | 1.72 | 1.72 | 0.65 | 0.65 | 1.30 | 0.00 | 0.77 | 0.77 | 0.68 | 1.37 | 2.05 |
| **Tg** | 0.00 | 1.06 | 1.06 | 0.71 | 0.71 | 1.42 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2.31 | 2.31 | 2.05 | 1.02 | 3.07 |
| **Total** | 6.03 | 47.52 | 53.55 | 12.05 | 22.70 | 34.75 | 1.06 | 25.00 | 34.38 | 17.24 | 1.72 | 18.97 | 18.18 | 16.24 | 34.42 | 16.15 | 43.85 | 60.00 | 13.31 | 22.87 | 36.18 |
| | 53.55 | | | 34.75 | | | 24.44 | | | | | | 41.07 | | | | | | | | |
| **Mixed species.** | 2.13 | 13.12 | 15.25 | 2.84 | 3.55 | 6.38 | 0.00 | 0.00 | 0.00 | 3.45 | 1.72 | 5.17 | 4.55 | 2.60 | 7.14 | 0.35 | 3.19 | 3.55 | 3.07 | 4.10 | 7.17 |
| | 15.25 | | | 6.38 | | | 3.33 | | | | | | 7.28 | | | | | | | | |

## A.10. The best fitting GLM of *T. congolense* status of all tsetse fly samples (Model 1.1).

### Backward elimination of model selection

Explanatory variables: subpopulation, sex and age
Response variables: *T. congolense* status in all tsetse species (N = 1090)

| Effect tested | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation * sex * age (Full model) | 28 | 424.8 | 905.6 | - | - |
| subpopulation * sex * age | 22 | 428.5 | 901.1 | 7.4977 | 0.2773 |
| sex * age | 21 | 428.7 | 899.3 | 0.2879 | 0.5915 |
| subpopulation:age | 15 | 431.5 | 893.0 | 5.6295 | 0.4659 |
| subpopulation * sex | 9 | 435.5 | 889.0 | 7.9623 | 0.2409 |
| age | 8 | 436.2 | 888.4 | 1.4336 | 0.2312 |
| sex | 7 | 436.4 | 886.7 | 0.3483 | 0.5551 |
| subpopulation | 1 | 893.5 | 445.5 | 18.7728 | **0.0046** |

A * B means an interaction between A and B.

### Summary of the best fitting model

Subpopulation was a significant predictor for *T. congolense* infection when analysed from 1090 tsetse flies using the binomial family of GLM analysis. Residual deviance from the best fitting model is indicated: minimum; the first quadrant; median (the second quadrant); the third quadrant; and maximum. Coefficients of factor levels are indicated with the estimate (which reflects the direction and magnitude of association compared to the reference level; in this case subpopulation BRGb), along with standard error (Std Error), Z value and Pr(>|Z|). Deviance of the model was 891.49 on 1089 degrees of freedom. Including the subpopulation variable decreased the deviance 18.77 points and 6 degree of freedom, a significant reduction in deviance. The Akaike Information Criterion (AIC) in this model was 886.72. Fisher's scoring Algorithm needed 6 iterations to perform the fit. An ANOVA table shows degrees of freedom (DF), sum of square (Sum Sq), mean square error (Mean Sq), the F value and Pr(>F) value of models with and without the subpopulation factor. For the post hoc results, each line shows the magnitude and direction (diff) of the difference between means, showing the lower (lwr) and upper (upr) confidence intervals and the significance level (P adj). Significant differences based on the post hoc test are indicated by shaded boxes.

```
Call:
glm(formula = TC.infection.ITS ~ Subpopulation, family = binomial,
    data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.6520  -0.6248  -0.4809  -0.4401   2.4339

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)       -1.53471    0.22050  -6.960  3.4e-12 ***
SubpopulationBRGp -0.05922    0.30805  -0.192    0.8476
SubpopulationMuGl -0.73397    0.64532  -1.137    0.2554
SubpopulationMuGp -0.56407    0.28931  -1.950    0.0512 .
SubpopulationSaGl -1.37401    0.63254  -2.172    0.0298 *
SubpopulationZuGa  0.09435    0.26744   0.353    0.7242
SubpopulationZuGp -0.75106    0.37474  -2.004    0.0450 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 891.49  on 1089  degrees of freedom
Residual deviance: 872.72  on 1083  degrees of freedom
AIC: 886.72

Number of Fisher Scoring iterations: 5
```

**ANOVA table**

```
                Df Sum Sq Mean Sq F value  Pr(>F)
Subpopulation    6   2.16  0.3601   2.981 0.00682 **
Residuals     1083 130.80  0.1208
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Post hoc test**

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = TCa1.6, data = TK, projections = FALSE, qr = TRUE, contrasts = NULL)

$Subpopulation
                    diff          lwr         upr     p adj
BRGp-BRGb -0.008473796 -0.128126289 0.111178698 0.9999930
MuGl-BRGb -0.083554965 -0.284565661 0.117455732 0.8833413
MuGp-BRGb -0.068089947 -0.173306038 0.037126143 0.4729605
SaGl-BRGb -0.125580827 -0.285714862 0.034553209 0.2368953
ZuGa-BRGb  0.014184397 -0.091696286 0.120065081 0.9997021
ZuGp-BRGb -0.084997272 -0.209817243 0.039822698 0.4078640
MuGl-BRGp -0.075081169 -0.274516356 0.124354019 0.9245346
MuGp-BRGp -0.059616152 -0.161790118 0.042557815 0.6003152
SaGl-BRGp -0.117107031 -0.275258866 0.041044804 0.3032043
ZuGa-BRGp  0.022658193 -0.080200024 0.125516410 0.9950267
ZuGp-BRGp -0.076523477 -0.198790066 0.045743113 0.5151873
MuGp-MuGl  0.015465017 -0.175658243 0.206588277 0.9999846
SaGl-MuGl -0.042025862 -0.268080245 0.184028521 0.9980596
ZuGa-MuGl  0.097739362 -0.093750570 0.289229293 0.7402883
ZuGp-MuGl -0.001442308 -0.204019947 0.201135331 1.0000000
SaGl-MuGp -0.057490879 -0.205022942 0.090041184 0.9118932
ZuGa-MuGp  0.082274345 -0.003361625 0.167910315 0.0690711
ZuGp-MuGp -0.016907325 -0.125086922 0.091272273 0.9992755
ZuGa-SaGl  0.139765224 -0.008241543 0.287771990 0.0786683
ZuGp-SaGl  0.040583554 -0.121513049 0.202680158 0.9900766
ZuGp-ZuGa -0.099181669 -0.208007763 0.009644424 0.1011714
```

## A.11. The best fitting GLM of *T. brucei* status of all tsetse fly samples (Model 1.2).

### Backward elimination of model selection

Explanatory variables: subpopulation, sex and age
Response variables: *T. brucei* status in all tsetse species (N = 1090)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation * sex * age (full model) | 28 | 260.7 | 577.3 | - | - |
| subpopulation * sex * age | 22 | 262.6 | 569.2 | 3.8810 | 0.6928 |
| sex * age | 21 | 262.8 | 567.5 | 0.3034 | 0.5818 |
| subpopulation * age | 15 | 265.6 | 561.2 | 5.6532 | 0.4631 |
| subpopulation * sex | 9 | 270.2 | 558.4 | 9.1979 | 0.1628 |
| age | 8 | 270.3 | 556.6 | 0.2534 | 0.6147 |
| sex | 7 | 271.1 | 556.2 | 1.5944 | 0.2067 |
| subpopulation | 1 | 275.7 | 553.4 | 9.1575 | 0.1649 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TB..ITS1. ~ 1, family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.3802  -0.3802  -0.3802  -0.3802   2.3079

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.5909     0.1189  -21.79   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 551.38  on 1089  degrees of freedom
Residual deviance: 551.38  on 1089  degrees of freedom
AIC: 553.38

Number of Fisher Scoring iterations: 5
```

### ANOVA table

```
             Df Sum Sq Mean Sq F value Pr(>F)
Residuals  1089   70.7 0.06492
```

## A.12. The best fitting GLM of *T. vivax* status of all tsetse fly samples (Model 1.3).

### Backword elimination of model selection

Explanatory variables: subpopulation, sex and age
Response variables: *T. vivax* status in all tsetse species (N = 1090)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation * sex * age (full model) | 28 | 567.3 | 1190.5 | - | - |
| subpopulation * sex * age | 22 | 569.1 | 1182.2 | 3.6546 | 0.7233 |
| subpopulation * sex | 21 | 572.9 | 1187.7 | 7.5158 | **0.0061** |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TV..ITS1. ~ Subpopulation * Sex, family = binomial,
    data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2278  -0.9005  -0.5365   1.1278   2.3272

Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.9841     0.3216  -6.170 6.83e-10 ***
SubpopulationBRGp          0.1183     0.4384   0.270   0.7874
SubpopulationMuGl          1.3780     0.6008   2.294   0.0218 *
SubpopulationMuGp          0.9186     0.3691   2.489   0.0128 *
SubpopulationSaGl         -0.6549     1.0839  -0.604   0.5457
SubpopulationZuGa          1.3739     0.3478   3.950 7.80e-05 ***
SubpopulationZuGp          1.7328     0.3818   4.539 5.65e-06 ***
SexMale                   -0.2131     0.5706  -0.373   0.7088
SubpopulationBRGp:SexMale  1.3858     0.7024   1.973   0.0485 *
SubpopulationMuGl:SexMale -1.8198     1.2863  -1.415   0.1571
SubpopulationMuGp:SexMale -0.3953     0.6442  -0.614   0.5394
SubpopulationSaGl:SexMale  0.5749     1.2933   0.444   0.6567
SubpopulationZuGa:SexMale  0.1767     0.6940   0.255   0.7990
SubpopulationZuGp:SexMale  0.5822     0.6971   0.835   0.4037
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1244.2  on 1089  degrees of freedom
Residual deviance: 1154.2  on 1076  degrees of freedom
AIC: 1182.2

Number of Fisher Scoring iterations: 5
```

### ANOVA table

```
                  Df Sum Sq Mean Sq F value   Pr(>F)
Subpopulation      6  13.55  2.2588  12.660 7.69e-14 ***
Sex                1   0.00  0.0016   0.009    0.926
Subpopulation:Sex  6   3.02  0.5027   2.817    0.010 *
Residuals       1076 191.99  0.1784
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Post hoc test** (Significant differences are indicated by shaded boxes.)

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = TVa1.4, data = TK, projections = FALSE, qr = TRUE, contrasts = NULL)

$Subpopulation
                    diff          lwr        upr     p adj
BRGp-BRGb   0.094317030 -0.051118675 0.23975274 0.4702556
MuGl-BRGb   0.105274823 -0.139050487 0.34960013 0.8643635
MuGp-BRGb   0.098128918 -0.029759569 0.22601741 0.2613615
SaGl-BRGb  -0.027268281 -0.221908657 0.16737210 0.9996139
ZuGa-BRGb   0.237588652  0.108892363 0.36628494 0.0000013
ZuGp-BRGb   0.348063284  0.196346593 0.49977998 0.0000000
MuGl-BRGp   0.010957792 -0.231452510 0.25336809 0.9999995
MuGp-BRGp   0.003811888 -0.120378946 0.12800272 1.0000000
SaGl-BRGp  -0.121585311 -0.313816355 0.07064573 0.5019562
ZuGa-BRGp   0.143271622  0.018249093 0.26829415 0.0130031
ZuGp-BRGp   0.253746254  0.105133157 0.40235935 0.0000111
MuGp-MuGl  -0.007145904 -0.239453191 0.22516138 1.0000000
SaGl-MuGl  -0.132543103 -0.407308615 0.14222241 0.7883516
ZuGa-MuGl   0.132313830 -0.100439140 0.36506680 0.6301698
ZuGp-MuGl   0.242788462 -0.003441441 0.48901836 0.0562061
SaGl-MuGp  -0.125397199 -0.304720078 0.05392568 0.3740805
ZuGa-MuGp   0.139459734  0.035370573 0.24354890 0.0015629
ZuGp-MuGp   0.249934366  0.118443784 0.38142495 0.0000005
ZuGa-SaGl   0.264856933  0.084957059 0.44475681 0.0003008
ZuGp-SaGl   0.375331565  0.178305718 0.57235741 0.0000005
ZuGp-ZuGa   0.110474632 -0.021801757 0.24275102 0.1724467

$Sex
                   diff         lwr        upr     p adj
Male-Female -0.0023748 -0.05560548 0.05085588 0.9302593
```

$`Subpopulation:Sex`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp:Female-BRGb:Female | 1.314150e-02 | -0.194073759 | 0.22035675 | 1.0000000 |
| MuGl:Female-BRGb:Female | 2.320621e-01 | -0.143097565 | 0.60722168 | 0.7156614 |
| MuGp:Female-BRGb:Female | 1.353709e-01 | -0.051054717 | 0.32179648 | 0.4510289 |
| SaGl:Female-BRGb:Female | -5.421245e-02 | -0.449885052 | 0.34146014 | 0.9999999 |
| ZuGa:Female-BRGb:Female | 2.311209e-01 | 0.057286426 | 0.40495533 | 0.0007278 |
| ZuGp:Female-BRGb:Female | 3.166209e-01 | 0.108883878 | 0.52435788 | 0.0000315 |
| BRGb:Male-BRGb:Female | -2.087912e-02 | -0.270829169 | 0.22907093 | 1.0000000 |
| BRGp:Male-BRGb:Female | 2.124542e-01 | -0.027386056 | 0.45229448 | 0.1479053 |
| MuGl:Male-BRGb:Female | -5.421245e-02 | -0.449885052 | 0.34146014 | 0.9999999 |
| MuGp:Male-BRGb:Female | 3.701562e-02 | -0.156148682 | 0.23017991 | 0.9999955 |
| SaGl:Male-BRGb:Female | -2.785587e-02 | -0.290608356 | 0.23489663 | 1.0000000 |
| ZuGa:Male-BRGb:Female | 2.228709e-01 | -0.068943198 | 0.51468496 | 0.3629441 |
| ZuGp:Male-BRGb:Female | 4.085326e-01 | 0.123139051 | 0.69392624 | 0.0001482 |
| MuGl:Female-BRGp:Female | 2.189206e-01 | -0.154408217 | 0.59224933 | 0.7855434 |
| MuGp:Female-BRGp:Female | 1.222294e-01 | -0.060483880 | 0.30494264 | 0.5936597 |
| SaGl:Female-BRGp:Female | -6.735395e-02 | -0.461291051 | 0.32658315 | 0.9999989 |
| ZuGa:Female-BRGp:Female | 2.179794e-01 | 0.048132243 | 0.38782652 | 0.0014361 |
| ZuGp:Female-BRGp:Female | 3.034794e-01 | 0.099067311 | 0.50789145 | 0.0000616 |
| BRGb:Male-BRGp:Female | -3.402062e-02 | -0.281214186 | 0.21317295 | 0.9999999 |
| BRGp:Male-BRGp:Female | 1.993127e-01 | -0.037653501 | 0.43627893 | 0.2123986 |
| MuGl:Male-BRGp:Female | -6.735395e-02 | -0.461291051 | 0.32658315 | 0.9999989 |
| MuGp:Male-BRGp:Female | 2.387412e-02 | -0.165709845 | 0.21345808 | 1.0000000 |
| SaGl:Male-BRGp:Female | -4.099736e-02 | -0.301129069 | 0.21913434 | 0.9999996 |
| ZuGa:Male-BRGp:Female | 2.097294e-01 | -0.079727158 | 0.49918592 | 0.4548687 |
| ZuGp:Male-BRGp:Female | 3.953911e-01 | 0.112408574 | 0.67837372 | 0.0002565 |
| MuGp:Female-MuGl:Female | -9.669118e-02 | -0.458893764 | 0.26551141 | 0.9997921 |
| SaGl:Female-MuGl:Female | -2.862745e-01 | -0.789258621 | 0.21670960 | 0.8192565 |
| ZuGa:Female-MuGl:Female | -9.411765e-04 | -0.356826873 | 0.35494452 | 1.0000000 |
| ZuGp:Female-MuGl:Female | 8.455882e-02 | -0.289059796 | 0.45817744 | 0.9999681 |
| BRGb:Male-MuGl:Female | -2.529412e-01 | -0.651577901 | 0.14569555 | 0.6782944 |
| BRGp:Male-MuGl:Female | -1.960784e-02 | -0.411984659 | 0.37276897 | 1.0000000 |
| MuGl:Male-MuGl:Female | -2.862745e-01 | -0.789258621 | 0.21670960 | 0.8192565 |
| MuGp:Male-MuGl:Female | -1.950464e-01 | -0.560763075 | 0.17067020 | 0.8793812 |
| SaGl:Male-MuGl:Female | -2.599179e-01 | -0.666704208 | 0.14686837 | 0.6676438 |
| ZuGa:Male-MuGl:Female | -9.191176e-03 | -0.435326914 | 0.41694456 | 1.0000000 |
| ZuGp:Male-MuGl:Female | 1.764706e-01 | -0.245294411 | 0.59823559 | 0.9812014 |

$`Subpopulation:Sex`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| SaGl:Female-MuGp:Female | -1.895833e-01 | -0.572992746 | 0.19382608 | 0.9285269 |
| ZuGa:Female-MuGp:Female | 9.575000e-02 | -0.048001141 | 0.23950114 | 0.6008821 |
| ZuGp:Female-MuGp:Female | 1.812500e-01 | -0.002054760 | 0.36455476 | 0.0561581 |
| BRGb:Male-MuGp:Female | -1.562500e-01 | -0.386295648 | 0.07379565 | 0.5679380 |
| BRGp:Male-MuGp:Female | 7.708333e-02 | -0.141935678 | 0.29610235 | 0.9962141 |
| MuGl:Male-MuGp:Female | -1.895833e-01 | -0.572992746 | 0.19382608 | 0.9285269 |
| MuGp:Male-MuGp:Female | -9.835526e-02 | -0.264963844 | 0.06825332 | 0.7774557 |
| SaGl:Male-MuGp:Female | -1.632267e-01 | -0.407121896 | 0.08066841 | 0.5929551 |
| ZuGa:Male-MuGp:Female | 8.750000e-02 | -0.187457140 | 0.36245714 | 0.9986289 |
| ZuGp:Male-MuGp:Female | 2.731618e-01 | 0.005028447 | 0.54129508 | 0.0409853 |
| ZuGa:Female-SaGl:Female | 2.853333e-01 | -0.092114269 | 0.66278094 | 0.3805470 |
| ZuGp:Female-SaGl:Female | 3.708333e-01 | -0.023378458 | 0.76504512 | 0.0902938 |
| BRGb:Male-SaGl:Female | 3.333333e-02 | -0.384665831 | 0.45133250 | 1.0000000 |
| BRGp:Male-SaGl:Female | 2.666667e-01 | -0.145366863 | 0.67870020 | 0.6476304 |
| MuGl:Male-SaGl:Female | 2.081668e-16 | -0.518464155 | 0.51846415 | 1.0000000 |
| MuGp:Male-SaGl:Female | 9.122807e-02 | -0.295502742 | 0.47795888 | 0.9999486 |
| SaGl:Male-SaGl:Female | 2.635659e-02 | -0.399421694 | 0.45213487 | 1.0000000 |
| ZuGa:Male-SaGl:Female | 2.770833e-01 | -0.167218062 | 0.72138473 | 0.7039478 |
| ZuGp:Male-SaGl:Female | 4.627451e-01 | 0.022634001 | 0.90285620 | 0.0286079 |
| ZuGp:Female-ZuGa:Female | 8.550000e-02 | -0.084983281 | 0.25598328 | 0.9208484 |
| BRGb:Male-ZuGa:Female | -2.520000e-01 | -0.471965712 | -0.03203429 | 0.0092813 |
| BRGp:Male-ZuGa:Female | -1.866667e-02 | -0.227073101 | 0.18973977 | 1.0000000 |
| MuGl:Male-ZuGa:Female | -2.853333e-01 | -0.662780935 | 0.09211427 | 0.3805470 |
| MuGp:Male-ZuGa:Female | -1.941053e-01 | -0.346494021 | -0.04171651 | 0.0016494 |
| SaGl:Male-ZuGa:Female | -2.589767e-01 | -0.493388261 | -0.02456523 | 0.0153602 |
| ZuGa:Male-ZuGa:Female | -8.250000e-03 | -0.274830837 | 0.25833084 | 1.0000000 |
| ZuGp:Male-ZuGa:Female | 1.774118e-01 | -0.082125112 | 0.43694864 | 0.5569742 |
| BRGb:Male-ZuGp:Female | -3.375000e-01 | -0.585131094 | -0.08986891 | 0.0004382 |
| BRGp:Male-ZuGp:Female | -1.041667e-01 | -0.341589257 | 0.13325592 | 0.9718801 |
| MuGl:Male-ZuGp:Female | -3.708333e-01 | -0.765045125 | 0.02337846 | 0.0902938 |
| MuGp:Male-ZuGp:Female | -2.796053e-01 | -0.469759353 | -0.08945117 | 0.0000782 |
| SaGl:Male-ZuGp:Female | -3.444767e-01 | -0.605024252 | -0.08392924 | 0.0008133 |
| ZuGa:Male-ZuGp:Female | -9.375000e-02 | -0.383580273 | 0.19608027 | 0.9983779 |
| ZuGp:Male-ZuGp:Female | 9.191176e-02 | -0.191453080 | 0.37527661 | 0.9983313 |
| BRGp:Male-BRGb:Male | 2.333333e-01 | -0.041784553 | 0.50845122 | 0.2011633 |
| MuGl:Male-BRGb:Male | -3.333333e-02 | -0.451332498 | 0.38466583 | 1.0000000 |
| MuGp:Male-BRGb:Male | 5.789474e-02 | -0.177644948 | 0.29343442 | 0.9999178 |
| SaGl:Male-BRGb:Male | -6.976744e-03 | -0.302282296 | 0.28832881 | 1.0000000 |
| ZuGa:Male-BRGb:Male | 2.437500e-01 | -0.077687323 | 0.56518732 | 0.3751778 |
| ZuGp:Male-BRGb:Male | 4.294118e-01 | 0.113791740 | 0.74503179 | 0.0004545 |
| MuGl:Male-BRGp:Male | -2.666667e-01 | -0.678700196 | 0.14536686 | 0.6476304 |
| MuGp:Male-BRGp:Male | -1.754386e-01 | -0.400221315 | 0.04934412 | 0.3264666 |
| SaGl:Male-BRGp:Male | -2.403101e-01 | -0.527109130 | 0.04648897 | 0.2176557 |
| ZuGa:Male-BRGp:Male | 1.041667e-02 | -0.303223701 | 0.32405703 | 1.0000000 |
| ZuGp:Male-BRGp:Male | 1.960784e-01 | -0.111597254 | 0.50375412 | 0.6716503 |
| MuGp:Male-MuGl:Male | 9.122807e-02 | -0.295502742 | 0.47795888 | 0.9999486 |
| SaGl:Male-MuGl:Male | 2.635659e-02 | -0.399421694 | 0.45213487 | 1.0000000 |
| ZuGa:Male-MuGl:Male | 2.770833e-01 | -0.167218062 | 0.72138473 | 0.7039478 |
| ZuGp:Male-MuGl:Male | 4.627451e-01 | 0.022634001 | 0.90285620 | 0.0286079 |
| SaGl:Male-MuGp:Male | -6.487148e-02 | -0.313955379 | 0.18421242 | 0.9998421 |
| ZuGa:Male-MuGp:Male | 1.858553e-01 | -0.093714713 | 0.46542524 | 0.6041094 |
| ZuGp:Male-MuGp:Male | 3.715170e-01 | 0.098655489 | 0.64437857 | 0.0004474 |
| ZuGa:Male-SaGl:Male | 2.507267e-01 | -0.080763523 | 0.58221701 | 0.3796166 |
| ZuGp:Male-SaGl:Male | 4.363885e-01 | 0.110536019 | 0.76224100 | 0.0006291 |
| ZuGp:Male-ZuGa:Male | 1.856618e-01 | -0.164047298 | 0.53537083 | 0.8829380 |

## A.13. The best fitting GLM of *T. congolense* status of *G. austeni* (Model 2.1).

### Backward elimination of model selection

Explanatory variables: sex and age
Response variables: *T. congolense* status in *G. austeni* (N = 282)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| sex * age (full model) | 4 | 136.2 | 280.5 | - | - |
| sex * age | 3 | 137.1 | 280.3 | 1.7729 | 0.1830 |
| sex | 2 | 137.6 | 279.1 | 0.8805 | 0.3481 |
| age | 1 | 137.7 | 277.4 | 0.3027 | 0.5822 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TC.infection.ITS ~ 1, family = binomial, data = Ga)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-0.652  -0.652  -0.652  -0.652   1.818

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.4404     0.1513  -9.517   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 275.44  on 281  degrees of freedom
Residual deviance: 275.44  on 281  degrees of freedom
AIC: 277.44

Number of Fisher Scoring iterations: 4
```

### ANOVA table

```
             Df Sum Sq Mean Sq F value Pr(>F)
Residuals   281  43.66  0.1554
```

## A.14. The best fitting GLM of *T. brucei* status of *G. austeni* (Model 2.2).

### Backward elimination of model selection

Explanatory variables: sex and age
Response variables: *T. brucei* status in *G. austeni* (N = 282)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| sex ∗ age (full model) | 4 | 71.3 | 150.7 | - | - |
| sex ∗ age | 3 | 71.4 | 148.7 | 0.0067 | 0.9345 |
| sex | 2 | 71.5 | 147.0 | 0.3313 | 0.5649 |
| 1 | 1 | 72.2 | 146.4 | 1.3632 | 0.2430 |

A ∗ B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TB..ITS1. ~ 1, family = binomial, data = Ga)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.3836  -0.3836  -0.3836  -0.3836   2.3005

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.573      0.232  -11.09   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 144.39  on 281  degrees of freedom
Residual deviance: 144.39  on 281  degrees of freedom
AIC: 146.39

Number of Fisher Scoring iterations: 5
```

### ANOVA table

```
             Df Sum Sq Mean Sq F value Pr(>F)
Residuals   281  18.58 0.06613
```

## A.15. The best fitting GLM of *T. vivax* status of *G. austeni* (Model 2.3).

### Backward elimination of model selection

Explanatory variables: sex and age
Response variables: *T.* vivax status in *G. austeni* (N = 282)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| sex * age (full model) | 4 | 180.97 | 369.95 | - | - |
| sex * age | 3 | 182.71 | 371.42 | 3.4705 | 0.06247 |
| sex | 2 | 182.8 | 369.52 | 0.1011 | 0.7506 |
| 1 | 1 | 182.77 | 367.53 | 0.1096 | 0.9467 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TV..ITS1. ~ 1, family = binomial, data = Ga)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-0.930  -0.930  -0.930   1.447   1.447

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.6144     0.1248  -4.924 8.47e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 365.53  on 281  degrees of freedom
Residual deviance: 365.53  on 281  degrees of freedom
AIC: 367.53

Number of Fisher Scoring iterations: 4
```

**ANOVA table**

```
             Df Sum Sq Mean Sq F value Pr(>F)
Residuals   281  64.24  0.2286
```

## A.16. The best fitting GLM of *T. brucei* status of *G. brevipalpis* (Model 3.2).

### Backward elimination of model selection

Explanatory variables: sex and age
Response variables: *T.* brucei status in *G. brevipalpis* (N = 141)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| sex ∗ age (full model) | 4 | 43.2 | 94.5 | - | - |
| sex ∗ age | 3 | 43.8 | 93.6 | 1.1901 | 0.2753 |
| sex | 2 | 44.1 | 92.2 | 0.5488 | 0.4588 |
| 1 | 1 | 45.6 | 93.2 | 3.0416 | 0.0812 |

A ∗ B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TB..ITS1. ~ 1, family = binomial, data = Gb)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.4573  -0.4573  -0.4573  -0.4573   2.1493

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.2051     0.2816  -7.831 4.83e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 91.233  on 140  degrees of freedom
Residual deviance: 91.233  on 140  degrees of freedom
AIC: 93.233

Number of Fisher Scoring iterations: 4
```

### ANOVA table

```
             Df Sum Sq Mean Sq F value Pr(>F)
Residuals   140  12.61 0.09007
```

## A.17. The best fitting GLM of *T. vivax* status of *G. brevipalpis* (Model 3.3).

### Backward elimination of model selection

Explanatory variables: sex and age
Response variables: *T. vivax* status in *G. brevipalpis* (N = 141)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| sex * age (full model) | 4 | 49.6 | 107.23 | - | - |
| sex * age | 3 | 49.8 | 105.6 | 0.3714 | 0.5422 |
| sex | 2 | 49.8 | 103.6 | 0.0101 | 0.9200 |
| 1 | 1 | 49.9 | 101.8 | 0.1423 | 0.7061 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TV..ITS1. ~ 1, family = binomial, data = Gb)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.4908  -0.4908  -0.4908  -0.4908   2.0862

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.0557     0.2655  -7.742 9.75e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 99.749  on 140  degrees of freedom
Residual deviance: 99.749  on 140  degrees of freedom
AIC: 101.75

Number of Fisher Scoring iterations: 4
```

### ANOVA table

```
           Df Sum Sq Mean Sq F value Pr(>F)
Residuals  140  14.18  0.1013
```

## A.18. The best fitting GLM of *T. congolense* status of *G. brevipalpis* (Model 3.1).

### Backward elimination of model selection

Explanatory variables: sex and age
Response variables: *T. congolense* status in *G. brevipalpis* (N = 141)

| Effected test s | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| sex ∗ age (full model) | 4 | 62.9 | 133.8 | - | - |
| sex ∗ age | 3 | 65.0 | 136.0 | 4.1177 | **0.0424** |

A ∗ B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TC.infection.ITS ~ Sex * Age_class, family = binomial,
    data = Gb)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.9349  -0.6887  -0.6308  -0.2622   2.0528

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)       -1.1226     0.7109  -1.579   0.1143
SexMale           -3.6067     1.7671  -2.041   0.0413 *
Age_class         -0.1955     0.4715  -0.415   0.6785
SexMale:Age_class  1.5715     0.8190   1.919   0.0550 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 131.77  on 140  degrees of freedom
Residual deviance: 125.83  on 137  degrees of freedom
AIC: 133.83

Number of Fisher Scoring iterations: 5
```

### ANOVA table

```
              Df Sum Sq Mean Sq F value Pr(>F)
Sex            1  0.108  0.1078   0.744  0.390
Age_class      1  0.154  0.1539   1.062  0.304
Sex:Age_class  1  0.457  0.4567   3.152  0.078 .
Residuals    137 19.849  0.1449
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

## Post hoc test

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Gb_TCa1a, data = Gb, projections = FALSE, qr = TRUE, contrasts = NULL)

  $Sex
                   diff        lwr        upr     p adj
  Male-Female -0.0578022 -0.1910781 0.07547371 0.3925609

  $Age
                        diff        lwr       upr     p adj
  Old-Juvenile    0.141514914 -0.1295668 0.4125966 0.4334625
  Young-Juvenile -0.003446724 -0.1655859 0.1586924 0.9986018
  Young-Old      -0.144961638 -0.4102544 0.1203312 0.4004934

  $`Sex:Age`
                                       diff        lwr        upr     p adj
  Male:Juvenile-Female:Juvenile  0.01041667 -0.2884395 0.30927288 0.9999985
  Female:Old-Female:Juvenile     0.09375000 -0.4931920 0.68069196 0.9973243
  Male:Old-Female:Juvenile       0.14375000 -0.2572077 0.54470769 0.9047831
  Female:Young-Female:Juvenile   0.06193182 -0.1841345 0.30799814 0.9782179
  Male:Young-Female:Juvenile    -0.15625000 -0.4951211 0.18262110 0.7661557
  Female:Old-Male:Juvenile       0.08333333 -0.5143791 0.68104577 0.9986046
  Male:Old-Male:Juvenile         0.13333333 -0.2832316 0.54989829 0.9393515
  Female:Young-Male:Juvenile     0.05151515 -0.2192392 0.32226952 0.9938981
  Male:Young-Male:Juvenile      -0.16666667 -0.5238682 0.19053483 0.7570612
  Male:Old-Female:Old            0.05000000 -0.6047612 0.70476116 0.9999266
  Female:Young-Female:Old       -0.03181818 -0.6049619 0.54132556 0.9999849
  Male:Young-Female:Old         -0.25000000 -0.8686911 0.36869115 0.8511253
  Female:Young-Male:Old         -0.08181818 -0.4622915 0.29865515 0.9892633
  Male:Young-Male:Old           -0.30000000 -0.7461445 0.14614453 0.3802233
  Male:Young-Female:Young       -0.21818182 -0.5325487 0.09618502 0.3441124
```

# A.19. The best fitting GLM of *T. brucei* status of *G. longipennis* (Model 4.2).

## Backward elimination of model selection

Explanatory variables: site, sex and age
Response variables: *T. brucei* status in *G. longipennis* (N = 90)

| Effective tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| site ∗ sex ∗ age (full model) | 8 | 18.5 | 52.9 | - | - |
| site ∗ sex ∗ age | 7 | 18.5 | 50.9 | <0.0001 | 0.9998 |
| sex ∗ age | 6 | 18.5 | 48.9 | <0.0001 | 0.9999 |
| site ∗ age | 5 | 19.0 | 48.1 | 1.1452 | 0.2846 |
| site ∗ sex | 4 | 19.0 | 46.1 | <0.0001 | 1.0000 |
| age | 3 | 19.1 | 44.3 | 0.1889 | 0.6638 |
| sex | 2 | 21.5 | 47.0 | 4.7058 | **0.0301** |
| site | 2 | 19.3 | 42.6 | 0.3208 | 0.5712 |
| sex | 1 | 22.0 | 46.1 | 5.5065 | **0.0190** |

A ∗ B means an interaction between A and B.

## Summary of the best fitting model

```
Call:
glm(formula = TB..ITS1. ~ Sex, family = binomial, data = Gl)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.46733  -0.46733  -0.46733  -0.00008   2.13011

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -19.57    1901.06  -0.010    0.992
SexMale        17.41    1901.06   0.009    0.993

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 44.087  on 89  degrees of freedom
Residual deviance: 38.581  on 88  degrees of freedom
AIC: 42.581

Number of Fisher Scoring iterations: 18
```

## ANOVA table

```
            Df Sum Sq Mean Sq F value Pr(>F)
Sex          1  0.286 0.28553    3.22 0.0749 .
Residuals  139 12.324 0.08866
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

## Post hoc test

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Gl_TBa1.7, data = Gl, projections = FALSE, qr = TRUE, contrasts = NULL)

  $Sex
                    diff          lwr       upr     p adj
  Male-Female 0.09406593 -0.009573488 0.1977054 0.0749003
```

## A.20. The best fitting GLM of *T. congolense* status of *G. longipennis* (Model 4.1).

### Backward elimination of model selection

Explanatory variables: site, sex and age
Response variables: *T.congolense* status in *G. longipennis* (N = 90)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| site ∗ sex ∗ age (full model) | 8 | 19.4 | 54.9 | - | - |
| site ∗ sex ∗ age | 7 | 19.4 | 52.9 | <0.0001 | 0.9999 |
| sex ∗ age | 6 | 19.4 | 50.9 | 0.0003 | 0.9852 |
| site ∗ age | 5 | 20.1 | 50.1 | 1.2450 | 0.2645 |
| site ∗ sex | 4 | 21.0 | 50.1 | 1.9270 | 0.1651 |
| age | 3 | 21.7 | 49.4 | 1.3524 | 0.2449 |
| sex | 2 | 21.8 | 47.5 | 0.1139 | 0.7357 |
| site | 1 | 22.0 | 46.1 | 0.56207 | 0.4534 |

A ∗ B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TC.infection.ITS ~ 1, family = binomial, data = Gl)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.3715  -0.3715  -0.3715  -0.3715   2.3272

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.6391     0.4226  -6.245 4.23e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 44.087  on 89  degrees of freedom
Residual deviance: 44.087  on 89  degrees of freedom
AIC: 46.087

Number of Fisher Scoring iterations: 5
```

### ANOVA table

```
            Df Sum Sq Mean Sq F value Pr(>F)
Residuals   89    5.6 0.06292
```

## A.21. The best fitting GLM of *T. vivax* status of *G. longipennis* (Model 4.3).

### Backward elimination of model selection

Explanatory variables: site, sex and age
Response variables: *T. vivax* status in *G. longipennis* (N = 90)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| site * sex * age (full model) | 8 | 29.9 | 75.9 | - | - |
| site * sex * age | 7 | 30.7 | 75.5 | 1.5684 | 0.2104 |
| sex * age | 6 | 30.9 | 73.7 | 0.2894 | 0.5906 |
| site * age | 5 | 31.2 | 72.4 | 0.6367 | 0.4249 |
| site * sex | 4 | 32.6 | 73.3 | 2.9252 | 0.0872 |
| age | 3 | 33.0 | 72.0 | 0.6638 | 0.4152 |
| sex | 2 | 33.8 | 71.7 | 1.7245 | 0.1891 |
| site | 1 | 35.3 | 72.7 | 2.9948 | 0.0835 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TV..ITS1. ~ 1, family = binomial, data = Gl)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-0.535  -0.535  -0.535  -0.535   2.007

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.8718     0.3101  -6.036 1.58e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 70.681  on 89  degrees of freedom
Residual deviance: 70.681  on 89  degrees of freedom
AIC: 72.681

Number of Fisher Scoring iterations: 4
```

### ANOVA table

```
           Df Sum Sq Mean Sq F value Pr(>F)
Residuals  89   10.4  0.1168
```

## A.22. The best fitting GLM of *T. brucei* status of *G. pallidipes* (Model 5.2).

### Backward elimination of model selection

Explanatory variables: site, sex and age
Response variables: *T. brucei* status in *G. pallidipes* (N = 577)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| site * sex * age (full model) | 12 | 127.6 | 279.3 | - | - |
| site * sex * age | 10 | 128.3 | 276.6 | 1.3678 | 0.5046 |
| sex * age | 9 | 129.1 | 276.3 | 1.6197 | 0.2031 |
| site * age | 7 | 130.5 | 275.0 | 2.721 | 0.2565 |
| site * sex | 5 | 131.1 | 272.3 | 1.2868 | 0.5255 |
| age | 4 | 131.6 | 271.2 | 0.89382 | 0.3444 |
| sex | 3 | 131.8 | 269.6 | 0.47261 | 0.4918 |
| site | 1 | 134.7 | 271.5 | 5.8277 | 0.0543 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TB..ITS1. ~ 1, family = binomial, data = Gp)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.3589  -0.3589  -0.3589  -0.3589   2.3556

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.7099     0.1721  -15.74   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 269.46  on 576  degrees of freedom
Residual deviance: 269.46  on 576  degrees of freedom
AIC: 271.46

Number of Fisher Scoring iterations: 5
```

### ANOVA table

```
            Df Sum Sq Mean Sq F value Pr(>F)
Residuals   576  33.75  0.0586
```

## A.23. The best fitting GLM of *T. congolense* status of *G. pallidipes* (Model 5.1).

### Backward elimination of model selection

Explanatory variables: site, sex and age
Response variables: *T. congolense* status in *G. pallidipes* (N = 577)

| Effected tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| site * sex * age (full model) | 12 | 206.2 | 436.4 | - | - |
| site * sex * age | 10 | 206.6 | 433.1 | 0.7510 | 0.6869 |
| sex * age | 9 | 207.1 | 432.3 | 1.1436 | 0.2849 |
| site * age | 7 | 208.4 | 430.8 | 2.5267 | 0.2827 |
| site * sex | 5 | 211.0 | 432.0 | 5.1885 | 0.0747 |
| age | 4 | 211.0 | 430.0 | 0.0056 | 0.9406 |
| sex | 3 | 211.0 | 428.0 | 0.0058 | 0.9392 |
| site | 1 | 213.2 | 428.5 | 4.4758 | 0.1067 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TC.infection.ITS ~ 1, family = binomial, data = Gp)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.5086  -0.5086  -0.5086  -0.5086   2.0539

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.9800     0.1275  -15.53   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 426.45  on 576  degrees of freedom
Residual deviance: 426.45  on 576  degrees of freedom
AIC: 428.45

Number of Fisher Scoring iterations: 4
```

### ANOVA table

```
            Df Sum Sq Mean Sq F value Pr(>F)
Residuals   576  61.51  0.1068
```

## A.24. The best fitting GLM of *T. vivax* status of *G. pallidipes* (Model 5.3).

### Backward elimination of model selection

Explanatory variables: site, sex and age
Response variables: *T. vivax* status in *G. pallidipes* (N = 577)

| Factors | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| site ∗ sex ∗ age (full model) | 12 | 306.8 | 637.5 | - | - |
| site ∗ sex ∗ age | 10 | 307.0 | 633.9 | 0.4228 | 0.8095 |
| sex ∗ age | 9 | 309.5 | 637.0 | 5.0480 | **0.0247** |
| sex ∗ age | 8 | 307.9 | 631.8 | 1.8623 | 0.3941 |
| site ∗ age | 6 | 640.7 | 640.7 | 12.9570 | **0.0015** |

A ∗ B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TV..ITS1. ~ Site + Sex + Age_class + Site:Sex +
    Sex:Age_class, family = binomial, data = Gp)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6526  -0.7783  -0.5454   1.0163   2.1510

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -1.75358    0.42965  -4.081 4.48e-05 ***
SiteMukinyo            0.79135    0.34958   2.264 0.023591 *
SiteZungu Luka         1.62137    0.36284   4.469 7.88e-06 ***
SexMale               -0.17401    0.66247  -0.263 0.792807
Age_class             -0.07683    0.21356  -0.360 0.719039
SiteMukinyo:SexMale   -1.75243    0.51480  -3.404 0.000664 ***
SiteZungu Luka:SexMale -0.66120   0.58756  -1.125 0.260444
SexMale:Age_class      0.75629    0.30679   2.465 0.013695 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 669.49  on 576  degrees of freedom
Residual deviance: 615.78  on 569  degrees of freedom
AIC: 631.78

Number of Fisher Scoring iterations: 4
```

### ANOVA table

```
             Df Sum Sq Mean Sq F value   Pr(>F)
Site          2   6.36   3.179  17.663 3.61e-08 ***
Sex           1   0.04   0.040   0.224  0.63635
Age_class     1   0.59   0.593   3.293  0.07012 .
Site:Sex      2   2.38   1.189   6.605  0.00146 **
Sex:Age_class 1   1.10   1.104   6.134  0.01355 *
Residuals   569 102.42   0.180
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Post hoc test** (Significant differences are indicated by shaded boxes.)

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Gp_TVa1.3a, data = Gp, projections = FALSE, qr = TRUE, contrasts = NULL)

$Site
                               diff        lwr       upr     p adj
Mukinyo-Buffalo Ridge    0.003811888 -0.0952837 0.1029075 0.9955055
Zungu Luka-Buffalo Ridge 0.253746254  0.1351634 0.3723291 0.0000020
Zungu Luka-Mukinyo       0.249934366  0.1450141 0.3548546 0.0000001

$Sex
                  diff         lwr        upr     p adj
Male-Female 0.01693078 -0.05415805 0.08801961 0.6401139

$Age
                     diff        lwr        upr     p adj
Old-Juvenile    0.02914977 -0.1217098 0.18000936 0.8926214
Young-Juvenile -0.05574574 -0.1457875 0.03429600 0.3136551
Young-Old      -0.08489551 -0.2289482 0.05915723 0.3495032

$`Site:Sex`
                                                      diff          lwr          upr     p adj
Mukinyo:Female-Buffalo Ridge:Female             0.12321539 -0.032688977  0.27911976 0.2124543
Zungu Luka:Female-Buffalo Ridge:Female          0.30030721  0.125887820  0.47472660 0.0000165
Buffalo Ridge:Male-Buffalo Ridge:Female         0.19818941 -0.004007567  0.40038638 0.0584395
Mukinyo:Male-Buffalo Ridge:Female               0.02177199 -0.139994963  0.18353895 0.9989043
Zungu Luka:Male-Buffalo Ridge:Female            0.40275816  0.161296656  0.64421967 0.0000345
Zungu Luka:Female-Mukinyo:Female                0.17709181  0.020682734  0.33350089 0.0160401
Buffalo Ridge:Male-Mukinyo:Female               0.07497401 -0.111909089  0.26185712 0.8613198
Mukinyo:Male-Mukinyo:Female                    -0.10144340 -0.243606074  0.04071928 0.3207281
Zungu Luka:Male-Mukinyo:Female                  0.27954277  0.050751735  0.50833380 0.0067776
Buffalo Ridge:Male-Zungu Luka:Female           -0.10211780 -0.304704187  0.10046859 0.7016237
Mukinyo:Male-Zungu Luka:Female                 -0.27853521 -0.440788646 -0.11628178 0.0000177
Zungu Luka:Male-Zungu Luka:Female               0.10245096 -0.139336735  0.34423865 0.8311933
Mukinyo:Male-Buffalo Ridge:Male                -0.17641741 -0.368218534  0.01538371 0.0917485
Zungu Luka:Male-Buffalo Ridge:Male              0.20456876 -0.057962730  0.46710024 0.2264212
Zungu Luka:Male-Mukinyo:Male                    0.38098617  0.148160670  0.61381167 0.0000528

$`Sex:Age`
                                        diff          lwr         upr     p adj
Male:Juvenile-Female:Juvenile   0.002023136 -0.173718309  0.17776458 1.0000000
Female:Old-Female:Juvenile     -0.186439382 -0.479961312  0.10708255 0.4560037
Male:Old-Female:Juvenile        0.156434326 -0.074945713  0.38781437 0.3828393
Female:Young-Female:Juvenile   -0.039280101 -0.178911322  0.10035112 0.9666857
Male:Young-Female:Juvenile     -0.091616282 -0.254348396  0.07111583 0.5923411
Female:Old-Male:Juvenile       -0.188462518 -0.490973236  0.11404820 0.4786018
Male:Old-Male:Juvenile          0.154411191 -0.088270358  0.39709274 0.4539965
Female:Young-Male:Juvenile     -0.041303236 -0.198956304  0.11634983 0.9755614
Male:Young-Male:Juvenile       -0.093639418 -0.272075268  0.08479643 0.6640455
Male:Old-Female:Old             0.342873708  0.004991478  0.68075594 0.0444013
Female:Young-Female:Old         0.147159282 -0.135903533  0.43022210 0.6730376
Male:Young-Female:Old           0.094823100 -0.200319950  0.38996615 0.9416925
Female:Young-Male:Old          -0.195714427 -0.413673498  0.02224464 0.1070342
Male:Young-Male:Old            -0.248050608 -0.481483724 -0.01461749 0.0297909
Male:Young-Female:Young        -0.052336182 -0.195343786  0.09067142 0.9019034
```

# Appendix B Supplementary Information for chapter 3

## B.1. Cloning and sequencing results of pSG2 fragments.

BLAST results for amplification products using pSG2 primers for *G. pallidipes* (Gp), *G. longipennis* (Gl) and *G. austeni* (Ga) from Zungu Luka (Zu) and Mukinyo (Mu), for which Hem and GPO1 showed no amplification. Note that S. glossinidius was only confirmed for the 120 bp bands but not all 120 bp bands were confirmed to be S. *glossinidius*.

| Samples | Sizes (bp) | Description | Query cover (%) | Identity (%) | Accession No. |
|---|---|---|---|---|---|
| ZuGp11HP_pSG2_M13F plasmid 1, 3, 4, 6-8 and 10 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGp11HP_pSG2_M13F plasmid 2 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 95 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 95 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 95 | AJ868435.1 |
| ZuGp11HP_pSG2_M13F plasmid 5 and 9 | 130 | unidentified | | | |
| ZuGp125HP_pSG2_M13F plasmid 1-6 | 130 | unidentified | | | |
| MuGp159HP_pSG2_M13F plasmid 1-6 | 130 | unidentified | | | |
| MuGp250HP_pSG2_M13F plasmid 1-6 | 130 | unidentified | | | |
| MuGl25HP_pSG2_M13F plasmid 1 | 115 | unidentified | | | |
| MuGl25HP_pSG2_M13F plasmid 2 | 130 | unidentified | | | |
| MuGl25HP_pSG2_M13F plasmid 3 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| MuGl25HP_pSG2_M13F plasmid 4 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 100 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 100 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 100 | AJ868435.1 |
| MuGl25HP_pSG2_M13F plasmid 5 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 98 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 98 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 98 | AJ868435.1 |
| MuGl25HP_pSG2_M13F plasmid 6 | 132 | unidentified | | | |
| SaGl24HP_pSG2_M13F plasmid 1-6 | 120 | unidentified | | | |
| ZuGa25HP_ pSG2_M13F plasmid 1 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 100 | AP008234.1 |

| Samples | Sizes (bp) | Description | Query cover (%) | Identity (%) | Accession No. |
|---|---|---|---|---|---|
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 100 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 100 | AJ868435.1 |
| ZuGa25HP_ pSG2_M13F plasmid 2 and 4-6 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGa25HP_ pSG2_M13F plasmid 2 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 98 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 98 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 98 | AJ868435.1 |
| ZuGa43HP_ pSG2_M13F plasmid 1 and 4-6 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGa43HP_ pSG2_M13F plasmid 2 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 98 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 98 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 98 | AJ868435.1 |
| ZuGa43HP_ pSG2_M13F plasmid 3 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 100 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 100 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 100 | AJ868435.1 |
| ZuGa88HP_ pSG2_M13F plasmid 1 and 4-5 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 100 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 100 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 100 | AJ868435.1 |
| ZuGa88HP_ pSG2_M13F plasmid 2-3 and 6 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGa94HP_ pSG2_M13F plasmid 1,2 and 6 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGa94HP_ pSG2_M13F plasmid 3-5 | 120 | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 100 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 100 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 100 | AJ868435.1 |

## B.2. Sequencing results of Hem and GPO1 positive PCR products.

The sequencing experiments were for investigation of genetic diversity among *G. austeni*, (Ga), *G. brevipalpis* (Gb) and *G. pallidipes* (Gp) from Buffalo Ridge (BR) and Zungu Luka (Zu).

| Samples | Sizes (bp) | Description | Query cover (%) | Identity (%) | Accession No. |
|---|---|---|---|---|---|
| **Hem primers** | | | | | |
| ZuGp 3 HP_HemF | 587 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| ZuGp 40 HP_HemF | 588 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| ZuGp51 HP_HemF/R | 640 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| ZuGp98 HP_HemF/R | 587 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| ZuGa35 HP_HemF/R | 625 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| ZuGa115 HP_HemF/R | 642 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 99 | 99 | P008232.1 |
| ZuGa145 HP_HemF | 599 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| ZuGa187 HP_HemF | 599 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| BRGb64 HP_HemF/R | 601 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 99 | 99 | P008232.1 |
| BRGp16 HP_HemF/R | 644 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 99 | 99 | P008232.1 |
| ZuGp14 HP_HemF/R | 643 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 99 | 99 | P008232.1 |
| BRGb70 HP_HemF/R | 601 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| ZuGa280 HP_HemF/R | 642 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 99 | 99 | P008232.1 |
| BRGp154 HP_HemF/R | 590 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| BRGp85 HP_HemF/R | 605 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 99 | 99 | P008232.1 |
| ZuGp32 HP_HemF | 593 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 99 | 99 | P008232.1 |
| ZuGp44 HP_HemR | 599 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| BRGb21 HP_HemF/R | 508 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| BRGp58 HP_HemF/R | 590 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| BRGp110 HP_HemF/R | 601 | *Sodalis glossinidius* str. 'morsitans' DNA, complete genome | 100 | 99 | P008232.1 |
| **GPO1 primers** | | | | | |
| BRGp16HP_GPO1F/R | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |

| Samples | Sizes (bp) | Description | Query cover (%) | Identity (%) | Accession No. |
|---|---|---|---|---|---|
| BRGp85HP_GPO1F/R | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| BRGp115HP_GPO1F/R | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| BRGp154HP_GPO1F/R | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGp3HP_GPO1F | 831 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGp32HP_GPO1F | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGp44HP_GPO1F/R | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGp51HP_GPO1F | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |

| Samples | Sizes (bp) | Description | Query cover (%) | Identity (%) | Accession No. |
|---|---|---|---|---|---|
| ZuGp98HP _GPO1F | 1050 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| BRGp280HP_ GPO1F/R | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGa40HP _GPO1F | 906 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| BRGb70HP _GPO1F/R | 1190 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGp14HP _GPO1F | 665 | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| ZuGa35HP _GPO1F | 1054 | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| ZuGa115HP_ GPO1F/R | 1190 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |

| Samples | Sizes (bp) | Description | Query cover (%) | Identity (%) | Accession No. |
|---|---|---|---|---|---|
| ZuGa145HP_ GPO1F/R | 830 | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| ZuGa187HP_ GPO1F/R | 831 | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 100 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 100 | AJ868435.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| BRGb21HP _GPO1F | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 98 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 98 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 98 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 98 | AJ868435.1 |
| BRGb64HP _GPO1F/R | 1200 | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 99 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 99 | AJ868435.1 |
| ZuGa35HP _GPO1F/R | 1200 | *Sodalis glossinidius* pSG2 plasmid from *Glossina austini* | 100 | 100 | AJ868436.1 |
| | | *Sodalis glossinidius* pSG2 plasmid from *Glossina palpalis palpalis* | 100 | 100 | AJ868435.1 |
| | | *Sodalis glossinidius* str. 'morsitans' plasmid pSG2, complete sequence | 100 | 99 | AP008234.1 |
| | | *Sodalis glossinidius* partial RepA gene for replication-associated protein, strain *pallidipes*-IAEA | 100 | 99 | LN887947.1 |

## B.3. Sample numbers of tsetse flies from different categories: subpopulation, sex, age trypanosome status and *Sodalis* status

| Subpopulation | Sex | Age group | Trypanosome status | | *Sodalis* status | |
|---|---|---|---|---|---|---|
| | | | positive | negative | positive | negative |
| BRGb (N = 141) | female | young | 16 | 45 | 42 | 19 |
| | | juvenile | 8 | 18 | 21 | 5 |
| | | old | 1 | 9 | 7 | 3 |
| | males | young | 13 | 13 | 18 | 8 |
| | | juvenile | 7 | 12 | 12 | 7 |
| | | old | 8 | 4 | 8 | 4 |
| BRGp (N = 154) | female | young | 21 | 34 | 55 | 0 |
| | | juvenile | 10 | 22 | 32 | 0 |
| | | old | 1 | 3 | 4 | 0 |
| | males | young | 4 | 12 | 14 | 2 |
| | | juvenile | 10 | 14 | 23 | 1 |
| | | old | 3 | 7 | 10 | 0 |
| ZuGp (N = 130) | female | young | 27 | 22 | 27 | 22 |
| | | juvenile | 26 | 15 | 29 | 12 |
| | | old | 4 | 2 | 1 | 5 |
| | males | young | 9 | 9 | 12 | 6 |
| | | juvenile | 9 | 2 | 5 | 6 |
| | | old | 3 | 2 | 3 | 2 |
| ZuGa (N = 282) | female | young | 71 | 63 | 20 | 114 |
| | | juvenile | 57 | 52 | 17 | 92 |
| | | old | 6 | 1 | 0 | 7 |
| | males | young | 7 | 10 | 5 | 12 |
| | | juvenile | 6 | 3 | 3 | 6 |
| | | old | 4 | 2 | 1 | 5 |
| MuGp (N = 293) | female | young | 39 | 69 | 1 | 107 |
| | | juvenile | 27 | 21 | 0 | 48 |
| | | old | 1 | 3 | 0 | 4 |
| | males | young | 15 | 48 | 0 | 63 |
| | | juvenile | 15 | 36 | 0 | 51 |
| | | old | 9 | 10 | 1 | 18 |
| MuGl (N = 32) | female | young | 3 | 2 | 0 | 5 |
| | | juvenile | 2 | 6 | 0 | 8 |
| | | old | 3 | 1 | 0 | 4 |
| | males | young | 2 | 6 | 0 | 8 |
| | | juvenile | 1 | 0 | 0 | 1 |
| | | old | 0 | 6 | 0 | 6 |
| SaGl (N = 58) | female | young | 1 | 7 | 0 | 8 |
| | | juvenile | 0 | 5 | 0 | 5 |
| | | old | 0 | 2 | 0 | 2 |
| | males | young | 3 | 12 | 0 | 15 |
| | | juvenile | 5 | 18 | 0 | 23 |
| | | old | 2 | 3 | 0 | 5 |

## B.4. The best fitting GLM model of *S. glossinidius* in all tsetse samples (Model 1).

### Backward elimination of model selection

Explanatoryvariables: subpopulation, sex, age and trypanosome status
Response variables: *Sodalis* status in all testse flies (N = 1090)

| Effect tested | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation * sex * age * trypanosome status (full model) | 55 | 316.6 | 743.2 | - | - |
| subpopulation * sex * age * trypanosome status | 50 | 318.5 | 736.9 | 3.6735 | 0.5973 |
| sex * age * trypanosome status | 49 | 319.0 | 736.0 | 1.0727 | 0.3003 |
| subpopulation * age * trypanosome status | 43 | 321.5 | 729.0 | 5.0442 | 0.5382 |
| subpopulation * sex * trypanosome status | 37 | 321.7 | 717.4 | 0.3686 | 0.9991 |
| subpopulation * sex * age | 31 | 323.8 | 709.5 | 4.1281 | 0.6594 |
| age * trypanosome status | 30 | 323.9 | 707.8 | 0.3106 | 0.5773 |
| sex * trypanosome status | 29 | 324.3 | 706.5 | 0.7095 | 0.3996 |
| subpopulation * trypanosome status | 23 | 326.3 | 698.6 | 4.0035 | 0.6762 |
| sex *age | 22 | 326.4 | 696.8 | 0.2716 | 0.6023 |
| subpopulation * age | 16 | 328.4 | 688.7 | 3.8997 | 0.6902 |
| subpopulation * sex | 10 | 333.5 | 686.9 | 10.2200 | 0.1157 |
| trypanosome status | 9 | 333.7 | 685.4 | 0.4533 | 0.5008 |
| age | 8 | 333.7 | 683.4 | 0.0183 | 0.8924 |
| sex | 7 | 333.7 | 681.4 | 0.0017 | 0.9674 |
| subpopulation | 1 | 1400 | 699.0 | 730.6 | **< 0.0001** |

A * B means an interaction between A and B.

### Summary of the best fitting model

The status of *Sodalis* and trypanosome infection were classified as binomial variables.

```
Call:
glm(formula = Sodalis ~ Subpopulation, family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7749  -0.5968  -0.1170   0.2074   3.1582

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)         3.8286     0.5836   6.560 5.36e-11 ***
SubpopulationBRGp  -2.9752     0.6096  -4.881 1.06e-06 ***
SubpopulationMuGl -22.3947  1153.0506  -0.019    0.985
SubpopulationMuGp  -8.8088     0.9187  -9.588  < 2e-16 ***
SubpopulationSaGl -22.3947   856.4646  -0.026    0.979
SubpopulationZuGa  -5.4638     0.6054  -9.025  < 2e-16 ***
SubpopulationZuGp  -3.4551     0.6103  -5.662 1.50e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1397.99  on 1089  degrees of freedom
Residual deviance:  667.41  on 1083  degrees of freedom
AIC: 681.41

Number of Fisher Scoring iterations: 17
```

**ANOVA table**

```
                Df Sum Sq Mean Sq F value Pr(>F)
Subpopulation    6  137.7  22.942   232.1 <2e-16 ***
Residuals     1083  107.1   0.099
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Post hoc test** (Significant differences are indicated by shaded boxes.)

$Subpopulation

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp-BRGb | -2.774247e-01 | -0.38568186 | -0.169167550 | 0.0000000 |
| MuGl-BRGb | -9.787234e-01 | -1.16059045 | -0.796856357 | 0.0000000 |
| MuGp-BRGb | -9.718975e-01 | -1.06709309 | -0.876701837 | 0.0000000 |
| SaGl-BRGb | -9.787234e-01 | -1.12360676 | -0.833840049 | 0.0000000 |
| ZuGa-BRGb | -8.156028e-01 | -0.91139976 | -0.719805909 | 0.0000000 |
| ZuGp-BRGb | -3.864157e-01 | -0.49934821 | -0.273483217 | 0.0000000 |
| MuGl-BRGp | -7.012987e-01 | -0.88174029 | -0.520857117 | 0.0000000 |
| MuGp-BRGp | -6.944728e-01 | -0.78691599 | -0.602029535 | 0.0000000 |
| SaGl-BRGp | -7.012987e-01 | -0.84438864 | -0.558208767 | 0.0000000 |
| ZuGa-BRGp | -5.381781e-01 | -0.63124045 | -0.445115822 | 0.0000000 |
| ZuGp-BRGp | -1.089910e-01 | -0.21961330 | 0.001631281 | 0.0565661 |
| MuGp-MuGl | 6.825939e-03 | -0.16609532 | 0.179747198 | 0.9999998 |
| SaGl-MuGl | 1.443290e-15 | -0.20452565 | 0.204525648 | 1.0000000 |
| ZuGa-MuGl | 1.631206e-01 | -0.01013244 | 0.336373578 | 0.0804076 |
| ZuGp-MuGl | 5.923077e-01 | 0.40902293 | 0.775592451 | 0.0000000 |
| SaGl-MuGp | -6.825939e-03 | -0.14030749 | 0.126655617 | 0.9999990 |
| ZuGa-MuGp | 1.562946e-01 | 0.07881437 | 0.233774888 | 0.0000001 |
| ZuGp-MuGp | 5.854818e-01 | 0.48760485 | 0.683358654 | 0.0000000 |
| ZuGa-SaGl | 1.631206e-01 | 0.02920952 | 0.297031618 | 0.0061640 |
| ZuGp-SaGl | 5.923077e-01 | 0.44564868 | 0.738966706 | 0.0000000 |
| ZuGp-ZuGa | 4.291871e-01 | 0.33072530 | 0.527648951 | 0.0000000 |

# B.5. The best fitting GLM model of association of *S. glossinidius* infection with *T. congolense* status and intrinsic factors of all tsetse samples (Model 2).

## Backward elimination of model selection

Explanatory variables: subpopulation, sex, age and *T. congolense* status
Response variables: *Sodalis* status

| Effect tested | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation ∗ sex ∗ age ∗ *T. congolense* status (full model) | 52 | 313.3 | 730.7 | - | - |
| subpopulation ∗ sex ∗ age ∗ *T. congolense* status | 49 | 313.3 | 724.7 | <0.0001 | 1 |
| sex ∗ age ∗ *T. congolense* status | 48 | 316.5 | 728.9 | 6.2813 | **0.0122** |
| subpopulation ∗ age ∗ *T. congolense* status | 43 | 313.7 | 713.5 | 0.7952 | 0.9922 |
| subpopulation ∗ sex ∗ *T. congolense* status | 38 | 314.5 | 705.0 | 1.5453 | 0.9078 |
| subpopulation ∗ sex ∗ age | 32 | 316.7 | 697.3 | 4.3218 | 0.6332 |

A ∗ B means an interaction between A and B.

## Summary of the best fitting model

```
Call:
glm(formula = Sodalis_status ~ Subpopulation + Sex + Age_class +
    TC.infection.ITS + Sex:Age_class + Sex:TC.infection.ITS +
    Age_class:TC.infection.ITS + Sex:Age_class:TC.infection.ITS,
    family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7516  -0.5757  -0.0964   0.2143   3.3494

Coefficients:
                                      Estimate Std. Error z value Pr(>|z|)
(Intercept)                            3.6206     0.6678   5.422 5.90e-08 ***
SubpopulationBRGp                     -3.0434     0.6135  -4.961 7.01e-07 ***
SubpopulationMuGl                    -22.6039  1122.7310  -0.020   0.9839
SubpopulationMuGp                     -9.1318     0.9822  -9.297  < 2e-16 ***
SubpopulationSaGl                    -22.9738   795.7821  -0.029   0.9770
SubpopulationZuGa                     -5.5371     0.6125  -9.041  < 2e-16 ***
SubpopulationZuGp                     -3.4745     0.6142  -5.657 1.54e-08 ***
SexMale                                0.2604     0.6161   0.423   0.6725
Age_class                              0.2029     0.2090   0.971   0.3316
TC.infection.ITSPos                    1.5936     0.8271   1.927   0.0540 .
SexMale:Age_class                     -0.3211     0.3442  -0.933   0.3508
SexMale:TC.infection.ITSPos           -2.5581     1.7211  -1.486   0.1372
Age_class:TC.infection.ITSPos         -1.2215     0.5276  -2.315   0.0206 *
SexMale:Age_class:TC.infection.ITSPos  2.9390     1.0716   2.742   0.0061 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1397.99  on 1089  degrees of freedom
Residual deviance:  650.97  on 1076  degrees of freedom
AIC: 678.97

Number of Fisher Scoring iterations: 17
```

**ANOVA table**

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Subpopulation | 6 | 137.65 | 22.942 | 233.406 | <2e-16 | *** |
| Sex | 1 | 0.00 | 0.000 | 0.001 | 0.9693 | |
| Age_class | 1 | 0.00 | 0.002 | 0.018 | 0.8929 | |
| TC.infection.ITS | 1 | 0.04 | 0.036 | 0.370 | 0.5433 | |
| Sex:Age_class | 1 | 0.00 | 0.000 | 0.002 | 0.9683 | |
| Sex:TC.infection.ITS | 1 | 0.58 | 0.579 | 5.890 | 0.0154 | * |
| Age_class:TC.infection.ITS | 1 | 0.28 | 0.282 | 2.864 | 0.0908 | . |
| Sex:Age_class:TC.infection.ITS | 1 | 0.41 | 0.409 | 4.159 | 0.0417 | * |
| Residuals | 1076 | 105.76 | 0.098 | | | |

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Post hoc test** (Significant differences are indicated by shaded boxes.)

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = Sod_TC3, data = TK, projections = FALSE, qr = TRUE, contrasts = NULL)

$Subpopulation

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp-BRGb | -2.774247e-01 | -0.385224983 | -0.169624423 | 0.0000000 |
| MuGl-BRGb | -9.787234e-01 | -1.159822927 | -0.797623882 | 0.0000000 |
| MuGp-BRGb | -9.718975e-01 | -1.066691345 | -0.877103586 | 0.0000000 |
| SaGl-BRGb | -9.787234e-01 | -1.122995315 | -0.834451493 | 0.0000000 |
| ZuGa-BRGb | -8.156028e-01 | -0.910995478 | -0.720210196 | 0.0000000 |
| ZuGp-BRGb | -3.864157e-01 | -0.498871603 | -0.273959821 | 0.0000000 |
| MuGl-BRGp | -7.012987e-01 | -0.880978777 | -0.521618626 | 0.0000000 |
| MuGp-BRGp | -6.944728e-01 | -0.786525857 | -0.602419669 | 0.0000000 |
| SaGl-BRGp | -7.012987e-01 | -0.843784760 | -0.558812643 | 0.0000000 |
| ZuGa-BRGp | -5.381781e-01 | -0.630847700 | -0.445508568 | 0.0000000 |
| ZuGp-BRGp | -1.089910e-01 | -0.219146445 | 0.001164427 | 0.0546370 |
| MuGp-MuGl | 6.825939e-03 | -0.165365550 | 0.179017427 | 0.9999998 |
| SaGl-MuGl | 1.443290e-15 | -0.203662499 | 0.203662499 | 1.0000000 |
| ZuGa-MuGl | 1.631206e-01 | -0.009401272 | 0.335642407 | 0.0779367 |
| ZuGp-MuGl | 5.923077e-01 | 0.409796442 | 0.774818943 | 0.0000000 |
| SaGl-MuGp | -6.825939e-03 | -0.139744169 | 0.126092292 | 0.9999990 |
| ZuGa-MuGp | 1.562946e-01 | 0.079141355 | 0.233447902 | 0.0000001 |
| ZuGp-MuGp | 5.854818e-01 | 0.488017919 | 0.682945589 | 0.0000000 |
| ZuGa-SaGl | 1.631206e-01 | 0.029774655 | 0.296466480 | 0.0058330 |
| ZuGp-SaGl | 5.923077e-01 | 0.446267616 | 0.738347769 | 0.0000000 |
| ZuGp-ZuGa | 4.291871e-01 | 0.331140832 | 0.527233418 | 0.0000000 |

$Sex

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| Male-Female | -0.0007274492 | -0.04018313 | 0.03872824 | 0.971148 |

```
$Age
                    diff         lwr        upr     p adj
Old-Juvenile   -0.04109884 -0.12311328 0.04091559 0.4677222
Young-Juvenile -0.01551861 -0.06298319 0.03194597 0.7231547
Young-Old       0.02558023 -0.05395501 0.10511547 0.7307633

$TC.infection.ITS
             diff         lwr        upr     p adj
Pos-Neg 0.01736404 -0.03591507 0.07064314 0.5226421

$`Sex:Age`
                                      diff         lwr        upr     p adj
Male:Juvenile-Female:Juvenile -0.042151965 -0.13574543 0.05144150 0.7929262
Female:Old-Female:Juvenile    -0.126309124 -0.28303752 0.03041928 0.1945295
Male:Old-Female:Juvenile      -0.014057093 -0.13916553 0.11105135 0.9995493
Female:Young-Female:Juvenile  -0.034007392 -0.10381019 0.03579540 0.7326147
Male:Young-Female:Juvenile    -0.018813531 -0.10753647 0.06990941 0.9906643
Female:Old-Male:Juvenile      -0.084157159 -0.24963605 0.08132173 0.6950216
Male:Old-Male:Juvenile         0.028094872 -0.10781537 0.16400512 0.9916986
Female:Young-Male:Juvenile     0.008144573 -0.07955892 0.09584807 0.9998230
Male:Young-Male:Juvenile       0.023338434 -0.08005998 0.12673685 0.9875757
Male:Old-Female:Old            0.112252032 -0.07288481 0.29738887 0.5114651
Female:Young-Female:Old        0.092301732 -0.06098218 0.24558565 0.5193383
Male:Young-Female:Old          0.107495593 -0.05527814 0.27026933 0.4117210
Female:Young-Male:Old         -0.019950299 -0.14071573 0.10081513 0.9971013
Male:Young-Male:Old           -0.004756438 -0.13735968 0.12784680 0.9999984
Male:Young-Female:Young        0.015193861 -0.06729206 0.09767978 0.9951502

$`Sex:TC.infection.ITS`
                           diff         lwr        upr     p adj
Male:Neg-Female:Neg   -0.018081033 -0.07356871 0.03740665 0.8361218
Female:Pos-Female:Neg -0.022640828 -0.10541647 0.06013482 0.8956068
Male:Pos-Female:Neg    0.099252116 -0.02783213 0.22633636 0.1850222
Female:Pos-Male:Neg   -0.004559795 -0.09297152 0.08385193 0.9991656
Male:Pos-Male:Neg      0.117333149 -0.01349205 0.24815835 0.0968804
Male:Pos-Female:Pos    0.121892943 -0.02263441 0.26642030 0.1322800

$`Age:TC.infection.ITS`
                               diff         lwr        upr     p adj
Old:Neg-Juvenile:Neg      -0.021430645 -0.13109485 0.08823357 0.9936043
Young:Neg-Juvenile:Neg    -0.019269412 -0.08148805 0.04294923 0.9503251
Juvenile:Pos-Juvenile:Neg  0.015254645 -0.11241562 0.14292491 0.9993917
Old:Pos-Juvenile:Neg      -0.118829523 -0.33486129 0.09720225 0.6183879
Young:Pos-Juvenile:Neg     0.023633569 -0.08713568 0.13440282 0.9903960
Young:Neg-Old:Neg          0.002161233 -0.10429008 0.10861254 0.9999999
Juvenile:Pos-Old:Neg       0.036685289 -0.11745886 0.19082944 0.9842091
Old:Pos-Old:Neg           -0.097398878 -0.33005839 0.13526063 0.8392962
Young:Pos-Old:Neg          0.045064214 -0.09540126 0.18552968 0.9425018
Juvenile:Pos-Young:Neg     0.034524056 -0.09039727 0.15944538 0.9694149
Old:Pos-Young:Neg         -0.099560111 -0.31397879 0.11485856 0.7708192
Young:Pos-Young:Neg        0.042902981 -0.06468637 0.15049233 0.8652713
Old:Pos-Juvenile:Pos      -0.134084168 -0.37575317 0.10758483 0.6093745
Young:Pos-Juvenile:Pos     0.008378925 -0.14655333 0.16331118 0.9999878
Young:Pos-Old:Pos          0.142463092 -0.09071931 0.37564550 0.5027150

$`Sex:Age:TC.infection.ITS`
                                                    diff         lwr         upr     p adj
Male:Juvenile:Neg-Female:Juvenile:Neg      -0.079103047 -0.19435573  0.03614964 0.5148666
Female:Old:Neg-Female:Juvenile:Neg         -0.078023074 -0.28668722  0.13064107 0.9869908
Male:Old:Neg-Female:Juvenile:Neg           -0.034836855 -0.18882509  0.11915138 0.9998651
Female:Young:Neg-Female:Juvenile:Neg       -0.049229311 -0.13600284  0.03754422 0.7848000
Male:Young:Neg-Female:Juvenile:Neg         -0.033598281 -0.14243151  0.07523495 0.9974618
Female:Juvenile:Pos-Female:Juvenile:Neg    -0.067649361 -0.24338339  0.10808467 0.9835843
Male:Juvenile:Pos-Female:Juvenile:Neg       0.142265483 -0.11551665  0.40004762 0.8138682
Female:Old:Pos-Female:Juvenile:Neg         -0.288679540 -0.61996981  0.04261073 0.1593521
Male:Old:Pos-Female:Juvenile:Neg            0.044662593 -0.32417816  0.41350335 0.9999998
Female:Young:Pos-Female:Juvenile:Neg       -0.013597362 -0.16040896  0.13321424 1.0000000
Male:Young:Pos-Female:Juvenile:Neg          0.011074314 -0.23995356  0.26210219 1.0000000
Female:Old:Neg-Male:Juvenile:Neg            0.001079973 -0.21718485  0.21934480 1.0000000
Male:Old:Neg-Male:Juvenile:Neg              0.044266192 -0.12250067  0.21103306 0.9993679
Female:Young:Neg-Male:Juvenile:Neg          0.029873736 -0.07796158  0.13770905 0.9990559
Male:Young:Neg-Male:Juvenile:Neg            0.045504766 -0.08076274  0.17177227 0.9903732
Female:Juvenile:Pos-Male:Juvenile:Neg       0.011453686 -0.17557909  0.19848646 1.0000000
```

$`Sex:Age:TC.infection.ITS`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| Male:Juvenile:Pos-Male:Juvenile:Neg | 0.221368530 | -0.04424480 | 0.48698186 | 0.2130464 |
| Female:Old:Pos-Male:Juvenile:Neg | -0.209576493 | -0.54699619 | 0.12784320 | 0.6697003 |
| Male:Old:Pos-Male:Juvenile:Neg | 0.123765640 | -0.25059023 | 0.49812151 | 0.9953605 |
| Female:Young:Pos-Male:Juvenile:Neg | 0.065505685 | -0.09465816 | 0.22566953 | 0.9737561 |
| Male:Young:Pos-Male:Juvenile:Neg | 0.090177361 | -0.16888596 | 0.34924068 | 0.9927942 |
| Male:Old:Neg-Female:Old:Neg | 0.043186219 | -0.19778832 | 0.28416076 | 0.9999871 |
| Female:Young:Neg-Female:Old:Neg | 0.028793763 | -0.17586693 | 0.23345445 | 0.9999990 |
| Male:Young:Neg-Female:Old:Neg | 0.044424793 | -0.17051944 | 0.25936903 | 0.9999451 |
| Female:Juvenile:Pos-Female:Old:Neg | 0.010373713 | -0.24504516 | 0.26579259 | 1.0000000 |
| Male:Juvenile:Pos-Female:Old:Neg | 0.220288557 | -0.09721417 | 0.53779128 | 0.4970644 |
| Female:Old:Pos-Female:Old:Neg | -0.210656466 | -0.59027416 | 0.16896123 | 0.8083242 |
| Male:Old:Pos-Female:Old:Neg | 0.122685668 | -0.29010951 | 0.53548084 | 0.9981974 |
| Female:Young:Pos-Female:Old:Neg | 0.064425712 | -0.17202724 | 0.30087866 | 0.9991916 |
| Male:Young:Pos-Female:Old:Neg | 0.089097388 | -0.22294643 | 0.40114121 | 0.9987509 |
| Female:Young:Neg-Male:Old:Neg | -0.014392456 | -0.16291064 | 0.13412573 | 1.0000000 |
| Male:Young:Neg-Male:Old:Neg | 0.001238574 | -0.16115810 | 0.16363524 | 1.0000000 |
| Female:Juvenile:Pos-Male:Old:Neg | -0.032812506 | -0.24590980 | 0.18028478 | 0.9999973 |
| Male:Juvenile:Pos-Male:Old:Neg | 0.177102338 | -0.10746687 | 0.46167155 | 0.6668567 |
| Female:Old:Pos-Male:Old:Neg | -0.253842685 | -0.60637814 | 0.09869277 | 0.4356742 |
| Male:Old:Pos-Male:Old:Neg | 0.079499448 | -0.30853600 | 0.46753490 | 0.9999497 |
| Female:Young:Pos-Male:Old:Neg | 0.021239493 | -0.16871222 | 0.21119120 | 0.9999999 |
| Male:Young:Pos-Male:Old:Neg | 0.045911169 | -0.23255427 | 0.32437661 | 0.9999946 |
| Male:Young:Neg-Female:Young:Neg | 0.015631030 | -0.08531423 | 0.11657629 | 0.9999972 |
| Female:Juvenile:Pos-Female:Young:Neg | -0.018420050 | -0.18938122 | 0.15254112 | 0.9999999 |
| Male:Juvenile:Pos-Female:Young:Neg | 0.191494794 | -0.06305757 | 0.44604715 | 0.3640724 |
| Female:Old:Pos-Female:Young:Neg | -0.239450229 | -0.56823361 | 0.08933316 | 0.4168869 |
| Male:Old:Pos-Female:Young:Neg | 0.093891904 | -0.27269884 | 0.46048265 | 0.9995498 |
| Female:Young:Pos-Female:Young:Neg | 0.035631949 | -0.10543158 | 0.17669548 | 0.9996057 |
| Male:Young:Pos-Female:Young:Neg | 0.060303625 | -0.18740642 | 0.30801367 | 0.9997237 |
| Female:Juvenile:Pos-Male:Young:Neg | -0.034051080 | -0.21719788 | 0.14909572 | 0.9999812 |
| Male:Juvenile:Pos-Male:Young:Neg | 0.175863764 | -0.08702772 | 0.43875524 | 0.5565355 |
| Female:Old:Pos-Male:Young:Neg | -0.255081259 | -0.59036255 | 0.08020003 | 0.3460906 |
| Male:Old:Pos-Male:Young:Neg | 0.078260874 | -0.29416872 | 0.45069047 | 0.9999352 |
| Female:Young:Pos-Male:Young:Neg | 0.020000918 | -0.13560739 | 0.17560923 | 0.9999996 |
| Male:Young:Pos-Male:Young:Neg | 0.044672595 | -0.21159932 | 0.30094451 | 0.9999903 |
| Male:Juvenile:Pos-Female:Juvenile:Pos | 0.209914843 | -0.08698532 | 0.50681501 | 0.4658992 |
| Female:Old:Pos-Female:Juvenile:Pos | -0.221030179 | -0.58359234 | 0.14153198 | 0.6957509 |
| Male:Old:Pos-Female:Juvenile:Pos | 0.112311954 | -0.28485499 | 0.50947890 | 0.9988575 |
| Female:Young:Pos-Female:Juvenile:Pos | 0.054051998 | -0.15391849 | 0.26202249 | 0.9994823 |
| Male:Young:Pos-Female:Juvenile:Pos | 0.078723675 | -0.21233143 | 0.36977878 | 0.9992454 |
| Female:Old:Pos-Male:Juvenile:Pos | -0.430945023 | -0.83962644 | -0.02226361 | 0.0283653 |
| Male:Old:Pos-Male:Juvenile:Pos | -0.097602889 | -0.53727409 | 0.34206831 | 0.9998881 |
| Female:Young:Pos-Male:Juvenile:Pos | -0.155862845 | -0.43661345 | 0.12488776 | 0.8078687 |
| Male:Young:Pos-Male:Juvenile:Pos | -0.131191169 | -0.47800849 | 0.21562615 | 0.9857211 |
| Male:Old:Pos-Female:Old:Pos | 0.333342133 | -0.15308604 | 0.81977031 | 0.5173972 |
| Female:Young:Pos-Female:Old:Pos | 0.275082178 | -0.07437814 | 0.62454250 | 0.2933907 |
| Male:Young:Pos-Female:Old:Pos | 0.299753854 | -0.10470116 | 0.70420887 | 0.3882476 |
| Female:Young:Pos-Male:Old:Pos | -0.058259956 | -0.44350375 | 0.32698383 | 0.9999978 |
| Male:Young:Pos-Male:Old:Pos | -0.033588279 | -0.46933376 | 0.40215720 | 1.0000000 |
| Male:Young:Pos-Female:Young:Pos | 0.024671677 | -0.24989028 | 0.29923363 | 1.0000000 |

## B.6. The best fitting GLM model of association of *S. glossinidius* infection with *T. brucei* status and intrinsic factors of all tsetse samples (Model 3).

### Backward elimination of model selection

Explanatory variables: subpopulation, sex, age and *T. brucei* status
Response variables: *Sodalis* status

| Effect tested | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation ∗ sex ∗ age ∗ *T. brucei* status (full model) | 50 | 316.1 | 732.1 | - | - |
| subpopulation ∗ sex ∗ age ∗ *T. brucei* status | 47 | 316.1 | 726.1 | <0.0001 | 1 |
| sex ∗ age ∗ *T. brucei* status | 46 | 317.8 | 727.6 | 3.4712 | 0.0625 |
| subpopulation ∗ age ∗ *T. brucei* status | 41 | 321.6 | 725.3 | 7.7087 | 0.1730 |
| subpopulation ∗ sex ∗ *T. brucei* status | 37 | 322.9 | 719.9 | 2.5682 | 0.6325 |
| subpopulation ∗ sex ∗ age | 31 | 325.0 | 711.9 | 4.0700 | 0.6672 |
| age ∗ *T. brucei* status | 30 | 325.5 | 712.0 | 1.0615 | 0.3029 |
| sex ∗ *T. brucei* status | 29 | 325.6 | 709.1 | 0.2558 | 0.6131 |
| subpopulation ∗ *T. brucei* status | 23 | 326.1 | 698.3 | 1.0060 | 0.9854 |
| sex ∗ age | 22 | 326.3 | 696.5 | 0.2707 | 0.6028 |
| subpopulation ∗ age | 16 | 328.2 | 688.3 | 3.8209 | 0.7009 |
| subpopulation ∗ sex | 10 | 333.2 | 686.4 | 10.082 | 0.1212 |
| *T. brucei* status | 9 | 333.7 | 685.4 | 0.9614 | 0.3268 |
| age | 8 | 333.7 | 683.4 | 0.0183 | 0.8924 |
| sex | 7 | 333.7 | 681.4 | 0.0017 | 0.9674 |
| subpopulation | 1 | 1400 | 699.0 | 730.58 | **< 0.0001** |

A ∗ B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = Sodalis ~ Subpopulation, family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7749  -0.5968  -0.1170   0.2074   3.1582

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)         3.8286     0.5836   6.560 5.36e-11 ***
SubpopulationBRGp  -2.9752     0.6096  -4.881 1.06e-06 ***
SubpopulationMuGl -22.3947  1153.0506  -0.019    0.985
SubpopulationMuGp  -8.8088     0.9187  -9.588  < 2e-16 ***
SubpopulationSaGl -22.3947   856.4646  -0.026    0.979
SubpopulationZuGa  -5.4638     0.6054  -9.025  < 2e-16 ***
SubpopulationZuGp  -3.4551     0.6103  -5.662 1.50e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1397.99  on 1089  degrees of freedom
Residual deviance:  667.41  on 1083  degrees of freedom
AIC: 681.41

Number of Fisher Scoring iterations: 17
```

**ANOVA table**

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|---|---|---|---|---|---|
| Subpopulation | 6 | 137.7 | 22.942 | 232.1 | <2e-16 *** |
| Residuals | 1083 | 107.1 | 0.099 | | |

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

**Post hoc test** (Significant differences are indicated by shaded boxes.)

$Subpopulation

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp-BRGb | -2.774247e-01 | -0.38568186 | -0.169167550 | 0.0000000 |
| MuGl-BRGb | -9.787234e-01 | -1.16059045 | -0.796856357 | 0.0000000 |
| MuGp-BRGb | -9.718975e-01 | -1.06709309 | -0.876701837 | 0.0000000 |
| SaGl-BRGb | -9.787234e-01 | -1.12360676 | -0.833840049 | 0.0000000 |
| ZuGa-BRGb | -8.156028e-01 | -0.91139976 | -0.719805909 | 0.0000000 |
| ZuGp-BRGb | -3.864157e-01 | -0.49934821 | -0.273483217 | 0.0000000 |
| MuGl-BRGp | -7.012987e-01 | -0.88174029 | -0.520857117 | 0.0000000 |
| MuGp-BRGp | -6.944728e-01 | -0.78691599 | -0.602029535 | 0.0000000 |
| SaGl-BRGp | -7.012987e-01 | -0.84438864 | -0.558208767 | 0.0000000 |
| ZuGa-BRGp | -5.381781e-01 | -0.63124045 | -0.445115822 | 0.0000000 |
| ZuGp-BRGp | -1.089910e-01 | -0.21961330 | 0.001631281 | 0.0565661 |
| MuGp-MuGl | 6.825939e-03 | -0.16609532 | 0.179747198 | 0.9999998 |
| SaGl-MuGl | 1.443290e-15 | -0.20452565 | 0.204525648 | 1.0000000 |
| ZuGa-MuGl | 1.631206e-01 | -0.01013244 | 0.336373578 | 0.0804076 |
| ZuGp-MuGl | 5.923077e-01 | 0.40902293 | 0.775592451 | 0.0000000 |
| SaGl-MuGp | -6.825939e-03 | -0.14030749 | 0.126655617 | 0.9999990 |
| ZuGa-MuGp | 1.562946e-01 | 0.07881437 | 0.233774888 | 0.0000001 |
| ZuGp-MuGp | 5.854818e-01 | 0.48760485 | 0.683358654 | 0.0000000 |
| ZuGa-SaGl | 1.631206e-01 | 0.02920952 | 0.297031618 | 0.0061640 |
| ZuGp-SaGl | 5.923077e-01 | 0.44564868 | 0.738966706 | 0.0000000 |
| ZuGp-ZuGa | 4.291871e-01 | 0.33072530 | 0.527648951 | 0.0000000 |

# B.7. The best fitting GLM model of association of *S. glossinidius* infection with *T. vivax* status and intrinsic factors of all tsetse samples (Model 4).

## Backward elimination of model selection

Explanatory variables: subpopulation, sex, age and *T. vivax* status
Response variables: *Sodalis* status

| Effect tested | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation * sex * age * *T. vivax* status (full model) | 54 | 319.1 | 746.3 | - | - |
| subpopulation * sex * age * *T. vivax* status | 50 | 320.6 | 741.1 | 2.8786 | 0.5783 |
| sex * age * *T. vivax* status | 49 | 320.6 | 739.2 | 0.00286 | 0.9574 |
| subpopulation * age * *T. vivax* status | 43 | 321.3 | 728.6 | 1.4238 | 0.9644 |
| subpopulation * sex * *T. vivax* status | 37 | 321.4 | 716.8 | 0.26888 | 0.9996 |
| subpopulation * sex * age | 31 | 323.8 | 709.6 | 4.7168 | 0.5806 |
| age * *T. vivax* status | 30 | 324.5 | 709.0 | 1.4494 | 0.2286 |
| sex * *T. vivax* status | 29 | 324.7 | 707.4 | 0.3652 | 0.5456 |
| subpopulation * *T. vivax* status | 23 | 326.2 | 698.3 | 2.9731 | 0.8122 |
| sex * age | 22 | 326.3 | 696.7 | 0.3394 | 0.5602 |
| subpopulation * age | 16 | 328.3 | 688.5 | 3.8160 | 0.7016 |
| subpopulation * sex | 10 | 333.4 | 686.7 | 10.2340 | 0.1151 |
| *T. vivax* status | 9 | 333.7 | 685.4 | 0.6562 | 0.4179 |
| age | 8 | 333.7 | 683.4 | 0.0183 | 0.8924 |
| sex | 7 | 333.7 | 681.4 | 0.0017 | 0.9674 |
| subpopulation | 1 | 699.0 | 1400 | 730.6 | < 0.0001 |

A * B means an interaction between A and B.

## Summary of the best fitting model

```
Call:
glm(formula = Sodalis ~ Subpopulation, family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7749  -0.5968  -0.1170   0.2074   3.1582

Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)          3.8286     0.5836   6.560 5.36e-11 ***
SubpopulationBRGp   -2.9752     0.6096  -4.881 1.06e-06 ***
SubpopulationMuGl  -22.3947  1153.0506  -0.019    0.985
SubpopulationMuGp   -8.8088     0.9187  -9.588  < 2e-16 ***
SubpopulationSaGl  -22.3947   856.4646  -0.026    0.979
SubpopulationZuGa   -5.4638     0.6054  -9.025  < 2e-16 ***
SubpopulationZuGp   -3.4551     0.6103  -5.662 1.50e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1397.99  on 1089  degrees of freedom
Residual deviance:  667.41  on 1083  degrees of freedom
AIC: 681.41

Number of Fisher Scoring iterations: 17
```

**ANOVA table**

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| Subpopulation | 6 | 137.7 | 22.942 | 232.1 | <2e-16 | *** |
| Residuals | 1083 | 107.1 | 0.099 | | | |

**Post hoc test** (Significant differences are indicated by shaded boxes.)

$Subpopulation

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp-BRGb | -2.774247e-01 | -0.38568186 | -0.169167550 | 0.0000000 |
| MuGl-BRGb | -9.787234e-01 | -1.16059045 | -0.796856357 | 0.0000000 |
| MuGp-BRGb | -9.718975e-01 | -1.06709309 | -0.876701837 | 0.0000000 |
| SaGl-BRGb | -9.787234e-01 | -1.12360676 | -0.833840049 | 0.0000000 |
| ZuGa-BRGb | -8.156028e-01 | -0.91139976 | -0.719805909 | 0.0000000 |
| ZuGp-BRGb | -3.864157e-01 | -0.49934821 | -0.273483217 | 0.0000000 |
| MuGl-BRGp | -7.012987e-01 | -0.88174029 | -0.520857117 | 0.0000000 |
| MuGp-BRGp | -6.944728e-01 | -0.78691599 | -0.602029535 | 0.0000000 |
| SaGl-BRGp | -7.012987e-01 | -0.84438864 | -0.558208767 | 0.0000000 |
| ZuGa-BRGp | -5.381781e-01 | -0.63124045 | -0.445115822 | 0.0000000 |
| ZuGp-BRGp | -1.089910e-01 | -0.21961330 | 0.001631281 | 0.0565661 |
| MuGp-MuGl | 6.825939e-03 | -0.16609532 | 0.179747198 | 0.9999998 |
| SaGl-MuGl | 1.443290e-15 | -0.20452565 | 0.204525648 | 1.0000000 |
| ZuGa-MuGl | 1.631206e-01 | -0.01013244 | 0.336373578 | 0.0804076 |
| ZuGp-MuGl | 5.923077e-01 | 0.40902293 | 0.775592451 | 0.0000000 |
| SaGl-MuGp | -6.825939e-03 | -0.14030749 | 0.126655617 | 0.9999990 |
| ZuGa-MuGp | 1.562946e-01 | 0.07881437 | 0.233774888 | 0.0000001 |
| ZuGp-MuGp | 5.854818e-01 | 0.48760485 | 0.683358654 | 0.0000000 |
| ZuGa-SaGl | 1.631206e-01 | 0.02920952 | 0.297031618 | 0.0061640 |
| ZuGp-SaGl | 5.923077e-01 | 0.44564868 | 0.738966706 | 0.0000000 |
| ZuGp-ZuGa | 4.291871e-01 | 0.33072530 | 0.527648951 | 0.0000000 |

## B.8. The best fitting GLM Model of trypanosome status in all tsetse fly samples (GLM Model 5).

### Backward elimination of model selection

Explanatory variables: subpopulation, sex, age and *Sodalis* status
Response variables: trypanosome status

| Effect tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation ∗ sex ∗ age ∗ *Sodalis* status (full model) | 44 | 688.0 | 1464.0 | - | - |
| subpopulation ∗ sex ∗ age ∗ *Sodalis* status | 42 | 690.2 | 1464.4 | 4.3473 | 0.1138 |
| sex ∗ age ∗ *Sodalis* status | 41 | 690.8 | 1463.6 | 1.2221 | 0.2690 |
| subpopulation ∗ age ∗*Sodalis* status | 38 | 693.1 | 1462.2 | 4.6263 | 0.2013 |
| subpopulation ∗ sex ∗ *Sodalis* status | 35 | 693.3 | 1456.6 | 0.3483 | 0.9507 |
| subpopulation ∗ sex ∗ age | 29 | 695.3 | 1448.6 | 3.9863 | 0.6785 |
| age ∗ *Sodalis* status | 28 | 695.4 | 1446.8 | 0.2116 | 0.6455 |
| sex ∗ *Sodalis* status | 27 | 695.7 | 1445.5 | 0.7060 | 0.4008 |
| subpopulation ∗ *Sodalis* status | 23 | 697.8 | 1441.6 | 4.1269 | 0.3891 |
| sex ∗ age | 22 | 698.4 | 1440.8 | 1.2188 | 0.2696 |
| subpopulation ∗ age | 16 | 701.0 | 1434.0 | 5.2116 | 0.5170 |
| subpopulation ∗ sex | 10 | 710.5 | 1441.0 | 18.944 | **0.0043** |
| *Sodalis* status | 15 | 701.3 | 1432.6 | 0.5920 | 0.4417 |
| age | 14 | 703.6 | 1435.3 | 4.6528 | **0.0310** |

A ∗ B means an interaction between A and B.

**Summary of the best fitting model**

```
Call:
glm(formula = Trypanosome_status ~ Subpopulation * Sex + Age_class,
    family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5220  -1.0112  -0.7421   1.1611   2.3830

Coefficients:
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                -0.92503    0.26411  -3.502 0.000461 ***
SubpopulationBRGp          -0.45633    0.32022  -1.425 0.154138
SubpopulationMuGl           0.38714    0.53696   0.721 0.470912
SubpopulationMuGp           0.30452    0.27247   1.118 0.263722
SubpopulationSaGl          -2.07024    1.05930  -1.954 0.050659 .
SubpopulationZuGa           0.74751    0.25401   2.943 0.003253 **
SubpopulationZuGp           0.97091    0.30298   3.205 0.001353 **
SexMale                    -0.14819    0.37424  -0.396 0.692125
Age_class                   0.21606    0.10020   2.156 0.031053 *
SubpopulationBRGp:SexMale   1.11534    0.51312   2.174 0.029731 *
SubpopulationMuGl:SexMale  -1.11636    0.89267  -1.251 0.211088
SubpopulationMuGp:SexMale  -0.47648    0.44755  -1.065 0.287044
SubpopulationSaGl:SexMale   1.56255    1.15890   1.348 0.177560
SubpopulationZuGa:SexMale   0.09416    0.53043   0.178 0.859098
SubpopulationZuGp:SexMale   0.23548    0.55517   0.424 0.671456
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1483.8  on 1089  degrees of freedom
Residual deviance: 1402.6  on 1075  degrees of freedom
AIC: 1432.6

Number of Fisher Scoring iterations: 5
```

**ANOVA table**

```
                    Df Sum Sq Mean Sq F value   Pr(>F)
Subpopulation        6  13.85  2.3089  10.057 7.83e-11 ***
Sex                  1   0.00  0.0043   0.019   0.8915
Age_class            1   1.04  1.0362   4.513   0.0339 *
Subpopulation:Sex    6   4.01  0.6690   2.914   0.0080 **
Residuals         1075 246.81  0.2296
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Post hoc test** (Significant differences are indicated by shaded boxes.)

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Tryp2, data = TK, projections = FALSE, qr = TRUE, contrasts = NULL)

$Subpopulation
                     diff         lwr         upr     p adj
BRGp-BRGb -0.0033618863 -0.16841292 0.16168914 1.0000000
MuGl-BRGb -0.0037677305 -0.28104589 0.27351042 1.0000000
MuGp-BRGb  0.0142570135 -0.13088016 0.15939418 0.9999515
SaGl-BRGb -0.1578625581 -0.37875463 0.06302952 0.3467402
ZuGa-BRGb  0.1879432624  0.04188934 0.33399718 0.0028906
ZuGp-BRGb  0.2524822695  0.08030312 0.42466142 0.0003246
MuGl-BRGp -0.0004058442 -0.27551071 0.27469902 1.0000000
MuGp-BRGp  0.0176188999 -0.12332190 0.15855970 0.9998007
SaGl-BRGp -0.1545006717 -0.37265846 0.06365712 0.3580928
ZuGa-BRGp  0.1913051488  0.04942048 0.33318982 0.0014127
ZuGp-BRGp  0.2558441558  0.08718719 0.42450112 0.0001665
MuGp-MuGl  0.0180247440 -0.24561448 0.28166397 0.9999943
SaGl-MuGl -0.1540948276 -0.46591874 0.15772909 0.7686407
ZuGa-MuGl  0.1917109929 -0.07243403 0.45585601 0.3276039
ZuGp-MuGl  0.2562500000 -0.02318963 0.53568963 0.0969923
SaGl-MuGp -0.1721195716 -0.37562823 0.03138909 0.1608024
ZuGa-MuGp  0.1736862489  0.05555829 0.29181421 0.0003084
ZuGp-MuGp  0.2382252560  0.08900017 0.38745034 0.0000558
ZuGa-SaGl  0.3458058205  0.14164234 0.54996930 0.0000136
ZuGp-SaGl  0.4103448276  0.18674555 0.63394411 0.0000015
ZuGp-ZuGa  0.0645390071 -0.08557787 0.21465589 0.8656008

$Sex
                   diff         lwr        upr     p adj
Male-Female -0.003938991 -0.06434895 0.05647097 0.8982186

$Age
                      diff         lwr        upr     p adj
Old-Juvenile     0.03702955 -0.08854137 0.16260046 0.7681637
Young-Juvenile -0.04984618 -0.12251840 0.02282604 0.2418863
Young-Old      -0.08687572 -0.20865079 0.03489934 0.2155418
```

$`Subpopulation:Sex`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp:Female-BRGb:Female | -0.0966311326 | -0.3317942787 | 0.13853201 | 0.9839476 |
| MuGl:Female-BRGb:Female | 0.1076257510 | -0.3181330373 | 0.53338454 | 0.9998870 |
| MuGp:Female-BRGb:Female | 0.0653482477 | -0.1462212591 | 0.27691775 | 0.9989952 |
| SaGl:Female-BRGb:Female | -0.2906297917 | -0.7396682182 | 0.15840863 | 0.6475491 |
| ZuGa:Female-BRGb:Female | 0.1805944181 | -0.0166857319 | 0.37787457 | 0.1140818 |
| ZuGp:Female-BRGb:Female | 0.2364496140 | 0.0006943542 | 0.47220487 | 0.0484809 |
| BRGb:Male-BRGb:Female | -0.0205932618 | -0.3042549991 | 0.26306848 | 1.0000000 |
| BRGp:Male-BRGb:Female | 0.1356296983 | -0.1365587151 | 0.40781811 | 0.9244216 |
| MuGl:Male-BRGb:Female | -0.1455925031 | -0.5946309296 | 0.30344592 | 0.9983379 |
| MuGp:Male-BRGb:Female | -0.0632937532 | -0.2825108354 | 0.15592333 | 0.9995098 |
| SaGl:Male-BRGb:Female | -0.1213983918 | -0.4195892849 | 0.17679250 | 0.9852110 |
| ZuGa:Male-BRGb:Female | 0.1810021656 | -0.1501699563 | 0.51217429 | 0.8588120 |
| ZuGp:Male-BRGb:Female | 0.2698293330 | -0.0540563509 | 0.59371502 | 0.2257125 |
| MuGl:Female-BRGp:Female | 0.2042568837 | -0.2194241258 | 0.62793789 | 0.9407256 |
| MuGp:Female-BRGp:Female | 0.1619793803 | -0.0453770957 | 0.36933586 | 0.3250153 |
| SaGl:Female-BRGp:Female | -0.1939986591 | -0.6410675133 | 0.25307020 | 0.9743480 |
| ZuGa:Female-BRGp:Female | 0.2772255507 | 0.0844705001 | 0.46998060 | 0.0001329 |
| ZuGp:Female-BRGp:Female | 0.3330807467 | 0.1010988626 | 0.56506263 | 0.0001382 |
| BRGb:Male-BRGp:Female | 0.0760378708 | -0.2044956088 | 0.35657135 | 0.9997540 |
| BRGp:Male-BRGp:Female | 0.2322608309 | -0.0366658959 | 0.50118756 | 0.1777755 |
| MuGl:Male-BRGp:Female | -0.0489613705 | -0.4960302247 | 0.39810748 | 1.0000000 |
| MuGp:Male-BRGp:Female | 0.0333373794 | -0.1818164750 | 0.24849123 | 0.9999997 |
| SaGl:Male-BRGp:Female | -0.0247672592 | -0.3199838926 | 0.27044937 | 1.0000000 |
| ZuGa:Male-BRGp:Female | 0.2776332983 | -0.0508633171 | 0.60612991 | 0.2058442 |
| ZuGp:Male-BRGp:Female | 0.3664604657 | 0.0453109861 | 0.68760995 | 0.0098280 |
| MuGp:Female-MuGl:Female | -0.0422775033 | -0.4533316953 | 0.36877669 | 1.0000000 |
| SaGl:Female-MuGl:Female | -0.3982555428 | -0.9690789839 | 0.17256790 | 0.5214839 |
| ZuGa:Female-MuGl:Female | 0.0729686670 | -0.3309166510 | 0.47685399 | 0.9999978 |
| ZuGp:Female-MuGl:Female | 0.1288238630 | -0.2951860839 | 0.55283381 | 0.9991553 |
| BRGb:Male-MuGl:Female | -0.1282190129 | -0.5806213490 | 0.32418332 | 0.9995990 |
| BRGp:Male-MuGl:Female | 0.0280039472 | -0.4172941832 | 0.47330208 | 1.0000000 |
| MuGl:Male-MuGl:Female | -0.2532182542 | -0.8240416953 | 0.31760519 | 0.9691989 |
| MuGp:Male-MuGl:Female | -0.1709195043 | -0.5859616969 | 0.24412269 | 0.9836353 |
| SaGl:Male-MuGl:Female | -0.2290241429 | -0.6906752029 | 0.23262692 | 0.9267945 |

$`Subpopulation:Sex`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| ZuGa:Male-MuGl:Female | 0.0733764146 | -0.4102338282 | 0.55698666 | 0.9999997 |
| ZuGp:Male-MuGl:Female | 0.1622035820 | -0.3164464246 | 0.64085359 | 0.9973986 |
| SaGl:Female-MuGp:Female | -0.3559780395 | -0.7910992999 | 0.07914322 | 0.2521128 |
| ZuGa:Female-MuGp:Female | 0.1152461704 | -0.0478932193 | 0.27838556 | 0.4996231 |
| ZuGp:Female-MuGp:Female | 0.1711013663 | -0.0369263853 | 0.37912912 | 0.2442015 |
| BRGb:Male-MuGp:Female | -0.0859415095 | -0.3470142658 | 0.17513125 | 0.9980594 |
| BRGp:Male-MuGp:Female | 0.0702814505 | -0.1782774666 | 0.31884037 | 0.9996090 |
| MuGl:Male-MuGp:Female | -0.2109407509 | -0.6460620113 | 0.22418051 | 0.9382107 |
| MuGp:Male-MuGp:Female | -0.1286420009 | -0.3177216984 | 0.06043770 | 0.5650495 |
| SaGl:Male-MuGp:Female | -0.1867466396 | -0.4635368337 | 0.09004355 | 0.5793092 |
| ZuGa:Male-MuGp:Female | 0.1156539179 | -0.1963877095 | 0.42769555 | 0.9937600 |
| ZuGp:Male-MuGp:Female | 0.2044810853 | -0.0998163664 | 0.50877854 | 0.5861096 |
| ZuGa:Female-SaGl:Female | 0.4712242098 | 0.0428688519 | 0.89957957 | 0.0162560 |
| ZuGp:Female-SaGl:Female | 0.5270794058 | 0.0796988098 | 0.97446000 | 0.0061124 |
| BRGb:Male-SaGl:Female | 0.2700365299 | -0.2043397303 | 0.74441279 | 0.8190766 |
| BRGp:Male-SaGl:Female | 0.4262594900 | -0.0413465274 | 0.89386551 | 0.1180960 |
| MuGl:Male-SaGl:Female | 0.1450372886 | -0.4433540468 | 0.73342862 | 0.9999150 |
| MuGp:Male-SaGl:Female | 0.2273360385 | -0.2115545905 | 0.66622667 | 0.9007927 |
| SaGl:Male-SaGl:Female | 0.1692313999 | -0.3139731777 | 0.65243598 | 0.9963918 |
| ZuGa:Male-SaGl:Female | 0.4716319574 | -0.0325940133 | 0.97585793 | 0.0949514 |
| ZuGp:Male-SaGl:Female | 0.5604591248 | 0.0609886138 | 1.05992964 | 0.0124427 |
| ZuGp:Female-ZuGa:Female | 0.0558551960 | -0.1376217970 | 0.24933219 | 0.9995104 |
| BRGb:Male-ZuGa:Female | -0.2011876799 | -0.4508209818 | 0.04844562 | 0.2749425 |
| BRGp:Male-ZuGa:Female | -0.0449647198 | -0.2814797016 | 0.19155026 | 0.9999959 |
| MuGl:Male-ZuGa:Female | -0.3261869212 | -0.7545422791 | 0.10216844 | 0.3679992 |
| MuGp:Male-ZuGa:Female | -0.2438881713 | -0.4168301655 | -0.07094618 | 0.0002086 |
| SaGl:Male-ZuGa:Female | -0.3019928099 | -0.5680202764 | -0.03596534 | 0.0105808 |
| ZuGa:Male-ZuGa:Female | 0.0004077476 | -0.3021278348 | 0.30294333 | 1.0000000 |
| ZuGp:Male-ZuGa:Female | 0.0892349150 | -0.2053066619 | 0.38377649 | 0.9991803 |
| BRGb:Male-ZuGp:Female | -0.2570428759 | -0.5380728928 | 0.02398714 | 0.1148832 |
| BRGp:Male-ZuGp:Female | -0.1008199158 | -0.3702645699 | 0.16862474 | 0.9931721 |
| MuGl:Male-ZuGp:Female | -0.3820421172 | -0.8294227132 | 0.06533848 | 0.1921441 |
| MuGp:Male-ZuGp:Female | -0.2997433673 | -0.5155442439 | -0.08394249 | 0.0002921 |
| SaGl:Male-ZuGp:Female | -0.3578480059 | -0.6535365209 | -0.06215949 | 0.0039917 |
| ZuGa:Male-ZuGp:Female | -0.0554474484 | -0.3843682042 | 0.27347331 | 0.9999991 |
| ZuGp:Male-ZuGp:Female | 0.0333797190 | -0.2882035914 | 0.35496303 | 1.0000000 |
| BRGp:Male-BRGb:Male | 0.1562229601 | -0.1560010945 | 0.46844701 | 0.9221458 |
| MuGl:Male-BRGb:Male | -0.1249992413 | -0.5993755016 | 0.34937702 | 0.9998201 |
| MuGp:Male-BRGb:Male | -0.0427004914 | -0.3100082859 | 0.22460730 | 0.9999995 |
| SaGl:Male-BRGb:Male | -0.1008051300 | -0.4359396352 | 0.23432938 | 0.9992404 |
| ZuGa:Male-BRGb:Male | 0.2015954275 | -0.1631953375 | 0.56638619 | 0.8487036 |
| ZuGp:Male-BRGb:Male | 0.2904225949 | -0.0677662717 | 0.64861146 | 0.2656572 |
| MuGl:Male-BRGp:Male | -0.2812222014 | -0.7488282188 | 0.18638382 | 0.7539708 |
| MuGp:Male-BRGp:Male | -0.1989234515 | -0.4540234479 | 0.05617654 | 0.3279368 |
| SaGl:Male-BRGp:Male | -0.2570280901 | -0.5825087926 | 0.06845261 | 0.3072509 |
| ZuGa:Male-BRGp:Male | 0.0453724674 | -0.3105697378 | 0.40131467 | 1.0000000 |
| ZuGp:Male-BRGp:Male | 0.1341996348 | -0.2149734102 | 0.48337268 | 0.9912279 |
| MuGp:Male-MuGl:Male | 0.0822987499 | -0.3565918791 | 0.52118938 | 0.9999966 |
| SaGl:Male-MuGl:Male | 0.0241941113 | -0.4590104662 | 0.50739869 | 1.0000000 |
| ZuGa:Male-MuGl:Male | 0.3265946688 | -0.1776313018 | 0.83082064 | 0.6463555 |
| ZuGp:Male-MuGl:Male | 0.4154218362 | -0.0840486748 | 0.91489235 | 0.2280574 |
| SaGl:Male-MuGp:Male | -0.0581046386 | -0.3407834041 | 0.22457413 | 0.9999897 |
| ZuGa:Male-MuGp:Male | 0.2442959189 | -0.0729806957 | 0.56157253 | 0.3491272 |
| ZuGp:Male-MuGp:Male | 0.3331230863 | 0.0234597012 | 0.64278647 | 0.0216585 |
| ZuGa:Male-SaGl:Male | 0.3024005575 | -0.0737990301 | 0.67860015 | 0.2790066 |
| ZuGp:Male-SaGl:Male | 0.3912277249 | 0.0214263030 | 0.76102915 | 0.0265779 |
| ZuGp:Male-ZuGa:Male | 0.0888271674 | -0.3080484520 | 0.48570279 | 0.9999720 |

## B.9. The best fitting GLM model of association of *T. vivax* status with *S. glossinidius* infection and intrinsic factors of all tsetse samples (Model 8).

### Backward elimination of model selection

Explanatory variables: subpopulation, sex, age and *Sodalis* status
Response variables: *T. vivax* status

| Effect tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation * sex *age * *Sodalis* status (full model) | 44 | 562.1 | 1212.2 | - | - |
| subpopulation * sex *age * *Sodalis* status | 42 | 563.4 | 1210.9 | 2.6277 | 0.2688 |
| sex * age * *Sodalis* status | 41 | 563.5 | 1209.0 | 0.1458 | 0.7026 |
| subpopulation * age * *Sodalis* status | 38 | 564.0 | 1204.0 | 1.0471 | 0.7899 |
| subpopulation * sex * *Sodalis* status | 35 | 564.2 | 1198.3 | 0.2867 | 0.9625 |
| subpopulation * sex * age | 29 | 566.1 | 1190.3 | 3.9191 | 0.6876 |
| age * *Sodalis* status | 28 | 566.8 | 1189.6 | 1.3343 | 0.2480 |
| sex * *Sodalis* status | 27 | 567.0 | 1187.9 | 0.3566 | 0.5504 |
| subpopulation * *Sodalis* status | 23 | 568.7 | 1183.4 | 3.3182 | 0.5904 |
| sex * age | 22 | 572.5 | 1189.0 | 7.5921 | **0.0059** |
| subpopulation * age | 17 | 571.7 | 1177.4 | 6.0173 | 0.4213 |
| subpopulation * sex | 11 | 581.1 | 1184.3 | 18.903 | **0.0043** |
| *Sodalis* status | 16 | 572.1 | 1176.3 | 0.8712 | 0.3506 |

A * B means an interaction between A and B.

## Summary of the best fitting model

```
Call:
glm(formula = TV_ITS1 ~ Subpopulation + Sex + Age_class + Subpopulation:Sex +
    Sex:Age_class, family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5732  -0.7956  -0.5213   1.0373   2.5622

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.8234     0.3850  -4.736 2.18e-06 ***
SubpopulationBRGp          0.1217     0.4386   0.277  0.78145
SubpopulationMuGl          1.4352     0.6061   2.368  0.01788 *
SubpopulationMuGp          0.9092     0.3693   2.462  0.01382 *
SubpopulationSaGl         -0.6380     1.0843  -0.588  0.55628
SubpopulationZuGa          1.3808     0.3480   3.968 7.26e-05 ***
SubpopulationZuGp          1.7468     0.3824   4.568 4.93e-06 ***
SexMale                   -1.4816     0.7222  -2.052  0.04021 *
Age_class                 -0.1128     0.1499  -0.752  0.45190
SubpopulationBRGp:SexMale  1.4858     0.7087   2.096  0.03605 *
SubpopulationMuGl:SexMale -1.9303     1.2978  -1.487  0.13692
SubpopulationMuGp:SexMale -0.2603     0.6494  -0.401  0.68856
SubpopulationSaGl:SexMale  0.6310     1.2964   0.487  0.62646
SubpopulationZuGa:SexMale  0.3312     0.7041   0.470  0.63805
SubpopulationZuGp:SexMale  0.7852     0.7096   1.106  0.26851
SexMale:Age_class          0.6687     0.2364   2.829  0.00467 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1244.2  on 1089  degrees of freedom
Residual deviance: 1144.3  on 1074  degrees of freedom
AIC: 1176.3

Number of Fisher Scoring iterations: 5
```

## ANOVA table

```
                    Df Sum Sq Mean Sq F value   Pr(>F)
Subpopulation        6  13.55  2.2588  12.740 6.22e-14 ***
Sex                  1   0.00  0.0016   0.009  0.92544
Age_class            1   0.31  0.3097   1.747  0.18656
Subpopulation:Sex    6   3.02  0.5039   2.842  0.00947 **
Sex:Age_class        1   1.25  1.2519   7.061  0.00800 **
Residuals         1074 190.42  0.1773
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Post hoc test** (Significant differences are indicated by shaded boxes.)

$Subpopulation

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp-BRGb | 0.094317030 | -0.050690546 | 0.23932461 | 0.4664914 |
| MuGl-BRGb | 0.105274823 | -0.138331249 | 0.34888089 | 0.8626969 |
| MuGp-BRGb | 0.098128918 | -0.029383095 | 0.22564093 | 0.2579935 |
| SaGl-BRGb | -0.027268281 | -0.221335681 | 0.16679912 | 0.9996073 |
| ZuGa-BRGb | 0.237588652 | 0.109271216 | 0.36590609 | 0.0000012 |
| ZuGp-BRGb | 0.348063284 | 0.196793211 | 0.49933336 | 0.0000000 |
| MuGl-BRGp | 0.010957792 | -0.230738910 | 0.25265449 | 0.9999995 |
| MuGp-BRGp | 0.003811888 | -0.120013357 | 0.12763713 | 1.0000000 |
| SaGl-BRGp | -0.121585311 | -0.313250471 | 0.07007985 | 0.4982314 |
| ZuGa-BRGp | 0.143271622 | 0.018617130 | 0.26792611 | 0.0125709 |
| ZuGp-BRGp | 0.253746254 | 0.105570639 | 0.40192187 | 0.0000103 |
| MuGp-MuGl | -0.007145904 | -0.238769332 | 0.22447752 | 1.0000000 |
| SaGl-MuGl | -0.132543103 | -0.406499769 | 0.14141356 | 0.7860057 |
| ZuGa-MuGl | 0.132313830 | -0.099753969 | 0.36438163 | 0.6268385 |
| ZuGp-MuGl | 0.242788462 | -0.002716597 | 0.48829352 | 0.0548625 |
| SaGl-MuGp | -0.125397199 | -0.304192193 | 0.05339779 | 0.3703479 |
| ZuGa-MuGp | 0.139459734 | 0.035676987 | 0.24324248 | 0.0014914 |
| ZuGp-MuGp | 0.249934366 | 0.118830862 | 0.38103787 | 0.0000005 |
| ZuGa-SaGl | 0.264856933 | 0.085486643 | 0.44422722 | 0.0002842 |
| ZuGp-SaGl | 0.375331565 | 0.178885717 | 0.57177741 | 0.0000004 |
| ZuGp-ZuGa | 0.110474632 | -0.021412366 | 0.24236163 | 0.1696779 |

$Sex

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| Male-Female | -0.0023748 | -0.05544863 | 0.05069902 | 0.9300535 |

$Age

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| Old-Juvenile | 0.02491397 | -0.08540777 | 0.13523571 | 0.8566179 |
| Young-Juvenile | -0.02517579 | -0.08902279 | 0.03867120 | 0.6242969 |
| Young-Old | -0.05008976 | -0.15707661 | 0.05689709 | 0.5149996 |

$`Subpopulation:Sex`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| BRGp:Female-BRGb:Female | 0.011654016 | -0.194951567 | 0.21825960 | 1.0000000 |
| MuGl:Female-BRGb:Female | 0.226394508 | -0.147661308 | 0.60045032 | 0.7455960 |
| MuGp:Female-BRGb:Female | 0.134394794 | -0.051482296 | 0.32027188 | 0.4585978 |
| SaGl:Female-BRGb:Female | -0.057816462 | -0.452324901 | 0.33669198 | 0.9999998 |
| ZuGa:Female-BRGb:Female | 0.229163128 | 0.055840135 | 0.40248612 | 0.0008128 |
| ZuGp:Female-BRGb:Female | 0.313616112 | 0.106490320 | 0.52074190 | 0.0000374 |
| BRGb:Male-BRGb:Female | -0.025652003 | -0.274866642 | 0.22356264 | 1.0000000 |
| BRGp:Male-BRGb:Female | 0.210412791 | -0.028721813 | 0.44954739 | 0.1554935 |
| MuGl:Male-BRGb:Female | -0.051399925 | -0.445908364 | 0.34310851 | 1.0000000 |
| MuGp:Male-BRGb:Female | 0.034461238 | -0.158134727 | 0.22705720 | 0.9999980 |
| SaGl:Male-BRGb:Female | -0.028881575 | -0.290860989 | 0.23309784 | 1.0000000 |
| ZuGa:Male-BRGb:Female | 0.223250556 | -0.067704937 | 0.51420605 | 0.3550194 |
| ZuGp:Male-BRGb:Female | 0.410545326 | 0.125991426 | 0.69509923 | 0.0001234 |
| MuGl:Female-BRGp:Female | 0.214740492 | -0.157489865 | 0.58697085 | 0.8044570 |
| MuGp:Female-BRGp:Female | 0.122740778 | -0.059434900 | 0.30491646 | 0.5816557 |
| SaGl:Female-BRGp:Female | -0.069470478 | -0.462248524 | 0.32330757 | 0.9999983 |
| ZuGa:Female-BRGp:Female | 0.217509112 | 0.048161702 | 0.38685652 | 0.0014155 |
| ZuGp:Female-BRGp:Female | 0.301962096 | 0.098151451 | 0.50577274 | 0.0000649 |
| BRGb:Male-BRGp:Female | -0.037306019 | -0.283772287 | 0.20916025 | 0.9999997 |
| BRGp:Male-BRGp:Female | 0.198758775 | -0.037510233 | 0.43502778 | 0.2121691 |
| MuGl:Male-BRGp:Female | -0.063053941 | -0.455831988 | 0.32972411 | 0.9999995 |
| MuGp:Male-BRGp:Female | 0.022807222 | -0.166218942 | 0.21183339 | 1.0000000 |
| SaGl:Male-BRGp:Female | -0.040535592 | -0.299901932 | 0.21883075 | 0.9999996 |
| ZuGa:Male-BRGp:Female | 0.211596540 | -0.077008353 | 0.50020143 | 0.4340104 |
| ZuGp:Male-BRGp:Female | 0.398891310 | 0.116741337 | 0.68104128 | 0.0001972 |

$`Subpopulation:Sex`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| MuGp:Female-MuGl:Female | -0.091999714 | -0.453136620 | 0.26913719 | 0.9998767 |
| SaGl:Female-MuGl:Female | -0.284210970 | -0.785715188 | 0.21729325 | 0.8237669 |
| ZuGa:Female-MuGl:Female | 0.002768619 | -0.352069980 | 0.35760722 | 1.0000000 |
| ZuGp:Female-MuGl:Female | 0.087221603 | -0.285297746 | 0.45974095 | 0.9999528 |
| BRGb:Male-MuGl:Female | -0.252046512 | -0.649510356 | 0.14541733 | 0.6791845 |
| BRGp:Male-MuGl:Female | -0.015981718 | -0.407204072 | 0.37524064 | 1.0000000 |
| MuGl:Male-MuGl:Female | -0.277794434 | -0.779298651 | 0.22370978 | 0.8465040 |
| MuGp:Male-MuGl:Female | -0.191933270 | -0.556573884 | 0.17270734 | 0.8894171 |
| SaGl:Male-MuGl:Female | -0.255276084 | -0.660865513 | 0.15031335 | 0.6904398 |
| ZuGa:Male-MuGl:Female | -0.003143952 | -0.428025902 | 0.42173800 | 1.0000000 |
| ZuGp:Male-MuGl:Female | 0.184150818 | -0.236373253 | 0.60467489 | 0.9723169 |
| SaGl:Female-MuGp:Female | -0.192211256 | -0.574492591 | 0.19007008 | 0.9193837 |
| ZuGa:Female-MuGp:Female | 0.094768333 | -0.048559859 | 0.23809653 | 0.6130498 |
| ZuGp:Female-MuGp:Female | 0.179221317 | -0.003544118 | 0.36198675 | 0.0610480 |
| BRGb:Male-MuGp:Female | -0.160046797 | -0.389415598 | 0.06932200 | 0.5212654 |
| BRGp:Male-MuGp:Female | 0.076017996 | -0.142356612 | 0.29439260 | 0.9966010 |
| MuGl:Male-MuGp:Female | -0.185794719 | -0.568076055 | 0.19648662 | 0.9370317 |
| MuGp:Male-MuGp:Female | -0.099933556 | -0.266051937 | 0.06618482 | 0.7535917 |
| SaGl:Male-MuGp:Female | -0.163276370 | -0.406453926 | 0.07990119 | 0.5874800 |
| ZuGa:Male-MuGp:Female | 0.088855762 | -0.185292391 | 0.36300392 | 0.9983438 |
| ZuGp:Male-MuGp:Female | 0.276150532 | 0.008806123 | 0.54349494 | 0.0351337 |
| ZuGa:Female-SaGl:Female | 0.286979590 | -0.089357476 | 0.66331665 | 0.3655828 |
| ZuGp:Female-SaGl:Female | 0.371432574 | -0.021619358 | 0.76448450 | 0.0866755 |
| BRGb:Male-SaGl:Female | 0.032164459 | -0.384604858 | 0.44893378 | 1.0000000 |
| BRGp:Male-SaGl:Female | 0.268229253 | -0.142591981 | 0.67905049 | 0.6335866 |
| MuGl:Male-SaGl:Female | 0.006416537 | -0.510522179 | 0.52335525 | 1.0000000 |
| MuGp:Male-SaGl:Female | 0.092277700 | -0.293315262 | 0.47787066 | 0.9999394 |
| SaGl:Male-SaGl:Female | 0.028934886 | -0.395590661 | 0.45346043 | 1.0000000 |
| ZuGa:Male-SaGl:Female | 0.281067018 | -0.161927142 | 0.72406118 | 0.6783867 |
| ZuGp:Male-SaGl:Female | 0.468361788 | 0.029545597 | 0.90717798 | 0.0238747 |
| ZuGp:Female-ZuGa:Female | 0.084452984 | -0.085528697 | 0.25443466 | 0.9259996 |
| BRGb:Male-ZuGa:Female | -0.254815131 | -0.474133653 | -0.03549661 | 0.0075499 |
| BRGp:Male-ZuGa:Female | -0.018750337 | -0.226543592 | 0.18904292 | 1.0000000 |
| MuGl:Male-ZuGa:Female | -0.280563053 | -0.656900118 | 0.09577401 | 0.4047099 |
| MuGp:Male-ZuGa:Female | -0.194701889 | -0.346642285 | -0.04276149 | 0.0014772 |
| SaGl:Male-ZuGa:Female | -0.258044703 | -0.491766528 | -0.02432288 | 0.0154960 |
| ZuGa:Male-ZuGa:Female | -0.005912572 | -0.271709068 | 0.25988392 | 1.0000000 |
| ZuGp:Male-ZuGa:Female | 0.181382198 | -0.077391062 | 0.44015546 | 0.5133473 |
| BRGb:Male-ZuGp:Female | -0.339268115 | -0.586170622 | -0.09236561 | 0.0003686 |
| BRGp:Male-ZuGp:Female | -0.103203321 | -0.339927360 | 0.13352072 | 0.9733244 |
| MuGl:Male-ZuGp:Female | -0.365016037 | -0.758067968 | 0.02803589 | 0.1010576 |
| MuGp:Male-ZuGp:Female | -0.279154873 | -0.468749487 | -0.08956026 | 0.0000757 |
| SaGl:Male-ZuGp:Female | -0.342497687 | -0.602278605 | -0.08271677 | 0.0008597 |
| ZuGa:Male-ZuGp:Female | -0.090365556 | -0.379343083 | 0.19861197 | 0.9988567 |
| ZuGp:Male-ZuGp:Female | 0.096929214 | -0.185601906 | 0.37946034 | 0.9970581 |
| BRGp:Male-BRGb:Male | 0.236064794 | -0.038243633 | 0.51037322 | 0.1822428 |
| MuGl:Male-BRGb:Male | -0.025747922 | -0.442517238 | 0.39102139 | 1.0000000 |
| MuGp:Male-BRGb:Male | 0.060113241 | -0.174733432 | 0.29495991 | 0.9998699 |
| SaGl:Male-BRGb:Male | -0.003229572 | -0.297666268 | 0.29120712 | 1.0000000 |
| ZuGa:Male-BRGb:Male | 0.248902559 | -0.071589022 | 0.56939414 | 0.3346824 |
| ZuGp:Male-BRGb:Male | 0.436197329 | 0.121505930 | 0.75088873 | 0.0003056 |
| MuGl:Male-BRGp:Male | -0.261812716 | -0.672633949 | 0.14900852 | 0.6716325 |

$`Subpopulation:Sex`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| MuGp:Male-BRGp:Male | -0.175951552 | -0.400072909 | 0.04816980 | 0.3167973 |
| SaGl:Male-BRGp:Male | -0.239294366 | -0.525249590 | 0.04666086 | 0.2194441 |
| ZuGa:Male-BRGp:Male | 0.012837765 | -0.299879801 | 0.32555533 | 1.0000000 |
| ZuGp:Male-BRGp:Male | 0.200132535 | -0.106637899 | 0.50690297 | 0.6348795 |
| MuGp:Male-MuGl:Male | 0.085861163 | -0.299731799 | 0.47145413 | 0.9999736 |
| SaGl:Male-MuGl:Male | 0.022518350 | -0.402007197 | 0.44704390 | 1.0000000 |
| ZuGa:Male-MuGl:Male | 0.274650481 | -0.168343679 | 0.71764464 | 0.7123887 |
| ZuGp:Male-MuGl:Male | 0.461945251 | 0.023129061 | 0.90076144 | 0.0281989 |
| SaGl:Male-MuGp:Male | -0.063342814 | -0.311693850 | 0.18500822 | 0.9998751 |
| ZuGa:Male-MuGp:Male | 0.188789318 | -0.089958100 | 0.46753674 | 0.5727920 |
| ZuGp:Male-MuGp:Male | 0.376084088 | 0.104025370 | 0.64814281 | 0.0003241 |
| ZuGa:Male-SaGl:Male | 0.252132132 | -0.078382816 | 0.58264708 | 0.3649406 |
| ZuGp:Male-SaGl:Male | 0.439426901 | 0.114533144 | 0.76432066 | 0.0005140 |
| ZuGp:Male-ZuGa:Male | 0.187294770 | -0.161385370 | 0.53597491 | 0.8736210 |

$`Sex:Age`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| Male:Juvenile-Female:Juvenile | 0.006446537 | -0.11945073 | 0.132343808 | 0.9999907 |
| Female:Old-Female:Juvenile | -0.105337049 | -0.31616030 | 0.105486201 | 0.7109069 |
| Male:Old-Female:Juvenile | 0.103847959 | -0.06444170 | 0.272137614 | 0.4911823 |
| Female:Young-Female:Juvenile | -0.004843907 | -0.09873916 | 0.089051342 | 0.9999904 |
| Male:Young-Female:Juvenile | -0.069144770 | -0.18849046 | 0.050200916 | 0.5627140 |
| Female:Old-Male:Juvenile | -0.111783586 | -0.33437756 | 0.110810392 | 0.7064113 |
| Male:Old-Male:Juvenile | 0.097401422 | -0.08541828 | 0.280221127 | 0.6507629 |
| Female:Young-Male:Juvenile | -0.011290444 | -0.12926482 | 0.106683936 | 0.9997948 |
| Male:Young-Male:Juvenile | -0.075591307 | -0.21467772 | 0.063495102 | 0.6307558 |
| Male:Old-Female:Old | 0.209185008 | -0.03985187 | 0.458221882 | 0.1577424 |
| Female:Young-Female:Old | 0.100493142 | -0.10569676 | 0.306683041 | 0.7322895 |
| Male:Young-Female:Old | 0.036192279 | -0.18276285 | 0.255147410 | 0.9970933 |
| Female:Young-Male:Old | -0.108691866 | -0.27113952 | 0.053755786 | 0.3963658 |
| Male:Young-Male:Old | -0.172992729 | -0.35136401 | 0.005378554 | 0.0633485 |
| Male:Young-Female:Young | -0.064300863 | -0.17525682 | 0.046655094 | 0.5624252 |

## B.10. The best fitting GLM model of association of *T. congolense* status with *S. glossinidius* infection and intrinsic factors of all tsetse samples (Model 6).

### Backward elimination of model selection

Explanatory variables: subpopulation, sex, age and *Sodalis* status
Response variables: *T. congolense* status

| Effect tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation * sex * age * *Sodalis* status (Full model) | 44 | 412.7 | 913.4 | - | - |
| subpopulation * sex * age * *Sodalis* status | 42 | 412.7 | 909.4 | <0.0001 | 1.0000 |
| sex * age * *Sodalis* status | 41 | 416.1 | 914.3 | 6.8425 | **0.0089** |
| subpopulation * age * *Sodalis* status | 39 | 413.4 | 904.8 | 1.3727 | 0.7120 |
| subpopulation * sex * *Sodalis* status + | 36 | 414.6 | 901.2 | 2.4087 | 0.4920 |
| subpopulation * sex *age | 30 | 417.2 | 894.3 | 5.1068 | 0.5302 |

A * B means an interaction between A and B.

### Summary of the best fitting model

```
Call:
glm(formula = TC_ITS1 ~ Sex + Age_class + Sodalis_status + Sex:Age_class +
    Sex:Sodalis_status + Age_class:Sodalis_status + Sex:Age_class:Sodalis_status,
    family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8266  -0.6396  -0.5076  -0.4079   2.3207

Coefficients:
                                Estimate Std. Error z value Pr(>|z|)
(Intercept)                      -2.5274     0.3465  -7.293 3.03e-13 ***
SexMale                           0.4389     0.6703   0.655  0.51264
Age_class                         0.5431     0.2008   2.705  0.00684 **
Sodalis_status                    1.6060     0.5952   2.698  0.00697 **
SexMale:Age_class                -0.7211     0.3809  -1.893  0.05835 .
SexMale:Sodalis_status           -1.4797     1.0364  -1.428  0.15336
Age_class:Sodalis_status         -1.1045     0.3889  -2.840  0.00451 **
SexMale:Age_class:Sodalis_status  1.5679     0.5925   2.646  0.00814 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 891.49  on 1089  degrees of freedom
Residual deviance: 869.35  on 1082  degrees of freedom
AIC: 885.35

Number of Fisher Scoring iterations: 5
```

### ANOVA table

```
                              Df Sum Sq Mean Sq F value  Pr(>F)
Sex                            1   0.32  0.3166   2.631 0.10510
Age_class                      1   0.17  0.1746   1.451 0.22865
Sodalis_status                 1   0.25  0.2473   2.054 0.15205
Sex:Age_class                  1   0.04  0.0416   0.345 0.55682
Sex:Sodalis_status             1   0.64  0.6428   5.341 0.02101 *
Age_class:Sodalis_status       1   0.31  0.3102   2.577 0.10870
Sex:Age_class:Sodalis_status   1   1.00  1.0024   8.329 0.00398 **
Residuals                   1082 130.22  0.1204
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Post hoc test** (Significant differences are indicated by shaded boxes.)

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = TCa, data = TK, projections = FALSE, qr = TRUE, contrasts = NULL)

$Sex
                 diff         lwr         upr    p adj
Male-Female -0.0361381 -0.07986773 0.007591525 0.105196

$Age
                      diff         lwr        upr     p adj
Old-Juvenile    0.050464651 -0.04043367 0.14136298 0.3937980
Young-Juvenile -0.004977293 -0.05758329 0.04762870 0.9731843
Young-Old      -0.055441944 -0.14359252 0.03270864 0.3028317

$Sodalis
            diff         lwr        upr     p adj
Pos-Neg 0.0318066 -0.01171887 0.07533206 0.1518994

$`Sex:Age`
                                       diff         lwr         upr      p adj
Male:Juvenile-Female:Juvenile -0.0250086865 -0.12874005 0.07872267 0.9832650
Female:Old-Female:Juvenile     0.1233765575 -0.05032842 0.29708153 0.3271630
Male:Old-Female:Juvenile      -0.0225048732 -0.16116487 0.11615512 0.9973356
Female:Young-Female:Juvenile  -0.0009447634 -0.07830849 0.07641896 1.0000000
Male:Young-Female:Juvenile    -0.0380089387 -0.13634221 0.06032433 0.8800285
Female:Old-Male:Juvenile       0.1483852440 -0.03501807 0.33178855 0.1908956
Male:Old-Male:Juvenile         0.0025038133 -0.14812802 0.15313565 1.0000000
Female:Young-Male:Juvenile     0.0240639231 -0.07313948 0.12126732 0.9811585
Male:Young-Male:Juvenile      -0.0130002522 -0.12759863 0.10159812 0.9995278
Male:Old-Female:Old           -0.1458814306 -0.35107201 0.05930915 0.3260573
Female:Young-Female:Old       -0.1243213209 -0.29420871 0.04556607 0.2936688
Male:Young-Female:Old         -0.1613854962 -0.34179063 0.01901963 0.1098333
Female:Young-Male:Old          0.0215601097 -0.11228644 0.15540666 0.9974298
Male:Young-Male:Old           -0.0155040655 -0.16247068 0.13146255 0.9996690
Male:Young-Female:Young       -0.0370641753 -0.12848484 0.05435649 0.8568462

$`Sex:Sodalis`
                              diff         lwr         upr     p adj
Male:Neg-Female:Neg   -0.070585851 -0.140571411 -0.00060029 0.0471557
Female:Pos-Female:Neg -0.003130403 -0.072487675  0.06622687 0.9994399
Male:Pos-Female:Neg    0.034954882 -0.057935914  0.12784568 0.7675669
Female:Pos-Male:Neg    0.067455448 -0.012017375  0.14692827 0.1283165
Male:Pos-Male:Neg      0.105540733  0.004871977  0.20620949 0.0356773
Male:Pos-Female:Pos    0.038085285 -0.062147698  0.13831827 0.7622257

$`Age:Sodalis`
                                diff         lwr        upr     p adj
Old:Neg-Juvenile:Neg         0.06915169 -0.06796992 0.20627330 0.7026454
Young:Neg-Juvenile:Neg      -0.02172826 -0.10063601 0.05717950 0.9698955
Juvenile:Pos-Juvenile:Neg    0.01016353 -0.09286527 0.11319233 0.9997618
Old:Pos-Juvenile:Neg         0.02589157 -0.15227746 0.20406060 0.9984304
Young:Pos-Juvenile:Neg       0.03926725 -0.05434034 0.13287485 0.8381299
Young:Neg-Old:Neg           -0.09087995 -0.22362704 0.04186714 0.3695577
Juvenile:Pos-Old:Neg        -0.05898816 -0.20734681 0.08937049 0.8667182
Old:Pos-Old:Neg             -0.04326012 -0.25096077 0.16444053 0.9914050
Young:Pos-Old:Neg           -0.02988444 -0.17186238 0.11209351 0.9909772
Juvenile:Pos-Young:Neg       0.03189179 -0.06523898 0.12902256 0.9367772
Old:Pos-Young:Neg            0.04761983 -0.12720482 0.22244447 0.9712847
Young:Pos-Young:Neg          0.06099551 -0.02607822 0.14806924 0.3429910
Old:Pos-Juvenile:Pos         0.01572804 -0.17122692 0.20268300 0.9998913
Young:Pos-Juvenile:Pos       0.02910372 -0.08030530 0.13851275 0.9741081
Young:Pos-Old:Pos            0.01337568 -0.16855728 0.19530865 0.9999442

$`Sex:Age:Sodalis`
                                                    diff         lwr         upr      p adj
Male:Juvenile:Neg-Female:Juvenile:Neg    -0.096904025 -0.24249273  0.048684684 0.5646149
Female:Old:Neg-Female:Juvenile:Neg        0.189411765 -0.05404052  0.432864054 0.3111370
Male:Old:Neg-Female:Juvenile:Neg         -0.095588235 -0.29531957  0.104143103 0.9204040
Female:Young:Neg-Female:Juvenile:Neg     -0.043315508 -0.15420240  0.067571386 0.9815815
Male:Young:Neg-Female:Juvenile:Neg       -0.074097007 -0.21168271  0.063488695 0.8372236
Female:Juvenile:Pos-Female:Juvenile:Neg  -0.059477124 -0.20316665  0.084212399 0.9712777
Male:Juvenile:Pos-Female:Juvenile:Neg     0.061969904 -0.13203921  0.255979019 0.9965662
Female:Old:Pos-Female:Juvenile:Neg       -0.087254902 -0.42673290  0.252223096 0.9995345
Male:Old:Pos-Female:Juvenile:Neg          0.046803069 -0.20570851  0.299314648 0.9999818
```

$`Sex:Age:Sodalis`

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| Female:Young:Pos-Female:Juvenile:Neg | 0.015618661 | -0.11286202 | 0.144099344 | 0.9999998 |
| Male:Young:Pos-Female:Juvenile:Neg | -0.027731092 | -0.21201634 | 0.156554159 | 0.9999979 |
| Female:Old:Neg-Male:Juvenile:Neg | 0.286315789 | 0.03083995 | 0.541791631 | 0.0134581 |
| Male:Old:Neg-Male:Juvenile:Neg | 0.001315789 | -0.21290722 | 0.215538795 | 1.0000000 |
| Female:Young:Neg-Male:Juvenile:Neg | 0.053588517 | -0.08166972 | 0.188846757 | 0.9794335 |
| Male:Young:Neg-Male:Juvenile:Neg | 0.022807018 | -0.13508128 | 0.180695313 | 0.9999986 |
| Female:Juvenile:Pos-Male:Juvenile:Neg | 0.037426901 | -0.12580780 | 0.200661601 | 0.9998461 |
| Male:Juvenile:Pos-Male:Juvenile:Neg | 0.158873929 | -0.05002419 | 0.367772050 | 0.3466558 |
| Female:Old:Pos-Male:Juvenile:Neg | 0.009649123 | -0.33855225 | 0.357850500 | 1.0000000 |
| Male:Old:Pos-Male:Juvenile:Neg | 0.143707094 | -0.12041596 | 0.407830144 | 0.8277276 |
| Female:Young:Pos-Male:Juvenile:Neg | 0.112522686 | -0.03749789 | 0.262543266 | 0.3688700 |
| Male:Young:Pos-Male:Juvenile:Neg | 0.069172932 | -0.13072690 | 0.269072763 | 0.9931462 |
| Male:Old:Neg-Female:Old:Neg | -0.285000000 | -0.57476618 | 0.004766183 | 0.0588281 |
| Female:Young:Neg-Female:Old:Neg | -0.232727273 | -0.47014618 | 0.004691632 | 0.0607565 |
| Male:Young:Neg-Female:Old:Neg | -0.263508772 | -0.51451008 | -0.012507460 | 0.0298194 |
| Female:Juvenile:Pos-Female:Old:Neg | -0.248888889 | -0.50328722 | 0.005509445 | 0.0618838 |
| Male:Juvenile:Pos-Female:Old:Neg | -0.127441860 | -0.41329387 | 0.158410147 | 0.9507509 |
| Female:Old:Pos-Female:Old:Neg | -0.276666667 | -0.67581218 | 0.122478851 | 0.4986710 |
| Male:Old:Pos-Female:Old:Neg | -0.142608696 | -0.47098947 | 0.185772075 | 0.9591343 |
| Female:Young:Pos-Female:Old:Neg | -0.173793103 | -0.41992136 | 0.072335154 | 0.4680571 |
| Male:Young:Pos-Female:Old:Neg | -0.217142857 | -0.49648653 | 0.062200815 | 0.3125033 |
| Female:Young:Neg-Male:Old:Neg | 0.052272727 | -0.14005857 | 0.244604022 | 0.9992106 |
| Male:Young:Neg-Male:Old:Neg | 0.021491228 | -0.18737535 | 0.230357811 | 1.0000000 |
| Female:Juvenile:Pos-Male:Old:Neg | 0.036111111 | -0.17682574 | 0.249047963 | 0.9999927 |
| Male:Juvenile:Pos-Male:Old:Neg | 0.157558140 | -0.09211156 | 0.407227842 | 0.6466741 |
| Female:Old:Pos-Male:Old:Neg | 0.008333333 | -0.36575320 | 0.382419867 | 1.0000000 |
| Male:Old:Pos-Male:Old:Neg | 0.142391304 | -0.15502679 | 0.439809396 | 0.9202153 |
| Female:Young:Pos-Male:Old:Neg | 0.111206897 | -0.09177761 | 0.314191403 | 0.8211032 |
| Male:Young:Pos-Male:Old:Neg | 0.067857143 | -0.17433385 | 0.310048137 | 0.9989515 |
| Male:Young:Neg-Female:Young:Neg | -0.030781499 | -0.15738539 | 0.095822391 | 0.9997272 |
| Female:Juvenile:Pos-Female:Young:Neg | -0.016161616 | -0.14937347 | 0.117050240 | 0.9999998 |
| Male:Juvenile:Pos-Female:Young:Neg | 0.105285412 | -0.08109660 | 0.291667429 | 0.7895454 |
| Female:Old:Pos-Female:Young:Neg | -0.043939394 | -0.37911700 | 0.291238216 | 0.9999995 |
| Male:Old:Pos-Female:Young:Neg | 0.090118577 | -0.15658127 | 0.336818429 | 0.9892406 |
| Female:Young:Pos-Female:Young:Neg | 0.058934169 | -0.05771061 | 0.175578952 | 0.8879259 |
| Male:Young:Pos-Female:Young:Neg | 0.015584416 | -0.16065342 | 0.191822247 | 1.0000000 |
| Female:Juvenile:Pos-Male:Young:Neg | 0.014619883 | -0.14151890 | 0.170758670 | 1.0000000 |
| Male:Juvenile:Pos-Male:Young:Neg | 0.136066911 | -0.06733461 | 0.339468432 | 0.5565590 |
| Female:Old:Pos-Male:Young:Neg | -0.013157895 | -0.35808970 | 0.331773914 | 1.0000000 |
| Male:Old:Pos-Male:Young:Neg | 0.120900076 | -0.13889742 | 0.380697573 | 0.9338771 |
| Female:Young:Pos-Male:Young:Neg | 0.089715668 | -0.05255143 | 0.231982768 | 0.6477289 |
| Male:Young:Pos-Male:Young:Neg | 0.046365915 | -0.14778273 | 0.240514557 | 0.9997707 |
| Male:Juvenile:Pos-Female:Juvenile:Pos | 0.121447028 | -0.08613195 | 0.329026006 | 0.7490151 |
| Female:Old:Pos-Female:Juvenile:Pos | -0.027777778 | -0.37518936 | 0.319633803 | 1.0000000 |
| Male:Old:Pos-Female:Juvenile:Pos | 0.106280193 | -0.15680077 | 0.369361156 | 0.9761075 |
| Female:Young:Pos-Female:Juvenile:Pos | 0.075095785 | -0.07308242 | 0.223273993 | 0.8858167 |
| Male:Young:Pos-Female:Juvenile:Pos | 0.031746032 | -0.16677487 | 0.230266935 | 0.9999960 |
| Female:Old:Pos-Male:Juvenile:Pos | -0.149224806 | -0.52028769 | 0.221838078 | 0.9769228 |
| Male:Old:Pos-Male:Juvenile:Pos | -0.015166835 | -0.30877278 | 0.278439109 | 1.0000000 |
| Female:Young:Pos-Male:Juvenile:Pos | -0.046351243 | -0.24370787 | 0.151005386 | 0.9998052 |
| Male:Young:Pos-Male:Juvenile:Pos | -0.089700997 | -0.32719501 | 0.147793017 | 0.9858960 |
| Male:Old:Pos-Female:Old:Pos | 0.134057971 | -0.27067679 | 0.538792729 | 0.9952862 |
| Female:Young:Pos-Female:Old:Pos | 0.102873563 | -0.23852856 | 0.444275691 | 0.9979585 |
| Male:Young:Pos-Female:Old:Pos | 0.059523810 | -0.30654883 | 0.425596454 | 0.9999953 |
| Female:Young:Pos-Male:Old:Pos | -0.031184408 | -0.28627694 | 0.223908124 | 0.9999998 |
| Male:Young:Pos-Male:Old:Pos | -0.074534161 | -0.36180749 | 0.212739171 | 0.9994909 |
| Male:Young:Pos-Female:Young:Pos | -0.043349754 | -0.23115592 | 0.144456412 | 0.9998356 |

# B.11. The best fitting GLM model of association of *T. brucei* status with *S. glossinidius* infection and intrinsic factors of all tsetse samples (Model 7).

## Backward elimination of model selection

Explanatory variables: subpopulation, sex, age and *Sodalis* status
Response variables: *T. brucei* status

| Effect tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation ∗ sex ∗age ∗ *Sodalis* status (full model) | 44 | 250.4 | 588.8 | - | - |
| subpopulation ∗ sex ∗age ∗ *Sodalis* status | 42 | 250.4 | 584.8 | 0.0000 | 1.0000 |
| sex ∗ age ∗ *Sodalis* status | 41 | 254.4 | 590.7 | 7.9383 | **0.0048** |
| subpopulation∗ age ∗ *Sodalis* status | 39 | 257.3 | 592.6 | 13.8440 | **0.0031** |
| subpopulation ∗ sex ∗ *Sodalis* status | 40 | 255.0 | 590.0 | 9.2688 | **0.0097** |
| subpopulation ∗ sex ∗ age | 36 | 256.4 | 584.9 | 12.1000 | 0.05977 |

A ∗ B means an interaction between A and B.

## Summary of the best fitting model

```
Call:
glm(formula = TB_ITS1 ~ Subpopulation + Sex + Age_class + Sodalis_status +
    Subpopulation:Sex + Subpopulation:Age_class + Sex:Age_class +
    Subpopulation:Sodalis_status + Sex:Sodalis_status + Age_class:Sodalis_status +
    Subpopulation:Sex:Sodalis_status + Subpopulation:Age_class:Sodalis_status +
    Sex:Age_class:Sodalis_status, family = binomial, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1774  -0.3964  -0.3451  -0.2910   2.9152

Coefficients: (8 not defined because of singularities)
                                      Estimate Std. Error z value Pr(>|z|)
(Intercept)                           18.1475  6522.6395   0.003    0.998
SubpopulationBRGp                     -5.8305  6622.9195  -0.001    0.999
SubpopulationMuGl                    -22.6222  6741.6311  -0.003    0.997
SubpopulationMuGp                    -21.0657  6522.6395  -0.003    0.997
SubpopulationSaGl                    -37.1366  6735.0394  -0.006    0.996
SubpopulationZuGa                    -21.6593  6522.6395  -0.003    0.997
SubpopulationZuGp                    -21.7779  6522.6396  -0.003    0.997
SexMale                                0.4186     1.9061   0.220    0.826
Age_class                            -18.3973  6522.6388  -0.003    0.998
Sodalis_status                       -20.8689  6522.6396  -0.003    0.997
SubpopulationBRGp:SexMale              0.6947     2.2465   0.309    0.757
SubpopulationMuGl:SexMale             14.9089  1404.7298   0.011    0.992
SubpopulationMuGp:SexMale             -0.8035     1.6935  -0.474    0.635
SubpopulationSaGl:SexMale             16.3835  1678.0737   0.010    0.992
SubpopulationZuGa:SexMale              0.9570     1.7313   0.553    0.580
SubpopulationZuGp:SexMale             -0.5995     1.0960  -0.547    0.584
SubpopulationBRGp:Age_class            3.1900  6622.9187   0.000    1.000
SubpopulationMuGl:Age_class            5.7674  6593.6555   0.001    0.999
SubpopulationMuGp:Age_class           18.8114  6522.6388   0.003    0.998
SubpopulationSaGl:Age_class           18.6552  6522.6388   0.003    0.998
SubpopulationZuGa:Age_class           18.8935  6522.6388   0.003    0.998
SubpopulationZuGp:Age_class           19.2365  6522.6388   0.003    0.998
```

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| SexMale:Age_class | -0.1687 | 0.5131 | -0.329 | 0.742 |
| SubpopulationBRGp:Sodalis_status | 4.3203 | 6622.9198 | 0.001 | 0.999 |
| SubpopulationMuGl:Sodalis_status | NA | NA | NA | NA |
| SubpopulationMuGp:Sodalis_status | -13.6391 | 6522.6390 | -0.002 | 0.998 |
| SubpopulationSaGl:Sodalis_status | NA | NA | NA | NA |
| SubpopulationZuGa:Sodalis_status | 39.0850 | 6650.9414 | 0.006 | 0.995 |
| SubpopulationZuGp:Sodalis_status | 21.0080 | 6522.6397 | 0.003 | 0.997 |
| SexMale:Sodalis_status | 1.2754 | 2.1652 | 0.589 | 0.556 |
| Age_class:Sodalis_status | 18.4461 | 6522.6388 | 0.003 | 0.998 |
| SubpopulationBRGp:SexMale:Sodalis_status | -1.0600 | 2.5941 | -0.409 | 0.683 |
| SubpopulationMuGl:SexMale:Sodalis_status | NA | NA | NA | NA |
| SubpopulationMuGp:SexMale:Sodalis_status | -37.3024 | 15977.1368 | -0.002 | 0.998 |
| SubpopulationSaGl:SexMale:Sodalis_status | NA | NA | NA | NA |
| SubpopulationZuGa:SexMale:Sodalis_status | -18.1933 | 1904.1998 | -0.010 | 0.992 |
| SubpopulationZuGp:SexMale:Sodalis_status | NA | NA | NA | NA |
| SubpopulationBRGp:Age_class:Sodalis_status | -3.2404 | 6622.9188 | 0.000 | 1.000 |
| SubpopulationMuGl:Age_class:Sodalis_status | NA | NA | NA | NA |
| SubpopulationMuGp:Age_class:Sodalis_status | NA | NA | NA | NA |
| SubpopulationSaGl:Age_class:Sodalis_status | NA | NA | NA | NA |
| SubpopulationZuGa:Age_class:Sodalis_status | -35.3812 | 6650.9405 | -0.005 | 0.996 |
| SubpopulationZuGp:Age_class:Sodalis_status | -18.7226 | 6522.6389 | -0.003 | 0.998 |
| SexMale:Age_class:Sodalis_status | -0.2673 | 0.8896 | -0.301 | 0.764 |

(Dispersion parameter for binomial family taken to be 1)

```
    Null deviance: 551.38  on 1089  degrees of freedom
Residual deviance: 512.87  on 1054  degrees of freedom
AIC: 584.87

Number of Fisher Scoring iterations: 17
```

**ANOVA table**

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|---|---|---|---|---|---|
| Subpopulation | 6 | 0.51 | 0.08525 | 1.314 | 0.248 |
| Sex | 1 | 0.10 | 0.10355 | 1.596 | 0.207 |
| Age_class | 1 | 0.02 | 0.01681 | 0.259 | 0.611 |
| Sodalis_status | 1 | 0.05 | 0.05466 | 0.843 | 0.359 |
| Subpopulation:Sex | 6 | 0.57 | 0.09537 | 1.470 | 0.185 |
| Subpopulation:Age_class | 6 | 0.30 | 0.05011 | 0.772 | 0.592 |
| Sex:Age_class | 1 | 0.05 | 0.05363 | 0.827 | 0.363 |
| Subpopulation:Sodalis_status | 4 | 0.07 | 0.01639 | 0.253 | 0.908 |
| Sex:Sodalis_status | 1 | 0.02 | 0.02441 | 0.376 | 0.540 |
| Age_class:Sodalis_status | 1 | 0.02 | 0.02453 | 0.378 | 0.539 |
| Subpopulation:Sex:Sodalis_status | 3 | 0.21 | 0.07030 | 1.084 | 0.355 |
| Subpopulation:Age_class:Sodalis_status | 3 | 0.39 | 0.12853 | 1.981 | 0.115 |
| Sex:Age_class:Sodalis_status | 1 | 0.00 | 0.00028 | 0.004 | 0.948 |
| Residuals | 1054 | 68.38 | 0.06487 | | |

## B.12. The association of *Sodalis* level (Model 9) and trypanosome level (Model 10).

### Introduction

Although the PCR screening method was not likely to reflect absolute abundance, the intensity of PCR bands on agarose gels probably could indicate levels of infection. In order to assess whether tsetse factors affected levels of infection with the endosymbiont, and whether this influenced to trypanosome infections, Sodalis and trypanosome statuses (see chapter 2) of individual flies were classified into six levels of PCR product intensity.

### Methods

GLM Models of *Sodalis* level and trypanosome level were for initial investigation whether tsetse factors affected levels of endosymbiont and trypanosome infections. *Sodalis* and trypanosome statuses of individual flies were considered based on band intensity of PCR reactions (chapter 2), which were classified into six levels of PCR product intensity: “0” for negative; “1” for very weak positive; “2” for weak positive; “3” for positive; “4” for strong positive; and “5” for very strong positive (Figure B.1). Generalised Linear Models (GLMs), as implemented in the programming environment R (version 3.1.2), were applied to test the effects of subpopulation (seven groups), sex (male and female), age (young, juvenile and old) of tsetse flies as fixed categorical variables. The GLM Model 9 analyses repeated the previous GLM Model anlyses but using levels of *Sodalis* as the response variable and considering levels of trypanosome infection, rather than just presence or absence, as the explanatory variable. For GLM Model 10, level of trypanosomes was treated as the response variable and levels of *Sodalis* was analysed as the explanatory variable. Subpopulation, Sex and age were fixed effect categorical variables in the GLMs analyses using the Poisson family since response variables were classified as levels of the *Sodalis* status and trypanosome in the GLM Model 9 and 10 (Table B.1). The best fitting model was The “step()” function and manual backward elimination technique was applied to the full models to determine the best fitting models using likelihood Ratio Tests (LRTs), at a significance level of alpha = 0.05.

Principle correspondence analysis (PCA) was used to perform graphic associations of trypanosome infection levels with level of Sodalis PCR screening and each intrinsic factor in all collected flies using Multiple Correspondence Analysis (MCA) MCA analyses included five categorical variables, comprised of 6 levels of *S. glossinidius*, *Trypanosoma spp.*, subpopulation, sex and age of tsetse flies. All factors were classified as categorical variables and analysed in the MCA B1 (Table B.2).

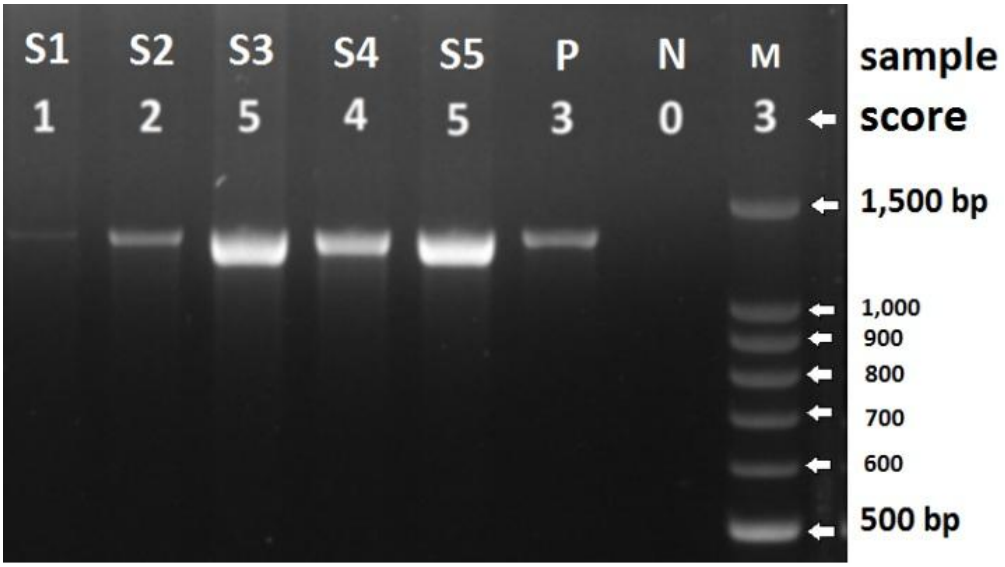



**Figure B.1** Scoring of *S. glossinidius* and *Trypanosoma spp.* presence based on relative intensity of PCR products.

Bands were classified into six categories based on running 3 µl of each PCR product alongside 2.5 µl of a 100 bp DNA ladder (Promega Corporation, Madison, U.S.A.). M: the 500 bp band of the DNA ladder was set as the "strong" amplification level (score = 4); the other ladder bands were considered "moderate" (score = 3). S1 (sample1): very weak amplification (score = 1); S2 (sample 2): weak amplification (score = 2); S3 (sample 3): very strong amplification (score = 5); S4 (sample 4): strong amplification (score = 4); S5 (sample 5): very strong amplification (score = 5); P (Positive control): moderate amplification (score = 3); N (Negative control): no amplification (score = 0). This picture was a gel document of GPO1 PCR products.

**Table B.1** Variables and types for statistical analysis in GLM Model 9 and 10 to determine associations of *Sodalis* level and trypanosome level.

| No | Intrinsic factors | Type of data | Compositions |
|---|---|---|---|
| **Model 9: An association of Sodalis levels with intrinsic factors and trypanosome levels in all tsetse flies (N = 1090)** | | | |
| **Explanatory variables** | | | |
| 1. | subpopulation | categorical | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl[a] |
| 2. | sex | categorical | male and female |
| 3. | age | categorical | young, juvenile and old |
| 4. | trypanosome level | integer | negative (0), very weak (1), weak (2), moderate (3), strong (4) and very strong (5) |
| 5. | subpopulation * sex | interactions between two categorical factors | |
| 6. | subpopulation * age | interactions between two categorical factors | |
| 7. | subpopulation * trypanosome level | interactions between categorical and integer factors | |
| 8. | sex * age | interactions between two categorical factors | |
| 9. | sex * trypanosome level | interactions between categorical and integer factors | |
| 10. | age * trypanosome level | interactions between categorical and integer factors | |
| 11. | subpopulation * sex * age | interactions among three categorical factors | |
| 12. | subpopulation * sex * trypanosome level | interactions among two categorical and one integer factors | |
| 13. | sex * age * trypanosome level | interactions among two categorical and one integer factors | |
| 14. | age * trypanosome level * subpopulation | interactions among two categorical and one integer factors | |
| 15. | subpopulation * sex * age * trypanosome level | interactions among three categorical and one integer factors | |
| **Response variable** | | | |
| 1. | *Sodalis* level | integer | negative (0), very weak (1), weak (2), moderate (3), strong (4) and very strong (5) |

| No | Intrinsic factors | Type of data | Compositions |
|---|---|---|---|
| **Model 10: An association of trypanosome levels with intrinsic factors and *Sodalis* levels in all tsetse flies (N = 1090)** | | | |
| **Explanatory variables** | | | |
| 1. | subpopulation | categorical | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl[a] |
| 2. | sex | categorical | male and female |
| 3. | age | categorical | young, juvenile and old |
| 4. | *Sodalis* level | integer | negative (0), very weak (1), weak (2), moderate (3), strong (4) and very strong (5) |
| 5. | subpopulation ∗ sex | interactions between two categorical factors | |
| 6. | subpopulation ∗ age | interactions between two categorical factors | |
| 7. | subpopulation ∗ *Sodalis* level | interactions between categorical and integer factors | |
| 8. | sex ∗ age | interactions between two categorical factors | |
| 9. | sex ∗ *Sodalis* level | interactions between categorical and integer factors | |
| 10. | age ∗ *Sodalis* level | interactions between categorical and integer factors | |
| 11. | subpopulation ∗ sex ∗ age | interactions among three categorical factors | |
| 12. | subpopulation ∗ sex ∗ *Sodalis* level | interactions among two categorical and one integer factors | |
| 13. | sex ∗ age ∗ *Sodalis* level | interactions among two categorical and one integer factors | |
| 14. | age ∗ *Sodalis* level ∗ subpopulation | interactions among two categorical and one integer factors | |
| 15. | subpopulation ∗ sex ∗ age ∗ *Sodalis* level | interactions among three categorical and one integer factors | |
| **Response variable** | | | |
| 1. | trypanosome level | integer | negative (0), very weak (1), weak (2), moderate (3), strong (4) and very strong (5) |

[a] BRGb is *G. brevipalpis* from Buffalo Ridge; BRGp is *G. pallidipes* from Buffalo Ridge: ZuGa is *G. austeni* from Zungu Luka; ZuGp is *G. pallidipes* from Zungu Luka; MuGl is *G. longipennis* from Mukinyo; MuGp is *G. pallidipes* from Mukinyo; and SaGl is *G. longipennis* from Sampu.

**Table B.2** Variables and types for statistical analysis in MCA B1 to determine associations among level of trypanosome and *Sodalis* andintrinsic factors of tsetse using the Multiple Correspondence Analysis (MCA)

| No | Intrinsic factors | Type of data | Compositions |
|---|---|---|---|
| 1. | subpopulation | categorical | BRGb, BRGp, ZuGa, ZuGp, MuGl, MuGp, SaGl |
| 2. | sex | categorical | male and female |
| 3. | age | categorical | young, juvenile and old |
| 4. | *Sodalis* level | categorical | negative, very weak, weak, moderate, strong and very strong |
| 5. | trypanosome level | categorical | negative, very weak, weak, moderate, strong and very strong |

**Results**

Using GLM analyses, the best fitting model of *Sodalis* level included a significant three-way interaction among subpopulation, sex, age and level of trypanosomes (Table B.3a and Table B.3c; Figure B.2). The relationship among these intrinsic factors and level of *Sodalis* present to infection level of trypanosomes in tsetse flies was very complicate with multiple three-way interaction among subpopulation, sex, age and *Sodalis* level (Table B.3b and Table B.3c; Figure B.3). Association of each trypanosome and tsetse flies species might make less complicate relationships among the factors but this study was limited with number of samples. A graphical MCA B1 (Table B. 4; Figure B.4) could not explan the association of levels of *Sodalis* and trypanosomes. Weak and moderate positive levels of *Sodalis* status, however, were strongly correlated with the BRGb subpopulation. Positive and negative of trypanosome infection status were correlated to sex: female vs strong positive; and male vs negative.

**Table B.3** Associations of *Sodalis glossinidius* and trypanosome levels in 1090 tsetse flies (Models 9 and 10). Levels were a qualitative classification based on the intensity of bands detected in PCR.

**Table B.3a** Model selctions of *Sodalis* level (Model 9).

| Effect tests | df | -logLik | AIC | $\chi^2$ | P-values |
|---|---|---|---|---|---|
| subpopulation ∗ sex ∗ age ∗ *Sodalis* level (full model) | 46 | 1444.2 | 2980.4 | - | - |
| subpopulation ∗ sex ∗ age ∗ *Sodalis* level | 43 | 1446.8 | 2979.7 | 5.2754 | 0.1527 |
| sex ∗ ages ∗*Sodalis* level | 43 | 1447.5 | 2978.8 | 1.0951 | 0.2953 |
| subpopulation ∗ age ∗ *Sodalis* level | 39 | 1452.7 | 2983.5 | 10.700 | **0.0135** |
| subpopulation ∗ sex ∗ *Sodalis* level | 39 | 1452.5 | 2982.9 | 10.131 | **0.0175** |
| subpopulation ∗ sex ∗ age | 36 | 1455.4 | 2982.8 | 16.028 | **0.0136** |

A ∗ B means an interaction between A and B

**Table B.3b** Model selctions of trypanosome level (Model10).

| Effect tests | df | logLik | AIC | $\chi$2 | P-values |
|---|---|---|---|---|---|
| subpopulation ∗ sex ∗ age ∗ trypanosome level (full model) | 55 | 1033.6 | 2177.2 | - | - |
| subpopulation ∗ sex ∗ age ∗ trypanosome level | 50 | 1038.6 | 2177.3 | 10.0660 | 0.0734 |
| sex ∗ ages ∗ trypanosome level | 49 | 1039.0 | 2175.9 | 0.6511 | 0.4197 |
| subpopulation ∗ age ∗ trypanosome level | 43 | 1042.7 | 2171.3 | 7.3949 | 0.2859 |
| subpopulation ∗ sex ∗ trypanosome level | 37 | 1051.8 | 2177.6 | 18.2630 | **0.0056** |
| subpopulation ∗ sex ∗ age | 37 | 1046.6 | 2167.2 | 7.9272 | 0.2435 |

**Table B.3c** Summary of the best fitting models of *Sodalis* and trypanosome levels (Models 9 and 10).

| Response variables | Explanatory variables of best fitting models | Full model | | | Best Fitting model | | |
|---|---|---|---|---|---|---|---|
| | | df | -logLik | AIC | df | -logLik | AIC |
| **Model 9** | Explanatory variables: subpopulation, sex, age and trypanosome level<br>Response variables: *Sodalis* level (N = 1090) | | | | | | |
| *Sodalis glossinidius* | subpopulation*sex* trypanosome level | 55 | 1033.6 | 2177.2 | 55 | 1033.6 | 2177.2 |
| **Model 10** | Explanatory variables: subpopulation, sex, age and *Sodalis* level<br>Response variables: trypanosome level (N = 1090) | | | | | | |
| Trypanosome level | Multiple three-way interaction among subpopulation, sex, age and *Sodalis* level | 46 | 1444.2 | 2980.4 | 42 | 1447.4 | 2978.8 |

A * B means interaction between A and B; AIC = Akaike information criterion; -logLik = - log(Likelihood); and df = Degree of freedom.

**Figure B.2** The best fitting GLM model of association of *Sodalis* levels with trypanosome levels and intrinsic factors of all tsetse samples (Model 9).

```
Call:
glm(formula = Sodalis_level ~ Subpopulation + Sex + Age_class +
    Trypanosome_level + Subpopulation:Sex + Subpopulation:Age_class +
    Subpopulation:Trypanosome_level + Sex:Age_class + Sex:Trypanosome_level +
    Age_class:Trypanosome_level + Subpopulation:Sex:Trypanosome_level,
    family = poisson, data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6224  -0.7622  -0.0865   0.1293   3.7148

Coefficients:
                                              Estimate Std. Error z value Pr(>|z|)
(Intercept)                                  9.380e-01  1.488e-01   6.305 2.88e-10 ***
SubpopulationBRGp                           -9.312e-02  1.902e-01  -0.490 0.624366
SubpopulationMuGl                           -2.034e+01  4.857e+03  -0.004 0.996658
SubpopulationMuGp                           -7.282e+00  1.995e+00  -3.650 0.000262 ***
SubpopulationSaGl                           -2.031e+01  4.080e+03  -0.005 0.996027
SubpopulationZuGa                           -2.267e+00  3.352e-01  -6.763 1.35e-11 ***
SubpopulationZuGp                           -7.963e-02  2.187e-01  -0.364 0.715756
SexMale                                     -4.311e-02  2.056e-01  -0.210 0.833946
Age_class                                    8.506e-02  8.924e-02   0.953 0.340508
Trypanosome_level                            8.325e-02  7.916e-02   1.052 0.292907
SubpopulationBRGp:SexMale                   -6.088e-02  1.811e-01  -0.336 0.736692
SubpopulationMuGl:SexMale                    1.279e-01  4.076e+03   0.000 0.999975
SubpopulationMuGp:SexMale                   -3.698e-01  1.475e+00  -0.251 0.802046
SubpopulationSaGl:SexMale                    1.058e-01  2.970e+03   0.000 0.999972
SubpopulationZuGa:SexMale                   -8.898e-01  5.939e-01  -1.498 0.134091
SubpopulationZuGp:SexMale                   -3.003e-02  2.287e-01  -0.131 0.895528
SubpopulationBRGp:Age_class                 -4.228e-02  1.085e-01  -0.390 0.696722
SubpopulationMuGl:Age_class                 -3.220e-02  2.022e+03   0.000 0.999987
SubpopulationMuGp:Age_class                  1.125e+00  9.634e-01   1.168 0.242963
SubpopulationSaGl:Age_class                 -3.996e-02  1.926e+03   0.000 0.999983
SubpopulationZuGa:Age_class                  7.653e-03  1.981e-01   0.039 0.969187
SubpopulationZuGp:Age_class                 -7.111e-02  1.271e-01  -0.560 0.575798
SubpopulationBRGp:Trypanosome_level         -6.323e-02  9.966e-02  -0.634 0.525782
SubpopulationMuGl:Trypanosome_level          1.516e-02  2.177e+03   0.000 0.999994
SubpopulationMuGp:Trypanosome_level         -1.238e+01  6.771e+02  -0.018 0.985411
SubpopulationSaGl:Trypanosome_level         -3.100e-02  2.460e+03   0.000 0.999990
SubpopulationZuGa:Trypanosome_level         -3.496e-02  9.386e-02  -0.372 0.709553
SubpopulationZuGp:Trypanosome_level         -6.706e-02  7.589e-02  -0.884 0.376861
SexMale:Age_class                           -4.120e-02  9.598e-02  -0.429 0.667748
SexMale:Trypanosome_level                    3.421e-02  9.014e-02   0.380 0.704254
Age_class:Trypanosome_level                 -4.493e-02  3.668e-02  -1.225 0.220614
SubpopulationBRGp:SexMale:Trypanosome_level  1.338e-01  1.311e-01   1.021 0.307178
SubpopulationMuGl:SexMale:Trypanosome_level -7.259e-02  3.582e+03   0.000 0.999984
SubpopulationMuGp:SexMale:Trypanosome_level -1.049e-01  1.079e+03   0.000 0.999922
SubpopulationSaGl:SexMale:Trypanosome_level  2.134e-02  2.791e+03   0.000 0.999994
SubpopulationZuGa:SexMale:Trypanosome_level  6.563e-01  1.751e-01   3.749 0.000178 ***
SubpopulationZuGp:SexMale:Trypanosome_level  5.485e-02  1.549e-01   0.354 0.723363
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2628.45  on 1089  degrees of freedom
Residual deviance:  990.46  on 1053  degrees of freedom
AIC: 2167.2

Number of Fisher Scoring iterations: 17
```

## ANOVA table

```
                                 Df Sum Sq Mean Sq F value   Pr(>F)
Subpopulation                     6 1474.1  245.69 214.638  < 2e-16 ***
Sex                               1    0.0    0.00   0.000 0.993375
Age_class                         1    0.1    0.12   0.101 0.750688
Trypanosome_level                 1    1.7    1.67   1.460 0.227203
Subpopulation:Sex                 6    8.0    1.33   1.162 0.324531
Subpopulation:Age_class           6    2.3    0.38   0.332 0.920380
Subpopulation:Trypanosome_level   6    6.5    1.09   0.953 0.455988
Sex:Age_class                     1    0.0    0.04   0.036 0.850143
Sex:Trypanosome_level             1   14.3   14.26  12.457 0.000434 ***
Age_class:Trypanosome_level       1    2.4    2.39   2.085 0.149004
Subpopulation:Sex:Trypanosome_level  6  15.7   2.61   2.279 0.034278 *
Residuals                      1053 1205.3    1.14
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Figure B.3** The best fitting GLM model of association of trypanosome level with *S. glossinidius* level and intrinsic factors of all tsetse samples (Model 10).

```
Call:
glm(formula = Trypanosome_level ~ Subpopulation + Sex + Age_class +
    Sodalis_level + Subpopulation:Sex + Subpopulation:Age_class +
    Sex:Age_class + Subpopulation:Sodalis_level + Sex:Sodalis_level +
    Age_class:Sodalis_level + Subpopulation:Sex:Age_class + Subpopulation:Sex:Sodalis_level +
    Subpopulation:Age_class:Sodalis_level, family = poisson,
    data = TK)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3017  -1.1787  -0.8952   0.6355   3.6658

Coefficients: (7 not defined because of singularities)
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -3.57114    1.71746  -2.079   0.0376 *
SubpopulationBRGp                          3.58218    1.79510   1.996   0.0460 *
SubpopulationMuGl                          3.04982    1.87561   1.626   0.1039
SubpopulationMuGp                          2.79504    1.73368   1.612   0.1069
SubpopulationSaGl                         16.20554  300.56161   0.054   0.9570
SubpopulationZuGa                          3.40414    1.72550   1.973   0.0485 *
SubpopulationZuGp                          3.39223    1.74618   1.943   0.0521 .
SexMale                                    0.46504    1.27438   0.365   0.7152
Age_class                                  1.05738    0.82533   1.281   0.2001
Sodalis_level                              1.00690    0.55550   1.813   0.0699 .
SubpopulationBRGp:SexMale                 -1.01738    1.41283  -0.720   0.4715
SubpopulationMuGl:SexMale                  0.28185    1.71575   0.164   0.8695
SubpopulationMuGp:SexMale                 -0.94731    1.33112  -0.712   0.4767
SubpopulationSaGl:SexMale                -15.46369  300.56029  -0.051   0.9590
SubpopulationZuGa:SexMale                 -1.49879    1.36688  -1.097   0.2729
SubpopulationZuGp:SexMale                 -1.18199    1.40463  -0.841   0.4001
SubpopulationBRGp:Age_class               -1.69540    0.90380  -1.876   0.0607 .
SubpopulationMuGl:Age_class               -0.82724    0.89446  -0.925   0.3550
SubpopulationMuGp:Age_class               -0.64557    0.83858  -0.770   0.4414
SubpopulationSaGl:Age_class              -14.38493  300.55659  -0.048   0.9618
SubpopulationZuGa:Age_class               -0.78361    0.83121  -0.943   0.3458
SubpopulationZuGp:Age_class               -0.73767    0.84173  -0.876   0.3808
SexMale:Age_class                          0.47327    0.33377   1.418   0.1562
SubpopulationBRGp:Sodalis_level           -1.23974    0.57525  -2.155   0.0312 *
SubpopulationMuGl:Sodalis_level                NA         NA      NA       NA
SubpopulationMuGp:Sodalis_level          -15.61397 1275.75392  -0.012   0.9902
SubpopulationSaGl:Sodalis_level                NA         NA      NA       NA
SubpopulationZuGa:Sodalis_level           -0.69603    0.57032  -1.220   0.2223
SubpopulationZuGp:Sodalis_level           -1.04013    0.56768  -1.832   0.0669 .
SexMale:Sodalis_level                     -0.42410    0.38172  -1.111   0.2665
Age_class:Sodalis_level                   -0.33123    0.26942  -1.229   0.2189
SubpopulationBRGp:SexMale:Age_class        0.26282    0.46481   0.565   0.5718
SubpopulationMuGl:SexMale:Age_class       -1.42878    0.74785  -1.911   0.0561 .
SubpopulationMuGp:SexMale:Age_class       -0.45730    0.39461  -1.159   0.2465
SubpopulationSaGl:SexMale:Age_class       13.72260  300.55582   0.046   0.9636
SubpopulationZuGa:SexMale:Age_class       -0.08883    0.40242  -0.221   0.8253
SubpopulationZuGp:SexMale:Age_class       -0.29825    0.42524  -0.701   0.4831
SubpopulationBRGp:SexMale:Sodalis_level    0.55135    0.40579   1.359   0.1742
SubpopulationMuGl:SexMale:Sodalis_level        NA         NA      NA       NA
SubpopulationMuGp:SexMale:Sodalis_level    0.69732 1804.18853   0.000   0.9997
SubpopulationSaGl:SexMale:Sodalis_level        NA         NA      NA       NA
SubpopulationZuGa:SexMale:Sodalis_level    0.87317    0.39519   2.210   0.0271 *
SubpopulationZuGp:SexMale:Sodalis_level    0.50551    0.39579   1.277   0.2015
SubpopulationBRGp:Age_class:Sodalis_level  0.46820    0.28366   1.651   0.0988 .
SubpopulationMuGl:Age_class:Sodalis_level      NA         NA      NA       NA
SubpopulationMuGp:Age_class:Sodalis_level      NA         NA      NA       NA
SubpopulationSaGl:Age_class:Sodalis_level      NA         NA      NA       NA
SubpopulationZuGa:Age_class:Sodalis_level  0.12103    0.27967   0.433   0.6652
SubpopulationZuGp:Age_class:Sodalis_level  0.31900    0.27692   1.152   0.2494
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1985.0  on 1089  degrees of freedom
Residual deviance: 1724.6  on 1048  degrees of freedom
AIC: 2978.8
```

**ANOVA table**

```
SubpopulationBRGp:Sodalis_level             -1.23974     0.57525  -2.155   0.0312 *
SubpopulationMuGl:Sodalis_level                   NA          NA      NA       NA
SubpopulationMuGp:Sodalis_level            -15.61397  1275.75392  -0.012   0.9902
SubpopulationSaGl:Sodalis_level                   NA          NA      NA       NA
SubpopulationZuGa:Sodalis_level             -0.69603     0.57032  -1.220   0.2223
SubpopulationZuGp:Sodalis_level             -1.04013     0.56768  -1.832   0.0669 .
SexMale:Sodalis_level                       -0.42410     0.38172  -1.111   0.2665
Age_class:Sodalis_level                     -0.33123     0.26942  -1.229   0.2189
SubpopulationBRGp:SexMale:Age_class          0.26282     0.46481   0.565   0.5718
SubpopulationMuGl:SexMale:Age_class         -1.42878     0.74785  -1.911   0.0561 .
SubpopulationMuGp:SexMale:Age_class         -0.45730     0.39461  -1.159   0.2465
SubpopulationSaGl:SexMale:Age_class         13.72260   300.55582   0.046   0.9636
SubpopulationZuGa:SexMale:Age_class         -0.08883     0.40242  -0.221   0.8253
SubpopulationZuGp:SexMale:Age_class         -0.29825     0.42524  -0.701   0.4831
SubpopulationBRGp:SexMale:Sodalis_level      0.55135     0.40579   1.359   0.1742
SubpopulationMuGl:SexMale:Sodalis_level           NA          NA      NA       NA
SubpopulationMuGp:SexMale:Sodalis_level      0.69732  1804.18853   0.000   0.9997
SubpopulationSaGl:SexMale:Sodalis_level           NA          NA      NA       NA
SubpopulationZuGa:SexMale:Sodalis_level      0.87317     0.39519   2.210   0.0271 *
SubpopulationZuGp:SexMale:Sodalis_level      0.50551     0.39579   1.277   0.2015
SubpopulationBRGp:Age_class:Sodalis_level    0.46820     0.28366   1.651   0.0988 .
SubpopulationMuGl:Age_class:Sodalis_level         NA          NA      NA       NA
SubpopulationMuGp:Age_class:Sodalis_level         NA          NA      NA       NA
SubpopulationSaGl:Age_class:Sodalis_level         NA          NA      NA       NA
SubpopulationZuGa:Age_class:Sodalis_level    0.12103     0.27967   0.433   0.6652
SubpopulationZuGp:Age_class:Sodalis_level    0.31900     0.27692   1.152   0.2494
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1985.0  on 1089  degrees of freedom
Residual deviance: 1724.6  on 1048  degrees of freedom
AIC: 2978.8
```

**Table B. 4** Adjusted eta squared of the combination of variables in dimensions 1-3 in MCA B1.

| Variables | Dimension 1 | Dimension 2 | Dimension 3 |
|---|---|---|---|
| *Sodalis*_level | 0.894 | 0.447 | 0.406 |
| subpopulation | 0.902 | 0.692 | 0.644 |
| sex | 0.000 | 0.239 | 0.254 |
| age | 0.004 | 0.053 | 0.144 |
| Trypanosome_level | 0.074 | 0.186 | 0.054 |

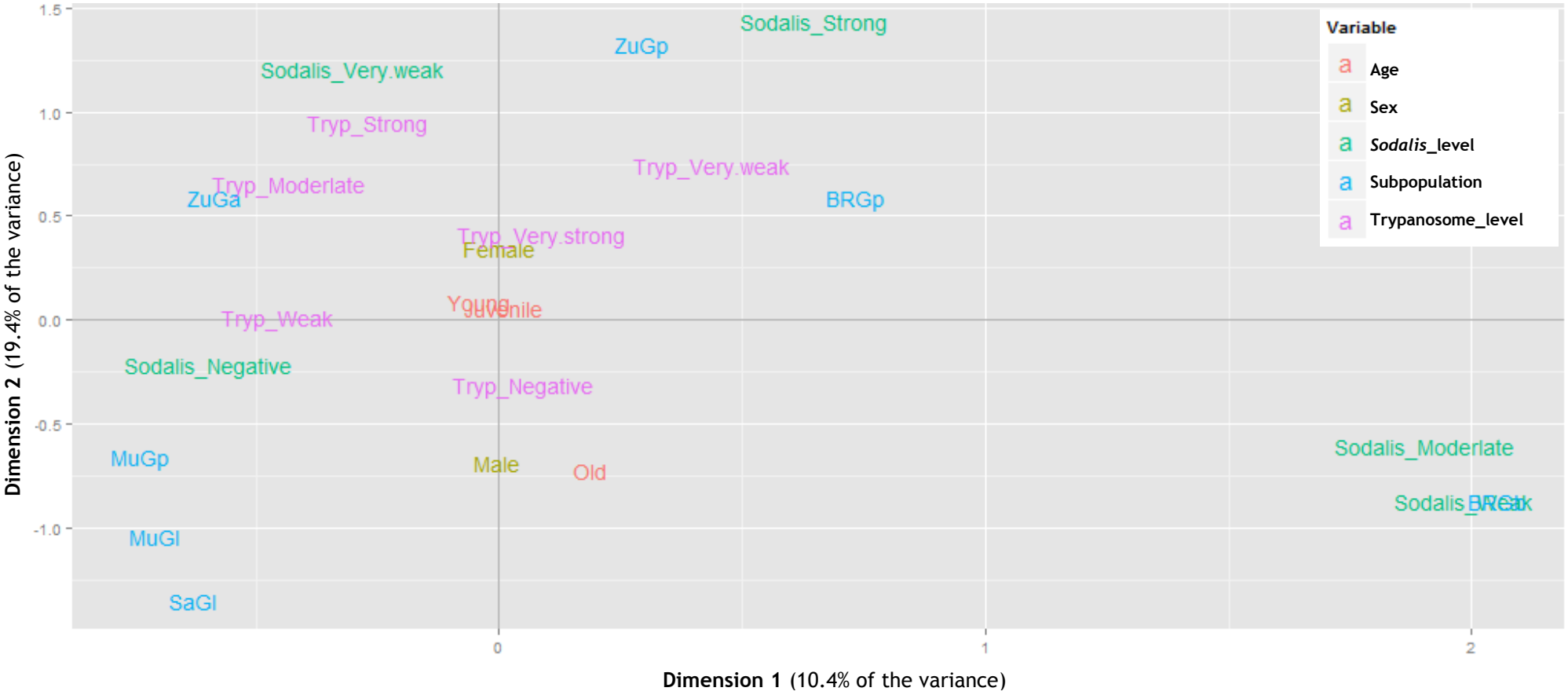



**Figure B.4** Dimension 1 and 2 of the MCA B1 for explaining association of level of trypanosome positive in tsetse flies from four sites (Buffalo Ridge, Zungu Luka, Mukinyo and Sampu) with species (*G. austeni* (Ga), *G. brevipalpis* (Gb), *G. longipennis* (Gl) and *G. pallidipes* (Gp)), sex (male and female), age (young, juvenile and old), *Sodalis* level (negative, very weak, weak, moderate, strong and very strong positive), as implemented in the R package FactoMineR statistical software programme

# Appendix C Supplementary Information for chapter 4

## C.1. Host determination of blood meal analysis from *G. pallidipes* based on direct sequencing of Cb PCR products.

| Samples | Host identification | Size (bp) | % Query cover | % Identity | Accession no. |
|---|---|---|---|---|---|
| ***G. pallidipes* from Buffalo Ridge** | | | | | |
| **African Buffalo** | | | | | |
| BRGp 1AB | African buffalo | 299 | 100 | 99 | JQ235516.1 |
| BRGp 9AB | African buffalo | 261 | 100 | 100 | JQ235516.1 |
| BRGp 10AB | African buffalo | 274 | 100 | 100 | JQ235527.1 |
| BRGp 12AB | African buffalo | 304 | 100 | 100 | JQ235544.1 |
| BRGp 16AB | African buffalo | 330 | 100 | 100 | JQ235544.1 |
| BRGp 23AB | African buffalo | 248 | 100 | 100 | JQ235544.1 |
| BRGp 25AB | African buffalo | 330 | 100 | 100 | JQ235544.1 |
| BRGp 32AB | African buffalo | 359 | 100 | 100 | JQ235544.1 |
| BRGp 41AB | African buffalo | 340 | 100 | 99 | JQ235516.1 |
| BRGp 49AB | African buffalo | 348 | 99 | 99 | JQ235526.1 |
| BRGp 53AB | African buffalo | 359 | 100 | 99 | JQ235544.1 |
| BRGp 64AB | African buffalo | 297 | 100 | 100 | JQ235544.1 |
| BRGp 77AB | African buffalo | 348 | 100 | 99 | JQ235544.1 |
| BRGp 96AB* | African buffalo | 339 | 100 | 99 | JQ235544.1 |
| BRGp 97AB | African buffalo | 339 | 100 | 99 | JQ235516.1 |
| BRGp 99AB | African buffalo | 337 | 99 | 100 | JQ235544.1 |
| BRGp 137AB | African buffalo | 336 | 100 | 99 | JQ235527.1 |
| BRGp 139AB | African buffalo | 339 | 100 | 99 | JQ235526.1 |
| BRGp 151AB | African buffalo | 348 | 100 | 99 | JQ235516.1 |
| **African elephants** | | | | | |
| BRGp 13AB | African elephants | 316 | 99 | 99 | JQ438758.1 |
| BRGp 70AB | African elephants | 342 | 100 | 100 | JQ438758.1 |
| **Antelopes** | | | | | |
| BRGp 17AB | bushbucks | 359 | 99 | 98 | HQ641317.1 |
| BRGp 74AB | bushbucks | 359 | 100 | 98 | JN632707.1 |
| **Humans** | | | | | |
| BRGp 66AB | humans | 359 | 99 | 99 | KC252520.1 |
| BRGp 118AB | humans | 334 | 100 | 99 | KP900938.1 |
| **Warthogs** | | | | | |
| BRGp 33AB | warthogs | 322 | 99 | 99 | FJ785389.1 |
| BRGp 52AB | warthogs | 297 | 95 | 95 | FJ785389.1 |

| Samples | Host identification | Size (bp) | % Query cover | % Identity | Accession no. |
|---|---|---|---|---|---|
| ***G. pallidipes* from Zungu Luka** | | | | | |
| **Warthogs** | | | | | |
| ZuGp 126AB | African elephants | 359 | 100 | 99 | JQ438674.1 |
| **Antelopes** | | | | | |
| ZuGp 10AB | bushbucks | 314 | 100 | 99 | JN632707.1 |
| ZuGp 58AB | bushbucks | 350 | 99 | 99 | JN632707.1 |
| ZuGp 61AB | bushbucks | 350 | 98 | 99 | JN632707.1 |
| ZuGp 107AB | bushbucks | 339 | 100 | 99 | JN632707.1 |
| **Humans** | | | | | |
| ZuGp 87AB | humans | 338 | 100 | 99 | KP900938.1 |
| ***G. pallidipes* from Mukinyo** | | | | | |
| **African buffalo** | | | | | |
| MuGp 1AB | African buffalo | 332 | 100 | 100 | JQ235538.1 |
| MuGp 4AB | African buffalo | 297 | 100 | 100 | JQ235538.1 |
| MuGp 7AB | African elephants | 359 | 99 | 100 | AY768855.1 |
| MuGp 15AB | African buffalo | 298 | 100 | 99 | JQ235516.1 |
| MuGp 16AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 17AB | African buffalo | 327 | 100 | 99 | JQ235544.1 |
| MuGp 18AB | African buffalo | 359 | 100 | 100 | JQ235538.1 |
| MuGp 19AB | African buffalo | 303 | 100 | 100 | JQ235538.1 |
| MuGp 21AB | African buffalo | 360 | 100 | 99 | JQ235544.1 |
| MuGp 28AB | African buffalo | 359 | 100 | 99 | JQ235544.1 |
| MuGp 38AB | African buffalo | 337 | 100 | 99 | JQ235538.1 |
| MuGp 42AB | African buffalo | 301 | 100 | 100 | JQ235538.1 |
| MuGp 43AB | African buffalo | 348 | 100 | 99 | JQ235538.1 |
| MuGp 54AB | African buffalo | 297 | 100 | 100 | JQ235538.1 |
| MuGp 59AB | African buffalo | 337 | 100 | 99 | JQ235539.1 |
| MuGp 60AB | African buffalo | 337 | 100 | 99 | JQ235538.1 |
| MuGp 61AB | African buffalo | 337 | 100 | 99 | JQ235544.1 |
| MuGp 62AB | African buffalo | 335 | 100 | 99 | JQ235516.1 |
| MuGp 64AB | African buffalo | 341 | 100 | 98 | JQ235538.1 |
| MuGp 66AB | African buffalo | 333 | 100 | 99 | JQ235516.1 |
| MuGp 67AB | African buffalo | 332 | 100 | 100 | JQ235516.1 |
| MuGp 68AB | African buffalo | 287 | 100 | 99 | JQ235544.1 |
| MuGp 71AB | African buffalo | 342 | 100 | 97 | JQ235516.1 |
| MuGp 72AB | African buffalo | 338 | 100 | 99 | JQ235538.1 |
| MuGp 73AB | African buffalo | 335 | 100 | 99 | JQ235538.1 |
| MuGp 76AB | African buffalo | 337 | 100 | 99 | JQ235538.1 |
| MuGp 79AB | African buffalo | 339 | 100 | 99 | JQ235538.1 |
| MuGp 80AB | African buffalo | 289 | 100 | 99 | JQ235538.1 |

| Samples | Host identification | Size (bp) | % Query cover | % Identity | Accession no. |
|---|---|---|---|---|---|
| ***G. pallidipes* from Mukinyo** | | | | | |
| **African buffalo** | | | | | |
| MuGp 81AB | African buffalo | 335 | 100 | 99 | JQ235516.1 |
| MuGp 84AB | African buffalo | 337 | 100 | 99 | JQ235538.1 |
| MuGp 89AB | African buffalo | 340 | 100 | 99 | JQ235516.1 |
| MuGp 97AB | African buffalo | 339 | 100 | 99 | JQ235516.1 |
| MuGp 107AB | African buffalo | 338 | 100 | 99 | JQ235516.1 |
| MuGp 110AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 135AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 136AB | African buffalo | 349 | 100 | 99 | JQ235516.1 |
| MuGp 139AB | African buffalo | 338 | 100 | 99 | JQ235538.1 |
| MuGp 143AB | African buffalo | 359 | 100 | 99 | JQ438758.1 |
| MuGp 145AB | African buffalo | 359 | 99 | 99 | JQ235538.1 |
| MuGp 147AB | African buffalo | 289 | 100 | 100 | JQ235538.1 |
| MuGp 148AB | African buffalo | 358 | 100 | 99 | JQ235538.1 |
| MuGp 149AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 151AB | African buffalo | 317 | 100 | 99 | JQ235538.1 |
| MuGp 152AB | African buffalo | 359 | 99 | 99 | JQ235516.1 |
| MuGp 153AB | African buffalo | 359 | 99 | 99 | JQ235538.1 |
| MuGp 154AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 158AB | African buffalo | 359 | 99 | 99 | JQ438758.1 |
| MuGp 161AB | African buffalo | 292 | 100 | 100 | JQ235538.1 |
| MuGp 168AB | African buffalo | 349 | 99 | 99 | JQ235516.1 |
| MuGp 170AB | African buffalo | 302 | 100 | 99 | JQ235538.1 |
| MuGp 176AB | African buffalo | 359 | 99 | 99 | JQ235538.1 |
| MuGp 179AB | African buffalo | 359 | 99 | 99 | JQ235538.1 |
| MuGp 180AB | African buffalo | 359 | 100 | 99 | JQ235544.1 |
| MuGp 181AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 184AB | African buffalo | 292 | 100 | 99 | JQ235516.1 |
| MuGp 186AB | African buffalo | 347 | 99 | 99 | JQ235538.1 |
| MuGp 187AB | African buffalo | 350 | 100 | 99 | JQ235516.1 |
| MuGp 189AB | African buffalo | 331 | 100 | 99 | JQ235538.1 |
| MuGp 193AB | African buffalo | 338 | 100 | 99 | JQ235538.1 |
| MuGp 195AB | African buffalo | 338 | 100 | 99 | JQ235538.1 |
| MuGp 203AB | African buffalo | 339 | 100 | 99 | JQ235516.1 |
| MuGp 206AB | African buffalo | 339 | 100 | 99 | JQ235538.1 |
| MuGp 207AB | African buffalo | 337 | 100 | 99 | JQ235538.1 |
| MuGp 210AB | African buffalo | 287 | 100 | 99 | JQ235547.1 |
| MuGp 212AB | African buffalo | 339 | 100 | 99 | JQ235516.1 |
| MuGp 215AB | African buffalo | 342 | 100 | 99 | JQ235516.1 |
| MuGp 229AB | African buffalo | 337 | 100 | 99 | JQ235544.1 |

| Samples | Host identification | Size (bp) | % Query cover | % Identity | Accession no. |
|---|---|---|---|---|---|
| ***G. pallidipes* from Mukinyo** | | | | | |
| **African buffalo** | | | | | |
| MuGp 232AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 233AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 234AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 237AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 240AB | African buffalo | 333 | 100 | 99 | JQ235516.1 |
| MuGp 241AB | African buffalo | 338 | 100 | 99 | JQ235516.1 |
| MuGp 243AB | African buffalo | 359 | 100 | 99 | JQ235544.1 |
| MuGp 244AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 245AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 246AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 248AB | African buffalo | 337 | 100 | 99 | JQ235516.1 |
| MuGp 250AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 252AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 254AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 260AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 265AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 266AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 268AB | African buffalo | 359 | 100 | 99 | JQ235544.1 |
| MuGp 269AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 273AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 279AB | African buffalo | 359 | 100 | 99 | JQ235538.1 |
| MuGp 288AB | African buffalo | 359 | 100 | 99 | JQ235544.1 |
| MuGp 289AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 292AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| MuGp 293AB | African buffalo | 359 | 100 | 99 | JQ235516.1 |
| **African elephants** | | | | | |
| MuGp 51AB | African elephants | 336 | 100 | 100 | AY768855.1 |
| MuGp 53AB | African elephants | 338 | 99 | 88 | JQ438758.1 |
| MuGp 65AB | African elephants | 350 | 99 | 99 | AY768855.1 |
| MuGp 70AB | African elephants | 359 | 99 | 99 | AY768855.1 |
| MuGp 77AB | African elephants | 285 | 100 | 99 | AY768855.1 |
| MuGp 83AB | African elephants | 337 | 100 | 100 | AY768855.1 |
| MuGp 93AB | African elephants | 359 | 100 | 99 | AY768855.1 |
| MuGp 95AB | African elephants | 359 | 100 | 99 | AY768855.1 |
| MuGp 98AB | African elephants | 359 | 99 | 99 | AY768855.1 |
| MuGp 111AB | African elephants | 359 | 100 | 100 | AY768855.1 |
| MuGp 138AB | African elephants | 359 | 99 | 100 | AY768855.1 |
| MuGp 157AB | African elephants | 359 | 99 | 99 | JQ438758.1 |
| MuGp 159AB | African elephants | 359 | 100 | 100 | AY768855.1 |

| Samples | Host identification | Size (bp) | % Query cover | % Identity | Accession no. |
|---|---|---|---|---|---|
| ***G. pallidipes* from Mukinyo** | | | | | |
| **African elephants** | | | | | |
| MuGp 163AB | African elephants | 359 | 99 | 99 | JQ438758.1 |
| MuGp 171AB | African elephants | 302 | 100 | 100 | AY768855.1 |
| MuGp 172AB | African elephants | 359 | 99 | 100 | AY768855.1 |
| MuGp 177AB | African elephants | 359 | 99 | 99 | JQ438758.1 |
| MuGp 178AB | African elephants | 359 | 99 | 99 | JQ438758.1 |
| MuGp 182AB | African elephants | 359 | 99 | 100 | AY768855.1 |
| MuGp 183AB | African elephants | 359 | 99 | 100 | AY768855.1 |
| MuGp 198AB | African elephants | 339 | 100 | 100 | AY768855.1 |
| MuGp 223AB | African elephants | 339 | 100 | 99 | JQ438758.1 |
| MuGp 224AB | African elephants | 359 | 100 | 100 | AY768855.1 |
| MuGp 251AB | African elephants | 359 | 100 | 100 | AY768855.1 |
| MuGp 259AB | African elephants | 359 | 100 | 99 | AY768855.1 |
| MuGp 261AB | African elephants | 359 | 100 | 100 | JQ438713.1 |
| MuGp 278AB | African elephants | 359 | 100 | 99 | AY768855.1 |
| MuGp 281AB | African elephants | 359 | 100 | 99 | AY768855.1 |
| **Antelopes** | | | | | |
| MuGp 88AB | lesser kudu | 338 | 99 | 99 | EF536356.1 |
| **Giraffes** | | | | | |
| MuGp 144AB | giraffes | 359 | 99 | 99 | EU088336.1 |
| MuGp 166AB | giraffes | 359 | 98 | 99 | EU088336.1 |
| MuGp 227AB | giraffes | 356 | 100 | 99 | EU088336.1 |
| MuGp 263AB | giraffes | 338 | 99 | 99 | EU088335.1 |
| **Humans** | | | | | |
| MuGp 13AB | humans | 361 | 100 | 98 | KC252520.1 |
| MuGp 90AB | humans | 334 | 100 | 99 | KP900938.1 |
| **Hyenas** | | | | | |
| MuGp 112AB | spotted hyenas | 340 | 100 | 100 | AF511064.1 |
| **Warthogs** | | | | | |
| MuGp 75AB | warthogs | 350 | 98 | 99 | DQ409327.1 |
| MuGp 255AB | warthogs | 346 | 100 | 99 | DQ409327.1 |
| MuGp 258AB | warthogs | 340 | 99 | 99 | DQ409327.1 |
| MuGp 262AB | warthogs | 282 | 96 | 99 | FJ785389.1 |

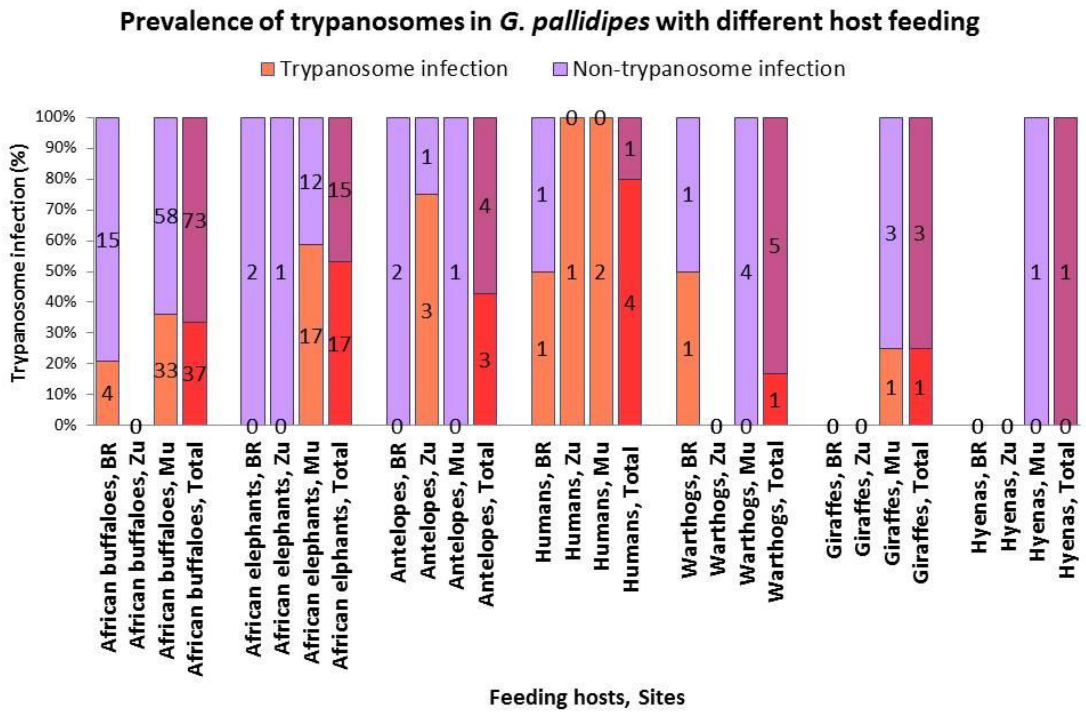
* heterozygous for two different African buffalo genotypes: JQ235516.1 and JQ235544.1.

## C.2. Host determination of *G. pallidipes* (Gp) from Buffalo Ridge (BR), Zungu Luka (Zu) and Mukinyo (Mu) using cloning and sequencing methods.

| Samples | Host determination (number of plasmids) | Size (bp) | % Query cover | % identity | Accession no. |
|---|---|---|---|---|---|
| ***G. pallidipes* from Buffalo Ridge (BRGp)** | | | | | |
| BRGp 34AB | humans (4) | 359 | 100 | 99 | KP900938.1 |
| | goats (1) | 359 | 99 | 99 | FM205715. |
| | antelopes (1) | 359 | 99 | 99 | JN632707.1 |
| BRGp108AB | mouse (2) | 359 | 99 | 99 | AP014941.1 |
| | goats (2) | 359 | 99 | 99 | FM205715.1 |
| | humans (1) | 359 | 100 | 99 | KP900938.1 |
| | African buffalo (1) | 359 | 100 | 99 | JQ235516.1 |
| BRGp 109AB | African elephants (3) | 359 | 100 | 99 | JQ438758.1 |
| | humans (1) | 359 | 100 | 99 | KP900938.1 |
| | mouse(1) | 359 | 99 | 95 | AB819918.1 |
| | goats (1) | 359 | 99 | 99 | FM205715.1 |
| BRGp 126AB | humans (2) | 359 | 100 | 99 | KP900938.1 |
| | goats (2) | 359 | 99 | 99 | FM205715.1 |
| | antelopes (1) | 359 | 97 | 96 | T290893.1 |
| BRGp 143AB | humans (3) | 359 | 100 | 99 | KP900938.1 |
| | goats (2) | 359 | 99 | 99 | FM205715.1 |
| | mouse (1) | 359 | 99 | 99 | AP014941.1 |
| BRGp 33AB | warthogs* (6) | 359 | 96 | 99 | FJ785389.1 |
| BRGp 52AB | warthogs* (6) | 359 | 96 | 99 | FJ785389.1 |
| ***G. pallidipes* from Zungu Luka (ZuGp)** | | | | | |
| ZuGp 2AB | humans (2) | 359 | 100 | 99 | KP900938.1 |
| | goat (2) | 359 | 99 | 99 | FM205715.1 |
| | mouse (1) | 359 | 99 | 99 | AP014941.1 |
| | warthogs (1) | 359 | 96 | 99 | FJ785389.1 |
| ZuGp 96AB | antelopes (4) | 359 | 99 | 99 | JN632707.1 |
| | humans (2) | 359 | 100 | 99 | KP900938.1 |
| ZuGp 102AB | chickens (3) | 359 | 100 | 100 | DQ512917.1 |
| | humans (2) | 359 | 100 | 99 | KP900938.1 |
| | fruit flies** (1) | 339 | 94 | 87 | BK006337.1 |
| ZuGp 54AB | antelopes (5) | 359 | 99 | 99 | JN632707.1 |
| ZuGp 104AB | warthogs (6) | 359 | 96 | 99 | FJ785389.1 |
| ***G. pallidipes* from Mukinyo (MuGp)** | | | | | |
| MuGp 9AB | African elephants (4) | 359 | 100 | 99 | JQ438758.1 |
| | African buffalo (2) | 359 | 100 | 99 | JQ235538.1 |
| MuGp 287AB | African elephant (2) | 359 | 100 | 99 | JQ235538.1 |
| | African buffalo (1) ) | 359 | 100 | 99 | JQ235538.1 |
| **Control** | | | | | |
| buccal cells | human*** | 359 | 100 | 99 | KC252520.1 |

* a sequence, which was identified as “single host feeding”; ** *Drosophila permilis*; ****Homo sapiens* isolate TDS41 mitochondrion, complete genome, Taiwan.

## C.3. Proportions of trypanosome infection of *G. pallidipes* that fed on different single blood meal sources.




Numbers of trypanosome positive and negative flies are represented in each bar.

## C.4. Summary of available wildlife in the Shimba Hills and Nguruman.

Light blue colour of geographic regions indicates that hosts are available in the area. Numbers of available hosts are shown for some animals.

| Wildlife | Geographic regions | |
|---|---|---|
| | the Shimba Hills* | Nguruman** |
| African Buffalo | **67** | 52 |
| African elephants | NI | 48 |
| Maasai giraffes | 3 | **237** |
| antelopes (Grants Gazelle, Common Wildebeest, Cape Eland, Impala, Fringe - Eared Oryx, Gerenuk, Bushbuck, Dik Dik, Common Waterbuck) | 50 | **3,315** |
| warthogs | 81 | 1 |
| primates (baboons, monkeys) | 2 (monkeys) | 121 (baboons) |
| many avian species (Ostrich Eagle, Falcon, Guinea-fowl,; Honey guide, Hornbill, Quail, Sunbird, Uluguru) | NI | 167 (ostriches) |
| lions | NI | NI |
| shrew (Oguge *et al.*, 2009) | NI | NI |
| rodent: East African Arvicanthis, Hinde's Pouched Rat, Black-tailed gerbil, Grammomys, Large Savannah African Dormouse, Typical Lemniscomys, Natal Mastomys, Southern African Pygmy Mouse, Gray-bellied Mouse, East African Praomys (Oguge *et al.*, 2009) | NI | |
| spotted hyenas | NI | |
| mongeese | NI | |
| bats | NI | |
| shrews | NI | |
| reptiles (Python, Cobra, Lizard, monitor lizard, and Gecko) | NI | |
| leopards | NI | |
| Genets | NI | |
| Civet cats/ Serval cats | NI | |
| Common zebras | | **2,399** |

* data provided from http://www.kws.go.ke/parks/parks_reserves/SHNR.html; http://www.kws.go.ke/content/shimba-hills-national-reserve; and private information from Priscillar Mumo Mutungi.

** data provided from http://www.kws.org/export/sites/kws/info/publications/census_reports/Amboseli_West_kili_Magadi_Natron_2010_census_report.pdf

NI = number of animals were not indicated.

# List of References


Abd-Alla, A. M., Bergoin, M., Parker, A. G., Maniania, N. K., Vlak, J. M., Bourtzis, K., Boucias, D. G. & Aksoy, S. 2013. Improving Sterile Insect Technique (SIT) for tsetse flies through research on their symbionts and pathogens. *Journal of Invertebrate Pathology,* 112 Suppl, S2-10.

Abubakar, L. U., Bulimo, W. D., Mulaa, F. J. & Osir, E. O. 2006. Molecular characterization of a tsetse fly midgut proteolytic lectin that mediates differentiation of African trypanosomes. *Insect Biochemistry and Molecular Biology,* 36, 344-352.

Adam, Y., Cecchi, G., Kgori, P. M., Marcotty, T., Mahama, C. I., Abavana, M., Anderson, B., Paone, M., Mattioli, R. & Bouyer, J. 2013. The sequential aerosol technique: a major component in an integrated strategy of intervention against Riverine Tsetse in Ghana. *PLoS Neglected Tropical Disease,* 7, e2135.

Adams, E. R., Hamilton, P. B. & Gibson, W. C. 2010. African trypanosomes: celebrating diversity. *Trends in Parasitology,* 26, 324-328.

Adams, E. R., Malele, I. I., Msangi, A. R. & Gibson, W. C. 2006. Trypanosome identification in wild tsetse populations in Tanzania using generic primers to amplify the ribosomal RNA ITS-1 region. *Acta Tropica,* 100, 103-109.

Adlington, D., Randolph, S. E. & Rogers, D. J. 1996. Flying to feed or flying to mate: gender differences in the flight activity of tsetse (*Glossina palpalis*). *Physiological Entomology,* 21, 85-92.

Afewerk, Y., Clausen, P. H., Abebe, G., Tilahun, G. & Mehlitz, D. 2000. Multiple-drug resistant *Trypanosoma congolense* populations in village cattle of Metekel district, north-west Ethiopia. *Acta Tropica,* 76, 231-238.

Ahmed, H. A., Picozzi, K., Welburn, S. C. & MacLeod, E. T. 2013. A comparative evaluation of PCR-based methods fro specific determination of African animals trypanosomes in Ugandan cattle. *Parasites & Vectors,* 6, 316.

Akane, A., MAtsubara, K., Nakamura, H., Takahashi, S. & Kimura, K. 1994. Identification of the heme compound copurified with deoxyribonucleic acid (DNA) from bloodstains, a major inhibitor of Polymerase Chain Reaction (PCR) amplification. *Journal of Forensic Sciences,* 39, 362-372.

Akoda, K. 2009. *Effect of nutritional stress on the tsetse fly's vector competence and its implications on trypanosome transmission in the field.* Ghent University.

Akoda, K., Van den Bossche, P., Lyaruu, E. A., De Deken, R., Marcotty, T., Coosemans, M. & Van den Abbeele, J. 2009a. Maturation of a *Trypanosoma brucei* infection to the infectious metacyclic stage is enhanced in nutritionally stressed tsetse flies. *Journal of Medical Entomology,* 46, 1446-1449.

Akoda, K., Van Den Bossche, P., Marcotty, T., Kubi, C., Coosemans, M., De Deken, R. & Van Den Abbeele, J. 2009b. Nutritional stress affects the tsetse fly's immune gene expression. *Medical and Veterinary Entomology,* 23, 195-201.

Aksoy, E., Telleria, E. L., Echodu, R., Wu, Y., Okedi, L. M., Weiss, B. L., Aksoy, S. & Caccone, A. 2014. Analysis of multiple tsetse fly populations in Uganda reveals limited diversity and species-specific gut microbiota. *Applied and Environmental Microbiology,* 80, 4301-12.

Aksoy, S. 1995a. Molecular analysis of the endosymbionts of tsetse flies: 16S rDNA locus and over-expression of a chaperonin. *Insect Molecular Biology,* 4, 23-29.

Aksoy, S. 1995b. *Wigglesworthia* gen. nov. and *Wigglesworthia glossinidia sp.* nov., Taxa consisting of the mycetocyte-associated, primary endosymbionts of tsetse flies. *International Journal of Systematic and Evolutionary Microbiology,* 45**,** 848-851.

Aksoy, S. 2000. Tsetse – A Haven for Microorganisms. *Parasitology Today,* 16**,** 114-118.

Aksoy, S. 2003. Control of tsetse flies and trypanosomes using molecular genetics. *Veterinary Parasitology,* 115**,** 125-145.

Aksoy, S., Chen, X. & Hypsa, V. 1997. Phylogeny and potential transmission routes of midgut-associated endosymbionts of tsetse (Diptera: *Glossinidae*). *Insect Molecular Biology,* 6**,** 183-190.

Aksoy, S., Gibson, W. C. & Lehane, M. J. 2003. Interactions between tsetse and trypanosomes with implications for the control of trypanosomiasis. *Advances in Parasitology.* Academic Press.

Aksoy, S., O'Neill, S. L., Maudlin, I., Dale, C. & Robinson, A. S. 2001. Prospects for control of African trypanosomiasis by tsetse vector manipulation. *Trends in Parasitology,* 17**,** 29-35.

Al-Soud, W. A., Jönsson, L. J. & Rådström, P. 2000. Identification and characterization of immunoglobulin G in blood as a major inhibitor of diagnostic PCR. *Journal of Clinical Microbiology,* 38**,** 345-350.

Al-Soud, W. A. & Radstrom, P. 2001. Purification and characterization of PCR-inhibitory components in blood cells. *Journal of Clinical Microbiology,* 39**,** 485-93.

Alaeddini, R. 2012. Forensic implications of PCR inhibition—a review. *Forensic Science International: Genetics,* 6**,** 297-305.

Alam, U., Galvani, A., Aksoy, S., Medlock, J., Brelsfoard, C., Pais, R., Lohs, C., Balmand, S., Carnogursky, J., Heddi, A. & Takac, P. 2011. *Wolbachia* symbiont infections induce strong cytoplasmic incompatibility in the tsetse fly *Glossina morsitans*. *PLoS Pathogens,* 7**,** e1002415.

Ali, B. H. & Hassan, T. 1986. Some observations on the toxicosis of isometamidium chloride (samorin) in camels *Veterinary & Human Toxicology,* 28**,** 424-426.

Alibu, V. P., Enyaru, J. C. K., Matovu, E., Malele, I. I., Chisi, J. E., Mbongo, N., Mansinsa, P., Intisar, E. R., Mohammed, Y., Abdelrahman, M. M., Ochi, E. B. & Lukaw, Y. S. 2015. Molecular xenomonitoring of trypanosomes in tsetse flies *Journal of Parasitology and Vector Biology,* 7**,** 108-114.

Allsopp, R. 1984. Control of tsetse flies (Diptera: *Glossinidae*) using insecticides: a review and future prospects. *Bulletin of Entomological Research,* 74**,** 1-23.

Allsopp, R., Baldry, D. A. T. & Rodrigues, C. 1972. The influence of game animals on the distribution and feeding habits of *Glossina pallidipes* in the Lambwe Valley. *Bulletin of the World Health Organization,* 47**,** 795-809.

Anderson, N. E., Mubanga, J., Fevre, E. M., Picozzi, K., Eisler, M. C., Thomas, R. & Welburn, S. C. 2011. Characterisation of the wildlife reservoir community for human and animal trypanosomiasis in the Luangwa Valley, Zambia. *PLoS Neglected Tropical Diseases,* 5**,** e1211.

Angara, T.-E. E., Ismail, A. A. & Ibrahim, A. M. 2014. An overview on the economic impacts of animal trypanosomiasis. *Global Journal for Research Analysis,* 3**,** 275-276.

Attardo, G. M., Lohs, C., Heddi, A., Alam, U. H., Yildirim, S. & Aksoy, S. 2008. Analysis of milk gland structure and function in *Glossina morsitans*: Milk protein production, symbiont populations and fecundity. *Journal of Insect Physiology,* 54**,** 1236-1242.

Auty, H., Cleaveland, S., Welburn, S. C., Anderson, N. E., Picozzi, K., Lembo, T., Mubanga, J., Hoare, R., Fyumagwa, R. D., Mable, B. & Hamill, L. 2012. Trypanosome diversity in wildlife species from the serengeti and Luangwa Valley ecosystems. *PLoS Neglected Tropical Diseases,* 6**,** e1828.

Balasegaram, M., Harris, S., Checchi, F., Ghorashian, S., Hamel, C. & Karunakara, U. 2006. Melarsoprol versus eflornithine for treating late-stage Gambian trypanosomiasis in the Republic of the Congo. *Bulletin of The World Health Organization,* 84**,** 783-791.

Balmand, S., Lohs, C., Aksoy, S. & Heddi, A. 2013. Tissue distribution and transmission routes for the tsetse fly endosymbionts. *Journal of Invertebrate Pathology,* 112 Suppl**,** S116-22.

Barry, J. D. & Turner, C. M. R. 1991. The dynamics of antigenic variation and growth of African trypanosomes. *Parasitology Today,* 7**,** 207-211.

Batista, J. S., Rodrigues, C. M. F., García, H. A., Bezerra, F. S. B., Olinda, R. G., Teixeira, M. M. G. & Soto-Blanco, B. 2011. Association of *Trypanosoma vivax* in extracellular sites with central nervous system lesions and changes in cerebrospinal fluid in experimentally infected goats. *Veterinary Research,* 42**,** 1-7.

Baylis, M. 1996. Effect of defensive behaviour by cattle on the feeding success and nutritional state of the tsetse fly, *Glossina pallidipes* (Diptera: *Glossinidae*). *Bulletin of Entomological Research,* 86**,** 329-336.

Baylis, M. & Nambiro, C. O. 1993. The effect of cattle infection by *Trypanosoma congolense* on the attraction, and feeding success, of the tsetse fly *Glossina pallidipes*. *Parasitology,* 106 ( Pt 4)**,** 357-361.

Belete, H., Tikubet, G., Petros, B., Oyibo, W. A. & Otigbuo, I. N. 2004. Control of human African trypanosomiasis: trap and odour preference of tsetse flies (*Glossina morsitans submorsitans*) in the upper Didessa river valley of Ethiopia. *Tropical Medicine & International Health,* 9**,** 710-714.

Bernhard, S. C., Nerima, B., Maser, P. & Brun, R. 2007. Melarsoprol- and pentamidine-resistant *Trypanosoma brucei rhodesiense* populations and their cross-resistance. *International Journal for Parasitology,* 37**,** 1443-1448.

Berret, M. P. 2001. Veterinary link to drug resistance in human African trypanosomiasis ? *The Lancet,* 358**,** 603-604.

Bett, B., Irungu, P., Nyamwaro, S. O., Murilla, G., Kitala, P., Gathuma, J., Randolph, T. F. & McDermott, J. 2008. Estimation of tsetse challenge and its relationship with trypanosomosis incidence in cattle kept under pastoral production systems in Kenya. *Veterinary Parasitology,* 155**,** 287-298.

Bitew, M., Amide, Y., Zenebe, T. & Degefu, H. 2011. Trypanosome infection rate in *Glossina pallidipes* and *Glossina fuscipes fuscipes* in Gojeb Valley, Southwest Ethiopia. *Global Veterinaria,* 6**,** 131-135.

Boakye, D. A., Tang, J., P., T., Merriweather, A. & Unnasch, T. R. 1999. Identification of bloodmeals in haematophagous Diptera by cytochrome B heterodupliex analysis. *Medical and Veterinary Entomology,* 13**,** 282-287.

Bonnet, J., Boudot, C. & Courtioux, B. 2015. Overview of the diagnostic methods used in the field for Human African Trypanosomiasis: what could change in the next years? *BioMed Research International,* 2015**,** 10.

Botero, A., Thompson, C. K., Peacock, C. S., Clode, P. L., Nicholls, P. K., Wayne, A. F., Lymbery, A. J. & Thompson, R. C. A. 2013. Trypanosomes genetic diversity, polyparasitism and the population decline of the critically endangered Australian marsupial, the brush tailed bettong or

woylie (Bettongia penicillata). *International Journal for Parasitology: Parasites and Wildlife,* 2**,** 77-89.

Boulanger, N., Hamilton, J. V., Vovelle, F., Brun, R., Lehane, M. J. & Bule, P. 2002. Epithelial innate immunity: a novel antimicrobial peptide with antiparasitic activity in the blood-sucking insect *Stomoxys calcitrans*. *The Journal of Biological Chemistry,* 277**,** 49921–49926.

Bouyer, J., Pruvot, M., Bengaly, Z., Guerin, P. M. & Lancelot, R. 2007. Learning influences host choice in tsetse. *Biology Letters,* 3**,** 113-116.

Boyle, J. A. 1971. Effect of blood intake of *Glossina austeni* Newst. on pupal weights in successive reproductive cycles. *Bulletin of Entomological Research,* 61**,** 1-5.

Brady, J. & Shereni, W. 1988. Landing responses of the tsetse fly *Glossina morsitans morsitans* Westwood and the stable fly *Stomoxys calcitrans* (L.) (Diptera: *Glossinidae* and *Muscidae*) to black-and-white patterns: a laboratory study. *Bulletin of Entomological Research,* 78**,** 301-311.

Bray, P. G., Barrett, M. P., Ward, S. A. & de Koning, H. P. 2003. Pentamidine uptake and resistance in pathogenic protozoa: past, present and future. *Trends in Parasitology,* 19**,** 232-239.

Bridges, D. J., Gould, M. K., Nerima, B., Maser, P., Burchmore, R. J. S. & de Koning, H. P. 2007. Loss of the high-affinity pentamidine transporter is responsible for high levels of cross-resistance between arsenical and diamidine drugs in African trypanosomes. *Molecular Pharmacology,* 71**,** 1098-1108.

Brightwell, R., Dransfield, R. D. & Kyorku, C. 1991. Development of a low-cost tsetse trap and odour baits for *Glossina pallidipes* and *G. longipennis* in Kenya. *Medical and Veterinary Entomology,* 5.

Brightwell, R., Dransfield, R. D., Stevenson, P. & Williams, B. 1997. Changes over twelve years in populations of *Glossina pallidipes* and *Glossina longipennis* (Diptera: *Glossinidae*) subject to varying trapping pressure at Nguruman, south-west Kenya. *Bulletin of Entomological Research,* 87**,** 349-370.

Brightwell, R., Dransfield, R. D. & Williams, B. G. 1992. Factors affecting seasional dispersal of the tsetse flies *Glossina pallidipes* and *G. longipennis* (Diptera: Glossinidae) at Nguruman, south-west Kenya. *Bulletin of Entomological Research,* 82**,** 167-182.

Bruce, D., Hamerton, A. E., Bateman, H. R. & Mackie, F. P. 1910. The development of trypanosomes in tsetse flies. *Proceedings of The Royal Society of London. Series B, Containing Papers of a Biological Character,* 82**,** 368-388.

Brun, R. 1979. Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. *Acta tropica,* 36**,** 289.

Brun, R., Blum, J., Chappuis, F. & Burri, C. 2009. Human African trypanosomiasis. *The Lancet,* 375**,** 148-159.

Bursell, E. 1957. The effect of humidity on the activity of tsetse flies. *Journal of Experimantal biology,* 34**,** 42-51.

Bursell, E. & Taylor, P. 1980. An energy budget for *Glossina* (Diptera: *Glossinidae*). *Bulletin of Entomological Research,* 70**,** 187-196.

Büscher, P. & Lejon, V. 2004. Diagnosis of human African trypanosomiasis. *In:* Maudlin, I., Holmes, P. H. & Miles, M. A. (eds.) *The trypanosomiases*. 1 ed. Oxfordshire, UK: CABI.

Castro, J. A., Mecca, M. M. D. & Bartel, L. C. 2006. Toxic side effects of drugs used treat Chagas' disease (American trypanosomiasis). *Human and Experimental Toxicology,* 25**,** 471-479.

Cattand, P., Jannin, J. & Lucas, P. 2001. Sleeping sickness surveillance: an essential step towards elimination. *Tropical medicine & international health,* 6**,** 348-361.

Cecchi, G., Mattioli, R. C., Slingenbergh, J. & De La Rocque, S. 2008. Land cover and tsetse fly distributions in sub-Saharan Africa. *Medical and Veterinary Entomology,* 22**,** 364-373.

Cecchi, G., Paone, M., Argilés Herrero, R., Vreysen, M. J. B. & Mattioli, R. C. 2015. Developing a continental atlas of the distribution and trypanosomal infection of tsetse flies (*Glossina species*). *Parasites & Vectors,* 8**,** 284.

Chapman, R. F. 1961. Some experiments to determine the methods used in host-finding by the tsetse fly, *Glossina medicorum* Austen. *Bulletin of Entomological Research,* 52**,** 83-97.

Chappuis, F., Loutan, L., Simarro, P., Lejon, V. & Büscher, P. 2005a. Options for field diagnosis of Human African Trypanosomiasis. *Clinical Microbiology Reviews,* 18**,** 133-146.

Chappuis, F., Udayraj, N., Stietenroth, K., Meussen, A. & Bovier, P. A. 2005b. Eflornithine is safer than melarsoprol for the treatment of second-stage Tr*ypanosoma brucei* gambiense human African trypanosomiasis. *Clinical Infectious Diseases,* 41**,** 748-751.

Checchi, F., Filipe, J. A. N., Haydon, D. T., Chandramohan, D. & Chappuis, F. 2008. Estimates of the duration of the early and late stage of gambiense sleeping sickness. *BMC Infectious Diseases,* 8**,** 16-16.

Cheng, Q. & Aksoy, S. 1999. Tissue tropism, transmission and expression of foreign genes in vivo in midgut symbionts of tsetse flies. *Insect Molecular Biology,* 8**,** 125.

Cheng, Q., Ruel, T. D., Zhou, W., Moloo, S. K., Majiwa, P., O'Neill, S. L. & Aksoy, S. 2000. Tissue distribution and prevalence of *Wolbachia* infections in tsetse flies, *Glossina spp. Medical and Veterinary Entomology,* 14**,** 44-50.

Chigusa, Y., Ohshita, M., Taya, J., Kirinoki, M., Yokoi, H., Kawai, S. & Matsuda, H. 1997. Relationship between longevity and body weight in *Glossina morsitans morsitans. Medical Entomology and Zoology,* 48**,** 91-96.

Chinery, W. A. 1965. Mechanical transmission of some disease pathogens by insects—a review. *Ghana Journal of Science,* 5**,** 249-263.

Clausen, P. H., Adeyemi, I., Bauer, B., Breloeer, M., Salchow, F. & Staak, C. 1998. Host preferences of tsetse (Diptera: *Glossinidae*) based on bloodmeal identifications. *Medical and Veterinary Entomology,* 12**,** 169-80.

Claxton, J. R., Faye, J. A. & Rawlings, P. 1992. Trypanosome infections in warthogs (*Phacochoerus aethiopicus*) in The Gambia. *Veterinary Parasitology,* 41**,** 179-187.

Cox, A., Tilley, A., McOdimba, F., Fyfe, J., Eisler, M., Hide, G. & Welburn, S. 2005. A PCR based assay for detection and differentiation of African trypanosome species in blood. *Experimental Parasitology,* 111**,** 24-29.

Cunningham, L. J., Lingley, J. K., Haines, L. R., Ndung'u, J. M., Torr, S. J. & Adams, E. R. 2016. Illuminating the prevalence of *Trypanosoma brucei* s.l. in *Glossina* using LAMP as a tool for xenomonitoring. *PLoS Neglected Tropical Diseases,* 10**,** e0004441.

D'Amico, F., Gouteux, J. P., Le Gall, F. & Cuisance, D. 1996. Are stable fly (Diptera: *Stomoxyinae*) vectors of *Trypanosoma vivax* in the Central African Republic. *Veterinary Research,* 27**,** 161-170.

Dagnachew, S. & Bezei, M. 2015. Review on *Trypanosoma vivax*. *African Journal of Bisic & Applied Sciences,* 7**,** 41-64.

Dale, C. & Maudlin, I. 1999. *Sodalis* gen. nov. and *Sodalis glossinidius sp*. nov., a microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. *International Journal of Systematic Bacteriology,* 49**,** 267-27.

Dale, C. & Welburn, S. C. 2001. The endosymbionts of tsetse flies: manipulating host–parasite interactions. *International Journal for Parasitology,* 31**,** 628-631.

Dale, C., Welburn, S. C., Maudlin, I. & Milligan, P. J. M. 1995. The kinetics of maturation of trypanosome infections in tsetse. *Parasitology,* 111**,** 187-191.

Dale, C., Young, S. A., Haydon, D. T. & Welburn, S. C. 2001. The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 1883-1888.

Damper, D. & Patton, C. L. 1976. Pentamidine transport and sensitivity in brucei-group trypanosomes. *Journal of Protozoology,* 23**,** 349-356.

Darby, A. C., Lagnel, J., Matthew, C. Z., Bourtzis, K., Maudlin, I. & Welburn, S. C. 2005. Extrachromosomal DNA of the symbiont *Sodalis glossinidius*. *The Journal of Bacteriology,* 187**,** 5003-5007.

de Koning, H. P. 2001. Transporters in African trypanosomes: role in drug action and resistance. *International Journal for Parasitology,* 31**,** 512-522.

De Vooght, L., Caljon, G., De Ridder, K. & Van Den Abbeele, J. 2014. Delivery of a functional anti-trypanosome Nanobody in different tsetse fly tissues via a bacterial symbiont, *Sodalis glossinidius*. *Microbiology Cell Fact,* 13**,** 156.

De Vooght, L., Caljon, G., Stijlemans, B., De Baetselier, P., Coosemans, M. & Van Den Abbeele, J. 2012. Expression and extracellular release of a functional anti-trypanosome Nanobody® in *Sodalis glossinidius*, a bacterial symbiont of the tsetse fly. *Microbial Cell Factories,* 11**,** 23-23.

De Vooght, L., Caljon, G., Van Hees, J. & Van Den Abbeele, J. 2015. Paternal transmission of a secondary symbiont during mating in the viviparous tsetse fly. *Molecular Biology and Evolution,* 32**,** 1977-80.

Dennis, J. W., MacLeod, E. T., Durkin, S. M., Downie, J. E. H., Hamill, L. C. & Anderson, N. E. 2014. *Sodalis glossinidius* prevalence and trypanosome presence in tsetse from Luambe National Park, Zambia. *Parasites & Vectors,* 7.

Desowitz, R. S. & Fairbairn, H. 1955. The influence of temperature on the length of the developmental cycle of *Trypanosoma vivax* in *Glossina palpalis*. *Annals of Tropical Medicine and Parasitology,* 49**,** 161-163.

Desquesnes, M., Bengaly, Z., Millogo, L., Meme, Y. & Sakande, H. 2001a. The analysis of the cross-reactions occurring in antibody-ELISA for the detection of trypanosomes can improve identification of the parasite species involved. *Annals of Tropical Medicine and Parasitology,* 95**,** 141-55.

Desquesnes, M. & Dia, M. L. 2003a. Mechanical transmission of *Trypanosoma congolense* in cattle by the African tabanid *Atylotus agrestis*. *Experimental Parasitology,* 105**,** 226-231.

Desquesnes, M. & Dia, M. L. 2003b. Trypanosoma vivax: mechanical transmission in cattle by one of the most common African tabanids, Atylotus agrestis. *Experimental Parasitology,* 103**,** 35-43.

Desquesnes, M. & Dia, M. L. 2004. Mechanical transmission of *Trypanosoma vivax* in cattle by the African tabanid, *Atylotus fuscipes*. *Veterinary Parasitology,* 119**,** 9-19.

Desquesnes, M., McLaughlin, G., Zoungrana, A. & Dávila, A. M. R. 2001b. Detection and identification of *Trypanosoma* of African livestock through a single PCR based on internal transcribed spacer 1 of rDNA. *International Journal for Parasitology,* 31**,** 610-614.

Desquesnes, M. & Tresse, L. 1996. Evaluation of the sensitivity of the Woo test for the detection of *Trypanosoma vivax*. *Revue d Elevage et de Medecine Veterinaire des Pays Tropicaux,* 49**,** 315-21.

Desta, M. 2014. Trypanosome infection rate of *Glossina morsitans* and trypanosomosis prevalence in cattle in upper Didessa valley western Ethiopia. *International Journal of Current Microbiology and Applied Sciences,* 3**,** 378-388.

Desta, M., Menkir, S. & Kebede, A. 2013. The study on tsetse fly (*Glossina species*) and their role in the trypanosome infection rate in Birbir valley, Akobo River system, western Ethiopia. *Journal of Veterinary Medicine and Animals Healths,* 5**,** 186-194.

Ministry of Livestock Development. 2011. Strategy for tsetse and trypanosomiasis eradication in Kenya 2011-2021. *In:* Services, Department of Veterinary Services. (ed.). Republic of Kenya.

Dobson, A. J. 1990. An introduction to generalized linear models. *In:* Chatfield, C. & Zidek, J. (eds.). London: Chapman and Hall.

Dolan, R. B., Okech, G., Alushula, H., Mutugi, M., Stevenson, P., Sayer, P. D. & Njogu, A. R. 1990. Homidium bromide as a chemoprophylactic for cattle trypanosomiasis in Kenya. *Acta Tropica,* 47**,** 137-144.

Douglas, A. E. 1989. Mycetocyte symbiosis in insects. *Biological reviews of the Cambridge Philosophical Society,* 64**,** 409-434.

Drevemo, S., Grootenhuis , J. G. & Karstad, L. 1974. Blood parameters in wild ruminants in Kenya. *Journal of Wildlife Disease,* 10**,** 327-334.

Drummond, R. A. & Brown, G. D. 2013. Signalling C-type lectins in antimicrobial immunity. *PLoS Pathogens,* 9**,** e1003417.

Duguma, R., Tasew, S., Olani, A., Damena, D., Alemu, D., Mulatu, T., Alemayehu, Y., Yohannes, M., Bekana, M., Hoppenheit, A., Abatih, E., Habtewold, T., Delespaux, V. & Duchateau, L. 2015. Spatial distribution of *Glossina sp*. and *Trypanosoma sp*. in south-western Ethiopia. *Parasites & Vectors,* 8.

Dukes, P. 1984. Arsenic and old taxa: subspeciation and drug sensitivity in *Trypanosoma brucei*. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 78**,** 711-725.

Dyer, N. A., Donnelly, M. J., Lawton, S. P., Ravel, S., Choi, K. S., Lehane, M. J., Robinson, A. S., Okedi, L. M., Hall, M. J. R. & Solano, P. 2008. Molecular phylogenetics of tsetse flies (Diptera: *Glossinidae*) based on mitochondrial (COI, 16S, ND2) and nuclear ribosomal DNA sequences, with an emphasis on the palpalis group. *Molecular Phylogenetics and Evolution,* 49**,** 227-239.

Dyer, N. A., Rose, C., Ejeh, N. O. & Acosta-Serrano, A. 2013. Flying tryps: survival and maturation of trypanosomes in tsetse flies. *Trends in Parasitology,* 29**,** 188-196.

Ebikeme, C. 2011. Trypanosomes, climate change and evolution; past and present.

Ejezie, G. C. & Davey, K. G. 1976. Some effects of allatectomy in the female tsetse, *Glossina austeni*. *Journal of Insect Physiology,* 22**,** 1743-1749.

Eloy, L. & Lucheis, S. 2009. Canine trypanosomiasis: etiology of infection and implications for public health. *Journal of Venomous Animals and Toxins including Tropical Diseases,* 15**,** 589-611.

England, E. C. & Baldry, D. A. T. 1972. The hosts and trypanosome infection rates of *Glossina pallidipes* in the Lambwe and Roo valleys. *Bulletin of The World Health Organization,* 47**,** 785-788.

Engstler, M. & Boshart, M. 2004. Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. *Genes & Development,* 18**,** 2798-2811.

Enyaru, J. C., Ouma, J. O., Malele, I. I., Matovu, E. & Masiga, D. K. 2010. Landmarks in the evolution of technologies for identifying trypanosomes in tsetse flies. *Trends in Parasitology,* 26**,** 388-394.

Esterhuizen, J., Rayaisse, J. B., Tirados, I., Mpiana, S., Solano, P., Vale, G. A., Lehane, M. J. & Torr, S. J. 2011. Improving the cost-effectiveness of visual devices for the control of riverine tsetse flies, the major vectors of Human African Trypanosomiasis. *PLoS Neglected Tropical Diseases,* 5**,** e1257.

Farikou, O., Njiokou, F., Cuny, G. & Geiger, A. 2011a. Microsatellite genotyping reveals diversity within populations of *Sodalis glossinidius*, the secondary symbiont of tsetse flies. *Veterinary Microbiology,* 150**,** 207-210.

Farikou, O., Njiokou, F., Mbida, J. A., Njitchouang, G. R., Djeunga, H. N., Asonganyi, T., Simarro, P. P., Cuny, G. & Geiger, A. 2010a. 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,* 10**,** 115-121.

Farikou, O., Njiokou, F., Simo, G., Asonganyi, T., Cuny, G. & Geiger, A. 2010b. Tsetse fly blood meal modification and trypanosome identification in two sleeping sickness foci in the forest of southern Cameroon. *Acta Tropica,* 116**,** 81-88.

Farikou, O., Thevenon, S., Njiokou, F., Allal, F., Cuny, G. & Geiger, A. 2011b. Genetic diversity and population structure of the secondary symbiont of tsetse flies, Sodalis glossinidius, in sleeping sickness foci in Cameroon. *PLoS Neglected Tropical Diseases,* 5**,** e1281.

Fávero, J. F., Da Silva, A. S., Biazus, A. H. & Volpato, A. 2016. *Trypanosoma vivax* infection in goat in west of Santa Catarina state, Brazil. *Comparative Clinical Pathology,* 25**,** 497-499.

Fèvre, E. M., Coleman, P., Welburn, S. & Maudlin, I. 2004. Reanalyzing the 1900–1920 sleeping sickness epidemic in Uganda. *Emerging Infectious Diseases journal,* 10**,** 567-573.

Fiske, W. F. 1920. Investigations into the bionomics of *Glossina palpalis*. *Bulletin of Entomological Research,* 10**,** 347-463.

Ford, J. 1971. Patterns of tsetse distribution and evolution. *The role of trypanosome in African ecology: a study of the tsetse fly problem*. London: George Allen & Unwin.

Foxman, B. & Riley, L. 2001. Molecular epidemiology: focus on infection. *American Journal of Epidemiology,* 153**,** 1135-1141.

Franco, J. R., Simarro, P. P., Diarra, A. & Jannin, J. G. 2014. Epidemiology of Human African Trypanosomiasis. *Clinical Epidemiology,* 6 257-275.

Gaston, K. A. & Randolph, S. E. 1993. Reproductive under-performance of tsetse flies in the laboratory, related to feeding frequency. *Physiological Entomology,* 18**,** 130-136.

Geerts, S., Geerts, S., Holmes, P. H., Eisler, M. C. & Diall, O. 2001. African bovine trypanosomiasis: the problem of drug resistance. *Trends in Parasitology,* 17**,** 25-28.

Geiger, A., Cuny, G. & Frutos, R. 2005a. Two tsetse fly species, *Glossina palpalis gambiensis* and *Glossina morsitans morsitans*, carry genetically distinct populations of the secondary symbiont *Sodalis glossinidius*. *Applied and Environmental Microbiology,* 71**,** 8941-8943.

Geiger, A., Fardeau, M.-L., Njiokou, F., Joseph, M., Asonganyi, T., Ollivier, B. & Cuny, G. 2011. Bacterial diversity associated with populations of *Glossina spp*. from Cameroon and distribution within the Campo sleeping sickness focus. *Microbial Ecology,* 62**,** 632-643.

Geiger, A., Fardeau, M.-L., Njiokou, F. & Ollivier, B. 2013. *Glossina spp*. gut bacterial flora and their putative role in fly-hosted trypanosome development. *Frontiers in Cellular and Infection Microbiology,* 3.

Geiger, A., Ravel, S., Frutos, R. & Cuny, G. 2005b. *Sodalis glossinidius* (*Enterobacteriaceae*) and vectorial competence of *Glossina palpalis gambiensis* and *Glossina morsitans morsitans* for *Trypanosoma congolense* savannah type. *Current Microbiology,* 51**,** 35-40.

Geiger, A., Ravel, S., Mateille, T., Janelle, J., Patrel, D., Cuny, G. & Frutos, R. 2007. Vector competence of *Glossina palpalis gambiensis* for *Trypanosoma brucei* s.l. and genetic diversity of the symbiont *Sodalis glossinidius*. *Molecular Biology and Evolution,* 24**,** 102-109.

Geigy, R., Kauffmann, M. & Steiger, R. 1971. Miscellanea: Influence of bloodmeals from different donors on the infection rates of "*Trypanosoma brucei*" in "*Glossina*". *Acta Tropica,* 28**,** 164-169.

Gibson, G. 1992. Do tsetse flies 'see' zebras? A field study of the visual response of tsetse to striped targets. *Physiological Entomology,* 17**,** 141-147.

Gibson, G. & Brady, J. 1988. Flight behaviour of tsetse flies in host odour plumes: the initial response to leaving or entering odour. *Physiological Entomology,* 13**,** 29-42.

Gibson, W. 2003. Species concepts for trypanosomes: from morphological to molecular definitions? *Kinetoplastid Biology and Disease,* 2**,** 10-10.

Gibson, W. & Bailey, M. 2003. The development of *Trypanosoma brucei* within the tsetse fly midgut observed using green fluorescent trypanosomes. *Kinetoplastid Biology and Disease,* 2**,** 1.

Glasgow, J. P. & Weitz, B. 1956. The natural hosts of some species of *Glossina* in East Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 50**,** 593-612.

Gooding, R. & McIntyre, G. 1998. *Glossina morsitans morsitans* and *Glossina palpalis palpalis*: dosage compensation raises questions about the milligan model for control of trypanosome development. *Experimental parasitology,* 90**,** 244-249.

Gottdenker, N. L., Chaves, L. F., Calzada, J. E., Saldaña, A. & Carroll, C. R. 2012. Host life history strategy, species diversity, and habitat influence *Trypanosoma cruzi* vector infection in changing landscapes. *PLoS Neglected Tropical Diseases,* 6**,** e1884.

Gould, E. A. & Higgs, S. 2009. Impact of climate change and other factors on emerging arbovirus diseases. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 103**,** 109-121.

Grant, I. F. 2001. Insecticides for tsetse and trypanosomiasis control: is the environmental risk acceptable? *Trends in Parasitology,* 17**,** 10 - 4.

Green, C. H. 1986. Effects of colours and synthetic odours on the attraction of *Glossina pallidipes* and *G. morsitans morsitans* to traps and screens. *Physiological Entomology,* 11**,** 411-421.

Green, C. H. & Flint, S. 1986. An analysis of colour effects in the performance of the F2 trap against *Glossina pallidipes* Austen and *G. morsitans morsitans* Westwood (Diptera: *Glossinidae*). *Bulletin of Entomological Research,* 76**,** 409-418.

Grootenhuis, J. G., Dwinger, R. H., Dolan, R. B., Moloo, S. K. & Murray, M. 1990. Susceptibility of African buffalo and Boran cattle to Trypanosoma congolense transmitted by Glossina morsitans centralis. *Veterinary Parasitology,* 35**,** 219-231.

Gutierrez, C., Corbera, J. A., Juste, M. C., Doreste, F. & Morales, I. 2005. An outbreak of abortions and high neonatal mortality associated with Trypanosoma evansi infection in dromedary camels in the Canary Islands. *Veterinary Parasitology,* 130**,** 163-168.

Hadaway, A. B. 1972. Toxicity of insecticides to tsetse flies. *Bulletin of the World Health Organization,* 46**,** 353-362.

Haines, L. R. 2013. Examining the tsetse teneral phenomenon and permissiveness to trypanosome infection. *Frontiers in Cellular and Infection Microbiology,* 3.

Haines, L. R., Jackson, A. M., Lehane, M. J., Thomas, J. M., Yamaguchi, A. Y., Haddow, J. D. & Pearson, T. W. 2005. Increased expression of unusual EP repeat-containing proteins in the midgut of the tsetse fly (*Glossina*) after bacterial challenge. *Insect Biochemistry and Molecular Biology,* 35**,** 413-423.

Haines, L. R., Lehane, S. M., Pearson, T. W. & Lehane, M. J. 2010. Tsetse EP protein protects the fly midgut from trypanosome establishment. *PLoS Pathogens,* 6**,** e1000793.

Hamill, L. C., Kaare, M. T., Welburn, S. C. & Picozzi, K. 2013. Domestic pigs as potential reservoirs of human and animal trypanosomiasis in Northern Tanzania. *Parasites & Vectors,* 6**,** 1-7.

Hao, Z., Kasumba, I. & Aksoy, S. 2003. Proventriculus (cardia) plays a crucial role in immunity in tsetse fly (Diptera: *Glossinidiae*). *Insect Biochemistry and Molecular Biology,* 33**,** 1155-1164.

Hao, Z., Kasumba, I., J., L. M., Gibson, W. C., Kwon, J. & Aksoy, S. 2001. Tsetse immune responses and trypanosome transmission: implications for the development of tsetse-based strategies to reduce trypanosomiasis. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 12648-12653.

Hargrove, J. W. 1976. The effect of human presence on the behaviour of tsetse (*Glossina spp.*) (Diptera, *Glossinidae*) near a stationary ox. *Bulletin of Entomological Research,* 66**,** 173-178.

Hargrove, J. W. 1994. Reproductive rates of tsetse flies in the field in Zimbabwe. *Physiological Entomology,* 19**,** 307-318.

Hargrove, J. W., Ouifki, R. & Ameh, J. E. 2011. A general model for mortality in adult tsetse (*Glossina spp.*). *Medical and Veterinary Entomology,* 25**,** 385-394.

Harley, J. M. B. 1968. The influence of temperature on reproduction and development in four species of *Glossina* (Diptera: *Muscidae*). *Proceedings of the Royal Entomological Society of London. Series A, General Entomology,* 43**,** 170-177.

Harmsen, R. 1970. Pteridines in flies' of the genus *Glossina* (*Diptera*). *Acta tropica,* 27**,** 165-172.

Harraca, V., Syed, Z. & Guerin, P. M. 2009. Olfactory and behavioural responses of tsetse fies, *Glossina spp.*, to rumen metabolites. *Journal of Comparative Physiology A,* 195**,** 815-824.

Hayes, E. J. & Wall, R. 1999. Age-grading adult insects: a review of techniques. *Physiological Entomology,* 24**,** 1-10.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences,* 270**,** 313-321.

Höft, R. & Höft, M. 1995. The differential effects of elephants on rain forest communities in the Shimba Hills, Kenya. *Biological Conservation,* 73**,** 67-79.

Hu, C. & Aksoy, S. 2006. Innate immune responses regulate trypanosome parasite infection of the tsetse fly *Glossina morsitans morsitans*. *Molecular Microbiology,* 60**,** 1194-1204.

Isaac, C., Ciosi, M., Hamilton, A., Scullion, K. M., Dede, P., Igbinosa, I. B., Nmorsi, O. P. G., Masiga, D. & Turner, C. M. R. 2016. Molecular identification of different trypanosome species and subspecies in tsetse flies of northern Nigeria. *Parasites & Vectors,* 9**,** 1.

Isaac, C., Hamilton, A., Scullion, K., Ciosi, M., Dede, P. M., Igbinosa, I. B. I., Nmorsi, O. P. G., Masiga, D. & Turner, C. M. R. 2013. Distribution of trypanosomes within tsetse in northern Nigeria. International Society for Infectious Disease.

Ivanova, N. V., Dewaard, J. R. & Hebert, P. D. N. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes,* 6**,** 998-1002.

Ivanova, N. V., Zemlak, T. S., Hanner, R. H. & Hebert, P. D. N. 2007a. Barcoding univeral primer cocktails for fish DNA barcoding. *Molecular Ecology Notes,* 7**,** 544-548.

Ivanova, N. V., zemlak, T. S., Hanner, r. H. & Hebert, P. D. N. 2007b. Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes,* 7**,** 544-548.

Jackson, C. H. N. 1946. An artificially isolated generation of tsetse flies (*Diptera*). *Bulletin of Entomological Research,* 37**,** 291-299.

Jaenson, T. G. T. 1986. Sex ratio distortion and reduced lifespan of *Glossina pallidipes* infected with the virus causing salivary gland hyperplasia. *Entomologia Experimentalis et Applicata,* 41**,** 265-271.

Jamonneau, V., Truc, P., Garcia, A., Magnus, E. & Buscher, P. 2000. Preliminary evaluation of LATEX/*T. b. gambiense* and alternative versions of CATT/*T. b. gambiense* for the serodiagnosis of Human African Trypanosomiasis of a population at risk in Cote d'Ivoire: considerations for mass-screening. *Acta Tropica,* 76**,** 175-83.

Jordan, A. M. 1976. Tsetse flies as vectors of trypanosomes. *Veterinary Parasitology,* 2**,** 143-152.

Jordan, A. M. 1986. *Trypanosomiasis control and african rural development*, Longman, London.

Kaaya, G. P. & Darjt, N. 1988. The humoral defense system in tsetse: Differences in response due to age, sex and antigen types. *Developmental & Comparative Immunology,* 12**,** 255-268.

Kaba, D., Zacarie, T., M'Pondi, A. M., Njiokou, F., Bosson-Vanga, H., Kröber, T., McMullin, A., Mihok, S. & Guerin, P. M. 2014. Standardising visual control

devices for tsetse flies: central and west African species *Glossina palpalis palpalis*. *PLoS Neglected Tropical Disease,* 8**,** e2601.

Kabayo, J. P. 2002. Aiming to eliminate tsetse from Africa. *Trends in Parasitology,* 18**,** 473-475.

Kaminsky, R. & Zweygarth, E. 1989. Feeder layer-free in vitro assay for screening antitrypanosomal compounds against *Trypanosoma bruce*i brucei and *T. b.* evansi. *Antimicrobial Agents and Chemotherapy,* 33**,** 881-885.

Kappmeier, K. 2000. A newly developed odour-bated "H trap" for the live collection of *Glossina brevipalpis* and *Glossina austeni* (Diptera: *Glossinidae*) in South Africa. *The Onderstepoort Journal of Veterinary Research,* 67**,** 15.

Karshima, S. N., Lawal, I. A. & Okubanjo, O. O. 2016. Feeding patterns and xenomonitoring of trypanosomes among tsetse flies around the Gashaka-Gumti National Park in Nigeria. *Journal of Parasitology Research,* 2016.

Kaushik, R. S., Uzonna, J. & Tabel, H. 1997. Effect of different cytokines on the growth of *Trypanosoma congolense* cultured under axenic conditions. *Veterinary Parasitology,* 70**,** 25-31.

Kay, B. H., Boreham, P. F. L. & Edman, J. D. 1979. Application of the "Feeding index" concept to studies of mosquito host-feeding patterns. *Mosquito News,* 39**,** 68-72.

Kennedy, P. G. E. 2013. Clinical features, diagnosis, and treatment of Human African Trypanosomiasis (sleeping sickness). *The Lancet Neurology,* 12**,** 186-194.

Kent, R. J. & Norris, D. E. 2005. Identification of mammalian blood meals in mosquitoes by a multiplexed Polymerase Chain Reaction targeting cytochrome B. *The American Journal of Tropical Medicine and Hygiene,* 73**,** 336-342.

Kinghorn, A. & Yorke, W. 1912. On the influence of meterological condition on the development of *Trypanosoma rhodesciense* in *Glossina morsitans*. *British Medical Journal,* 2**,** 835-837.

Koch, A. 1967. Endosymbiosis of Animals with Plant Micro-Organisms. The University of Chicago Press.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X. & Willson, A. C. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America,* 86**,** 6196-6200.

Kubi, C., Van den Abbeele, J., De Deken, R., Marcotty, T., Dorny, P. & Van den Bossche, P. 2006. The effect of starvation on the susceptibility of teneral and non-teneral tsetse flies to trypanosome infection. *Medical and Veterinary Entomology,* 20**,** 388-392.

Kyorku, C., Brightwell, R. & dransfield, R. D. 1990. Traps and odour baits for the tsetse fly, *Glossina longipennis* (Diptera: *Glossinidae*). *Bulletin of Entomological Research,* 80**,** 405-415.

Lah, E. F. C., Ahamad, M., Haron, M. S. & Ming, H. T. 2012. Establishment of a molecular tool for blood meal identification in Malaysia. *Asian Pacific Journal of Tropical Biomedicine,* 2**,** 223-227.

Laird, M. 1977. *Tsetse: The future for biological methods in intergrated control,* Ottawa, Canada, International Development Research Centre.

Langley, P. A. 1972. The role of physical and chemical stimuli in the development of in vitro feeding techniques for tsetse flies Glossina spp.(Dipt., Glossinidae). *Bulletin of Entomological Research,* 62**,** 215-228.

Langley, P. A., Hall, M. J. R., Felton, T. & Ceesay, M. 1988. Determining the age of tsetse flies, *Glossina spp*. (Diptera: *Glossinidae*): an appraisal of the pteridine fluorescence technique. *Bulletin of Entomological Research,* 78**,** 387-395.

Langley, P. A. & Stafford, K. 1990. Feeding frequency in relation to reproduction in *Glossina morsitans morsitans* and *G. pallidipes*. *Physiological Entomology,* 15**,** 415-421.

Langley, P. A. & Wall, R. 1990. The implications of hunger in the tsetse fly, *Glossina pallidipes,* in relation to its availability to trapping techniques. *Journal of Insect Physiology,* 36**,** 903-908.

Langridge, W. P., Kernaghan, R. J. & Glover, P. E. 1963. A review of recent knowledge of the ecology of the main vectors of trypanosomiasis. *Bulletin of The World Health Organization,* 28**,** 671-701.

Leak, S. G. A. 1998. Tsetse biology and ecology. *The role in the epidemiology and control of trypanosmiasis*. London: CABI Publishing.

Lehane, M. J. 1997. Peritrophic matrix structure and function. *Annual Review of Entomology,* 42**,** 525-550.

Lehane, M. J., Aksoy, S. & Levashina, E. 2004. Immune responses and parasite transmission in blood-feeding insects. *Trends in Parasitology,* 20**,** 433-439.

Lehane, M. J. & Mail, T. S. 1985. Determining the age of adult male and female Glossina morsitans morsitans using a new technique. *Ecological Entomology,* 10**,** 219-224.

Lehane, M. J., Msangi, A. R., Whitaker, C. J. & Lehane, S. M. 2000. Grouping of trypanosome species in mixed infections in *Glossina pallidipes*. *Parasitology,* 120**,** 583-592.

Lindh, J. M. & Lehane, M. J. 2011. The tsetse fly *Glossina fuscipes fusipes* (Diptera: *Glossina*) harbours a surprising diversity of bacteria other than symbionts. *Antonie van Leeuwenhoek,* 99**,** 711-720.

Liniger, M., Acosta-Serrano, A., Van Den Abbeele, J., Renggli, C. K., Brun, R., Englund, P. T. & Roditi, I. 2003. Cleavage of trypanosome surface glycoproteins by alkaline trypsin-like enzyme(s) in the midgut of *Glossina morsitans*. *International Journal for Parasitology,* 33**,** 1319-1328.

Lisulo, M., Sugimoto, C., Kajino, K., Hayashida, K., Mudenda, M., Moonga, L., Ndebe, J., Nzala, S. & Namangala, B. 2014. Determination of the prevalence of African trypanosome species in indigenous dogs of Mambwe district, eastern Zambia, by loop-mediated isothermal amplification. *Parasites & Vectors,* 7**,** 19.

Liu, R., Lehane, S. M., He, X., Lehane, M. J., Hertz-Fowler, C., Berriman, M., Pickett, J. A., Field, L. M. & Zhou, J. 2010. Characterisations of odorant-binding proteins in the tsetse fly *Glossina morsitans morsitans*. *Cellular and Molecular Life Sciences,* 67**,** 919-929.

Lloyd, L. & Johnson, W. B. 1924. The trypanosome infections of tsetse-flies in northern Nigeria and a new method of estimation. *Bulletin of Entomological Research,* 14**,** 265-288.

Losos, G. J. & Ikede, B. O. 1972. *Veterinary Pathology,* Switzerland, Boehm Co, Basel.

Lyimo, I. N. & Ferguson, H. M. 2009. Ecological and evolutionary determinants of host species choice in mosquito vectors. *Trends in Parasitology,* 25**,** 189-196.

Lyimo, I. N., Haydon, D. T., Russell, T. L., Mbina, K. F., Daraja, A. A., Mbehela, E. M., Reeve, R. & Ferguson, H. M. 2013. The impact of host species and vector control measures on the fitness of African malaria vectors. *Proceedings of the Royal Society of London B: Biological Sciences,* 280.

Lymbery, A. J. & Thompson, R. C. A. 2012. The molecular epidemiology of parasite infections: Tools and applications. *Molecular and Biochemical Parasitology,* 181**,** 102-116.

Macleod, E. T., Darby, A. C., Maudlin, I. & Welburn, S. C. 2007a. Factors affecting trypanosome maturation in tsetse flies. *PloS ONE,* 2**,** e239.

MacLeod, E. T., Maudlin, I., Darby, A. C. & Welburn, S. C. 2007b. Antioxidants promote establishment of trypanosome infections in tsetse. *Parasitology,* 134**,** 827-831.

Majekodunmi, A., Fajinmi, A., Dongkum, C., Picozzi, K., Thrusfield, M. & Welburn, S. 2013. A longitudinal survey of African animal trypanosomiasis in domestic cattle on the Jos Plateau, Nigeria: prevalence, distribution and risk factors. *Parasites & Vectors,* 6**,** 239.

Makumi, J. N., Stevenson, P. & Green, C. H. 2000. Control of *Glossina longipennis* (Diptera: *Glossinidae*) by insecticide-treated targets at Galana Ranch, Kenya, and confirmation of the role of *G. longipennis* as a vector of cattle trypanosomiasis. *Bulletin of Entomological Research,* 90**,** 397-406.

Maleki-Ravasan, N., Oshaghi, M. A., Javadian, E., Rassi, Y., Sadraei, J. & Mohtarami, F. 2009. Blood meal identification in field-captured sand flies: comparison of PCR-RFLP and ELISA Assays. *Iranian Journal of Arthropod-borne Diseases,* 3**,** 8-18.

Malele, I., Craske, L., Knight, C., Ferris, V., Njiru, Z., Hamilton, P., Lehane, S., Lehane, M. & Gibson, W. 2003. The use of specific and generic primers to identify trypanosome infections of wild tsetse flies in Tanzania by PCR. *Infection, Genetics and Evolution* 3**,** 271-279.

Malele, I. I., Ouma, J. O., Nyingilili, H. S., Kitwika, W. A., Malulu, D. J., Magwisha, H. B. & Kweka, E. J. 2016. Comparative performance of traps in catching tsetse flies (Diptera: *Glossinidae*) in Tanzania. *Onderstepoort Journal of Veterinary Resesearch,* 83**,** 7 pages.

Mamoudou, A., Zoli, A., Hamadama, H., Abah, S., Geerts, S., Clausen, P. H., Zessin, K. H., Kyule, M. & Van Den Bossche, P. 2008. Seasonal distribution and abundance of tsetse flies (*Glossina spp*.) in the Faro and Deo Division of the Adamaoua Plateau in Cameroon. *Medical and Veterinary Entomology,* 22**,** 32-36.

Marc, D., Ketsarin, K., Sarawut, Y., Cristina, M., Sophie, R., Ming-Hui, W., Zhao-Rong, L., Serge, M. & Sathaporn, J. 2011. Specific primers for PCR amplification of the ITS1 (ribosomal DNA) of *Trypanosoma lewisi*. *Infection, Genetics and Evolution***,** 1361-1367.

Marcello, L. & Barry, J. D. 2007. Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favoured by archive substructure. *Genome research,* 17**,** 1344-1352.

Masake, R. A., Majiwa, P. A. O., Moloo, S. K., Makau, J. M., Njuguna, J. T., Maina, M., Kabata, J., ole-MoiYoi, O. K. & Nantulya, V. M. 1997. Sensitive and specific detection of *Trypanosoma vivax* using the Polymerase Chain Reaction. *Experimental Parasitology,* 85**,** 193-205.

Masake, R. A., Nantulya, V. M., Pellé, R., Makau, J. M., Gathuo, H. & ole-MoiYoi, O. K. 1994. A species-specific antigen of *Trypanosoma* (Duttonella) *vivax* detectable in the course of infection is encoded by a differentially expressed tandemly reiterated gene. *Molecular and Biochemical Parasitology,* 64**,** 207-218.

Masiga, D. K., Mcnamara, J. J., Laveissiere, C., Truc, C. & Gibson, W. C. 1996. A high prevalence of mixed trypanosome infections in tsetse flies in Sinfra, Cote d'lvoire, detected by DNA amplification. *Parasitology,* 112**,** 75-80.

Masiga, D. K., Smyth, A. J., Hayes, P., Bromidge, T. J. & Gibson, W. C. 1992. Sensitive detection of trypanosomes in tsetse flies by DNA amplification. *International Journal for Parasitology,* 22**,** 909-918.

Masumu, J., Marcotty, T., Ndeledje, N., Kubi, C., Geerts, S., Vercruysse, J., Dorny, P. & Van den Bossche, P. 2006. Comparison of the transmissibility of *Trypanosoma congolense* strains, isolated in a trypanosomiasis endemic area of eastern Zambia, by *Glossina morsitans morsitans*. *Parasitology,* 133**,** 331-334.

Matovu, E., Seebeck, T., Enyaru, J. C. K. & Kaminsky, R. 2001. Drug resistance in *Trypanosoma brucei* spp., the causative agents of sleeping sickness in man and nagana in cattle. *Microbes and Infection,* 3**,** 763-770.

Matthew, C. Z. 2007. *Biological and molecular aspects of Sodalis glossinidius*. University of Edinburgh.

Matthew, C. Z., Darby, A. C., Young, S. A., Hume, H. L. & Welburn, S. C. 2005. The rapid isolation and growth dynamics of the tsetse symbiont *Sodalis glossinisius*. *FEMS Microbiology Letters,* 248**,** 69-74.

Mattioli, R. C., Faye, J. A. & Jaitner, J. 2001. Estimation of trypanosomal status by the buffy coat technique and an antibody ELISA for assessment of the impact of trypanosomosis on health and productivity of N'Dama cattle in The Gambia. *Veterinary Parasitology,* 95**,** 25-35.

Maudlin, I. & Ellis, D. S. 1985. Association between intracellular rickettsial-like infections of midgut cells and susceptibility to trypanosome infection in *Glossina spp*. *Parasitology Research,* 71**,** 683-687.

Maudlin, I., Holmes, P. H. & Miles, M. A. 2004. *The trypanosomiases,* Wallingford, UK, CABI Publishing.

Maudlin, I., Welburn, S. C. & Milligan, P. J. 1998. Trypanosome infections and survival in tsetse. *Parasitology,* 116 Suppl**,** S23-8.

Mbahin, N., Affognon, H., Andoke, J., Tiberius, M., Mbuvi, D., Otieno, J., Muasa, P. & Saini, R. K. 2013. Parasitological prevalence of bovine trypanosomiasis in Kubo Division of Kwale Country of Coastal: baseline survey *American Journal of Animal and Veterinary Sciences,* 8**,** 28-36.

McDermott, J., Woitag, T., Sidibe´, I., Bauer, B., Diarra, B., Denis, O., Kamuanga, M., Peregrine, A., Eisler, M., Zessin, K.-H., Mehlitz, D. & Clausen, P.-H. 2003. Field studies of drug-resistant cattle trypanosomes in Ke´ne´dougou Province, Burkina Faso. *Acta Tropica,* 86**,** 93-103.

McLntyre, G. S. & Gooding, R. H. 1996. Variation in the pteridine content in the heads of tsetse flies (Diptera: *Glossinidae*: *Glossina Wiedemann*): evidence for genetic control. *Canadian Journal of Zoology,* 74**,** 621-626.

Mekata, H., Konnai, S., Simuunza, M., Chembensofu, M., Kano, R., Witola, W. H., Tembo, M. E., Chitambo, H., Inoue, N., Onuma, M. & Ohashi, K. 2008. Prevalence and source of trypanosome infections in field-captured vector flies (*Glossina pallidipes*) in southeastern Zambia. *Journal of Veterinary Medical Science,* 70**,** 923-928.

Mellanby, H. 1937. Experimental work with the tsetse-fly, *Glossina palpalis*, in Uganda. *Parasitology,* 29**,** 131-141.

Mhindurwa, A. 1994. Field observations of tsetse flies (*Glossina spp*.(Diptera: *Glossinidae*)) with new odour-baited trapping devices. *Bulletin of Entomological Research,* 84**,** 529-532.

Mihok, S. 2002. The development of a multipurpose trap (the Nzi) for tsetse and other biting flies. *Bulletin of Entomological Research,* 92**,** 385-403.

Mihok, S., Carlson, D. A. & Ndegwa, P. N. 2007. Tsetse and other biting fly responses to Nzi traps baited with octenol, phenols and acetone. *Medical and Veterinary Entomology,* 21**,** 70-84.

Mihok, S., Maramba, O., Munyoki, E. & Kagoiya, J. 1995. Mechanical transmission of *Trypanosoma spp.* by African *Stomoxyinae* (Diptera: *Muscida*e). *Tropical Medical Parasitology,* 46**,** 103-5.

Mihok, S., Otieno, L. H., Darji, N. & Munyinyi, D. 1992. Influence of d(+)-glucosamine on infection rates and parasite loads in tsetse flies (*Glossina spp.*) infected with *Trypanosoma brucei*. *Acta Tropica,* 51**,** 217-228.

Milord, F., Pépin, J., Ethier, L., Milord, F., Loko, L., Ethier, L. & Mpia, B. 1992. Efficacy and toxicity of eflornithine for treatment of *Trypanosoma brucei gambiense* sleeping sickness. *The Lancet,* 340**,** 652-655.

Moggridge, J. Y. 1949. Climate and the activity of the Kenya coastal *Glossina*. *Bulletin of Entomological Research,* 40**,** 307-321.

Moloo, S. K. & Gooding, R. H. 2000. Long-term study on the susceptibility to *Trypanosoma congolense* infections and genetics of colonized *Glossina pallidipes* from allopatric populations in Kenya. *Canadian Journal of Zoology,* 78**,** 1289-1292.

Moloo, S. K., Kutuza, S. B. & Desal, J. 1987. Comparative study on the infection rates of different *Glossina species* for East and West African *Trypanosoma vivax* stocks. *Parasitology,* 95**,** 537-542.

Moloo, S. K., Sabwa, C. L. & Kabata, J. M. 1992. Vector competence of *Glossina pallidipes* and *G. morsitans centralis* for *Trypanosoma vivax*, *T. congolense* and *T. b. brucei*. *Acta Tropica,* 51**,** 271-280.

Moloo, S. K. & Shaw, M. K. 1989. Rickettsial infections of midgut cells are not associated with susceptibility of *Glossina morsitans centralis* to *Trypanosoma congolense* infection. *Acta Tropica,* 46**,** 223-227.

Moody, A. H. & Chiodini, P. L. 2000. Methods for the detection of blood parasites. *Clinical & Laboratory Haematology,* 22**,** 189-201.

Moore, S., Shrestha, S., Tomlinson, K. W. & Vuong, H. 2012. Predicting the effect of climate change on African trypanosomiasis: integrating epidemiology with parasite and vector biology. *Journal of The Royal Society Interface,* 9**,** 817-830.

Morlais, I., Grebaut, P., Bodo, J. M., Djoha, S., Cuny, G. & Herder, S. 1998. Detection and identification of trypanosomes by Polymerase Chain Reaction in wild tsetse flies in Cameroon. *Acta Tropica* 70**,** 109-117.

Morrison, W. I. & Murray, M. 1985. The role of humoral immune responses in determining susceptibility of A/J and C57BL/6 mice to infection with *Trypanosoma congolense*. *Parasite Immunology,* 7**,** 63-79.

Moser, D. R., Cook, G. A., Ochs, D. E., Bailey, C. P., McKane, M. R. & Donelson, J. E. 1989. Detection of *Trypanosoma congolense* and *Trypanosoma brucei* subspecies by DNA amplification using the Polymerase Chain Reaction. *Parasitology,* 99**,** 57-66.

Motloang, M., Masumu, J., Mans, B., Van den Bossche, P. & Latif, A. 2012. Vector competence of *Glossina austeni* and *Glossina brevipalpis* for *Trypanosoma congolense* in KwaZulu-Natal, South Africa. *The Onderstepoort Journal of Veterinary Research,* 79**,** E1.

Muhanguzi, D., Picozzi, K., Hattendorf, J., Thrusfield, M., Kabasa, J. D., Waiswa, C. & Welburn, S. C. 2014. The burden and spatial distribution of bovine African trypanosomes in small holder crop-livestock production systems in Tororo District, south-eastern Uganda. *Parasites & Vectors,* 7**,** 1-10.

Munks, R. J. L., Sant'Anna, M. R. V., Grail, W., Gibson, W., Igglesden, T., Yoshiyama, M., Lehane, S. M. & Lehane, M. J. 2005. Antioxidant gene expression in the blood-feeding fly *Glossina morsitans morsitans*. *Insect Molecular Biology,* 14**,** 483-491.

Murphy, N. B. & Welburn, S. C. 1997. Programmed cell death in procyclic *Trypanosoma brucei rhodesiense* is associated with differential expression of mRNAs. *Cell Death and Differentiation,* 4**,** 365-370.

Murray, M., Murray, P. K. & McIntyre, W. I. M. 1977. An improved parasitological technique for the diagnosis of African trypanosomiasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 71**,** 325-326.

Murray, M., Trail, J. C. M., Turner, D. A. & Wissocq, Y. 1983a. *Lifestock productivity and trypanotolerance* [Online]. Addis Ababa, Ethiopia: International Laboratory for Research on Animal Diseases, The International Centre of Insect Physiology and Ecology 2015].

Murray, M., Trail, J. C. M., Turner, D. A. & Wissocq, Y. 1983b. Network Training Manual. *Lifestock Productivity and Trypanotolerance*. Addis Ababa, Ethiopia: International Livestock Centre for Africa.

Muturi, C. N., Ouma, J. O., Malele, I. I., Ngure, R. M., Rutto, J. J., Mithöfer, K. M., Enyaru, J. & Masiga, D. K. 2011. Tracking the feeding patterns of tsetse flies (*Glossina* genus) by analysis of bloodmeals using mitochondrial cytochromes genes. *PloS ONE,* 6**,** e17284.

Muzari, M. O. & Hargrove, J. W. 2005. Artificial larviposition sites for field collections of the puparia of tsetse flies *Glossina pallidipes* and *G. m. morsitans* (Diptera: *Glossinidae*). *Bulletin of Entomological Research,* 95**,** 221-229.

Nagamune, K., Acosta-Serrano, A., Uemura, H., Brun, R., Kunz-Renggli, C., Maeda, Y., Ferguson, M. A. & Kinoshita, T. 2004. Surface sialic acids taken from the host allow trypanosome survival in tsetse fly vectors. *The Journal of Experimental Medicine,* 199**,** 1445-1450.

Namangala, B., Hachaambwa, L., Kajino, K., Mweene, A. S., Hayashida, K., Simuunza, M., Simukoko, H., Choongo, K., Chansa, P., Lakhi, S., Moonga, L., Chota, A., Ndebe, J., Nsakashalo-Senkwe, M., Chizema, E., Kasonka, L. & Sugimoto, C. 2012. The use of Loop-mediated Isothermal Amplification (LAMP) to detect the re-emerging Human African Trypanosomiasis (HAT) in the Luangwa and Zambezi Valleys. *Parasites & Vectors,* 5**,** 1-5.

Namangala, B., Oparaocha, E., Kajino, K., Hayashida, K., Moonga, L., Inoue, N., Suzuki, Y. & Sugimoto, C. 2013. Short report: preliminary investigation of trypanosomosis in exotic dog breeds from Zambia's Luangwa and Zambezi Valleys Using LAMP. *The American Journal of Tropical Medicine and Hygiene,* 89**,** 116-118.

Nantulya, V. M. 1990. Trypanosomiasis in domestic animals: the problems of diagnosis. *Revue scientifique et technique (International Office of Epizootics),* 9 357-367.

Nash, T. A. M. 1936. The relationship between the maximum temperature and the seasonal longevity of *Glossina submorsitans*, Newst., and *G. tachinoides*, Westw., in Northern Nigeria. *Bulletin of Entomological Research,* 27**,** 273-279.

Nash, T. A. M., Jordan, A. M. & Boyle, J. A. 1968. The large-scale rearing of *Glossina austeni* Newst in the laboratory. *Annals of Tropical Medicine and Parasitology,* 62**,** 336-341.

Ndegwa, P. N., Irungu, L. W. & Moloo, S. K. 1992. Effect of puparia incubation temperature: increased infection rates of *Trypnosoma congolense* in

*Glossina morsitans centralis*, *G. fuscipes fuscipes* and *G. brevipalpis*. *Medical and Veterinary Entomology*, 127-130.

Nilson, T. L., Sinclair, B. J. & Roberts, S. P. 2006. The effects of carbon dioxide anesthesia and anoxia on rapid cold-hardening and chill coma recovery in *Drosophila melanogaster*. *Journal of Insect Physiology,* 52, 1027-1033.

Nitcheman, S. 1987. Comparison of the longevity of tsetse flies (*Glossina morsitans morsitans* Westwood, 1850) infected with trypanosomes (*Trypanosoma nannomonas congolense* Broden, 1904) and uninfected tsetse flies. *Annales de parasitologie humaine et comparée,* 63, 163-164.

Njenga, S. M., Mwandawiro, C. S., Muniu, E., Mwanje, M. T., Haji, F. M. & Bockarie, M. J. 2011. Adult population as potential reservoir of NTD infections in rural villages of Kwale district, Coastal Kenya: implications for preventive chemotherapy interventions policy. *Parasites & Vectors,* 4, 1-6.

Njiokou, F., Simo, G., Mbida Mbida, A., Truc, P., Cuny, G. & Herder, S. 2004. A study of host preference in tsetse flies using a modified heteroduplex PCR-based method. *Acta Tropica,* 91, 117-120.

Njiru, B. N. 2014. *The physiological status of the tsetse fly, Glossina fuscipes fuscipes, attracted to different hosts and control devices and its implications for control of Human and Animal African Trypanosomiasis.* Faculty of Science, University of the Witwatersrand, Johannesburg.

Njiru, Z. K., Constantine, C. C., Guya, S., Crowther, J., Kiragu, J. M., Thompson, R. C. & Dávila, A. M. 2005. The use of ITS1 rDNA PCR in detecting pathogenic African trypanosomes. *Parasitology Research,* 95, 186-192.

Njiru, Z. K., Makumi, J. N., Okoth, S., Ndungu, J. M., Gibson, W. C. & Bailey, M. 2004. Identification of trypanosomes in *Glossina pallidipes* and *G. longipennis* in Kenya. *Infection, Genetics and Evolution* 4, 29-35.

Njiru, Z. K., Traub, R., Ouma, J. O., Enyaru, J. C. & Matovu, E. 2011. Detection of Group 1 *Trypanosoma brucei gambiense* by Loop-Mediated Isothermal Amplification. *Journal of Clinical Microbiology,* 49, 1530-1536.

Nogge, G. 1976. Sterility in tsetse flies (*Glossina morsitans* Westwood) caused by loss of symbionts. *Experimental Parasitology,* 32, 995-996.

Nogge, G. & Gerresheim, A. 1982. Experiments on the elimination of symbionts from the tsetse fly, *Glossina morsitans morsitans* (Diptera: *Glossinidae*), by antibiotics and lysozyme. *Journal of Invertebrate Pathology,* 40, 166-179.

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N. & Hase, T. 2000. Loop-mediated isothermal amplification of DNA. *Nucleic Acids Research,* 28, E63.

Nthiwa, D. M., Odongo, D. O., Ochanda, H., Khamadi, S. & Gichimu, B. M. 2015. Trypanosoma infection rates in *Glossina species* in Mtito Andei Division, Makueni County, Kenya. *Journal of Parasitology Research*.

Nyawira, N. C. 2009. *Analysis of tsetse fly (Diptera: Glossinidae) blood meals using mitochondrial cytochrome genes for vertebrate host identification.* Master Degree, Ugerton University.

O'Neill, S. L., Gooding, R. H. & Aksoy, S. 1993. Phylogenetically distant symbiotic microorganisms reside in *Glossina* midgut and ovary tissues. *Medical and Veterinary Entomology,* 7, 377-383.

Odiit, M., Kansiime, F. & Enyaru, J. C. 1997. Duration of symptoms and case fatality of sleeping sickness caused by *Trypanosoma brucei rhodesiense* in Tororo, Uganda. *East African Medical Journal,* 74, 792-5.

Oguge, N. O., Durden, L. A., Keirans, J. E., Balami, H. D. & Schwan, T. G. 2009. Ectoparasites (sucking lice, fleas and ticks) of small mammals in southeastern Kenya. *Medical and Veterinary Entomology,* 23**,** 387-392.

World Organisation for Animal Health 2013. Trypanosomosis (tsetse-transmitted). April 2013 ed.

Okech, G., Dolan, R. B., Stevenson, P., Alushula, H., Watson, E. D., Luckins, A. G. & Omuse, J. K. 1996. The effect of trypanosomosis on pregnancy in trypanotolerant Orma Boran cattle. *Theriogenology,* 46**,** 441-447.

Okoth, S. O., Kokwaro, E. D., Kirugu, J. M. & Murila, G. A. 2007. *Glossina pallidipes* and host interactions: implications of host preference on tranmission Rick of Rhodesian Sleeping Sickness in Kenya. *Trends in Applied Sciences Research,* 2**,** 386-394.

Olubayo, R. O., Moloo, S. K. & Naessens, J. 1996. Comparative parasite development in African buffalo and N'Dama cattle infected with either *Trypanosoma congolense* or *T. vivax*. *Bulletin of Animal Health and Production in Africa,* 44**,** 23-32.

Onyiah, J. A. 1980. Mechanism of host selection by tsetse flies. *International Journal of Tropical Insect Science,* 1**,** 31-34.

Opel, K. L., Chung, D. & McCord, B. R. 2009. A study of PCR inhibition mechanisms using real time PCR. *Journal of Forensic Sciences***,** 1-9.

OsórioI, A. L. A. R., Madruga, C. R., Desquesnes, M., SoaresII, C. O., Ribeiro, L. R. R. & da Costa, S. C. G. 2008. *Trypanosoma* (Duttonella) *vivax*: its biology, epidemiology, pathogenesis, and introduction in the new world - a review. *Memórias do Instituto Oswaldo Cruz* 103**,** 1.

Otter, C. J. D. & Naters, W. M. V. D. G. V. 1992. Single cell recordings from tsetse (*Glossina m. morsitans*) antennae reveal olfactory, mechano - and cold receptors. *Physiological Entomology,* 17**,** 33-42.

Ouma, J. O., Masake, R. A., Masiga, D. K., Moloo, S. K., Njuguna, J. T. & Ndung'u, J. M. 2000. Comparative sensitivity of dot-ELISA, PCR and dissection method for the detection of trypanosome infections in tsetse flies (Diptera: *Glossinidae*). *Acta Tropica,* 75**,** 315-321.

Pais, R., Lohs, C., Wu, Y., Wang, J. & Aksoy, S. 2008. The obligate mutualist *Wigglesworthia glossinidia* influences reproduction, digestion, and immunity processes of its host, the tsetse fly. *Applied and Environmental Microbiology,* 74**,** 5965-5974.

Panton, E. 2015. *A comparative evaluation of generalised and species-specific primers in PCR-based identification of trypanosomes, for development of efficient screening protocols.* Master of Science, Institute of Biodiversity Animal Health and Comparative Medicine.

Pathak, A. K. 2009. Effect of *Trypanosoma spp.* on nutritional status and performance of livestock. *Veterinary World,* 2**,** 435-438.

Peacock, L., Cook, S., Ferris, V., Bailey, M. & Gibson, W. 2012a. The life cycle of *Trypanosoma* (Nannomonas) *congolense* in the tsetse fly. *Parasites & Vectors,* 5**,** 1-13.

Peacock, L., Ferris, V., Bailey, M. & Gibson, W. 2012b. The influence of sex and fly species on the development of trypanosomes in tsetse Flies. *PLoS Neglected Tropical Diseases,* 6**,** e1515.

Pepin, J. & Milord, F. 1994. The treatment of human trypanosomiasis. *Advances in Parasitology,* 160**,** 100-160.

Pepin, J., Milord, F., Guern, C. & Schechter, P. J. 1987. Difluoromethylornithine for arseno-resistant Trypanosoma brucei gambiense sleeping sickness. *The Lancet,* 2**,** 1431–1433.

Picozzi, K., Carrington, M. & Welburn, S. C. 2008. A multiplex PCR that discriminates between *Trypanosoma brucei brucei* and zoonotic *T. b. rhodesiense*. *Experimental Parasitology,* 118**,** 41-46.

Pollock, J. N. 1982. *Training manual for tsetse control personnel volume 1* [Online]. Rome: Food and Africulture Organisation of The United Nations, Agriculture and Consumer Protection. Available: http://www.fao.org/docrep/009/p5178e/P5178E06.htm [Accessed 5 January 2014].

Potts, W. H. 1933. Observations on *Glossina morsitans*, Westw., in East Africa. *Bulletin of Entomological Research***,** 293-300.

Pradhan, V., Kamble, Y., Ladniya, V. & Mogul, M. 2015. A overview of species identification by DNA barcoding. *International Journal of Current Microbiology and Applied Sciences* 4**,** 127-140.

Priotto, G., Fogg, C., Balasegaram, M., Erphas, O., Louga, A., Checchi, F., Ghabri, S. & Piola, P. 2006. Three drug combinations for late-stage *Trypanosoma brucei gambiense* sleeping sickness: a randomized clinical trial in Uganda. *PLoS Clinical Trials,* 1**,** e39.

Radwanska, M., Chamekh, M., Vanhamme, L., Claes, F., Magez, S., Magnus, E., de Baetselier, P., Buscher, P. & Pays, E. 2002a. The serum resistance-associated gene as a diagnostic tool for the detection of *Trypanosoma brucei rhodesiense*. *American Journal of Tropical Medicine and Hygiene,* 67**,** 684-90.

Radwanska, M., Claes, F., Magez, S., Magnus, E., Perez-Morga, D., Pays, E. & Büscher, P. 2002b. Novel primer sequences for Polymerase Chain Reaction-based detection of *Trypanosoma brucei gambiense*. *The American Journal of Tropical Medicine and Hygiene,* 67**,** 289-95.

Randolph, S. E. & Rogers, D. J. 1978. Feeding cycles and flight activity in field populations of tsetse (Diptera: *Glossinidae*). *Bulletin of Entomological Research,* 68**,** 655-671.

Rio, R. V. M., Wu, Y.-n., Filardo, G. & Aksoy, S. 2006. Dynamics of multiple symbiont density regulation during host development: tsetse fly and its microbial flora. *Proceedings of the Royal Society of London B: Biological Sciences,* 273**,** 805-814.

Rita, V. M. R., Yi-neng, W., Giovanni, F. & Serap, A. 2006. Dynamics of multiple symbiont density regulation during host development: tsetse fly and its microbial flora. *Proceedings of the Royal Society of London B: Biological Sciences,* 273**,** 805-814.

Robays, J., Bilengue, M. M. C., Stuyft, P. V. D. & Boelaert, M. 2004. The effectiveness of active population screening and treatment for sleeping sickness control in the Democratic Republic of Congo. *Tropical Medicine & International Health,* 9**,** 542-550.

Roberts, L. W., Wellde, B. T., Reardon, M. J. & Onyango, F. K. 1989. Mechanical transmission of *Trypanosoma brucei* rhodesiense by *Glossina morsitans morsitans* (Diptera:*Glossinidae*). *Annual of Troical Medical Parasitology,* 83 Suppl 1**,** 127-31.

Roditi, I. & Lehane, M. J. 2008. Interactions between trypanosomes and tsetse flies. *Current Opinion in Microbiology,* 11**,** 345-351.

Rogers, D. J. & Randolph, S. E. 1993. Distribution of tsetse and ticks in Africa: past, present and future. *Parasitology Today,* 9**,** 266-271.

Rose, C., Belmonte, R., Armstrong, S. D., Molyneux, G., Haines, L. R., Lehane, M. J., Wastling, J. & Acosta-Serrano, A. 2014. An investigation into the protein composition of the teneral *Glossina morsitans morsitans* Peritrophic Matrix. *PLoS Neglected Tropical Diseases,* 8**,** e2691.

Ross, A. & Novoa-Montero, D. 1993. Comparability and reliability of ELISA, Immunofluorescence, and Indirect Hemagglutination Assays for *Trypanosoma cruzi* and *Trypanosoma rangeli*. *Journal of Infectious Diseases,* 168**,** 1581-1584.

Rowlands, G. J., Mulatu, W., Authié, E., d'Ieteren, G. D. M., Leak, S. G. A., Nagda, S. M. & Peregrine, A. S. 1993. Epidemiology of bovine trypanosomiasis in the Ghibe valley, southwest Ethiopia. 2. Factors associated with variations in trypanosome prevalence, incidence of new infections and prevalence of recurrent infections. *Acta Tropica,* 53**,** 135-150.

Ruiz, J. P., Nyingilili, H. S., Mbata, G. H. & Malele, I. I. 2015. The role of domestic animals in the epidemiology of human African trypanosomiasis in Ngorongoro conservation area, Tanzania. *Parasites & Vectors,* 8**,** 510.

Sahaw, A. P. M. 2009. Assessing the economics of animal trypanosomosis in Africa-history and current perspectives. *Onderstepoort Journal of Veterinary Research,* 76**,** 27-32.

Salekwa, L. P., Nnko, H. J., Ngonyoka, A., Estes, A. B., Agaba, M. & Gwakisa, P. S. 2014. Relative abundance of tsetse fly species and their infection rates in Simanjiro, northern Tanzania. *Livestock Research for Rural Development,* 26.

Sant'Anna, M. R. V., Jones, N. G., Hindley, J. A., Mendes-Sousa, A. F., Dillon, R. J., Cavalcante, R. R., Alexander, B. & Bates, P. A. 2008. Blood meal identification and parasite detection in laboratory-fed and field-captured *Lutzomyia longipalpis* by PCR using FTA databasing paper. *Acta Tropica,* 107**,** 230-237.

Saunders, D. S. 1962. Age determination for female tsetse flies and the age compositions of samples of *Glossina pallidipes* Aust., *G. palpalis fuscipes* Newst. and *G. brevipalpis* Newst. *Bulletin of Entomological Research,* 53**,** 579-595.

Schofield, S. & Torr, S. J. 2002. A comparison of the feeding behaviour of tsetse and stable flies. *Medical and Veterinary Entomology,* 16**,** 177-185.

Seck, M. T., Pagabeleguem, S., Bassene, M. D., Fall, A. G., Diouf, T. A., Sall, B., Vreysen, M. J., Rayaisse, J. B., Takac, P., Sidibe, I., Parker, A. G., Mutika, G. N., Bouyer, J. & Gimonneau, G. 2015. Quality of sterile male tsetse after long distance transport as chilled, irradiated pupae. *PLoS Neglected Tropical Diseases,* 9**,** e0004229.

Shaw, A., Torr, S., Waiswa, C., Cecchi, G., Wint, G. & Mattioli, R. 2013. Estimating the costs of tsetse control options: an example for Uganda. *Preventive Veterinary Medicine,* 110**,** 290 - 303.

Shaw, A. P. M. 2004. Economics of African Trypanosomiasis. *In:* Maudlin, I., Holmes, P. H. & Miles, M. A. (eds.) *The Trypanosomiases*. Wallingford, UK: CABI Publishing.

Shaw, M. K. & Moloo, S. K. 1991. Comparative study on Rickettsia-like organisms in the midgut epithelial cells of different *Glossina species*. *Parasitology,* 102 Pt 2**,** 193-199.

Simarro, P. P., Cecchi, G., Paone, M., Franco, J. R., Diarra, A., Ruiz, J. A., Fèvre, E. M., Courtin, F., Mattioli, R. C. & Jannin, J. G. 2010. The atlas of Human African Trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *International Journal of Health Geogrphics,* 9.

Simarro, P. P., Diarra, A., Postigo, J. A. R., Franco, J. R. & JAnnin, J. G. 2011. The Human African Trypanosomiasis control and surveillance programme

of the World Health Organization 2000–2009: the way forward. *PLoS Neglected Tropical Diseases,* 5.

Simarro, P. P., Franco, J., Diarra, A., Ruiz Postigo, J. A. & Jannin, J. 2012. Update on field use of the available drugs for the chemotherapy of Human African Trypanosomiasis. *Parasitology*.

Simo, G., Fongho, P., Farikou, O., Ndjeuto-Tchouli, P. I. N., Tchouomene-Labou, J., Njiokou, F. & Asonganyi, T. 2015. Trypanosome infection rates in tsetse flies in the "silent" sleeping sickness focus of Bafia in the Centre Region in Cameroon. *Parasites & Vectors,* 8**,** 1-11.

Simo, G., Njitchouang, G. R., Njiokou, F., Cuny, G. & Asonganyi, T. 2011. Trypanosoma brucei s.l.: microsatellite markers revealed high level of multiple genotypes in the mid-guts of wild tsetse flies of the Fontem sleeping sickness focus of Cameroon. *Experimental Parasitology,* 128**,** 272-278.

Simo, G., Silatsa, B., Flobert, N., Lutumba, P., Mansinsa, P., Madinga, J., Manzambi, E., Deken, R. D. & Asonganyi, T. 2012. Identification of different trypanosome species in the mid-guts of tsetse flies of the Malanga (Kimpese) sleeping sickness focus of the Democratic Republic of Congo. *Parasites & Vectors,* 5**,** 1-9.

Smith, C. L., Weiss, B. L., Aksoy, S. & Runyen-Janecky, L. J. 2013. Characterization of the achromobactin iron acquisition operon in *Sodalis glossinidius*. *Applied and Environmental Microbiology,* 79**,** 2872-2881.

Smith, D. H., Pepin, J. & Stich, A. H. R. 1998. Human African trypanosomiasis: an emerging public health crisis. *British Medical Bulletin,* 54**,** 341-355.

Snow, W. F., Tarimo, S. A., Staak, C. & Butler, L. 1988. The feeding habits of the tsetse, *Glossina pallidipes* Austen on the south Kenya coast, in the context of its host range and trypanosome infection rates in other parts of east Africa. *Acta Tropica,* 45**,** 339-49.

Soumana, I. H., Berthier, D., Tchicaya, B., Thevenon, S., Njiokou, F., Cuny, G. & Geiger, A. 2013a. Population dynamics of *Glossina palpalis gambiensis* symbionts, *Sodalis glossinidius*, and *Wigglesworthia glossinidia*, throughout host-fly development. *Infection, Genetics and Evolution,* 13**,** 41-48.

Soumana, I. H., Simo, G., Njiokou, F., Tchicaya, B., Abd-Alla, A. M., Cuny, G. & Geiger, A. 2013b. The bacterial flora of tsetse fly midgut and its effect on trypanosome transmission. *Journal of Invertebrate Pathology,* 112 Suppl**,** S89-93.

Soumana, I. H., Tchicaya, B. L., Loriod, B., Rihet., P. & Geiger, A. 2014. Identification of overexpressed genes in *Sodalis glossinidius* inhabiting trypanosome-infected self-cured tsetse flies. *Frontiers in Microbiology,* 5**,** 1-8.

Spath, J. 1995. Olfactory attractants for West African tsetse flies, *Glossina spp.* (Diptera: *Glossinidae*). *Tropical Medicine and Parasitology,* 46**,** 253-7.

Steverding, D. 2008. The history of African trypsnosomiasis. *Parasites & Vectors,* 1.

Stouthamer, R., Breeuwer, J. A. J. & Hurst, G. D. D. 1999. *Wolbachia Pipientis*: microbial manipulator of arthropod reproduction. *Annual Review of Microbiology,* 53**,** 71-102.

Sumba, A. L., Mihok, S. & Oyieke, F. A. 1998. Mechanical transmission of *Trypanosoma evansi* and *T. congolense* by *Stomoxys niger* and *S. taeniatus* in a laboratory mouse model. *Medical and Veterinary Entomology,* 12**,** 417-22.

Sutton, P., Maskall, J. & Thornton, I. 2002. Concentrations of major and trace elements in soil and grass at Shimba Hills National Reserve, Kenya. *Applied Geochemistry,* 17**,** 1003-1016.

Swallow, B. M. 1999. Impacts of trypanosomiasis on African agriculture. International Centre for Research in Agroforestry, Nairobi, Kenya: International Livestock Research Institute.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution,* 30**,** 2725-2729.

Tarimo-Nesbitt, R. A., Golder, T. K., Dransfield, R. D., Chaudhury, M. F. & Brightwell, R. 1999. Trypanosome infection rate in cattle at Nguruman, Kenya. *Veterinary Parasitology,* 81**,** 107-117.

Taylor, A. W. 1930. Experiments on the mechanical transmission of West African strains of *Trypanosoma brucei* and *T. gambiense* by *Glossina* and other biting flies. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 24**,** 289-303.

Terblanche, J. S., Clusella-Trullas, S., Deere, J. A. & Chown, S. L. 2008. Thermal tolerance in a south-east African population of the tsetse fly *Glossina pallidipes* (Diptera, *Glossinidae*): Implications for forecasting climate change impacts. *Journal of Insect Physiology,* 54**,** 114-127.

Thumbi, S. M., McOdimba, F. A., Mosi, R. O. & Jung'a J. O. 2008. Comparative evaluation of three PCR base diagnostic assays for the detection of pathogenic trypanosomes in cattle blood. *Parasites & Vectors,* 46**,** 1-7.

Tirados, I., Esterhuizen, J., Rayaisse, J. B., Diarrassouba, A., Kaba, D., Mpiana, S., Vale, G. A., Solano, P., Lehane, M. J. & Torr, S. J. 2011. How do tsetse recognise their hosts? The role of shape in the responses of tsetse (*Glossina fuscipes* and *G. palpalis*) to artificial hosts. *PLoS Neglected Tropical Diseases,* 5**,** e1226.

Toh, H., Weiss, B. L., Perkin, S. A. H., Yamashita, A., Oshima, K., Hattori, M. & Aksoy, S. 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Research,* 16**,** 149-156.

Torr, S. J. 1989. The host-orientated behaviour of tsetse flies (*Glossina*): the interaction of visual and olfactory stimuli. *Physiological Entomology,* 14**,** 325-340.

Torr, S. J. 1994. Responses of tsetse flies (Diptera: *Glossinidae*) to warthog (*Phacochoerus aethiopicus Pallas*). *Bulletin of Entomological Research,* 84**,** 411-419.

Torr, S. J. & Mangwiro, T. N. C. 2000. Interactions between cattle and biting flies: effects on the feeding rate of tsetse. *Medical and Veterinary Entomology,* 14**,** 400-409.

Torr, S. J., Maudlin, I. & Vale, G. A. 2007. Less is more: restricted application of insecticide to cattle to improve the cost and efficacy of tsetse control. *Medical and Veterinary Entomology,* 21**,** 53-64.

Torr, S. J. & Vale, G. A. 2015. Know your foe: lessons from the analysis of tsetse fly behaviour. *Trends in Parasitology,* 31**,** 95-99.

Truc, P., Lejon, V., Magnus, E., Jamonneau, V., Nangouma, A., Verloo, D., Penchenier, L. & Buscher, P. 2002. Evaluation of the micro-CATT, CATT/T*rypanosoma brucei gambiense*, and LATEX/*T. b. gambiense* methods for serodiagnosis and surveillance of Human African Trypanosomiasis in west and central Africa. *Bulletin of The World Health Organization,* 80**,** 882-6.

Turner, D. A. 1987. The population ecology of *Glossina pallidipes* Austen (Diptera: *Glossinidae*) in the Lambwe Valley, Kenya. I. Feeding behaviour and activity patterns. *Bulletin of Entomological Research,* 77**,** 317-333.

Turner, D. A. & Invest, J. F. 1973. Laboratory analyses of vision in tsetse flies (Dipt., *Glossinidae*). *Bulletin of Entomological Research,* 62**,** 343-357.

Uilenberg, G. 1998. *A field guide for the diagnosis, treatment and preventiom of African Animal Trypanosomosis,* Rome, Italy.

Vale, G. A. 1977. The flight of tsetse flies (Diptera: *Glossinidae*) to and from a stationary ox. *Bulletin of Entomological Research,* 67**,** 297-303.

Vale, G. A. & Hall, D. R. 1985. The use of 1-octen-3-ol, acetone and carbon dioxide to improve baits for tsetse flies, *Glossina spp.* (Diptera: *Glossinidae*). *Bulletin of Entomological Research,* 75**,** 219-232.

Vale, G. A., Hargrove, J. W., Solano, P., Courtin, F., Rayaisse, J.-B., Lehane, M. J., Esterhuizen, J., Tirados, I. & Torr, S. J. 2014. Explaining the host-finding behavior of blood-sucking insects: computerized simulation of the effects of habitat geometry on tsetse fly movement. *PLoS Neglected Tropical Diseases,* 8**,** e2901.

Vale, G. A., Lovemore, D. F., S., F. & Cockbilla, G. F. 1988. Odour-baited targets to control tsetse flies, *Glossina spp.* (Diptera: *Glossinidae*), in Zimbabwe. *Bulletin of Entomological Research,* 78.

Van Den Abbeele, J., Claes, Y., Van Bockstaele, D., Leray, D. & Coosemans, M. 1999. *Trypanosoma brucei spp.* development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis. *Parasitology,* 118**,** 469-478.

Velasco, C., Guarneri, L., Leonardi, A. & Testa, R. 2003. Effects of intravenous and infravesical administration of suramin, terazosin and BMY 7378 on bladder instability in conscious rats with bladder outlet obstruction. *British Journal of Urology International,* 92**,** 131-136.

Vickerman, K. 1973. The mode of attachment of *Trypanosoma vivax* in the proboscis of the tsetse fly *Glossina fuscipes*: an ultrastructural study of the epimastigote stage of the trypanosome. *Journal of Protozoology,* 20**,** 394-404.

Vickerman, K. 1985. Developmental cycles and biology of pathogenic trypanosomes. *British Medical Bulletin,* 41**,** 105-114.

Vincent, I. M., Creek, D., Watson, D. G., Kamleh, M. A., Woods, D. J., Wong, P. E., Burchmore, R. J. S. & B., M. P. 2010. A molecular mechanism for eflornithine resistance in African Trypanosomes. *PLoS Pathogens,* 6**,** 1-9.

Vreysen, M. J. 2001. Principles of area-wide integrated tsetse fly control using the sterile insect technique. *Médecine tropicale,* 61**,** 397-411.

Wall, R. 1990. Ovarian ageing of tsetse flies (*Diptera*: *Glossinidae*) - interspecific differences. *Bulletin of Entomological Research,* 80**,** 85-90.

Walshe, D. P., Lehane, M. J. & Haines, L. R. 2011. Post eclosion age predicts the prevalence of midgut trypanosome infections in *Glossina*. *PLoS ONE,* 6**,** e26984.

Wamwiri, F. N., Alam, U., Thande, P. C., Aksoy, E., Ngure, R. M., Aksoy, S., Ouma, J. O. & Murilla, G. A. 2013. *Wolbachia, Sodalis* and trypanosome co-infections in natural populations of *Glossina austeni* and *Glossina pallidipes*. *Parasites & Vectors,* 6**,** 232.

Wamwiri, F. N., Ndungu, K., Thande, P. C., Thungu, D. K., Auma, J. E. & Ngure, R. M. 2014. Infection with the secondary tsetse-endosymbiont *Sodalis glossinidius* (Enterobacteriales: *Enterobacteriaceae*) influences parasitism in *Glossina pallidipes* (*Diptera: Glossinidae*). *Journal of Insect Science,* 14.

Wang, C. C. 1995. Molecular mechanisms and therapeutic approaches to the treatment of African trypanosomiasis. *Annual Review of Pharmacology and Toxicology* 35, 93-127.

Wang, J., Brelsfoard, C., Wu, Y. & Aksoy, S. 2013a. Intercommunity effects on microbiome and GpSGHV density regulation in tsetse flies. *Journal of invertebrate pathology,* 112, S32-S39.

Wang, J., Hu, C., Wu, Y., Stuart, A., Amemiya, C., Berriman, M., Toyoda, A., Hattori, M. & Aksoy, S. 2008. Characterization of the antimicrobial peptide attacin loci from *Glossina morsitans*. *Insect Molecular Biology,* 17, 293-302.

Wang, J., Weiss, B. L. & Aksoy, S. 2013b. Tsetse fly microbiota: form and function. *Frontiers in Cellular and Infection Microbiology,* 3, 69.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. N. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of The Royal Society B: Biological Sciences,* 360, 1847-1857.

Warnes, M. L. 1995. Field studies on the effect of cattle skin secretion on the behaviour of tsetse. *Medical and Veterinary Entomology,* 9, 284-288.

Wastling, S. L. & Welburn, S. C. 2011. Diagnosis of human sleeping sickness: sense and sensitivity. *Trends in Parasitology,* 27, 394-402.

Wedman, F. E. 1980. African Animal Trypanosomiasis.

Weiss, B. L., Wang, J. & Aksoy, S. 2011. Tsetse immune system maturation requires the presence of obligate symbionts in larvae. *PLoS Biology,* 9, e1000619.

Weiss, B. L., Wang, J., Maltz, M. A., Wu, Y. & Aksoy, S. 2013. Trypanosome infection establishment in the tsetse fly gut is influenced by microbiome-regulated host immune barriers. *PLoS Pathogens,* 9.

Weiss, J. B. 1995. DNA probes and PCR for diagnosis of parasitic infections. *Clinical Microbiology Reviews,* 8, 113-130.

Weitz, B. 1963. The feeding habits of *Glossina*. *Bulletin of World Health Organization,* 28, 711-729.

Welburn, S. C., Arnold, K., Maudlin, I. & Gooday, G. W. 1993. Rickettsia-like organisms and chitinase production in relation to transmission of trypanosomes by tsetse flies. *Parasitology,* 107 (Pt 2), 141-145.

Welburn, S. C., Coleman, P. G., Maudlin, I., Fevre, E. M., Odiit, M. & Eisler, M. C. 2006. Crisis, what crisis? Control of Rhodesian sleeping sickness. *Trends in Parasitology,* 22, 123 - 8.

Welburn, S. C. & Maudlin, I. 1990. Haemolymph lectin and the maturation of trypanosome infections in tsetse. *Medical and Veterinary Entomology,* 4, 43-48.

Welburn, S. C. & Maudlin, I. 1991. Rickettsia-like organisms, puparial temperature and susceptibility to trypanosome infection in *Glossina morsitans*. *Parasitology,* 102 Pt 2, 201-206.

Welburn, S. C. & Maudlin, I. 1999. Tsetse-trypanosome interactions: rites of passage. *Parasitology Today,* 15, 399-403.

Welburn, S. C., Picozzi, K., Fèvre, E. M., Coleman, P. G., Odiit, M., Carrington, M. & Maudlin, I. 2001. Identification of human-infective trypanosomes in animal reservoir of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. *The Lancet,* 358, 2017-2019.

Werren, J. H. 1997. Biology of Wolbachia. *Annual Review of Entomology,* 42, 587.

Whiteside, E. F. 1962. *Drugs, Parasite and Hosts,* Churchill Livingstone, London.

World Health Organization. 2016a. *Human African Trypanosomiasis* [Online]. Available: http://www.who.int/trypanosomiasis_african/vector_control/en/.

World Health Organization. 2016b. *Human AfricanTrypanosomiasis (sleeping sickness)* [Online]. Available: http://www.who.int/mediacentre/factsheets/fs259/en/ [Accessed 14 May 2016 2016].

Wigglesworth, V. B. 1929. Digestion in the tsetse-fly: a study of structure and function. *Parasitology,* 21**,** 288-321.

Willemse, L. P. M. & Takken, W. 1994. Odor-induced host location in tsetse flies (Diptera: *Glossinidae*). *Journal of Medical Entomology,* 31**,** 775-794.

Wilson, S. G., Morris, K. R. S., Lewis, I. J. & Krog, E. 1963. The effects of trypanosomiasis on rural economy: with special reference to the Sudan, Bechuanaland and West Africa. *Bulletin of the World Health Organization,* 28**,** 595-613.

Woolhouse, M. E. J., Bealby, K., McNamara, J. J. & Silutongwe, J. 1994. Trypanosome infections of the tsetse fly *Glossina pallidipes* in the Luangwa Valley, Zambia. *International Journal for Parasitology,* 24**,** 987-993.

Woolhouse, M. E. J. & Hargrove, J. W. 1998. On the interpretation of age-prevalence curves for trypanosome infections of tsetse flies. *Parasitology,* 116**,** 149-156.

Woolhouse, M. E. J., Hargrove, J. W. & Mcnamara, J. J. 1993. Epidemiology of trypanosome infections of the tsetse fly *Glossina pallidipes* in the Zambezi Valley. *Parasitology,* 106**,** 479-485.

Yuval, B. 2006. Mating systems of blood-feeding flies. *Annu. Rev. Entomol.,* 51**,** 413-440.