
**UNIVERSITY OF NAIROBI**

# INTRASPECIFIC GENETIC DIVERSITY AMONG TRYPANOSOMES FROM KWALE AND MARSABIT COUNTIES IN KENYA

**NAOMI NYAWIRA KIMENYI (BSc.)**

**156/12241/2018**

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Award of the Degree of Master of Science in Genetics in the School of Biological Sciences of The University of Nairobi.**

**6/7/2021**

## DECLARATION

I declare that this thesis is my original work and has not been submitted elsewhere for examination, award of a degree or publication. Where other people's work, or my own work has been used, this has properly been acknowledged and referenced in accordance with the University of Nairobi requirements

Signature _____ Date 8/6/2022

**Naomi Nyawira Kimenyi, B.S.C**

I56/12241/2018

Department of Biology

University of Nairobi

This thesis is submitted for examination with our approval as research supervisors:

Signature _____ Date 8/06/2022

**Dr. Nelson Amugune**

Department of Biology

P. O. Box 30197- 00100

Nairobi, Kenya

amugunenelson@gmail.com

Signature _____ Date 8/6/2022

**Dr. Merid Getahun**

Animal Health, International Center for Inset Physiology and Ecology (*icipe*)

P. O. Box 30772-00100

Nairobi, Kenya

mgetahun@icipe.org

**DEDICATION**

I dedicate this thesis to my parents Mr. Patrick Kimenyi and Mrs. Jane Kimenyi and my siblings Kelvin Muteru and Esther Wambui without whom I would neither have started nor persevered through my masters.

**ACKNOWLEDGEMENT**

I wish to thank the following who contributed immensely in the birth of this thesis; my supervisors Dr. Nelson Amugune and Dr. Merid Getahun whose knowledge, guidance and patience empowered me to begin and complete the work. I am also grateful to retired Prof. Titus Mukiama who supervised formulation of the study design and writing of the proposal. I am grateful to the International Center for Insect Physiology and Ecology (*icipe*) where I undertook the research and the Department of Biology at the University of Nairobi where I was registered for my master's degree programme. Further, I am very grateful to Kelvin Muteru for his support with data analysis.

I am also thankful to my colleagues and friends at the Molecular Biology and Bioinformatics Unit (MBBU) at *icipe,* Mr James Kabii of MBBU, *icipe* and others within *icipe* particularly Allan Okwaro who supported me and made my research memorable.

Finally, I'm grateful to my late grandfather Mr. Charles Muteru for his financial support. I salute you!

Last but not least, glory is to God!

## TABLE OF CONTENTS

**LIST OF TABLES**

**LIST OF FIGURES**

## LIST OF APPENDICES

## LIST OF ABBREVIATIONS

AMOVA- Analysis of Molecular Variance

Ar –Allelic Richness

CK – Coastal Kenya

$F_{IS}$ – Fisher's Inbreeding Coefficient

$F_{ST}$ – Wright's Fixation Index

$H_E$ – Expected Heterozygosity

$H_0$ – Observed Heterozygosity

HWE – Hardy Weinberg Equilibrium

I- Shannon's Information Index

IBD- Isolation by Distance

IBE- Isolation by Environment

*Icipe*- International Center for Insect Physiology and Ecology

ITS-1 – Internal Transcribe Spacer Unit 1

NK – Northern Kenya

PATTEC- Pan-African Tsetse and Trypanosome Eradication Campaign

PCA- Principle Component Analysis

PCR- Polymerase Chain Reaction

Q Value- STRUCTURE software cluster assignment value

SSR- Simple Sequence Repeats

TAE- Tris base, Acetic acid and EDTA buffer

UPGMA- Unweighted Pair Group Method with Arithmetic Mean

VSG- Variable Surface Glycoprotein

**ABSTRACT**

African Animal Trypanosomiasis caused by trypanosomes is a great challenge to livestock keeping. The parasites are transmitted biologically by tsetse flies *(Glossina* spp.) and mechanically by stable flies like *Stomoxys* spp., horse fly (*Tabanus* spp.) and camel fly (*Hippobosca camelina*). While *Trypanosoma brucei* and *T. congolense* require cyclical transmission by tsetse flies for maintenance over long periods in a region, their occurrence in parts of tsetse free Northern Kenya, including Marsabit County has been reported. That this could be due to the presence of unknown biological vectors and reservoirs or continuous introduction of trypanosomes due to continuous movement of livestock between tsetse free and infested areas. The aim of the study was to carry out microsatellite genotyping to explore intra-specific genetic diversity between trypanosomes from Shimba Hill, Kwale County tsetse fly belt, and those from Laisamis, Shurr and Ngurunit, tsetse free sites from Marsabit County.

Trypanosome DNA was obtained from camel blood, cattle blood and tsetse flies. Molecular detection of trypanosomes was done through the ITS1 PCR diagnostic test. Microsatellite genotyping was done through PCR amplification of seventeen microsatellite loci using fluorescently labeled forward primers. Microsatellite genotyping showed geographical origin based structuring among *Trypanozoon* isolates.

There was clear separation between isolates from the two regions signaling the potential of microsatellite markers as diagnostic markers for *T. brucei* and *T. evansi* isolates. *Trypanosoma vivax* isolates also clustered largely based on sampling location with moderate differentiation (P > 0.5) between the two locations. In addition, the results revealed significant differentiation between isolates from Marsabit and Kwale regions thus revealing that genetic heterogeneity is linked to biological transmission. The results also revealed that *T. congolense* isolates from

Marsabit County are not genetically separated from those from Kwale County. Therefore, these isolates are likely introduced in the region through animal movement.

These results have shed new light on the population structure and genetic range of trypanosome species from two distinct ecological settings. There is however need to increase the sample size and area coverage in order to explore the genetic diversity and differentiation between *T. congolense* isolated from tsetse free areas and those isolated from tsetse endemic area.

# CHAPTER ONE: INTRODUCTION

## 1.1 Background information

Trypanosomes are protozoan parasites belonging to the genus *Trypanosoma*. They are the causative agents of trypanosomiasis which are a group of veterinary and human diseases. In Africa, *Trypanosoma brucei gambiense* (West Africa) and *T. b. rhodesiense* (East Africa) cause human trypanosomiasis (sleeping sickness) (Echodu *et al.*, 2015). Animal trypanosomiasis is caused by *T. b. brucei*, *T. evansi*, *T. vivax*, *T. congolense* and *T. equiperdum* (Morrison *et al.*, 2016). Trypanosomes (other than *T. evansi* and *T. equiperdum*) are transmitted cyclically (biologically) by tsetse flies (*Glossina* spp.) and/or non-cyclically (mechanically) by horseflies (*Tabanus* spp.), camel flies (*Hippobosca camelina),* stable flies (*Stomoxys* spp.), tsetse flies (*Glossina* spp.) and vampire bats (*Desmodus rotundus*) (for *T. evansi*) (Mihok *et al.*, 1995; Desquesnes and Dia, 2003; Desquesnes *et al.*, 2013; Kamidi *et al.*, 2017; Getahun *et al.*, 2020). Vampire bats in South America are both vectors and reservoirs of *T. evansi* (Brun *et al.*, 1998). Cyclical trypanosomes undergo development in the vector midgut and mouthparts while non-cyclical trypanosomes remain in developmental forms that are similar to bloodstream forms of *T. brucei* and do not develop further in the vector (Lai *et al.*, 2008). Tsetse flies transmit *T. congolense, T. vivax, T. brucei* biologically and mechanically while other biting flies transmit them only mechanically (Desquesnes and Dia, 2003; Desquesnes *et al*., 2013;Duffy *et al*., 2009 Osório *et al*., 2008; Takeet *et al*., 2017).

*Trypanosoma brucei* and *T. congolense* have been isolated in tsetse free areas of Northern Kenya from as far back as the eighties to as soon as 2017 (Gibson and Wilson, 1983; Njiru *et al.*, 2006; Getahun *et al.*, 2020). Though the species are transmitted mechanically by non-tsetse biting flies,

they survive for short periods of time in the vector mouthparts before inoculation into the next host and they do not undergo any development (Desquesnes and Dia, 2003). It is because of this reason that the parasites are not established outside Africa unlike *T. evansi* and *T. vivax* (Jones and Alberto, 2001). Occurrence of *T. b. brucei* and *T. congolense* may be due to continuous introduction from neighboring tsetse infested areas through animal movement or the presence of tsetse flies in these areas but at very low hard to detect densities (Wells, 1972). However, isolation of the parasites from parts of Northern Kenya far away from tsetse fly belts is evidence that the species are actively maintained in these areas (Gibson and Wilson, 1983; Getahun *et al.*, 2020). An explanation for this may be the possibility of biological transmission by an unidentified non-tsetse vector (Gibson and Wilson, 1983; Getahun *et al.*, 2020). Maintenance of these species in tsetse free areas by unidentified mechanism raises concerns on the possibility of their (plus human infective trypanosomes) spread beyond tsetse fly belts of Africa and beyond African borders.

Multilocus genotype analysis of *T. vivax* strains has revealed that South American *T. vivax* strains are genetically homogeneous and are differentiated from African strains (Garcia *et al.*, 2014). Contrarily, African *T. vivax* strains are genetically heterogeneous and occasionally, new genotypes are discovered in East Africa (Rodrigues *et al.*, 2008, Rodrigues *et al*., 2017; Adams *et al.*, 2010). Both biological and mechanical vectors of *T. vivax* co-exist in Africa and thus the parasites are transmitted both biologically and mechanically. However, tsetse flies are absent in South America and therefore the *T. vivax* strains found there are only transmitted mechanically (Rodrigues *et al.*, 2008). Biological transmission of *T. vivax* strains in Africa may allow genetic exchange and consequently, genetic heterogeneity among the strains may be attributed to mode of transmission (Fikru *et al.*, 2016). The correlation between genetic heterogeneity and biological

transmission can be investigated by comparing *T. vivax* isolates from tsetse free areas and those from tsetse infested areas (Rodrigues *et al.*, 2008). However, although African *T. vivax* strains have been shown to have a clonal population structure (Duffy *et al.*, 2009; Garcia *et al.*, 2014), the presence of eight genes associated with meiosis in the parasite may be evidence of the potential of the species to reproduce sexually (Duffy *et al.*, 2009).

Thus, we compared Northern Kenyan (Marsabit County) and Coastal Kenyan (Kwale County) *Trypanozoon* isolates (while ignoring established nomenclature and treating the group as one), *T. vivax* and *T. congolense* isolates from tsetse free and tsetse endemic areas of Kenya, to explore their intraspecific population structure, genetic differentiation, diversity, and gene flow rates. This is the first intra-species comparative population genetics study on trypanosomes from tsetse free and tsetse endemic areas of Africa. Ribosomal DNA sequence analysis may fail to inform on interspecies genetic separation among a group of trypanosome isolates (Fikru *et al.*, 2014; Getahun *et al.*, 2020). Microsatellite analysis, which is a powerful tool for population genetics studies (Senan *et al.*, 2014; Fikru *et al.*, 2016) has the potential for being used to study genetic separation of trypanosome spp. Therefore, we carried out microsatellite genotyping on Trypanozoon, T. vivax and T. congolense isolates from tsetse free and tsetse endemic areas of Kenya, and explored their intraspecific population structure, genetic differentiation, diversity and gene flow rates.

**1.2 Problem statement**

Our specific knowledge on the genetic diversity of the clinically important trypanosomes (*T. congolense, T. vivax* and *T. b. brucei*) from outside tsetse fly belts is limited compared to the data available on the same species from tsetse infested areas (African Union, 2018). In addition, it was noted with a lot of concern that PATTEC (Pan-African Tsetse and Trypanosome Eradication

Campaign) countries reports did not have adequate detailed information on non-tsetse transmitted trypanosomiasis. It was therefore suggested that adequate prominence should be given to this very important area (African Union, 2018). Therefore, the objective of this research project was to address this knowledge gap by assessing the population structure, genetic diversity and differentiation of trypanosomes from tsetse free and tsetse infested areas.

Genetic heterogeneity among African *T. vivax* strains hinders unambiguous identification of *T. vivax* isolates because a majority of PCR diagnostic tests are based on West African *T. vivax* DNA sequences and a number of them fail to identify some East African genotypes (Adams *et al.,* 2010; Fikru *et al.,* 2014). There is therefore need to understand the level of heterogeneity among African *T. vivax* strains. One way to achieve this is to carry out a comparative population genetic study between *T. vivax* strains from tsetse free and tsetse infested areas of Africa to find out whether the heterogeneity could be stemming from biological transmission. (Rodrigues *et al.*, 2008; Fikru *et al.,* 2014).

**1.3 Justification**

Getahun *et al.,* 2020 isolated presumed tsetse transmitted species; *T. congolense* and *T. b. brucei* in other biting flies such as *Stomoxys calcitrans* and in camels reared outside tsetse fly belts of Kenya, specifically in Marsabit County. These trypanosomes showed remarkable complexity and diversity. In addition, the parasites' degree of virulence in camels varied showing that camel trypanosomiasis is caused by multiple trypanosomes unlike previously thought. Intraspecific genetic variation between *T. congolense* and *T. b. brucei*, within and without tsetse fly belts supports the hypothesis that occurrence of these species in tsetse free areas may be due to the presence of other biological vectors (other than tsetse flies). However, distinguishing between taxa within the *Trypanozoon* subgenus with molecular and morphological methods is hard as

genetic data analysis of *T. evansi* and *T. brucei* has shown that some *T. evansi* isolates are genetically closer to *T. brucei* isolates than other *T. evansi* isolates (Carnes *et al.*, 2015; Kamidi *et al.*, 2017). In addition, commonly used rDNA molecular markers for *Trypanozoon* characterization have not conclusively distinguished between the two species even with genetic sequencing (Büscher *et al.*, 2019; Getahun *et al.*, 2020). This has rendered the taxonomic ranking within the *Trypanozoon* subgenus problematic with a need for review.

The aim of this study was to find out how *T. b. brucei* and *T. congolense* are maintained and are able circulate among domestic animals and biting flies in the absence of their biological vectors (tsetse flies) and also to find out whether *T. vivax* isolates from tsetse free areas of Kenya are genetically separated from *T. vivax* from tsetse endemic areas. Isolates rom Marsabit County were used in this study because that is where the initial isolate used in the Getahun *et al*., 2020 study were sampled. These were compared with isolates from Kwale County because the region is tsetse infested and close to Marsabit County compared to counties in Western Kenya. The findings of my study will help contribute towards the management of trypanosomiasis since intraspecific genetic diversity has vital implications on virulence and the design of preventive strategies.

### 1.4 Research Questions

- Is the occurrence of *T. brucei* and *T. congolense* outside Kenyan tsetse fly belts due to the presence of other biological vectors and reservoirs?
- Are *T. vivax* isolates from tsetse free regions in Kenya genetically separated from *T. vivax* isolates from tsetse endemic areas?

## 1.5 Objectives

### 1.5.1 General objective

To determine intraspecific genetic diversity within *T. brucei. brucei, T. evansi, T. congolense* and *T. vivax* isolates detected in tsetse flies and livestock from a tsetse endemic area in Kwale County and tsetse free areas in Marsabit County of Kenya.

### 1.5.2 Specific Objectives

i. To identify trypanosomes in camels, cattle and tsetse flies from Shurr, Laisamis and Ngurunit of Marsabit County (tsetse free areas) and Shimba Hills, Kwale County (tsetse endemic area) through ITS-1 (Internal Transcribed Spacer – 1) based PCR diagnostic tests.

ii. To investigate the genetic diversity of the identified trypanosomes using microsatellite analysis.

# CHAPTER TWO: LITERATURE REVIEW

## 2.1 Trypanosome biology

Trypanosomes belong to the domain eukaryote (Hampl *et al.*, 2009; Burki *et al.*, 2020). This domain is divided into the super groups Excavata, Amoebozoa, Chromalveolata, Archaeplastida, Opisthokonta and Rhizaria (Burki *et al.*, 2020). The super group Excavata is further divided into the clades Discicristata, Preaxostyla and Fornicata. The clade Discicristata is further divided into the phyla Euglenozoa and Percolozoa (Hampl *et al.*, 2009). Phylum Euglenozoa is made up of three orders; Kinetoplastida, Diplonemida and Euglenida. The suborders Trypanosomatina and Bodomina make up the order Kinetoplastida which are characterized by the presence of a mass of mitochondrial kDNA (kinetoplast) (Hampl *et al.*, 2009; Burki *et al.*, 2020). The suborder Trypanosomastina forms one family known as Trypanosomastidae. Members of this family are uniflagellate with a small kinetoplast. This family is made up of the genera *Leishmania* and *Trypanosoma* among others (Maslov *et al*., 2001).

Members of the genus *Trypanosoma* are transmitted by biting flies (among other modes of transmission) and are divided into two clades depending on the site of development in the digestive tract of the insect. These clades are Stercoraria (the parasites develop in the posterior part) and Salivaria (the parasites develop in the anterior part) (Desquesnes *et al.*, 2013). The Salivaria group is made up of the African pathogenic trypanosomes which are further divided into four sub-generas: Nannomonas, Duttonella, Trypanozoon and Pycnomonas. The subgenera Nannomonas is made up of the animal infective *Trypanosoma congolense*, *Trypanosoma simiae* and *Trypanosoma godfreyi*. Duttonella includes *T. vivax* and *T. uniforme* (Desquesnes *et al.*, 2013; Radwanska *et al.*, 2018). The sub-genera Trypanozoon is made up of *Trypanosoma brucei*

*brucei, T. b. gambiense, T. b. rhodesiense, T. evansi* and *T. equiperdum*. *Trypanosoma b. brucei, T. b. gambiense* and *T. b. rhodesiense* are found in Sub-Saharan Africa while *T. evansi* and *T. equiperdum* are found in and outside Africa (Desquesnes *et al.*, 2013).

*Trypanosoma* cells are elongated with single copy organelles (nucleus, mitochondria, flagella pocket and kinetoplast) located strategically within them (Matthews, 2005) (Figure 1). The flagella pocket, from which a flagellum exits the cell, is located at the posterior end. The trypanosome cells depends on the flagellum for motility in the bloodstream of the mammalian host (Bargul *et al.*, 2016). The mitochondrion of a trypanosome is made up of catenated (linked DNA strands) DNA known as kinetoplast. In mammalian bloodstream, the mitochondrion is simple, tubular in shape and lacks cristae. This is because in this stage, the mitochondrion is not respiring due to the presence of glucose in the bloodstream which is broken down in glycolytic reactions that happen in the glycosomes (Parsons, 2004). The kinetoplast comprises of two different types of circular DNA known as maxicircles and minicircles. Maxicircles have mitochondrial proteins coding genes and minicircles have genes coding for short guide RNAs through which maxicircle RNAs are post-transcriptionally edited by addition or removal of uridines (Aphasizhev and Aphasizheva, 2011).

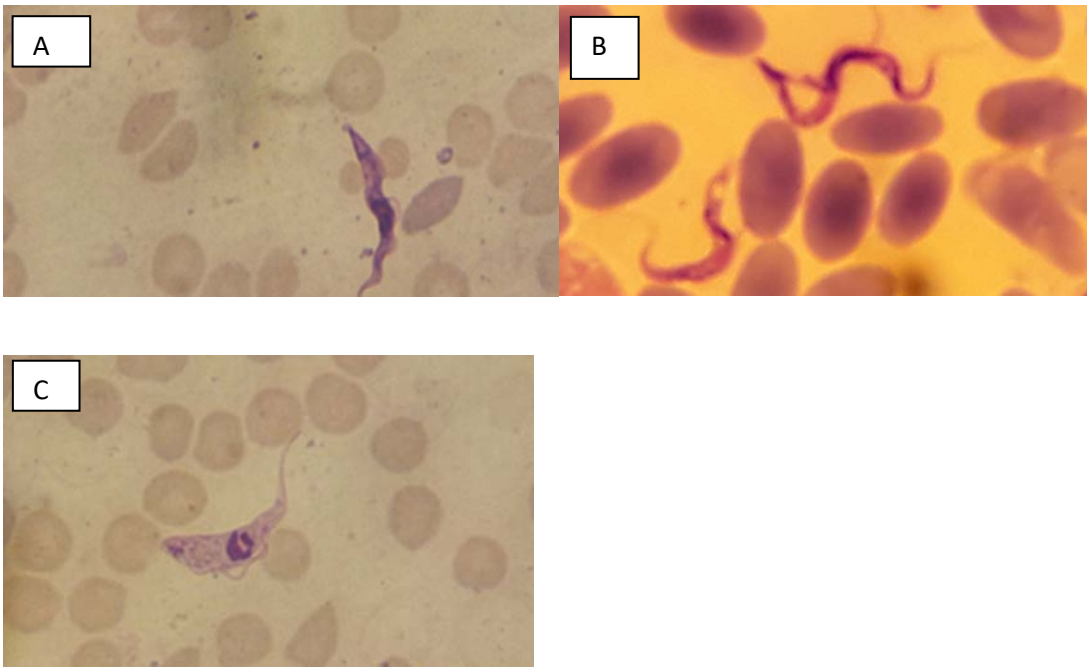



**Figure 1**: Microscope images showing the cells of (A) *Trypanosoma brucei*, (B) *T. vivax* and (C) *T. congolense* (C) in mammalian blood in the midst of blood cells (source: *icipe* Molecular Biology and Bioinformatics Unit (MBBU) laboratory.

### 2.2 Trypanosome Life cycle and Transmission

While *T. equiperdum* is transmitted sexually, other Salivarian trypanosomes are predominantly transmitted either cyclically and/ or non-cyclically by invertebrate vectors (Carnes *et al.*, 2015). Tsetse flies transmit Salivarian trypanosomes other than *T. evansi* cyclically while other non-tsetes biting flies transmit them non-cyclically. As outlined in Figure. 1, when feeding on the mammalian host, the tsetse deposit metacyclic trypanosomes on the host's dermal tissues from where they enter the bloodstream (Radwanska *et al.*, 2018). Here, the parasites multiply rapidly as morphologically slender forms before differentiating into non-replicative morphologically stumpy forms. This process involves activation of mitochondrial DNA with formation of cristae, proline and α-ketogluterate oxidases which are utilized in the tsetse midgut for proline

metabolism (Matthews, 2005). Due to antigenic variation, the parasitemia fluctuates in an effort to survive the host's immune responses (Radwanska *et al.*, 2018). Ascending parasitemia are made up of proliferative slender parasites and are replaced by non-proliferative stumpy parasites during decline which continue their developmental cycle in the vector (Radwanska *et al.*, 2018).Tsetse flies take up blood stream trypomastigotes of *T. brucei* and *T. congolense* into the midgut where they transform into procyclic trypomastigotes. These parasites subsequently develop into epimastigotes then migrate into salivary glands and proboscis in that order. Here, they evolve into metacyclic trypomastigotes (Mathews, 2005; Osório *et al.*, 2008; Morrison *et al.*, 2009; Radwanska *et al.*, 2018).

*Trypanosoma vivax* unlike *T. brucei* and *T. congolense* does not have a procyclic stage within the tsetse midgut and does not migrate within the insect. In contrast, the trypomastigotes develop directly into epimastigotes within the vector proboscis. For this reason, in addition to cyclical transmission, the parasite is successfully transmitted non-cyclically (Jackson *et al.*, 2015).

Non-cyclical transmission of trypanosomes is carried out by biting flies of the genera *Stomoxys, Tabanus* and *Hippobosca. Trypanosoma evansi* exists as developmental forms of *T. brucei* with no further development in the vector (Kamidi *et al.*, 2017). Due to diskinetoplastidy, *T. evansi* cannot utilize proline in the tsetse midgut as an energy source and are thus incapable of cyclical development in the vector and consequently, cyclical transmission (Schnaufer, 2010).

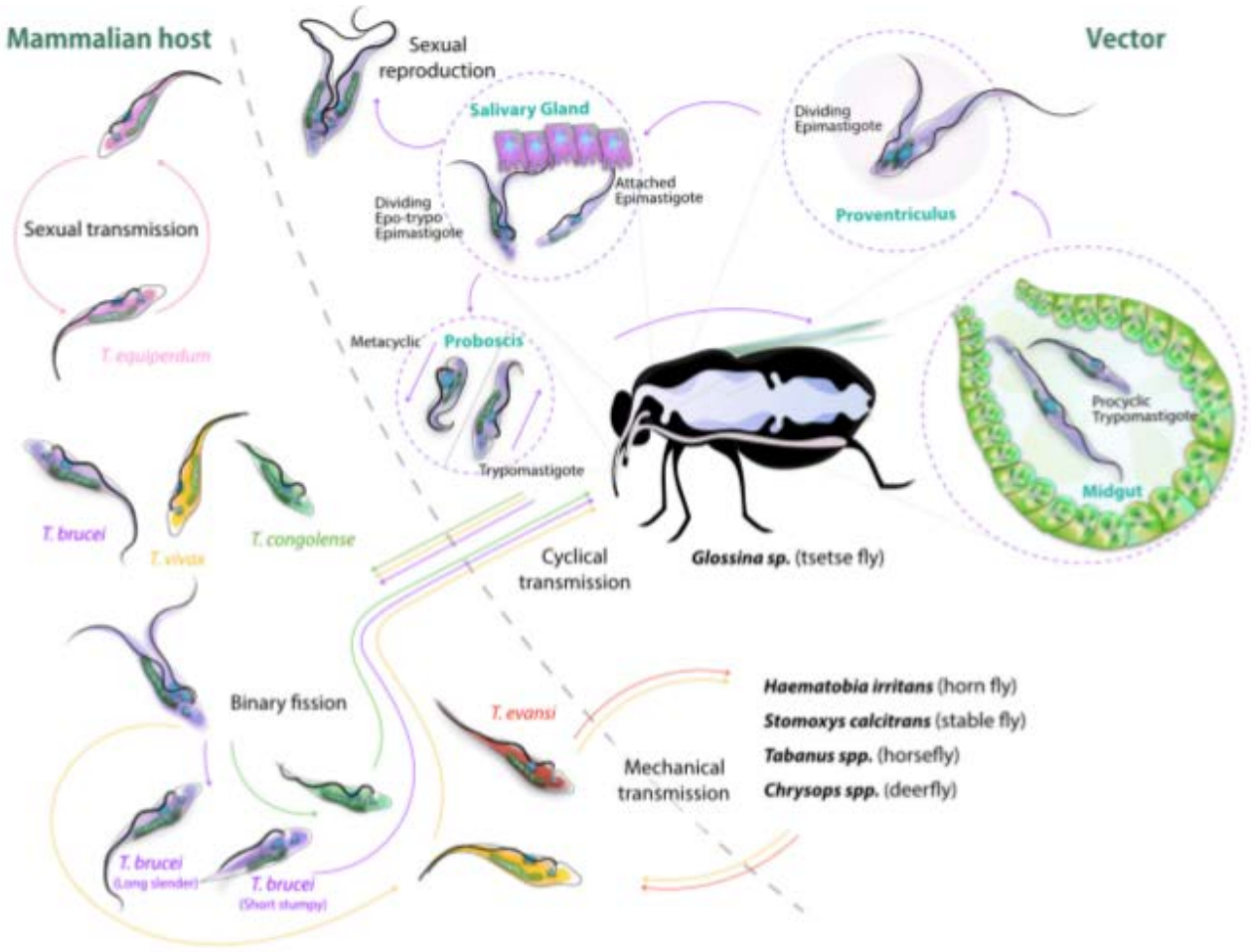



**Figure 2**: Cyclic and mechanical transmission of Salivarian trypanosomes (other than *T. vivax*) by tsetse flies and other biting flies (Source: Radwanska et al., 2018).

### 2.3 Economic Importance

Livestock keepers in sub-Saharan Africa face a great set back due to African Animal trypanosomiasis which costs them up to US$ 4.5 billion annually. It's made up of a number of veterinary diseases with nagana and surra being the most common (Shaw *et al.*, 2014). Nagana in cattle, sheep, goats and wild life results mainly from *T. congolense*, *T. vivax*, and *T. brucei* (Auty *et al*., 2015). The disease lowers the yield of over 260 million goats and sheep and 150 million cattle within Sub-Saharan Africa (Shaw *et al.*, 2014). The parasites reach the brains of the

animals through the lymphatic system and blood stream and cause a number of symptoms including: emaciation, severe anemia, feticide, infertility and is fatal if not treated (Yaro *et al.*, 2016).

*Trypanosoma evansi* causes surra that mainly affects camels and horses and occasionally cattle and buffalos among other animals (Desquesnes *et al.*, 2013). The disease causes death to thousands of animals annually in Africa and beyond in addition to the economic burden on livestock breeders due to loss in investment in chemotherapeutic interventions (Kamidi *et al.*, 2017). Surra is fatal if left untreated and its symptoms include loss of appetite and productivity, fever, anemia and abortion (Desquesnes *et al.*, 2013; Getahun *et al.*, 2020).

**2.4 Molecular methods for trypanosome detection**

The Internal Transcribed Spacers-1 (ITS-1) region of ribosomal DNA (rDNA) is the most commonly targeted regions for identification of several trypanosome species at a go (Desquesnes *et al*., 2001). This is so because the ITS-1 flanking regions are highly conserved in addition to their size variability among trypanosome species and sub-groups (Musaya *et al*., 2017). The rDNA locus has 100 to 200 copies where each transcribed unit is made up of 18S, 5.8S and 28S rRNA genes separated by two ITS regions (ITS1 and ITS2). A pair of PCR primers designed from the conserved regions of rDNA genes are used to amplify the ITS1 region. The primer sequences targeting ITS-1 CF (5'CCGGAAGTTCACCGATATTG and ITS1 BR (5'TTGCTGCGTTCTTCAACGAA) (Desquesnes *et al*., 2001; Njiru *et al*., 2005). The ITS CF anneals to 18S while the ITS BR anneals to the 5.8S regions of rDNA thus leading to the amplification of ITS-1 (Botelho *et al.*, 2005). The ITS1 region varies is size within the trypanosome species and is thus used to tell trypanosome species apart based on their sizes (Desquesnes *et al*., 2001; Njiru *et al*., 2005). The length of ITS1 sequences for *Trypanosoma*

(Trypanozoon) *brucei* is 480bp, 700bp for *T. congolense* Savannah, 710bp for *T. congolense* Forest, 620bp for *T. congolense* Kilifi and 250bp for *T. vivax*. (Njiru *et al*., 2005). The technique allows species determination in mixed infections. However, this is impossible in cases where the length of the amplicon is equal for two species or where there exists intra-species variation. Also, it is impossible to tell apart species within the Trypanozoon group using ITS typing. Sequencing the region does not always distinguish *T. brucei brucei*, *T. evansi* and *T. equiperdum* either (Wen *et al.*, 2016; Büscher *et al.*, 2019; Getahun *et al.*, 2020). In addition, the ITS region has a relatively low copy number (100-200bp), which limits its sensitivity during tests (Hutchinson and Stevens, 2018). Additional PCR assays have been developed for the detection of *T. evansi* isolates. One of these assays targets a 488bp fragment of the Variable Surface Glycoprotein (VSG) antigen type RoTat 1.2 variant. This gene occurs in most *T. evansi* strains but is absent in some (Kamidi *et al.*, 2017).

### 2.5 Characterization of trypanosomes based on microsatellite DNA

Microsatellites, also known as Simple Sequence Repeats (SSR) or Short Tandem Repeats (STR) are tandem nucleotide repeats that are found all-over eukaryotic genomes (Andrews, 2013). They are made up of repeated motifs of 1-6 nucleotides (Guichoux *et al.*, 2011). Microsatellite data is used commonly in genetic mapping, parentage analysis, fingerprinting, genetic structure analysis and genetic diversity studies (Sheriff and Alemayehu, 2018). Their high allelic diversity, high level of repeat-number polymorphism, co-dominant inheritance, high distinction between related individuals, high rate of mutations, and abundance are some of the qualities that make them suitable for use in population genetic studies (Abdul-Muneer, 2014; Sheriff and Alemayehu, 2018). To genotype microsatellite loci, fluorescently labeled forward primers plus non-labeled reverse primers are used to amplify the microsatellite region. This produces two fluorescently

labeled allelic products with differing lengths depending on the microsatellite repeating units. To analyze the allelic lengths, the PCR products are separated by capillary electrophoresis(Andrews, 2013).

Microsatellites have been used previously to study *Trypanozoon*, *T. vivax* and *T. congolense* populations in Africa. Kamidi *et al* 2017 studied the genetic diversity between *T. evansi* and *T. brucei* isolates from Northern Kenya using fifteen microsatellite loci where it was revealed that *T. evansi* isolates have multiple evolutionary origins from *T. brucei.* Population genetics studies on *T. vivax* using microsatellite typing have also revealed that the species is clonal and differentiated into genetically distinct groups which may be totally different species(Duffy *et al.*, 2009; Garcia *et al.*, 2014). Microsatellite typing in *T. congolense* has revealed that the species reproduces both clonally and sexually and that there is often low sub-structuring among the three different *T. congolense* sub-species (Morrison *et al.*, 2009; Simo *et al.*, 2013; Simo et al., 2014a; Simo, *et al.*, 2014b). However, no studies have compared the populations differentiation based on mode of transmission i.e. from tsetse infested and tsetse free areas, which is the objective of this study.

**2.6 Genetic diversity**

Genetic diversity is the variation of alleles and genotypes in a population and is directly linked to the ability of a population to evolve (Greenbaum *et al.*, 2014; Sheriff and Alemayehu, 2018). To measure genetic diversity, the average observed and expected heterozygosity (probability that two alleles chosen at random are different), allele frequencies (mean number of alleles), deviation from the Hardy-Weinberg Equilibrium are tested, Fisher's inbreeding coefficient (probability that two alleles at a locus are identical by descent) and the Simpson's diversity index

(measures number of species present and their relative abundance) are calculated (Toro and Caballero, 2005; Sheriff and Alemayehu, 2018).

Two of the measures of genetic diversity are allelic richness and measures based on heterozygosities. Allelic richness is the mean number of alleles at a locus (Greenbaum *et al.*, 2014). Reduction in the number of alleles in a population lowers its potential to adapt to changing environments and thus prevents evolution by natural selection. In addition, high allelic richness promotes evolution in a population by making the genotypic space available to mutation, even when the levels are low (Greenbaum *et al.*, 2014). The measures based on heterozygosities are calculation of observed and expected heterozygosities. Observed heterozygosity is the frequency of individuals that are heterozygous in a population while expected heterozygosity is the probability that gametes chosen at random in a population are of different alleles (Toro *et al* 2009. A decrease in the number of observed heterozygotes may lead to a decrease in the fitness of individuals in a population (Greenbaum *et al.*, 2014)

**2.7 Population structure**

Population structuring analysis is done to figure out whether a population is genetically homogeneous or whether the population is made up of demes(sub-populations) that are genetically distinct (Patterson *et al.*, 2006). Two of the major methods used are Distance methods and the Model based methods.

The model based methods make the assumption that individuals in a cluster are drawn randomly from a parametric model (Jonathan, 2000). An example of these methods is the Bayesian clustering method (implemented in STRUCTURE) that uses two models; the allele frequency model (independent of correlated) and the reproduction model (admixture or no-admixture) (Jonathan, 2000; Heller, 2005). This method groups individuals in a population into K clusters

and makes the assumptions that markers are at linkage equilibrium and are unlinked and that the markers are at Hardy Weinberg Equilibrium.

The distance methods are based on phylogenetic algorithms and include Principle component analysis (PCA). Firstly, genetic distances between pairs of individuals are used to construct a genetic distance matrix which is represented as a dendrogram or a multidimensional scaling plot (Jonatha, 2000). However, while distance methods are easy to do and are visually appealing, it is often not easy to ascertain that the clusters obtained are not too heavily dependent on distance measures ( Jonathan *et al*., 2000; Patterson *et al*., 2006, Price *et al*., 2006).

**2.8 Genetic differentiation and migration (gene flow) analysis**.

Genetic differentiation is the buildup in allele frequency differences between isolated populations. One of the major drivers of genetic differentiation are local adaptations arising from spatial and temporal heterogeneity in selection pressures acting on heritable traits (Merilä and Crnokrak, 2001). In addition, random genetic drift, immigration and mutation are forces of genetic differentiation.

To quantify genetic differentiation, a fixation index (Wright's $F_{ST}$) is calculated. The parameter $F_{ST}$ quantifies genetic differentiation at a biallelic locus and ranges from 0-1 (Verity and Nichols, 2014). $F_{ST}$ is thus a characteristic of the demes (sub-populations) being studied. The inbreeding coefficient F in $F_{ST}$ is the association between yoking gametes that make up a diploid organism. In a situation where gametes forming a diploid individual have no correlation, then F would be equal to zero. However, where parents of the subject individual are closely related, the F value would be positive because homologous gene copies from the mother and father are likely to match and thus the individual has higher chances of being homogenous. In this case also, the individual is more likely to be genetically differentiated from the entire sub-population. The F

value is a property of individuals in a population however; the same logic can be used to explain genetic differentiation of demes from each other. Therefore, the inbreeding coefficient $F_{ST}$ is used to define the relationship among a pair of gametes drawn at random from a sub-population (deme). Subscripts S and T in $F_{ST}$ show that both the sub-populations and whole population are compared in the correlation (Balding, 2003; Verity and Nichols, 2014).

Genetic differentiation is also estimated through Analysis of Molecular Variance (AMOVA). With AMOVA, genetic distances (squared Euclidian distances) are used to calculate an $F_{ST}$ analogue known as $\Omega_{ST}$ ( MeirMans, 2006). While an AMOVA makes it possible to determine the reasons for hierarchical structuring in a population, it requires that the population's hierarchical structure is known as a prerequisite. Thus, the results of a clustering analysis such as STRUCTURE are used as the foundation of an AMOVA analysis. (MeirMans, 2012).

Evolutionary forces leading to genetic differentiation between populations include migration (gene flow), mutation, genetic drift and natural selection. Population differentiation indicators like Wright's $F_{ST}$ can be used to indirectly estimate migration rates between populations ($4N_em = 1/F_{ST} - 1$, with $N_em$ as gene flow rate) (Wilson and Rannala, 2003). However, this makes the assumptions that the populations are in an island model (an island model has a mainland population and a few island populations with migration being directional from the mainland population to the island populations) with constant population sizes and that migration from the mainland to the islands is symmetrical in every direction (Jost, 2008). Yet, in natural conditions, migration within structured populations is often asymmetrical particularly in systems where migration is caused by physical transport such as animal movement from one location to another (Sundqvist *et al.*, 2016). It then becomes crucial to determine the rate of directional migration in such cases so as to fully understand the drivers of genetic structuring within the population.

## 2.9 Landscape genetics

Landscape genetics is a field of study that blends population ecology, population genetics and spatial statistics. In this field, population geneticists study the influence of environmental factors on population structure and genetic variation (Storfer *et al.*, 2007). In addition, the field explains the relationship between landscape topographies and evolutionary process like selection, gene flow and genetic drift (Manel *et al.*, 2003). Landscape genetics generally involves correlating genetic data, geographical data and environmental data through measurement of dissimilarity. Population genetics models include isolation by distance (IBD), isolation by environment (IBE) / Genotype by environment association (GAE), Isolation by resistance (IBR), landscape genetic simulation modeling and Pattern-process modeling (Schwabl *et al.*, 2017).

Isolation by distance (IBD) is the decrease in genetic similarity between populations as the geographical distance between them increases (Dario *et al.*, 2017). It is a technique used among others to determine whether populations from different geographical origins are genetically separated (Bohonak, 2002). Populations' Isolation by distance is determined through plotting genetic distances or similarity between populations $(1/F_{ST} - 1)/4)$ against geographical distances between the populations. Isolation by distance can also be used to test whether population structure models are valid (Bohonak, 2002; Dario *et al.*, 2017).

Both geographical distances and environmental differences affect differentiation and population structure among organisms from two different geographical locations. Isolation by environment model (IBE) tests whether populations from different geographies have different genotypes (Jiang *et al.*, 2019). Isolation by environment directly affects adaptation of a population to their particular environment. Therefore, IBE positive population's resilience is determined by adaptability of the populations and the degree of environmental variations.

A Mantel test is one of the tests used to explore the relationship between genetic distances and geographical distances and genetic distances and environmental distances. The test is a landscape genetics tool that evaluates the amount of spatial structure in a genetic distance matrix (Diniz-Filho *et al.*, 2013).A partial Mantel test, compares two matrices while holding the effects of a third matrix constant. While studying the effects of geographical and environmental distances on genetic distances, a partial mantel test is used to detect the significance of both IBD and IBE simultaneously (Sexton *et al*., 2014) .

In parasitic populations, cases (for example outbreaks) in different geographies occur. Intense local transmission is encouraged by vector, host and parasite-related factors. Therefore, understanding how environmental and biological factors influence the movement of vectors, hosts and parasites is vital for disease control (Schwabl *et al.*, 2017).

## CHAPTER THREE: MATERIALS AND METHODS

### 3.1 Description of the study area

Samples analyzed in this study were obtained from Shurr (N02°.08', E038°.27'), Ngurunit (N01°.74', E 037.29'), and Laisamis (N 01° 23' 11" E 37° 57' 11.7") all in Marsabit County, Northern Kenya and Shimba Hills (Lat -4.243 and long 39.403) in Kwale County, Coastal Kenya (Figure 3). The main livestock in Shurr, Ngurunit and Laisamis are camels, cattle, goats and sheep while in Shimba Hills there are only cattle, goats and sheep. These livestock are the hosts and reservoirs of the trypanosome species in these areas. Shimba Hills in Kwale County falls within the Coastal tsetse fly belt. Ngurunit, Laisamis and Shurr are tsetse free areas and vectors of trypanosomes found there are camel flies, tabanids and horseflies (Getahun *et al.*, 2020).

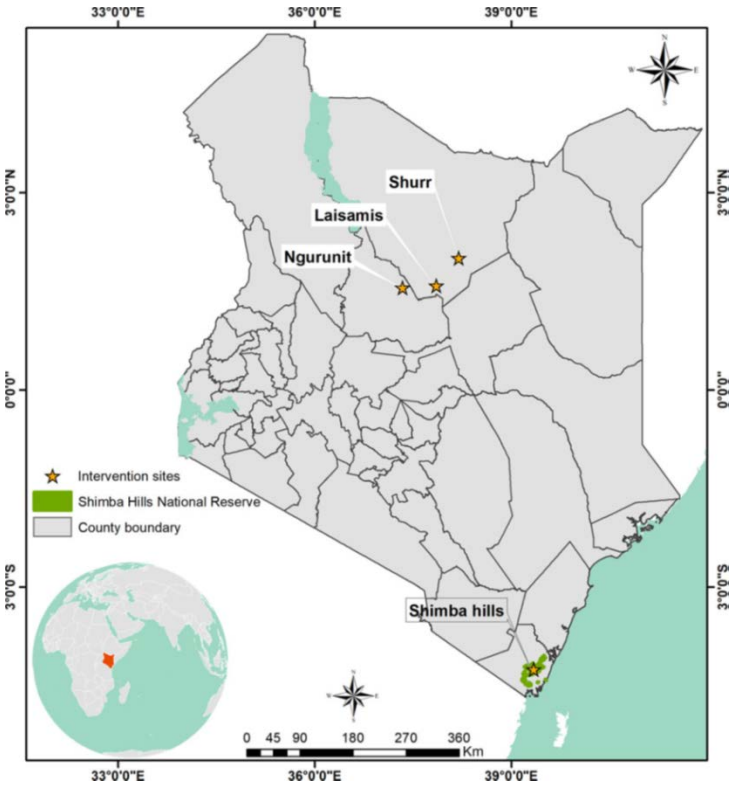



**Figure 3**: A map of Kenya showing locations of Shimba Hills, Ngurunit and Shurr from where samples used in the study were collected (Kimenyi *et al.*, 2021).

### 3.2 Samples

Trypanosome isolates from camel and cattle blood collected in Shurr, Laisamis and Ngurunit, in Marsabit County and cattle blood and tsetse flies collected in Shimba Hills, Kwale County were used. The blood and tsetse fly samples were collected between 2017 and 2019. For samples collected in 2017, DNA was extracted and stored at $-20^0$ C at the International Center for Inset Physiology and Ecology (*icipe*). All samples from Marabit County were isolated from camel and cattle blood. A majority of Kwale County isolates came from tsetse flies and a few from cattle blood. A convenient sample size was used to determine the sample size. Therefore, DNA was extracted from blood and insect samples until a sample size of 30 isolates from each region (Marsabit and Kwale Counties) was obtained.

### 3.3 Molecular Characterization

#### 3.3.1 DNA extraction

Blood samples from the field were stored at -80℃ at the International Center for Insect Physiology and Ecology (*icipe*). The blood samples (contained in vials) were first transferred into a 4℃ refrigerator overnight for thawing and left to complete thawing at room temperature right before DNA extraction for about thirty minutes. From there, 100µl of the blood sample were then transferred into a sterile micro-centrifuge tube and DNA extracted using a Qiagen DNeasy Blood and Tissue Kit (Qiagen, Germany) following the manufacturer's protocol (Getahun et al., 2020).

Tsetse flies' samples from Shimba Hills were stored in 70% ethanol. The tsetse flies were air dried in a fume chamber as a prerequisite to DNA extraction. Each fly was then chilled in liquid nitrogen and transferred into a micro- centrifuge tube into which silicon beads and lysis buffers were added. The micro-centrifuge tubes were then placed in Tissue Lyser II (Marogen) and the

contents homogenized at 30 $H_Z$ for three minutes. Subsequently, DNA was extracted from these tsetse fly samples using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Germany) and following the manufacturer's protocol (Getahun *et al*., 2020).

The quality of trypanosome positive archived DNA samples from previous work (Getahun *et al.*, 2020) (stored at -80℃ at the International Center for Insect Physiology and Ecology (*icipe*) was checked through gel electrophoresis. For samples with poor DNA quality, blood samples were traced back from the -80℃ freezer and DNA re-extracted.

Following each process of DNA extraction, gel electrophoresis was done with a 0.8% agarose gel to confirm the quality of genomic DNA obtained.

### 3.3.2 ITS Typing

The ITS-1 forward and reverse primers (ITS1 CF 5'CCGGAAGTTCACCGATATTG and ITS1 BR 5'TTGCTGCGTTCTTCAACGAA) were used for PCR based diagnostic test for trypanosome detection and classification as described previously by Njiru et al., 2005. A 10 µl PCR reaction mixture was set up with 1µl of the template, 0.5µl each of the forward and backward ITS1 primers, 5µl DreamTaq DNA Polymerase and 3µl of nuclease free water. The amplification reactions were carried out using the following PCR cycling profile: 95℃ for 3 minutes for the initial step, 35 cycles at 95℃ for 30 seconds, annealing at 61℃ for 30 seconds, extension at 72℃ for 30 seconds and final extension at 72℃ for 10 minutes (Getahun *et al*.,, 2020). For visualization, gel electrophoresis was carried out on the PCR products with 2% agarose in 100ml of Tris-Acetate-EDTA (TAE buffer) (0.4M Tris acetate and 10Mm EDTA, pH 8.3).

Thirty trypanosome positive samples (ten for each species) were then amplified with ITS primers (ITS CF 5' GCTGTAGGTGAACTTGCAGCAGCTGGATCATT and ITS BR 5'

GCGGGTAGTCCTGCCAAACACTCAGGTCTG) that target the whole ITS region in trypanosomes in its entirety as described by (Mossaad *et al*., 2017) for sequencing purposes. The amplification reactions were carried out using the following PCR cycling profile: 95℃ for 3 minutes for the initial step, 35 cycles at 95℃ for 30 seconds, annealing at 65℃ for 30 seconds, extension at 72℃ for 30 seconds and final extension at 72℃ for 10 minutes.

Gel electrophoresis was carried out on all PCR products with a 2% agarose gel in 100ml of Tris-Acetate-EDTA (TAE buffer). For visualization, the gel was stained with 0.5µg/ml ethidium bromide and observed under ultraviolet (UV) light.

### 3.4 Microsatellite Genotyping

#### 3.4.1 Fragment analysis

Fluorescently labeled forward primers (6- FAM, HEX, ROX) of previously described 17 microsatellite loci were used for microsatellite genotyping (Echodu *et al*., 2015; Kamidi *et al.*, 2017).

All 17 primer pairs were optimized for amplification of microsatellite loci in *Trypanozoon*, *T. congolense* and *T. vivax.* A 10 µl PCR reaction mixture was set up with 1µl of the template, 0.5µl each of the forward and backward microsatellite primers, 2µl Blend Taq$^{TM}$ DNA Polymerase and 6µl of nuclease free water. The PCR conditions used were: Initial denaturation at 95℃ for 15 minutes, 35 amplification cycles of 95℃ for 30 seconds, primer specific annealing temperature for 30 seconds, elongation at 72℃ for 30 seconds, and final elongation at 72℃ for 7 minutes. The quality of a number of PCR amplicons was then checked though gel electrophoresis and visualized under ultra violet light.

The PCR amplicons in 96 well plates were sent to University of Illinois in The United States of America for fragment sizing by capillary electrophoresis, with the Applied Biosystems

(Waltham, MA, USA) 3730 DNA Analyzer and 500 LIZ as the size standard. Allele scoring was done using Geneious Prime 2020.2.2 (https://www.geneious.com) software. Fragment sizes were then exported into an excel sheet and edited manually and with an excel plug; Microsatellite Toolkit.

### 3.4.2 Population structure

The STRUCTURE v2.3.4 software (Jonathan, 2000) was used to infer the population's structures for *Trypanozoon, T. vivax* and *T. congolense* isolates through the Bayesian clustering method. Ten independent runs each for K (genetic clusters) = 1-7 were performed with a burnin of 50,000 and 100,000 MCMC reps for 1,000 iterations. The optimal value of K was determined through the ad hoc statistic "ΔK"(Evanno *et al*., 2005) in Structure Harvester v0.6.94 (Earl and vonHoldt, 2012). STRUCTURE runs for *Trypanozoon* and *T. congolense* isolates were done with independent allele frequencies and admixture models while runs for *T. vivax* isolates were ran with correlated allele frequencies and no admixture models (Porras-Hurtado *et al.*, 2013). The STRUCTURE membership coefficients (Q-values) were used to assess probability of assignment of each isolate to a specific cluster. Sub-populations within *Trypanozoon* and *T. vivax* and *T. congolense* populations were also identified through multivariate analysis using principal component analysis (PCA), in R package Adegenet (Solymos *et al*., 2020). Multivariate analysis was done to complement Bayesian analysis and unlike the latter, it does not make assumptions on Hardy Weinberg equilibrium or linkage disequilibrium as it is not model based (Jombart *et al*., 2010). The optimal number of clusters was determined through Bayesian information criterion (BIC) (Jombart, 2008).

In addition, genetic clustering within populations of the three trypanosome species studied was also explored through hierarchical clustering. Cavalli-Sforze and Edwards genetic distances were

calculated and an UPGMA (Unweighted Pair Group Method with Arithmetic Mean) dendrogram constructed in Populations v1.2.32 (http://bioinformatics.org/populations/). The dendrogram was then viewed and edited in FigTree v1.31 (http://tree.bio.ed.ac.uk/software/figtree/).

### 3.4.3 Genetic diversity

Genetic diversity was estimated based on allele frequencies. Total number of different alleles, the effective number of alleles (no of equally frequent alleles required to achieve the same expected heterozygosity as the studied population), number of private alleles and Shannon's Information Index (I) were determined in GenALEX v6.5 (Peakall and Smouse, 2006). Allelic richness ($A_R$) was calculated in PopGenReport package in R (Adamack, 2014), observed heterozygosity ($H_O$), expected heterozygosity ($H_E$)( Nei's gene diversity) and Fisher's inbreeding coefficient ($F_{IS}$) were estimated in GenALEX v6.5 (Peakall and Smouse, 2006). As a test to the non-random association of alleles within diploid individuals and at different loci, agreement with Hardy-Weinberg equilibrium and linkage disequilibrium were estimated in R package Genepop (Rousset *et al*., 2020).

*Trypanosoma congolense* alleles were rarefied for the smallest samples described by (Kalinowski, 2004) upon which allelic richness and private allelic richness were determined in HP-RARE (Kalinowski, 2005).

### 3.4.4 Genetic Differentiation and Isolation by distance analysis

Pairwise $F_{ST}$ and their related P values for *Trypanozoon*, *T. vivax* and *T. congolense* isolates were calculated so as to estimate population differentiation at two levels; among the STRUCTURE inferred clusters and among the populations derived from sampling localities. Between and within population variance at the two levels was determined through analysis of molecular variance (AMOVA) in FSTAT v2.9.4 (Goudet, 2003). Isolation by distance (IBD) and isolation

by environment (IBE) were estimated by plotting among population genetic distances ((1/$F_{ST}$ − 1)/4) against between populations geographical (km) and environmental distance using a partial Mantel test (Nathaniel Mantel, 1967) analysed in R package vegan (Oksanen *et al.*, 2020)

### 3.4.5 Migration rate and Gene flow analysis

Gene flow ($N_e$m) within the three trypanosome species was inferred indirectly from the allele frequency data based on Wright's equation $FST = 1/(4Nem + 1)$ ($N_e$ is the effective population size and M is migration rate) (Wright, 1990; Whitlock, 1999). The BayesAss Edition 3.0software was used to estimate migration rate between populations of the three trypanosome species based on Bayesian inference (Wilson and Rannala, 2003).

# CHAPTER FOUR: RESULTS

## 4.1 Trypanosome isolates identified in blood and insect samples

### 4.1.1 DNA Extraction and gel electrophoresis

A total of five hundred and sixty seven DNA samples yielded good quality DNA. The quality and quantity of extracted DNA in each blood and tsetse fly sample was confirmed through gel electrophoresis as shown in Figure 4.

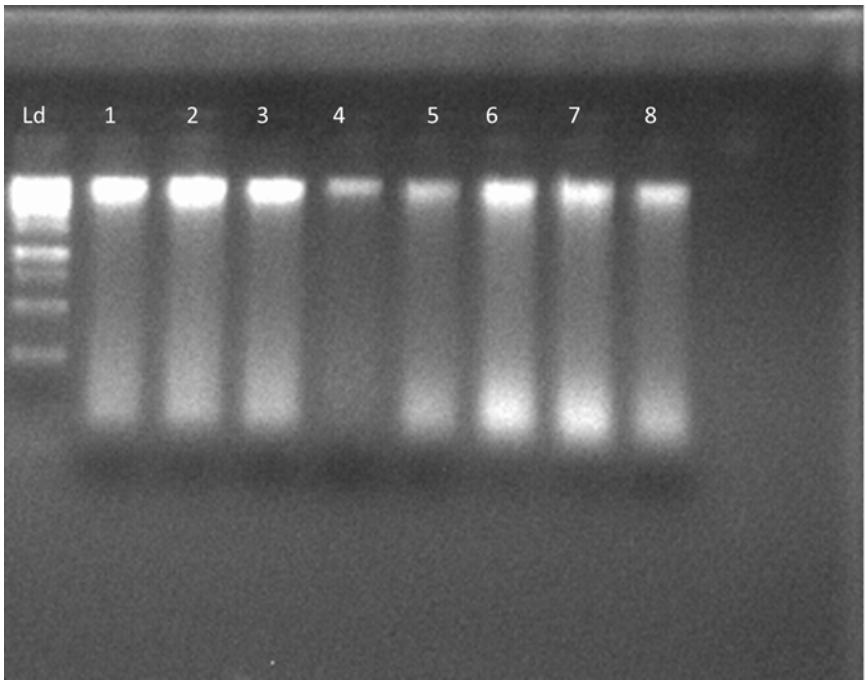



**Figure 4**: Genomic DNA of trypanosome blood samples.

Key: Ld- 1000bp ladder, Lane 1- 8: Extracted whole DNA samples.

### 4.1.2 PCR Amplification of Trypanosome isolates on ITS-1 markers

A total of one hundred and forty-four samples (Table 1) out of the five hundred and sixty seven were successfully amplified by ITS-1 primers as demonstrated in Figure 5. Of the one hundred and forty-four trypanosome positive samples, fifty-two (n= 52) were *Trypanozoon* isolates, thirty of them being from Marsabit County sites and twenty-two of them being from Kwale County.

Sixty *T. vivax* isolates (n= 60) were identified, thirty being from Marsabit County and thirty being from Kwale County. Out of thirty-two *T. congolense* isolates (n=32), only three were detected in samples from Marsabit County and the other twenty-nine from Kwale County.

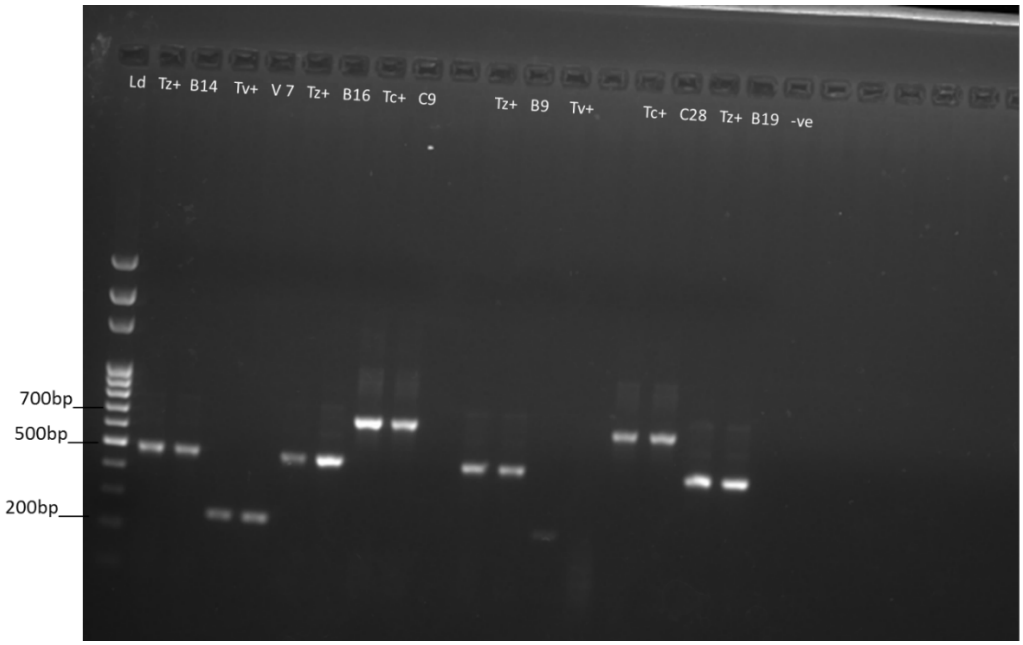



**Figure 5**: A gel image showing the different sizes of amplified fragment using ITS-1 primers. Key: Ld - 100 bp ladder; Tz+ - *Trypanozoon* positive control; Tv+ - *T. vivax* positive control; Tc - *T. congolense* positive control; B14 - 133R; V7 - 97R; B16 - 58C; C9 - Tangini 22; B9 - 65F; C28 - 140 23A; B19 - 36SH (Appendix I). B16 - *Trypanozoon*; V7- *T. vivax*; C9 - *T. congolense*.

**Table 1: Sample information of Trypanosome isolates used in the study.**

| Code | Sample ID | Species | Host | Sampling Location | Collection date |
|---|---|---|---|---|---|
| NB1 | 1 | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB2 | 15 | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB3 | 48 | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB4 | 45 | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB5 | 46 | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB6 | 48/23 | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB7 | 15F | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB8 | 97F | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB9 | 65F | *Trypanozoon* | Camel | Ngurunit | 2019 |
| NB10 | 1-Jun | *Trypanozoon* | Camel | Shurr | 2018 |
| NB11 | 2-Jun | *Trypanozoon* | Camel | Shurr | 2018 |
| NB12 | 3-Jun | *Trypanozoon* | Camel | Shurr | 2018 |
| NB13 | 36SH | *Trypanozoon* | Camel | Shurr | 2018 |
| NB14 | 37SH | *Trypanozoon* | Camel | Shurr | 2018 |
| NB15 | J SH | *Trypanozoon* | Camel | Shurr | 2018 |
| NB16 | 58C | *Trypanozoon* | Camel | Shurr | 2018 |
| NB17 | 45SN | *Trypanozoon* | Camel | Shurr | 2018 |
| NB18 | 30 SH | *Trypanozoon* | Camel | Shurr | 2018 |
| NB19 | 1D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB20 | 6D | *Trypanozoon* | Camel | Laisamis | 2019 |

| | | | | | |
|---|---|---|---|---|---|
| NB21 | 7D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB22 | 19D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB23 | 38D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB24 | 40D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB25 | 50D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB26 | 67D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB27 | 75D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB28 | 78D | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB29 | 18A | *Trypanozoon* | Camel | Laisamis | 2019 |
| NB30 | 50A | *Trypanozoon* | Camel | Laisamis | 2019 |
| CB1 | Mawia 6 | *Trypanozoon* | Cattle | Shimba Hills | 2018 |
| CB2 | 122 23A | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB3 | 123 23A | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB4 | 133 23A | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB5 | 166 23A | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB6 | 193 23B | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB7 | 203 23B | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB8 | 198 27B | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB9 | 202 27B | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB10 | 202 28A | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB11 | 240 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB12 | 247 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |

| | | | | | |
|---|---|---|---|---|---|
| CB13 | 302 28B | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB14 | 133 R | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB15 | 152 R | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB16 | 567 20-2 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB17 | 609 20-2 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB18 | 446 17-2 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB19 | 767 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB20 | 0-25 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB21 | 0-46 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| CB22 | 0-48 | *Trypanozoon* | Tsetse flies | Shimba Hills | 2020 |
| NV1 | 23 23-8 | *T. vivax* | Camel | Ngurunit | 2019 |
| NV2 | 95 23-8 | *T. vivax* | Camel | Ngurunit | 2019 |
| NV3 | 73 23-8 | *T. vivax* | Camel | Ngurunit | 2019 |
| NV4 | 97 23-8 | *T. vivax* | Camel | Ngurunit | 2019 |
| NV5 | 95 R | *T. vivax* | Camel | Ngurunit | 2019 |
| NV6 | 73 R | *T. vivax* | Camel | Ngurunit | 2019 |
| NV7 | 97 R | *T. vivax* | Camel | Ngurunit | 2019 |
| NV8 | 48 SH | *T. vivax* | Camel | Shurr | 2018 |
| NV9 | 31 SH | *T. vivax* | Camel | Shurr | 2018 |
| NV10 | ML 49 | *T. vivax* | Camel | Shurr | 2018 |
| NV11 | 318 P | *T. vivax* | Camel | Shurr | 2018 |
| NV12 | 312 P | *T. vivax* | Camel | Shurr | 2018 |

| | | | | | |
|---|---|---|---|---|---|
| NV13 | 3 ML | *T. vivax* | Camel | Shurr | 2018 |
| NV14 | 49 ML | *T. vivax* | Camel | Shurr | 2018 |
| NV15 | 22 ML | *T. vivax* | Camel | Shurr | 2018 |
| NV16 | 23 20-8 | *T. vivax* | Camel | Shurr | 2018 |
| NV17 | C1 SH | *T. vivax* | Camel | Shurr | 2018 |
| NV18 | C2 SH | *T. vivax* | Camel | Shurr | 2018 |
| NV19 | 9A | *T. vivax* | Camels | Laisamis | 2019 |
| NV20 | 50A | *T. vivax* | Camels | Laisamis | 2019 |
| NV21 | 52A | *T. vivax* | Camels | Laisamis | 2019 |
| NV22 | 19A | *T. vivax* | Camels | Laisamis | 2019 |
| NV23 | 36 A | *T. vivax* | Camels | Laisamis | 2019 |
| NV24 | 44A | *T. vivax* | Camels | Laisamis | 2019 |
| NV25 | 37A | *T. vivax* | Camels | Laisamis | 2019 |
| NV26 | 21A | *T. vivax* | Camels | Laisamis | 2019 |
| NV27 | 38A | *T. vivax* | Camels | Laisamis | 2019 |
| NV28 | 6A | *T. vivax* | Camels | Laisamis | 2019 |
| NV29 | 16A | *T. vivax* | Camels | Laisamis | 2019 |
| NV30 | 49A | *T. vivax* | Camels | Laisamis | 2019 |
| CV1 | Mangawani 3 | *T. vivax* | Cattle | Shimba Hills | 2018 |
| CV2 | Mangawani 6 | *T. vivax* | Cattle | Shimba Hills | 2018 |
| CV3 | Mangawani 8 | *T. vivax* | Cattle | Shimba Hills | 2018 |
| CV4 | 33 Glossina | *T. vivax* | Tsetse flies | Shimba Hills | 2018 |

| | | | | | |
|---|---|---|---|---|---|
| CV5 | KN139 | *T. vivax* | Cattle | Shimba Hills | 2018 |
| CV6 | KN22 | *T. vivax* | Cattle | Shimba Hills | 2018 |
| CV7 | 158 23A | *T. vivax* | Tsetse flies | Shimba Hills | 2018 |
| CV8 | 169 23A | *T. vivax* | Tsetse flies | Shimba Hills | 2018 |
| CV9 | 129 23A | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV10 | 183 23A | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV11 | 197 28B | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV12 | 214 27C | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV13 | 260 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV14 | 315 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV15 | 193 R | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV16 | 164 R | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV17 | 153 R | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV18 | 104 R | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV19 | 587 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV20 | 591 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV21 | 606 20-2 | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV22 | 434 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV23 | 439 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV24 | 522 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV25 | 62 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV26 | 132 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |

| | | | | | |
|---|---|---|---|---|---|
| CV27 | 26 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV28 | 722 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV29 | 753 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| CV30 | 757 F | *T. vivax* | Tsetse flies | Shimba Hills | 2020 |
| NC1 | 45 F | *T. congolense* | Camel | Ngurunit | 2019 |
| NC2 | 5F | *T. congolense* | Camel | Ngurunit | 2019 |
| NC3 | 7 F | *T. congolense* | Camel | Ngurunit | 2019 |
| CC1 | KN 62 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC2 | Mawia 2 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC3 | KN 42 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC4 | Mawia 7 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC5 | KN 63 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC6 | Mawia 4 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC7 | Mangawani 2 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC8 | Coast 1 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC9 | Tangini 22 | *T. congolense* | Cattle | Shimba Hills | 2018 |
| CC10 | 206 27B | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC11 | 211 27B | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC12 | 437 17TH | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC13 | 451 17TH | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC14 | 516 S | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC15 | 526 S | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |

| | | | | | |
|---|---|---|---|---|---|
| CC16 | 760 S | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC17 | 767 S | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC18 | 786 S | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC19 | 49 D1 | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC20 | 567 S | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC21 | 145 S | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC22 | 125 23A | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC23 | 303 28B | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC24 | 264 28B | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC25 | 119 L | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC26 | 121 23A | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC27 | 133 23A | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC28 | 140 23A | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |
| CC30 | 147 23A | *T. congolense* | Tsetse flies | Shimba Hills | 2020 |

**4. 2 Allele scoring**

All seventeen primers amplified microsatellite DNA in *Trypanozoon* isolates (Sample gel image in Figure 6) but only fourteen were successfully scored and alleles generated. For *T. vivax* isolates, only nine of the seventeen microsatellite primers amplified microsatellite loci and of this, six loci were scored and alleles generated while six microsatellite primers amplified loci in *T. congolense* and five scored with alleles.

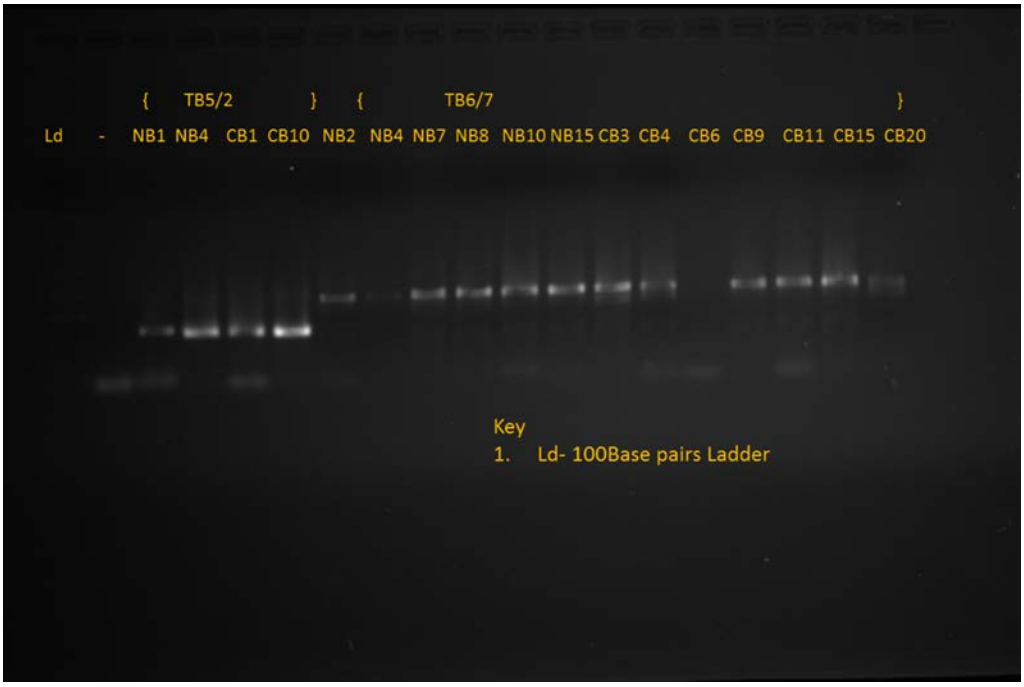



**Figure 6**: A gel image of microsatellite loci TB5/2 (83-107 base pairs) and TB6/7 (104-134bp) amplification for a number of *Trypanozoon* isolates.

Microsatellite primers that were successfully scored for each trypanosome species and their annealing temperature are described in Table 2.

All one hundred and forty-five Trypanosome isolates were successfully genotyped. However, there was 8.65%, 9.17% and 2.08% missing data, in the *Trypanozoon*, *T. vivax* and *T. congolense* data sets respectively (Appendices VIII, IX and X).

**Table 2: Microsatellite loci primers details.** Columns *T. e, T. v* and *T. c* are of annealing temperatures for *Trypanozoon, T. vivax and T. congolense* respectively.

Loc - Location; ℃ - Annealing temperature.

| Locus | Forward Primer | Reverse Primer | Motif | Size | Loc | (℃) | Reference |
|---|---|---|---|---|---|---|---|
| *Trypanozoon* | | | | | | | |
| TB 8/11 | [FAM]-TGTAGCAGTGGTACGCAC | CACCCAACGCATGTAAGC | AT | 97-127 | 8 | 52 | (Balmer *et al*., 2006) |
| TB 2/19 | [HEX]-CTGGTGCGTGTAACTGTG | GAAGTGAGGACATGCACG | AT | 84-104 | 2 | 57 | (Balmer *et al*., 2006) |
| TB 11/13 | [FAM]-CAAGAACTCTGCATTGAGC | ATCTGTTGGCGATGGTGA | AT | 125-161 | 11 | 57 | (Balmer *et al*., 2006) |
| TB 6/7 | [HEX]-AAGCTGACAGGTGGTTGA | GAACATGCGTGCGTGTG | AT | 104-136 | 6 | 54 | (Balmer *et al*., 2006) |
| TB 1/8 | [FAM]-AGGTTTAGTGCATGTCGGA | CCTGTTGTACGGAGGTCA | CA | 97-117 | 1 | 59 | (Balmer *et al*., 2006) |
| TB10/5 | [FAM]-AAAGGCGATATGTTATTATTGA | ATTGGGTATACTGTCCCTCA | TA | 79-115 | 10 | 54 | (Balmer *et al*., 2006) |
| TB 9/6 | [HEX]-TGATTCATTGGTTAAGACAGG | AATGATAACTGCGGATTACAC | AC | 124-158 | 9 | 50 | (Balmer *et al*., 2006) |
| Tryp 62 | [ROX]-AAGGCGACCAACTTCAACC | GTTGTCATCGGCTTGCTCC | AC | 153-177 | 11 | 61 | Sistrom *et al*., 2013 |
| Tryp 67 | [FAM]-GTTGCTGAGGTGCAACTGG | GTCGTCAGGCACCAAAACG | GTT | 151-178 | 7 | 61 | Sistrom *et al*., 2013 |
| TB 3/3 | [HEX]CATTCGAAGTAAATGCGCGTATAAC | GGTTGGAGCTTTCGACACAAGCG | AT | 72-132 | 3 | 59 | (Salim *et al*., 2011) |
| TB | [FAM]- | GGTCGGTGTTGGCAGTGTG | GT | 170-230 | 7 | 63 | (Salim, *et al*., |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 7/12 | CATGGCGTACGTTGCTTCGGTTTC | CATAG | | 123-183 | | | 2011) |
| TB 8/1 | [HEX]-CCAAATATGCGATTAGTTTCC | TGTTTATGTGGAAGGAAATGAA | TA | | 8 | 55 | (Salim, *et al*., 2011) |
| TB 11/29 | [FAM]-AATGAGTGATACTATGAAAGTGT | CACCATCACTGCTCTTATCA | CA | 122-152 | 11 | 54 | (Salim, *et al*., 2011) |
| *T. vivax* | | | | | | | |
| TB 8/11 | [FAM]-TGTAGCAGTGGTACGCAC | CACCCAACGCATGTAAGC | AT | 97-127 | 8 | 58 | (Balmer *et al*., 2006) |
| TB 1/8 | [FAM]-AGGTTTAGTGCATGTCGGA | CCTGTTGTACGGAGGTCA | CA | 97-117 | 1 | 52 | (Balmer *et al*., 2006) |
| Tryp 67 | [FAM]-GTTGCTGAGGTGCAACTGG | GTCGTCAGGCACCAAAACG | GTT | 151-178 | 7 | 59 | Sistrom *et al*., 2013 |
| TB 11/29 | [FAM]-AATGAGTGATACTATGAAAGTGT | CACCATCACTGCTCTTATCA | CA | 122-152 | 11 | 52 | (Salim, *et al*., 2011) |
| Tryp 54 | [ROX]-AGTCGGCGTGATGGTACTC | TTCAGCCCACAAACAACCG | AAAT | 144-176 | 10 | 58 | Sistrom *et al*., 2013 |
| Tryp 55 | [FAM]-AATTCAACCCCAACAGCCC | CTCGTTCAATGACTTGCCCC | GT | 208-246 | 5 | 52 | Sistrom *et al*., 2013 |
| *T. congolense* | | | | | | | |
| TB 8/11 | [FAM]-TGTAGCAGTGGTACGCAC | CACCCAACGCATGTAAGC | AT | 97-127 | 8 | 52 | (Balmer *et al*., 2006) |
| TB 6/7 | [HEX]-AAGCTGACAGGTGGTTGA | GAACATGCGTGCGTGTG | AT | 104-136 | 6 | 54 | (Balmer *et al*., 2006) |
| TB 1/8 | [FAM]-AGGTTTAGTGCATGTCGGA | CCTGTTGTACGGAGGTCA | CA | 97-117 | 1 | 59 | (Balmer *et al*., 2006) |
| TB 5/2 | [HEX]-CAACCGAAAGTAAGGGGAAC | TCTCGCCTTCTTTGCCC | AT | 83-107 | 11 | 55 | (Balmer *et al*., 2006) |

### 4.3 *Trypanozoon* Population Genetics Parameters

#### 4.3.1 Population structure

Structure harvester results indicated the best K value as K = 2 for Bayesian clustering analysis with STRUCTURE v2.3.4 (Figure 7A) thus indicating that two distinct genetic sub-populations are the most likely hierarchical level of population structure. One sub-population includes all isolates from Marsabit County (tsetse free) and the other sub-population includes all sub-populations from Kwale County which is tsetse infested. The next best fit of K =3 revealed sub-structuring within isolates from Kwale County (Figure 7B). One sub-population (red) includes thirty-one *Trypanozoon* isolates, all but one being from Marsabit County and another sub-population (green) consists of four isolates all from Kwale County region. The final sub-population (blue) is made up of seventeen isolates, all from the Kwale County. Assignment of all isolates to these three clusters based on Q values is displayed in Appendix II. Of the fifty-two *Trypanozoon* isolates, six Kwale County isolates and one Marsabit County isolate showed uncertain assignment to either of the three clusters ($Q < 0.8$) and these were excluded from further STRUCTURE based analysis.

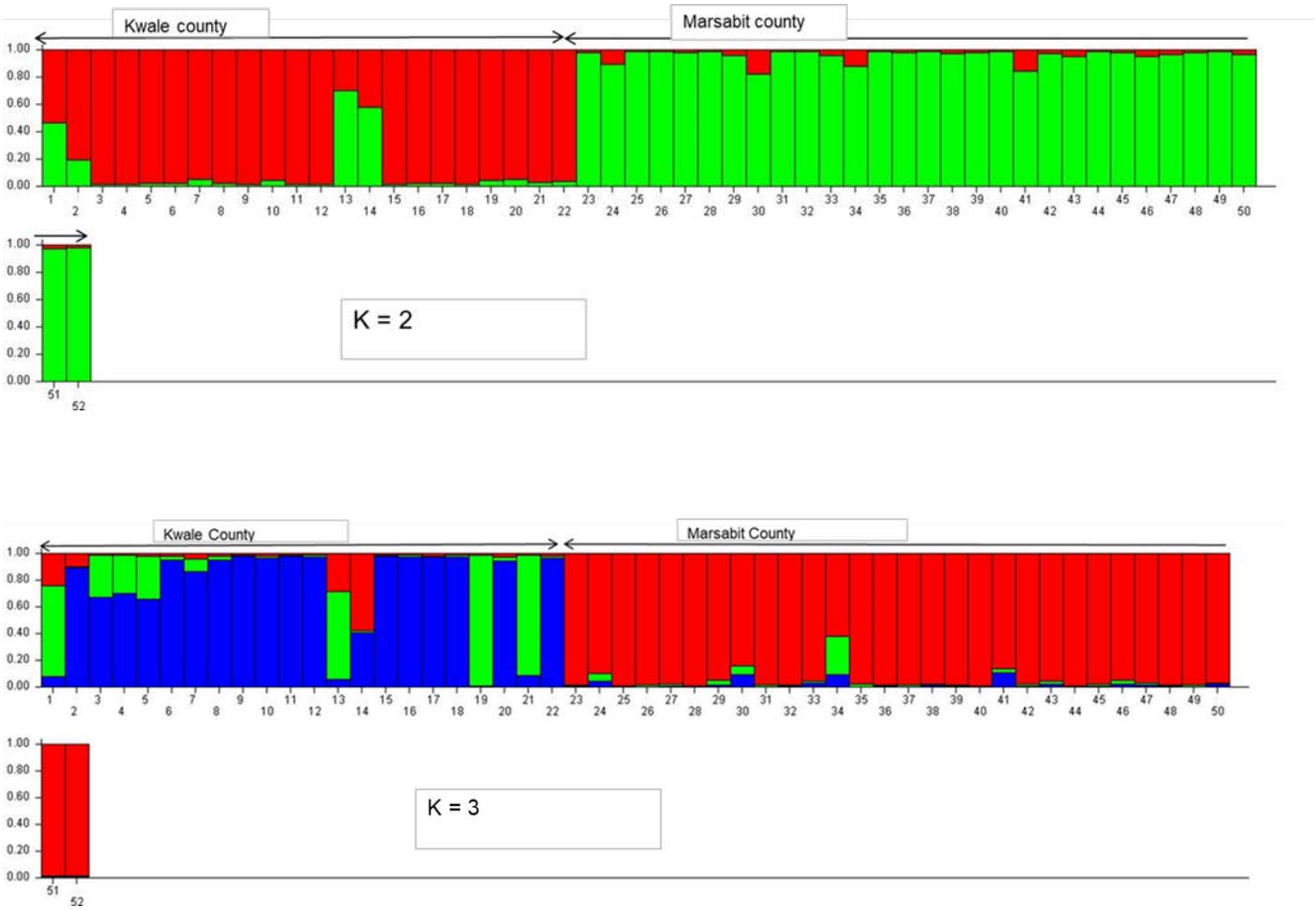



**Figure 7**: Bayesian clustering based on STURUCTURE results on fifty-two *Trypanozoon* isolates (*T. brucei* and *T. evansi*) from Marsabit and Kwale Counties in Kenya based on fourteen microsatellite loci. STRUCTURE harvester results based on the adhoc statistic indiated K=2 as the best K value and the next best fit of K was K =3 which revealed substracturing within iolats from shimba hills in Kwale County.

Multivariate analysis confirmed Bayesian clustering results by revealing two distinct genetic clusters that were clearly differentiated based on geography (Figure 8). Each cluster was dominated by isolates from the two different sampling localities indicating the relevance of geographical isolation in clustering by multivariate analysis of these isolates. Cluster a in the PCA includes two isolates from Kwale Kenya in addition to all isolates from Marabit County while cluster b includes twenty isolates from Marsabit County.

Hierarchical clustering by a UPGMA (unweighted pair group method with arithmetic mean) (Figure 9) further confirmed the results of Bayesian and multivariate analysis by revealing a population structured according to geography. Though not strong (bootstrap = 0.4023), the first level of separation clearly separated Marsabit County samples from Kwale County samples. However, unlike other clustering methods where two isolates (CB13 and CB14) from Kwale County grouped among Marsabit County isolates, all isolates from both locations grouped separately.

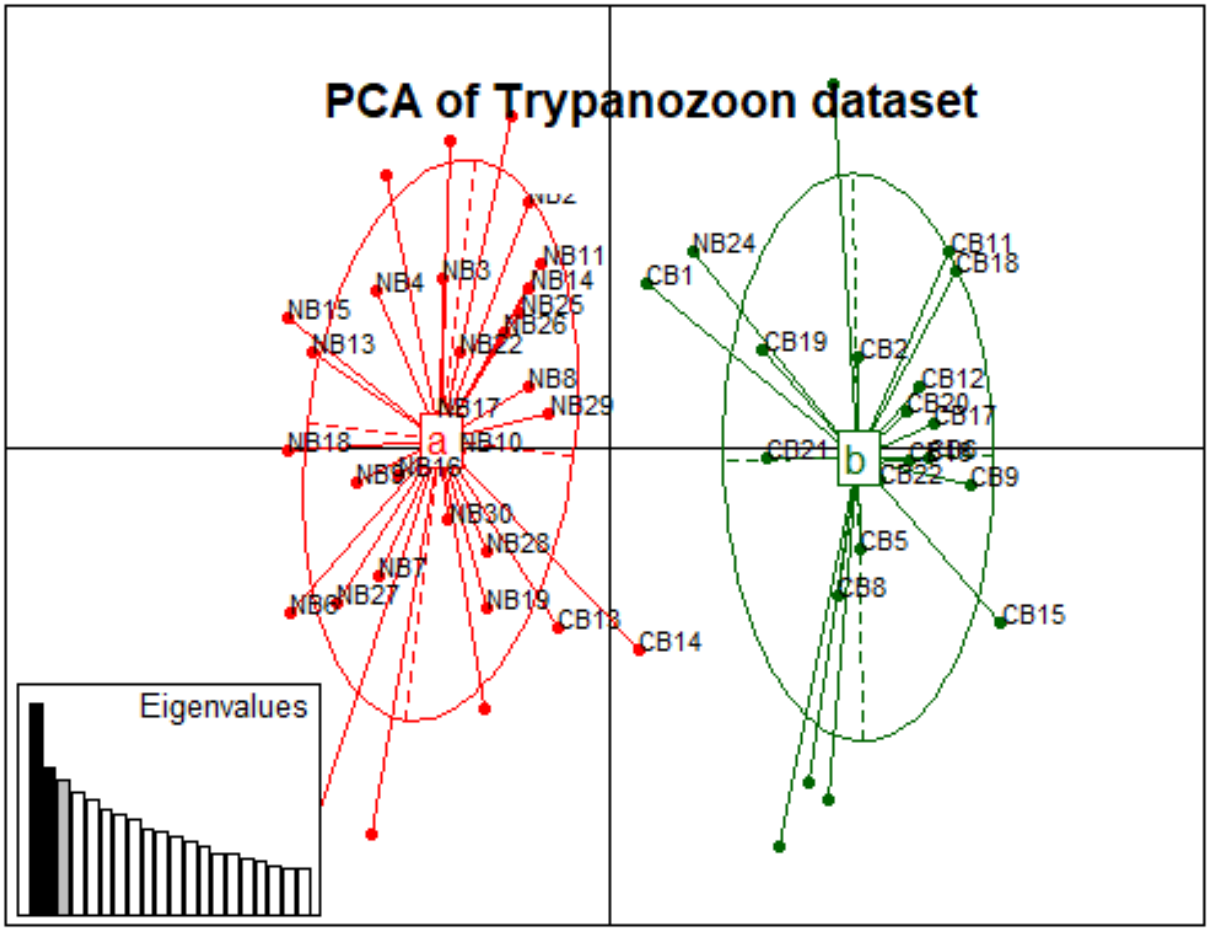



**Figure 8**: Principal Component Analysis (PCA) showing genetic population structure of the *Trypanozoon* isolates. Cluster a, is consists mainly of isolates from Marsabit County other than isolates CB13 and CB14, while Cluster b consists mainly of isolates from Kwale County.

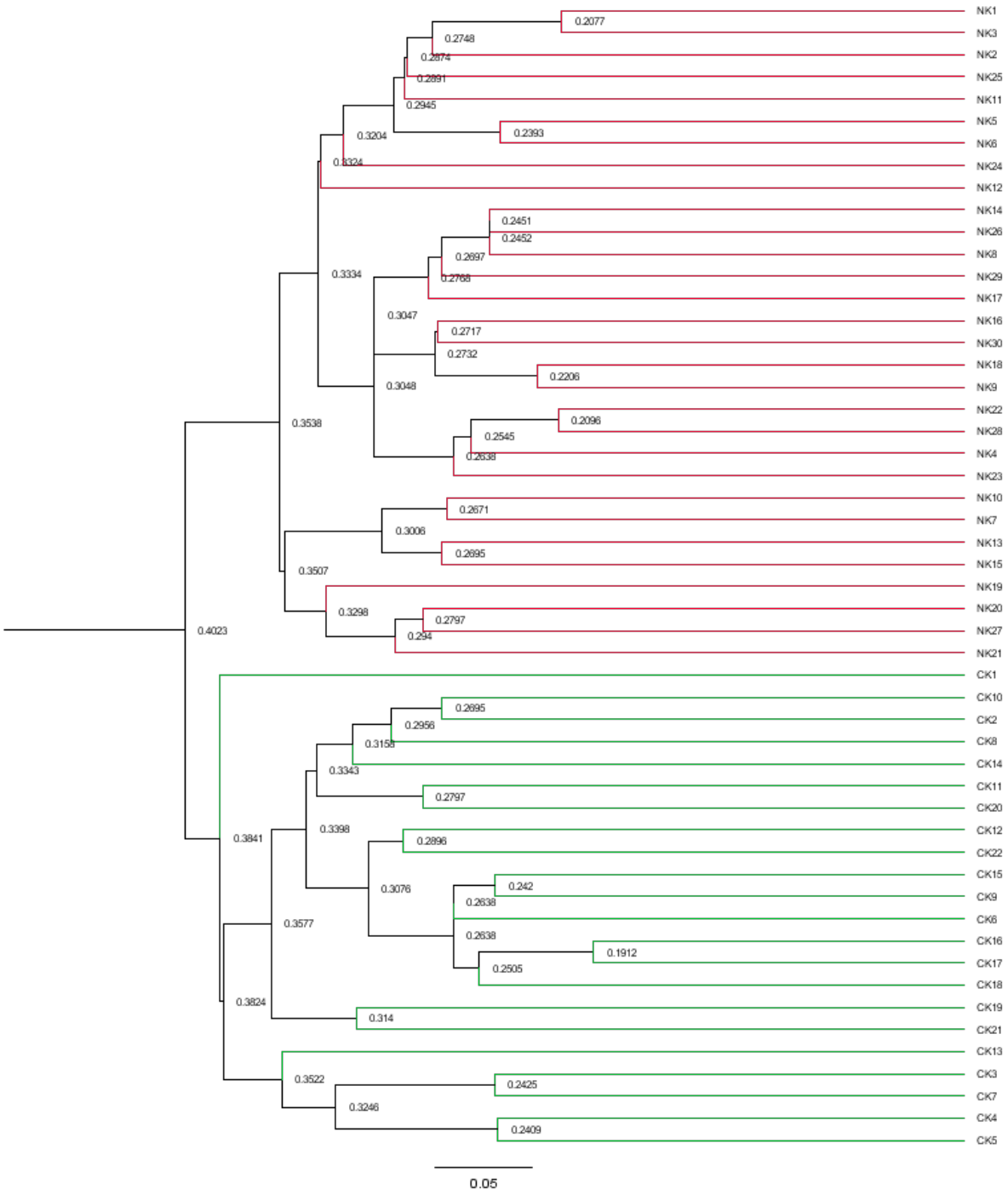



**Figure 9**: An UPGMA dendrogram for *Trypanozoon* isolates constructed based on 14 microsatellite loci and 1000 Bootstraps using Cavalli-Sforza and Edwards distances. The red branches are of isolates rom Marsabit County while the green branches are of isolates from Shimba Hills in Kwale County.

### 4.3.2 Genetic diversity

Only 4% of all the loci combinations were in linkage disequilibrium. However, no loci showed conformity to HWE in both populations. Among location of origin-based populations, 171 different alleles were identified with different number of alleles estimated at 9.714 and 9.286 for Kwale County and Marsabit County populations respectively. Shannon's Information Index (I) for the both populations were 0.114 and 0.084 respectively. The average number of private alleles for the coastal and Marsabit County was 2.929 and 2.50 respectively.

Allelic richness detected was 5.08 and 5.33 for the Marsabit County and Kwale County populations respectively, showing relatively greater diversity among Kwale County isolates. Observed heterozygosity (Ho) was 0.79 and 0.835 while expected heterozygosity (He) was 0.819 and 0.813 for Marsabit and Kwale County populations respectively (Table 3). The Fisher inbreeding coefficient ($F_{IS}$) for the Marsabit County isolates was -0.044 and that for the Kwale County isolates was 0.023. Among the structure-defined clusters, allelic richness ranged from 4.142 in cluster c to 3.695 in cluster b. Observed heterozygosity ranged from 0.836 to 0.759 while expected heterozygosity ranged from 0.807 to 0.772. Inbreeding coefficient ($F_{IS}$) values ranged between -0.053 and 0.011 (Table 3).

**Table 3:** Genetic diversity parameters among *Trypanozoon* isolates for both origin based populations and STRUCTURE clusters based populations

N-Sample size; Na-Average sample size; Ne-Effective sample size; Ho-Observed; Heterozygosities; uHe-unbiased expected heterozygosity; F-Fixation Index; I- Shannon's; Information Index; Ar- Allelic richness

| Pop | N | Na | Ne | Ho | uHe | F | I | Ar |
|---|---|---|---|---|---|---|---|---|
| red | 29 | 9.0 | 5.292 | 0.836 | 0.807 | -0.053 | 1.804 | 4.487 |
| green | 2 | 2.571 | 2.4 | 0.75 | 0.726 | -0.444 | 0.882 | - |
| Blue | 14 | 7.286 | 4.604 | 0.759 | 0.772 | 0.011 | 1.625 | 4.201 |
| Overall | 52 | 6.286 | 4.099 | 0.809 | 0.815 | -0.148 | 1.625 | 4.344 |
| Pop | N | Na | Ne | Ho | uHe | F | I | Ar |
| Kwale | 22 | 8.429 | 5.082 | 0.79 | 0.819 | 0.023 | 2.026 | 5.33 |
| Marsabit | 30 | 10.857 | 6.701 | 0.835 | 0.813 | -0.044 | 1.752 | 5.08 |
| Overall | 52 | 9.643 | 5.892 | 0.812 | 0.816 | -0.01 | 1.889 | 5.205 |

### 4.3.3 Population differentiation

Fixation index ($F_{ST}$) value between sampling localities was 0.0617 and revealed significant differentiation between the populations ($P< 0.05$). The small $F_{ST}$ value is an indication of moderate differentiation between Marsabit and Kwale County. Among STRUCTURE based cluster's $F_{ST}$ values were 0.117 between red sub-population and green sub-population, 0.0965 between the red sub-population and the green sub-population, 0.1062 between the blue sub-population and the green sub-population indicating rather moderate differentiation among all

clusters. Differentiation between the blue sub-population and the red sub-population was significant (p=0.012).

AMOVA results (Figure 10) revealed that 99% of the variation was caused by differences in genotypes within isolates and differences between populations accounted for 1% of the variation.

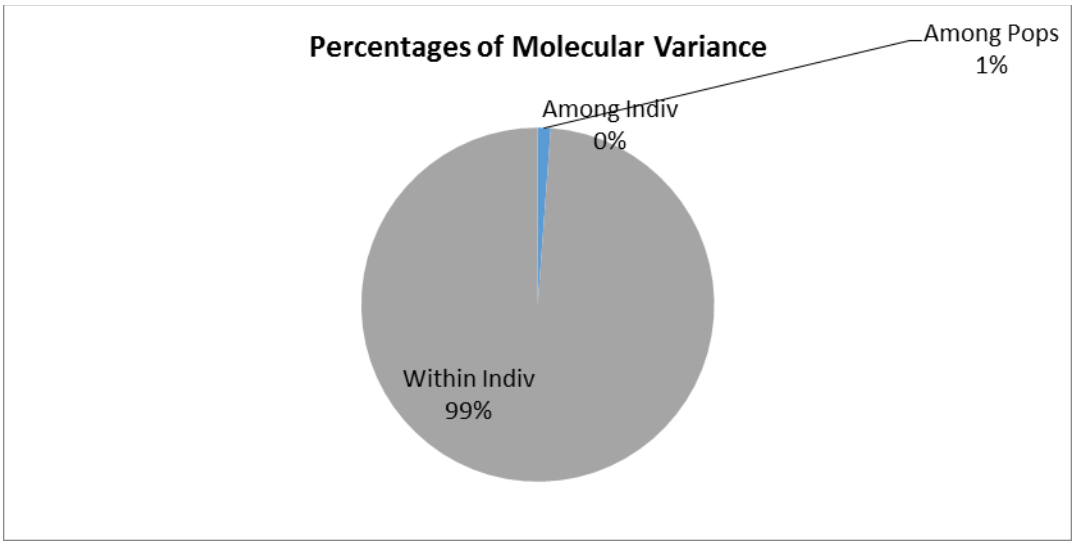



**Figure 10**: An AMOVA diagram of *Trypanozoon* isolates revealing that differences between genotypes within individuals were the greatest factor contributing towards genetic differentiation.

The partial Mantel Test results revealed a significant correlation (r = 0.3493, p =1e-04) between pairwise genetic distances $F_{ST}$ (1/1-$F_{ST}$) and geographical distances (Haversine distances) while controlling for environmental distances (Euclidian distances).

Recent emigration and migration rates between coastal and northern populations were 0.0113 and 0.0391 respectively. The high rate of self-distribution within both populations was an indication of asymmetric gene flow within the populations. Gene flow ($N_em$) or the effective number of migrants in the coastal and northern populations was 3.801 and 3.803 respectively.

## 4.4 *Trypanosoma vivax* Population Genetics Parameters

### 4.4.1 Population structure

A K value of three (K=3) (number of clusters) was suggested as the most probable level of hierarchy of population structure by Bayesian analysis (Figure 11). Cluster a(red) is made up of nineteen *T. vivax* isolates all of which were from Marsabit County. Cluster b (green) includes twenty-one isolates, thirteen of which are from Kwale County and eight from Marsabit County. Cluster c(blue) includes twenty isolates, three of which are from Marsabit County and seventeen from Kwale County. Assignment of the *T. vivax* isolates to specific clusters based on the Q-values is displayed in Appendix II. Ten *T. vivax* isolates from Kwale County and eight from Marsabit County showed uncertain assignment to either of three clusters ($Q<0.8$) and these were excluded from further (STRUCTURE based populations) analysis.

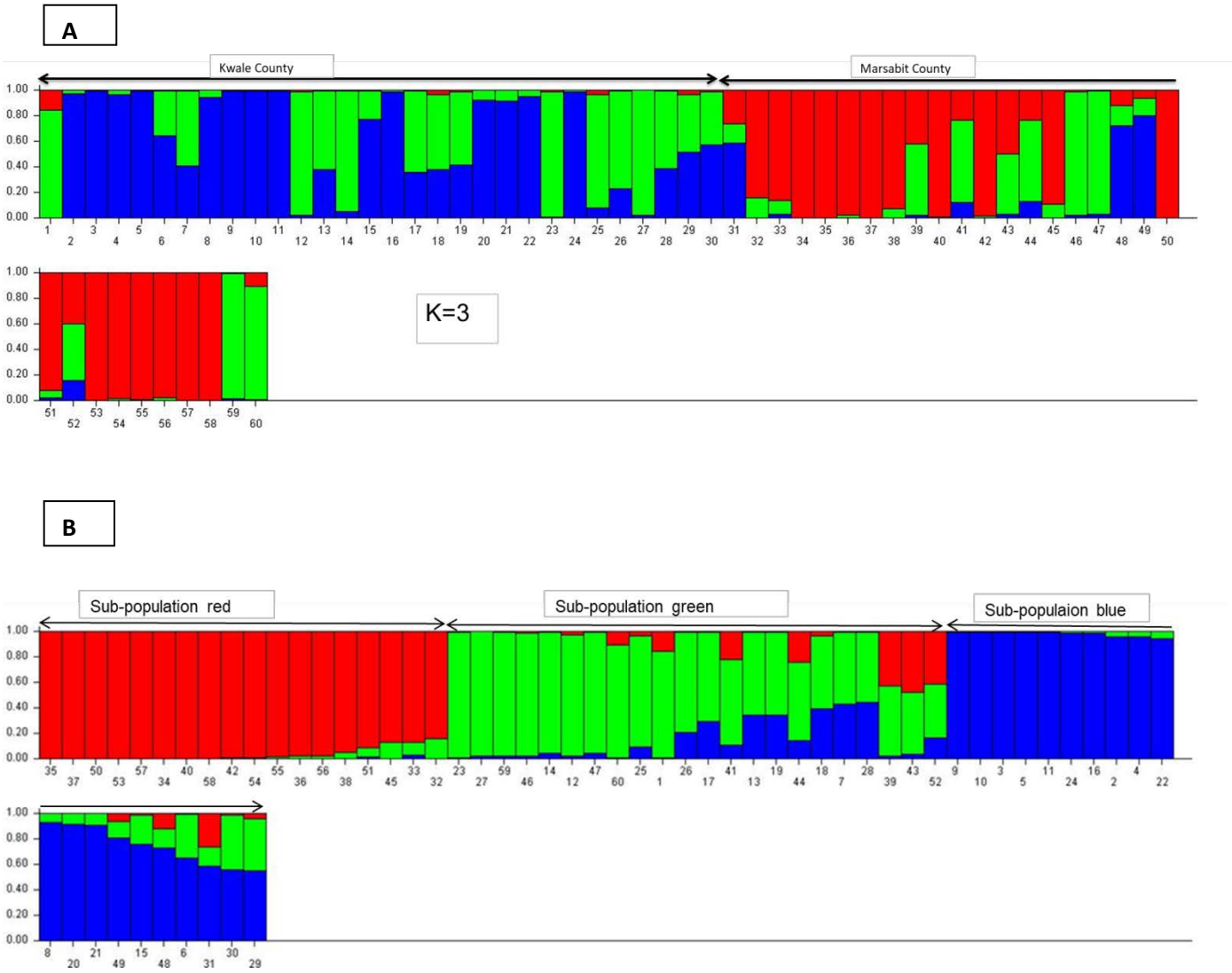



**Figure 11:** STRUCTURE generated image of best fit for K (K=3) on sixty *T. vivax* isolates from Marsabit County and Kwale County regions based on six microsatellites loci. Figure 11.A Shows assignment of samples (within their location of origin based populations) into different sub-populations. Figure 11.B shows the classification of samples into their STRUCTURE sub-populations regardless of their location of origin.

Multivariate analysis for *T. vivax* isolates confirmed presence of three distinct genetic clusters, which agrees with Bayesian analysis (Figure 12). None of the three clusters are found in the

same multivariate space. PC axis 1 separates cluster a from clusters b and c while PC axis 2 separates cluster b from cluster c. Cluster a includes eighteen isolates from Marsabit County, cluster b includes eleven isolates from Marsabit County and seventeen isolates from Kwale County. Cluster c includes only one isolate from Marsabit County and thirteen isolates from Kwale County. Multivariate analysis like Bayesian analysis revealed separation of a group of Marsabit County isolates from Kwale County isolates and also clustered the Kwale County isolates into two clusters.

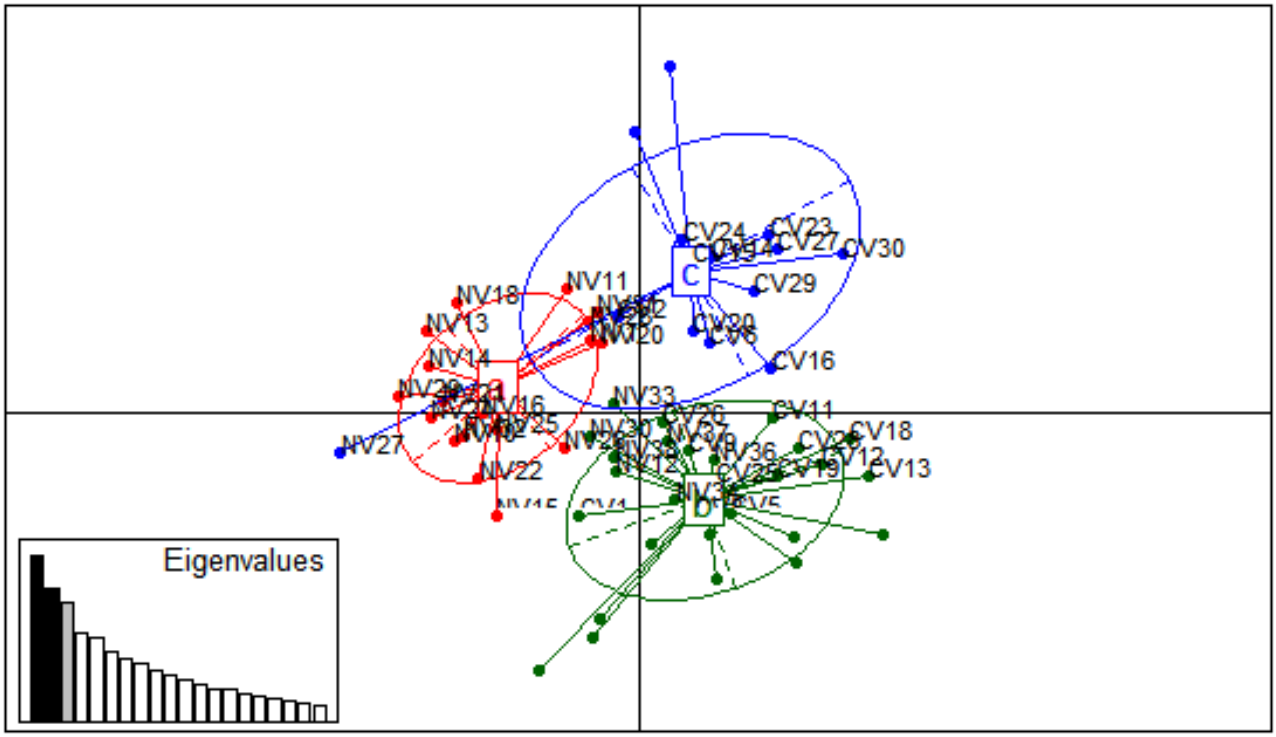



**Figure 12**: A Principal Component Analysis (PCA) showing how *T. vivax* samples from both sampling locations are genetically structured. Cluster a is made up of isolates from Marsabit County, cluster b has samples from both locations while cluster c has samples from only Kwale County.

Further, we constructed an UPGMA dendrogram to show the relationship between *T. vivax* isolates from Marsabit County and Kwale County (Figure 13). The dendrogram composed of two

distinct genetic clusters among all isolates. In addition, there was further sub-structuring in one of the two clusters. This tree agreed with Bayesian and multivariate analysis results by revealing that *T. vivax* isolates from the same sampling location did not clustered together. The tree also confirmed the presence of a group of Marsabit County *T. vivax* isolates that formed a unique cluster away from Kwale County isolates. However, the strength of the bootstrap values among the clusters was low therefore the inference drawn from the dendrogram results is limited.

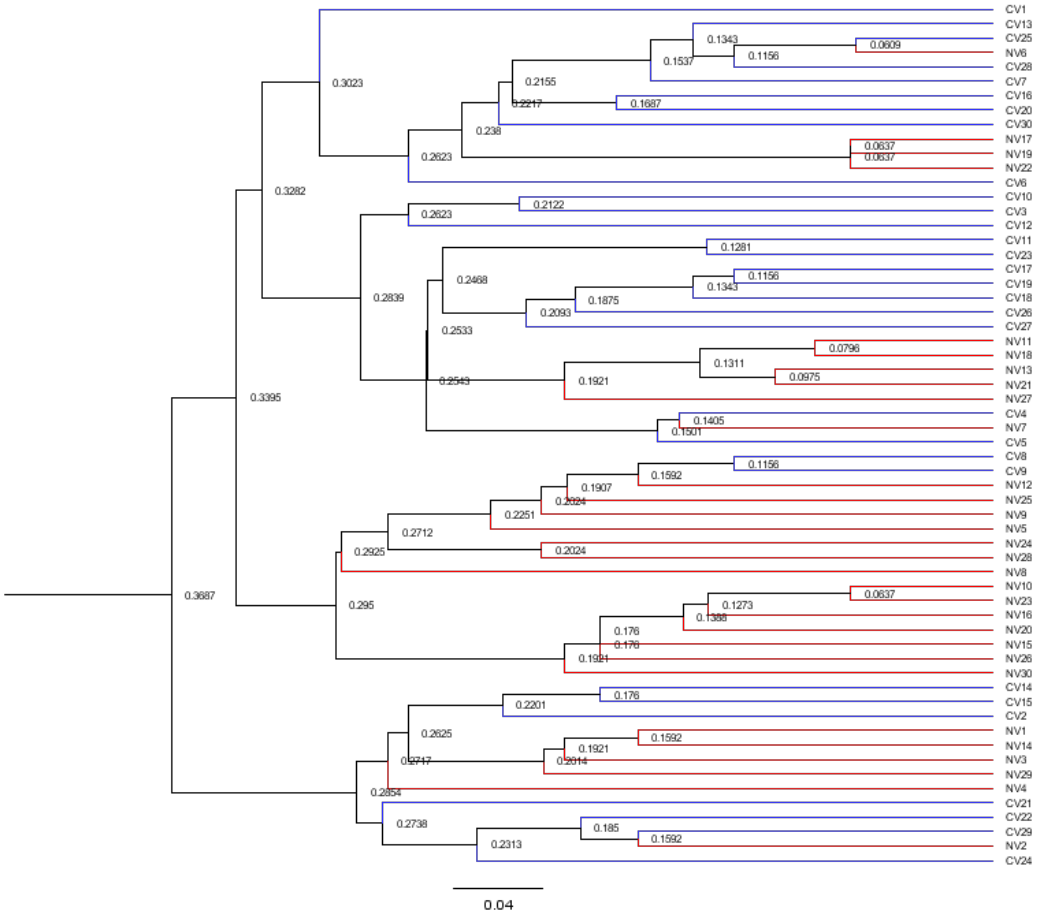



**Figure 13***:* A UPGMA tree based on six microsatellite marker, constructed using Cavalli-Sforza and Edwards distances showing hierarchical structuring of *T. vivax* isolates. The blue branches are of isolates from Kwale County while the red branches are of isolates from Marsabit County.

### 4.4.2 Linkage disequilibrium and genetic diversity

Only one (TB8/11 and TB1/8) of thirty loci combinations was in linkage disequilibrium ($p < 0.05$). Yet, all but loci TB1/8 showed deviation from Hardy Weinberg Equilibrium in at least one of the sampling localities. There were 69 different alleles identified with the effective allele size estimated at 3.635. Within sampling localities, allelic richness was 6.094 and 6.011 for Marsabit County and Kwale County populations respectively. The average number of private alleles for Marsabit County and Kwale County populations was 3.33 and 3.667 respectively. Shannon's Information index (I) of the populations was estimated at 1.502 and 1.531 respectively. Observed heterozygosities were 0.734 and 0.732 and expected heterozygosities (Nei's gene diversity) were 0.698 and 0.718 for the northern and coastal populations, respectively. Inbreeding coefficient ($F_{IS}$) values were -0.082 and -0.065. Among STRUTURE-defined sub-populations, allelic richness ranged from 4.097 in sub-population green to 5.958 in sub-population blue. Observed heterozygosity ranged from 0.667 in sub-population blue to 0.832 in sub-population green and expected heterozygosity ranged from 0.662 in sub-population blue forto 0.806 in sub-population green (Table 4). Average numbers of private alleles were 2.167, 2.833 and 1.167 for sub-populations red, green and blue respectively. Lastly, the average numbers of different alleles were and 5.33, 6.667 and4.167 for sub-populations red, green and blue respectively.

**Table 4:** Population diversity indices calculated from *T. vivax* microsatellite alleles data.

N-Sample size; Na-Average sample size; Ne-Effective sample size; Ho-Observed Heterozygosities; uHe-unbiased expected hetrozygosity; F-Fixation Index; I- Shannon's Information Index; Ar- Allelic richness

| **Pop** | **N** | **Na** | **Ne** | **Ho** | **uHe** | **F** | **I** | **Ar** |
|---|---|---|---|---|---|---|---|---|
| **Kwale** | 30 | 8.167 | 3.673 | 0.732 | 0.718 | -0.065 | 1.531 | 6.011 |
| **Marsabit** | 30 | 7.833 | 3.597 | 0.734 | 0.698 | -0.082 | 1.502 | 6.094 |
| **Overall** | 60 | 8 | 3.635 | 0.733 | 0.708 | -0.073 | 1.52 | 6.053 |
| **Red** | 18 | 5.333 | 3.11 | 0.763 | 0.652 | -0.211 | 1.253 | 3.872 |
| **Green** | 10 | 6.667 | 4.941 | 0.789 | 0.798 | -0.067 | 1.639 | 5.21 |
| **Blue** | 13 | 4.167 | 2.776 | 0.61 | 0.634 | -0.109 | 1.14 | 3.499 |
| **Overall** | 41 | 5.389 | 3.61 | 0.744 | 0.695 | -0.129 | 1.344 | 4.145 |

### 4.4.3 Population differentiation

Fixation index value of *T. vivax* isolates between the sampling localities based populations was 0.955 indicating moderate significant differentiation (P=0.005). Between STRUCTURE-defined clusters, $F_{ST}$ values were 0.98 between clusters a(red) and b(green), 0.172 between clusters a(red) and c(blue) and 0.621 between clusters c(blue) and b(green), and all clusters were significantly differentiated (P<0.05).

AMOVA results for *T. vivax* isolates revealed that 84% of the variation was due to differences in genotypes within individuals while differences between sub-populations accounted for 5% of the

variance. Differences between isolates in a population accounted for 11% of the total variance (Figure 14).

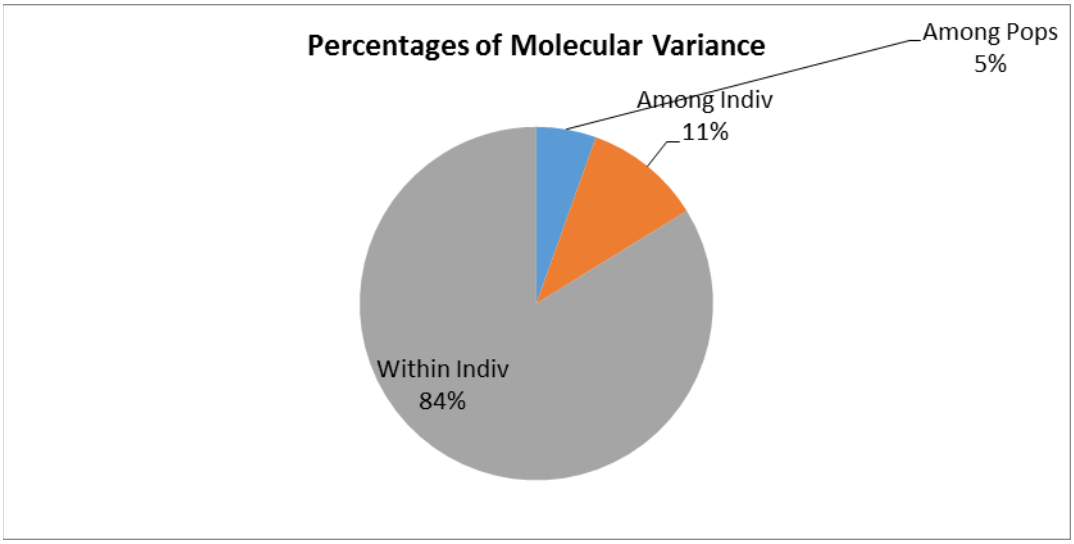



**Figure 14:** AMOVA analyses results on six microsatellite loci on T. vivax isolates from Marsabit County and Kwale County.

According to the partial Mantel test results, there was significant correlation between genetic distances and geographical distances while controlling for environmental conditions (r = 0.3846, p = 1e-04).

Emigration and immigration rates between the Kwale County and Marsabit County populations were 0.0238 and 0.0356 respectively. Asymmetric gene flow within both populations was evident due to the high self-distribution recorded. Asymmetric gene flow hinders adaptation by opposing natural selection while maintaining genetic diversity (Telschow *et al.*, 2006). The effective number of migrants ($N_em$) was equal in both populations at 2.275. Being that one<$N_em$<four, the two populations are genetically close but gene flow between them is limited likely due to the barrier caused by the geographical distance between them.

### 4.5 *Trypanosoma congolense* Population Genetics Parameters

#### 4.5.1 Population Structure

Bayesian clustering analysis suggested a K value of 3 as the most likely level of hierarchy of population structure (Figure 14). However, none of the isolates had a Q value (assignment value) greater than 0.4 (Appendix II) and therefore STRUCTURE based clusters were not analyzed further. However, it is important to note that all three *T. congolense* isolates from Marsabit County clustered together with Kwale County isolates.

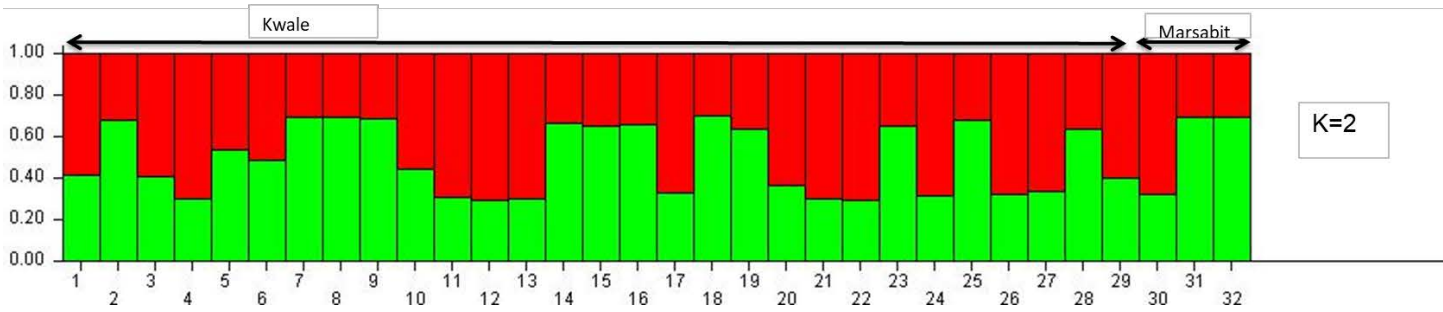



**Figure 15**: Bayesian clustering based on STRUCTURE results on thirty three *T. congolense* isolates from Marsabit and Kwale regions in Kenya based on five microsatellite loci. STRUCTURE harvester results based on the adhoc statistic indiated K=3 as the best K value. All isolates had Q valuse below 0.4 which may be an indiation of a hiigh level of genetic admiture within the population.

Multivariate clustering results were harmonious with those obtained by the Bayesian analysis method on the presence of three distinct genetic clusters. PCA results show Marsabit County isolates clustering among Kwale County isolates which is an indication of no genetic separation (Figure 16). However, unlike in Bayesian analysis, clusters in PCA were clearly differentiated. All three clusters were found in different multivariate spaces. Clusters orange and green were differentiated by PC axis 1 while PC axis 2 differentiated clusters orange and blue.

To further assess the level of population structure based on genetic distance matrix, an UPGMA dendrogram of similarity based on Cavalli-Sforza and Edwards (DC) pairwise genetic distances was constructed (Figure 17). The results confirmed Bayesian and multivariate analysis results on the absence of genetic separation of Marsabit County isolates from the Kwale County isolates. However, unlike the former methods, the dendrogram revealed the presence of two genetic clusters with sub-structuring in one of the clusters and two isolates (C10 and C1) sharing a multilocus genotype (MLG). The bootstrap values of these clusters are however weak (bootstrap=0.2) therefore inferences made from the dendrogram are rather limited.

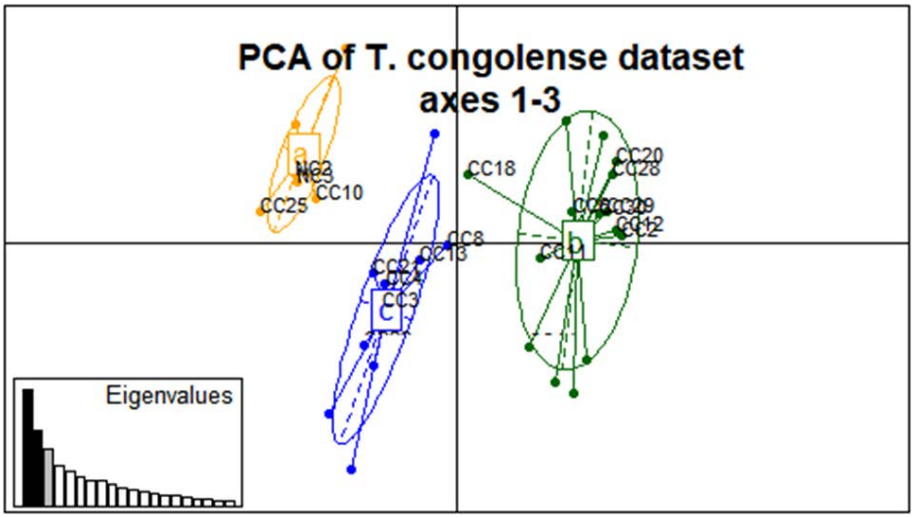



**Figure 16**: A Principal Component Analysis (PCA) showing spatial structuring of *T. congolense* isolates.

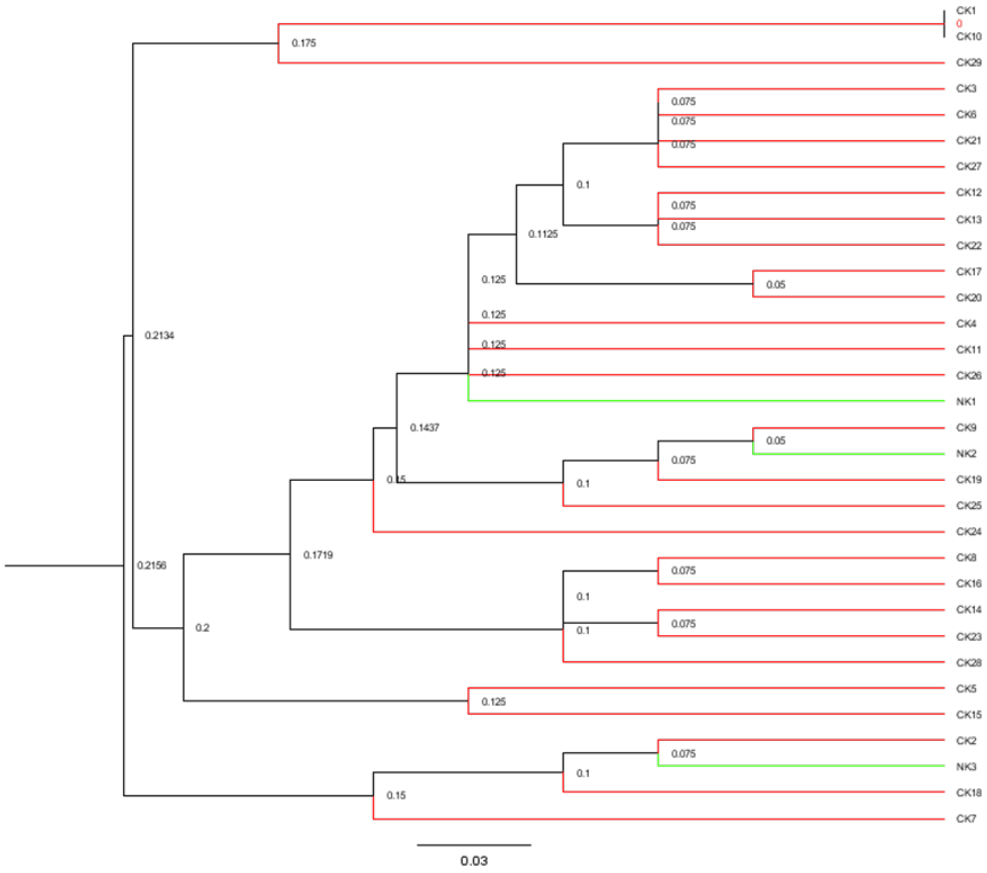



**Figure 17**: A UPGMA dendrogram based on Cavalli-Sforze distances showing no separation of *T. congolense* between Kwale County and Marsabit County *T. congolense* isolates.

Red branches -Kwale County isolates, green branches- Marsabit County isolates.

### 4.5.2 Genetic diversity

Hardy Weinberg proportions from allele frequency data showed deviation from Hardy Weinberg Equilibrium (HWE) in at least one population of all loci. Linkage disequilibrium analysis revealed that only one (TB8/11 and TB6/7) of ten loci combinations was in linkage disequilibrium.

Genetic diversity analysis indicated the presence of 42 alleles across the two populations. Allelic richness (with rarefaction) for the Marsabit and Kwale populations was 3.56 and 2.80 respectively thus indicating greater genetic diversity within the Kwale County population. Mean observed heterozygosities were 0.668 and 0.667 while expected heterozygosities were 0.710 and 0.653 for the Marsabit and Kwale County populations respectively. Inbreeding coefficient (Fis) was -0.195 and 0.009 respectively (Table 5).

**Table 5:** Population diversity indices calculated from *T. congolense* microsatellite alleles data.

N-Sample size; Na-Average sample size; Ne-Effective sample size; Ho-Observed Heterozygosities; uHe-unbiased expected hetrozygosity; F-Fixation Index; I- Shannon's Information Index; $A_R$*- Allelic Richness with rarefaction

| **Pop** | **N** | **Na** | **Ne** | **Ho** | **uHe** | **F** | **I** | **$A_R$*** |
|---|---|---|---|---|---|---|---|---|
| **Kwale** | 28.6 | 8.6 | 3.617 | 0.668 | 0.711 | 0.01 | 1.562 | 2.8 |
| **Marsabit** | 3 | 2.6 | 2.354 | 0.667 | 0.653 | -0.195 | 0.861 | 3.56 |
| **Overall** | 15.8 | 5.6 | 2.986 | 0.667 | 0.682 | -0.093 | 1.212 | 3.18 |

### 4.5.3 Genetic differentiation

We estimated fixation index ($F_{ST}$) between the Kwale and Marsabit populations at -0.028 with a p value of 0.713. Partial Mantel test results revealed a positive non-significant correlation between geographical and genetic distances while controlling for environmental conditions (r = 0.311, p = 0.3219).

Immigration and emigration rates from the Kwale and Marsabit population were 0.2612 and 0.033 respectively. Self-distribution rate within the Kwale and Marsabit populations was 0.9667 and 0.788.

# CHAPTER FIVE: DISCUSSION, CONCLUSSION AND RECOMMENDATIONS

## 5.1 Discussion

This study comprising of population genetic differentiation of *Trypanozoon* (*T. brucei T. evansi*), *T. congolense* and *T. vivax* from the two distinct localities, i.e., from tsetse free and tsetse infested areas has demonstrated close intra-species phylogenetic relationships, for example among *T. congolense* regardless of the difference in the mode of transmission, climate and vectors present. However, we have also seen some trypanosomes like *Trypanozoon* are distinct between the two sites demonstrating local adaptation and evolution. Individual trypanosome species were found to be distributed across the different ecological settings, spanning from wet Kwale County infested with tsetse flies to dry Marsabit County where there are no tsetse flies. The influence of transmission on certain individuals, locations, host infectious states, or parasite strains shows significant heterogeneity in most host-parasite systems. It was noted in this study that both sites have domestic animals and blood feeding insects, the only differences being the absence of camels in Kwale County, which are abundant in Marsabit County. All trypanosome species (*Trypanozoon, T. congolense* and *T. vivax*) were however encountered in both sites.

*Trypanozoon* isolates clustered into two sub-populations by Bayesian clustering in STRUCTURE. A majority of isolates in each sub-population were from one of the two sampling locations (that is one sub-population had all samples from Marsabit County and the other sub-population had a majority isolates from Kwale County) indicating genetic distinctness of isolates from each population. These results are consistent with both multivariate and hierarchical clustering, both of which grouped the isolates into two groups each dominated by isolates from each sampling location. Clustering of these isolates mainly by geographical location is due to the

presumed presence of different *Trypanozoon* species: *T. evansi* and *T. brucei* in the two geographical regions. *Trypanosoma evansi*, which is exclusively transmitted mechanically, is mainly found in tsetse free Northern Kenya (Marsabit County), while *T. brucei* is present in the Coastal region (Kwale County) where tsetse flies are abundant. This is consistent with a previous study that reported *T. brucei* to be the dominant *Trypanozoon* species in Shimba Hills, Kwale County (Coastal Kenya) (Kulohoma *et al*., 2020). Because of its diskinetoplastidy, *T. evansi* is incapable of completing cyclical development in the tsetse flies (Lai *et al.*, 2008). However, we cannot rule out its circulation in tsetse infested areas as tsetse flies do transmit *T. evansi* mechanically and the parasite can survive in various domestic and wild animals besides camels. Nonetheless, previous genomic analyses of *T. evansi* and *T. brucei* isolates have shown that the species are not always distinguishable even with DNA sequencing (Büscher *et al.*, 2019; Getahun *et al.*, 2020).

The next best fit of K (K =3) in STRUCTURE revealed a level of sub-structuring within isolates from Kwale County and genetic homogeneity among isolates from Marsabit County. In addition, greater genetic diversity among Kwale County *Trypanozoon* (*T. brucei*) evidenced by grater allelic richness, Shannon's Information Index (I), and average private alleles values (Table 3) compared to Marsabit County Trypanozoon (*T. evansi*). However, there were no Multi Locus Genotypes (MLGs) in the population and this casts doubt on the idea of absolute clonal reproduction (where the offspring are genetically identical to their parents) among *T. evansi* isolates as reported by (Meeûs, *et al.*, 2011).

Wright's fixation index ($F_{ST}$) between the populations of origin based populations was small yet statistically significant thus revealing moderately significant genetic differentiation between Marsabit County (*T. evansi*) and Kwale County (*T. brucei*) *Trypanozoon* isolates. Further, the

low gene flow rate ($N_em < 4$) is a clear indication that the two populations are not panmitic, thus there are obstacles to gene flow. However, the effective number of migrants is greater than one in both populations, which is more evidence that isolates from the two location of origin-based populations are genetically closely related with low genetic differentiation. Nonetheless, isolation by distance analysis revealed that structuring and genetic differentiation within the *Trypanozoon* isolates was strongly dependent on location of origin and AMOVA results showed that while the isolates are mainly structured based on location of origin, differences in genotypes within isolates was the greatest contributing factor to genetic differentiation and sub-structuring. This indicates that though they are very closely genetically related, the two species are fundamentally genetically distinct.

This may be evidence against the need for revision of the entire taxonomic unit (*Trypanozoon: T. brucei* and *T. evansi*) due to their close genetic relationships as proposed by (Büscher *et al.*, 2019). Classification of trypanosome DNA is mainly based on Ribosomal DNA Genetic markers (ITS-1 markers) (Njiru *et al.*, 2004, 2005, 2006). Telling apart *T. brucei* and *T. evansi* with these markers is impossible with gel electrophoresis and at times difficult after sequencing (Büscher *et al.*, 2019; Getahun *et al.*, 2020). A large number of microsatellite loci that amplify *Trypanozoon* genotypes have been identified (Balmer *et al.*, 2006; Salim *et al.*, 2011a; Sistrom *et al.*, 2013). Identification of private microsatellite marker bands (Gel electrophoresis bands unique in each taxon) within the *Trypanozoon* group may present a potential for application of microsatellite loci as species identification markers to remove the problem of precise species identification.

*Trypanosoma vivax* is widespread in both Africa and South America. Genetic analysis has demonstrated that the South American *T. vivax* is closely related to West African *T. vivax*, showing possible introduction of the species into South America from West Africa but far from

East African strains (Cortezi et al., 2006). Analysis of *T. vivax* microsatellite data from Brazil, Venezuela and Nigeria in West Africa corroborated prior studies that *T. vivax* isolates from South America originated from West Africa. However, isolates from Kenya in East Africa separated from isolates from both West Africa and South America by large genetic distances. However, *T. vivax* isolates from Brazil, isolated from asymptomatic cattle and cattle showing different pathologies were genetically homogeneous (Rodrigues et al., 2008). Results obtained from this and other studies, has revealed that the genetic diversity of South American *T. vivax* isolates is much lower than that of African strains (Desquesnes and Dia, 2004; Rodrigues *et al.*, 2008). This may likely due to the lack of genetic recombination, which occurs only in tsetse flies (Gibson and Stevens, 1999). Genetic homogeneity among South American *T. vivax* isolates is a likely explanation for how cattle populations there are able to temporarily control *T. vivax* infections (Rodrigues *et al*., 2008).

From this study, STRUCTURE and PCA results suggests that *T. vivax* isolates are clustered into three distinct genetic groups. The results are consistent with previous phylogenetic studies on *T. vivax* which indicated genetic heterogeneity within the species especially in East Africa (Fasogbon *et al*., 1990; Craig *et al.,* 2009). These results also suggest that *T. vivax* isolates within clusters were more similar in geographical origin which might be affected by local adaptation such as mode of transmission, vertebrate hosts and environmental factors. The clustering of a group of Marsabit County isolates separately shows an independent evolution due to differences in geographical factors, hosts and vectors which may be a driving force of genetic separation. However, the identification of two distinct clusters in Kwale County in the same ecosystem indicates that the existence of different *T. vivax* strains demonstrating genetic differentiation can result in the same habitat, which may be an indication that other factors, other than geographical

variation, can contribute to genetic heterogeneity. In addition, the possibility of strains specific to different wildlife host species cannot be ruled out.

Five out of six loci showed deviation from HWE which may be due to predominant clonal reproduction of *T. vivax* populations. Genetic diversity in the Marsabit County population was relatively lower compared to the Kwale County population. This was inferred from results of genetic diversity measures: average number of alleles, Shannon's Information Index, Nei's gene diversity ($H_E$) and allelic richness, all of which had lower values of the Marsabit County population. Among STRUCTURE defined clusters (Figure 7), sub-population red with isolates from Marsabit County had higher genetic diversity compared to the blue sub-population with the highest number of isolates from Kwale County which may be attributed to disparity in population size which leads to bias in allele frequencies. The green sub-population with isolates from both sampling locations however revealed the greatest genetic diversity. This results of higher genetic diversity among tsetse borne *T. vivax* agree with findings from previous studies of greater diversity among *T. vivax* from tsetse endemic areas (Rodrigues *et al.*, 2008). The greater observed heterozygosities, deviation from HWE and negative $F_{IS}$ among Kwale County isolates suggests that the population is clonal. These also results suggest that the *T. vivax* populations studied here are predominantly clonal and agree with previous studies that revealed clonal reproduction in tsetse borne and non-tsetse borne *T. vivax* (Duffy *et al.*, 2009) . However, previous studies on *T. vivax* isolates have revealed that meiosis associated genes in *T. brucei* (which has been shown to reproduce sexually) are highly conserved in *T. vivax* and therefore we cannot completely eliminate the possibility of genetic recombination among *T. vivax* strains (El-Sayed *et al.*, 2005). Being that these results suggests that *T. vivax* populations studied here are

clonal; the high number of unique genotypes in the populations may be due to amplification failure, null alleles or dropout alleles.

The Marsabit County isolates are significantly differentiated from the Kwale County isolates based on Wright's $F_{ST}$. Among STRUCTURE-defines clusters, the greatest differentiation observed was between the red sub-population, with all Marsabit County isolates and the green sub-population with a mixture of isolates from both populations. The least differentiation was between the red sub-population and the blue sub-population with a majority of isolates from the Kwale County population. This revealed that the green sub-population is significantly differentiated from the other sub-population. These results agree with AMOVA results that revealed that differences between isolates in a subpopulation accounted for more variance than did differences between sub-populations. In addition, isolation by distance results reveal further evidence that the structuring of the *T. vivax* isolates was dependent on geographical origin. However, the findings suggest moderate gene flow between the populations regardless of isolation by distance. This is an indication that the populations are genetically close with the limited gee flow likely due to the barrier caused by the geographical distance between them.

However, a group of isolates from Marsabit County cluster together with a group of isolates from Kwale County in STRUTURE and this sub-population(green) shows relatively greater significant differentiation from other clusters dominated by isolates from both sampling locations. Therefore, it may be unlikely that *T. vivax* isolates from tsetse free Matsabit County in the sub-population were introduced in Marsabit County from tsetse endemic areas by animal movement. These results therefore agree with the idea that heterogeneity among African *T. vivax* isolates is linked to biological transmission by tsetse flies. This is because *T. vivax* isolates from Kwale County that are tsetse borne are significantly differentiated from isolates from Marsabit

County that are not tsetse borne. These observations therefore disagree with sequence analysis results of the rDNA of *T. vivax* strains from tsetse free and tsetse infested areas of Ethiopia that indicated that genetic heterogeneity among the strains is not linked to their geographical origin (Fikru *et al*., 2016). This then means that grater genetic heterogeneity among African *T. vivax* strains compared to Latin American strains (Garcia et al., 2014) is associated to mode of transmission.

Bayesian, multivariate and hierarchical analysis classified *T. congolense* isolates from both sites into three genetic clusters without clear separation by ecology which agrees with previous findings of *T. congolense* classification into three sub-species in East Africa: *T. congolense* Savannah, *T. congolense* Kilifi and *T. congolense* Forest (Pereira *et al*., 2022). The microsatellite genetic analysis revealed no separation between *T. congolense* from tsetse free areas and tsetse endemic areas which is evidence of free genetic admixture between *T. congolense* populations or migration of *T. congolense* from one geographical location to another. All loci revealed deviation from HWE in at least one population. This may be attributed to the clonal nature of *T. congolense* strains (Tibayrenc and Ayala, 2019) However, 90% of loci combinations were in linkage equilibrium. This is an indication that loci used in the study were evenly distributed in the genome. However, linkage disequilibrium results among loci within Marsabit County isolates were inconclusive due to the small number of isolates in the population. The Kwale County isolates revealed greater diversity with rarefaction. This is expected due to biological transmission by tsetse flies in the region which gives way for genetic exchange (Gibson and Stevens, 1999).

The findings revealed no significant differentiation between the Kwale and Marsabit populations. The negative $F_{ST}$ value we speculate is a result of sampling bias within the population.

Differentiation between the isolates is not significant which means that there is no differentiation between the Kwale and Marsabit County isolates. Furthermore, isolation by distance analysis revealed that genetic structuring within the population was independent of origin of sample. The high gene flow rate between Marsabit County and Kwale County populations further revealed the absence of geographical barrier to genetic exchange between the two populations.

Trypanosomes undergo genetic adaptation events that allow them to utilize different energy sources in vectors and hosts (Ooi *et al.*, 2016; Szöőr et al., 2020). *Trypanosoma brucei* for example has mitochondrial genes that allows it to utilize α-ketoglutarate as source of energy in the tsetse midgut (Szöőr *et al*, 2020). *Trypanosoma evansi* lacks these genes and thus it has lost its ability for biological transmission. Therefore, genetic data would reflect intra-species differences in isolates with different biological vectors. Therefore, these results do not support the hypothesis of presence of an unknown biological vector for *T. congolense* isolates from tsetse free areas, first proposed by Gibson et al., 1983 because the three clustering analysis methods used failed to separate *T. congolense* isolates from Marsabit County with those from Kwale County. However, the small number of samples in the Marsabit County population challenges the accuracy of these results and requires more population genetics studies on *T. congolense* from tsetse free and tsetse infested areas with more samples.

*Trypanosoma congolense* isolates have however been isolated in livestock blood and biting flies from tsetse free areas of Northern Kenya for a long period of time (Gibson and Wilson, 1983; Getahun *et al.*, 2020). Yet, these results show no genetic separation between Marsabit County and Kwale County (tsetse borne) isolates which suggests that the parasite is probably introduced in these areas from livestock that travels to tsetse-inhabited areas for pasture and water. Also, these results may suggest presence of tsetse flies in Marsabit County but at very low hard to

detect densities that need detailed wide survey of tsetse flies. In addition, most Marsabit County isolates were sampled from camels and very few from cows and therefore the results could be biased towards isolates selected based on one host. Host selection has been shown to affect population structuring in *T. brucei* (Simo, 2014a; Simo, *et al.*, 2014b).

### 5.2 Conclusion

1. This study reveals novel insights into the intraspecific genetic connectivity among *T. vivax, Trypanozoon and T. congolense* isolates from tsetse endemic and tsetse free areas of Africa.

2. *Trypanozoon* microsatellite analysis results show clear separation of Marsabit County and Kwale County isolates. These results also reveal the need to explore application of microsatellite loci in *Trypanozoon* taxa identification.

3. In addition, the results revealed no genetic separation of *T. congolense* strains from tsetse free and tsetse endemic areas.

### 5.3 Recommendations

1. Since microsatellite loci have shown clear separation of Marsabit County *Trypanozoon (T. evannsi)* and Kwale County Trypanozoon (*T. brucei*), there is need to explore their potential in application as diagnostic markers for Trypanosomes. In future studies, optimizing microsatellite based markers to differentiate *T. brucei* and *T. evansi* needs to be investigated.
2. *Trypanosoma congolense* isolates from Marsabit County analyzed in this study were too few for conclusive results. There is therefore need to increase the sample size and area of coverage in order to investigate the genetic diversity and differentiation between *T. congolense* isolated from tsetse free areas and those isolated from tsetse endemic areas.

**FUNDING**

This research project was funded by BMZ/GIZ- GIZ project number and contract number 81219442 and MPI-icipe partner group to MNG.

**REFERENCES**


Abdul-Muneer, P. M. (2014). *Application of Microsatellite Markers in Conservation Genetics and Fisheries Management: Recent Advances in Population Structure Analysis and Conservation Strategies*. Genetics Research International, 2014, pp. 1–11. doi: 10.1155/2014/691759.

Adamack, G. & (2014). *Introduction to PopGenReport using PopGenReport Ver. 2.0.* Available at: http://www.tug.org/mactex/.

Adams, E. R., Hamilton , P. and Rodrigues A. (2010). *New Trypanosoma (Duttonella) vivax genotypes from tsetse flies in East Africa*. Parasitology, 137(4), pp. 641–650. doi: 10.1017/S0031182009991508.

African Union. (2008). *The PAN-AFRICAN Tsetse and Trypanosomiasis Eradication Campaign (PATTEC) Discusses Strategies and Shares Best Practices.* Available at: https://au.int/en/pressreleases/20181126/ (Accessed on: 22-10-2020)

Andrews, L. V (2013). *Chapter 21 Scoring Microsatellite Loci.* (April). doi: 10.1007/978-1-62703-389-3.

Aphasizhev, R. and Aphasizheva, I. (2011). *Mitochondrial RNA processing in trypanosomes*. Research in Microbiology, 162(7), pp. 655–663. doi: 10.1016/j.resmic.2011.04.015.

Auty, H., Torr, S. T., Michael. T., Jayaraman, S. and Morrison, L. J. (2015). *Cattle trypanosomosis: the diversity of trypanosomes and implications for disease epidemiology and control.* Revue scientifique et technique (International Office of Epizootics), 34(2), pp. 587–98. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26601459.

Balding, D. J. (2003). *Likelihood-based inference for genetic correlation coefficients.* Theoretical Population Biology, 63(3), pp. 221–230. doi: 10.1016/S0040-5809(03)00007-8.

Balmer, O., Christopher, P., Macleod, A. and Caccone, A. (2006). *Characterization of di-, tri- and tetranucleotide microsatellite markers with perfect repeats for Trypanosoma brucei and related species*. Molecular Ecology Notes, 6(2), pp. 508–510. doi: 10.1111/j.1471-8286.2006.01297.x.

Bargul, J. L., Jung, J.,McOdimba, F. A., Omogo, C. O., Adunga, V. O., Kruger, T., Masiga, D. K. and Engstler M. (2016). *Species-Specific Adaptations of Trypanosome Morphology and Motility to the Mammalian Host.* PLoS Pathogens, 12(2), pp. 1–30. doi: 10.1371/journal.ppat.1005448.

Bohonak, A. J. (2002). *IB*D (isolation by distance): *A program for analyses of isolation by distance.* Journal of Heredity, 93(2), pp. 153–154. doi: 10.1093/jhered/93.2.153.

Brun, R., Hecker, H. and Lun, Z. R. (1998). *Trypanosoma evansi and T. equiperdum:* Distribution, biology, treatment and phylogenetic relationship (a review). Veterinary Parasitology, 79(2), pp. 95-107. doi: 10.1016/S0304-4017(98)00146-0.

Beltrame-Botelho, I.T., Gaspar-Silva, D., Steindel, M., Dávila, A.M.R. and Grisard, E.C., 2005. Internal transcribed spacers (ITS) of Trypanosoma rangeli ribosomal DNA (rDNA): a useful marker for inter-specific differentiation. Infection, Genetics and Evolution, 5(1), pp.17-28.

Burki, F., Rodger, A. J., Brown, M. W. and Simpson, A. G. B. (2020). *The New Tree of Eukaryotes.* Trends in Ecology and Evolution, 35(1), pp. 43–55. doi: 10.1016/j.tree.2019.08.008.

Büscher, P. Gonzatti, I. M., Hébert, L., Inoue, N., Pascucci, I., Schnaufer, A., Suganuma4, Louis Touratier, K. and Reet N. V. (2019). *Equine trypanosomosis: Enigmas and diagnostic challenges.* Parasites and Vectors, 12(1), pp. 1–8. doi: 10.1186/s13071-019-3484-x.

Carnes, J. Anupama, A., Balmer, O., Jackson, A., Lewis, M., Brown, R., Cestari, I., Desquesnes, M., Gendrin, C., Hertz-Fowler, C., Imamura, H., Ivens, A., Kořený, L., Lai, D. H., MacLeod, A., McDermott, S. M., Merritt, C., Monnerat, S., Moon, W., Myler, P., Phan, I., Ramasamy, G., Sivam, D., Lun, Z. R., Lukeš, J., Stuart, K., Schnaufer, A., (2015). *Genome and Phylogenetic Analyses of Trypanosoma evansi Reveal Extensive Similarity to T. brucei and Multiple Independent Origins for Dyskinetoplasty*. PLoS Neglected Tropical Diseases, 9*(*1). doi: 10.1371/journal.pntd.0003404.

Claes, F., Bu, P., Touratier, L. and Goddeeris, B. M. (2005). *Trypanosoma equiperdum : master of disguise or historical mistake ?*. Trends in Parasitology, 21(7). doi: 10.1016/j.pt.2005.05.010.

Cortez, A. P., Venturai R. M., Rodrigues, A. C., Batista J. S., Paiva, F., EZ4, N., Machado R. Z., Gibson, W. C. and Teixeira, M. M. G. (2006). *The taxonomic and phylogenetic relationships of Trypanosoma vivax from South America and Africa*. Parasitology, pp. 159–169. doi: 10.1017/S0031182006000254.

Dario, M. A., Varella, L. C., Costa, L., Moratelli, R., Nascimento, M., Luiz, Y., Leite, R., Roque, L. R., Llewellyn, S., Cristina, S. and Jansen, M. (2017). *High Trypanosoma spp . diversity is maintained ´ rito Santo by bats and triatomines in Espı state , Brazil',* Plos One, pp. 1–22.Desquesnes, M., Dargantes, A., Lai, D., Lun, Z., Holzmuller, P. and Jittapalapong, S. (2013). *Trypanosoma evansi and surra: A review and perspectives*

*on origin, history, distribution, taxonomy, morphology, hosts, and pathogenic effects*. BioMed Research International, 2013. doi: 10.1155/2013/194176.

Desquesnes, M. and Dia, M. L. (2003). *Mechanical transmission of Trypanosoma congolense in cattle by the African tabanid Atylotus agrestis*. Experimental Parasitology, 105(3–4), pp. 226–231. doi: 10.1016/j.exppara.2003.12.014.

Desquesnes, M. and Dia, M. L. (2004). *Mechanical transmission of Trypanosoma vivax in cattle by the African tabanid Atylotus fuscipes.* Veterinary Parasitology, 119(1), pp. 9–19. doi: 10.1016/j.vetpar.2003.10.015.

Diniz-Filho, J. A. F., Soares, T. N., Lima, J. S., Dobrovolski, R., Landeiro, V. L., Telles, M., Rangel, T and Bini, L. M. (2013). *Mantel test in population genetics.* Genetics and Molecular Biology, 36(4), pp. 475–485. doi: 10.1590/S1415-47572013000400002.

Duffy, C. W., Morrison, L. J., Black, A., Pinchbeck, G. L., Christley, R. M., Schoenefeld, A., Tait, A., Turner, C. M. R. and Macleod, A. (2009). *Trypanosoma vivax displays a clonal population structure.* International Journal for Parasitology, 39(13), pp. 1475–1483. doi: 10.1016/j.ijpara.2009.05.012.

Earl, D. A. and vonHoldt, B. M. (2012). *STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method*. Conservation Genetics Resources, 4(2), pp. 359–361. doi: 10.1007/s12686-011-9548-7.

Echodu, R., Sistrom, M., Bateta, R., Murilla, G. and Okedi, L. (2015). *Genetic Diversity and Population Structure of Trypanosoma brucei in Uganda : Implications for the Epidemiology of Sleeping Sickness and Nagana*. PLoS Neglected Tropical Diseases, pp. 1–19. doi: 10.5061/dryad.m7q4c.

Echodu, R., Sistrom, M., Bateta, R., Murilla, G., Okedi, L., Aksoy, S., Enyioha, C., Enyaru, J.,

Opiyo, E.,Gibson, W. and Caccone, A. (2015). *Genetic Diversity and Population Structure of Trypanosoma brucei in Uganda: Implications for the Epidemiology of Sleeping Sickness and Nagana*', PLoS Neglected Tropical Diseases, 9(2), pp. 1–18. doi: 10.1371/journal.pntd.0003353.

El-Sayed, N. M. et al. (2005) 'Comparative genomics of trypanosomatid parasitic protozoa', *Science*, 309(5733), pp. 404–409. doi: 10.1126/science.1112181.

Evanno, G., Regnaut, S. and Goudet, J. (2005). *Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study*. Molecular Ecology, 14(8), pp. 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x.

Fasogbon, A. I., Knowles, G. and Gardiner, P. R. (1990). *A comparison of the isoenzymes of Trypanosoma (Duttonella) vivax isolates from East and West Africa. International Journal for Parasitology, 20(3), pp. 389–394. doi: 10.1016/0020-7519(*90)90156-H.

Fikru, R.., Hagos, A., Rogé, S., Reyna-Bello, A., Gonzatti, M. I., Merga, B., Goddeeris, B. M., and Büscher, P. (2014). *A proline racemase based PCR for identification of Trypanosoma vivax in cattle blood.* PLoS ONE, 9(1), pp. 1–7. doi: 10.1371/journal.pone.0084819.

Fikru, R., Matetovici, I., Rogé, S., Merga, B., Goddeeris, B., Büscher, P. and Van Reet, N. (2016). *Ribosomal DNA analysis of tsetse and non-tsetse transmitted Ethiopian Trypanosoma vivax strains in view of improved molecular diagnosis*. Veterinary Parasitology, 220, pp. 15–22. doi: 10.1016/j.vetpar.2016.02.013.

Garcia, H. A., Rodrigues, A. C., Rodrigues, C. M.F. Bengaly, Z., Minervino, A. H.H., Riet-Correa, F., Machado, R. Z., Paiva, F., Batista, J. S., Neves, L., Hamilton, P. B. and Teixeira, M M.G. (2014). *Microsatellite analysis supports clonal propagation and*

*reduced divergence of Trypanosoma vivax from asymptomatic to fatally infected livestock in South America compared to West Africa*', Parasites and Vectors, 7(1), pp. 1–13. doi:10.1186/1756-3305-7-210.

Getahun, M., Villinger, J., Bargul, J., Orone, A., Ngiela, J., Ahuya, P., Muema, J., Saini, R., Torto, B. and Masiga, D. (2020). *Molecular characterization of pathogenic African trypanosomes in biting flies and camels in surra-endemic areas outside the tsetse fly belt in Kenya*', BioRxiv, pp. 1–37. doi: 10.1101/2020.06.18.156869.

Gibson, W.C., Wilson, A.J. and Moloo, S.K., 1983. Characterisation of Trypanosoma (Trypanozoon) evansi from camels in Kenya using isoenzyme electrophoresis. Research in veterinary science, 34(1), pp.114-118.

Gibson, W. and Stevens, J. (1999). Genetic exchange in the Trypanosomatidae, *Advances in Parasitology*. doi: 10.1016/s0065-308x(08)60240-7.

Goudet, J. (2003). FSTAT (version 2.9.4), a program (for Windows 95 and above) to estimate and test population genetics parameters', pp. 1–54. Available at: http://www.unil.ch/izea/softwares/fstat.html.

Greenbaum, G., Templeton, A. R., Zarmi, Y. and Bar-David, S. (2014) '*Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift*', PLoS ONE, 9(12), pp. 1–23. doi: 10.1371/journal.pone.0115203.

Guichoux, E., Lepoittevin, C. and Revardel, E. (2011). *Current trends in microsatellite genotyping*. Molecular Ecology Resources, 11(4), pp. 591–611. doi: 10.1111/j.1755-0998.2011.03014.x.

Hampl, V., Hug, L., Leigh, J. W., Dacks, J. B., Lang, B. F., Simpson, A. G. B. and Roger, A. J.

(2009). *Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic 'supergroups*'. Pnas, 106(10).

Heller, K. A. (2005). *Bayesian Hierarchical Clustering*. 22nd International Confer- ence on Machine Learning, Bonn, Germany, 2005.

Hutchinson, R. and Stevens, J.R., 2018. Barcoding in trypanosomes. Parasitology, 145(5), pp.563-573.

Jackson, A. P., Goyard, S., Xia, D., Foth, B. J., Sanders, M., Wastling, J. M., Minoprio, P. and Berriman, M. (2015). *Global gene expression profiling through the complete life cycle of Trypanosoma vivax*. PLoS Neglected Tropical Diseases, 9(8), pp. 1–29. doi: 10.1371/journal.pntd.0003975.

Jiang, S., Luo, M. X., Gao, R. H., Zhang, W., Yang, Y. Z., Li, Y. J. and Liao, P. C. (2019). *Isolation-by-environment as a driver of genetic differentiation among populations of the only broad-leaved evergreen shrub Ammopiptanthus mongolicus in Asian temperate deserts.* Scientific Reports, 9(1), pp. 1–14. doi: 10.1038/s41598-019-48472

Jombart, T. (2008). *Adegenet: A R package for the multivariate analysis of genetic markers*. Bioinformatics, 24(11), pp. 1403–1405. doi: 10.1093/bioinformatics/btn129.

Jombart, T., Devillard, S. and Balloux, F. (2010). *Discriminant analysis of principal components : a new method for the analysis of genetically structured populations*. BioMed entral Genetics, doi: 10.1186/1471-2156-11-94.

Jonathan K. Pritchard, M. S. and P. D. (2000). *Inference of Population Structure Using Multilocus Genotype Data*. Genetics, 155(1), pp. 945–959. doi: 10.1007/s10681-008-9788-0. Jost, L. (2008). *GST and its relatives do not measure differentiation*. Molecular Ecology, 17(18), pp. 4015–4026. doi: 10.1111/j.1365-294X.2008.03887.x.

Jones, T. and Alberto, M. (2001) *Trypanosoma vivax-out of Africa.* Trends in Parasitology, 17(2), pp 99-101

Kalinowski, S. T. (2004). *Counting alleles with rarefaction: Private alleles and hierarchical sampling designs*. Conservation Genetics, 5(4), pp. 539–543. doi: 10.1023/B:COGE.0000041021.91777.1a.

Kalinowski, S. T. (2005). *HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness*. Molecular Ecology Notes, 5(1), pp. 187–189. doi: 10.1111/j.1471-8286.2004.00845.x.

Kamidi, C. M., Saarman, N. P., Dion, K., Mireji, P. O., Ouma, C., Murilla, G., Aksoy, S., Schnaufer, A. and Caccone, A. (2017). *Multiple evolutionary origins of Trypanosoma evansi in Kenya.* Plos Negleted Tropical Diseases, pp. 1–21. doi: 10.5061/dryad.8g678.

Kimenyi, N. N., Kimenyi, K. M., Amugune, N. O. and Getahun, M. N. (2021) *Genetic connectivity of trypanosomes between tsetse-infested and tsetse-free areas of Kenya*. Parasitology, pp.1-13.

Kulohoma, B. W., Wamwenje, S. A. O., Wangwe, I. I., Masila, N., Mirieri, C. K. & Wambua, L. (2020). *Prevalence of Trypanosomes Associated with Drug Resistance in Shimba Hills, Kwale County, Kenya.* BMC Research Notes, pp. 13:234. https://link.springer.com/article/10.1186/s13104-020-05077-3

Lai, D. H., Hashimi, H., Lun, Z.-R., Ayala, F. J. and Lukes, J. (2008). *Adaptations of Trypanosoma brucei to gradual loss of kinetoplast DNA: Trypanosoma equiperdum and Trypanosoma evansi are petite mutants of T. bruce*i. Proceedings of the National Academy of Sciences, 105(6), pp. 1999–2004. doi: 10.1073/pnas.0711799105.

Li, M. M., Li, B. L., Jiang, S. X. Zhao, Y. W., Xu, X. L. and Wu, J. X. (2019). *Microsatellite-*

*based analysis of genetic structure and gene flow of Mythimna separata (Walker) (Lepidoptera: Noctuidae) in China*. Ecology and Evolution, 9(23), pp. 13426–13437. doi: 10.1002/ece3.5799.

Whitlock, M. C (1999). *Indirect measures of gene flow and migration: FST ?1/(4Nmˆ1)*. Journal of Revenue and Pricing Management, 82(4), pp. 117–125. doi: 10.1057/s41272-017-0102-4.

Manel, S.. Schwartz, M. K., Luikart, G. and Taberlet, P. (2003). *Landscape genetics: Combining landscape ecology and population genetics*. Trends in Ecology and Evolution, 18(4), pp. 189–197. doi: 10.1016/S0169-5347(03)00008-9.

Maslov, D. A., Podlipaev, S. A. and Luke, J. (2001). *Phylogeny of the Kinetoplastida : Taxonomic Problems and Insights into the Evolution of Parasitism*. Mem Inst Oswaldo Cruz, Rio de Janeiro, v96(April), pp. 397–402.

Matthews, K. R. (2005). *The developmental cell biology of Trypanosoma brucei*. Journal of Cell Science, 118(2), pp. 283–290. doi: 10.1242/jcs.01649.

MeirMans, P. G. (2006). *Using the Amova Framework To Estimate a Standardized Genetic Differentiation Measure*. Evolution, 60(11), p. 2399. doi: 10.1554/05-631.1.

Merilä, J. and Crnokrak, P. (2001). *Comparison of genetic differentiation at marker loci and quantitative traits.* Journal of Evolutionary Biology, 14(6), pp. 892–903. doi: 10.1046/j.1420-9101.2001.00348.x.

Mihok, S., Maramba, O., Munyoki, E. and Kagoiya, J. (1995). *Mechanical transmission of Trypanosoma spp. by African Stomoxyinae (Diptera: Muscidae).*Tropical medicine and parasitology : official organ of Deutsche Tropenmedizinische Gesellschaft and of Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ*)*, 46(2), pp. 103–5.

Available at: http://www.ncbi.nlm.nih.gov/pubmed/8525279.

Morrison, L. J., Tweedie, A., Black, A., Pinchbeck, G. L., Christley, R. M. and Schoenefeld, A. (2009), *Discovery of mating in the major African livestock pathogen Trypanosoma congolense.* PLoS ONE, 4(5), pp. 1–8. doi: 10.1371/journal.pone.0005564.

Morrison, L. J., Vezza, L., Rowan, T. and Hope, J. C. (2016). *Animal African Trypanosomiasis: Time to Increase Focus on Clinically Relevant Parasite and Host Species*. Trends in Parasitology, 32(8), pp. 599–607. doi: 10.1016/j.pt.2016.04.012.

Mossaad, E., Salim, B., Suganuma, K., Musinguzi, P., Hassan, M. A., Elamin, E. A., Mohammed, G. E., Bakhiet, A. O., Xuan, X., Satti, R. A. and Inoue, N. (2017). *Trypanosoma vivax is the second leading cause of camel trypanosomosis in Sudan after Trypanosoma evansi*. Parasites and Vectors, 10(1), pp. 1–10. doi: 0.1186/s13071-017-2117-5.

Musaya, J., Chisi, J., Senga, E., Nambala, P., Maganga, E., Matovu, E. and Enyaru, J., 2017. Polymerase chain reaction identification of Trypanosoma brucei rhodesiense in wild tsetse flies from Nkhotakota Wildlife Reserve, Malawi. Malawi Medical Journal, 29(1), pp.11-15.

Nathaniel Mantel (1967). *The Detection of Disease Clustering and a Generalized Regression Approach.* Cancer Research, 27(1), pp. 209–220. doi: 10.5694/j.1326-5377.1952.tb83832.x.

Njiru, Z. K.. Constantine, C. C., Ndung'u, J. M., Robertson, I., Okaye, S., Thompson, R. C. A. and Reid, S. A. (2004). *Detection of Trypanosoma evansi in camels using PCR and CATT/T. evansi tests in Kenya*. Veterinary Parasitology, 124(3–4), pp. 187–199. doi:

10.1016/j.vetpar.2004.06.029.

Njiru, Z. K., Constantine, C. C., Guya, S., Crowther, J., Kiragu, J. M., Thompson, R. C.A. and Dávila, A. M.R. (2005). *The use of ITS1 rDNA PCR in detecting pathogenic African trypanosomes*. Parasitology Research, 95(3), pp. 186–192. doi: 10.1007/s00436-004-1267-5.

Njiru, Z. K., Constantine, C. C., Masiga, D. K.,Reid, S. A., Thompson, R. C. A. and Gibson, W. C. (2006). *Characterization of Trypanosoma evansi type* B §'. Infection, Genetis and Evolution, 6, pp. 292–300. doi: 10.1016/j.meegid.2005.08.002.

Oksanen, A. J. *et al.* (2020) *'Package " vegan "'*. https://cran.r-project.org, https://github.com/vegandevs/vegan NeedsCompilation

Ooi, C. P., Schuster, S., Cren-Travaillé, C., Bertiaux, E., Cosson, A., Goyard, S., Perrot, S. and Rotureau, Brice (2016) '', *Frontiers in Cellular and Infection Microbiology*, 6(SEP), pp. 1–16. doi: 10.3389/fcimb.2016.00115.

Osório, A. L. A. R., Madruga, C. R., Desquesnes, M., Soares, C. O., Ribeiro, L. R. R. and Sylvio C. G. (2008). *Trypanosoma (Duttonella) vivax: Its biology, epidemiology, pathogenesis, and introduction in the New World - A review*. Memorias do Instituto Oswaldo Cruz, 103(1), pp. 1–13. doi: 10.1590/S0074-02762008000100001.

MeirMans, P. G. (2012). *Amova-Based Clustering of Population Genetic Data*. Journal of Heredity, 103(5), pp. 744–750. doi: 10.1093/jhered/ess047.

Patterson, N., Price, A. L. and Reich, D. (2006). *Population structure and eigenanalysis*. PLoS Genetics, 2(12), pp. 2074–2093. doi: 10.1371/journal.pgen.0020190.

Peakall, R. and Smouse, P. E. (2006). *GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research*. Molecular Ecology Notes, 6(1), pp. 288–

295. doi: 10.1111/j.1471-8286.2005.01155.x.

Pereira, S.S., Mathenge, K., Masiga, D. and Jackson, A., 2022. Transcriptomic Profiling of Trypanosoma Congolense Mouthpart Parasites from Naturally Infected Flies.

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á. and Lareu, M. V. (2013). *An overview of STRUCTURE: Applications, parameter settings, and supporting software*. Frontiers in Genetics, 4(MAY), pp. 1–13. doi: 10.3389/fgene.2013.00098.

Prugnolle, F. and De Meeûs, T. (2008). *The impact of clonality on parasite population genetic structure*. Parasite, 15(3), pp. 455–457. doi: 10.1051/parasite/2008153p455.

Radwanska, M., Vereecke, N., Deleeuw, V. and Pinto, J. (2018). Salivarian Trypanosomosis : A Review of Parasites Involved , Their Global Distribution and Their Interaction With the Innate and Adaptive Mammalian Host Immune System', Frontiers in Immunology, 9(October), pp. 1–20. doi: 10.3389/fimmu.2018.02253.

Rodrigues, A. C., Neves, L., Garcia, H. A., Viola, L. B., Marcili, A., Da Silva, F. M., Sigauque, I., Batista, J. S., Paiva, F. and Teixeira, M. M.G. (2008). *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(11), pp. 1317–1328. doi: 10.1017/S0031182008004848.

Rodrigues, C. M., Garcia, H. A., Rodrigues, A. C., Costa-Martins, A. G., Pereira, C. L., Pereira, D. L., Bengaly, Z., Neves, L., Camargo, E. P., Hamilton, P. B. and Teixeira, M. M. (2017). *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*. Parasites and Vectors, 10(1), pp. 1–16. doi: 10.1186/s13071-017-2241-2.

Rousset, F., Lopez, L. and Belkhir, K. (2020). R package: genepop, p. 16. Available at: http://kimura.univ-montp2.fr/~rousset/Genepop.htm.

Salim, B., Bakheit, M. A., Kamau, J., Nakamura, I. and Sugimoto, C. (2011b). *Molecular epidemiology of camel trypanosomiasis based on ITS1 rDNA and RoTat 1.2 VSG gene in the Sudan*. Parasites and Vectors, 4(1), pp. 2–6. doi: 10.1186/1756-3305-4-31.

Schnaufer, A. (2010). *Evolution of dyskinetoplastic trypanosomes: how, and how often*?. Trends in Parasitology, 26(12), pp. 557–558. doi: 10.1016/j.pt.2010.08.001.

Schwabl, P., Llewellyn, M. S., Landguth, E. L., Andersson, B., Kitron, U., Costales, J. A., Ocaña, S. and Grijalva, Ma. J. (2017a). *Prediction and Prevention of Parasitic Diseases Using a Landscape Genomics Framework.* Trends in Parasitology, 33(4), pp. 264–275. doi: 10.1016/j.pt.2016.10.008.

Schwabl, P., Llewellyn, M. S., Landguth, E. L., Andersson, B., Kitron, U., Costales, J. A., Ocaña, S. and Grijalva, M. J. (2017b). *Prediction and Prevention of Parasitic Diseases Using a Landscape Genomics Framework.* Trends in Parasitology, 33(4), pp. 264–275. doi: 10.1016/j.pt.2016.10.008.

Senan, S., Kizhakayil, D., Sikumar, B.and Sheena, T. E. (2014). *Methods for Development of Microsatellite Markers: An Overview*. Notulae Scientia Biologicae, 6(1), pp. 1–13. doi: 10.15835/nsb619199.

Sexton, J. P., Hangartner, S. B. and Hoffmann, A. A. (2014). *Genetic isolation by environment or distance: Which pattern of gene flow is most common?*. Evolution, 68(1), pp. 1–15. doi: 10.1111/evo.12258.

Shaw, A. P. M., Cecchi, G., Wint, G. R.W., Mattioli, R. C. and Robinson, T. P. (2014). *Mapping the economic benefits to livestock keepers from intervening against bovine*

*trypanosomosis in Eastern Africa*. Preventive Veterinary Medicine, 113(2), pp. 197–210. doi: 10.1016/j.prevetmed.2013.10.024.

Sheriff, O. and Alemayehu, K. (2018). *Genetic diversity studies using microsatellite markers and their contribution in supporting sustainable sheep breeding programs : A review Genetic diversity studies using microsatellite markers and their contribution in supporting sustainable sheep breedi*. Cogent Food & Agriculture, 1(1). doi: 10.1080/23311932.2018.1459062.

Simo, G., Sobgwi, P. F., Njitchouang, G. R., Njiokou, F., Kuiate, J. R. and Cuny, G. A. T. (2013). *Identification and genetic characterization of Trypanosoma congolense in domestic animals of Fontem in the South-West region of Cameroon*. Infection, Genetics and Evolution, 18, pp. 66–73. doi: 10.1016/j.meegid.2013.04.019.

Simo, G., Fogue, P. S., Melachio, T. T. T. Njiokou, F., Kuiate, J. R. and Asonganyi, T. (2014) *Population genetics of forest type of Trypanosoma congolense circulating in Glossina palpalis palpalis of Fontem in the South-West region of Cameroon*. Parasites and Vectors, 7(1), pp. 1–10. doi: 10.1186/1756-3305-7-385.

Simo, G., Njitchouang, G. R., Melachio, T. T. T., Njiokou, F., Cuny, G. and Tazoacha, A. (2014). *Population genetics of Trypanosoma brucei circulating in Glossina palpalis palpalis and domestic animals of the Fontem sleeping sickness focus of Cameroon*. Parasites and Vectors, 7(1), pp. 1–11. doi: 10.1186/1756-3305-7-156.

Sistrom, M. Echodu, R., Hyseni, C., Enyaru, J., SerapAksoy, C. A (2013). *Taking advantage of genomic data to develop reliable microsaatellite loci in Trypanosoma brucei*. Moleular Eology Resources.33(0), pp. 1–9.

Solymos, P., Cori, A. and Calboli, F. (2020) 'Package " adegenet " R topics documented :'

Storfer, A.., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L.and Waits, L. P. (2007). *Putting the "landscape" in landscape genetics.* Heredity, 98(3), pp. 128–142. doi: 10.1038/sj.hdy.6800917.

Sundqvist, L., Keenan, K., Zackrisson, M., Prodöhl, P. and Kleinhans, D. (2016). *Directional genetic differentiation and relative migration*. Ecology and Evolution, 6(11), pp. 3461–3475. doi: 10.1002/ece3.2096.

Szöőr, B., Silvester, E. and Matthews, K. R. (2020). *A Leap Into the Unknown – Early Events in African Trypanosome Transmission.* Trends in Parasitology, 36(3), pp. 266–278. doi: 10.1016/j.pt.2019.12.011.

Takeet, M. I., Fagbemi, B. O., Peters, S. O., DeDonato, M. Y., Abdul M. Wheto, M. and Imumorin, I. G. (2017). *Genetic diversity among Trypanosoma vivax strains detected in naturally infected cattle in Nigeria based on ITS1 of rDNA and diagnostic antigen gene sequences.* Journal of Parasitic Diseases, 41(2), pp. 433–441. doi: 10.1007/s12639-016-0822-1.

Telschow, A., Engelstädter, J., Yamamura, N., Hammerstein, P. and Hurst, G. D.D. (2006). *Asymmetric gene flow and constraints on adaptation caused by sex ratio distorters*. Journal of Evolutionary Biology, 19(3), pp. 869–878. doi: 10.1111/j.1420-9101.2005.01049.x.

Tibayrenc, M. and Ayala, F.J., 2019. How clonal is Trypanosoma congolense? A necessary clarification of the predominant clonal evolution model. Acta tropica, 190, pp.28-29.

Toro, M. A. and Caballero, A. (2005). *Characterization and conservation of genetic diversity in subdivided populations*. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1459), pp. 1367–1378. doi: 10.1098/rstb.2005.1680.

Toro, M. A., Fernández, J. and Caballero, A. (2009). *Molecular characterization of breeds and its use in conservation*. Livestock Science, 120(3), pp. 174–195. doi: 10.1016/j.livsci.2008.07.003.

*measure it - G ST, D, neither or both?*. Molecular Ecology, 23(17), pp. 4216–4225. doi: 10.1111/mec.12856.

Wells, E. A. (1972). *The importance of mechanical transmission in the epidemiology of nagana: A review*. Tropical Animal Health and Production, 4(2), pp. 74–89. doi: 10.1007/BF02359739.

Wen, Y. Z., Lun, Z. R., Zhu, X. Q., Hide, G. and Lai, D. H. (2016). *Further evidence from SSCP and ITS DNA sequencing support Trypanosoma evansi and Trypanosoma equiperdum as subspecies or even strains of Trypanosoma brucei*. Infection, Genetics and Evolution, 41, pp. 56–62. doi: 10.1016/j.meegid.2016.03.022.

Wilson, G. A. and Rannala, B. (2003). *Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes*. Genetics society of America, 1191(March), pp. 1177–1191.

Wright, S. (1990). *Evolution in mendelian populations.* Bulletin of Mathematical Biology, 52(1–2), pp. 241–295. doi: 10.1007/BF02459575.

Yaro, M., Munyard, K. A., Stear, M. J. and Groth, D. M. (2016). *Combatting African Animal Trypanosomiasis (AAT) in livestock: The potential role of trypanotolerance*. Veterinary Parasitology, 225, pp. 43–52. doi: 10.1016/j.vetpar.2016.05.003.

## APPENDICES

**Appendix I: *Trypanozoon* samples Q values obtained from clustering analysis with STRUCTURE.**

| Sample | Population | a | b | c |
|---|---|---|---|---|
| CB1 | Kwale | 0.234 | 0.686 | 0.08 |
| CB10 | Kwale | 0.017 | 0.041 | 0.942 |
| CB11 | Kwale | 0.01 | 0.012 | 0.979 |
| CB12 | Kwale | 0.009 | 0.028 | 0.964 |
| CB13 | Kwale | 0.276 | 0.673 | 0.051 |
| CB14 | Kwale | 0.526 | 0.013 | 0.46 |
| CB15 | Kwale | 0.011 | 0.007 | 0.982 |
| CB16 | Kwale | 0.011 | 0.016 | 0.972 |
| CB17 | Kwale | 0.012 | 0.012 | 0.975 |
| CB18 | Kwale | 0.01 | 0.016 | 0.974 |
| CB19 | Kwale | 0.009 | 0.97 | 0.02 |
| CB2 | Kwale | 0.103 | 0.007 | 0.89 |
| CB20 | Kwale | 0.037 | 0.034 | 0.93 |
| CB21 | Kwale | 0.013 | 0.803 | 0.183 |
| CB22 | Kwale | 0.025 | 0.013 | 0.962 |
| CB3 | Kwale | 0.01 | 0.3 | 0.69 |
| CB4 | Kwale | 0.011 | 0.302 | 0.687 |
| CB5 | Kwale | 0.015 | 0.314 | 0.67 |
| CB6 | Kwale | 0.017 | 0.026 | 0.957 |
| CB7 | Kwale | 0.038 | 0.144 | 0.818 |
| CB8 | Kwale | 0.014 | 0.044 | 0.942 |
| CB9 | Kwale | 0.009 | 0.008 | 0.983 |
| NB1 | Marsabit | 0.984 | 0.008 | 0.007 |
| NB10 | Marsabit | 0.982 | 0.009 | 0.008 |
| NB11 | Marsabit | 0.955 | 0.018 | 0.027 |
| NB12 | Marsabit | 0.592 | 0.342 | 0.066 |

| | | | | |
|---|---|---|---|---|
| NB13 | Marsabit | 0.976 | 0.017 | 0.007 |
| NB14 | Marsabit | 0.985 | 0.006 | 0.008 |
| NB15 | Marsabit | 0.983 | 0.01 | 0.007 |
| NB16 | Marsabit | 0.975 | 0.012 | 0.013 |
| NB17 | Marsabit | 0.984 | 0.007 | 0.009 |
| NB18 | Marsabit | 0.989 | 0.006 | 0.05 |
| NB19 | Marsabit | 0.858 | 0.028 | 0.114 |
| NB2 | Marsabit | 0.911 | 0.044 | 0.045 |
| NB20 | Marsabit | 0.977 | 0.01 | 0.013 |
| NB21 | Marsabit | 0.948 | 0.025 | 0.027 |
| NB22 | Marsabit | 0.987 | 0.005 | 0.008 |
| NB23 | Marsabit | 0.978 | 0.013 | 0.009 |
| NB24 | Marsabit | 0.885 | 0.065 | 0.05 |
| NB25 | Marsabit | 0.97 | 0.009 | 0.021 |
| NB26 | Marsabit | 0.983 | 0.005 | 0.012 |
| NB27 | Marsabit | 0.983 | 0.01 | 0.006 |
| NB28 | Marsabit | 0.961 | 0.011 | 0.028 |
| NB29 | Marsabit | 0.978 | 0.01 | 0.012 |
| NB3 | Marsabit | 0.986 | 0.005 | 0.008 |
| NB30 | Marsabit | 0.982 | 0.007 | 0.0111 |
| NB4 | Marsabit | 0.983 | 0.011 | 0.006 |
| NB5 | Marsabit | 0.972 | 0.016 | 0.012 |
| NB6 | Marsabit | 0.99 | 0.006 | 0.005 |
| NB7 | Marsabit | 0.939 | 0.039 | 0.022 |
| NB8 | Marsabit | 0.855 | 0.061 | 0.084 |
| NB9 | Marsabit | 0.981 | 0.013 | 0.006 |

**Appendix II: *Trypanosoma vivax* samples Q values obtained from clustering analysis with STRUCTURE.**

| Sample ID | Population | a | b | c |
|---|---|---|---|---|
| CV1 | Kwale | 0.133 | 0.858 | 0.008 |
| CV10 | Kwale | 0 | 0.039 | 0.961 |
| CV11 | Kwale | 0 | 0.003 | 0.997 |
| CV12 | Kwale | 0 | 0.033 | 0.967 |
| CV13 | Kwale | 0 | 0.004 | 0.996 |
| CV14 | Kwale | 0.001 | 0.353 | 0.646 |
| CV15 | Kwale | 0.002 | 0.577 | 0.422 |
| CV16 | Kwale | 0 | 0.061 | 0.939 |
| CV17 | Kwale | 0 | 0.004 | 0.996 |
| CV18 | Kwale | 0 | 0.003 | 0.997 |
| CV19 | Kwale | 0 | 0.005 | 0.994 |
| CV2 | Kwale | 0.015 | 0.954 | 0.031 |
| CV20 | Kwale | 0.002 | 0.614 | 0.384 |
| CV21 | Kwale | 0.002 | 0.884 | 0.114 |
| CV22 | Kwale | 0.007 | 0.207 | 0.786 |
| CV23 | Kwale | 0 | 0.009 | 0.991 |
| CV24 | Kwale | 0.003 | 0.594 | 0.404 |
| CV25 | Kwale | 0.028 | 0.676 | 0.296 |
| CV26 | Kwale | 0.009 | 0.512 | 0.479 |
| CV27 | Kwale | 0 | 0.05 | 0.95 |
| CV28 | Kwale | 0 | 0.104 | 0.896 |
| CV29 | Kwale | 0 | 0.06 | 0.939 |
| CV3 | Kwale | 0.012 | 0.959 | 0.029 |
| CV30 | Kwale | 0 | 0.009 | 0.991 |
| CV4 | Kwale | 0.028 | 0.887 | 0.085 |
| CV5 | Kwale | 0.001 | 0.758 | 0.24 |
| CV6 | Kwale | 0 | 0.969 | 0.03 |
| CV7 | Kwale | 0.001 | 0.664 | 0.354 |
| CV8 | Kwale | 0.026 | 0.507 | 0.467 |
| CV9 | Kwale | 0.011 | 0.411 | 0.577 |
| NV1 | Marsabit | 0.263 | 0.172 | 0.565 |
| NV10 | Marsabit | 0.998 | 0.002 | 0 |
| NV11 | Marsabit | 0.914 | 0.059 | 0.027 |
| NV12 | Marsabit | 0.358 | 0.492 | 0.15 |
| NV13 | Marsabit | 0.996 | 0.004 | 0 |
| NV14 | Marsabit | 0.957 | 0.042 | 0.001 |

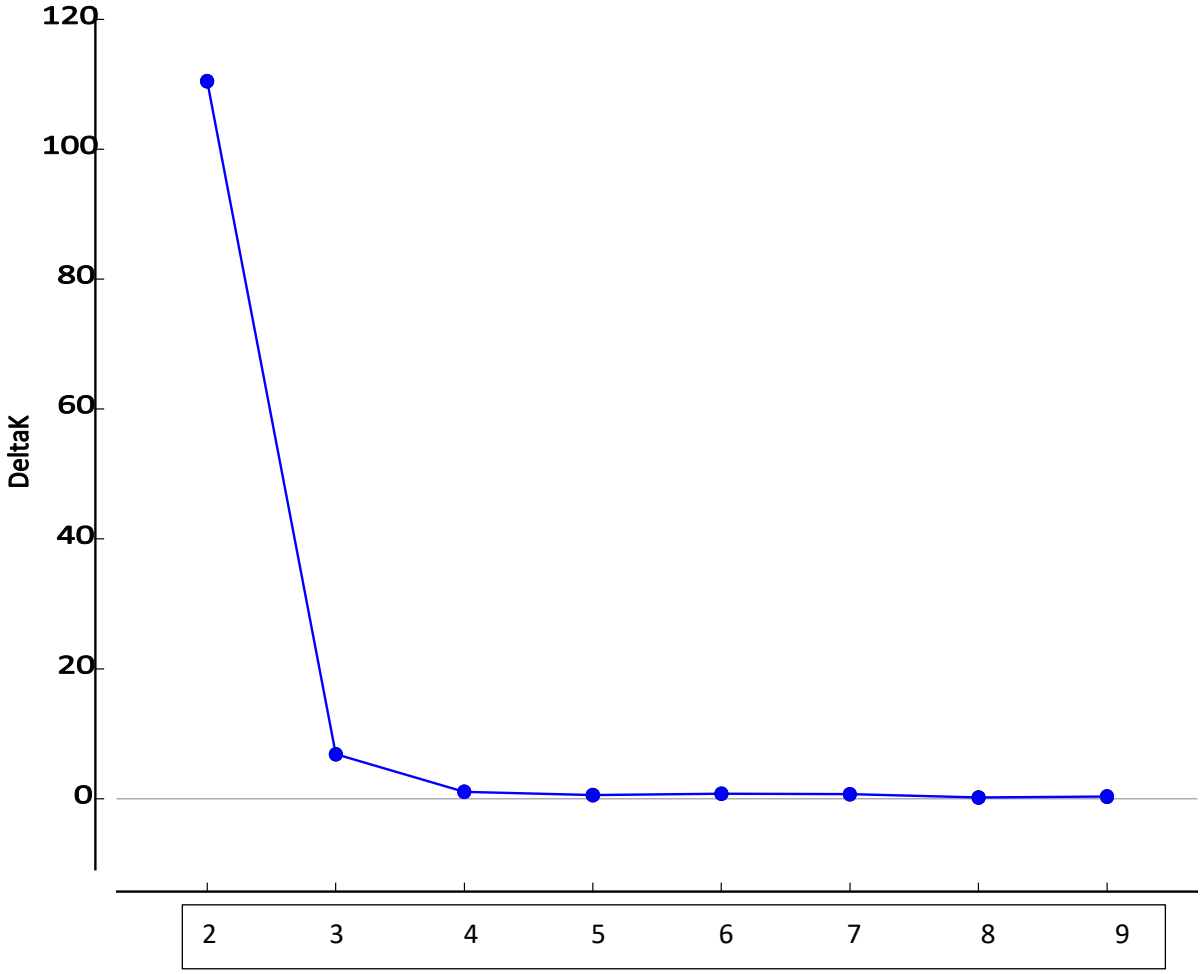
| | | | | |
|---|---|---|---|---|
| NV15 | Marsabit | 0.98 | 0.02 | 0 |
| NV16 | Marsabit | 0.977 | 0.023 | 0 |
| NV17 | Marsabit | 0.998 | 0.002 | 0 |
| NV18 | Marsabit | 0.998 | 0.012 | 0 |
| NV19 | Marsabit | 0.821 | 0.178 | 0.001 |
| NV2 | Marsabit | 0.852 | 0.111 | 0.038 |
| NV20 | Marsabit | 0.995 | 0.005 | 0 |
| NV21 | Marsabit | 0.999 | 0.001 | 0 |
| NV22 | Marsabit | 0.979 | 0.021 | 0 |
| NV23 | Marsabit | 0.999 | 0.001 | 0 |
| NV24 | Marsabit | 0.91 | 0.089 | 0.01 |
| NV25 | Marsabit | 0.397 | 0.58 | 0.023 |
| NV26 | Marsabit | 0.985 | 0.015 | 0 |
| NV27 | Marsabit | 0.241 | 0.612 | 0.148 |
| NV28 | Marsabit | 0.976 | 0.024 | 0 |
| NV29 | Marsabit | 0.551 | 0.404 | 0.044 |
| NV3 | Marsabit | 0.247 | 0.612 | 0.141 |
| NV30 | Marsabit | 0.887 | 0.111 | 0.001 |
| NV4 | Marsabit | 0.01 | 0.96 | 0.03 |
| NV5 | Marsabit | 0.006 | 0.949 | 0.044 |
| NV6 | Marsabit | 0.12 | 0.194 | 0.686 |
| NV7 | Marsabit | 0.063 | 0.162 | 0.774 |
| NV8 | Marsabit | 0.03 | 0.981 | 0.015 |
| NV9 | Marsabit | 0.089 | 0.899 | 0.013 |

**Appendix III: *Trypanosoma congolense* samples Q values obtained from clustering analysis with STRUCTURE**

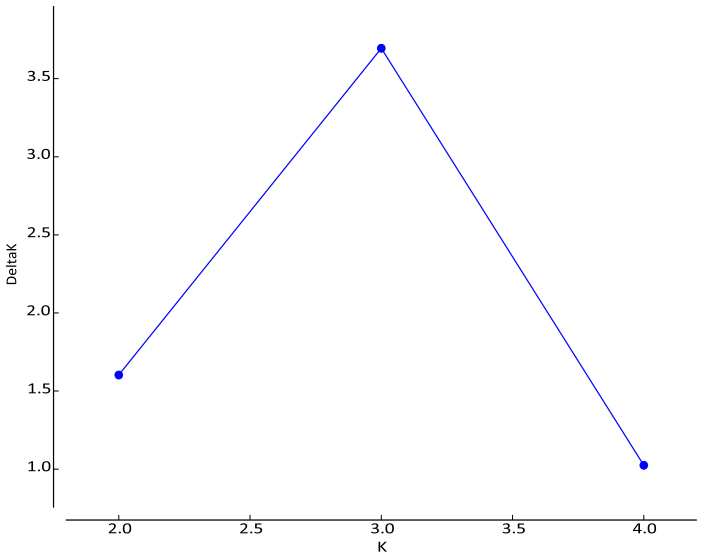
| SampleID | Population | a | b |
|---|---|---|---|
| CC1 | Kwale | 0.262 | 0.738 |
| CC10 | Kwale | 0.877 | 0.123 |
| CC11 | Kwale | 0.333 | 0.667 |
| CC12 | Kwale | 0.064 | 0.936 |
| CC13 | Kwale | 0.508 | 0.492 |
| CC14 | Kwale | 0.504 | 0.496 |
| CC15 | Kwale | 0.917 | 0.083 |
| CC16 | Kwale | 0.931 | 0.069 |
| CC17 | Kwale | 0.9 | 0.1 |
| CC18 | Kwale | 0.321 | 0.679 |

| | | | |
|---|---|---|---|
| CC19 | Kwale | 0.093 | 0.097 |
| CC2 | Kwale | 0.061 | 0.939 |
| CC20 | Kwale | 0.082 | 0.918 |
| CC21 | Kwale | 0.83 | 0.17 |
| CC22 | Kwale | 0.856 | 0.144 |
| CC23 | Kwale | 0.814 | 0.186 |
| CC24 | Kwale | 0.202 | 0.798 |
| CC25 | Kwale | 0.938 | 0.062 |
| CC26 | Kwale | 0.786 | 0.214 |
| CC27 | Kwale | 0.271 | 0.729 |
| CC28 | Kwale | 0.077 | 0.923 |
| CC29 | Kwale | 0.064 | 0.936 |
| CC3 | Kwale | 0.816 | 0.184 |
| CC30 | Kwale | 0.113 | 0.887 |
| CC4 | Kwale | 0.902 | 0.098 |
| CC5 | Kwale | 0.15 | 0.85 |
| CC6 | Kwale | 0.135 | 0.865 |
| CC7 | Kwale | 0.794 | 0.206 |
| CC8 | Kwale | 0.278 | 0.722 |
| NC1 | Marsabit | 0.105 | 0.895 |
| NC2 | Marsabit | 0.928 | 0.072 |
| NC3 | Marsabit | 0.928 | 0.072 |

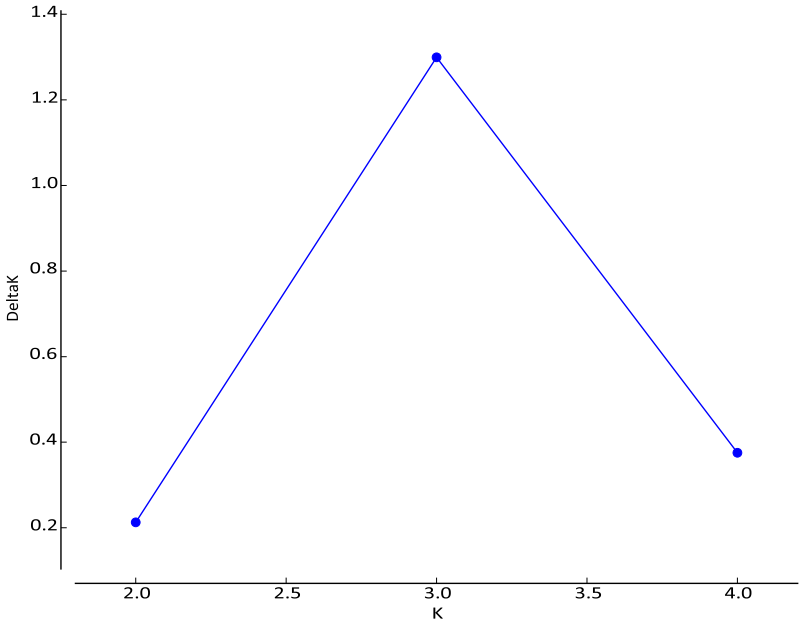
**Appendix IV: DeltaK image (mean(|L''(K)|) / sd(L(K)) obtained from STRUTUTE Harvester software showing the best fit for K for *Trypanozoon* isolates at K=2 and the next best fit at K=3**



**Appendix V: DeltaK image (mean(|L''(K)|) / sd(L(K)) obtained from STRUTUTE Harvester software showing the best fit for K for T. vivax isolates at K=3**




**Appendix VI: DeltaK image (mean(|L''(K)|) / sd(L(K for K for *T. congolense* isolates at K=3**



**Appendix VII: Migration rate values for *Trypanozoon*, *T. congolense* and *T. vivax* between their Marsabit County and Kwale County populations generated by BA3 software with their corresponding P values in brackets**

| | | Marsabit County | Kwale County |
|---|---|---|---|
| *Trypanozoon* | Marsabit County | 0.99 | 0.01 |
| | Kwale County | 0.04 | 0.96 |
| *T. vivax* | Marsabit County | 0.96 | 0.36 |
| | Kwale County | 0.02 | 0.97 |
| *T. congolense* | Marsabit County | 0.97 | 0.26 |
| | Kwale County | 0.03 | 0.74 |

**Appendix VIII: *Trypanozoon* samples microsatellite loci fragment sizes**

| Name | TB11/13 | | TB6/7 | | TB10/5 | | TB11/29 | | TB8/1 | | TB1/8 | | TB5/2 | | Tryp 62 | | Tryp 67 | | TB3/3 | | TB7/12 | | TB2/19 | | TB9/6 | | TB8/11 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CB11 | 125 | 157 | 108 | 114 | 76 | 112 | 162 | 166 | 132 | 137 | 95 | 101 | 83 | 105 | 161 | 161 | 169 | 172 | 0 | 0 | 185 | 187 | 90 | 101 | 133 | 143 | 91 | 117 |
| CB110 | 129 | 165 | 130 | 134 | 76 | 76 | 0 | 0 | 130 | 134 | 95 | 97 | 83 | 83 | 171 | 174 | 166 | 169 | 99 | 99 | 185 | 187 | 90 | 97 | 135 | 139 | 0 | 0 |
| CB111 | 157 | 163 | 118 | 134 | 76 | 82 | 136 | 140 | 132 | 142 | 95 | 99 | 79 | 85 | 171 | 171 | 166 | 169 | 79 | 99 | 185 | 187 | 90 | 99 | 135 | 141 | 89 | 89 |
| CB112 | 157 | 165 | 114 | 134 | 86 | 90 | 144 | 162 | 124 | 132 | 95 | 99 | 85 | 101 | 169 | 171 | 169 | 172 | 101 | 101 | 185 | 187 | 83 | 83 | 135 | 155 | 0 | 0 |
| CB113 | 147 | 157 | 108 | 118 | 78 | 100 | 150 | 152 | 124 | 150 | 97 | 105 | 85 | 87 | 161 | 163 | 169 | 172 | 67 | 79 | 181 | 185 | 79 | 81 | 125 | 131 | 0 | 0 |
| CB114 | 133 | 165 | 118 | 134 | 80 | 86 | 148 | 148 | 122 | 130 | 95 | 97 | 79 | 101 | 165 | 169 | 172 | 172 | 79 | 79 | 181 | 181 | 90 | 97 | 125 | 131 | 0 | 0 |
| CB115 | 127 | 165 | 110 | 134 | 82 | 86 | 152 | 152 | 124 | 174 | 95 | 95 | 79 | 101 | 169 | 169 | 172 | 175 | 79 | 79 | 182 | 185 | 93 | 99 | 127 | 127 | 0 | 0 |
| CB116 | 131 | 143 | 134 | 136 | 78 | 82 | 152 | 162 | 124 | 140 | 95 | 99 | 79 | 101 | 171 | 174 | 0 | 0 | 75 | 99 | 185 | 187 | 0 | 0 | 127 | 127 | 0 | 0 |
| CB117 | 0 | 0 | 114 | 134 | 78 | 82 | 0 | 0 | 130 | 132 | 95 | 99 | 105 | 105 | 171 | 171 | 169 | 172 | 79 | 79 | 185 | 187 | 93 | 97 | 127 | 127 | 0 | 0 |
| CB118 | 127 | 165 | 114 | 134 | 76 | 82 | 140 | 154 | 144 | 148 | 97 | 97 | 79 | 101 | 171 | 174 | 163 | 166 | 0 | 0 | 185 | 189 | 93 | 93 | 127 | 127 | 0 | 0 |
| CB119 | 125 | 147 | 106 | 112 | 76 | 88 | 126 | 152 | 142 | 142 | 105 | 107 | 83 | 87 | 169 | 172 | 157 | 163 | 0 | 0 | 185 | 189 | 0 | 0 | 125 | 133 | 0 | 0 |
| CB12 | 133 | 165 | 130 | 134 | 76 | 76 | 144 | 152 | 130 | 137 | 95 | 99 | 79 | 101 | 167 | 171 | 172 | 175 | 79 | 99 | 185 | 187 | 95 | 97 | 129 | 139 | 87 | 87 |
| CB1 | 13 | 16 | 12 | 13 | 82 | 86 | 14 | 15 | 12 | 14 | 92 | 95 | 85 | 89 | 16 | 16 | 16 | 16 | 0 | 0 | 18 | 18 | 9 | 90 | 12 | 13 | 89 | 89 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 7 | 5 | 2 | 4 | | | 0 | 2 | 8 | 0 | | | | | 9 | 9 | 6 | 9 | | | 5 | 7 | 0 | | 5 | 1 | | |
| CB121 | 141 | 143 | 106 | 114 | 82 | 86 | 128 | 128 | 122 | 142 | 107 | 111 | 105 | 105 | 165 | 167 | 157 | 163 | 0 | 0 | 181 | 185 | 90 | 101 | 127 | 133 | 0 | 0 |
| CB122 | 133 | 165 | 114 | 134 | 86 | 94 | 144 | 162 | 122 | 124 | 97 | 99 | 79 | 101 | 167 | 169 | 166 | 169 | 0 | 0 | 181 | 185 | 93 | 101 | 129 | 135 | 89 | 89 |
| CB13 | 143 | 159 | 106 | 108 | 86 | 112 | 132 | 136 | 142 | 178 | 95 | 99 | 79 | 101 | 167 | 167 | 154 | 172 | 67 | 79 | 181 | 185 | 95 | 99 | 125 | 127 | 0 | 0 |
| CB14 | 143 | 165 | 110 | 132 | 82 | 112 | 152 | 162 | 122 | 126 | 99 | 103 | 83 | 89 | 165 | 167 | 172 | 175 | 67 | 67 | 185 | 189 | 97 | 99 | 127 | 139 | 0 | 0 |
| CB15 | 149 | 151 | 108 | 132 | 80 | 82 | 142 | 152 | 122 | 134 | 99 | 107 | 83 | 89 | 165 | 167 | 172 | 175 | 79 | 79 | 185 | 185 | 99 | 101 | 127 | 139 | 0 | 0 |
| CB16 | 157 | 165 | 130 | 130 | 76 | 94 | 140 | 142 | 124 | 134 | 95 | 99 | 79 | 101 | 163 | 165 | 172 | 175 | 79 | 79 | 185 | 185 | 93 | 101 | 125 | 139 | 0 | 0 |
| CB17 | 153 | 165 | 118 | 132 | 78 | 82 | 152 | 154 | 142 | 178 | 99 | 99 | 79 | 101 | 167 | 167 | 154 | 169 | 67 | 79 | 182 | 182 | 90 | 101 | 135 | 139 | 0 | 0 |
| CB18 | 127 | 165 | 112 | 134 | 78 | 91 | 182 | 182 | 178 | 180 | 99 | 101 | 79 | 101 | 167 | 171 | 166 | 169 | 67 | 79 | 187 | 189 | 95 | 97 | 135 | 139 | 0 | 0 |
| CB19 | 129 | 165 | 114 | 134 | 0 | 0 | 0 | 0 | 130 | 142 | 95 | 99 | 79 | 101 | 169 | 169 | 169 | 172 | 79 | 101 | 185 | 185 | 97 | 99 | 135 | 153 | 0 | 0 |
| NB11 | 129 | 133 | 116 | 122 | 90 | 94 | 128 | 164 | 132 | 137 | 97 | 99 | 79 | 99 | 0 | 0 | 163 | 166 | 99 | 99 | 0 | 0 | 90 | 97 | 129 | 129 | 91 | 120 |
| NB110 | 127 | 129 | 114 | 116 | 78 | 78 | 148 | 152 | 132 | 137 | 95 | 97 | 103 | 105 | 0 | 0 | 172 | 175 | 79 | 99 | 0 | 0 | 79 | 83 | 131 | 137 | 87 | 91 |
| NB111 | 129 | 133 | 110 | 114 | 76 | 80 | 148 | 150 | 132 | 137 | 97 | 99 | 79 | 101 | 0 | 0 | 172 | 175 | 99 | 99 | 0 | 0 | 83 | 85 | 125 | 133 | 87 | 87 |
| NB112 | 131 | 147 | 112 | 116 | 76 | 80 | 130 | 152 | 132 | 134 | 92 | 97 | 83 | 95 | 171 | 174 | 0 | 0 | 99 | 99 | 0 | 0 | 83 | 85 | 129 | 129 | 111 | 111 |
| NB113 | 127 | 131 | 114 | 118 | 103 | 105 | 146 | 152 | 132 | 134 | 97 | 101 | 99 | 103 | 153 | 153 | 172 | 175 | 99 | 99 | 181 | 181 | 79 | 81 | 127 | 139 | 0 | 0 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NB114 | 129 | 133 | 118 | 122 | 76 | 88 | 148 | 152 | 130 | 137 | 95 | 99 | 91 | 99 | 0 | 0 | 154 | 169 | 79 | 99 | 185 | 185 | 97 | 101 | 133 | 139 | 91 | 120 |
| NB115 | 127 | 135 | 112 | 114 | 80 | 90 | 148 | 152 | 132 | 137 | 97 | 101 | 99 | 103 | 0 | 0 | 169 | 172 | 79 | 99 | 181 | 181 | 95 | 97 | 133 | 139 | 108 | 120 |
| NB116 | 127 | 137 | 116 | 118 | 74 | 103 | 144 | 146 | 140 | 154 | 95 | 99 | 99 | 99 | 0 | 0 | 166 | 169 | 0 | 0 | 181 | 182 | 90 | 99 | 133 | 139 | 111 | 120 |
| NB117 | 131 | 135 | 114 | 116 | 74 | 105 | 148 | 152 | 137 | 137 | 99 | 101 | 99 | 99 | 0 | 0 | 154 | 166 | 79 | 79 | 181 | 185 | 99 | 101 | 127 | 139 | 91 | 120 |
| NB118 | 127 | 131 | 116 | 118 | 74 | 105 | 148 | 152 | 132 | 137 | 95 | 97 | 91 | 99 | 0 | 0 | 169 | 172 | 75 | 79 | 181 | 182 | 90 | 101 | 133 | 139 | 91 | 120 |
| NB119 | 127 | 131 | 114 | 118 | 74 | 78 | 148 | 150 | 132 | 137 | 95 | 99 | 79 | 99 | 0 | 0 | 172 | 175 | 105 | 115 | 181 | 189 | 93 | 97 | 125 | 139 | 89 | 89 |
| NB12 | 127 | 135 | 114 | 118 | 88 | 94 | 128 | 162 | 128 | 132 | 92 | 97 | 99 | 103 | 0 | 0 | 163 | 166 | 79 | 99 | 0 | 0 | 90 | 93 | 129 | 135 | 91 | 120 |
| NB120 | 127 | 129 | 112 | 122 | 78 | 82 | 144 | 148 | 130 | 137 | 92 | 99 | 79 | 99 | 0 | 0 | 172 | 175 | 67 | 67 | 181 | 181 | 97 | 101 | 125 | 151 | 87 | 91 |
| NB121 | 129 | 147 | 116 | 122 | 78 | 80 | 136 | 148 | 130 | 154 | 92 | 95 | 101 | 105 | 165 | 165 | 172 | 175 | 67 | 67 | 181 | 185 | 93 | 97 | 133 | 139 | 108 | 120 |
| NB122 | 133 | 133 | 116 | 122 | 91 | 94 | 150 | 152 | 132 | 137 | 92 | 97 | 79 | 99 | 165 | 169 | 172 | 175 | 79 | 79 | 185 | 185 | 97 | 101 | 129 | 151 | 120 | 120 |
| NB123 | 131 | 131 | 114 | 122 | 76 | 84 | 166 | 170 | 132 | 137 | 92 | 97 | 91 | 99 | 165 | 165 | 154 | 169 | 79 | 79 | 185 | 187 | 87 | 97 | 133 | 155 | 120 | 120 |
| NB124 | 123 | 153 | 116 | 122 | 76 | 80 | 166 | 168 | 128 | 144 | 97 | 99 | 97 | 103 | 161 | 165 | 169 | 169 | 79 | 79 | 182 | 185 | 93 | 93 | 127 | 133 | 91 | 120 |
| NB125 | 129 | 133 | 112 | 118 | 74 | 90 | 124 | 130 | 128 | 130 | 97 | 99 | 79 | 99 | 161 | 165 | 169 | 172 | 79 | 99 | 182 | 185 | 90 | 99 | 131 | 139 | 89 | 120 |
| NB126 | 127 | 131 | 116 | 122 | 76 | 78 | 136 | 148 | 130 | 137 | 95 | 99 | 79 | 99 | 161 | 165 | 169 | 172 | 79 | 99 | 185 | 187 | 90 | 99 | 133 | 139 | 87 | 87 |
| NB1 | 12 | 14 | 11 | 12 | 74 | 90 | 15 | 16 | 13 | 16 | 97 | 99 | 79 | 99 | 16 | 16 | 16 | 16 | 67 | 67 | 18 | 18 | 9 | 10 | 12 | 15 | 87 | 87 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 27 | 9 | 7 | 6 | 2 | | | 2 | 6 | 2 | 0 | | | | | 1 | 3 | 9 | 9 | | | 1 | 7 | 0 | 1 | 5 | 1 | | |
| NB128 | 133 | 133 | 116 | 122 | 76 | 78 | 144 | 152 | 132 | 137 | 92 | 95 | 79 | 99 | 0 | 0 | 172 | 175 | 67 | 101 | 0 | 0 | 0 | 0 | 127 | 151 | 87 | 111 |
| NB129 | 129 | 133 | 116 | 122 | 76 | 76 | 166 | 166 | 132 | 140 | 95 | 99 | 97 | 103 | 167 | 167 | 154 | 169 | 79 | 79 | 181 | 182 | 87 | 99 | 133 | 139 | 87 | 87 |
| NB13 | 131 | 133 | 114 | 122 | 74 | 105 | 148 | 164 | 132 | 137 | 95 | 99 | 79 | 99 | 0 | 0 | 169 | 172 | 99 | 99 | 0 | 0 | 90 | 97 | 127 | 133 | 91 | 120 |
| NB130 | 127 | 135 | 110 | 118 | 74 | 90 | 164 | 164 | 132 | 137 | 95 | 99 | 99 | 99 | 167 | 169 | 169 | 172 | 79 | 79 | 181 | 185 | 90 | 99 | 125 | 131 | 87 | 120 |
| NB14 | 127 | 133 | 112 | 122 | 76 | 91 | 152 | 164 | 132 | 137 | 92 | 97 | 79 | 99 | 0 | 0 | 169 | 172 | 69 | 69 | 0 | 0 | 97 | 99 | 133 | 137 | 91 | 120 |
| NB15 | 153 | 153 | 116 | 122 | 78 | 78 | 146 | 152 | 132 | 174 | 95 | 99 | 99 | 99 | 0 | 0 | 172 | 175 | 67 | 67 | 181 | 191 | 77 | 79 | 129 | 135 | 91 | 120 |
| NB16 | 133 | 147 | 116 | 122 | 91 | 94 | 148 | 152 | 132 | 132 | 95 | 97 | 79 | 99 | 0 | 0 | 172 | 175 | 67 | 67 | 182 | 182 | 77 | 79 | 129 | 131 | 91 | 120 |
| NB17 | 151 | 157 | 116 | 122 | 86 | 88 | 148 | 152 | 130 | 146 | 95 | 97 | 79 | 99 | 0 | 0 | 172 | 175 | 79 | 99 | 181 | 181 | 79 | 101 | 131 | 137 | 111 | 120 |
| NB18 | 129 | 147 | 118 | 122 | 76 | 82 | 148 | 152 | 126 | 142 | 97 | 99 | 91 | 99 | 0 | 0 | 169 | 172 | 79 | 99 | 0 | 0 | 79 | 79 | 131 | 139 | 89 | 93 |
| NB19 | 127 | 135 | 116 | 118 | 74 | 78 | 146 | 150 | 134 | 160 | 97 | 105 | 79 | 97 | 0 | 0 | 169 | 172 | 79 | 99 | 181 | 181 | 90 | 101 | 133 | 139 | 91 | 120 |

**Appendix IX: *Trypanosoma vivax* microsatellite loci fragment sizes**

| ind | Population | TB8/11 | | Tryp 54 | | Tryp 67 | | Tryp 55 | | TB1/8 | | TB11/29 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CV1 | CV | 89 | 119 | 0 | 0 | 165 | 167 | 214 | 246 | 95 | 99 | 134 |
| CV10 | CV | 89 | 125 | 160 | 164 | 161 | 165 | 212 | 214 | 95 | 99 | 138 |
| CV11 | CV | 89 | 103 | 160 | 164 | 163 | 167 | 214 | 214 | 99 | 103 | 138 |
| CV12 | CV | 89 | 89 | 164 | 168 | 161 | 165 | 212 | 212 | 97 | 103 | 124 |
| CV13 | CV | 89 | 89 | 164 | 168 | 161 | 163 | 214 | 214 | 95 | 99 | 138 |
| CV14 | CV | 89 | 121 | 164 | 168 | 0 | 0 | 214 | 216 | 97 | 101 | 140 |
| CV15 | CV | 85 | 89 | 164 | 168 | 159 | 159 | 214 | 216 | 95 | 97 | 140 |
| CV16 | CV | 89 | 125 | 164 | 168 | 161 | 161 | 214 | 214 | 97 | 99 | 140 |
| CV17 | CV | 89 | 89 | 160 | 164 | 161 | 161 | 212 | 212 | 97 | 99 | 138 |
| CV18 | CV | 89 | 103 | 160 | 164 | 161 | 161 | 212 | 212 | 97 | 103 | 138 |
| CV19 | CV | 89 | 89 | 160 | 164 | 161 | 165 | 214 | 216 | 97 | 99 | 138 |
| CV2 | CV | 89 | 115 | 0 | 0 | 167 | 169 | 214 | 246 | 97 | 101 | 140 |
| CV20 | CV | 89 | 115 | 164 | 168 | 161 | 167 | 214 | 214 | 97 | 99 | 172 |
| CV21 | CV | 89 | 93 | 164 | 168 | 159 | 159 | 218 | 218 | 97 | 97 | 166 |
| CV22 | CV | 85 | 85 | 164 | 168 | 163 | 167 | 214 | 218 | 97 | 97 | 138 |
| CV23 | CV | 89 | 103 | 164 | 168 | 163 | 167 | 214 | 214 | 97 | 101 | 138 |
| CV24 | CV | 89 | 127 | 164 | 168 | 163 | 167 | 218 | 218 | 97 | 99 | 172 |
| CV25 | CV | 89 | 89 | 164 | 168 | 163 | 167 | 214 | 214 | 95 | 99 | 160 |
| CV26 | CV | 89 | 127 | 160 | 164 | 165 | 167 | 0 | 0 | 97 | 99 | 138 |
| CV27 | CV | 89 | 105 | 160 | 164 | 161 | 167 | 218 | 218 | 97 | 101 | 138 |
| CV28 | CV | 89 | 89 | 164 | 168 | 163 | 167 | 214 | 216 | 95 | 99 | 122 |
| CV29 | CV | 87 | 89 | 164 | 168 | 163 | 167 | 214 | 216 | 97 | 99 | 138 |
| CV3 | CV | 87 | 89 | 160 | 164 | 165 | 169 | 214 | 238 | 95 | 99 | 140 |
| CV30 | CV | 89 | 105 | 164 | 168 | 163 | 167 | 216 | 216 | 95 | 97 | 138 |
| CV4 | CV | 89 | 89 | 160 | 164 | 163 | 167 | 212 | 244 | 95 | 99 | 144 |
| CV5 | CV | 89 | 89 | 160 | 164 | 163 | 167 | 226 | 240 | 95 | 99 | 140 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CV6 | CV | 87 | 89 | 164 | 168 | 159 | 161 | 214 | 246 | 95 | 97 | 126 |
| CV7 | CV | 89 | 89 | 164 | 168 | 159 | 161 | 212 | 244 | 95 | 99 | 0 |
| CV8 | CV | 89 | 89 | 160 | 164 | 161 | 163 | 212 | 214 | 97 | 99 | 178 |
| CV9 | CV | 87 | 89 | 160 | 164 | 161 | 163 | 212 | 214 | 97 | 99 | 142 |
| NV1 | NV | 89 | 89 | 0 | 0 | 165 | 167 | 214 | 216 | 97 | 97 | 0 |
| NV10 | NV | 85 | 89 | 0 | 0 | 165 | 169 | 214 | 222 | 97 | 99 | 134 |
| NV11 | NV | 85 | 89 | 0 | 0 | 163 | 167 | 214 | 222 | 97 | 101 | 0 |
| NV12 | NV | 89 | 89 | 0 | 0 | 163 | 165 | 214 | 246 | 97 | 99 | 0 |
| NV13 | NV | 89 | 89 | 0 | 0 | 163 | 167 | 214 | 222 | 97 | 97 | 134 |
| NV14 | NV | 89 | 89 | 0 | 0 | 159 | 165 | 214 | 220 | 97 | 97 | 134 |
| NV15 | NV | 89 | 89 | 0 | 0 | 157 | 165 | 214 | 222 | 97 | 99 | 132 |
| NV16 | NV | 89 | 109 | 0 | 0 | 167 | 169 | 214 | 222 | 97 | 99 | 134 |
| NV17 | NV | 89 | 89 | 156 | 160 | 157 | 165 | 214 | 222 | 95 | 99 | 144 |
| NV18 | NV | 85 | 89 | 156 | 160 | 163 | 167 | 214 | 218 | 97 | 101 | 134 |
| NV19 | NV | 89 | 89 | 156 | 160 | 157 | 165 | 214 | 230 | 95 | 99 | 0 |
| NV20 | NV | 85 | 89 | 0 | 0 | 163 | 167 | 214 | 220 | 97 | 99 | 0 |
| NV21 | NV | 85 | 89 | 154 | 158 | 161 | 163 | 214 | 222 | 97 | 99 | 134 |
| NV22 | NV | 89 | 89 | 156 | 156 | 163 | 167 | 214 | 222 | 97 | 99 | 134 |
| NV23 | NV | 89 | 89 | 0 | 0 | 165 | 171 | 214 | 222 | 95 | 99 | 144 |
| NV24 | NV | 85 | 89 | 154 | 158 | 165 | 167 | 214 | 222 | 97 | 99 | 134 |
| NV25 | NV | 89 | 113 | 0 | 0 | 159 | 171 | 212 | 220 | 97 | 99 | 134 |
| NV26 | NV | 87 | 89 | 160 | 164 | 167 | 171 | 212 | 214 | 97 | 99 | 134 |
| NV27 | NV | 87 | 89 | 160 | 160 | 165 | 169 | 214 | 222 | 97 | 101 | 134 |
| NV28 | NV | 87 | 121 | 160 | 164 | 163 | 167 | 214 | 222 | 95 | 97 | 0 |
| NV29 | NV | 89 | 113 | 160 | 160 | 157 | 163 | 212 | 214 | 97 | 97 | 134 |
| NV30 | NV | 85 | 89 | 0 | 0 | 165 | 173 | 214 | 216 | 97 | 99 | 144 |
| NV31 | NV | 89 | 89 | 0 | 0 | 159 | 163 | 214 | 234 | 97 | 97 | 0 |
| NV32 | NV | 0 | 0 | 160 | 160 | 153 | 167 | 214 | 218 | 97 | 99 | 134 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NV33 | NV | 89 | 99 | 176 | 176 | 161 | 165 | 214 | 244 | 97 | 101 | 0 |
| NV34 | NV | 89 | 89 | 176 | 176 | 161 | 163 | 230 | 244 | 97 | 99 | 122 |
| NV35 | NV | 89 | 89 | 0 | 0 | 163 | 167 | 212 | 214 | 95 | 99 | 0 |
| NV36 | NV | 89 | 89 | 0 | 0 | 163 | 167 | 212 | 212 | 95 | 97 | 0 |
| NV37 | NV | 89 | 111 | 0 | 0 | 167 | 175 | 212 | 244 | 97 | 99 | 160 |
| NV38 | NV | 89 | 89 | 0 | 0 | 167 | 169 | 214 | 230 | 97 | 99 | 146 |

**Appendix X: *Trypanosoma congolense* microsatellite loci fragment sizes**

| Sample | Pop | TB 5/2 | | TB 8/11 | | | TB 1/8 | | TB 6/7 | | TB 4/2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CC1 | CK | 80 | 80 | 88 | 106 | 97 | 99 | 115 | 121 | 100 | 100 |
| CC8 | CK | 80 | 83 | 88 | 88 | 97 | 99 | 115 | 123 | 104 | 104 |
| CC26 | CK | 80 | 80 | 88 | 94 | 95 | 97 | 115 | 117 | 106 | 106 |
| CC17 | CK | 80 | 80 | 88 | 102 | 97 | 99 | 115 | 119 | 106 | 106 |
| CC10 | CK | 85 | 89 | 88 | 106 | 97 | 99 | 121 | 133 | 106 | 106 |
| CC15 | CK | 80 | 87 | 92 | 104 | 97 | 99 | 113 | 177 | 106 | 106 |
| CC9 | CK | 80 | 89 | 88 | 88 | 95 | 95 | 117 | 133 | 106 | 106 |
| CC23 | CK | 85 | 86 | 88 | 96 | 97 | 99 | 107 | 119 | 108 | 108 |
| CC21 | CK | 80 | 80 | 88 | 94 | 97 | 99 | 115 | 119 | 108 | 108 |
| CC16 | CK | 85 | 87 | 88 | 102 | 97 | 99 | 115 | 119 | 108 | 108 |
| CC7 | CK | 83 | 101 | 88 | 106 | 97 | 99 | 119 | 119 | 108 | 108 |
| CC20 | CK | 80 | 80 | 88 | 104 | 97 | 99 | 105 | 105 | 109 | 109 |
| CC29 | CK | 80 | 103 | 88 | 106 | 97 | 99 | 105 | 107 | 109 | 109 |
| CC30 | CK | 80 | 83 | 88 | 94 | 95 | 97 | 109 | 109 | 109 | 109 |
| CC28 | CK | 80 | 80 | 88 | 88 | 97 | 99 | 107 | 115 | 109 | 109 |
| CC19 | CK | 80 | 80 | 88 | 94 | 105 | 111 | 107 | 115 | 109 | 109 |
| CC27 | CK | 101 | 101 | 88 | 102 | 97 | 99 | 115 | 119 | 109 | 109 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CC11 | CK | 80 | 87 | 88 | 104 | 97 | 99 | 115 | 119 | 109 | 109 |
| CC2 | CK | 80 | 83 | 88 | 88 | 97 | 99 | 105 | 117 | 110 | 110 |
| CC6 | CK | 80 | 85 | 88 | 106 | 97 | 99 | 115 | 121 | 110 | 110 |
| CC12 | CK | 80 | 99 | 86 | 88 | 97 | 99 | 121 | 121 | 110 | 110 |
| CC24 | CK | 101 | 101 | 88 | 98 | 97 | 99 | 119 | 123 | 110 | 110 |
| CC5 | CK | 105 | 105 | 88 | 88 | 97 | 99 | 119 | 125 | 110 | 110 |
| CC14 | CK | 85 | 87 | 88 | 96 | 97 | 99 | 119 | 133 | 110 | 110 |
| CC22 | CK | 80 | 89 | 88 | 102 | 95 | 99 | 119 | 119 | 112 | 112 |
| CC13 | CK | 80 | 93 | 88 | 102 | 97 | 99 | 121 | 125 | 112 | 112 |
| CC25 | CK | 85 | 87 | 94 | 106 | 97 | 99 | 115 | 119 | 106 | 106 |
| CC3 | CK | 80 | 83 | 88 | 104 | 97 | 99 | 115 | 119 | 108 | 118 |
| CC4 | CK | 80 | 81 | 88 | 102 | 97 | 99 | 115 | 119 | 109 | 118 |