Estimating the Genetic Diversity and Population Structure of Tsetse Flies and African Trypanosomes in Zambia and Malawi

(ザンビア、マラウイにおけるツェツェバエとアフリカトリパノソーマ原虫の遺伝的多様性と集団構造の推定)

Yukiko Nakamura

Table of Contents

List of Abbreviations

AAT	African animal trypanosomiasis*
AMOVA	analysis of molecular variance
AMPtk	Amplicon Tool Kit
BLAST	Basic Local Alignment Search Tool
CATL	cathepsin L-like cysteine protease
CI	confidence interval
CO1	cytochrome oxidase 1
CSF	cerebrospinal fluid
DDT	dichlorodiphenyltrichloroethane
DNA	deoxyribonucleic acid
DNPW	Department of National Parks and Wildlife
FFLB	fluorescent fragment length barcoding
GMA	game management area
HAT	human African trypanosomiasis*
ITS	internal transcribed spacer
KNP	Kasungu National Park
LZNP	Lower Zambezi National Park
MGMA	Musalangu Game Management Area
MLEE	multilocus enzyme electrophoresis
NP	national park
NWR	Nkhotakota Wildlife Reserve
OTU	operational taxonomic unit
PCA	principal component analysis
PCR	polymerase chain reaction
PCV	packed cell volume
RFLP	restriction fragment length polymorphisms
SHKB	Shikabeta
SIT	sterile insect technique
SRA	serum resistance-associated
TPM	two-phase mutation
VSG	variant surface glycoprotein
WHO	World Health Organization
ZAWA	Zambia Wildlife Authority
gGAPDH	glyceraldehyde 3-phosphate dehydrogenase

gHAT	*Trypanosoma brucei gambiense*-associated HAT
rHAT	*Trypanosoma brucei rhodesiense*-associated HAT
rRNA	ribosomal ribonucleic acid
s.d.	standard deviation

* Although there are differing opinions against the use of the term “trypanosomiasis” and “trypanosomosis” for AAT, we have used the term “trypanosomiasis” throughout this thesis to achieve consistency.

Technical Glossary

π	average nucleotide diversity within locations
A_S	allelic size range
F_{IS}	fisher's inbreeding coefficient
F_{ST}	A fixation index. The information statistic for examining the overall level of genetic divergence among subpopulations
H	number of haplotypes
H_E	expected heterozygosity among polymorphic loci
H_O	observed heterozygosity among polymorphic loci
Hd	haplotype diversity
LD	linkage disequilibrium
N_A	mean number of alleles
Nd	average number of nucleotide differences within locations
Ne	Effective population size
Φ_{ST}	An analogue of F_{ST}.

List of Tables

List of Figures

Notes

The contents of Chapter One have been published in PLOS Neglected Tropical Diseases.

Nakamura Y, Yamagishi J, Hayashida K, Osada N, Chatanga E, Mweempwa C, Chilongo K, Chisi J, Musaya J, Inoue N, Namangala B, Sugimoto C. Genetic diversity and population structure of *Glossina morsitans morsitans* in the active foci of human African trypanosomiasis in Zambia and Malawi. PLoS Negl Trop Dis 13, e0007568, 2019.

The contents of Chapter Two have been accepted by Frontiers in Veterinary Science, section Parasitology.

Nakamura Y, Hayashida K, Delesalle V, Qiu Y, Omori R, Simuunza M, Sugimoto C, Namangala B, Yamagishi J. Genetic diversity of African trypanosomes in tsetse flies and cattle from the Kafue ecosystem. Front Vet Sci In press.

General Introduction

African trypanosomes

Trypanosomes are single-celled parasitic flagellate protozoa, and its systematic position is Superkingdom: Eukaryota, Clade: Discoba, Phylum: Euglenozoa, Class: Kinetoplastea, Subclass: Metakinetoplastina, Order: Typanosomatida, Family: *Trypanosomatidae*, Genus: *Trypanosoma* (Levine et al., 1980). The salivaria group *Trypanosoma* contains subgenera *Duttonella*, *Nannomonas*, *Pycnomonas*, and *Trypanozoon*, in which trypanosomes seen in Africa are commonly called the African trypanosomes (Table 1). In the salivaria group, transmission from the tsetse fly (*Glossina* spp.) vector to mammalian hosts occur among the anterior station of the vector and the parasite is inoculated into the mammalian host on the course of bloodmeal acquisition (Stevens and Brisse, 2004). An exception is *T. vivax*, which allows mechanical transmission by other biting flies such as *Tabanus*, *Chrysops*, *Haematopota*, and *Stomoxus* spp (Desquesnes and Dia 2004).

Salivaria trypanosomes in the mammalian stage have a clonal surface layer of variant surface glycoprotein (VSG), which is highly immunogenic, leading to efficient destruction of the bloodstream parasite population. Each trypanosome has an archive of thousands of VSGs, but only a single VSG is expressed at a time. The transcribed VSG is spontaneously switched at a high frequency (0.1-1% switch/parasite/generation), and thus enables exhibiting antigenic variation to evade immune response in mammalian hosts (Robinson et al., 1999). This immune evasion mechanism has prevented the development of effective vaccines for African trypanosomiasis.

Table 1. Classification of African trypanosomes

Subgenus	Species	Subgroup/lineages	Host range and pathogenicity
Duttonella	*T. vivax*	*T. vivax*	High pathogenicity in cattle
		T. vivax-like	Unknown
Nannomonas	*T. congolense*	Savannah	High pathogenicity in cattle
		Forest	Low pathogenicity in cattle
		Kilifi	Low pathogenicity in cattle
	T. simiae	*T. simiae*	Acute and fatal infection in pigs
		T. simiae Tsavo	Moderate pathogenicity in pigs
	T. godfreyi		Chronic and occasionally fatal infection in pigs
Pycnomonas	*T. suis*		Moderate pathogenicity in pigs
Trypanozoon	*T. b. brucei*		Low pathogenicity in cattle High pathogenicity in horses, camels and dogs
	T. b. rhodesiense		Acute infection in humans
	T. b. gambiense		Chronic infection in humans

Human African trypanosomiasis and African animal trypanosomiasis

African trypanosomes are of significant burden in terms of public health and economy, due to their development of human African trypanosomiasis (HAT) and African animal trypanosomiasis (AAT) in human and livestock, respectively. Parasites responsible for HAT include *T. brucei rhodesiense* and *T. b. gambiense* from the subgenera *Trypanozoon*. While *T. b. gambiense*-associated HAT (gHAT) is distributed in West Africa and has a chronic disease feature, *T. b. rhodesiense*-associated HAT (rHAT) is distributed in East and southern Africa and has an acute disease manifestation (Figure 1). The main reservoir is considered to be human for *T. b. gambiense*, and wild animals for *T. b. rhodesiense*, respectively. Another parasite from the subgenus *Trypanozoon*, *T. b. brucei*, cannot infect human but infects a variety of other mammalian species and causes AAT. Other African trypanosomes responsible for AAT include *T. congolense*, *T. simiae* and *T. godfreyi* of subgenus *Nannomonas*, and *T. vivax* of subgenus *Duttonella* (Table 1). *T. congolense* and *T. vivax* are recognized as the major agent of AAT in bovids. On the other hand, *T. simiae* and *T. godfreyi* are primarily found in suids (Connor and Van Den Bosshe, 2004; Auty et al., 2012).

Clinical symptoms associated with the early stage (hemolymphatic stage) HAT correspond to the invasion of lymph, bloodstream, and other tissues by the trypanosomes. The signs include skin sore (chancre) at the site of tsetse-bite, skin rash and itching, fever and malaise, adenopathy, local edema, and cardiovascular disturbances (Stich et al., 2002). However, these signs are variable and inconsistent, which makes the patient unaware of infection. It may also lead to misdiagnosis to other febrile disease such as malaria. Clinical signs associated with the late stage (meningoencephalitic stage) are disturbances of consciousness and sleep, disorders including tonus, motility and abnormal movements, and mental changes (Blum et al., 2006). Progression into the late stage occurs 300 to 500 days after infection in gHAT, and 3 weeks to 2 months after infection in rHAT. For primary diagnosis and staging of both forms of HAT, direct or indirect parasite detection is critical, since not all current drugs reach the blood-brain barrier and severe side effects are reported (Blum et al., 2001). After the parasitological diagnosis of blood films and cerebrospinal fluid (CSF), a lumbar puncture is conducted for staging.

The clinical signs of AAT are non-specific, including symptoms such as anemia, relapsing fever, enlarged lymph nodes, reduced milk yield, reduced productivity, decreased fertility, and in severe cases, abortion, emaciation, and eventual death (Gardiner, 1989). The diagnosis for AAT also requires direct or indirect parasite detection. However, in many settings, veterinary services and tools for diagnosis is lacking. As a result, it is common for farmers to treat to their cattle on their own decision without diagnosis.

Genetics and molecular epidemiology of African trypanosomes

Prior to the advent in molecular tools, detection and taxonomic groupings of trypanosomes relied on morphological similarities among species in the combination with host range, disease pathology, and geographical distribution. Further on, isoenzyme-based approach (MLEE: multilocus enzyme electrophoresis) and DNA-based molecular approach (RFLP: restriction fragment length polymorphisms, and PCR-based methods) were established to characterize the variation of trypanosome populations (Tibayrenc et al., 1993). PCR-based methods including both the universal detection of all trypanosome species, and methods that aim for the detection of specific trypanosome species and subspecies, have been developed. The former methods are based on the amplification of target regions such as the internal transcribed spacer (ITS) region of ribosomal genes (Desquesnes et al., 2001; Njiru et al., 2005), the glyceraldehyde 3-phosphate dehydrogenase (gGAPDH) gene (Fermino et al., 2015), and the catalytic domain of the cathepsin L-like cysteine protease (CATL) sequence (Lima et al., 1994). The latter includes methods such as the human serum resistance associated (SRA) gene (Welburn et al., 2001; Radwanska et al., 2002), which specifically detects human-infective *T. b. rhodesiense* and is the only method to distinguish from *T. b. brucei*. Assessing the genetic diversity is important in identifying any specific taxonomic unit that may be of epidemiological significance, such as host-specific variants, correlation between disease pathogenesis and geographical distribution.

Treatment and Control of HAT and AAT

- **Treatment and control strategies for HAT**

The drugs available are pentamidine for early stage gHAT, and a combination therapy of nifurtimox and eflornithine for the late stage. Although the nifurtimox-eflornithine combination therapy is an improved treatment compared to melarsoprol or eflornithine monotherapy, it requires hospitalization of patients and cannot be administered at the primary health-care level. Fexinidazole, an oral monotherapy, is now included in the first-line treatment options for gHAT (Lindner et al., 2020).

Fexinidazole can be used in both stages of gHAT, which minimizes the need for the systematic lumbar puncture and CSF examination to discriminate between the stages. The use of fexinidazole against rHAT patients is now in an on-going clinical trial, and for the time being, the treatment for early and late stage rHAT is suramin and melarsoprol, respectively (DNDi). All drugs used for the treatment of HAT is available free of charge through WHO Geneva. However, for many endemic countries, there are problems such as lack of diagnostic infrastructure, and serious side-effects and drug resistance with the available treatments.

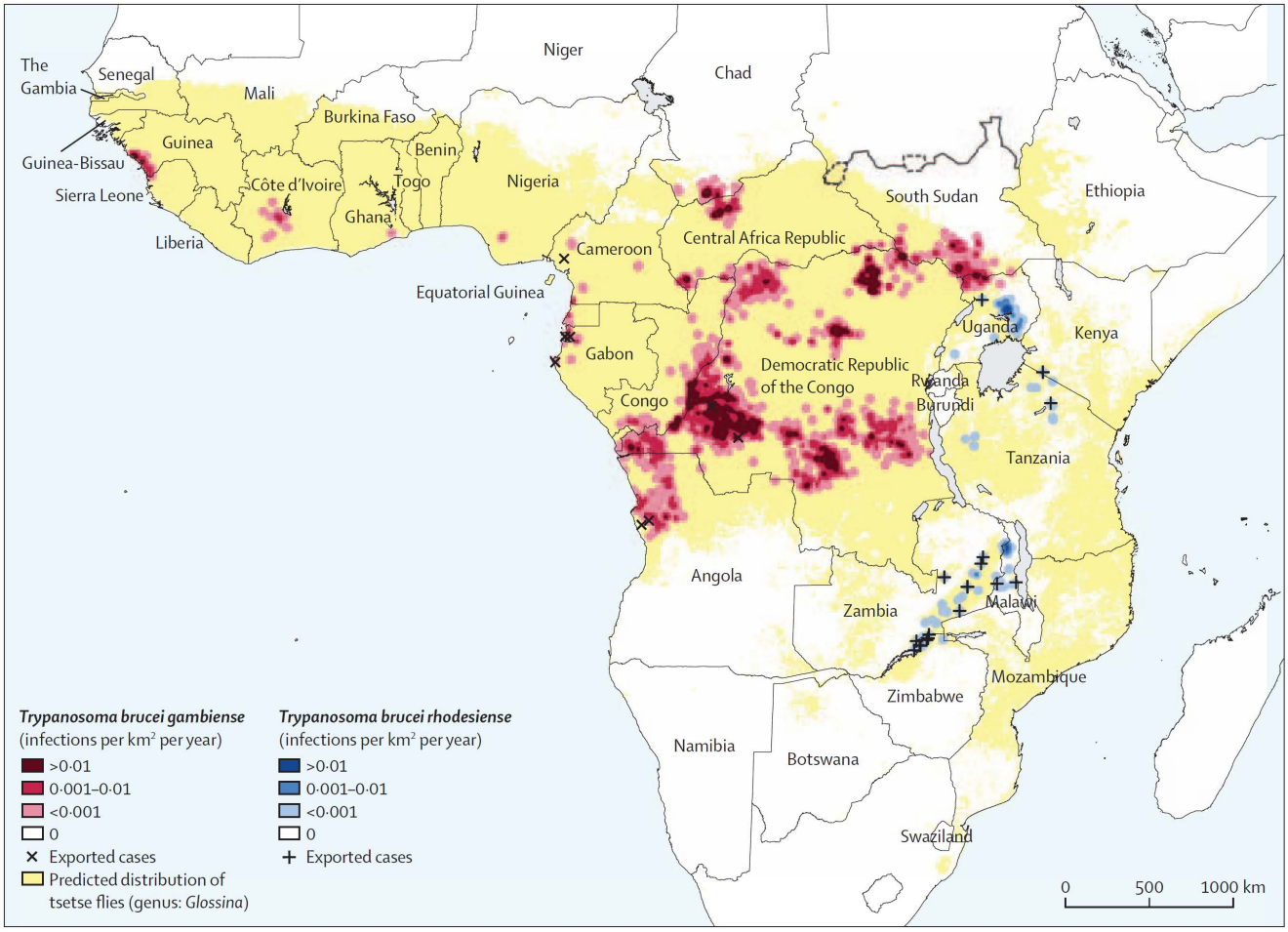


Figure 1. Geographic distribution of reported human African trypanosomiasis cases during 2010– 2014

This figure was reprinted from Büscher et al., Human African trypanosomiasis. Lancet 390, 2397-409, Copyright (2017), with permission from Elsevier.

Due to its anthroponotic nature, the control of gHAT has mainly focused on active and passive surveillance and treatment programs and have succeeded in reducing the newly reported cases of gHAT from 9,680 in 2009 to 953 in 2018. Therefore, the 2020 goal of HAT elimination as a public health problem is considered achievable, and the 2030 goal of the elimination of transmission feasible for gHAT (Franco et al., 2020). On the other hand, because of the zoonotic nature of rHAT, reducing the reservoir of infection is difficult for rHAT, and hence the elimination of rHAT transmission is presently considered unfeasible (Simarro et al., 2015; Aksoy et al., 2017; Franco et al., 2020). Therefore, control strategies for rHAT targets the reduction of transmission instead of complete elimination. In this objective, vector control is the primary strategy, which aims to reduce the vector population and minimize contact between the vector and humans.

- **Treatment and control strategies for AAT**

In most areas where bovine trypanosomiasis is endemic, control is primarily conducted by the use of trypanocides. Three compounds: isometamidium chloride (e.g. Samorin®), homidium (e.g. Novidium®, Ehidium®), and diminazene aceturate (e.g. Berenil®) are available and have been on the market for over 50 years (Geerts et al., 2001). Among these, isometamidium chloride and homidium have prophylactic properties, while diminazene aceturate has therapeutic properties. In addition to the risk of drug resistance, the spread of poor-quality generic products with no or reduced trypanocidal activity makes the treatment less effective and leads to under- or over- dosing, which is also considered as a cause for drug resistance to (Geerts et al., 2001).

The main control strategy for AAT is the combined use of trypanocides, insecticides, and other strategies to avoid contact with tsetse flies. Another approach is exploiting trypanotolerance trait, either by rearing the trypanotolerant breeds or by introducing the genes for trypanotolerance into susceptible cattle (Leak, 1999d). In areas where cattle are constantly kept in tsetse fly infested area, the strategic use of trypanocides will be to start the injection of isometamidium chloride during the start of the rainy season, when the tsetse fly population size starts to increase, and to continue the injection every 2 to 3 months until it reaches the dry season, when the risk of infection is relatively low. For the strategy to be effective, persistent monitoring of the herd and elimination of any established infection by diminazene aceturate is necessary. The use of insecticides on cattle for fly and tick control is also a promising control method (Okelo et al., 2020). Restricted Application Protocol, in which insecticide is applied to a proportionate of the herd to gain benefit for the whole herd, has shown sufficient protection to the herd by spraying as small as 25% of the population (Muhanguzi et al., 2015). However, careful planning should be held and the effectiveness of this method should be assessed in individual situations prior to implementation, since the insecticides valid for both flies and ticks are relatively expensive compared to common acaricide, and long-term use may develop resistance in both flies and ticks.

Tsetse fly classification and distribution

Tsetse flies are under Order: Diptera, Family: *Glossinidae*, Genus: *Glossina*. At present, there are 31 species and subspecies identified, and distributed among approximately 10 million km^2 of the sub-Saharan African countries (Leak, 1998c). The classification of tsetse flies is mainly based on the structural differences of the genitalia. Three subgenera, *Glossina* (Morsitans group), *Nemorhina* (Palpalis group), and *Austenina* (Fusca group), are distinguished base on these morphological differences, and also differ in their ecological characteristics (Table 2). Briefly, Morsitans group tsetse flies distribute in arid to semi-arid savannah woodlands, Palpalis group tsetse flies distribute across semi-arid to sub-humid forests, and Fusca group tsetse flies favor sub-humid to humid forests. In addition to the ecological characteristics, the distribution of Morsitans group tsetse flies often corresponds with the distribution of wild animals (Leak, 1998e).

Life cycle and biology of tsetse flies

A single tsetse fly egg is produced and hatches to first-stage larva in the adult female uterus. Following the development of the second-stage larva, the third-stage larva is deposited on the ground. After the full-grown larva pupate, the adult fly will emerge after 30 to 40 days. The lifespan of an adult fly is up to 3 to 4 months after pupation (Leak, 1998b). An adult female fly will produce one full-grown larva every 9 to 10 days. In a colony, female flies produce at least 10 offspring during their reproductive lives. In the wild, it is likely that fewer numbers of offspring are reproduced. Therefore, tsetse flies are termed as k-strategists with a very low rate of reproduction. Due to their low reproduction rate, little sustained mortality pressure is thought to be required for extinction. In simulations of *G. pallidipes* and *G. morsitans morsitans*, daily elimination of even 1% of a population was suggested to provide effective control (Vale et al., 1985), and mortality of 4% per day imposed on female flies are estimated to be sufficient for extinction if there is no migration (Hargrove, 1988). In addition to the low reproductivity rate, the low dispersal capacity, the low genetic variability, and the exceptional susceptibility to synthetic pyrethroid insecticides, tsetse flies appear to be highly vulnerable targets for control measures.

Table 2. Tsetse fly groups and ecological characteristics

Group	Species/subspecies	Ecological zones	Rainfall (mm)	Main vegetation types
Morsitans (*Glossina*)	*G. morsitans morsitans* *G. morsitans submorsitans* *G. morsitans centralis* *G. pallidipes* *G. austeni* *G. swynnertoni* *G. longipalpis*	Arid ~ Semi-arid	<500 - 1000	Savanna woodland
Palpalis (*Nemorhina*)	*G. palpalis palpalis* *G. palpalis gambiensis* *G. fuscipes fuscipes* *G. fuscipes quanzensis* *G. fuscipes martini* *G. tachinoides* *G. pallicera pallicera* *G. pallicera* *G. caliginea*	Semi-arid ~ Humid	500 - 1500<	Riverine
Fusca (*Austenina*)	*G. fusca fusca* *G. nigrofusca nigrofusca* *G. nigrofusca hopkinsi* *G. medicorum* *G. tabaniformis* *G. brevipalpis* *G. longipennis* *G. frezili* *G. severini* *G. haningtoni* *G. fuscipleuris* *G. vanhoofi*	Sub-humid ~ Humid	1000 - 1500<	Forest

Tsetse control

A summary of tsetse control methods is shown in Table 3. Ground spraying using insecticides have been the first attempts to control tsetse (Leak, 1998f). DDT and dieldrin are residual insecticides which have been used for spraying the resting sites of tsetse flies. Aerial spraying was developed to cover larger areas in tsetse control. Same insecticides as ground spraying were used from fixed-wing aircraft and helicopters. However, due to the direct ecological and environmental impact, and the high cost required, the use of spraying methods have nowadays ceased. Another insecticide-based method, the bait technology uses artificial or biological traps and targets to visually attract tsetse flies, and the use of odor attractants will enhance the trapping efficiency (Leak, 1998h). Examples of bait include traps, targets, and pour-on on cattle. Trap efficiency and attractiveness are largely affected by color, where black and blue is the most efficient and attractive colors. Other non-insecticide-base methods include bush clearing and wild animal elimination which are no longer conducted due to environmental and ethical reasons, sterile insect technique, and biological control. Successful tsetse control requires careful planning using the combination of these techniques. Considering the possibility and magnitude of reinvasion is crucial in succeeding an efficient and effective tsetse control. Using population genetics as means to elucidate the genetic diversity and population structure, several studies have been successful in identifying the population structure and the connectivity among populations in tsetse-infested countries (Beadell et al., 2010; Opiro et al., 2016; Hyseni et al., 2012; Solano et al., 2010).

rHAT and AAT in Zambia and Malawi

Zambia and Malawi are two countries with persistent rHAT cases (Figure 2), with a few to a hundred new reported rHAT cases annually (Simarro et al., 2012; World Health Organization, 2020). In both countries, tsetse flies are distributed mainly inside the protected areas, and intensive cattle rearing is conducted in adjacent areas. A molecular survey conducted in East Zambia have revealed up to 30% prevalence of AAT in cattle, where the majority of infection was by *T. congolense* (Simukoko et al., 2007). In Malawi, 14.2% of the total cattle examined had anti-trypanosomal antibodies (Van den Bossche et al., 2000). Since cattle is the major source of nutrition and income, cost for prophylaxis and treatment and the reduction in livestock products are of great economic loss to the community.

In conclusion, tsetse control is an important strategy for both HAT and AAT. For effective and successful tsetse control, elucidating the population structure of tsetse flies prior to implementation is necessary. In addition, since the diseases occur in remote areas away from the city, lack of information, infrastructure, and early detection methods are burden to successful control of HAT and AAT. Therefore, conducting molecular epidemiology studies on the spot of disease occurrence is essential in understanding the disease epidemiology and constructing effective control.

In Chapter One, the genetic diversity and population structure of tsetse flies from the major rHAT foci were assessed using population genetic methods, with the aim to determine the degree of connectivity between populations and to detect isolated populations for possible tsetse control candidate areas.

In Chapter Two, a molecular epidemiology study was done to detect and characterize the genetic diversity of the trypanosome population found in the same ecosystem where wildlife, livestock, human, and tsetse flies co-exist.

Table 3. Summary of tsetse control methods

	Method	Description	Advantages	Disadvantages
Insecticide-based methods	Ground spraying	Resting sites of tsetse flies are sprayed with insecticides	Relatively low-cost	Labor-intensive. Negative environmental effect.
	Aerial spraying	Insecticides are sprayed from fixed-wing aircraft and helicopters	Can be applied to large-scale areas	Costly. Negative environmental effect.
	Traps and targets	Also known as bait technology, the use of traps and targets are used in combination of odor attractants and insecticides.	Efficient for depopulation of tsetse flies and also works as barriers against re-invasion	Requires maintenance and security against robbery.
	Pour-ons	Cattle or other livestock aretreated with insecticides to kill tsetse flies that feed on them.	Effective in both decreasing the tsetse fly population and preventing AAT	Less effective in areas with low population of livestock. Cannot prevent reinvasion.
Non-insecticide-based methods	Wild animal elimination	Conducted up to the early 1900s, as means to depopulate the blood meal source of tsetse flies		Ethically and environmentally unaccepted.
	Bush clearing	Clearing the vegetation applicable to tsetse fly resting and breeding sites	Still an option for development strategy where land settlement for agriculture is considered	Destructs environment.
	Sterile insect technique	Mass rearing of sterilized (irradiation, chemo sterilization, physiological sterilization) male flies, marking and releasing them into the wild to mate with wild females	The female fly will produce unviable offspring throughout life after it is inseminated by a sterile male,	Costly. Depopulation of non-sterile wild males are needed prior to implementation. Only effective in a closed population without reinvasion.
	Biological control	Parasitic wasps (*Nesolynx* spp.), parasitic mites (*Leptus* spp.), bacteria (*Serratia marcescens*), symbiotic bacteria (*Sodalis*, *Wigglesworthia*, *Wolbachia*), etc.		Careful considerations for the effects against biodiversity of the ecosystem and pathogenicity against other organisms is needed.

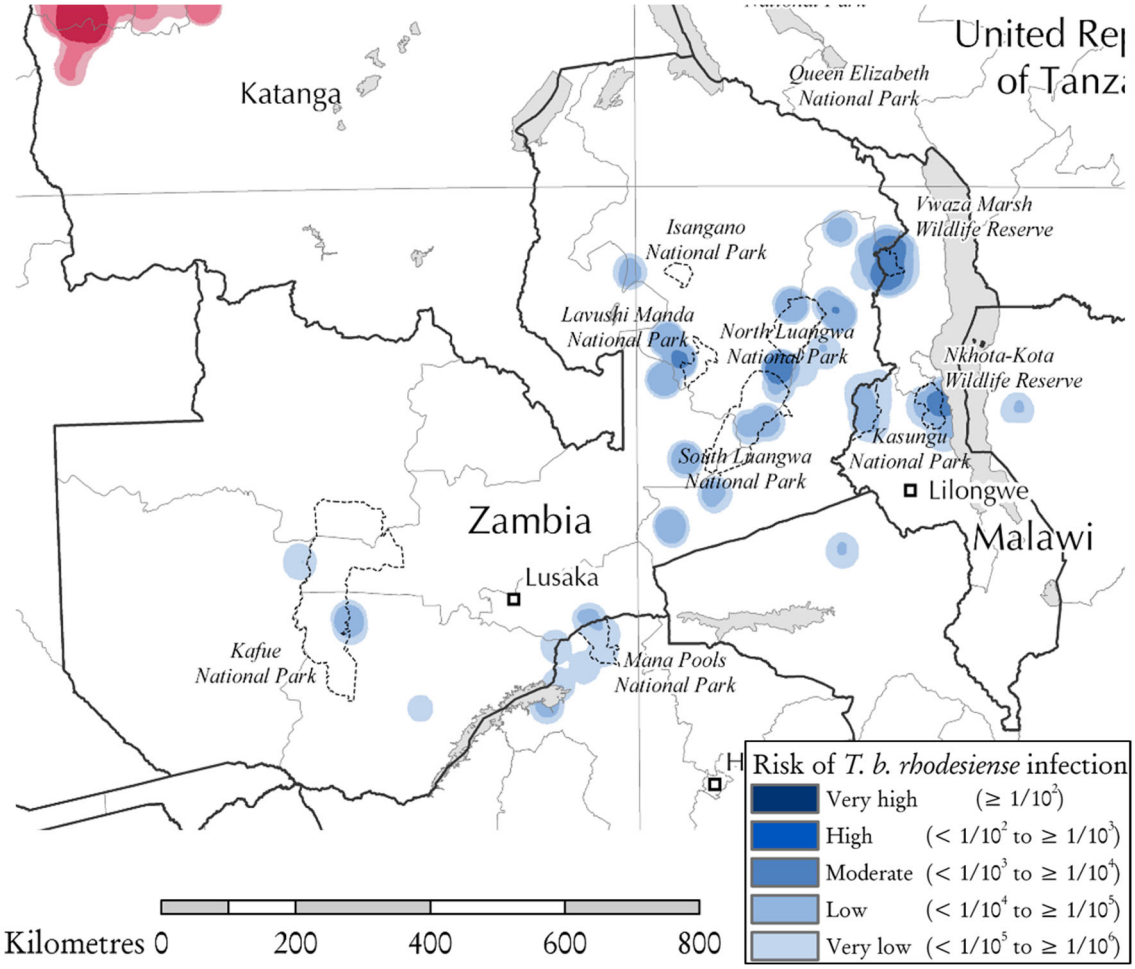


Figure 2. The risk of *Trypanosoma brucei rhodesiense* infection in Zambia and Malawi (2000-2009)

Obtained and modified from Simarro et al., 2012 under the Creative Commons Attribution (CC BY) license *(https://journals.plos.org/plosone/s/licenses-and-copyright)*.

Chapter One

Population genetics of *Glossina morsitans morsitans* in Zambia and Malawi

1.1 Summary

Tsetse flies are distributed across several sub-Saharan African countries between 15ºN and 29ºS. Both female and male tsetse flies are hematophagous and transmit African trypanosomes to mammalian hosts when they acquire bloodmeal. Infection of African trypanosomes cause HAT to human, and AAT to livestock. Due to the lack of vaccines, limited control options, and its large variety of wildlife hosts, prevention and control of both diseases rely largely on controlling the tsetse fly population. For conducting effective tsetse control, prior knowledge against the connectivity of discrete tsetse fly populations is fundamental to prevent re-invasion of the control-implemented area. In Chapter One, the genetic diversity and population structure of *G. m. morsitans* was assessed in the active HAT foci of Zambia and Malawi. One hundred eight and 99 *G. m. morsitans* samples were collected from three locations in Zambia and two locations in Malawi, respectively. Population genetic analyses were conducted using partial mitochondrial cytochrome oxidase gene sequences and ten microsatellite loci. The analyses revealed high level of genetic differentiation between one location in Malawi and the other four locations, suggesting the presence of two different genetic clusters which exists on the East and West side of the escarpment of the Great Rift Valley.

1.2 Introduction

Tsetse flies are the sole vector of the *Trypanosoma* parasites that cause HAT and AAT. Both diseases present a significant burden in terms of public health and economy in sub-Saharan Africa (World Health Organization, 2013). AAT is a wasting disease of livestock causing weight loss, anemia, and in severe cases, abortion and eventual death (Gardiner, 1989). It is estimated to cost livestock producers and consumers US$1.3 billion annually for the treatment or by the loss in livestock products (Kristjanson et al., 1999). HAT, which is also called sleeping sickness, has two forms with different causative parasite species. rHAT is caused by the infection of *T. b. rhodesiense*, and generally has an acute disease progression, in which 80% will become fatal within the first 6 months if left untreated (Brun et al., 2001). Although rHAT is responsible for less than 2% of the overall reported HAT cases, the eradication of this form is considered to be unfeasible due to the presence of wild and domestic animal reservoirs and limited availability of control tools. Moreover, the lack of prophylactic drugs and vaccines, the high cost, severe side effects and development of drug resistance against the available drugs, makes tsetse control the most theoretically desirable method for controlling the disease (Solano et al., 2010) . Suppressing the tsetse fly population by insecticide traps and targets and conducting sterile insect technique (SIT) program have been successful in eradicating *G. austeni* in Zanzibar island, Tanzania (Msangi et al., 1998). The success of this program largely relied on the isolated environment in which reinvasion was impossible. In the case of most land-to-land countries, complete elimination of tsetse fly populations is unfeasible due to reinvasion. Therefore, control, and not elimination, of the tsetse fly population would be the goal of tsetse control. For effective control, it is necessary to identify the extent of the tsetse fly distribution and its connectivity with residing populations. In this objective, several population genetics studies have been successful in identifying population structure and the extent to which the discrete populations are connected by dispersal and migration in several tsetse-infested African countries (Beadell et al., 2010; Opiro et al., 2016; Hyseni et al., 2012; Solano et al., 2010).

Zambia and Malawi, along with other southern African countries, lie within the “common fly belt” in southern Africa (Robinson et al., 1997). In these areas, three subspecies or species of the Morsitans group: *G. m. morsitans*, *G. m. centralis*, and *G. pallidipes*; one species from the Fusca group; *G. brevipalpis*; and one species from the Palpalis group; *G. fuscipes* are found (Wint and Rogers., 2000). Among them, *G. m. morsitans* and *G. pallidipes* are considered the major vectors of AAT and HAT. In fact, the prevalence of *T. b. rhodesiense*, as assessed by the presence of the *T. b. rhodesiense*-specific human serum resistance-associated (SRA) gene, was highest in *G. m. morsitans* (Laohasinnarong et al., 2015). *G. m. morsitans* inhabits savanna woodlands. In East and southern Africa, they are distributed across Tanzania, Mozambique, Zimbabwe, Zambia and Malawi, constituting four allopatric belts (Rogers and Robinson, 2004). Among the four allopatric belts, the genetic diversity and population structure of *G. m. morsitans* in two has been explored using mitochondrial and microsatellite markers.

The studies revealed significant genetic differences among five populations from Zambia, Mozambique and Zimbabwe (Wohlford et al., 1999; Krafsur and Endsley, 2002). Further research involving four populations in Tanzania has also revealed high differentiation and low rates of gene flow, suggesting that the discontinuous distribution has overwhelmed gene flow and lead to genetic drift (Ouma et al., 2007). Although it has been suggested that conducting further studies from other allopatric belts would increase our understanding in the population structure of *G. m. morsitans* (Wohlford et al., 1999), to date, this is the first research including populations from Malawi.

Zambia and Malawi are two countries in East and southern Africa with persisting HAT burden. (World Health Organization, 2013). The Lower Zambezi region and the Luangwa valley in Zambia (Simarro et al., 2012), and Kasungu National Park and Nkhotakota Wildlife Reserve in Malawi (Chisi et al., 2011) are the major HAT foci in these countries (Figure 2). The two countries share foci in the northern region: The Chama district of Zambia and the Rumphi district of Malawi. To investigate the genetic diversity and population structure of *G. m. morsitans* in these major HAT foci, a population genetics study was conducted using a partial 407-bp mitochondrial cytochrome oxidase 1 gene (CO1) sequence and 10 microsatellite loci.

1.3 Materials and Methods

1.3.1 Tsetse sample collection and sample preparation

G. m. morsitans were collected from Lower Zambezi National Park (LZNP, S15º 37.577, E29º 36.132), Shikabeta (SHKB, S14º 57.208, E29º 49.924), and the Musalangu Game Management Area (MGMA, S11º 09.807, E33º 24.224) in Zambia, and Kasungu National Park (KNP, S13º 01.289, E33º 08.517) and the Nkhotakota Wildlife Reserve (NWR, S12º 52.212, E34º 08.283) in Malawi (Figure 3). Sampling within the national parks and GMAs was conducted with permission from the Zambian Wildlife Authority (ZAWA) and the Department of National Parks and Wildlife (DNPW) of Malawi. Samples were collected between May 2012 and February 2018, using mobile traps attached to the rear end of the car. Each trapping drive was conducted within a one-kilometer diameter, a distance which is within the average lifetime dispersal of *G. morsitans* (Jackson, 1948). The captured flies were inspected using light microscopy for morphological identification of species and sexing (Pollock, 1982). Apparently pregnant females were excluded from the analysis. The flies were put into individual 2 mL sample tubes with silica beads to dry. The dried flies were then transferred into new tubes with beads and smashed using a Beads Cell Disrupter (Micro Smash MS-100, Tomy, Tokyo, Japan) at 3,000 rpm for 45 s. DNA was extracted using a modified protocol of DNA Isolation Kit for Mammalian Blood (Roche, Basel, Switzerland). Briefly, 330 μl of white cell lysis buffer was added to each tube, vortexed, and heated at 37°C for 30 min. Subsequently, 170 µl of protein precipitation solution was added, vortexed, and centrifuged at 15,000 rpm for 20 min. Finally, DNA was precipitated by the addition of ethanol. Extracted DNA was stored at −30°C until further use.

1.3.2 Mitochondrial CO1 amplification, sequencing, and analyses

New primers were designed (Gmm_CO1F: 5′–CTT TAC CTG TAT TAG CCG GAG C–3′, Gmm_CO1R: 5′–ACT CCT GTT AAA CCT CCT ACT G–3′) for the amplification of mitochondrial CO1 gene. Reactions included 1–10 ng of template DNA, 1×Ampdirect Plus (Shimadzu, Kyoto, Japan), 0.25 units of BioTaq HS DNA Polymerase (Bioline, Memphis, TN, USA), 10 mM primers, and PCR grade water up to a total volume of 10 μl. Amplification were conducted in an initial denaturation step at 95°C for 10 min, followed by 30 cycles each for 94°C 30 s, 60°C 30 s, 72°C 1 min, and a final extension step at 72°C for 7 min. The PCR products were purified using ExoSAP-IT (Applied Biosystems, Beverly, MA, USA), following the manufacturer's protocol. Purified PCR products were sequenced with ABI 3130/3500xl sequencers (Applied Biosystems, Beverly, MA, USA). The chromatograms were visually inspected, and poor-quality data were trimmed using ApE Plasmid Editor v2.0.51 (M. Wayne Davis, Univ. Utah, USA). The polished forward and reverse sequences were used to create a consensus sequence for each sample, and multiple sequence alignment was made using online MAFFT v7 (Katoh et al., 2017), resulting in 108 fully aligned 407-bp sequences.

The summary statistics calculated were as follows: the average number of nucleotide differences between individuals within locations (*Nd*), the average nucleotide diversity within locations (π), the number of haplotypes within locations (*H*), and haplotype diversity within locations (*Hd*). The summary statistics were analyzed using DnaSP v6.11.01 (Rozas et al., 2017). Neutrality tests, using Tajima's *D*: based on the difference between expected segregating sites, and Fu's F_s: based on the degree of excess of rare alleles, population pairwise Φ_{ST} (an analogue of F_{ST}, which estimates the deviation from random mating among demes) and genetic diversity within and among populations evaluated by the analysis of molecular variance (AMOVA) using the haplotype frequencies distance method, were analyzed using ARLEQUIN v3.5.2.2 (Excoffier et al., 1992; Excoffier and Lischer, 2010). Bonferroni correction was applied for multiple-comparison correction of population pairwise comparisons. A median-joining haplotype network was constructed using POPART v1.7 (Leigh and Bryant, 2015) to infer and visualize the evolutionary relationships of the haplotypes.

1.3.3 Microsatellite amplification, marker validation, and analyses

Twelve autosomal microsatellite loci were used, which have been described in other studies (Baker and Krafsur, 2001; Hyseni et al., 2011; Ouma et al., 2003; Ouma et al., 2006b). Fluorescently-labeled forward primers (FAM, VIC, NED, and PET) were used for amplification, in a reaction containing 1–10 ng template DNA, 0.05 µl Multiplex PCR mix 1 and 5.0 µl Multiplex PCR mix 2 (Multiplex PCR Assay Kit, Takara, Shiga, Japan), 10 mM primers, and PCR grade water up to 10 μl. Amplification conditions were 94°C for 1 min, 30 cycles each of 94°C 30 s, 57°C 90 s, and 72°C 90 s, and a final extension step at 72°C for 10 min.

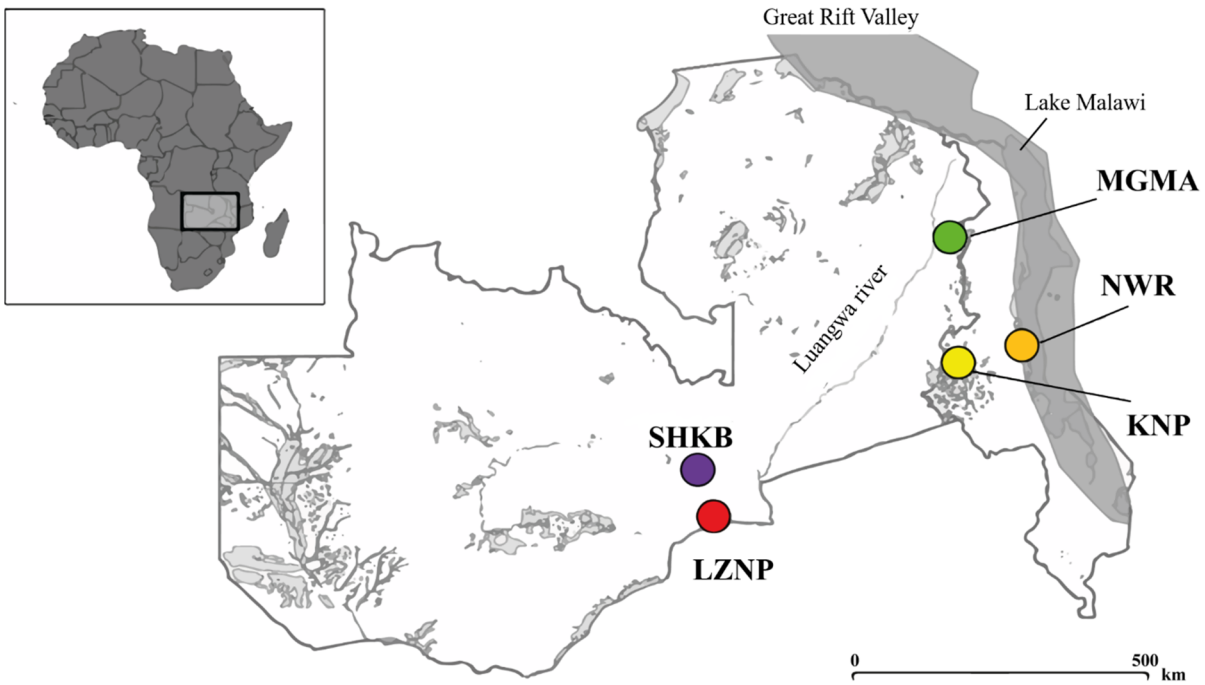


Figure 3. Sampling locations for *Glossina morsitans morsitans*

Markers show the sampling locations for *G. m. morsitans*. The full names of the location codes can be found in Table 6. The map layers were obtained from MapCruzin.com *(https://mapcruzin.com/)*, and the figure was created using the QGIS v3.0 software *(https://qgis.org/en/site/)*. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

PCR products were diluted from ×1 to ×20 according to the thickness of the band from the electrophoresis results, and four different fluorescence-labeled samples were pooled into one to be genotyped on an ABI 3130 sequencer. Alleles were scored using Peak Scanner v1.5 (Applied Biosystems, Beverly, MA, USA) and the scored peaks were manually edited. Micro-Checker v2.2.3 was used to check for null alleles, and two loci were dropped due to null alleles (Van Oosterhout et al., 2004), resulting in 10 valid loci for further analyses.

Genepop files were generated in MSA v4.05 (Dieringer and Schlotterer, 2003) to be used in GENEPOP v4.7 (Rousset, 2008). Test for deviation from Hardy-Weinberg equilibrium and linkage disequilibrium (LD) using the Markov chain method was conducted with 10,000 dememorizations, 1,000 batches and 10,000 iterations per batch (Raymond and Rousset, 1995). For genetic diversity indices, the mean number of alleles (N_A), allelic size range (A_S, the nucleotide length range among microsatellite alleles), expected heterozygosity among polymorphic loci (H_E), observed heterozygosity among polymorphic loci (H_O), estimation of the inbreeding coefficients (F_{IS}, estimation of the deviation from random mating within demes), and population pairwise F_{ST} based on number of different alleles were calculated using ARLEQUIN v3.5.2.2. The genetic diversity within and among populations or groups were calculated by AMOVA analysis as implemented in ARLEQUIN v3.5.2.2 (Excoffier and Lischer, 2010). Bonferroni correction was used for multiple-comparison correction of the population pairwise comparison.

Bayesian clustering method in STRUCTURE v2.3.4 (Pritchard et al., 2000) was used to determine the genetic structure of *G. m. morsitans*. For each K = 1–10, ten replicate runs were carried out with a burn-in length of 20,000 followed by 200,000 iterations. The most likely value of K was determined using the Evanno method (Evanno et al., 2005) implemented in STRUCTURE HARVESTER v0.6.94 (Earl and Von Holdt, 2012). The replicates for the most likely K were aligned using CLUMPP v1.1.2 (Jakobsson and Rosenberg, 2007), and the aligned cluster assignments were visualized using DISTRUCT v1.1 (Rosenberg, 2004). For a different clustering method, the principal component analysis (PCA) was conducted using the "dudi.pca" function in R v3.6.1 (R Core Team, 2013).

The effective population size (*Ne*), including both female and male samples, was estimated for each population using the LD method implemented in NeEstimator v2.1 (Do et al., 2014), and tests for population bottleneck was conducted using the two-phase mutation (TPM) model in BOTTLENECK v1.2.02 (Cornuet and Luikart, 1996), as recommended for microsatellite loci (Di Rienzo et al., 1994). Statistical significance was assessed using Wilcoxon's signed-rank test. Tests for recent bottleneck events were carried out using a mode-shift indicator of allele frequency distributions (Luikart et al., 1998).

Table 4. Primer sequences used throughout the thesis

Method	Primer names	Sequences (5′→3′)	Reference
CO1 PCR	Gmm_CO1F	CTTTACCTGTATTAGCCGGAGC	Designed in this study
	Gmm_CO1R	ACTCCTGTTAAACCTCCTACTG	Designed in this study
Microsatellite	GpCAG133	ATTTTTGCGTCAACGTGA ATGAGGATGTTGTCCAGTTT	Baker and Krafsur, 2001; Ouma et al., 2006b; Hyseni et al., 2011
	GpC101	CCTCAATACAGCAGCAGATG CAAGGTGTGTTGTCGTCTTC	
	GmmF10	TGCCTTTCGATAGAGAAACCATC ACCTGGACACTTATACCGCTC	
	GpC10b	GTTGATGTTGTGATGGTAATGA GCTGGCAAAGAAACTAATGA	
	GmsCAG2	GCTTTTCTCGTCCCATAA GCGTTGTTGATGACTGTG	
	GmmK22	ACGCTTACGTTTCCGTTACAC AAGCTAACCGAACCAGCAC	
	GmmC15	ACTGCATCTGCCTCTGTCG TGAACGAGAAAATGTGAATGGTAAG	
	GmmL11	CCACCACTAACAACGACAGC TGGCTGGTTACAAGATTGCAC	
	GmmA06	ACTTCCATGTTATGTTCGTTGC TGCCTTAGTTGAGAAACTCTGC	
	GmmC17	TGCGCTTTGAACGGAACG CTATGCCGCCTGGCTTATC	
	GmmP07	ACTGACATATTGAGTTGAAAGGGG TCTTCCGTTAAATACAGAGTGCAG	
	GmmL17	CGTACATGCAAGGCAGAGC TCAACTGAAACCGAAAGAGC	
Glossina ITS PCR	GlossinaITS1_for	GTGATCCACCGCTTAGAGTGA	Dyer et al., 2008
	GlossinaITS1_rev	GCAAAAGTTGACCGAACTTGA	
SRA PCR	SRA284F	ATAGTGACAAGATGCGTACTCAACGC	Radwanska et al., 2002
	SRA284R	AATGTGTTCGAGTACTTCGGTCACGCT	
ITS1 PCR	AITSF	CGGAAGTTCACCGATATTGC	Gaithuma et al., 2019
	AITSR	AGGAAGCCAAGTCATCCATC	
CATL PCR	DTO154illumina	ACACTCTTTCCCTACACGACGCTCTTCC GATCTNNACAGAATTCCAGGGCCAATG CGGCTCGTGCTGG	Modified from Lima et al., 1994
	DTO155illumina	GTGACTGGAGTTCAGACGTGTGCTCTT CCGATCTNNTTAAAGCTTCCACGAGTT CTTGATGATCCAGTA	

1.4 Results

1.4.1 CO1 genetic diversity and haplotype diversity

Sixteen haplotypes, Hap_1 to Hap_16 (Figure 4, Figure 5, Table 5), were identified as a result of CO1 gene analysis of 108 *G. m. morsitans* flies from five locations. All haplotype sequences have been uploaded to GenBank through accession numbers LC455935-51 and LC458946-53. The number of haplotypes found within each location in Zambia and Malawi varied from three to nine (Table 5). Haplotype diversity (*Hd*) ranged between 0.582 in SHKB and 0.801 in KNP, and nucleotide diversity (π) ranged between 0.002 in SHKB and 0.007 in NWR (Table 6). The most common haplotype, Hap_1, was found in 33 individuals (30.56%) from all sampling locations except NWR, indicating a common ancestry of the *G. m. morsitans* from four locations. The second most common haplotype Hap_3 was found in 18 individuals (16.67%) from four sampling locations. Seven other haplotypes occurred in two or three sampling locations, and the other seven haplotypes were unique to specific localities ("private haplotypes"). Among the private haplotypes, four were singletons, where the haplotype was found in only a single individual. Interestingly, none of the haplotypes were shared between the Malawian locations, KNP and NWR, despite their relatively close geographic distance. The only haplotype that was not unique to NWR was Hap_9 which occurred in both NWR and MGMA (Figure 4, Figure 5, Table 5). On the other hand, KNP shared many common haplotypes with the Zambian locations and seems to be a part of the major haplogroup. The other four haplotypes (Hap_13, Hap_14, Hap_15, and Hap_16) observed in NWR formed a distinctive haplogroup that was separated from the major haplogroup (Figure 4). As a result of neutrality tests, KNP showed significant *p* values for both Tajima's *D* and Fu's F_S ($p < 0.05$, $p < 0.02$ respectively), suggesting the effects of either positive selection or population expansion (Table 7).

The AMOVA analysis indicated that the overall genetic variation within populations was larger (84.29%) than the variation among populations (15.71%), suggesting little genetic structure (Table 8). Population pairwise Φ_{ST} values showed significantly different values after Bonferroni correction ($p < 0.05$) between NWR and other locations, with values ranging from 0.203 to 0.307. Comparisons between LZNP and other locations also showed significant Φ_{ST}, ranging from 0.177 to 0.273 (Table 9).

Table 5. Haplotype distributions among *Glossina morsitans morsitans* populations

The number of sequences from each haplotype are shown for each location. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

	LZNP	SHKB	MGMA	KNP	NWR	Total
Hap_1	4	8	11	10	0	33
Hap_2	13	1	0	0	0	14
Hap_3	6	5	3	4	0	18
Hap_4	1	0	0	0	0	1
Hap_5	1	0	2	1	0	4
Hap_6	0	0	1	2	0	3
Hap_7	0	0	4	3	0	7
Hap_8	0	0	1	1	0	2
Hap_9	0	0	2	0	4	6
Hap_10	0	0	1	1	0	2
Hap_11	0	0	0	1	0	1
Hap_12	0	0	0	1	0	1
Hap_13	0	0	0	0	7	7
Hap_14	0	0	0	0	5	5
Hap_15	0	0	0	0	1	1
Hap_16	0	0	0	0	3	3
	25	14	25	24	20	108

Table 6. Tsetse fly sampling locations and genetic diversity indices

LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve, *n*: number of individuals, *Nd:* average number of nucleotide differences within locations, π: average nucleotide diversity within locations, *H*: number of haplotypes, *Hd*: haplotype diversity, *N*: number of gene copies, N_A: mean number of alleles, A_S: allelic size range, H_E: expected heterozygosity among polymorphic loci, H_O: observed heterozygosity among polymorphic loci, F_{IS}: Fisher's inbreeding coefficient.

Sampling location	Sampling date	CO1					Microsatellite					
		n	***Nd***	***π***	***H***	***Hd***	***N***	***N_A***	***A_S***	***H_E***	***H_O***	***F_{IS}***
LZNP	2012 May	25	1.213	0.003	5	0.670	40	10.1	36.5	0.835	0.900	−0.086
SHKB	2017 Oct	14	0.637	0.002	3	0.582	38	8.8	34.9	0.768	0.890	−0.163
MGMA	2017 Oct	25	1.540	0.004	8	0.780	40	9.8	36.3	0.815	0.860	−0.057
KNP	2018 Mar	24	1.301	0.003	9	0.801	40	8.8	41.8	0.798	0.935	−0.177
NWR	2018 Feb	20	2.868	0.007	5	0.790	40	9.0	40.6	0.817	0.905	−0.111
	Total	108			16		198					
	Mean		1.512	0.004		0.725		9.3	38.0	0.807	0.898	−0.119

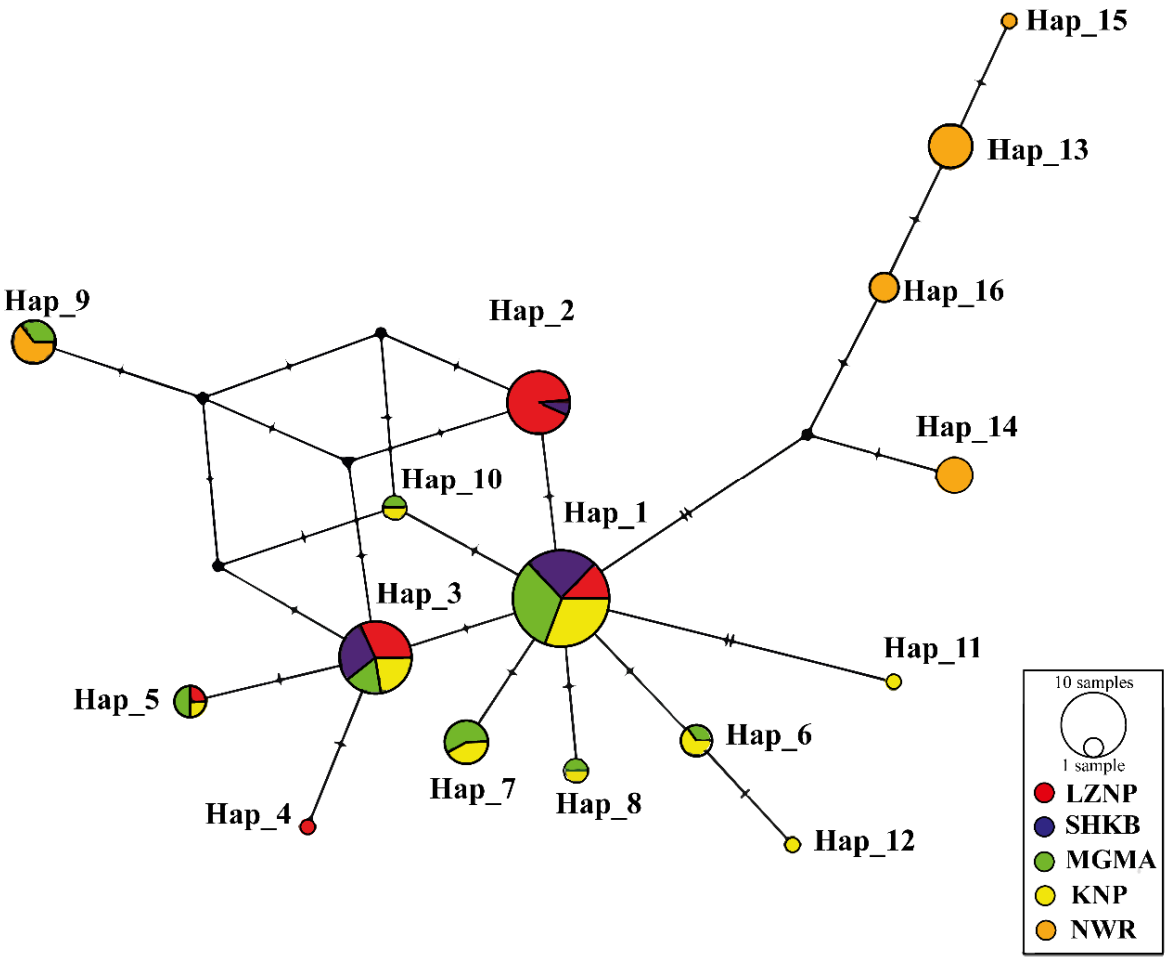


Figure 4. Median-joining haplotype network of CO1 sequences

The median-joining network of 108 CO1 sequences. The network was visualized using POPART v1.7. Each circle represents a haplotype, separated by the number of nucleotide differences shown in hatch marks. The size of a circle is proportional to the number of sequences assigned to that haplotype. The location from which the sequence was obtained is indicated by color in the legend. The black dot represents an intermediate missing haplotype. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

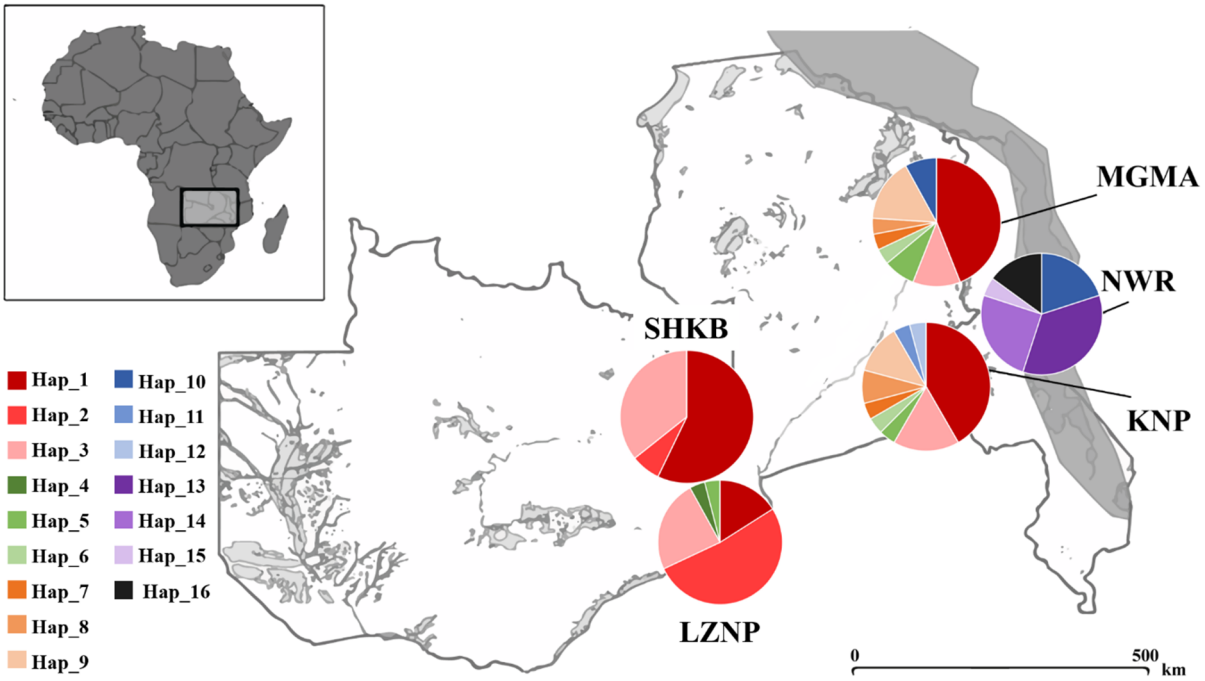


Figure 5. Haplotype distribution

The haplotype (Hap_1 to Hap_16) distribution for each location is shown as a pie graph. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

Table 7. Neutrality test results

* Statistically significant values at Tajima's *D* ($p < 0.05$) and Fu's *F*s ($p < 0.02$). LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve. s.d.: standard deviation.

Statistics	LZNP	SHKB	MGMA	KNP	NWR	Mean	s.d.
Sample size	25	14	25	24	20	21.6	4.722
Tajima's *D*	−0.400	0.037	−0.867	−1.509	0.917	−0.364	0.917
Tajima's *D p*-value	0.443	0.655	0.227	0.041*	0.846	0.442	0.322
No. of alleles	5	3	8	9	5	6.0	2.449
Theta pi	1.213	0.637	1.540	1.301	2.868	1.512	0.828
Exp. no. of alleles	4.270	2.588	4.907	4.396	6.420	4.516	1.376
Fu's *Fs*	−0.335	−0.040	−2.610	−4.683	1.773	−1.179	2.502
Fs p-value	0.420	0.350	0.045	0.004*	0.841	0.332	0.338

Table 8. AMOVA results for comparison of five locations

The genetic diversity within and among populations was evaluated for CO1 sequences and microsatellite allele data using the haplotype frequencies distance method, implemented in ARLEQUIN v3.5.2.2 (Excoffier et al., 1992; Excoffier and Lischer, 2010). *d.f.*: degree of freedom.

	Source of variation	*d.f.*	Sum of squares	Variance components	Percentage of variation	P-values
CO1	Among populations	5	8.431	0.069	15.71	< 0.001
	Within populations	111	41.005	0.369	84.29	
	Total	116	49.436	0.438		
Microsatellite	Among populations	4	32.736	0.106	2.56	< 0.001
	Within populations	191	770.983	4.037	97.44	< 0.001
	Total	195	809.389	4.142		

Table 9. Population pairwise comparison results

Lower diagonal: population pairwise Φ_{ST} comparisons among CO1 sequences based on haplotype differences. Upper diagonal: population pairwise F_{ST} comparisons based on the number of different alleles among microsatellite samples. * $p < 0.05$ after Bonferroni correction. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

	LZNP	**SHKB**	**MGMA**	**KNP**	**NWR**
LZNP		0.015*	0.005	0.009	0.024*
SHKB	0.198*		0.002	0.002	0.075*
MGMA	0.195*	0.029		0.013	0.032*
KNP	0.177*	0.010	−0.031		0.065*
NWR	0.273*	0.307*	0.203*	0.205*	

1.4.2 Microsatellite genetic diversity and population structure

A total of 99 *G. m. morsitans* samples were genotyped using 10 microsatellite loci. None of the microsatellite loci pairs showed significant results of the LD tests. N_A was highest in LZNP (10.100) and lowest in SHKB and KNP (8.800) in the five locations (Table 6). The lowest H_E was observed in SHKB (0.768), and the highest was observed in LZNP (0.835).

The percentage of variation within populations was highest (97.44%) compared to the percentage of variation among populations (2.56%) in the AMOVA analysis (Table 8). The F_{ST} estimate among locations was 0.028 ($p < 0.001$). The pairwise F_{ST} values between NWR and other populations were all statistically significant (Table 9), which ranged from 0.024 (NWR vs. LZNP) to 0.075 (NWR vs. SHKB), indicating small to moderate genetic distance according to Wright's criteria (Wright, 1984). The other pairwise comparison (SHKB vs. LZNP) which showed statistical significance at $p < 0.05$ had a value of 0.015, indicating low genetic distance.

Including all five locations in the STRUCTURE analysis, the Evanno method resulted in the identification of K = 2 genetic clusters (Figure 6). NWR was the only location in which the majority of the genetic cluster is shown as red, whereas the majority of the clusters in the other locations were green (Figure 7). LZNP and MGMA also had small proportion of red cluster, which indicates little genetic mixture between these locations (Figure 7). The PCA showed similar results with STRUCTURE, where NWR shared clusters with LZNP and MGMA (Figure 8).

Ne was estimated using the LD method. SHKB and KNP had 95% confidence intervals which included infinity. LZNP, MGMA, and NWR showed *Ne* of 121.1 (95% CI: 65.0–606.2), 133.5 (95% CI: 65.7–2763.0), and 32.1 (95% CI: 24.2–45.5), respectively. NWR had relatively low *Ne* values compared to the other locations (Table 10). None of the locations were positive using BOTTLENECK analysis under the TPM model (Table 10), and all five populations were approximated to have the expected normal L-shaped allele frequency distributions and are therefore likely to be near mutation-drift equilibrium (Luikart et al., 1998).

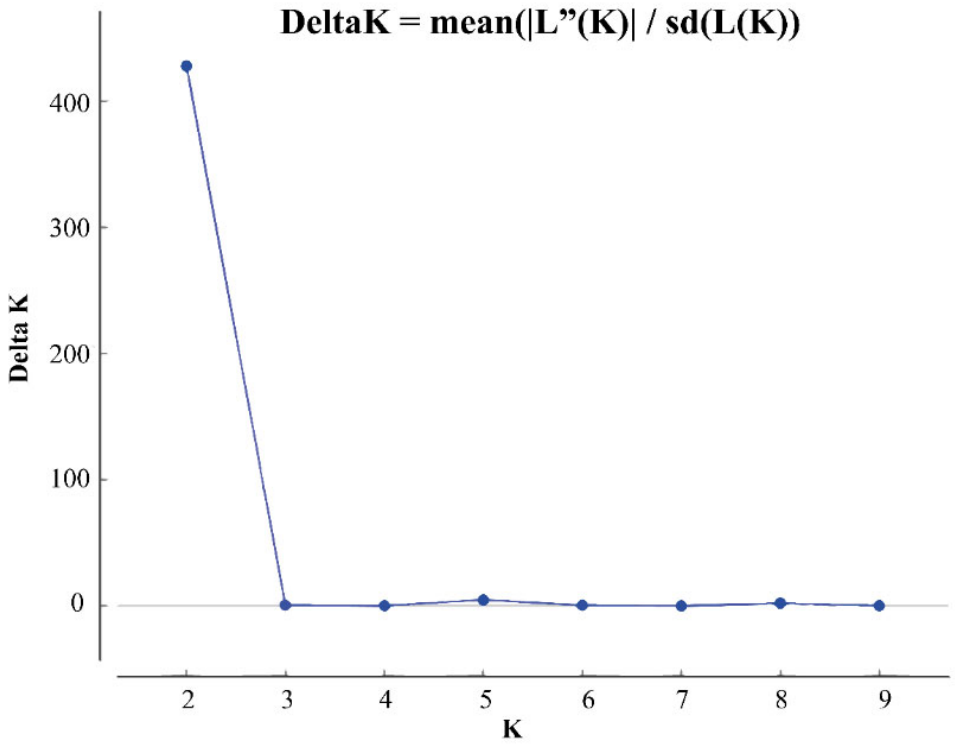


Figure 6. Delta K plot for detecting the most likely K

For each K = 1–10, ten replicate runs were carried out with a burn-in length of 20,000 followed by 200,000 iterations. The most likely value of K was estimated by calculating deltaK, the second order rate of change of the likelihood (Evanno et al., 2005), implemented in STRUCTURE HARVESTER v0.6.94 (Earl and Von Holdt, 2012).

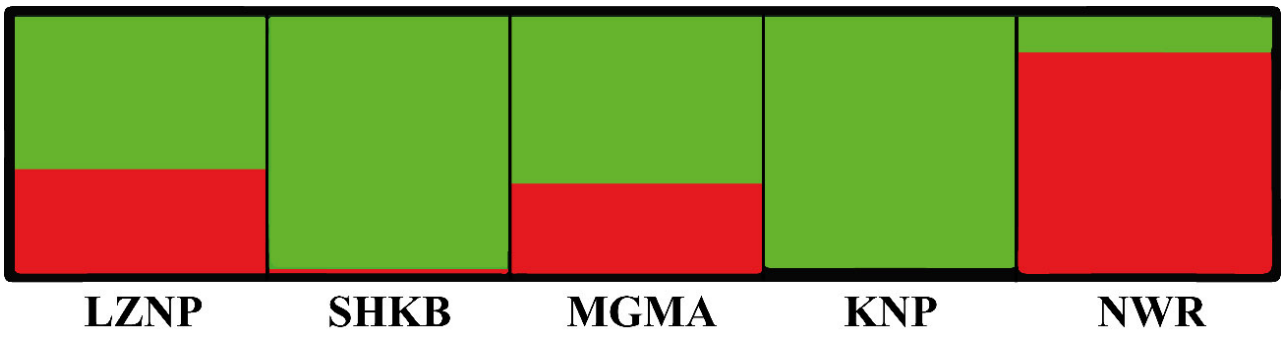


Figure 7. STRUCTURE result

Structure plot for K = 2 (inferred by deltaK). The proportion of the individuals that belong to each cluster is shown in different colors for each sampling location. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

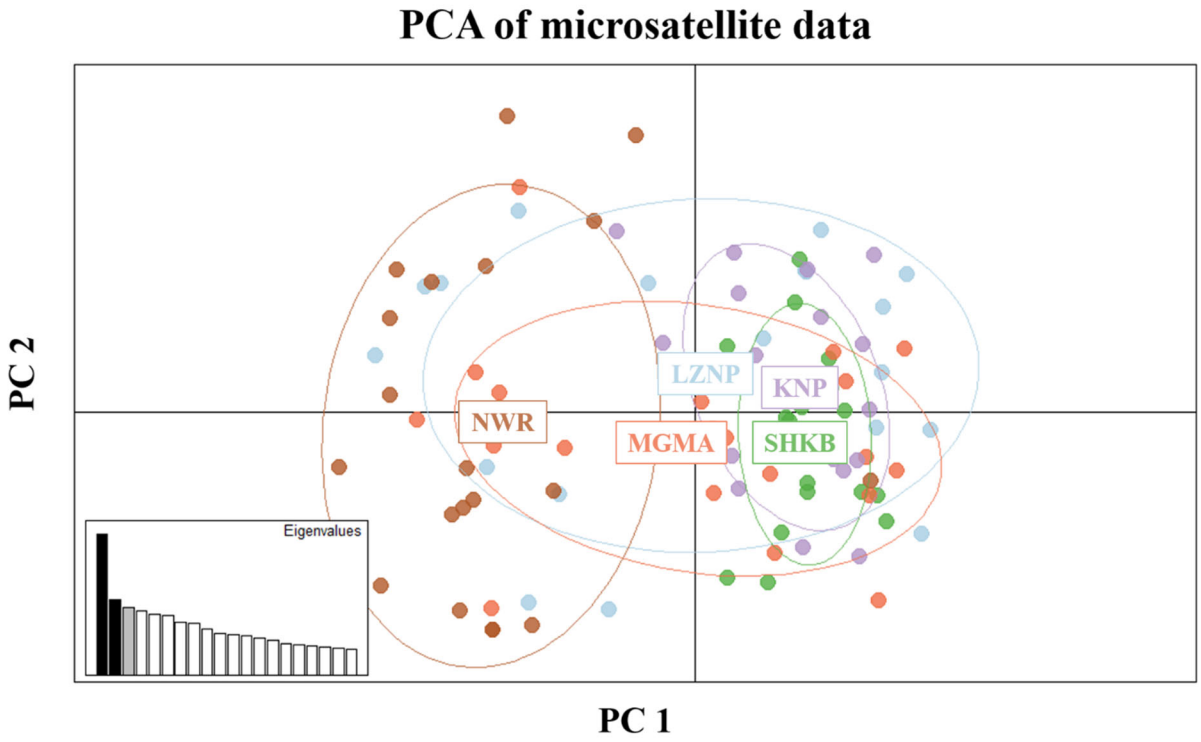


Figure 8. PCA of microsatellite data

The samples are colored according to its sampled location. The eigenvalues are shown in the bottom-left insert. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

Table 10. Effective population size estimates and tests for bottlenecks

LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve. *Ne*: Effective population size, TPM: two-phase mutation model.

	Sample size	***Ne***	**95% Cl**	**TPM P-value**	**Mode-shift**
LZNP	20	121.1	65.0–606.2	0.322	Normal L-shaped
SHKB	19	infinite	183.8–infinite	0.131	Normal L-shaped
MGMA	20	133.5	65.7–2763.0	0.160	Normal L-shaped
KNP	20	infinite	191.7–infinite	0.193	Normal L-shaped
NWR	20	32.1	24.2–45.5	0.625	Normal L-shaped

1.5 Discussion

1.5.1 Population structure

The haplotype network analysis and the pairwise population comparison of the CO1 sequences revealed high differentiation between NWR and the other locations (Figure 4, Table 9). The microsatellite STRUCTURE and PCA analyses also supported this tendency, although a slight genetic mixture between LZNP, MGMA and NWR was observed (Figure 7, Figure 8, Table9). The pairwise population comparison of microsatellite alleles also showed significant F_{ST} values between NWR and the other locations but had relatively low values, ranging from 0.024 to 0.075 (Table 9). Pairwise F_{ST} values ranged between 0.027 and 0.161 in a previous study of *G. m. morsitans* among one population from Zambia, three populations from Zimbabwe, one population from Mozambique, and four populations from Tanzania, and pairwise genetic differences were considered to be high when the F_{ST} value exceeded 0.030 (Ouma et al., 2007). When the same criteria were applied, the overall pairwise genetic differences were high, except for the comparisons of LZNP with SHKB ($F_{ST} = 0.015$) and LZNP with NWR ($F_{ST} = 0.024$). These results suggest that there is limited gene flow between *G. m. morsitans* in NWR from the same subspecies found in other locations, a finding that corresponds well with the estimation of *G. m. morsitans* distribution conducted in previous studies (Robinson et al., 1997; Rogers and Robinson, 2004; Ouma et al., 2007). NWR is adjacent to Lake Malawi, a southern constituent of the Great Rift Valley. This area extends from approximately 3,000 m high escarpment in the West and stretches East up to a narrow plain beside Lake Malawi, which has an altitude of around 500 m (Figure 9). This escarpment is a potential geographical barrier preventing gene flow between NWR and other locations studied, since these areas applies to the upper altitudinal limit for the survival of tsetse flies of 1,600 m to 2,200 m (Tikubet and Gemetchu, 1984; Langridge, 1976). However, from the result of microsatellite analyses, it is also suspected that human activities including tsetse control, change in human demographics, movement of human and livestock, and movement of wildlife may have altered the clear separation between NWR and other locations (especially between MGMA and LZNP).

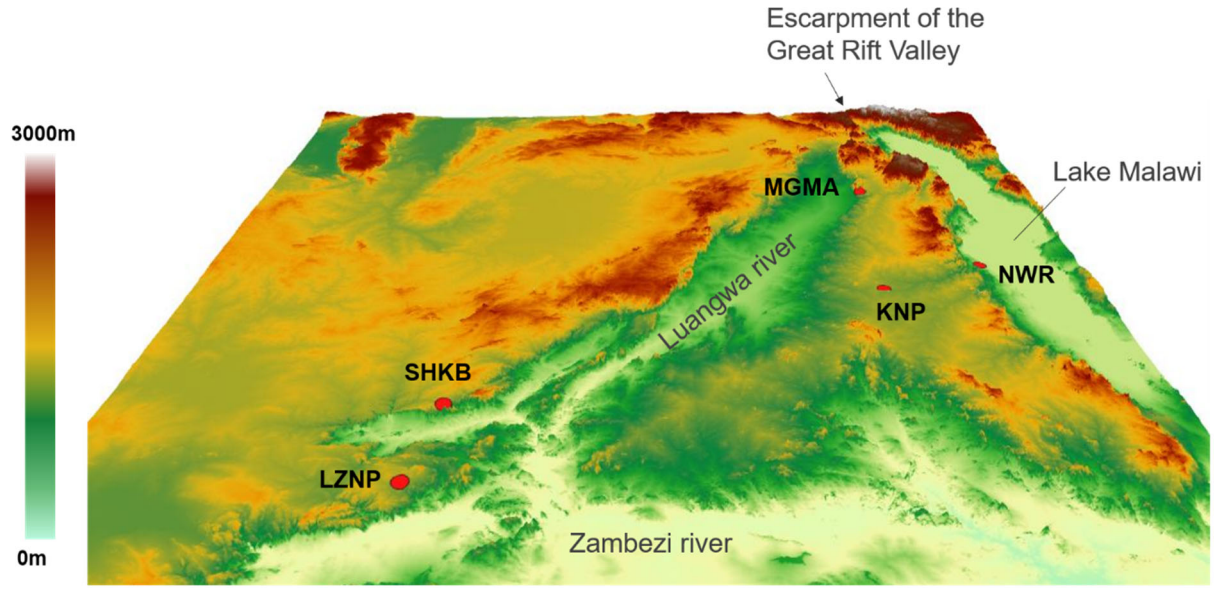


Figure 9. 3D map of sampling locations with altitude

The figure was generated using QGIS Desktop 3.8.0 with GRASS 7.6.1 (*http://qgis.org*), and the shuttle radar topography mission (SRTM) data was downloaded from NASA server (*https://urs.earthdata.nasa.gov/*) using the SRTM Downloader Plugin. LZNP: Lower Zambezi National Park, SHKB: Shikabeta, MGMA: Musalangu Game Management Area, KNP: Kasungu National Park, NWR: Nkhotakota Wildlife Reserve.

1.5.2 Population size changes

High *Hd* and low π were inferred from the results of the CO1 analysis (Table 6). This pattern of high *Hd* and low π is consistent with observations of other *Glossina* species in Uganda and Kenya (Kato et al., 2015; Ouma et al., 2011), and suggests that the populations have experienced a significant population decline resulting in the loss of genetic diversity and subsequently diverged into different haplotypes as the population size recovered. The tsetse fly population in southern Africa is known to have experienced severe decreases due to a rinderpest epizootic in the 1890s, leading up to 90% decline in the size of wildlife populations (Ford, 1971). Since the major blood meal source of *G. m. morsitans* are wildlife, this event is possibly the cause of the historic loss of genetic diversity in southern African countries such as Zambia and Malawi. Subsequently, the *G. m. morsitans* population may have expanded as wildlife recovered from the rinderpest epidemic. This hypothesis, that the high *Hd* and low π observed in this study reflects the state of recovery from the rinderpest epidemic, remains speculative since the BOTTLENECK analysis did not detect a positive event in any of the locations included in this study (Table 10). It is likely that the sample size used in this study was insufficient to detect a bottleneck event 120 years ago, or that sufficient generations have passed between the presumed bottleneck event and the sampling generation to allow re-establishment of a mutation-drift equilibrium. This method is known to be able to detect bottleneck events up to 40–80 generations in the past (Ouma et al., 2007; Luikart et al., 1998), and 120 years is approximately 973 tsetse fly generations (the generations were estimated by using a lifecycle of 45 days per fly; Leak, 1998b). Relatively large *Ne* have been estimated for SHKB and KNP, both including infinity within the 95% confidence interval. Population expansion or selection was suggested for KNP, since this site showed statistically significant negative values in both Tajima's *D* and Fu's F_s (Table 7). Both tests will evaluate the deviation from equilibrium expectations based on the infinite-site model without recombination. However, in order to differentiate population expansion from selection, factors such as the evenness of the distribution of mutation across the whole genome and the ratio between nonsynonymous and synonymous mutations must be explored (Hahn et al., 2002). Therefore, further investigations are required to obtain definite conclusion for population history in KNP. The *Ne* of LZNP, MGMA, and NWR were relatively small. NWR had the smallest *Ne* at 32.1 with a 95% confidence interval of 24.3–45.5 (Table 10). This *Ne* estimate is lower than any *Ne* estimate observed in tsetse flies in the same Morsitans group tsetse fly, *G. pallidipes*. Estimated *Ne* using microsatellites were 180 and 551 in different regions of Kenya (Ouma et al., 2006a). The low *Ne* of NWR and the restricted gene flow between NWR and the other locations indicate that NWR is potentially an isolated location with relatively small population size. Therefore, cost-effective tsetse control activities may be possible. However, the evidence presented in this study is not strong enough to definitely draw this conclusion, since the other locations surrounding Lake Malawi have not been explored in this study. In addition, the *Ne* estimates should be handled with care, since the census population densities are usually much greater than the estimated *Ne*, and the *Ne* estimation is affected

by a number of demographic and genetic phenomena, such as sex ratio and temporal variation in population size (Krafsur and Maudlin, 2018). In this study, the estimation of *Ne* was conducted using both female and male flies. Although care was taken to make the sex ratio 50:50, the possibility of the biasing affect leading to smaller *Ne* estimates cannot be ruled out.

In conclusion, population genetic analyses of a combination of mitochondrial CO1 sequences and nuclear microsatellite markers were successful in illustrating an inter-country population structure of *G. m. morsitans.* One location, NWR in Malawi, showed to be potentially isolated from the other locations, which indicates its potential as a target for cost-effective tsetse control program. On the other hand, the other population including 3 locations from Zambia and one location from Malawi had a high rate of gene flow, indicating the difficulty of accomplishing successful tsetse control. Therefore, in these regions, a more fine-scale analyses should be conducted, focusing on the direction of migration and the degree of spill over to neighboring locations.

Chapter Two

Molecular epidemiology of African animal trypanosomiasis in a human-wildlife-livestock interface in Zambia

2.1 Summary

The genetic diversity of *Trypanosoma* spp. found in cattle and tsetse flies within the Kafue ecosystem in Zambia was assessed, using PCR targeting the ITS1 and CATL region. 12.65% of the total cattle and 26.85% of the total tsetse flies were positive for at least one species of Trypanosomes. Cattle positive for *T. vivax* had significantly lower packed cell volume compared to negative cattle, suggesting that *T. vivax* is the dominant *Trypanosoma* spp. causing anemia. Among the twelve *T. vivax* CATL operational taxonomic units (OTUs) detected, one was from a previously reported *T. vivax* lineage, two OTUs were from previously reported *T. vivax*-like lineages, and nine OTUs were considered to be novel *T. vivax*-like lineages. These findings support previous reports that indicate *T. vivax*-like lineages to be highly diverse. In addition, human-infective *T. b. rhodesiense* was detected in 5.42% of the cattle, indicating the risk of cattle acting as a reservoir of rHAT in the community. Altogether, the findings support that combining CATL PCR with next generation sequencing is useful in assessing *Trypanosoma* spp. diversity.

2.2 Introduction

Bovine trypanosomiasis in Zambia is mainly caused by *T. congolense* and *T. vivax*, and to a lesser extent, by *Trypanozoon* (*T. b. brucei*). High prevalence of these *Trypanosoma* spp. has been reported in eastern and southern Zambia (33.5% prevalence among all cattle and 29.3% prevalence among anemic cattle, respectively), where the majority of infections were by *T. congolense* (Simukoko et al., 2007; Mbewe et al., 2015). Since the clinical symptoms are non-specific, diagnosis of bovine trypanosomiasis requires direct or indirect detection of the parasite, its antigen, or antibodies against the parasite. The most widely used technique is direct parasite detection by wet blood films or Giemsa-stained thick and thin fixed blood films. Although increased sensitivity can be obtained when hematocrit centrifugation technique or the buffy coat technique is used to concentrate the parasite load (Woo, 1969), these techniques require specialized skills to differentiate species, cannot differentiate the subgroups within the *Trypanosoma* spp., and cannot detect parasites in low-parasitemia cases (Uilenberg and Boyt, 1998). To overcome these issues, molecular methods have been established to enable species-specific and cross-species detection and to detect within-species diversity. An example is the PCR targeting the ITS region of ribosomal genes (Desquesnes et al., 2001; Njiru et al., 2005). This method has been commonly used on account of its high sensitivity attributed to the high copy number, and its feasibility of the inter-species discrimination by gel electrophoresis. Sequencing the PCR amplicon can further elevate the sensitivity and allow the differentiation of subgroups within *T. congolense* and, to some extent, *T. vivax* (Gaithuma et al., 2019). However, to elucidate the highly complex disease epidemiology and genetic diversity of *T. congolens*e and *T. vivax*, molecular methods with the resolution to distinguish beyond the subgroups or lineages are needed.

T. congolense comprises of three subgroups (savannah, forest, and kilifi) that differ in pathogenicity, virulence, and geographical distribution (Bengaly et al., 2002). These three subgroups have been reported to coexist in Zambia (Mekata et al., 2008). Experimental infections of susceptible zebu cattle (*Bos indicus*) have revealed kilifi as non-pathogenic, forest as poorly pathogenic, and savannah as the most virulent subgroup (Bengaly et al., 2002). The savannah subgroup is reported to be the most genetically divergent and widespread subgroup across sub-Saharan Africa. Moreover, difference in virulence and drug resistance have been observed in savannah subgroup parasites from the same location (Masumu et al., 2006; Van den Bossche et al., 2011).

T. vivax can be separated into East and West African isolates, and West African *T. vivax* are believed to be more pathogenic than the other (Stephen, 1986). However, severe hemorrhagic outbreaks caused by *T. vivax* with high mortality have also been reported in East Africa (Gardiner et al., 1989), and the correlation between the isolates and disease severity in cattle remains unclear. In addition, great genetic diversity among *T. vivax* isolates from East Africa has been observed (Rodrigues et al., 2008; Cortez et al., 2009; Adams et al., 2010a; Adams et al., 2010b; Auty et al., 2012; Fikru et al., 2014; Fikru et al., 2016). Isolates that are genetically and morphologically distinct from the reference strain *T.*

vivax Y486 have been reported (Cortez et al., 2006; Rodrigues et al., 2008; Auty et al., 2012). Therefore, parasites within the subgenus *Duttonella* separated by relative genetic distance from *T. vivax* Y486 have been widely termed as "*T. vivax*-like" lineage (Rodrigues et al., 2008; Cortez et al., 2009; Adams et al., 2010b; Rodrigues et al., 2017). The taxonomic position of these trypanosomes is still unclear. Notably, a study in Mozambique suggested that a *T. vivax*-like lineage, TvL-Gorongosa, should be elevated to the species status (Rodrigues et al., 2017).

Genetic markers that are used to assess the intraspecific diversity of *T. congolense* and *T. vivax* include the CATL sequences (Lima et al., 1994), the ITS rRNA (ITS1, 5.8S, and ITS2) sequences (Cortez et al., 2006), fluorescent fragment length barcoding (FFLB) using the 18S and 28S ribosomal RNA regions (Hamilton et al., 2008), and the gGAPDH sequences (Fermino et al., 2015). While the gGAPDH sequence is a reliable marker for the identification of species and major intraspecific groups, the relationships within a subgroup of *T. congolense* or among the two *T. vivax* lineages (*T. vivax* and *T. vivax*-like) have not been well-resolved (Rodrigues et al., 2017). FFLB displayed the highest sensitivity in detecting *Duttonella* trypanosomes from tsetse fly samples but could not distinguish within the two *T. vivax* lineages (Hamilton et al., 2008). ITS rRNA PCR and sequencing have high resolution and have been successful in distinguishing within the *T. vivax* and *T. vivax*-like lineage. However, the sensitivity of this approach is inferior to FFLB and requires cloning and sequencing of several clones from each sample (Rodrigues et al., 2017). CATL PCR displayed comparable sensitivity to FFLB (Rodrigues et al., 2017), and the sequences have been used for genotyping a variety of *Trypanosoma* spp. (Lima et al., 1994; Cortez et al., 2009; Ortiz et al., 2009; Rodrigues et al., 2010; Garcia et al., 2011). Therefore, it has promising potential to enable fine-scale genotyping within the *T. vivax* lineages. In fact, CATL genes have been successful in identifying nine clades of *T. vivax* (TviCatL1-9) throughout West Africa, East Africa, and South America (Cortez et al., 2009; Nakayima et al., 2013). *T. vivax* isolates from West Africa shared CATL clades with South America, including TviCatL1-4. *T. vivax* CATL clades that include East and southern Africa isolates showed divergent sequences, including TviCatL5–7 for isolates from Mozambique and TviCatL8–9 from Kenya (Cortez et al., 2009). *T. vivax* from Zambian cattle also clustered with East and southern African sequences (Nakayima et al., 2013). Extending this method to assess the genetic diversity of trypanosomes found in tsetse flies and cattle from the same ecosystem can enable a more in-depth understanding of the parasite community.

In this study, *Trypanosoma* spp. prevalence in tsetse flies and cattle from the Kafue ecosystem was assessed. Combined molecular methods for comprehensive detection and genotyping of African trypanosomes were used. ITS1 PCR was used to detect *Trypanosoma* spp. at the species level, and CATL PCR was used together with next generation sequencing to genotype *Trypanosoma* spp. below the species level.

2.3 Materials and methods

2.3.1 Sample collection and DNA extraction

This study was performed in the Itezhi-Tezhi District, located in the south-western region of the Central Province of Zambia. This area is adjacent to the Kafue National Park (Kafue NP) and Nkala Game Management Area (Nkala GMA). Kafue NP is the oldest and largest NP in Zambia, covering an area of roughly 22,400 km^2 (Mwima, 2001). Kafue NP is surrounded by several GMAs, which act as buffer zones between the protected areas and human settlement, mitigating the adverse effects of human activities. Some human activities, such as licensed safari, subsistence hunting, and agricultural activities are permitted in the GMAs for the local population. The Nkala GMA covers an area of approximately 194 km^2. There is no local population living in this area (Mkanda et al., 2014). Kafue NP and the surrounding GMAs are components of the Kafue ecosystem, a home to a wide variety of flora and fauna, including tsetse flies. *G. m. centralis* is the dominant tsetse fly species in this area, with a lower distribution of *G. pallidipes* (Rogers and Robinson, 2004). Therefore, in many adjacent communities, human settlement, cattle grazing areas, and tsetse infested areas overlap, creating a human-wildlife-livestock interface.

Cattle blood sampling was conducted under ethics approval Ref. No. 2019-Feb-081 (ERES Converge IRB, Lusaka, Zambia). The samples were collected during the end of rainy season (April and May) of 2019 in five villages: Ntubya, Kaminza, Iyanda, New Ngoma, and Basanga (Figure 10). The major cattle breeds in these villages were crosses between local breeds (Tonga and Baila) and exotic breeds (mostly Boran and Brahman). The estimated cattle population was 15,000 heads. Using Cochran's formula (Cochran, 1977) with a confidence interval (CI) of 5 and a 95% confidence level, the required sample size was estimated to be a minimum of 375 heads. Prior to blood sampling, a written consent was provided from the owner of the animal. Subsequently, questionnaire was taken, concerning the basic information of the animal (individual identification, sex, age, and breed), past drug histories, tick control measures, and grazing areas. The information was used to feedback the result to the farmers, and the private data were then anonymized for statistical analysis. Blood samples were collected from the jugular vein and transferred into heparin-lithium tubes. A total of 498 blood samples were randomly collected from 65 farmers. The samples were subjected to micro-hematocrit centrifugation (12,000 rpm, 5 min) to obtain packed cell volume (PCV) values. Thin blood smears were made and subjected to Giemsa staining. The remaining blood samples were preserved at 4°C until DNA extraction. Four epsilon traps were set once for 3 constructive days per each location to catch tsetse flies. The traps were visited two times a day in the morning and afternoon to collect the captured flies. A mobile trap was also used to supply the number of tsetse flies in areas with insufficient catch. Mobile trapping was conducted within a one-kilometer radius, which is within the average lifetime dispersal of morsitans group of tsetse flies (Jackson, 1948). In total, 298 tsetse flies were captured from 10 sampling locations (Figure 10). The captured flies were inspected using a stereo-microscope for morphological

identification of species and sex. The flies were stored in 2 mL sample tubes with silica beads to dry and used for DNA extraction.

DNA was extracted from cattle blood samples using QuickGene DNA whole blood kit S (Kurabo, Osaka, Japan), following the manufacturer's protocol. The dried tsetse flies were smashed using a Micro Smash MS–100 bead cell disrupter (Tomy, Tokyo, Japan) at 3,000 rpm for 45 sec. DNA was extracted using a modified protocol of the DNA Isolation Kit for Mammalian Blood (Roche, Basel, Switzerland). Briefly, 330 µL of white cell lysis buffer was added to each tube, vortexed, and heated at 37°C for 30 min. Subsequently, 170 µL of protein precipitation solution was added, vortexed thoroughly, and centrifuged at 20,400 g (High Speed Refrigerated Micro Centrifuge MX-305, Tomy, Tokyo, Japan) for 20 min. The samples were ethanol precipitated and stored at -30°C until further use.

2.3.2 Taxonomical identification of tsetse flies using Glossina ITS PCR

For molecular identification of the tsetse fly species, Glossina ITS PCR was conducted using GlossinaITS1_for (5'–GTG ATC CAC CGC TTA GAG TGA–3') and GlossinaITS1_rev (5'–GCA AAA GTT GAC CGA ACT TGA–3') primers (Dyer et al., 2008). Reactions contained 1–10 ng DNA, 1× Ampdirect Plus (Shimadzu, Kyoto, Japan), 0.25 U BioTaq HS DNA Polymerase (Bioline, Memphis, TN, USA), 0.2 mM primers, and distilled water to a total volume of 10 µL. Amplification included an initial denaturation step at 95°C for 10 min, followed by 30 cycles each of 94°C for 30 sec, 62°C for 1 min, 72°C for 2 min, and a final extension step at 72°C for 7 min. The bands were visually inspected by gel electrophoresis (*G. pallidipes*: 920 bp, *G. m. centralis*: 800 bp and 150 bp).

2.3.3 *Trypanosoma. brucei rhodesiense* detection using SRA PCR

SRA284F (5'–ATA GTG ACA AGA TGC GTA CTC AAC GC–3') and SRA284R (5'–AAT GTG TTC GAG TAC TTC GGT CAC GCT–3') primers were used to detect *T. b. rhodesiense* (Radwanska et al., 2002). The PCR reagents used were the same as those described above for Glossina ITS PCR. The amplification condition were as follows: an initial denaturation step at 95°C for 10 min, followed by 40 cycles each of 94°C for 30 sec, 60°C for 1 min, 72°C for 1 min, and a final extension step at 72°C for 2 min.

2.3.4 Trypanosome species detection using ITS1 PCR

AITSF (5′–CGG AAG TTC ACC GAT ATT GC–3′) and AITSR (5′–AGG AAG CCA AGT CAT CCAT C–3′) primers were used for ITS1 PCR (Gaithuma et al., 2019). The PCR reagents used were the same as those of the Glossina ITS PCR described above. The amplification condition were as follows: an initial denaturation step at 95°C for 10 min, followed by 37 cycles of 94°C for 30 sec, annealing at 58°C for 90 sec, 72°C for 2 min, and a final extension step at 72°C for 7 min. Trypanosoma spp. were identified based on the different band size of gel electrophoresis.

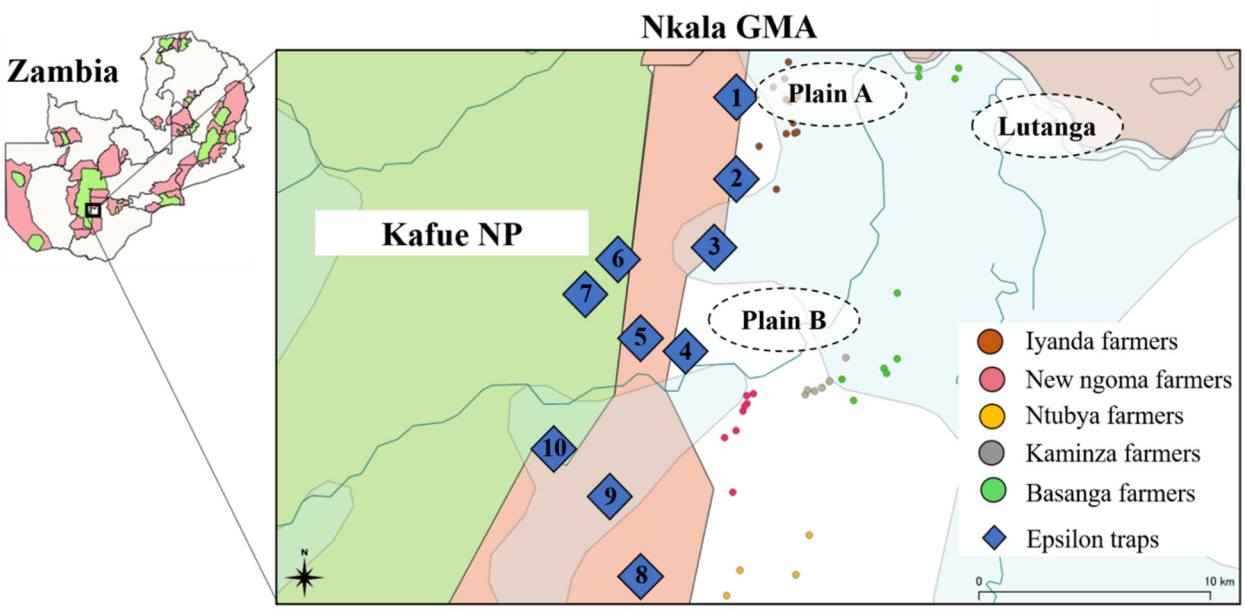


Figure 10. Sampling locations for cattle blood and tsetse fly samples

The map shows the locations of cattle blood sampling among five villages and tsetse fly sampling in ten locations. The approximate locations of the major grazing areas (Plain A, Plain B, Lutanga) are also shown. NP: national park (green), GMA: game management area (pink).

2.3.5 CATL PCR and MiSeq sequencing for trypanosome genotyping

CATL PCR and MiSeq sequencing was carried out to further genotype the ITS1 positive samples using DTO154/DTO155 primers (Lima et al., 1994), which were customized by attaching Illumina adapter sequences (Illumina, San Diego, CA, USA) to the 5′ ends (5'–DTO154illumina: ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT NNA CAG AAT TCC AGG GCC AAT GCG GCT CGT GCT GG–3', 5'–DTO155illumina: GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC TNN TTA AAG CTT CCA CGA GTT CTT GAT GAT CCA GTA–3′). Reactions included 1–10 ng DNA, 0.25 mM primers, 1× Ampdirect Plus (Shimadzu, Kyoto, Japan), 0.25 U BioTaq HS DNA Polymerase (Bioline, Memphis, TN, USA), and distilled water up to 10 µL. PCR grade water was used in place of genomic DNA as a negative control. Amplification included an initial denaturation step at 94°C for 10 min, followed by 37 cycles each of 94°C for 1 min, 60°C for 1 min, 72°C for 90 sec, and a final extension step at 72°C for 10 min. The final PCR was done to attach a unique illumina index to each sample to enable multiplexed sequencing. Reactions contained 1 mM Illumina dual-index primer mix (Illumina, Hayward, CA, USA), KAPA Taq EXtra buffer (Kapa Biosystems, Wilmington, MA, USA), 1 U KAPA Taq EXtra (Kapa Biosystems, Wilmington, MA, USA), 0.2 mM KAPA dNTP Mix (Kapa Biosystems, Wilmington, MA, USA), 1.5 mM $MgCl_2$ (Kapa Biosystems, Wilmington, MA, USA) and distilled water to a total volume of 20 µL. The PCR products were pooled in equal amounts into one library and were analyzed using 2% agarose gel electrophoresis. The band of interest was cut and purified using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). The library was quantified using a Qubit dsDNA HS assay kit and a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), and adjusted to 4 nM using nuclease-free water as the final library. The final library was then sequenced using MiSeq. MiSeq Reagent Kit v3 (Illumina, Hayward, CA, USA) was used for 300 base pairs, paired-end sequencing. 25% PhiX DNA spike-in control was added to increase diversity.

2.3.6 Bioinformatics

The raw reads were polished and grouped into OTUs at 95% identity using the Amplicon Tool Kit (AMPtk) pipeline with default parameters of "*amptk illumina*", "*amptk dada2*", "*amptk filter*", and "*amptk lulu*" (Palmer et al., 2018). The extracted OTUs were manually filtered by excluding reads equal to or less than the negative control. Taxonomy was assigned to the OTUs by using the top hit result of BLASTn (Altschul et al., 1990) above 90% identity and 100% query coverage.

2.3.7 Phylogenetic analyses

The OTU sequences and reference sequences were aligned using MAFFT online v7 (Katoh et al., 2017). The alignment included 257 nucleotides and were used to construct neighbor-joining trees

(Saitou and Nei, 1987) using MEGA X (Kumar et al., 2018). The evolutionary distances were computed by the Maximum Composite Likelihood method (Tamura et al., 2018), using default parameters with 10,000 bootstraps. The tree was then visualized and annotated using iTOL v5.5 (Letunic and Bork, 2019).

2.3.7 Statistical analyses

All statistical analyses and visualization were done in R v3.6.1 (R Core Team, 2013). Fisher's exact tests were conducted to test possible influence of age, sex, village, grazing area, latest injection of isometamidium chloride (Samorin®) and diminazene aceturate (Berenil®), tick control method, and the knowledge about HAT and AAT against the observed positive rate of each parasite species. For pairwise comparisons of *Trypanosoma* spp. prevalence and mean PCV values, statistical significance ($p < 0.05$) was assessed using the Wilcoxon rank sum test with Bonferroni correction.

2.4 Results

2.4.1 Summary of cattle information and questionnaire results

The cattle age distributed between 2 months and 15 years, with a mean of 3.96 years. The total female and male samples were 306 and 191 samples, respectively. The number of cattle included from each village was 105, 85, 90, 102, and 116 for Ntubya, Kaminza, Iyanda, New ngoma, and Basanga, respectively. The cattle were grazing in four types of areas: Lutanga, Plain A, Plain B, and Nkala GMA (Figure 10). All farmers from Basanga grazed their cattle in Lutanga (n = 116), and all farmers from Iyanda grazed their cattle in plain A (n = 90). Farmers from Ntubya, Kaminza and New ngoma shared grazing areas in plain B (n = 73) and inside the Nkala GMA (n = 203). 66 animals have been injected with Berenil® within 2 weeks, and 313 animals have been injected with Samorin® within 3 months. Most of the farmers answered that they regularly use acaricides by either spraying (n = 194) or dipping (n = 283). From the result of Fisher's exact tests, sex, village, grazing area, and tick was significantly correlated with the *T. vivax* prevalence. In addition, the injection of Samorin® had a significant influence on the prevalence of *T. congolense* (Table 11).

2.4.2 African trypanosome prevalence in cattle and tsetse flies

Seven of 498 (1.41%) thin blood smears were positive for *Trypanosoma* spp. by microscopy, which were all from Ntubya village (Table 12). All cattle from the other villages were microscopically negative, and Ntubya had significantly higher prevalence of *Trypanosoma* spp. by microscopy compared to New ngoma and Basanga (Table 12, Table 13). 63 out of 498 cattle (12.65%) were positive for *Trypanosoma* spp. by ITS1 PCR (Table 12). Among villages, Ntubya had the highest prevalence (32 of 105 cattle, 30.48%), followed by New Ngoma (12/102, 11.77%), Kaminza (8/85, 9.41%), Basanga (7/116, 6.03%), and Iyanda (4/90, 4.44%). Ntubya had significantly higher prevalence of the total *Trypanosoma* spp. and of *T. vivax* detected by ITS1 PCR compared to every other villages (Table

12, Table 13). Overall, the most abundant species detected in cattle samples was *T. vivax* (50/498, 10.04%), followed by *T. congolense* (8/498, 1.61%), and *Trypanozoon* (7/498, 1.41%). *T. godfreyi* and *T. simiae* were not detected from any of the cattle samples. All were single infections, in exception of two multiple infections of *T. vivax*/*Trypanozoon* and *T. vivax*/*T. congolense* found in Ntubya (Table 12).

A total of 212 *G. pallidipes* and 86 *G. m. centralis* samples were confirmed by Glossina ITS PCR. As a result of ITS1 PCR, 37 of 86 (43.02%) *G. m. centralis* were positive for one or more *Trypanosoma* spp. (Table 14). Among the 37 positive samples, seven were multiple infections of either *T. vivax*/*T. simiae*, *T. vivax*/*T. congolense*, *T. simiae*/*T. congolense*, or *T. vivax/T. simiae*/*T. congolense* (Table 13). On the other hand, 43 of 212 (20.28%) *G. pallidipes* were positive for one or more *Trypanosoma* spp. Of the 43 positive samples, five were multiple infections of either *T. vivax*/T. congolense, *T. godfreyi*/*T. simiae* and *Trypanozoon*/*T. congolense* (Table 14). In both tsetse fly species included, *T. vivax* was most abundant (52/298, 17.45%), followed by *T. congolense* (26/298, 8.73%), *T. simiae* (9/298, 3.02%), *Trypanozoon* (5/298, 1.68%), and *T. godfreyi* (1/298, 0.34%). *G. m. centralis* had significantly higher prevalence of the total *Trypanosoma* spp., *T. vivax*, and *T. simiae* by ITS1 PCR compared to *G. pallidipes* (Table 14, Table 15). The proportion of multiple infections in the positive samples was comparable among both tsetse fly species (18.92% and 11.63% for *G. m. centralis* and *G. pallidipes*, respectively) (Table 14).

Twenty-seven out of 498 cattle (5.42%) were positive for SRA PCR (Table 12). In descending order, the prevalence in each village was 7.84% at New Ngoma (8/102), 6.67% at Ntubya (7/105), 5.56% at Iyanda (5/90), 5.17% at Basanga (6/116), and 1.18% at Kaminza (1/85). Unfortunately, multiple false-positive bands were identified from the tsetse fly samples. Since no negative controls with PCR grade water turned positive in any batch, it was concluded that contamination during PCR and PCR preparation had not occurred. Therefore, contamination of the SRA gene during DNA extraction was suggested, and all tsetse fly samples were excluded from the SRA PCR experiments.

Table 11. Variables with significant influence against trypanosome prevalence

Variables which resulted to have significant influence on *T. vivax* and *T. congolense* prevalence are shown. (Fisher's exact test, $p < 0.05$)

		No. of *T. vivax* positive samples	No. of *T. vivax* negative samples	*T. vivax* prevalence	Fisher's exact P-value
Sex	Female	23	283	7.52%	0.044
	Male	25	166	13.09%	
Village	Ntubya	29	76	27.62%	0.000
	Kaminza	6	79	7.06%	
	Iyanda	1	89	1.11%	
	New ngoma	11	91	10.78%	
	Basanga	2	114	1.72%	
Grazing area	Plain A	1	89	1.11%	0.000
	Plain B	12	61	16.44%	
	Lutanga	2	114	1.72%	
	GMA	31	172	15.27%	
Tick control	Spraying	34	160	17.53%	0.000
	Dipping	14	269	4.95%	
		No. of *T. congolense* positive samples	**No. of *T. congolense* negative samples**	***T. congolense* prevalence**	**Fisher's exact P-value**
Samorin® injection within 3 months	Yes	2	311	0.64%	0.027
	No	6	168	3.45%	

2.4.3 CATL OTU diversity

A total of 143 ITS1 PCR-positive cattle and tsetse fly samples were subjected to CATL PCR and MiSeq amplicon sequencing. This resulted in obtaining 33 OTUs from 5 *Trypanosoma* spp. (Figure 11). *T. vivax* OTUs were assigned to their lineages (*T. vivax* and *T. vivax*-like) according to the gGAPDH group (Rodrigues et al., 2017) and CATL clades (Cortez et al., 2009) of the reference sequences they cluster with (Figure 12). Nine OTUs, OTU_Tv4–12, did not cluster with any of the reference sequences and were considered to be included in novel clades of the *T. vivax*-like lineage (Figure 11, Figure 12). The phylogenetic relationship of the OTUs and reference sequences among all *Trypanosoma* spp. are shown in Figure 13. The number of OTUs assigned to *T. vivax*, *T. congolense*, *T. simiae*, *T. godfreyi*, and *Trypanozoon* were twelve, fifteen, four, one, and one, respectively (Figure 11). Among them, OTU_Tv1, 2, 3, 4, 6, 7, and 10 (*T. vivax*), OTU_Tc1, 2, 3, and 11 (*T. congolense*), and OTU_Ts1 (*T. simiae*) were found in both cattle and tsetse fly samples. The most abundant OTU in both cattle and tsetse flies was OTU_Tv2 (Figure 11). OTU_Tb1 (*Trypanozoon*) was only detected in cattle OTUs found only in tsetse flies included OTU_Tv5, 8, 9, 11, and 12 (*T. vivax*), OTU_Tc4, 5, 6, 7, 8, 9, 10, 12, 13, 14, and 15 (*T. congolense*), OTU_Ts2, 3, and 4 (*T. simiae*), and OTU_Tg1 (*T. godfreyi*). *G. m. centralis* had the highest diversity of OTUs, in which all OTUs, except OTU_Tv10 and OTU_Tb1, were detected (Figure 11). All OTU sequences have been deposited to GenBank under accession numbers MT673751 to MT673783.

Table 12. Microscopy, ITS1 PCR, and SRA PCR results for cattle samples

The number of cattle blood samples, number of samples positive for each *Trypanosoma* spp. based on microscopic observation, and gel electrophoresis results of ITS1 PCR and SRA PCR are shown in the table. The numbers are the total of individual samples positive with one (single infection) or more (multiple infection) *Trypanosoma* spp. The prevalence by each test was compared between each village by Wilcoxon rank sum test. P-values were adjusted by Bonferroni correction. (*$p < 0.05$. P-values for all pairwise comparisons can be found in Table 13)

Village	no. of samples	Microscopy		ITS1 PCR										SRA PCR	
				T. vivax		*T. godfreyi*		*T. simiae*		*Trypanozoon*		*T. congolense*		Total			
		no. of positives (*n*)	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)
Ntubya	105	7	6.67* (2.72 - 13.25)	29	27.62* (19.34 - 37.2)	0	0.00 (0.00 - 3.45)	0	0.00 (0.00 - 3.45)	1	0.95 (0.02 - 5.19)	4	3.81 (1.05 - 9.47)	32	30.48* (21.87 - 40.22)	7	6.67 (2.72 - 13.25)
Kaminza	85	0	0.00 (0.00 - 4.25)	7	8.24 (3.38 - 16.23)	0	0.00 (0.00 - 4.25)	0	0.00 (0.00 - 4.25)	1	1.18 (0.03 - 6.38)	0	0.00 (0.00 - 4.25)	8	9.41 (4.15 - 17.71)	1	1.18 (0.03 - 6.38)
Iyanda	90	0	0.00 (0.00 - 4.02)	1	1.11 (0.03 - 6.04)	0	0.00 (0.00 - 4.02)	0	0.00 (0.00 - 4.02)	2	2.22 (0.27 - 7.80)	1	1.11 (0.03 - 6.04)	4	4.44 (1.22 - 10.99)	5	5.56 (1.83 - 12.49)
New ngoma	102	0	0.00 (0.00 - 3.55)	11	10.78* (5.51 - 18.48)	0	0.00 (0.00 - 3.55)	0	0.00 (0.00 - 3.55)	0	0.00 (0.00 - 3.55)	1	0.98 (0.02 - 5.34)	12	11.76 (6.23 - 19.65)	8	7.84 (3.45 - 14.87)
Basanga	116	0	0.00 (0.00 - 3.13)	2	1.72 (0.21 - 6.09)	0	0.00 (0.00 - 3.13)	0	0.00 (0.00 - 3.13)	3	2.59 (0.54 - 7.37)	2	1.72 (0.57 - 2.87)	7	6.03 (2.46 - 12.04)	6	5.17 (1.92 - 10.92)
Total	498	7	1.41 (0.57 - 2.87)	50	10.04 (7.54 - 13.02)	0	0.00 (0.00 - 0.74)	0	0.00 (0.00 - 0.74)	7	1.41 (0.57 - 2.87)	8	1.61 (0.70 - 3.14)	63	12.65 (9.86 - 15.89)	27	5.42 (3.60 - 7.79)

Table 13. Result of Wilcoxon rank sum test P-values for each pairwise comparison of *Trypanosoma* spp. detection between villages

The statistical significance of pairwise comparisons of the outcome for each test was assessed by Wilcoxon rank sum test with Bonferroni correction. * $p <$ 0.05

NA: Comparison between zero positive samples

Microscopy

	Basanga	Iyanda	Kaminza	New ngoma
Iyanda	NA			
Kaminza	NA	NA		
New ngoma	NA	NA	NA	
Ntubya	0.019*	0.052	0.063*	0.033*

ITS1 PCR (*T. congolense*)

	Basanga	Iyanda	Kaminza	New ngoma
Iyanda	1.000			
Kaminza	1.000	1.000		
New ngoma	1.000	1.000	1.000	
Ntubya	1.000	1.000	0.705	1.000

ITS1 PCR (*T.vivax*)

	Basanga	Iyanda	Kaminza	New ngoma
Iyanda	1.000			
Kaminza	0.281	0.248		
New ngoma	0.050*	0.059	1.000	
Ntubya	0.000*	0.000*	0.007*	0.022*

ITS1 PCR (Total)

	Basanga	Iyanda	Kaminza	New ngoma
Iyanda	1.000			
Kaminza	1.000	1.000		
New ngoma	1.000	0.682	1.000	
Ntubya	0.000*	0.000*	0.004*	0.010*

ITS1 PCR (*Trypanozoon*)

	Basanga	Iyanda	Kaminza	New ngoma
Iyanda	1.000			
Kaminza	1.000	1.000		
New ngoma	1.000	1.000	1.000	
Ntubya	1.000	1.000	1.000	1.000

SRA PCR

	Basanga	Iyanda	Kaminza	New ngoma
Iyanda	1.000			
Kaminza	1.000	1.000		
New ngoma	1.000	1.000	0.347	
Ntubya	1.000	1.000	0.385	1.000

Table 14. ITS1 PCR results for tsetse fly samples

The number of tsetse fly DNA samples, number of samples positive for each *Trypanosoma* spp. based on the gel electrophoresis result of ITS1 PCR, total of individual samples positive with one or more *Trypanosoma* spp., and the number of samples positive for one (single infection) and more than one (multiple infection) *Trypanosoma* spp. for each tsetse fly species (Gmc: *G. m. centralis*, Gp: *G. pallidipes*) is shown in the table. The prevalence by each test was compared between tsetse fly species by Wilcoxon rank sum test. P-values were adjusted by Bonferroni correction. (* $p < 0.05$. P-values for all pairwise comparisons can be found in Table 15)

Tsetse fly species	no. of samples	ITS1 PCR											
		T. vivax		*T. godfreyi*		*T. simiae*		*Trypanozoon*		*T. congolense*		Total	
		no. of positives (*n*)	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)	*n*	% (95% CI)
Gmc	86	29	33.72* (23.88 - 44.72)	0	0.00 (0.00 - 4.20)	7	8.14* (3.34 - 16.05)	1	1.16 (0.03 - 6.31)	8	9.30 (4.10 - 17.51)	37	43.02* (32.39 - 54.15)
Gp	212	23	10.85 (7.00 - 15.83)	1	0.47 (0.01 - 2.6)	2	0.94 (0.11 - 3.37)	4	1.89 (0.52 - 4.76)	18	8.49 (5.11 - 13.09)	43	20.28 (15.09 - 26.33)
Total	298	52	17.45 (13.32 - 22.24)	1	0.34 (0.01 - 1.86)	9	3.02 (1.39 - 5.66)	5	1.68 (0.55 - 3.87)	26	8.72 (5.78 - 12.52)	80	26.85 (21.9 - 32.26)

Table 15. Result of Wilcoxon rank sum test P-values for each pairwise comparison of *Trypanosoma* spp. detection between tsetse fly species

The statistical significance of pairwise comparisons of the outcome for each test was assessed by Wilcoxon rank sum test with Bonferroni correction.

* $p < 0.05$

ITS1 PCR (*T. vivax*)		*G. m. centralis*
	G. pallidipes	0.000*

ITS1 PCR (*Trypanozoon*)		*G. m. centralis*
	G. pallidipes	0.662

ITS1 PCR (*T. godfreyi*)		*G. m. centralis*
	G. pallidipes	0.529

ITS1 PCR (*T. congolense*)		*G. m. centralis*
	G. pallidipes	0.823

ITS1 PCR (*T. simiae*)		*G. m. centralis*
	G. pallidipes	0.001*

ITS1 PCR (Total)		*G. m. centralis*
	G. pallidipes	0.000*

Subgenus	Species	*T. vivax* lineage / *T. congolense* subgroup	CATL clade	OTU	Cattle								Tsetse fly					
					Ntubya	Kaminza	Iyanda	New ngoma	Basanga	Total by OTU	Total by CATL clade	Total by species	Gp	Total by CATL clade	Total by species	Gmc	Total by CATL clade	Total by species
Duttonella	*T. vivax*	*T. vivax*	TviCatL7	OTU Tv1	7	1	0	1	0	9	9	35	5	5	9	9	9	14
		T. vivax-like	TviCatL8	OTU Tv2	16	1	1	9	0	27	27		7	7		10	10	
			TviCatL9	OTU Tv3	14	1	0	5	1	21	21		5	5		3	3	
			novel	OTU Tv4	1	0	0	0	0	1	1		4	4		7	7	
			novel	OTU Tv5	0	0	0	0	0	0	0		1	1		1	1	
			novel	OTU Tv6	8	0	0	0	0	8	8		5	5		2	2	
			novel	OTU Tv7	4	0	0	0	0	4	4		5	5		4	4	
			novel	OTU Tv8	0	0	0	0	0	0	0		0	0		3	3	
			novel	OTU Tv9	0	0	0	0	0	0	0		0	0		2	2	
			novel	OTU Tv10	1	0	0	0	0	1	1		1	1		0	0	
			novel	OTU Tv11	0	0	0	0	0	0	0		0	0		2	2	
			novel	OTU Tv12	0	0	0	0	0	0	0		0	0		1	1	
Nannomonas	*T. congolense*	savannah	SAV1	OTU Tc1	3	1	1	1	0	6	6	6	1	1	1	4	4	4
				OTU Tc2	1	0	0	0	0	1			0			4		
				OTU Tc3	1	0	0	0	0	1			0			3		
				OTU Tc4	0	0	0	0	0	0			0			3		
				OTU Tc5	0	0	0	0	0	0			0			2		
				OTU Tc6	0	0	0	0	0	0			0			3		
				OTU Tc7	0	0	0	0	0	0			0			3		
				OTU Tc8	0	0	0	0	0	0			0			1		
				OTU Tc9	0	0	0	0	0	0			0			2		
			SAV2	OTU Tc10	0	0	0	0	0	0	0		0	0		1	1	
			SAV3	OTU Tc11	1	0	0	0	0	1			0			4		
				OTU Tc12	0	0	0	0	0	0	1		0	0		4	4	
				OTU Tc13	0	0	0	0	0	0			0			2		
		forest		OTU Tc14	0	0	0	0	0	0	0		0	0		1	1	
		kilifi		OTU Tc15	0	0	0	0	0	0	0		0	0		1	1	
	T. simiae	*T. simiae*		OTU Ts1	1	0	0	0	0	1	1	1	0	0	0	2	2	4
		T. simiae Tsavo		OTU Ts2	0	0	0	0	0	0			0			4		
				OTU Ts3	0	0	0	0	0	0	0		0	0		3	4	
				OTU Ts4	0	0	0	0	0	0			0			3		
	T. godfreyi			OTU Tg1	0	0	0	0	0	0	0	0	1	1	1	2	2	2
Trypanozoon				OTU Tb1	0	0	1	0	0	1	2	1	0	0	0	0	0	0
Total no. of positive samples					23	3	2	9	1				9			16		
Total no. of samples included					32	8	4	12	7				43			37		

Figure 11. The number of cattle and tsetse fly samples positive for the CATL OTUs detected in this study

The number of samples positive for each CATL OTUs are shown. The results are grouped per villages and tsetse fly species for cattle and tsetse flies, respectively. Darkness level 1: the maximum value of each column to up to 50% of the maximum value, level 2: 49% to 25% of the maximum value, level 3: 24% to 5 % of the maximum value. Gp: *G. pallidipes*, Gmc: *G. m. centralis*.

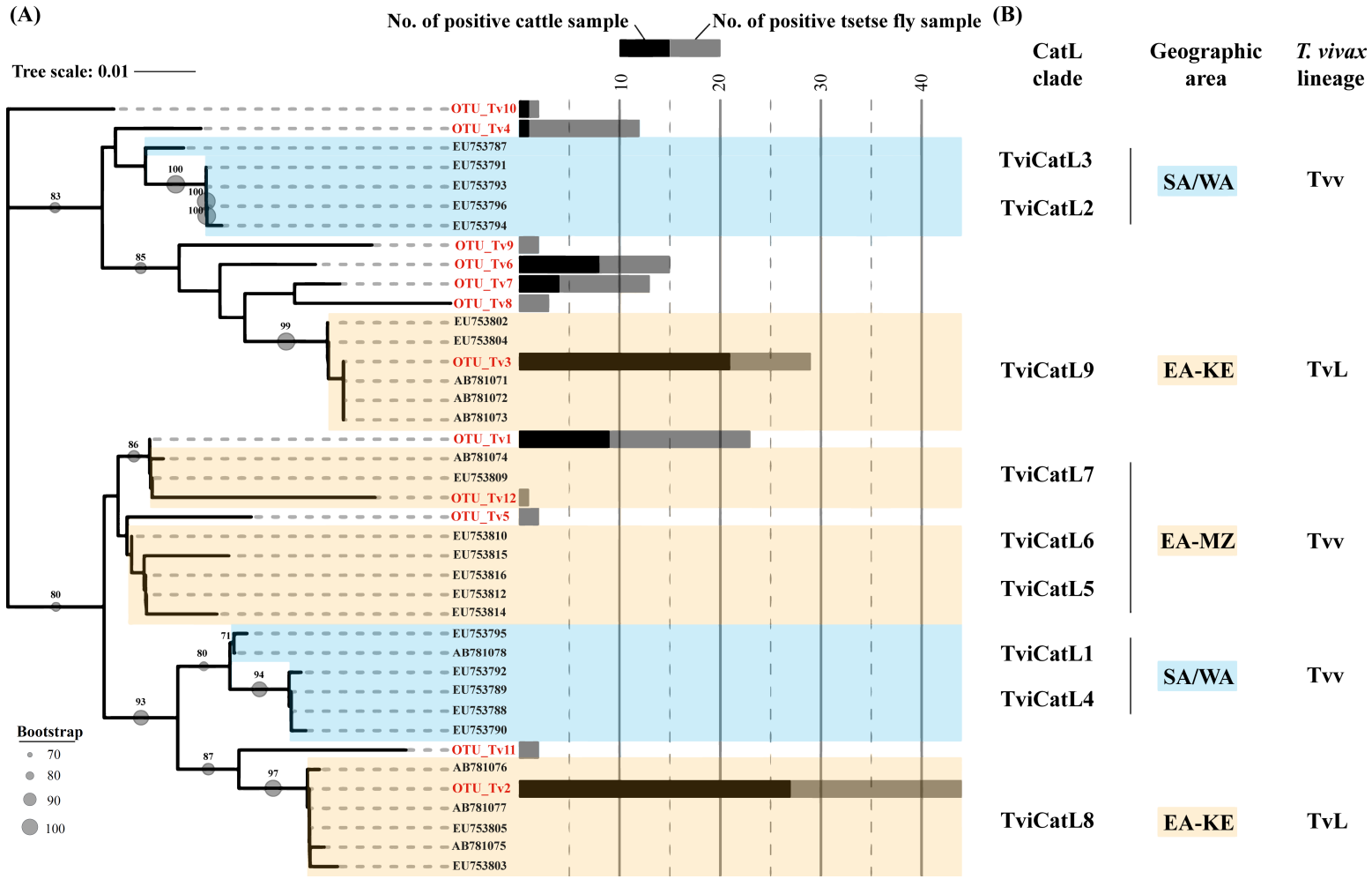


Figure 12. Phylogenetic relationship of *Trypanosoma vivax* CATL OTUs

(A) Subtree of *T. vivax* OTU neighbor-joining tree with reference sequences. Bootstrap values larger than 0.70 are shown in relative size (B) Geographic area and *T. vivax lineages* were assigned to each CATL clade with reference to Cortez et al., 2009 and Rodrigues et al., 2017, respectively. SA: South America, WA: West Africa, EA: East Africa, KE: Kenya, MZ: Mozambique. Tvv: *T. vivax*, TvL: *T. vivax*-like.

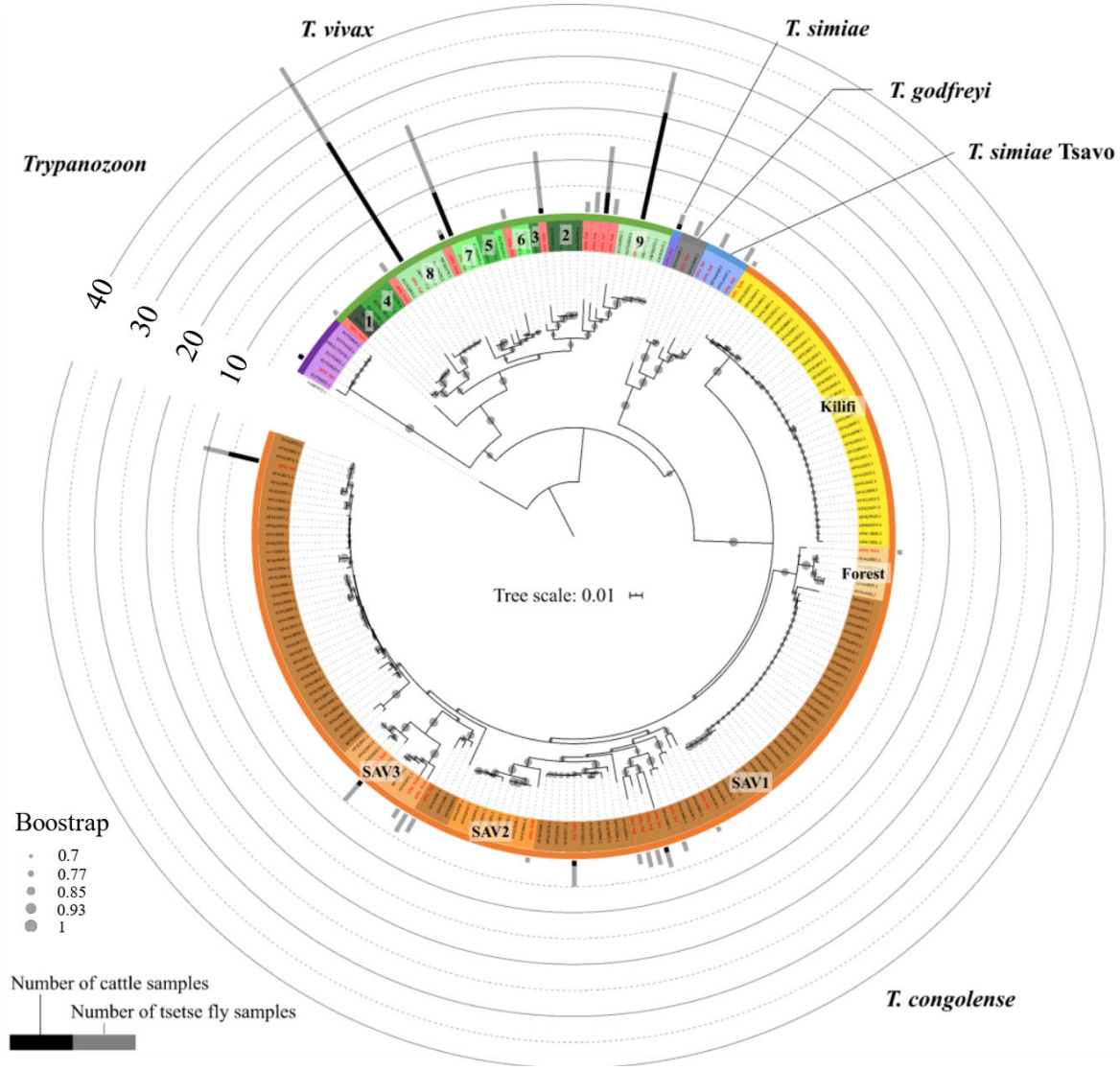


Figure 13. Phylogenetic relationship of CATL OTUs

Neighbor-joining tree of CATL OTUs from all *Trypanosoma* spp. detected in this study with reference sequences. Bootstrap values larger than 0.7 are shown. The outer layer is colored by *Trypanosoma* spp., and the inner layer is colored according to the reported CATL clades (Cortez et al., 2009, Nakayima et al., 2013).

2.4.4 Trypanosome infection and PCV of cattle

The mean PCV value of cattle from each village, regardless of the result of ITS1 PCR, was Ntubya 29.57 (s.d.: 6.80, 95% CI: 28.27-30.87), Kaminza 32.54 (s.d.: 5.40, 95% CI: 31.39-33.69), Iyanda 31.76 (s.d.: 4.31, 95% CI: 30.83-32.67), New Ngoma 33.18 (s.d.: 5.63, 95% CI: 32.08-34.27), and Basanga 32.75 (s.d.: 5.76, 95% CI: 31.70-33.80). Ntubya had the lowest mean PCV, which was lower than other villages when compared in pairwise manners (Table 16). ITS1 PCR-positive cattle had significantly lower mean PCV in Ntubya and New Ngoma (Figure 14). When the cattle samples were grouped by each *Trypanosoma* spp. they were positive to by the result of ITS1 PCR, regardless of the village it was sampled in, animals positive for *T. vivax* had significantly lower mean PCV compared to negative cattle (Figure 15). To check for confounding effects, multiple regression analysis was conducted including village, sex, age, and the outcome of ITS1 PCR for each *Trypanosoma* spp. as coefficients. As a result, *T. vivax* infection significantly decreased PCV values even after excluding all other coefficients ($p = 0.002$). In addition, cattle with a single infection of OTU_Tv2 (*T. vivax*-like, OTU pattern 2) had a tendency of having lower mean PCV compared to cattle negative for all OTUs (OTU pattern 0), but was not statistically significant after correction of multiple comparisons by Bonferroni correction (Figure 16).

Table 16. Result of Wilcoxon rank sum test P-values for each pairwise comparison of cattle PCV between villages

* $p < 0.05$

	Basanga	Iyanda	Kaminza	New ngoma
Iyanda	0.326			
Kaminza	1.000	0.920		
New ngoma	1.000	0.771	1.000	
Ntubya	0.002*	0.315	0.011*	0.002*

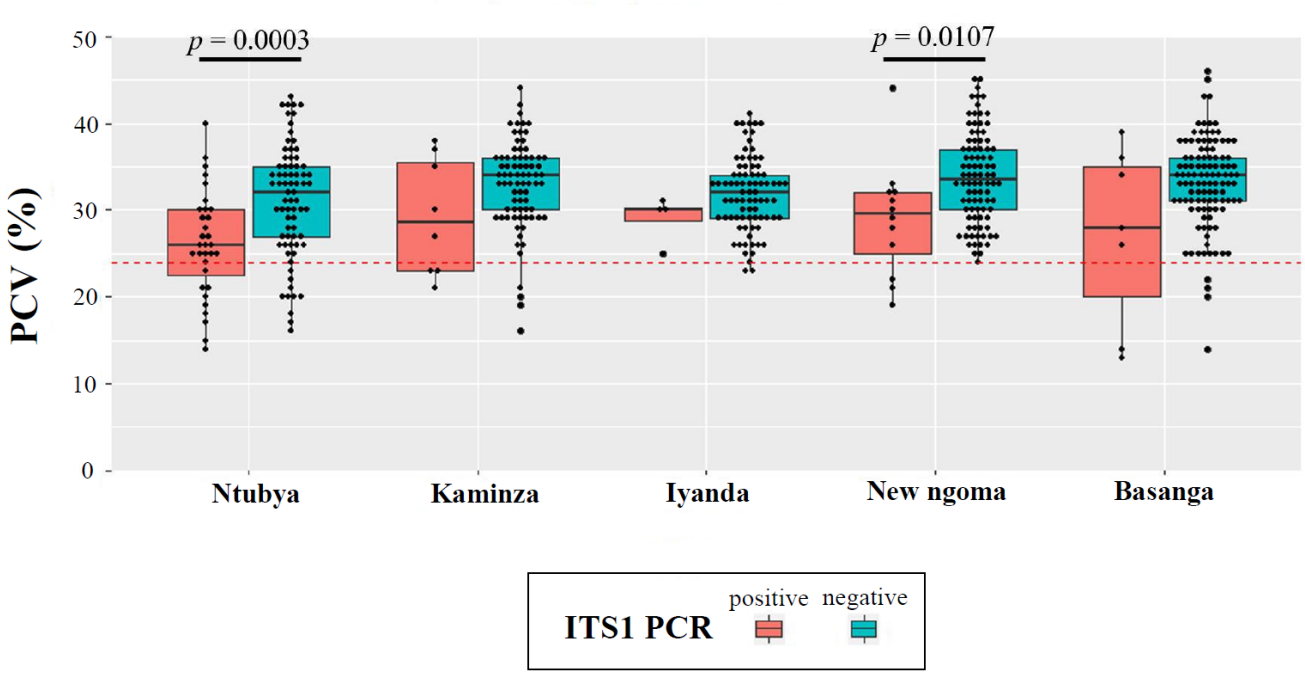


Figure 14. Mean PCV comparisons between villages

The PCV for ITS1 PCR positive and negative cattle are shown for each village. Statistical significance was assessed by Wilcoxon rank sum test, and only the *p* values lower than 0.05 are shown. The dotted red line is drawn at PCV 24% which is commonly used as an indicator for anemia (Mbewe et al., 2015).

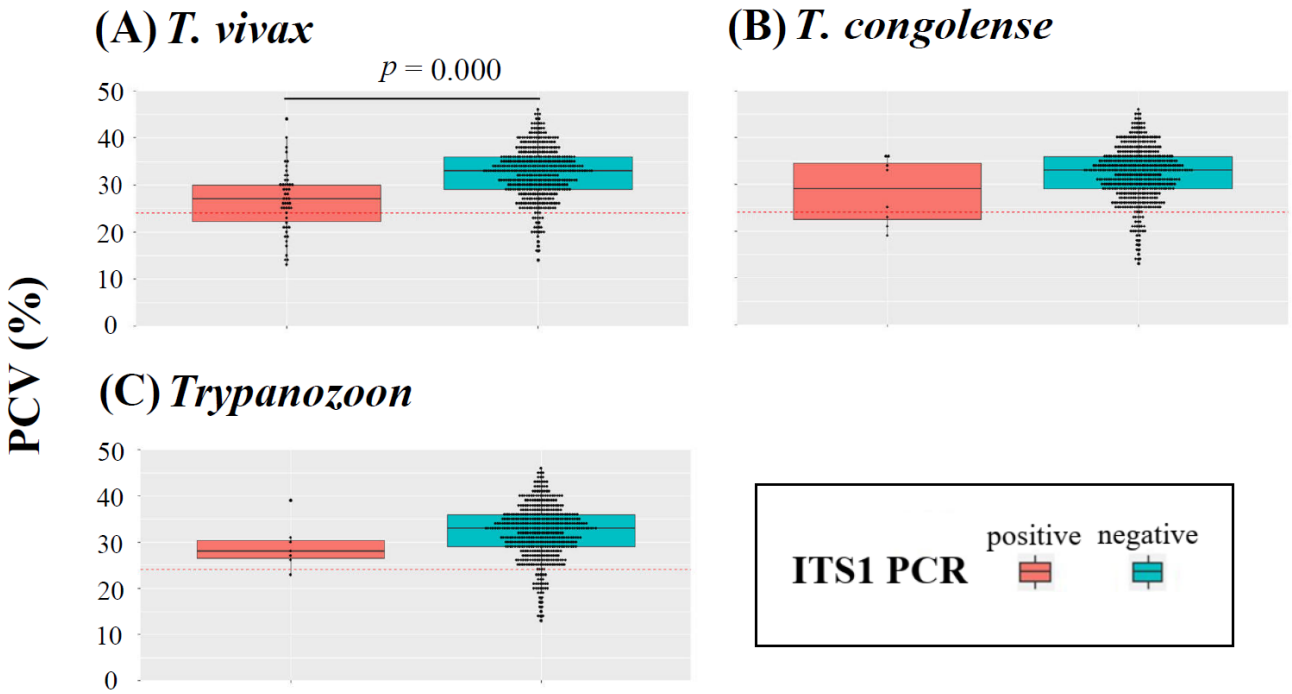


Figure 15. Mean PCV comparisons between ITS1 PCR positive and negative cattle per *Trypanosoma* spp.

The boxplot shows the comparison of the PCV between (A) *T. vivax*, (B) *T. congolense*, (C) *Trypanozoon*-positive and negative cattle samples (confirmed by ITS1 PCR). Statistical significance was assessed by Wilcoxon rank sum test, and only the *p* values lower than 0.05 are shown. The dotted red line is drawn at PCV 24% which is commonly used as an indicator for anemia (Mbewe et al., 2015).

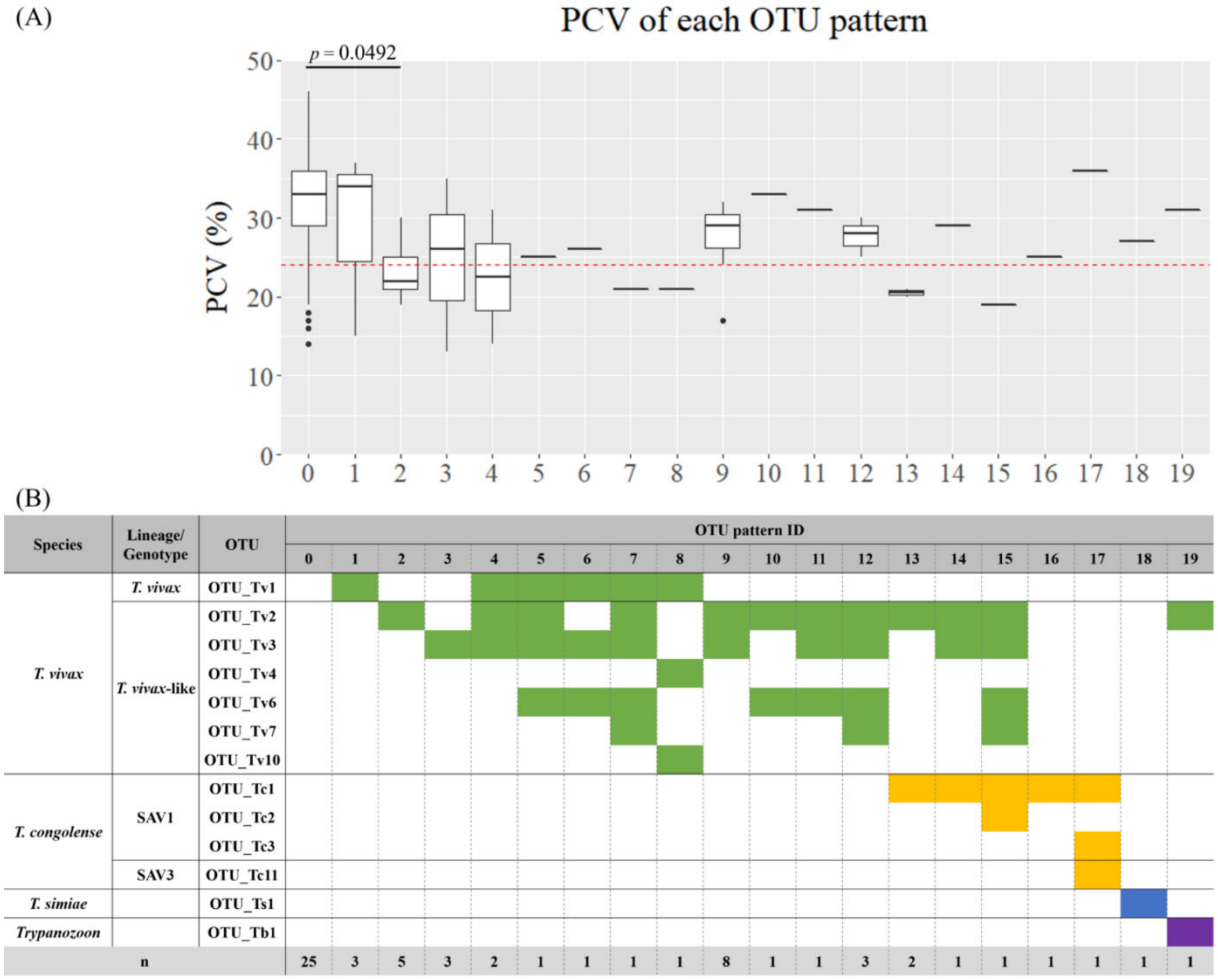


Figure 16. Mean PCV comparisons between OTU pattern groups

Each cattle were assigned to a group according to a pattern of which OTU they were positive for (confirmed by CATL PCR and sequencing). (A) Comparison of the PCV between different OTU pattern groups. (B) The table shows which pattern is positive to which OTU(s). OTU pattern 0 indicates samples which were negative for all OTUs. Statistical significance was assessed by Wilcoxon rank sum test, and only the p values lower than 0.05 are shown. All p values ended non-significant ($p > 0.05$) after correction for multiple comparisons. The dotted red line is drawn at PCV 24% which is commonly used as an indicator for anemia (Mbewe et al., 2015).

2.5 Discussion

T. vivax was the most prevalent *Trypanosoma* spp. found in this study (Table 12, Table 14). Cattle that are male, being reared in Ntubya village, grazing in plain B or in the GMA, or being used spraying rather than dipping as tick control were more likely to be infected with *T. vivax* (Table 11). These results are assumed to be largely influenced by the frequency of the opportunity of cattle encountering tsetse fly bites. For example, among the villages, Ntubya is in the closest range from the GMA, and tsetse flies are present even within the village. In addition, among the four grazing areas, plain B and the GMA are the only areas where tsetse flies are observed. On the other hand, the only factor that had influence in *T. congolense* was the injection of Samorin® (Table 11), where cattle that did not have Samorin® injected within 3 months were more likely to be infected with *T. congolense*. Therefore, to prevent *T. congolense* infection, the use of prophylactic trypanocides seems to be effective. However, the use of Berenil® and Samorin® did not influence the prevalence of *T. vivax* (Fisher's exact test: $p = 0.507$, $p = 0.530$, respectively), and efforts to minimize contact with tsetse flies seems to be important for prevention.

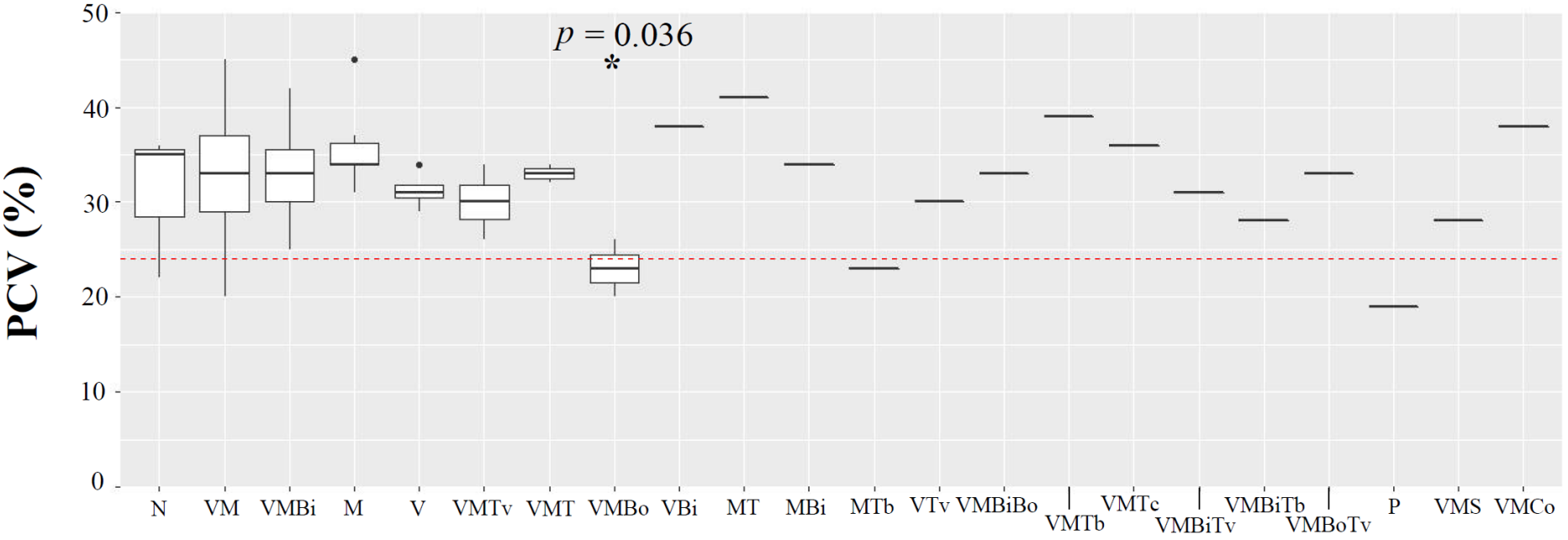
In this study, cattle infected with *T. vivax* had significantly lower mean PCV than non-infected cattle (Figure 15). This indicates that *T. vivax* is the major *Trypanosoma* spp. causing anemia in this area. Moreover, since only seven cattle had a detectable level of parasitemia by thin blood smears (Table 12), it is assumed that *T. vivax* is causing chronic bovine trypanosomiasis with low parasitemia. This assumption agrees with the previous report that *T. vivax* infections are prolonged with long aparasitemic intervals (Wilson et al., 1976). On the other hand, there was a tendency toward a lower mean PCV for cattle infected with *T. congolense* compared to non-infected cattle, but no significant difference was observed (Table 12). In East and southern Africa, *T. congolense* has been reported to be more prevalent and to be more pathogenic than *T. vivax* (Wilson et al., 1976; Simukoko et al., 2007; Mbewe et al., 2015), and there has been more focus on *T. congolense* as the agent of AAT. The results obtained in this study re-emphasize the importance of *T. vivax* as an agent of anemia, along with *T. congolense*. Monitoring *T. vivax* would be especially crucial, since it can be mechanically transmitted by other biting flies and become widespread in regions without tsetse fly infestation where the consciousness against AAT is expected to be lower. Furthermore, bovine trypanosomiasis research and control programs in tsetse infested areas should also be enlightened to include other biting flies.

Notably, a variety of CATL OTU sequences were detected (Figure 11, Figure 12), with pronounced diversity in *T. vivax* (*T. vivax* and *T. vivax*-like). Some of the *T. vivax* CATL clades detected in this study (TviCatL7, TviCatL8, TviCatL9) were consistent with those reported in other East and southern African countries (Rodrigues et al., 2008; Cortez et al., 2009; Nakayima et al., 2013), which were expected considering the geographical location of the study area. TviCatL7 contains isolates from Mozambique nyala and Zambian cattle, which were confirmed to be closely related to West African and South American isolates (Rodrigues et al., 2008; Nakayima et al., 2013). One OTU was detected from

this clade (OTU_Tv1), which was found in both cattle and tsetse flies (Figure 11, Figure 12). TviCatL8 (OTU_Tv2) and TviCatL9 (OTU_Tv3) contains *T. vivax*-like isolates, which to date have only been detected from tsetse fly infested areas in East Africa. OTU_Tv2 and OTU_Tv3 were the most abundant OTUs found in both cattle and tsetse flies, which indicates that these *T. vivax*-like lineages are widespread across East and southern Africa (Figure 11, Figure 12). In addition, cattle with single infection of OTU_Tv2 (*T. vivax*-like lineage, TviCatL8 clade) tended to have lower PCV values than those that were negative for all OTUs (Figure 16). Since cattle with single infection of OTU_Tv1 (*T. vivax* lineage, TviCatL7 clade) had a mean PCV value that was comparable with cattle that were negative for all OTUs, parasites of different lineages and CATL clades may differ in their pathogenesis against cattle.

To elucidate other agents with potential influence against anemia, PCR amplification (RLB PCR) and amplicon sequencing of the 18S rRNA V4 region was conducted to detect *Piroplasma* spp. (Squarre et al., 2020). The test detected five species from the *Theileria* genus (*Theileria velifera*, *T. mutans*, *T. parva*, *T. sp* (buffalo), and *T. taurotragi*), two species from the *Babesia* genus (*Babesia bigemina* and *B. occultans*), and an unknown species from the *Colpodella* genus outside *Piroplasmida*. Only the samples that were included in both ITS1 PCR and RLB PCR were grouped according to the infection pattern of the parasites, and their mean PCV were compared (Figure 17). Statistical significance was not observed in all pairwise comparison after correction of multiple comparisons by Bonferroni correction. However, VMBo (a multiple infection of *T. velifera*, *T. mutans* and *B. occultans*) had significantly lower mean PCV compared to the base-mean of all samples (Kruskal-Wallis test: $p = 0.036$). In addition, P (a single infection of *T. parva*) also had relatively low PCV of 19% which statistical significance was unable to be assessed because of the low sample size. Therefore, along with *T. vivax*, *B.occultans* and *T. parva* are also thought to be important agents for anemia in this community.

The rich OTU diversity within *G. m. centralis* was remarkable, since all OTUs were detected except for OTU_Tv10 and OTU_Tb1 (Figure 11). Similar results were shown in Mozambique, where the highest diversity of gGAPDH and ITS rDNA genotypes were identified in tsetse flies (*G. m. morsitans* and *G. pallidipes*) samples compared to cattle and wild animal samples (Rodrigues et al., 2017). These results indicate the possibility of a variety of different genotypes emerging through recombination inside tsetse flies, of which some genotypes adapted to cattle and have spread across East and West Africa, together with the movement of human and cattle migration (Rodrigues et al., 2017). In addition, different host susceptibility between *Trypanosoma* spp. and subgroups have been reported (Conner and Van den Bossche, 2004), which may be occurring within the different parasite lineages or CATL OTUs. The difference in susceptibility may be caused by the host preference of the parasite or the immune system of the animal host. However, more stringent classification of *T. vivax* taxonomy and population genomic studies on isolates will be needed to further assess these possibilities of recombination within tsetse flies and adaptation of certain genotypes to cattle. On the contrary, the OTU



	N	VM	VMBi	M	V	VMTv	VMT	VMBo	VBi	MT	MBi	MTb	VTv	VMBiBo	VMTb	VMTc	VMBiTv	VMBiTb	VMBoTv	P	VMS	VMCo
T. velifera																						
T. mutans																						
T. taurotragi																						
T. parva																						
T. sp buffalo																						
B. bigemina																						
B. occultans																						
Colpodella sp.																						
T. vivax																						
Trypanozoon																						
T. congolense																						
No. of samples	3	94	19	7	4	4	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 17. Mean PCV comparisons between hemoparasite infection groups

Each cattle were grouped according to which hemoparasite they were positive for (confirmed by RLB PCR and sequencing). (A) Comparison of the mean PCV between different infection groups. Statistical significance for pairwise comparisons were assessed by Wilcoxon rank sum test. All *p* values ended non-significant ($p > 0.05$) after correction for multiple comparisons. The statistical significance against the overall mean was assessed by Kruskal-Wallis test (* $p < 0.05$). The dotted red line is drawn at PCV 24% which is commonly used as an indicator for anemia (Mbewe et al., 2015). (B) The table shows which pattern is positive to which hemoparasite species. N: negative, VM: positive for *T. velifera* and *T. mutans*, VMBi: *T. velifera*, *T. mutans* and *B. bigemina*, M: *T. mutans*, V: positive for *T. velifera*, VMTv: *T. velifera*, *T. mutans* and *T. vivax*, VMT: *T. velifera*, *T. mutans* and *T. taurotragi*, VMBo: *T. velifera*, *T. mutans* and *B. occultans*, VBi: *T. velifera* and *B. bigemina*, MT: *T. mutans* and *T. taurotragi*, MBi: *T. mutans* and *B. bigemina*, MTb: *T. mutans* and *Trypanozoon*, VTv: *T. velifera* and *T. vivax*, VMBiBo: *T. velifera*, *T. mutans*, *B. bigemina* and *B. occultans*, VMTb: *T. velifera*, *T. mutans* and *Trypanozoon*, VMTc: *T. velifera*, *T. mutans* and *T. congolense*, VMBiTb: *T. velifera*, *T. mutans*, *B. bigemina* and *Trypanozoon*, VMBoTv: *T. velifera*, *T. mutans*, *B. occultans* and *T. vivax*, P: *T. parva*, VMS: *T. velifera*, *T. mutans* and *T.* sp buffalo, VMCo: *T. velifera*, *T. mutans* and *Colpodella* sp.

diversity in *G. pallidipes* was low and comparable with what was observed in cattle (Figure 11). Some hypotheses of the reason of the observed discrepancy between the two species are: the difference in bloodmeal host preference, host abundance, distribution, and dispersal rate. Different bloodmeal preferences have been reported, where *G. pallidipes* was more likely to take bloodmeal from bovids such as bushbuck, buffalo, and eland, and *G. m. morsitans* preferred suids (warthog) over bovids (bushbuck) (Leak, 1998a). The distribution of *G. pallidipes* in the Kafue area are predicted to be strictly restricted (Rogers and Robinson, 2004), and the dispersal rates may be small compared to *G. m. centralis* (Leak, 1998g). Therefore, *G. pallidipes* sampled in this study may have had limited access to the variety of wildlife, and preferentially took their bloodmeal from cattle, which resulted in the similar OTU diversity observed between *G. pallidipes* and cattle. However, careful consideration is needed, since the decreased OTU diversity observed in *G. pallidipes* may have been affected by the bias associated with trapping methods. *G. pallidipes* samples were all trapped by epsilon traps, whereas a majority of *G. m. centralis* samples were trapped by the mobile traps. Furthermore, other hematophagous flies, such as Tabanids or *Stomoxys* spp., were not included in this study, which are also likely to be responsible of the mechanical transmission of *T. vivax* between cattle. Clarifying the OTU diversity in these biting flies would increase the understanding of the dynamics of OTUs within the ecosystem.

The overall prevalence of *T. b. rhodesiense* confirmed by SRA PCR was 5.42% (Table 12), with the highest value in New Ngoma (7.84%) and the lowest in Kaminza (1.18%). The prevalence in this study was relatively high compared to other *T. b. rhodesiense* prevalence studies conducted in Zambia (Laohasinnarong et al., 2015). Although there has been no official report of HAT cases in the villages, the Kafue ecosystem is considered as an old HAT focus with a re-emerging risk since the last case in 2016 (Squarre et al., 2016). A study conducted in the same Kafue ecosystem detected *T. b. rhodesiense* in buffalos, a sable antelope, and a vervet monkey using SRA PCR, of which the prevalence were 9.40% (5/53), 12.50% (1/8), and 100% (1/1), respectively (Squarre et al., 2020). The collective findings indicate that *T. b. rhodesiense* is circulating within free-ranging wildlife and human-owned cattle. This could be a concern to the community and requires careful monitoring.

In summary, combining CATL PCR and next generation sequencing proved to be useful in elucidating the *Trypanosoma* spp. diversity, especially for *T. vivax* and *T. vivax*-like trypanosomes. Further studies using isolates with different OTUs, such as in vivo pathogenesis tests, could clarify the epidemiology and relationship of these parasites with disease manifestation in cattle.

General Conclusions

Human African trypanosomiasis (HAT) and African animal trypanosomiasis (AAT) are both protozoan diseases that are transmitted between mammalian hosts via blood sucking activity of tsetse flies (*Glossina* spp.). HAT is caused by *Trypanosoma brucei gambiense* and *T. b. rhodesiense*, whereas AAT is mainly caused by *T. congolense*, *T. vivax*, and *T. b. brucei*. *T. b. rhodesiense*-associated HAT (rHAT) occur in East Africa and has a zoonotic nature, which complicates the disease ecology and makes the complete elimination of rHAT unfeasible. In addition, the absence of effective vaccine is also a drawback to effective rHAT and AAT control. Therefore, controlling the tsetse fly distribution and suppressing the population size are considered to be an important measure for both rHAT and AAT control. Moreover, there is little information about the abundance and genetic diversity of the trypanosome population found in mammalian hosts and tsetse flies from the same ecosystem. By increasing the understanding against the ecology of rHAT and AAT, more efficient decision can be made against control methods such as drug administration against livestock. The aim of this thesis was to elucidate the genetic diversity and population structure of the tsetse fly vector and the trypanosome parasite to suggest effective control strategies against rHAT and AAT.

In Chapter One, partial mitochondrial CO1 sequences and 10 microsatellite loci of *G. m. morsitans* collected from three locations of Zambia and two locations from Malawi were analyzed, and two genetically separated clusters were identified. There appears to be restricted gene flow between Nkhotakota Wildlife Reserve in Malawi and the other four locations, and it was hypothesized that the escarpment of the Great Rift Valley acts as an environmental barrier, since its high altitude is at the limit of the tsetse fly's biological habitat range. In addition to its apparently restricted gene flow, the small effective population size indicates that Nkhotakota Wildlife Reserve may be a population where tsetse control activities can be applied at a lower cost compared to non-isolated populations. However, further research will be needed to identify the genetic population structure in other low-altitude sites around Lake Malawi that have not been included in this study, in order to confirm that there are no re-invasions into Nkhotakota Wildlife Reserve from adjacent areas. Kasungu National Park is also located in Malawi, but probably included in the major tsetse belt in the Luangwa river basin in Zambia. This emphasizes that tsetse control programs should be organized across multiple countries with shared distribution of tsetse flies. Kasungu National Park also had a relatively large effective population size with signals of population size expansion. In the aspect of tsetse and disease control, further study is needed to elucidate the extent of the tsetse belt and assess the degree of migration into adjacent reservoir communities.

In Chapter Two, the genetic diversity of *Trypanosoma* spp. within the Kafue ecosystem was assessed, using comprehensive sequencing methods targeting ITS1 and the CATL region. The outcome has re-emphasized the significance of *T. vivax* as an agent of chronic bovine trypanosomiasis. Furthermore, possible agents were narrowed down into specific genotypes, and their circulation in the

ecosystem was speculated. This study has shown that combining CATL PCR with next generation sequencing is useful in illustrating *Trypanosoma* spp. diversity, especially for *T. vivax* and *T. vivax*-like trypanosomes. Further studies using isolates, such as population genomic studies or *in vivo* pathogenesis tests, could clarify the epidemiology and relationship of these parasites with disease manifestation in cattle. In addition, human-infective *T. b. rhodesiense* was detected in 5.42% of the cattle included in the study, presenting the risk of cattle acting as a reservoir of rHAT in the community. Since the drug efficacy against *T. brucei* parasites are lower compared to *T. congolense* and *T. vivax*, it is assumed that the general drug usage is insufficient to clear *T. b. rhodesiense* from cattle. While also considering the toxic effect of the drugs, strategic clearance of *T. b. rhodesiense*, followed by careful and persistent monitoring is necessary for prevention of rHAT.

Overall, genetic diversity and population structure was assessed for both the tsetse fly vector and the *Trypanosoma* spp. parasite, using a combination of molecular methods. Extending these methods across geographic areas, potential vectors, and mammalian hosts have the potential to expand the knowledge against the disease ecology of both rHAT and AAT.

Acknowledgements

I am greatly indebted to Prof. Junya Yamagishi and Dr. Kyoko Hayashida for their continuous support and guidance throughout my research. I am especially thankful to them for providing me freedom to explore research ideas and carry out fieldwork activities. I would also like to thank Emeritus Prof. Chihiro Sugimoto for his guidance and support to conduct research activities in Zambia. I am grateful for Ms. Naoko Kawai for her assistance in laboratory experiments and her warm company during my student life. I am thankful for the help and inspiration given to me by all my fellow lab-mates in the Division of Collaboration and Education.

I would also like to thank Prof. Kazuhiko Ohashi and Prof. Satoru Konnai (Laboratory of Infectious Diseases), Prof. Yasuhiko Suzuki (Division of Bioresources), and Prof. Ryosuke Omori (Division of Bioinformatics) from the Graduate School of Infectious Diseases, Hokkaido University for invaluable comments and suggestions on this thesis. I would also like to send sincere gratitude toward Dr. Naoki Osada from the Graduate School of Information Science and Technology, Hokkaido University, for his supervision in the analyses in Chapter One. I am grateful to Dr. Martin Simuunza and Prof. Boniface from the University of Zambia for their assistance in ethical submission, research permit, and to the warm support during the stay in Zambia. I am also grateful to Dr. Yongjin Qiu, Dr. Hayato Harima, and Mrs. Ami Soda from Hokudai Center for Zoonosis Control in Zambia (HUCZCZ) for their kind support in procurement and providing laboratory capacity.

I would like to send special thanks to Dr. Victoire Delesalle from Melindika, NGO of International Solidarity, for providing accommodation and support throughout the field sampling in the Itezhi-Tezhi, and also to her inspiring spirit against community veterinary practices. I also thank Mr. Jerry Banda from the Itezhi-Tezhi District Veterinary Office, and Dr. Cornelius Mweempwa from the University of Pretoria, for their reassuring assistance in fieldwork.I would also like to thank Olaf, Mirriam, Boyd, Kennedy, Felix, and Sylvester for organizing visits and interpreting between the farmers, and to all the farmers who kindly cooperated in the study. Finally, I would like to thank my family, friends, and to everyone who I have met during my student life; all interactions were influential and have motivated me to keep on with the work.

The studies in Chapter One and Chapter Two were supported by “Japan Program for Infectious Diseases Research and Infrastructure” from Japan Agency for Medical Research and Development (AMED) under grant number JP20w m0125008. The expenses for travelling and stay for the fieldwork activities in Chapter Two were supported by the Global Leadership Programme in Africa, United Nations University.

References

Adams ER, Hamilton PB, Gibson WC. African trypanosomes: celebrating diversity. Trends Parasitol 26, 324-328, 2010.

Adams ER, Hamilton PB, Rodrigues AC, Malele II, Delespaux V, Teixeira MM, Gibson W. New *Trypanosoma* (*Duttonella*) *vivax* genotypes from tsetse flies in East Africa. Parasitology 137, 641-650, 2010.

Aksoy S, Büscher P, Lehane M, Solano P, Van den Abbeele J. Human African trypanosomiasis control: achievements and challenges. PLoS Negl Trop Dis 11, e0005454, 2017.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 215, 403-410, 1990.

Auty H, Anderson NE, Picozzi K, Lembo T, Mubanga J, Hoare R, Fyumagwa RD, Mable B, Hamill L, Cleaveland S, Welburn SC. Trypanosome diversity in wildlife species from the Serengeti and Luangwa Valley ecosystems. PLoS Negl Trop Dis 6, e1828, 2012.

Baker MD, Krafsur ES. Identification and properties of microsatellite markers in tsetse flies *Glossina morsitans sensu lato* (Diptera: *Glossinidae*). Mol Ecol Notes 1, 234-236, 2001.

Beadell JS, Hyseni C, Abila PP, Azabo R, Enyaru JCK, Ouma JO, Mohammed YO, Okedi LM, Aksoy S, Caccone A. Phylogeography and population structure of *Glossina fuscipes fuscipes* in Uganda: implications for control of tsetse. PLoS Negl Trop Dis 4, e636, 2010.

Bengaly Z, Sidibé I, Ganaba R, Desquesnes M, Boly H, Sawadogo L. Comparative pathogenicity of three genetically distinct types of *Trypanosoma congolense* in cattle: clinical observations and haematological changes. Vet Parasitol 108, 1-9, 2002.

Blum J, Nkunku S, Burri C. Clinical description of encephalopathic syndromes and risk factors for their occurrence and outcome during melarsoprol treatment of human African trypanosomiasis. Trop Med Int Health 6, 390-400, 2001.

Blum J, Schmid C, Burri C. Clinical aspects of 2541 patients with second stage human African trypanosomiasis. Acta Trop 97, 55-64, 2006.

Brun R, Schumacher R, Schmid C, Kunz C, Burri C. The phenomenon of treatment failures in human African trypanosomosis. Trop Med Int Health 6, 906-914, 2001.

Büscher P, Cecchi G, Jamonneau V, Priotto G. Human African trypanosomiasis. Lancet 390, 2397-2409, 2017.

Chisi JE, Muula AS, Ngwira B, Kabuluzi S. A retrospective study of human African trypanosomiasis in three Malawian districts. Tanzan J Health Res 13, 79-86, 2011.

Cochran WG. Estimation of sample size. In: Sampling techniques, 3rd ed. Cochran WG. eds. John Wiley & Sons, Inc., New York. pp. 72-88, 1977.

Connor RJ, Van den Bossche P. African animal trypanosomoses. In: Infectious diseases of livestock, 2nd ed. Vol. 1. 2nd ed. Coetzer JAW, Tustin RC. eds. Oxford University Press Southern Africa, South Africa. pp. 251-296, 2004.

Cornuet JM, Luikart G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001-2014, 1996.

Cortez AP, Rodrigues AC, Garcia HA, Neves L, Batista JS, Bengaly Z, Paiva F, Teixeira MM. Cathepsin L-like genes of *Trypanosoma vivax* from Africa and South America–characterization, relationships and diagnostic implications. Mol Cell Probes 23, 44-51, 2009.

DNDi. Efficacy and safety of fexinidazole in patients with human African trypanosomiasis (HAT) due to *Trypanosoma brucei rhodesiense*. Available from: *https://clinicaltrials.gov/ct2/show/NCT03974178*.

Desquesnes M, Dia ML. Mechanical transmission of *Trypanosoma vivax* in cattle by the African tabanid *Atylotus fuscipes*. Vet Parasitol 119, 9-19, 2004.

Desquesnes M, McLaughlin G, Zoungrana A, Dávila AM. Detection and identification of *Trypanosoma* of African livestock through a single PCR based on internal transcribed spacer 1 of rDNA. Int J Parasitol 31, 610-614, 2001.

Di Rienzo A, Peterson AC, Garzat JC, Valdes AM, Slatkint M, Freimer NB. Mutational processes of simple-sequence repeat loci in human populations. Genetics 91, 3166-3170, 1994.

Dieringer D, Schlotterer C. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes 3, 167-169, 2003.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour 14, 209-214, 2014.

Dyer NA, Lawton SP, Ravel S, Choi KS, Lehane MJ, Robinson AS, Okedi LM, Hall MJ, Solano P, Donnelly MJ. 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. Mol Phylogenet Evol 49, 227-239, 2008.

Earl DA, von Holdt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4, 359–361, 2012.

Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14, 2611-2620, 2005.

Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10, 564-567, 2010.

Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479-491, 1992.

Fermino BR, Paiva F, Soares P, Tavares LE, Viola LB, Ferreira RC, Botero-Arias R, de-Paula CD, Campaner M, Takata CS, Teixeira MM. Field and experimental evidence of a new caiman trypanosome species closely phylogenetically related to fish trypanosomes and transmitted by leeches. Int J Parasitol Parasites Wildl 4, 368-378, 2015.

Fikru R, Hagos A, Roge S, Reyna-Bello A, Gonzatti MI, Merga B, Goddeeris BM, Büscher P. A proline racemase based PCR for identification of *Trypanosoma vivax* in cattle blood. PLoS One 9, e84819, 2014.

Fikru R, Matetovici I, Rogé S, Merga B, Goddeeris BM, Büscher P, Van Reet N. Ribosomal DNA analysis of tsetse and non-tsetse transmitted Ethiopian *Trypanosoma vivax* strains in view of improved molecular diagnosis. Vet Parasitol 220,15-22, 2016.

Ford J. The role of the trypanosomiases in African ecology: a study of the tsetse fly problem. Clarendon Press, Oxford. pp.698, 1971.

Franco JR, Cecchi G, Priotto G, Paone M, Diarra A, Grout L, Simarro PP, Zhao W, Argaw D. Monitoring the elimination of human African trypanosomiasis at continental and country level: update to 2018. PLoS Negl Trop Dis 14, e0008261, 2020.

Gaithuma AK, Yamagishi J, Martinelli A, Hayashida K, Kawai N, Marsela M, Sugimoto C. A single test approach for accurate and sensitive detection and taxonomic characterization of trypanosomes by comprehensive analysis of internal transcribed spacer 1 amplicons. PLoS Negl Trop Dis 13, e0006842, 2019.

Garcia HA, Rodrigues AC, Martinkovic F, Minervino AH, Campaner M, Nunes VL, Paiva F, Hamilton PB, Teixeira MM. Multilocus phylogeographical analysis of *Trypanosoma* (*Megatrypanum*) genotypes from sympatric cattle and water buffalo populations supports evolutionary host constraint and close phylogenetic relationships with genotypes found in other ruminants. Int J Parasitol 41, 1385-1396, 2011.

Gardiner PR. Recent studies of the biology of *Trypanosoma vivax*. Adv Parasitol 28, 229-317, 1989.

Gardiner PR, Assoku RK, Whitelaw DD, Murray M. Haemorrhagic lesions resulting from *Trypanosoma vivax* infection in Ayrshire cattle. Vet Parasitol 31, 187-197, 1989.

Geerts S, Holmes PH, Eisler MC, Diall O. African bovine trypanosomiasis: the problem of drug resistance. Trends Parasitol 17, 25-28, 2001.

Hahn MW, Rausher MD, Cunningham CW. Distinguishing between selection and population expansion in an experimental lineage of bacteriophage T7. Genetics 161, 11-20, 2002.

Hamilton PB, Adams ER, Malele II, Gibson WC. A novel, high-throughput technique for species identification reveals a new species of tsetse-transmitted trypanosome related to the *Trypanosoma brucei*

subgenus, *Trypanozoon*. Infect Genet Evol 8, 26-33, 2008.

Hargrove JW. Tsetse: the limits to population growth. Med Vet Entomol 2, 203-217, 1988.

Hyseni C, Beadell JS, Ocampo-Gomez J, Okedi LM, Gaunt M, Caccone A. The *G. m. morsitans* (Diptera: *Glossinidae*) genome as a source of microsatellite markers for other tsetse fly (*Glossina*) species. Mol Ecol Resour 11, 586-589, 2011.

Hyseni C, Kato AB, Okedi LM, Masembe C, Ouma JO, Aksoy S, Caccone A. The population structure of *Glossina fuscipes fuscipes* in the Lake Victoria basin in Uganda: implications for vector control. Parasit Vectors 5, 222, 2012.

Jackson CHN. The analysis of a tsetse-fly population III. Ann Eugen Cambridge 14, 91-108, 1948.

Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801-1806, 2007.

Jordan AM. Tsetse-flies (*Glossinidae*). Lane RP, Crosskey RW. eds. Chapman & Hall, London. pp. 333-388, 1993.

Kato AB, Hyseni C, Okedi LM, Ouma JO, Aksoy S, Caccone A, Masembe C. Mitochondrial DNA sequence divergence and diversity of *Glossina fuscipes fuscipes* in the Lake Victoria basin of Uganda: implications for control. Parasit Vectors 8, 385, 2015.

Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20, 1160-1166, 2019.

Krafsur ES, Endsley MA. Microsatellite diversities and gene flow in the tsetse fly, *Glossina morsitans s.l.* Med Vet Entomol 16, 292-300, 2002.

Krafsur ES, Maudlin I. Tsetse fly evolution, genetics and the trypanosomiases-a review. Infect Genet Evol 64, 185-206, 2018.

Kristjanson PM, Swallow BM, Rowlands GJ, Kruska RL, De Leeuw PN. Measuring the costs of African animal trypanosomosis, the potential benefits of control and returns to research. Agric Syst 59, 79-98, 1999.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35, 1547-1549, 2018.

Langridge WP. A Tsetse and trypanosomiasis survey of Ethiopia. Ministry of Overseas Development (UK) and Ministry of Agriculture of the Ethiopia. pp. 1-40, 1976.

Laohasinnarong D, Goto Y, Asada M, Nakao R, Hayashida K, Kajino K, Kawazu SI, Sugimoto C, Inoue N, Namangala B. Studies of trypanosomiasis in the Luangwa valley, north-eastern Zambia. Parasit Vectors 8, 555, 2015.

Leak SGA. Behavioural ecology. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 104-146, 1998a.

Leak SGA. Biology. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 17-24, 1998b.

Leak SGA. Classification and anatomy. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 7-16, 1998c.

Leak SGA. Control of trypanosomosis in domestic livestock. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 423-432, 1998d.

Leak SGA. Ecology – distribution and habitats. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 79-103, 1998e.

Leak SGA. Insecticidal spraying. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 345-354, 1998f.

Leak SGA. Population dynamics. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 147-167, 1998g.

Leak SGA. Traps and targets. In: Tsetse biology and ecology: their role in the epidemiology and control of trypanosomosis. Leak SGA. eds. CABI Publishing, Wallingford. pp. 355-370, 1998h.

Leigh JW, Bryant D. POPART: full-feature software for haplotype network construction. Methods Ecol Evol 6, 1110-1116, 2015.

Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments, Nucleic Acids Res 47, 256-259, 2019.

Levine ND, Corliss JO, Cox FE, Deroux G, Grain J, Honigberg BM, Leedale GF, Loeblich AR 3rd, Lom J, Lynn D, Merinfeld EG, Page FC, Poljansky G, Sprague V, Vavra J, Wallace FG. A newly revised classification of the protozoa. J Protozool 27, 37-58, 1980.

Lima AP, Tessier DC, Thomas DY, Scharfstein J, Storer AC, Vernet T. Identification of new cysteine protease gene isoforms in *Trypanosoma cruzi*. Mol Biochem Parasitol 67, 333-338, 1994.

Lindner AK, Lejon V, Chappuis F, Seixas J, Kazumba L, Barrett MP, Mwamba E, Erphas O, Akl EA, Villanueva G, Bergman H. New WHO guidelines for treatment of gambiense human African trypanosomiasis including fexinidazole: substantial changes for clinical practice. Lancet Infect Dis 20, e38-e46, 2020.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89, 238-247, 1998.

Masumu J, Marcotty T, Geysen D, Geerts S, Vercruysse J, Dorny P, Van den Bossche P. Comparison of the virulence of *Trypanosoma congolense* strains isolated from cattle in a trypanosomiasis endemic area of eastern Zambia. Int J Parasitol 36, 497-501, 2006.

Mbewe NJ, Namangala B, Sitali L, Vorster I, Michelo C. Prevalence of pathogenic trypanosomes in anaemic cattle from trypanosomosis challenged areas of Itezhi-tezhi district in central Zambia. Parasit Vectors 8, 638, 2015.

Mekata H, Konnai S, Simuunza M, Chembensofu M, Kano R, Witola WH, Tembo ME, Chitambo H, Inoue N, Onuma M, Ohashi K. Prevalence and source of trypanosome infections in field-captured vector flies (*Glossina pallidipes*) in southeastern Zambia. J Vet Med Sci 70, 923-928, 2008.

Mkanda FX, Mwakifwamba A, Simpamba T. Traditional stewardship and conservation in the game management areas of Nkala and Namwala, Zambia. Oryx 48, 514-521, 2014.

Msangi AR, Saleh KM, Kiwia N, Malele II, Mussa WA, Mramba F, Juma KG, Dyck VA, Vreysen MJ, Parker AG, Feldmann U. Success in Zanzibar: eradication of tsetse. In: Area-wide control of fruit flies and other insect pests. Joint proceedings of the international conference on area-wide control of insect pests, 28 May-2 June, 1998 and the Fifth International Symposium on Fruit Flies of Economic Importance, Penang, Malaysia, Penerbit Universiti Sains Malaysia. pp. 57-66, 1998.

Muhanguzi D, Okello WO, Kabasa JD, Waiswa C, Welburn SC, Shaw APM. Cost analysis of options for management of African animal trypanosomiasis using interventions targeted at cattle in Tororo district; south-eastern Uganda. Parasit Vectors 8, 387, 2015.

Mwima HK. A Brief History of Kafue National Park, Zambia. Koedoe 44, 57-72, 2001.

Nakayima J, Nakao R, Alhassan A, Hayashida K, Namangala B, Mahama C, Afakye K, Sugimoto C. Genetic diversity among *Trypanosoma* (*Duttonella*) *vivax* strains from Zambia and Ghana, based on cathepsin L-like gene. Parasite 20, 24, 2013.

Njiru ZK, Constantine CC, Guya S, Crowther J, Kiragu JM, Thompson RC, Dávila AM. The use of ITS1 rDNA PCR in detecting pathogenic African trypanosomes. Parasitol Res 95, 186-192, 2005.

Okelo WO, MacLeod ET, Muhanguzi D, Shaw A, Welburn SC. Controlling human and animal African trypanosomiasis using insecticide treated cattle: What are the costs and benefits? 2020, PREPRINT (Version 3). Available from: *https://doi.org/10.21203/rs.2.24102/v1*.

Opiro R, Saarman NP, Echodu R, Opiyo EA, Dion K, Halyard A, Aksoy S, Caccone A. Evidence of temporal stability in allelic and mitochondrial haplotype diversity in populations of *Glossina fuscipes fuscipes* (Diptera: *Glossinidae*) in northern Uganda. Parasit Vectors 9, 258, 2016.

Ortiz PA, Da Silva FM, Cortez AP, Lima L, Campaner M, Pral EM, Alfieri SC, Teixeira MM. Genes of cathepsin L-like proteases in *Trypanosoma rangeli* isolates: markers for diagnosis, genotyping and phylogenetic relationships. Acta Trop 112, 249-259, 2009.

Ouma JO, Beadell JS, Hyseni C, Okedi LM, Krafsur ES, Aksoy S, Caccone A. Genetic diversity and

population structure of *Glossina pallidipes* in Uganda and western Kenya. Parasit Vectors 4, 122, 2011.

Ouma JO, Cummings MA, Jones KC, Krafsur ES. Characterization of microsatellite markers in the tsetse fly, *Glossina pallidipes* (Diptera: *Glossinidae*). Mol Ecol Notes 3, 450-453, 2003.

Ouma JO, Marquez JG, Krafsur ES. Microgeographical breeding structure of the tsetse fly, *Glossina pallidipes* in south-western Kenya. Med Vet Entomol 20, 138-149, 2006.

Ouma JO, Marquez JG, Krafsur ES. New polymorphic microsatellites in *Glossina pallidipes* (Diptera: *Glossinidae*) and their cross-amplification in other tsetse fly taxa. Biochem Genet 44, 471-477, 2006.

Ouma JO, Marquez JG, Krafsur ES. Patterns of genetic diversity and differentiation in the tsetse fly *Glossina morsitans morsitans* Westwood populations in East and southern Africa. Genetica 130, 139-151, 2007.

Palmer JM, Jusino MA, Banik MT, Lindner DL. Non-biological synthetic spike-in controls and the AMPtk software pipeline improve mycobiome data. PeerJ 6, e4925, 2018.

Pollock JN. Training manual for tsetse control personnel. Vol.1: Tsetse biology, systematics an d distribution, techniques. FAO, 1982. Available from: *http://www.fao.org/documents/card/en/c/4f831550-7aa2-57f8-8a54-b5714b0abb2b/*

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 155, 945-959, 2000.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. Available from: *http://www.R-project.org/*

Radwanska M, Chamekh M, Vanhamme LU, Claes F, Magez S, Magnus E, De Baetselier P, Büscher P, Pays E. The serum resistance-associated gene as a diagnostic tool for the detection of *Trypanosoma brucei rhodesiense*. Am J Trop Med Hyg 67, 684-690, 2002.

Raymond M, Rousset F. An exact test for population differentiation. Evolution 49, 1280-1283, 1995.

Robinson NP, Burman N, Melville SE, Barry JD. Predominance of duplicative VSG gene conversion in antigenic variation in African trypanosomes. Mol Cell Biol 19, 5839-5846, 1999.

Robinson T, Rogers D, Williams B. Mapping tsetse habitat suitability in the common fly belt of southern Africa using multivariate analysis of climate and remotely sensed vegetation data. Med Vet Entomol 11, 235-245, 1997.

Rodrigues AC, Garcia HA, Ortiz PA, Cortez AP, Martinkovic F, Paiva F, Batista JS, Minervino AH, Campaner M, Pral EM, Alfieri SC. Cysteine proteases of *Trypanosoma* (*Megatrypanum*) *theileri*: cathepsin L-like gene sequences as targets for phylogenetic analysis, genotyping diagnosis. Parasitol Int 59, 318-325, 2010.

Rodrigues AC, Neves L, Garcia HA, Viola LB, Marcili A, Da Silva FM, Sigauque I, Batista JS, Paiva F,

Teixeira MM. Phylogenetic analysis of *Trypanosoma vivax* supports the separation of South American/West African from East African isolates and a new *T. vivax*-like genotype infecting a nyala antelope from Mozambique. Parasitology 135, 1317-1328, 2008.

Rodrigues CM, Garcia HA, Rodrigues AC, Costa-Martins AG, Pereira CL, Pereira DL, Bengaly Z, Neves L, Camargo EP, Hamilton PB, Teixeira MM. New insights from Gorongosa National Park and Niassa National Reserve of Mozambique increasing the genetic diversity of *Trypanosoma vivax* and *Trypanosoma vivax*-like in tsetse flies, wild ungulates and livestock from East Africa. Parasit Vectors 10, 337, 2017.

Rogers DJ, Robinson TP. Tsetse distribution. In: The trypanosomiasis. Maudlin I, Holmes PH, Miles MA. eds. CABI Publishing, Wallingford. pp. 139-179, 2004.

Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4, 137-138, 2004.

Rousset F. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8, 103-106, 2008.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34, 3299-3302, 2017.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4, 406-425, 1987.

Simarro PP, Cecchi G, Franco JR, Paone M, Diarra A, Priotto G, Mattioli RC, Jannin JG. Monitoring the progress towards the elimination of gambiense human African trypanosomiasis. PLoS Negl Trop Dis 9, e0003785, 2015.

Simarro PP, Cecchi G, Franco JR, Paone M, Diarra A, Ruiz-Postigo JA, Fèvre EM, Mattioli RC, Jannin JG. Estimating and mapping the population at risk of sleeping sickness. PLoS Negl Trop Dis 6, e1859, 2012.

Simukoko H, Marcotty T, Phiri I, Geysen D, Vercruysse J, Van den Bossche P. The comparative role of cattle, goats and pigs in the epidemiology of livestock trypanosomiasis on the plateau of eastern Zambia. Vet Parasitol 147, 231-238, 2007.

Solano P, Kaba D, Ravel S, Dyer NA, Sall B, Vreysen MJ, Seck MT, Darbyshir H, Gardes L, Donnelly MJ, De Meeûs T, Bouyer J. Population genetics as a tool to select tsetse control strategies: suppression or eradication of *Glossina palpalis gambiensis* in the Niayes of Senegal. PLoS Negl Trop Dis 4, e692, 2010.

Squarre D, Hayashida K, Gaithuma A, Chambaro H, Kawai N, Moonga L, Namangala B, Sugimoto C, Yamagishi J. Diversity of trypanosomes in wildlife of the Kafue ecosystem, Zambia. Int J Parasitol Parasites Wildl 12, 34-41, 2020.

Squarre D, Kabongo I, Munyeme M, Mumba C, Mwasinga W, Hachaambwa L, Sugimoto C, Namangala B. Human African trypanosomiasis in the Kafue National Park, Zambia. PLoS Negl Trop Dis 10, e0004567, 2016.

Squarre D, Nakamura Y, Hayashida K, Kawai N, Chambaro H, Namangala B, Sugimoto C, Yamagishi J. Investigation of the piroplasm diversity circulating in wildlife and cattle of the greater Kafue ecosystem, Zambia. Parasit Vectors 13, 599, 2020.

Stephen LE. Trypanosomiasis. A veterinary perspective. Pergamon Press, Oxford, 1986.

Stevens JR and Brisse S. Systematics of trypansomes of medical and veterinary importance. In: The trypanosomiasis. Maudlin I, Holmes PH, Miles MA. eds. CABI publishing, Wallingford. pp. 1-23, 2004.

Stich A, Abel PM, Krishna S. Human African trypanosomiasis. BMJ 325, 203-206, 2002.

Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101, 11030-11035, 2018.

Tibayrenc M, Neubauer K, Barnabe C, Guerrini F, Skarecky D, Ayala FJ. Genetic characterization of six parasitic protozoa: parity between random-primer DNA typing and multilocus enzyme electrophoresis. Proc Natl Acad Sci U S A 90, 1335-1339, 1993.

Tikubet G, Gemetchu T. Altitudinal distribution of tsetse in the Finchaa river valley (western part of Ethiopia). Int J Trop Insect Sci 5, 389-395, 1984.

Uilenberg G, Boyt WP. A field guide for the diagnosis, treatment and prevention of African animal trypanosomosis. FAO, 1998. Available from: *http://www.fao.org/3/X0413E/X0413E00.htm*

Vale GA, Bursell E, Hargrove JW. Catching-out the tsetse fly. Parasitology Today 1, 106-110, 1985.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. MICRO - CHECKER : software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4, 535-538, 2004.

Van den Bossche P, Chitanga S, Masumu J, Marcotty T, Delespaux V. Virulence in *Trypanosoma congolense* savannah subgroup. A comparison between strains and transmission cycles. Parasite Immunol 33, 456-460, 2011.

Van den Bossche P, Shumba W, Makhambera P. The distribution and epidemiology of bovine trypanosomosis in Malawi. Vet Parasitol 88, 163-176, 2000.

Welburn SC, Picozzi K, Fèvre EM, Coleman PG, Odiit M, Carrington M, Maudlin I. Identification of human-infective trypanosomes in animal reservoir of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. Lancet 358, 2017-2019, 2001.

Wilson AJ, Paris J, Luckins AG, Dar FK, Gray AR. Observations on a heard of beef cattle maintained in a tsetse area. Trop Anim Health Prod 8, 1-12, 1976.

Wint W, Rogers D. Predicted distributions of tsetse in Africa. FAO Consult Rep, 2000. Available from: *http://www.fao.org/ag/againfo/programmes/documents/livatl2/text/tsepredsum.htm*

Wohlford DL, Krafsur ES, Griffiths NT, Marquez JG, Baker MD. Genetic differentiation of some *Glossina morsitans morsitans* populations. Med Vet Entomol 13, 377-385, 1999.

Woo PT. The haematocrit centrifuge for the detection of trypanosomes in blood. Can J Zoo 47, 921-923, 1969.

World Health Organization. Control and surveillance of human African trypanosomiasis. World Health Organ Tech Rep Ser 984, 1-237, 2013.

World Health Organization. Number of new reported cases of human African trypanosomiasis (*T. b. rhodesiense*). 2020. Available from: *https://www.who.int/data/gho/data/indicators/indicator-details/GHO/lymphatic-filariasis*

Wright S. Variability within and among natural populations. In: Evolution and the genetics of populations. Vol.4, University of Chicago Press, Chicago, 1984.

Japanese abstract (和文要旨)

ヒトアフリカトリパノソーマ症 (HAT: human African trypanosomiasis)および家畜トリパノソーマ症 (AAT: African animal trypanosomiasis)の両者はツェツェバエの吸血によって伝播する原虫病である。前者は *Trypanosoma brucei rhodesiense* および *T. b. gambiense*、後者は *T. congolense*、 *T. vivax*、 *T. b. brucei* などによって引き起こされる。東南部アフリカにおける HAT および AAT は野生動物も保有宿主となること、そしてワクチンが存在しないことから、その疾病の予防対策には媒介昆虫であるツェツェバエの分布域や個体数を制御することが最も有効な手段であると考えられている。さらに、両感染症は遠隔地での発生が多く、それぞれの地域での原虫保有状況や維持されている原虫集団の遺伝的特徴に関して不明な点が多い。そこで本研究では HAT、AAT 両感染症の実効的な対策手段の立案を目的に、下記の研究を行った。

第 1 章ではザンビアおよびマラウイにおける HAT 高度浸淫地域 5 か所を対象に、地域間のツェツェバエの遺伝子流動を特定することを目的に集団遺伝学手法を用いてその遺伝的多様性および集団構造を解析した。集団遺伝学解析にはミトコンドリア CO1 領域およびマイクロサテライト座位 10 か所をマーカーとして用いた。CO1 のハプロタイプネットワーク、およびマイクロサテライトの STRUCTURE 解析の結果から、マラウイの一地域 (Nkhotakota Wildlife Reserve; NWR) がその他の地域とは異なる集団構造であることが示唆された。NWR とその他の地域間の遺伝子流動が限られていた原因としては、大地溝帯に起因する NWR 周辺の地形が推察された。NWR の集団有効サイズも小さい傾向にあったころから、他の集団と比較して効果的なツェツエコントロールを行いやすいことが示唆された。一方で、その他の 4 地域では自由な遺伝子流動が起こっており、大規模なコントロールを実施しても周辺地域からのハエの再流入によって集団の再拡大が起こる可能性が高い。よってこれらの地域では、それぞれの地域間における遺伝子流動の方向性を解析し、集団の拡大を抑制する対策が必要であると考えられる。

第 2 章では野生動物、家畜、ヒトのインターフェイスとなっている Itezhi Tezhi 地域を対象にトリパノソーマ原虫の分子疫学調査を行った。トリパノソーマ原虫を種レベルで網羅的に検出する ITS1 PCR と、種以下の遺伝子型別が可能な CATL PCR と次世代シーケンス解析を組合わせて、同一エコシステム内のツェツェバエとウシが保有するトリパノソーマ原虫の遺伝的多様性を解析した。ITS 1 PCR の結果、ツェツェバエとウシの両方で *T. vivax* が最も多く検出され、*T. vivax* に感染したウシでは非感染牛と比較して顕著に低いヘマトクリット値が観察された。顕微鏡下で原虫が観察された検体数は限られていたことから、本地域では *T. vivax* が低い parasitemia で維持されており、慢性的な AAT を起こしていることが示唆された。さらに、CATL 配列を用いた遺伝子型別により 33 operational taxonomic units (OTUs)が検出された。ウシとツェツェバエで最も多く検出された OTU は共通していた一方で、ウシと比較してツェツェバエの方が全体の OTU の多様性が高いことがわかった。よって本地域ではツェツェバエで OTU の多様性の高い原虫集団が維持されており、ウシではその内の限られた原虫集団が維持されていることが示唆された。また、ウシの 5.42%がヒト感染性 *T. b. rhodesiense* 特異的 SRA PCR 陽性であったことから、ウシが HAT の保有宿主として重要であり、今後のモニタリングが重要であると考えられる。

以上、第 1 章、第 2 章により、トリパノソーマ原虫およびツェツェバエの集団構造を遺伝的多様性から推測した。今後はそれぞれの解析対象を拡大、および両者を統合することでより実効的な HAT/AAT 対策の立案が可能であると考える。