University of Glasgow

Capewell, Paul (2011) *Human serum resistance in Trypanosoma brucei.*
PhD thesis.

http://theses.gla.ac.uk/2404/

Human Serum Resistance in *Trypanosoma brucei*

Paul Capewell

**Submitted in fulfilment of the requirements for
the Degree of Doctor of Philosophy**

**Institute of Infection, Immunity and Inflammation
College of Medical, Veterinary and Life Sciences
University of Glasgow**

September 2010

Abstract

Trypanosoma brucei is the causative agent of both sleeping sickness in humans and the related veterinary disease, Nagana. Both diseases have a wide distribution across sub-Saharan Africa and affect some of the poorest areas of the world. *T. brucei* can be segregated into three morphologically identical sub-species based on host, geography and pathology. *T. b. brucei* is limited to domestic and wild animals throughout sub-Saharan Africa and is non-infective to humans due to trypanosome lytic factors found in human serum. *T. b. gambiense* and *T. b. rhodesiense* are human infective sub-species, named due to their relative geographic locations. *T. b. gambiense* is the dominant form of the disease, causing over 90% of reported cases. Study of *T. b. gambiense* is complicated in that there are two distinct groups. Group 1 is invariably resistant to lysis and by far the more prevalent group. Group 2 *T. b. gambiense* exhibit a variable resistance phenotype and are only found at a small number of Côte d'Ivoire disease foci. There are two trypanosome lytic factors in human serum (TLF-1 & 2), both containing the proteins Apolipoprotein L1 (ApoL1) and Haptogoblin-related protein (Hpr). It has been conclusively demonstrated that the lytic component of TLF is ApoL1, although Hpr is required for maximal lysis by facilitating uptake of TLF particles via the HpHbR cell surface receptor.

This thesis has exposed several features of the human infectivity phenotype in both groups of *T. b. gambiense*, an area of research for which data has been lacking due to the difficulty of working with the organism. Fluorescence microscopy indicated that group 1 *T. b. gambiense* exhibit avoidance of TLF-1 particles by down-regulating HpHbR receptor expression and function. However, they are also able to resist the effects of recombinant ApoL1, suggesting an additional neutralisation or compensatory mechanism. Due to group 1 *T. b. gambiense* avoidance of TLF-1, TLF-2 is the more important lytic particle for this sub-species group and future research must take this into consideration. Unlike group 1, group 2 *T. b. gambiense* displays a variable human serum resistance phenotype that involves a neutralisation or compensatory mechanism for ApoL1, with no significant avoidance of lytic particles. Despite the high variability of the phenotype of group 2 *T. b. gambiense*, Quantitative Trait Analysis (QTL) using twenty-five F1 progeny from a *T. b. brucei* / group 2 *T. b. gambiense* cross indicated a strong heritable component to human serum resistance largely determined by a 30 gene locus on chromosome 8. Finally, a six multi-locus genotype population analysis of a Côte d'Ivoire *T. b. gambiense* focus was conducted, revealing little relationship between the two groups of *T. b. gambiense* in the field. The differences in the human serum resistance phenotypes and population genetics of both groups of *T. b. gambiense* revealed both prior and during this study make it appear likely that the two groups have evolved distinct human serum resistance strategies.

Acknowledgments

I would first like to thank my supervisors Dr. Annette Macleod and Prof. Mike Turner for their unwavering support and advice during my thesis. Despite a long litany of issues, they have stood by me, and more importantly, believed in me. Thank you.

I would also particularly like to thank Dr. Nicola Veitch, who has been a constant source of guidance and strength throughout the difficult times of my thesis. I am grateful for the patience she has shown me, even on the occasions when it was not deserved.

Additionally, I would like to thank Dr. Liam Morrison for sharing both his in-depth knowledge of trypanosomes and his love of IPA and I would also like to show my gratitude to Dr. Caroline Clucas for imparting a small piece of her considerable technical knowledge to me. Less specifically, I would like to thank all of the past and present members of the Trypanosome Genetics Group for the opportunity to work with such committed scientists and the camaraderie they have shown me.

I would also like to show my appreciation to my parents, Graham and Sue Capewell, for the decades of untiring support they have shown me. It has not been easy (nor cheap) but they have never missed an opportunity to lend aid or encouragement.

Finally, I would like to thank Esmeralda for her love and support. She has made my life complete.

Declaration

I declare that this thesis and the results presented within it are entirely my own work, with the following exceptions: the phenotype assays for the progeny of the *T. b. brucei* / *T. b. gambiense* cross were performed with the assistance of Dr. Nicola Veitch, University of Glasgow, who performed approximately half of the assays. In addition, a statistical analysis performed by Paul Johnson, University of Glasgow, is also presented regarding the degree of heritability for the human serum resistance trait.

No part of thesis has been previously submitted for a degree at any other institution.

Paul Capewell

Supporting Publications

KIEFT, R., CAPEWELL, P., TURNER, C. M., VEITCH, N. J., MACLEOD, A. and HAJDUK, S. (2010). Mechanism of *Trypanosoma brucei gambiense* (group 1) resistance to human trypanosome lytic factor. *Proc Natl Acad Sci U S A*.

Abbreviations

μg	Microgram
μl	Microlitre
μM	Micromolar
AFLP	Amplifed Fragment Length Polymorphisms
ANOVA Analysis of Variance	
ApoA1	Apolipoprotein A1
ApoL1	Apolipoprotein L1
BIIT	Blood Incubation Infectivty Test
Bp	Base Pair
CATT	Card Agglutination Test for Trypanosomiasis
cDNA	Complementary DNA
cM	Centimorgan
DABCO	1,4-diazabicyclo[2.2.2]octane
DAPI	4',6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
ES	Expression Site
ESAG	Expression Site Associated Gene
FBS	Foetal Bovine Serum
FITC	Fluorescein isothiocyanate
G	Gram
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GLM	General Linear Model
Hb	Haem
HDL	High Density Lipoprotein
HMI-9	Hirumi Modfied Iscove's Medium 9
Hp	Haptoglobin
Hpr	Haptoglobin Related Protein
HS	Human Serum
HSR	Human Serum Resistant
HSS	Human Serum Sensitive
Kb	Kilobase
L	Litre
LOD	Logarithm of the Odds

M	Molar
mg	Milligram
ml	Millilitre
MLG	Multilocus genotype
mM	Millimolar
mRNA	Messenger RNA
ORF	Open Reading Frame
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
QPCR	Quantitative PCR
QTL	Quantitative Trait Loci
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
RT-PCR	Reverse Transcriptase PCR
SRA	Serum Resistance Associated
Tbb	*Trypansoma brucei brucei*
Tbg	*Trypansoma brucei rhodesiense*
Tbr	*Trypansoma brucei gambiense*
TgsGP	*T. b. gambiense*-Specific Glycoprotein
TIM	Triose-phosphate isomerase
TLF	Trypanosome lytic Factor
UTR	Untranslated Region
VAT	Variable Antigen Type
VSG	Variant Surface Glycoprotein
WHO	World Health Organisation

Contents

Chapter 1: Introduction to *Trypanosoma brucei* and human infectivity

1.1 General Introduction

The African trypanosome *Trypanosoma brucei* is a flagellated protozoan of the class kinetoplastida and is a major parasite of humans and animals in sub-Saharan Africa. *Trypanosoma*tids are an evolutionary distinct order of eukaryotes, useful in the study of general eukaryote biology (Fernandes *et al.*, 1993; Stevens *et al.*, 2001). However, it is their capacity to be pathogens of humans and animals that make them especially important for study - the various subspecies of *T. brucei* are the causative agents of sleeping sickness in humans and the related veterinary disease, Nagana. Both diseases have a wide distribution across sub-Saharan Africa and deeply affect some of the poorest areas of the world (WHO, 2006).

There have been numerous outbreaks of African sleeping sickness during the twentieth century but due to implementation of a number of control measures, the disease was almost eradicated during the 1960s. Civil unrest and disease control failure since then have led to a resurgence of infection. It is estimated that there are half a million human infections and 70,000 deaths a year due to the parasite (WHO, 2006). This number is likely to be an underestimate however due to several African nations lacking a sleeping sickness monitoring program. In addition, recent studies have also indicated that the level of under-reporting of sleeping sickness deaths in endemic areas can be as high as 85% (Fevre *et al.*, 2005; Odiit *et al.*, 2005). Sleeping sickness exacts a huge cost on affected areas, placing a burden on local finances and healthcare systems. Few direct studies have been made on the size of this burden but at the Buma focus in the Democratic Republic of Congo it was estimated that 43% of household income was lost due to the disease, despite subsidisation (Lutumba *et al.*, 2007). In addition to the direct human cost, *T. brucei* in conjunction with *T. congolense* and *T. vivax* also have an impact on livestock mortality and morbidity. For communities already suffering under the burden of HIV and malaria, sleeping sickness and Nagana has a devastating effect on agricultural and economic development throughout sub-Saharan Africa (WHO, 2006). The study of *T. brucei* biology and how certain sub-species can infect humans and persist in non-human reservoirs will have important roles in controlling the disease.

1.2 Trypanosome morphology

Trypanosomes have a variable morphology depending on the life cycle stage of the organism. There are however several features common to each stage that can be typified in the bloodstream and procyclic form (Matthews, 2005). Trypanosomes contain the majority of the organelles seen in other eukaryotes, such as a nucleus, a mitochondrion, a lysosome, endoplasmic reticulum and Golgi apparatus. There are however some distinct differences. Most eukaryotic cells possess several lysosome organelles, while trypanosomes only possess one large lysosomal structure. There are also several organelles unique to the class, including the glycosomes, involved in glycolysis, and the kinetoplast. The kinetoplast is diagnostic of the entire class and is visible as a geimsa stainable structure separate from the nucleus at the base of the flagellum. It was originally thought to be involved in movement of the trypanosome, but it was later shown that this structure is actually the genome of the mitochondrion. The flagellum is attached to the DNA of the mitochondria via the basal body.

The mitochondrial genome itself is arranged in a highly esoteric way when compared to the familiar genomes of other eukaryotes. It consists of mini-circles and maxi-circles of DNA arranged in a linked chain-like mass. The maxi-circles encode several mitochondrial genes while the mini-circles encode guide RNA. The guide RNA is important for the RNA editing that takes place in the mitochondria of kinetoplastida to ensure correct gene expression (Benne *et al.*, 1986). The flagellum and basal body are important mediators of cell division and it is possible that the mechanical connection aids the separation of the complex mitochondrial genome during the process. The flagellum enters the cell body through the flagellar pocket, an invagination in the cell surface. This area is the primary site of endocytosis and the area at which transfer of substances across the cell membrane occurs (Overath & Engstler, 2004). Trypanosomes require a specific site of endocytosis because the outer cell membrane is covered in a dense, uniform coat of variant surface glycoprotein (VSG) (Vickerman, 1969). This masks the cell surface receptors normally involved in the transfer of substances involved in endocytosis. The VSG is a highly expressed, uniform protein that forms a thick layer over the majority of the cell surface. This presents a single dominant antigen to the host immune response that can be changed over the course of an infection in the process of antigenic variation (Section 1.7).

1.3 Lifecycle

It seems likely from molecular analysis of GAPDH genes that African trypanosomes are monophyletic and are likely to have originated as insect gut parasites (Hamilton *et al.*, 2007; Hamilton *et al.*, 2004), diverging from the American trypanosome *T. cruzi* approximately 100 million years ago (Stevens *et al.*, 1999). They have since evolved a complex lifecycle to exploit several definitive hosts and vectors (**Figure 1.1**).

Stumpy Bloodstream Form Form
VSG
Repressed mitochondrion
Non-Proliferative

Procyclic Form
No VSG
Activated mitochondrion
Proliferative

Density Dependant Differentiation

Bloodstream

Midgut

Salivary Gland

Mammal

Tsetse

Epimastigotes
Attached
Proliferative

Slender Bloodstream Form
VSG
Repressed mitochondrion
Proliferative

Metacyclics
Expresses VSG
Non-proliferative

Figure 1.1 Schematic representation of the lifecycle of all sub-species of *Trypanosoma brucei*, indicating the key features of each life cycle stage.

Although technically the life cycle can start anywhere, it is customary in parasite biology to begin descriptions at the inoculation stage of the mammalian host. Metacyclic form trypanosomes are transferred from the tsetse vector to the tissue of the mammalian host as the insect feeds. After a brief time in the lymph and interstitial fluid, the metacyclics differentiate into the bloodstream form and begin to replicate asexually by binary fission. It is at this lifecycle stage that the infamous antigenic variation used by trypanosomes to evade host immune defences is expressed (**Section 1.7**). The mammalian bloodstream is a homeostatically stable environment for a trypanosome, with high levels of glucose and other accessible nutrients. The abundance of glucose allows the trypanosome to reduce activity of the mitochondrion and instead rely on a high rate of glycolysis

in the glycosomes for its metabolism. Most bloodstream form trypanosomes have a 'long-slender' morphology but as the infection progresses 'short-stumpy' forms develop that are non-dividing but are infective to the tsetse.

The trypanosome enters the vector when a tsetse ingests the stumpy form parasites with the blood meal. These stumpy bloodstream forms differentiate into procyclic forms in the gut of the tsetse due to cues from temperature, pH change and the presence of enzymes such as trypsin (Matthews, 1999). Procyclics lose their VSG coat and instead express a similarly structured but less variable procyclin coat (Roditi & Clayton, 1999). As there is much less glucose in the tsetse gut then the mammalian bloodstream, mitochondrion activity increases and more cristae develop to facilitate this. Procyclic form cells undergo multiple rounds of replication within the gut of the tsetse before individuals migrate to the salivary gland as non-dividing proventricular forms. After reaching the salivary glands the proventricular forms morph into epimastigotes and attach to epithelial cells with their flagella. The epimastigotes undergo further asexual replication within the salivary gland but it is also at this approximate stage that sexual recombination is likely to occur (Tait *et al.*, 2007). The epimastigotes are non-infective for the mammalian host and they must first change form one final time to metacyclics - the VSG coat re-develops and mitochondrion activity decreases. Finally, the trypanosomes detach from the salivary gland wall and the metacyclic form cells are free to infect a new primary host the next time the tsetse feeds and thus continue the cycle.

1.4 Vectors

African trypanosomes are transmitted exclusively by flies of the *Glossina* genus (or tsetse - 'fly destructive to cattle' in the Tswana dialect of South Africa). Due to this fact, the geographic range of human infective trypanosomes largely overlaps with areas of high tsetse infestation (**Figure 1.2**). This has made targeting the tsetse a possible avenue for disease control with some early successes in isolated foci (Barrett *et al.*, 2003). Unlike many other vector species, both male and female tsetse feed exclusively on blood. They tend to feed on large mammals and do so in a telmophagic manner, using their mouthparts to tear host skin and feed on the resulting pool of blood and lymph. This feeding mechanism makes the bite of a tsetse quite painful. To assist with feeding, the flies secrete chemicals in their saliva to dilate blood vessels and slow coagulation (Parker & Mant, 1979). It is at this stage that an infected tsetse can transmit trypanosomes to the host. Although there are around thirty species of tsetse (Lane & Crosskey, 1993), they fall into

two broad groups - the riverine *palpalis* group usually found along the waterways of Western Africa and the *morsitans* group that prefer the drier savannah in Eastern Africa. The differences in the ecologies of the two tsetse groups correlate with the two distinct sub-species of human infective trypanosomes. *Trypanosoma brucei gambiense* is transmitted largely by *palpalis* tsetse in the West of Africa and *T. b. rhodesiense* is transmitted by *morsitans* in the East.
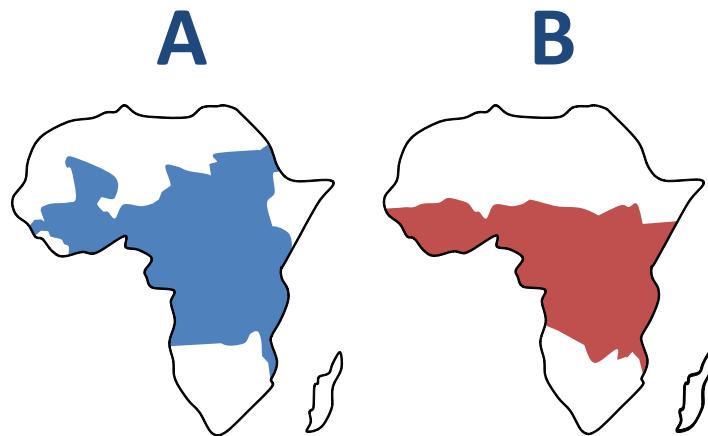
Figure 1.2 Distribution of countries with African sleeping sickness (A) and areas with high tsetse infestation (B). Adapted from (WHO, 2006) & (Rogers & Robinson, 2004).

1.5 The three *T. brucei* sub-species

Trypanosoma brucei can be segregated into three morphologically identical sub-species based on host species, geography and pathology. *T. b. brucei* is limited to non-human animals throughout sub-Saharan Africa and is non-infective to humans due to trypanosome lytic factors found in the serum of some old world primates (Section 1.13). *T. b. brucei* cells exposed to human serum quickly swell and lyse in a consistent manner (Pays *et al.*, 2006). *T. b. gambiense* and *T. b. rhodesiense* are human infective sub-species due to their ability to resist lysis by human serum. The two sub-species are defined by their relative geographic locations (Figure 1.3). *T. b. rhodesiense* is found in eastern sub-Saharan Africa (Rhodesia is the former name for Zimbabwe) and is primarily a disease of animals rather than humans, with little human to human transmission (Gibson *et al.*, 2002; Radwanska *et al.*, 2002; Welburn *et al.*, 2001). The human serum resistance phenotype in this sub-species shows variable expression between passages, although the

capability to resist lysis is not completely lost despite repeated rodent passage (Willett & Fairbairn, 1955). The disease typically progresses rapidly and is always fatal. Due to its amenability to laboratory manipulation this particular human infective sub-species has seen intensive study.
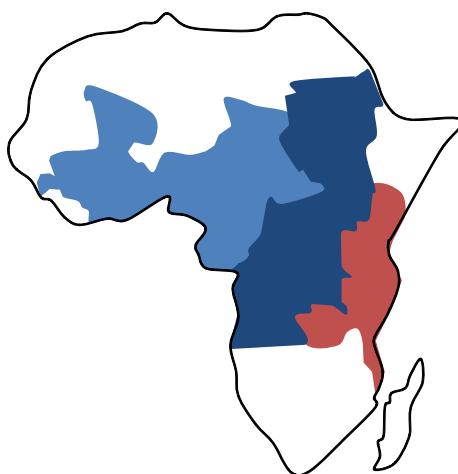
Figure 1.3 The distribution of countries reporting *T. b. gambiense* (blue) and *T. b. rhodesiense* (red) across sub-Saharan Africa. The dark blue area indicates the countries in which Sleeping Sickness is epidemic (adapted from (WHO, 2006).

A serum resistance associated (*SRA*) gene has been described in *T. b. rhodesiense* and is used as a diagnostic test to differentiate it from morphologically identical *T. b. brucei* sharing the same range (Enyaru *et al.*, 2006). There are however several human infective isolates that are geographically defined as *T. b. rhodesiense* but do not possess *SRA* (De Greef *et al.*, 1989; Enyaru *et al.*, 2006). What these particular lines are and how they relate to local populations has not been investigated.

As previously mentioned, *T. b. gambiense* is transmitted primarily by *G. palpalis* vectors found along waterways, often close to human settlements. It is by far the most prevalent causative human infective trypanosome and is responsible for over 90% of reported sleeping sickness cases (WHO, 2006). The disease is typically endemic and although it is believed to be primarily a human disease, domestic pigs may serve as a reservoir (Gibson *et al.*, 1978; Gibson, 1986; Paindavoine *et al.*, 1986). Other wild fauna have been shown to carry *T. b. gambiense*, albeit in limited numbers (Njiokou *et al.*, 2006). The *T. b. gambiense* form of the disease progresses more slowly than with

T. b. rhodesiense and tends to be either self-limiting or develop into a chronic disease of the lymphatic and nervous systems (Checchi *et al.*, 2008). Research with *T. b. gambiense* is complicated in that there are two discernable groups of *T. b gambiense,* identifiable by both genetic methods and various displayed phenotypes (Gibson, 1986). Isoenzyme and restriction fragment length polymorphisms (RFLP) profiles of isolates reveal a dominant group with low genetic variability termed group 1, and a second smaller group found only in Côte d'Ivoire. These group 2 *T. b. gambiense* display more genetic variability than group 1 (Gibson, 1986). The two groups also differ in human serum resistance phenotype. Group 1 *T. b. gambiense* possess an invariant infective phenotype while group 2 *T. b. gambiense* show a variable phenotype similar to *T. b. rhodesiense* (Mehlitz *et al.*, 1982).

1.6 Pathology

The pathology of the infection of both types of sleeping sickness has been well documented over the years and is delineated into two distinct stages (Burri & Brun, 2003; Stich *et al.*, 2002). The initial haemolymphatic stage of infection begins within the first few weeks following inoculation. The parasites initially divide asexually in the tissue surrounding the tsetse bite, often characterised by a chancre. This is a large inflammatory structure that forms in response to the insect bite and the presence of the parasites, although it is usually painless. The trypanosomes then invade the capillaries and enter the circulatory system where they continue to replicate within the blood of the host. The acute blood stage infection is often characterised by flu-like episodes of fever and headache. In *T. b. gambiense* the number of parasites in the blood tends to be very low and can often be undetectable for extended periods. Patients infected with *T. b. rhodesiense* exhibit comparatively higher parasitaemia and much more defined periods of fever. Peaks in parasitaemia are often associated with these symptomatic periods. As with other *T. brucei* infections, there is often an invasion of the lymphatic system. In the often long-term and chronic *T. b. gambiense* infections this can lead to numerous physiological changes including enlarged lymph nodes (Winterbottom's sign), weight loss, weakness and oedema.

In the second encephalitic stage of the infection, trypanosomes cross the blood-brain barrier and invade the central nervous system. The ensuing meningoencephalitis causes numerous neurological symptoms that get progressively worse as the infection reaches its conclusion - apathy, fatigue, confusion, motor changes and extreme changes in sleep patterns have all been reported. These neurological symptoms can occur within weeks of *T. b. rhodesiense* infection or

months or even years in *T. b. gambiense* infections. If left untreated, the infected person will experience convulsions, coma and an eventual painful death. The distinction is made between the two stages of the disease as different drugs are needed to combat each. Extremely toxic drugs capable of crossing the blood brain barrier must be employed if the disease proceeds to the encephalitic stage and should be avoided if possible (**section 1.8**).

An important feature of trypanosome infection is that parasitaemia peaks and troughs with some regularity. As the trypanosomes multiply, the host immune system responds by producing antibodies targeted to the cell surface VSG and after several days the parasitaemia decreases. Following the decrease in parasite number, a new wave of parasitaemia forms. The new parasites are unaffected by the previous immune response and the host must initiate a new one. This process appears to continue indefinitely (until death of the host or treatment). Trypanosomes perform this cyclic evasion of the immune response by varying the VSG being expressed on the cell surface in the process of antigenic variation (Barry & McCulloch, 2001; Pays *et al.*, 2001). By presenting different antigens to the immune system on a regular cycle, the acquired immune system is unable to completely clear an infection.

1.7 Antigenic variation

Some of the mechanisms of antigenic variation have been elucidated for trypanosomes by categorising the antibodies produced by animals inoculated with trypanosomes (Lumsden *et al.*, 1968). It became clear that trypanosomes could modify the VSG antigen that they presented to the immune system to create different variable antigen types (VAT). Changes in VAT correlated with new peaks of parasitaemia (Balber, 1972). While there were many antibodies that clustered into discrete families, the true diversity of VSG expression has been revealed by genetic techniques. *VSG* genes are found throughout the genome, usually at the telomeres or in clustered subtelomeric arrays and are transcriptionally silent. There are also *VSG* arrays on the mini-chromosomes (Vanderploeg *et al.*, 1984). Most *VSG* genes are represented as pseudo-genes and contribute to antigenic variation by serving as a reservoir to create new mosaic *VSG*s by recombination (Marcello & Barry, 2007). All in all, there are many hundreds of *VSG* genes in *T. brucei*, and nearly 10% of the coding regions of the trypanosome genome are *VSG* genes (Berriman *et al.*, 2005). In order for a *VSG* gene to be active it must sit in one of the telomeric expression sites that possess the expression site associated genes (*ESAG*) (Pays *et al.*, 1989). A similar contingent of *ESAG*s is found in each expression site, although few have been fully

characterised. There are approximately twenty bloodstream expression sites found in the *T. b. brucei* genome but only one is active at a time (Becker *et al.*, 2004). This is likely to be due to a physical position that the expression site needs to be in to be transcribed (Navarro & Gull, 2001). The expression sites contribute to the antigenic variation of trypanosomes in that each expression site is likely to contain a different *VSG*. By changing expression sites via *in situ* transcriptional switching, they can quickly vary the *VSG* presented to the immune system.

Varying expression sites may also contribute to the host adaptability of the parasite by modifying the *ESAG*s being expressed. Although the *ESAG* complement of each expression site is similar, there are subtle differences between genes in different expression sites. The most studied of these *ESAG*s is the transferrin receptor formed from *ESAG6/7* (Ligtenberg *et al.*, 1994; Salmon *et al.*, 1994; Steverding *et al.*, 1994; Steverding *et al.*, 1995). These two genes have hypervariable regions that differ between copies of each gene in different expression sites and the receptor expressed from them appear to have different affinities to various animal transferrins (Bitter *et al.*, 1998; Gerrits *et al.*, 2002). It is proposed that expression site switching aids host adaptation as *T. brucei* are able to infect a wide range of definitive hosts (Pays *et al.*, 2001; van Luenen *et al.*, 2005). The variation found in *ESAG6/7* between expression sites can also be exploited to assess when an expression site switch has occurred by sequencing the transcript of the expressed *ESAG*s, particularly the sequence of the hypervariable regions of *ESAG6/7* (Hertz-Fowler *et al.*, 2008).

In addition to *in situ* transcriptional switching, the *VSG* at the active expression site can be swapped for another either using homologous recombination (McCulloch & Barry, 1999). This system utilises the homologous 70bp repeat regions found upstream of the 5' end of all *VSG*s and in the expression sites. The *VSG* can either enter the expression site by gene conversion with another *VSG* from the library or by reciprocal exchange from another expression site (Conway *et al.*, 2002; Stockdale *et al.*, 2008). Switching occurs spontaneously within a small number of trypanosomes and these are selected for by the host immune response. There is a pattern of *VSG* expression that ensures that gene switching occurs consecutively and some *VSG*s are more likely to be selected for than others (Barry & Turner, 1991). The sequence homology between some *VSG*s can also lead to partial gene recombination with members of the library of pseudo-*VSG*s, creating entirely new 'mosaic' *VSG*s, further increasing the potential for antigenic variation (Marcello & Barry, 2007). By changing the *VSG*s or modifying them by recombination in varying expression sites, trypanosomes have an almost limitless capacity to vary the antigen they expose to the host immune system. The most important feature of antigenic variation is that it renders

the acquired immune system of mammals largely ineffective. To overcome this, innate resistance features targeting the parasites specifically are employed to combat infection (Section 1.13).

1.8 Drug treatment in humans

Several drugs are used to treat African sleeping sickness, depending on the stage of the disease and the sub-species (Leach & Roberts, 1981; Nok, 2003). Suramin and pentamidine are administered during the early stages of the infection before the central nervous system is invaded. Both drugs are relatively toxic and surprisingly old – suramin was introduced in 1920 and pentamidine in 1941. Pentamidine is less toxic than suramin and has traditionally been used to treat *T. b. gambiense* rather than *T. b. rhodesiense* despite working efficiently against both sub-species (Nok, 2003). If the drugs are administered in the initial stages of the disease the chances of survival are high. Rapid and accurate disease identification is therefore essential. This is hampered by poor equipment and resources found in endemic sleeping sickness regions, the difficulties in defining the disease from clinical symptoms and also the lack of robust methods for diagnosing infection.

For late stage infections, melarsoprol or eflornithine are used. Melarsoprol is extremely toxic, requiring hospitalisation and monitoring for at least ten days (Burri *et al.*, 2000). Eflornithine can be used for *T. b. gambiense* infections but not *T. b. rhodesiense*. Despite the expense of manufacture, the WHO has recently negotiated deals with the companies involved to produce all of the mentioned drugs at a reduced cost to the organisation. These are now provided free at the point of care to patients in affected area (http://www.who.org/). However, there are increasing incidents of drug resistance being reported in the field (de Koning, 2001; Maser *et al.*, 2003; Matovu *et al.*, 2001). Human trypanomiasis treatment failure has reached 30% in some areas where melarsoprol is employed (Legros *et al.*, 1999).

1.9 Trypanosome genome organisation

Trypanosomes possess an unusual genomic structure when compared to other eukaryotes, reflecting the large evolutionary distance between them. Approximately 20% of the DNA component is within the kinetoplast, with the remaining 80% found in the nucleus (Simpson, 1972). Despite the fact that trypanosome chromosomes do not condense during mitosis, pulse field gel electrophoresis has revealed that there are several size classes of chromosome in the nucleus of the cell; approximately one hundred mini-chromosomes (50-150kb), one to six

intermediate-chromosomes depending on the strain (200-1000kb) and eleven mega-chromosomes (>1000kb) (Gottesdiener *et al.*, 1990; Melville *et al.*, 1998; Vanderploeg *et al.*, 1984). The mega-chromosomes contain most of the metabolic and *VSG* genes while the smaller chromosomes contain predominantly *VSG* genes and pseudo-genes (Melville *et al.*, 1998). The non-telomeric regions of the mega-chromosomes have been sequenced for both the *T. b. brucei* strain TREU927 and the *T. b. gambiense* strain DAL972 (Berriman *et al.*, 2005; Jackson *et al.*, 2010). The complete complement of *VSG*s has largely been sequenced in the *T. b. brucei* strain Lister 427 and efforts are underway to replicate this in the *T. b. brucei* genome strain TREU927 (http://www.sanger.ac.uk).

An important aspect of parasite biology is the ploidy of the organism. This attribute has been shown to have implications in the emergence of new traits in parasites, such as drug resistance in *Plasmodium falciparum* (Sibley & Hunt, 2003). By comparing isoenzymes, RFLP profiles and total DNA content it was shown that the mega-chromosomes of bloodstream trypanosomes are diploid (Gibson *et al.*, 1980; Tait, 1980). The ploidy of the intermediate and mini-chromosomes is still undetermined, although they are divided equally between daughter cells during binary fission and the patterns of certain features in the mini-chromosome are highly stable over multiple generations (Alsford *et al.*, 2001). Cellular DNA content is similar to blood stage forms in procyclic and metacyclic forms so it can be extrapolated that the organism is also diploid in these stages (Kooy *et al.*, 1989; Tait *et al.*, 1989). In fact, no haploid cells have been described despite population genetics studies revealing that genetic exchange occurs in trypanosomes (Gibson *et al.*, 1980; Tait, 1980; Tait, 1983). Several triploid cells lines have been described, always arising in genetic crosses performed in the laboratory (Gibson *et al.*, 1992; Hope *et al.*, 1999). The prevalence of triploids is related to the strains involved in the crosses and is more likely to occur in strains less related to each other (Hope *et al.*, 1999). The existence of triploid cells has not yet been demonstrated to occur in the field (MacLeod *et al.*, 2005a).

The study of trypanosome genetics has been greatly enhanced by the publication of the *T. b. brucei* genome sequence of the coding areas of the mega-chromosomes (Berriman *et al.*, 2005) and the subsequent publication of the genome of the group 1 *T. b. gambiense* DAL972 (Jackson *et al.*, 2010) showing that the two subspecies are extremely alike, with an average of 99.2% identity between genes (Jackson *et al.*, 2010). There were no differences in gene complement or order. There are however important caveats before discounting the potential existence of a gene specific to group 1 *T. b. gambiense* conferring human infectivity. Sub-telomeric regions were not

included in the sequencing process due to limitations in the approaches used. The *T. b. gambiense* reads were also assembled using the *T. b. brucei* genome as a scaffold – it is possible that novel regions of the genome could have been discarded by employing this technique. It is also important to remember that non-coding regions were omitted for comparative analysis and it is possible that resistance mechanisms may involve the modification of a non-coding region that affects of expression of a gene, rather than the presence or absence of that gene in the genome.

1.10 Experimental genetics of *T. brucei*

Sexual recombination has been described under laboratory conditions and shown to occur in the tsetse (Jenni *et al.*, 1986). It is not yet known exactly when recombination occurs within the fly, although it is likely to occur before the epimastigotes differentiate to metacyclics in the salivary glands (Gibson *et al.*, 2006; Tait *et al.*, 2007). Recent antibody based assays have identified attached cells expressing proteins that preceed meiosis, although no haploid stage have yet been identified (Peacock *et al.*, 2011). Recombination is not obligatory in *T. brucei* and the majority of trypanosomes passing through the tsetse possess the same genotype as the parents (Jenni *et al.*, 1986). Additionally, self-fertilisation can occur within a single parental line (Gibson & Stevens, 1999; Tait *et al.*, 1996; Tait & Turner, 1990). There have been numerous population studies and several further crosses confirming that sexual recombination is a common phenomenon in *T. brucei* but rates can vary greatly between strains (Gibson, 2009; Gibson & Garside, 1991; Gibson *et al.*, 1992; MacLeod *et al.*, 2005a; Schweizer *et al.*, 1994; Sternberg *et al.*, 1988; Tait *et al.*, 1996; Turner *et al.*, 1990). The segregation of alleles in three laboratory derived *T. brucei* crosses correlates closely with a Mendelian model and alleles on different chromosomes segregate independently (MacLeod *et al.*, 2005a; Turner *et al.*, 1990).

Mendelian inheritance requires that meiosis and cell fusion are necessary but does not stipulate the order of events. The discovery of triploid cells in association with genetic crosses provided possible evidence for fusion preceding meiosis, although recent data suggest that fusion occurs after meiosis (Peacock *et al.*, 2011). The first hypothesis proposed was that two parental cells fuse to form a tetraploid which then pares down DNA content by an unknown method to create a normal diploid cell (Paindavoine *et al.*, 1986). Incorrect destruction of DNA would result in triploidy. In this fusion-loss model however the assortment of the parental alleles would be non-Mendelian. It has been shown conclusively that *T. brucei* reproduces in a Mendelian manner so the model is unsuitable for understanding genetic exchange in this species (MacLeod *et al.*,

2005a) A different fusion model was also postulated in which the cells fuse to form tetraploid cells as before, but the nuclei then undergo meiosis to create eight haploid daughter nuclei (Gibson *et al.*, 1995). Two of these haploid nuclei fuse to form the normal diploid cell and the other haploid nuclei are destroyed. This model can result in Mendelian inheritance but no cells with a physiology featuring eight haploid nuclei have been described. Triploidy may arise if three haploid nuclei fused accidently or more likely if a haploid cell fused with a normal diploid cell. In addition, trisomy has been described for individual chromosomes although this is most likely as a result of non-disjunction during meiosis rather than fusion (Hope *et al.*, 1999). The final model proposed to explain sexual reproduction in *T. brucei* is that like many other protozoan parasites, there is a haploid gamete stage produced via meiosis. These then fuse to create standard diploid cells (Tait & Turner, 1990). As with the fusion hypothesis, no cells have yet been identified that feature the characteristics of such a haploid cell. At present it is impossible to determine which of the proposed models is correct (or if both models are wrong) but efforts are underway with laboratory crosses using fluorescent trypanosomes in an attempt to clarify the situation (Gibson *et al.*, 2006).

The publication of the *T. brucei* genome and the revelation that trypanosome sub-species are genetically similar has allowed the development of forward genetic techniques for both *T. b. brucei* and group 2 *T. b. gambiense*. One such approach is to measure the link between the alleles of different genes with a phenotypic trait in the progeny of a genetic cross. Firstly a genetic map must be created indicating how close genetic markers are to each other along a chromosome. This is estimated by genetic distance in centiMorgans and can be inferred by how often alleles of genes are inherited relative to neighbouring alleles in the progeny of isolates. Genes physically close to each are more likely to have alleles inherited together. When the distance between genes increases, the frequency of a crossover between them also increases. The distance between genes can be estimated with no prior knowledge of the genomic sequence of an isolate based on the likelihood of recombination between markers (Doerge, 2002). The pattern of inheritance reveals the frequency of crossovers between markers and this is used to infer genetic distance and produce a genetic map of the chromosomes. Two genetic maps have been created for *T. brucei*. The first map was for the fully sequenced *T. brucei* reference strain TREU927 and originated from a cross between TREU927 and the *T. b. brucei* strain STIB247 (MacLeod *et al.*, 2005b). A second map has also been created for the group 2 *T. b. gambiense* strain STIB386 by also crossing it with STIB247 (Cooper *et al.*, 2008). STIB247 was selected to be the second parent in each cross due to the fact it is predominantly homozygous along each of the mega-

chromosomes while STIB386 and TREU927 are largely heterozygous (MacLeod *et al.*, 2005b). Differences in the phenotype of the progeny can be attributed to whichever of the chromosome homologues the hybrid has inherited from the heterozygous parent.

Using this methodology, the relationship between the phenotype and the inherited haplotype of the hybrids can be ascertained. Relevant alleles are either inherited by the progeny or not and likewise the phenotype will either be present or not. By comparing the parental markers inherited by the progeny, markers more likely to be possessed by progeny displaying the phenotype suggest that genes close to those markers are involved with the phenotype. It is important that the distance between markers is also considered to take into account random crossover events (Churchill & Doerge, 1994). Using this methodology Quantitative trait loci (QTL) analysis is ideal for the study of complex phenotypes that arise due to multi-gene interactions. This is revealed as continuous phenotypic response in the progeny, rather than the discrete presence or absence of the phenotype in question. In this case the degree that the phenotype is displayed in each progeny is also considered in conjunction with the proportion of progeny inheriting the marker. Using these methods, several experiments have identified key genes involved with cellular processes, despite being part of a complex multi-gene response. For example, several regions of the *Toxoplasma gondii* genome were implicated by QTL analysis as being involved with the virulence of the parasite (Su *et al.*, 2002). One locus however contributed over 50% of the phenotype and further detailed QTL mapping and experimental analysis of the area identified a kinase that is vital to virulence (Taylor *et al.*, 2006). QTL analysis has also been successfully utilised in *T. brucei,* revealed a primary locus implicated in host pathogenesis and suggesting several other regions involved in the phenotype (Morrison *et al.*, 2009). QTL analysis can also be applied to phenotypes showing low heritability or penetrance. This is necessary as an investigation of human serum resistance using the group 2 *T. b. gambiense* strain STIB386 would need to overcome the fact the phenotype is variable, despite clearly possessing a genetic component. Most studies of variable or low penetrance traits are concerned with human diseases, predominantly diabetes and cancer (Glazier *et al.*, 2002). In many diabetes studies for example, the phenotype has both an environmental and genetic component and the degree of expression varies greatly. Despite this, several related diabetes related loci have been identified by designing studies with a large enough number of test subjects able to overcome the inherent variance (Glazier *et al.*, 2002). With a large number of progeny and sufficient heritability, the methodology applied in these human studies can be transferred to trypanosome phenotypes.

1.11 Mating in natural *T. brucei* populations

It has been shown that mating can occur in *T. brucei* under laboratory conditions but it is also important to understand whether the phenomenon can occur in the field. Three theoretical positions have emerged to explain the population structures of *T. brucei* disease foci. At one end of the spectrum there is the suggestion that the foci are panmictic populations with high levels of genetic exchange and sexual recombination. This theory was originally based on analysis of allele frequencies for several polymorphic isoenzyme markers for parasite populations (Gibson *et al.*, 1980; Tait, 1980). At the other end of the spectrum is the clonal theory whereby sexual recombination is rare or absent and genetic variation arises exclusively from mutation (Tibayrenc *et al.*, 1990). A third option is that populations follow an epidemic structure in that there is some background recombination but a successful genotype can become dominant, as might be expected in an epidemic (Maynard-Smith *et al.*, 1993).

Understanding the population structure of a disease focus has important implications for the control and treatment of disease. In a clonal population for example, double-drug resistance would require two separate mutation events simultaneously, which is an unlikely occurrence. On the other hand, if sex is present in a population, double-resistance can arise from recombination of two single-resistant parents. A randomly mating population therefore has the variability and capability to evolve more quickly in response to the selective pressure of drug treatment. This rapid evolution is not limited to drug resistance and can also occur with other genetic traits, such as virulence, pathogenesis or human infectivity.

Several techniques have been utilised to investigate the population genetics of *T. brucei* populations. All are based on detecting the distribution of alleles within a population to infer relationships between isolates. Initially, assays using isoenzymes were utilised to investigate the relationships present in *T. brucei* stocks (Gibson *et al.*, 1978; Godfrey *et al.*, 1987; Kilgour, 1980). This technique utilises the differing characteristics of allozymes (such as charge, size and conformation) and the affect these have on their rate of passage through starch gels. Polymorphisms between enzymes were detected as bands at differing positions in gels from which the alleles for single copy genes could be inferred. There are many downsides to this technique, such as requiring vast amounts of starting material and synonymous mutations in a gene are not detected, even though these would be informative to infer relationships. Parallel to this research there have been limited studies utilising restriction fragment length polymorphisms (RFLP) to investigate *T. brucei* populations (Paindavoine *et al.*, 1986; Paindavoine *et al.*, 1989).

This method used combinations of restriction enzymes and the unique fragments generated from cut genomic DNA when visualised on Southern blots. Relationships between parasite isolates were then inferred by the fragments that each possessed. Point mutations in genome sequence can lead to the loss or gain of an enzyme restriction site and insertions, deletions or duplications also result in changes in the fragment size. Obviously this method misses polymorphisms that do not lead to changes in restriction sites or the size of fragments. Additionally, substantial amounts of starting material are required and this precluded the use of this method for many field studies, especially with *T. b. gambiense* which has low parasite densities in the host. There is a benefit to both methods in that knowledge of the genome sequence of the organisms is not required. The development of the polymerase chain reaction (PCR) has overcome the limitations of needing large initial sample amounts somewhat, allowing amplification of very small amounts of DNA. It has also allowed development of newer techniques for studying population genetics.

One technique which utilises PCR is the amplified fragment length polymorphism (AFLP) technology (Agbo *et al.*, 2001; Agbo *et al.*, 2002). This system first digests genomic DNA with a selection of restriction enzymes. Adapters are ligated to the ends of these genomic fragments and PCR targeted to these adapters amplify the fragments. The amplified fragments are then visualised on polyacrylamide gels and the pattern of sizes in the fragment lengths can be used to infer relationships and inheritance. The use of PCR allows much smaller starting amounts of material to be used compared to previous techniques, but it is still not ideal for parasites that show low density in the host. Also the genome positions of polymorphic loci are unknown and the potential for mixed-genotype infections cannot be discerned.

Further refinement of PCR technology has led to the development of highly sensitive mini and microsatellites techniques. These have been made possible by the publication of the *T. brucei* genome (Berriman *et al.*, 2005). Mini and microsatellites are heritable sequences of tandemly repeated DNA found throughout eukaryotic genomes (Jeffreys *et al.*, 1990; Jeffreys *et al.*, 1985; Weissenbach *et al.*, 1992). Minisatellites are usually between 10-100bp in length and microsatellites are between 2-6bp. Microsatellites are far more abundant then minisatellites and show a larger degree of variation in copy number making them ideal to determine relationships between individuals of a population – the unique variation in the copy number of each microsatellite acts as a genetic fingerprint. The *T. brucei* genome sequencing project revealed the presence of many microsatellites of larger than 12 repeats equally spaced throughout the mega-chromosomes (Berriman *et al.*, 2005). By using PCR amplification of a DNA sequence containing a

microsatellite and gel electrophoresis of the amplicon in a high percentage agarose gel, it is possible to distinguish between microsatellite repeat sizes and assign alleles to individuals. This technique has been utilised in several modern studies of trypanosome population genetics, including *T. b. gambiense* (Morrison *et al.*, 2008). Band ambiguity can be alleviated using technology such as Genescan® which fluorescently tags the PCR product, allowing it to be accurately sized against standard markers using capillary electrophoresis

By employing these various techniques, population structures have been described for several *T. brucei* populations. *T. brucei* has different population structures in different geographical foci; clonal (Koffi *et al.*, 2009; Mathieu-daude & Tibayrenc, 1994; Morrison *et al.*, 2008), epidemic (Hide *et al.*, 1994; MacLeod *et al.*, 2000; MacLeod *et al.*, 2001; Stevens & Tibayrenc, 1996) and panmictic (Tait, 1980). Epidemic and panmictic populations structures strongly indicate that mating occurs at these foci (MacLeod *et al.*, 2000). With regards to the human infective trypanosomes, microsatellite studies with *T. b. rhodesiense* isolates have shown that although the overall *T. brucei* population in a given geographic location appears to be epidemic in eastern Africa; the *T. b. rhodesiense* population itself is clonal. This population has remained stable for several decades, separate from the *T. b. brucei* population (MacLeod *et al.*, 2001). *T. b. gambiense* have been less thoroughly studied than other *T. brucei* sub-species but several group 1 foci exhibit clonal population structures (Koffi *et al.*, 2009; Morrison *et al.*, 2008; Tibayrenc *et al.*, 1990; Truc & Tibayrenc, 1993). It is also known that they show a much less diverse *VSG* and *ESAG* repertoire then other *T. brucei* subspecies (Dero *et al.*, 1987; Hertz-Fowler *et al.*, 2008). Group 2 *T. b. gambiense* are far less prevalent and even less well characterised then group 1. RFLP and isoenzyme studies indicate that group 2 are genetically distinct from group 1 and are more closely related to *T. b. brucei* then group 1 *T. b. gambiense* (Gibson, 1986; Hide *et al.*, 1994). Despite this, the two groups of *T. b. gambiense* are still placed into the same sub-species based on historical and geographical grounds. If group 1 *T. b gambiense* are truly clonal then they may have lost the ability to undergo genetic recombination – no genetic cross with a group 1 *T. b. gambiense* has yet been performed in the laboratory and this would preclude forward genetic analysis. However, the well characterised group 2 *T. b. gambiense* strain STIB386 has been shown to undergo sexual recombination in the laboratory in a Mendelian manner and would be ideal for such analysis (MacLeod *et al.*, 2005a; Turner *et al.*, 1990).

1.12 Primate resistance to trypanosomes

Due to antigenic variation, *T. brucei* can largely avoid the effects of the mammalian adaptive immune system indefinitely. However, some species of primate have evolved an innate resistance mechanism to combat infection. This innate immune mechanism quickly lyses *T. b. brucei*, although *T. b. gambiense* and *T. b rhodesiense* are able to resist lysis and still infect humans. The lytic effect of human serum on non-human infective trypanosomes was first shown in the early 20th century (Laveran & Mesnil, 1912). When observed by microscopy, trypanosome lysis by human serum proceeds rapidly along a well defined path. After a lag phase of approximately thirty minutes, the cell experiences gross changes to morphology as it swells rapidly until it ruptures within one to two hours (Rifkin, 1978; Rifkin, 1984). Microscopic evidence indicates that the swelling of the cell and lysis is due to the lysosome swelling (Pays & Vanhollebeke, 2009; Pays *et al.*, 2006). The serum from several old world primates, including gorillas, baboons and mandrills were also shown to lyse non-human infective trypanosomes but with differing rates of activity (Laveran & Mesnil, 1912).

There have been several studies on primate serum and lysis (Laveran & Mesnil, 1912; Lugli *et al.*, 2004; Seed *et al.*, 1990). The ability to lyse trypanosomes is confined to old world primates and may have originated on two different occasions (Poelvoorde *et al.*, 2004) – once in the *Cercopithecidae group (*baboons & mandrills) and once in the *Hominidae* group. The high similarity of the genes involved in trypanosome lysis make it more likely that the trait evolved prior to the split between the two groups of primates, approximately 25 million years ago. The serum of the *Papio* genus (baboon) primates has a much higher lysis efficiency compared to *Hominidae* and also lyses human infective *T. b. rhodesiense*. It is interesting to note that *Hominidae* of the *Pan* genus (chimpanzees) do not posses trypanocidal serum despite the close evolutionary links with the *Gorilla gorilla* and *Homo sapiens*. The simplest explanation is that chimpanzees have lost the ability to combat trypanosome infection rather than the lytic effect arising separately in other *Hominidae*.

1.13 Trypanosome lytic factors, TLF-1 & TLF-2

The earliest observations of trypanosome lysis indicated that the mechanism was not associated with the classical immune system components of immunoglobin or complement but rather uses some unknown factor in human serum (Laveran & Mesnil, 1912). It has also been known for nearly a century that the serum of humans is highly toxic to multiple trypanosome species (Hajduk

et al., 1994; Laveran & Mesnil, 1912). While this might have evolved as a specific anti-trypanosome measure, it may also have arisen as a general protozoa defence strategy as the trypanolytic component of human serum also produces measurable physiological effects in *Plasmodium falciparum* (Imrie *et al.*, 2004) and two *Leishmania* species (Samanovic *et al.*, 2009). Nevertheless, most work on elucidating the mechanisms behind the lytic effect has been performed using *T. brucei*. By testing the trypanocidal effect of separated fractions of HDL on *T. b. brucei*, it was shown that the majority of the lytic effect was associated with the densest fraction of human serum – the high density lipoprotein (HDL) component (Rifkin, 1978). Purified human HDL from the same fraction size also clears *T. b. brucei* infections in mice (Rifkin, 1978).

HDL bodies are spherical, soluble amalgams of phospholipids, cholesterol and apolipoproteins (Pays *et al.*, 2006) and in addition to a trypanocidal effect, appear to have an important role in controlling both blood pressure and cholesterol levels (Castelli *et al.*, 1977; Forde *et al.*, 1977; Miller, 1978). HDLs are a diverse family of particles, containing several different protein families (Blanche *et al.*, 1981). The importance of HDL to trypanosome lysis was confirmed in experiments that showed that the level of HDL in serum correlates with the efficiency of lysis, depletion of HDL reduces lytic activity and that anti-human HDL antibodies can inhibit lysis (Gillett & Owen, 1991; Hajduk *et al.*, 1989). Physiological evidence was provided by the characterisation of the human condition Tangiers disease in which patients have a reduced HDL component in their serum and no corresponding trypanolytic effect (Rifkin, 1978).

Most HDL bodies are not trypanocidal but one particular particle found in separation studies was shown to be highly lytic. This was termed the Trypanosome Lytic Factor (subsequently TLF-1 after the discovery of a second TLF) (Hajduk *et al.*, 1989). Western blots identified two proteins unique to TLF-1 not found in other HDL particles. In the same series of experiments it was noted that lysis was inhibited when trypanosomes were incubated below 17°C, the same temperature at which endocytosis is greatly reduced in eukaryotes (Dunn *et al.*, 1980). This fact, coupled with the initial thirty minute lag phase before observable lysis effects provided compelling evidence that TLF-1 is internalised by endocytosis before it can mediate lysis. Two different receptors for TLF-1 have been suggested to exist based on data from low-temperature binding assays - a high affinity receptor that has approximately 350 copies per parasite and a low affinity receptor that has 60,000 receptors per cell (Drain *et al.*, 2001). In addition to TLF-1, a second TLF has since been described in a larger sized fraction of HDL (Raper *et al.*, 1999; Tomlinson *et al.*, 1997). The larger size of the TLF-2 particle is due to its association with IgM molecules on the particle surface. TLF-2

contains much less lipid then TLF-1 but possesses similar lytic characteristics and is found at a similar concentration in blood (Raper *et al.*, 2001). It does however possess a much shorter half-life then TLF-1 and so is difficult to characterise further. The receptor for TLF-2 is less well characterised, although it is known that it has low affinity, high turnover and haptoglobin does not compete for the receptor, unlike TLF-1 (Vanhollebeke *et al.*, 2007b). It has been suggested that IgM is the ligand for entry of TLF-2 into the cell as this particular immunoglobin has been shown to weakly bind VSG on the trypanosome surface with a characteristic low affinity (Vanhollebeke & Pays, 2010). TLF-2 would therefore be internalised as part of the normal surface recycling mechanism found in all trypanosome species.

1.14 ApoA1, Hpr and ApoL1

The protein component of TLF-1 and TLF-2 consists of the apolipoproteins L1 (ApoL1) and A1 (ApoA1) (Tytler *et al.*, 1995) and haptoglobin-related protein (Hpr) (Smith & Hajduk, 1995; Tomlinson *et al.*, 1997). TLF-2 is also found bound to IgM, contributing to the larger size of the particle (Raper *et al.*, 2001). ApoA1 is a ubiquitous apolipoprotein contained within 70% of all HDL particles and is predominantly involved in providing structure. Both recombinant ApoA1 and transgenic mice expressing human ApoA1 are unable to kill *T. b. brucei* indicating that direct involvement in lysis can be discounted (Owen *et al.*, 1992). However, the binding of TLF-1 particles to the low affinity receptor was competed by other HDL particles , including non-lytic bovine partiles, suggesting that the ligand could be ApoA1 (Vanhamme & Pays, 2004). As ApoA1 is often bound to cholesterol, a receptor to scavenge this protein is unsurprising due to the fact trypanosomes are cholesterol auxotrophs but have extremely high cholesterol requirements due to rapid turnover of the VSG coat (Seyfang *et al.*, 1990).

Of the remaining two proteins, haptoglobin-related protein appeared to be the most likely candidate to be the lytic factor. Hpr is a close paralogue to haptoglobin protein (Hp) which has been shown to possess a toxic peroxidase effect when bound to haemoglobin in low pH conditions similar to those within the lysosome (Bonkovsky, 1991; Connell & Smithies, 1959). It was hypothesised that a similar toxicity may be present in Hpr and several experiments have implicated Hpr in trypanosome lysis. It was shown that the then unknown high affinity TLF-1 receptor is not competed for by other HDLs but by Haptoglobin bound to Haem, indicating Hpr is the likely ligand in this case (Drain *et al.*, 2001; Smith & Hajduk, 1995). It can be conjectured that a Haptoglobin receptor would be useful for trypanosomes to take up bound haemoglobin for iron,

with the binding of Hpr a cross-reaction. Trypanosome lines that do not have efficient uptake of TLF-1 do not survive if transferrin is removed from serum, highlighting the importance of iron to the parasite and why they would need to bind Haptoglobin bound to haem (Hager & Hajduk, 1997). Anti-haptoglobin antibodies have been shown to cross react with Hpr and also protect trypanosomes from lysis in the presence of TLF-1 (Smith & Hajduk, 1995). It was also noted that non-lytic serum from chimpanzees does not contain Hpr due to a frame-shift mutation in the gene, adding further weight to the argument that Hpr is the lytic factor (McEvoy *et al.*, 1988).

It was proposed that Hpr was the trypanolytic component of HDL with a working hypothesis that haemoglobin-bound Hpr enters the parasitic cell via endocytosis. After Hpr enters the low pH environment of the lysosome, an iron-based Fenton peroxidase reaction occurs which damages the membrane and initiates lysis (Smith & Hajduk, 1995). Adding catalase to the parasite to reduce peroxide concentration also reduced lysis, seemingly supporting the peroxidase hypothesis. It was also shown that the end-products of lipid peroxidation were found in lysed cells and application of DPPD (an inhibitor of the lipid degrading peroxidase effect) prevented lysis (Bishop *et al.*, 2001). These conclusions that Hpr is the lytic component of TLF are however disputed. It has since been shown that Hpr may not bind haemoglobin after all so there can be no associated peroxidase Fenton reaction (Muranjan *et al.*, 1997). Also, no peroxidase intermediaries were shown to be formed during interaction with TLF-1 so it may be that the observed end-products of peroxidation are formed coincidentally (Portela *et al.*, 2000). It has been insinuated that the observed peroxidase effect was due to contamination by haemoglobin bound to Haptoglobin (Vanhamme *et al.*, 2003). More recent studies of the effect of catalase on lysis have also been unable to repeat earlier results of lysis inhibition (Raper *et al.*, 2001). Finally, recombinant Hpr has not been shown to be lytic (Vanhollebeke *et al.*, 2007b).

An alternate hypothesis was put forward that the trypanocidal protein of TLF was a different common component of both TLF particles - ApoL1. ApoL1 arose as a possible lytic factor candidate when it was shown to co-localise with the serum resistance associated (SRA) protein of *T. b. rhodesiense* (Section 1.17) (Vanhamme *et al.*, 2003). ApoL1 possesses a pore forming domain under low pH conditions that is similar to that of bacterial colicins and the mammalian Bcl2 proteins involved in apoptosis (Perez-Morga *et al.*, 2005). This domain is essential to the lytic function of the protein (Perez-Morga *et al.*, 2005). The protein localises with the lysosome membrane marker p67 after endocytosis, suggesting that it embeds in the lysosomal membrane after entry (Vanhamme *et al.*, 2003). Both naturally derived and recombinant ApoL1 lyses

trypanosomes with a swollen lysosome morphology similar to that seen after exposure to human serum, although unbound ApoL1 lyses at a slower rate than normal human serum (Vanhamme *et al.*, 2003; Vanhamme & Pays, 2004; Vanhollebeke *et al.*, 2007b). It was proposed that ApoL1 induces the formation of pores in the lysosome that in turn create an influx of chloride ions (Vanhamme *et al.*, 2003). The influx of ions upsets the osmotic balance of the lysosome, causing it to rapidly swell and rupture.

Like Hpr, ApoL1 is present in higher primates but due to a premature stop codon it is inactive in chimpanzees, explaining the lack of lytic activity in this species (Poelvoorde *et al.*, 2004). Hpr has also recently been revealed to be present in all old world primates, even those with non-lytic serum (Vanhollebeke & Pays, 2010) suggesting it evolved before ApoL1 or innate immunity to trypanosomes. Strong evidence for ApoL1 being the lytic factor of TLF came from the discovery of a unique medical case of a human in India found to be infected by the animal trypanosome *T. evansi*. Normally this trypanosome species is not human infective although it is closely related to *T. brucei*. The evidence indicates that the patient's infection is likely due to two frame-shift mutations, one in each allele of their ApoL1 gene. This has rendered the protein non-functional. The patient's Hpr gene has no major deviations in sequence from that of other humans (Vanhollebeke & Pays, 2006). The addition of natural or recombinant ApoL1 to the patient's serum fully rescued the lytic phenotype.

1.15 A consensus hypothesis for the mechanism of trypanosome lysis

In light of the mounting evidence that ApoL1 and Hpr are both vital for efficient trypanosome lysis, a consensus involving both proteins has begun to form (**Figure 1.4**). Reconstituted HDL containing both purified ApoL1 and Hpr causes 99% lysis while those that contain Hpr or ApoL1 separately are twenty times less effective over the same time-span (Shiflett *et al.*, 2005). It has been argued that both Hpr peroxidation and ApoL1 pH dependent pore generation are needed for the full lytic effect due to the killing effect seen in naturally derived Hpr (Shiflett *et al.*, 2007). There is some evidence that the killing effect of naturally derived Hpr is actually due to the uncleaved signal peptide which exhibits trypanocidal effects by disrupting the fluidity of the plasma membrane. This effect is only seen with bloodstream form trypanosomes. The killing effect may explain the disparity between recombinant Hpr assays (in which the signal peptide is cleaved) and naturally derived Hpr where it remains uncleaved (Harrington *et al.*, 2010). However,

the morphology of trypanosomes killed by Hpr signal peptide does not match that of exposure to human serum or TLF (Harrington *et al.*, 2010). Experiments investigating recombinant ApoL1 with serum from the patient with a defective ApoL1 gene have come to a similar conclusion that both ApoL1 and Hpr are necessary for efficient lysis. Recombinant ApoL1 on its own lysed trypanosomes at a slower rate than normal human serum but when added to the serum from the patient with defective ApoL1, normal human serum activity was restored (Vanhollebeke *et al.*, 2007b). The rate at which ApoL1 lysed parasites was also increased when added to recombinant Hpr, although they found no evidence that Hpr has a direct killing effect on trypanosomes due to the fact that the patient with non-functional ApoL1 had perfectly normal Hpr. This suggests that Hpr has no direct impact on combating trypanosome infection unless paired with ApoL1. This is not the case with ApoL1, as even transgenic animals expressing human ApoL1 are resistant to *T. b. brucei* infection, even without co-expression of Hpr (Molina-Portela *et al.*, 2008; Thomson *et al.*, 2009). It appears likely that Hpr is the ligand that facilitates the uptake of TLF-1 and ApoL1 and a receptor that binds both haptoglobin bound to haem and Hpr has recently been described, termed *HpHbR* (Vanhollebeke *et al.*, 2008).

The HpHbR receptor is a GPI anchored protein of approximately 300 copies found on the cell surface of the parasite. This closely matches the characteristics of the proposed high affinity receptor for TLF-1 (Drain *et al.*, 2001). Down-regulation of the receptor in *T. b. brucei* leads to complete protection against the lytic effects of TLF-1, but not normal human serum or TLF-2. Therefore, the current hypothesis of innate immunity to trypanosomes in humans is that the TLF particles bind to receptors on the trypanosome surface (Drain *et al.*, 2001; Vanhollebeke *et al.*, 2008), are internalised by endocytosis (Hager *et al.*, 1994) and transported to the lysosome (Shimamura *et al.*, 2001) where they cause disruption of the lysosomal membrane and eventual cell death (Hager *et al.*, 1994; Lorenz *et al.*, 1994). ApoL1 is the component of TLF-1 that causes lysosome disruption. The protein undergoes a conformational change when exposed to the low pH environment of the lysosome, embeds in the lysosomal membrane and forms channels that allow chloride ions to enter. This upsets the osmotic balance of the organelle, causing expansion and eventual rupture of the lysosome. Hpr facilitates lysis by efficiently binding TLF-1 to the HpHbR receptor so that ApoL1 can enter the lysosome. There may also be secondary toxic effects caused by the Hpr signal peptide that lower the fluidity of the lysosomal membrane, making it more susceptible to ApoL1. The presence of ApoL1 in TLF-2 makes it likely that this particle also uses the protein as a lytic mechanism, although uptake of TLF-2 is not by Hpr and may instead involve the weak binding of IgM to VSG (Vanhollebeke & Pays, 2010).
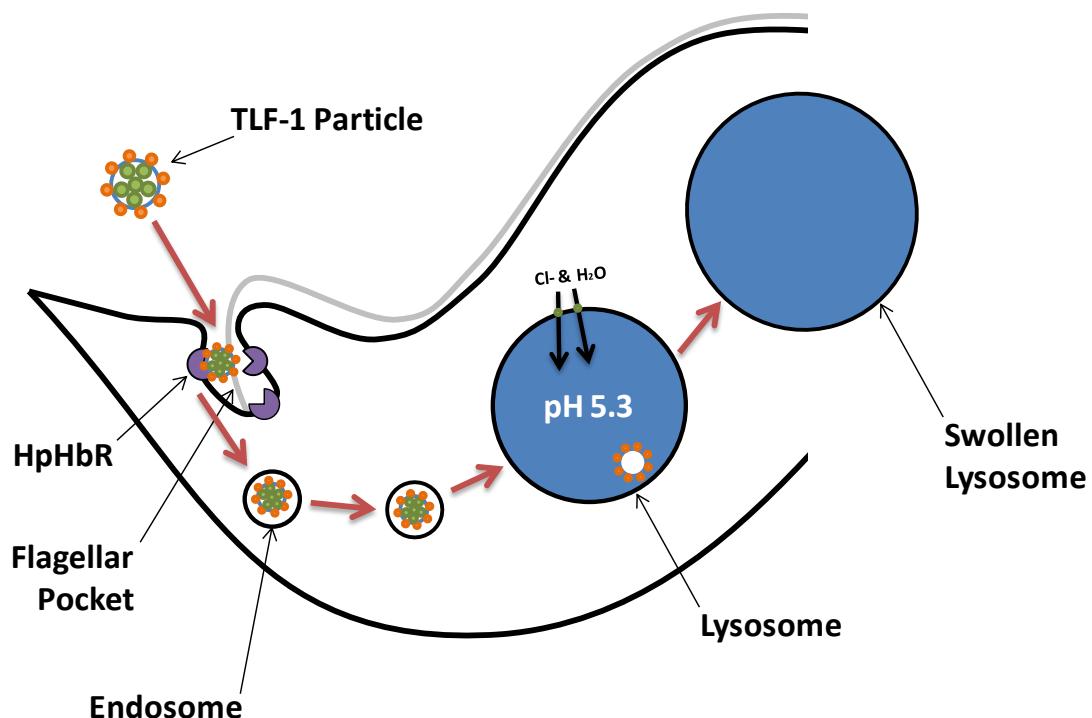
Figure 1.4 An overview of the current model of TLF-1 mediated lysis in *T. b. brucei* (adapted from (Pays & Vanhollebeke, 2009; Pays *et al.*, 2006). The TLF particle contains Hpr (Orange), ApoL1 (Green) and ApoA1. The TLF particle binds to the HpHbR receptor with the Hpr protein as the ligand. It is then internalised by endocytosis and transferred into early endosomes (pH 7). As the endosomes mature and eventually integrate with the lysosome, the pH of the organelle falls to 5.3 causing the release of ApoL1 (green) from the TLF particle. ApoL1 undergoes a pH mediated conformational change, exposing the pore-forming and membrane addressing domains. This new structure allows the integration of ApoL1 into the plasma membrane of the lysosome. When integrated within the membrane, ApoL1 acts as an ion channel that allows Cl^- to freely enter the organelle and disrupt osmotic balance. Water then enters the lysosome along the new osmotic gradient, causing uncontrolled swelling of the organelle and eventually lysing the trypanosome. TLF-2 acts in a similar manner to TLF-1, although the ligand is not Hpr and is instead believed to be internalised due to weak interactions with TLF-2 bound IgM and the surface VSG.

1.16 Measuring human serum resistance

Understanding the animal and vector reservoir of human infective trypanosomes and distinguishing human infective sub-species from *T. b. brucei* are important factors in sleeping sickness control. To this end, several tests have been formulated over the years to identify human infective trypanosomes. The most decisive test for human infectivity is clearly whether a trypanosome can infect a human, and such a methodology was used in the famous Tinde experiments (Willett & Fairbairn, 1955). There are however obvious ethical and legal reservations with inoculating viable, deadly parasites into humans. As a substitute, assays have been developed based on the long known fact that the trypanolytic factor is a component of human serum.

Early diagnostic tests used whole blood or purified serum and direct observation of lysis with limited numbers of trypanosomes (Laveran & Mesnil, 1912). This crude analysis led to a standardised test for human serum resistance (HSR) in the Blood Incubation Infectivity Test (BIIT) (Rickman & Robson, 1970). BIIT initially assessed parasite survival rates in mice injected with human serum and was primarily designed for identification of *T. b. rhodesiense* from *T. b. brucei* (Rickman & Robson, 1970). The BIIT was then later modified to assess parasite survival by incubation with human serum *ex vivo* and then injecting the cells into a rodent model (Targett & Wilson, 1973). Although the technique is expensive and time consuming, it has helped confirm numerous animal reservoirs of human infective trypanosomes (Geigy *et al.*, 1973; Gibson & Wellde, 1985).

Over the years, alternate methods have been developed to assess serum resistance in trypanosomes in more quantitative ways. For example, an assay of the number of metacyclics that convert to bloodstream forms in the presence of serum (Brun & Jenni, 1987). The test is laborious, takes several days and there is inherent inaccuracy due to the fact that only a minority of metacyclics exhibit a resistant phenotype (Brun & Jenni, 1987). Another option is to use blood-stream form trypanosomes and simply observing cell death or gross changes in morphology over time when exposed to lytic components. This method is favoured by most researchers (Hajduk *et al.*, 1994; Jenni & Brun, 1982; Vanhollebeke *et al.*, 2007a). Although highly accurate with modern microscopes and culturing techniques, the assay is labour intensive and requires highly trained observers. These increasingly accurate techniques have also revealed that the simple human serum resistant/sensitive phenotype is too crude a distinction - there are also intermediate levels of resistance. Although noted in BIIT tests (Rickman & Robson, 1970) it was unclear if this was due

to mixtures of resistant and sensitive lines or whether resistance had a continuous distribution. Assays using cloned lines showed that an intermediate phenotype of slower lysis is a true phenomenon in some lines (Turner *et al.*, 2004). Interestingly, the *T. b. brucei* reference strain TREU927 possesses an intermediate phenotype (Turner *et al.*, 2004; Vanhamme *et al.*, 2004).

More modern serum resistance assays use automated procedures to estimate cell lysis by employing dyes and fluorescence. A method based on the release of intracellular enzymes from lysed trypanosomes and the cleavage of Calcein-AM into fluorescent products has been successfully employed (Tomlinson *et al.*, 1995). Unfortunately, when this method was compared to microscope determined survival controls there was a degree of variability and discrepancy. An alternate approach is to use two dyes with differing affinities to live or lysed trypanosomes. For example, the dyes Syto 10 and DEAD Red used in conjunction. Syto 10 is a membrane permeable dye that stains nucleic acid with a fluorescent marker visible in the green spectrum. DEAD Red stains nucleic material with a red marker but is not membrane permeable. Thus trypanosomes that have compromised membranes due to the interactions with TLF will show both red and green fluorescence, while healthy cells will only show green. This allows for an accurate estimation of the percentage death and the resistance to lysis of an isolate using fluorometers or FACS (Turner *et al.*, 2004).

An interesting facet of human serum resistance that has been revealed by these resistance assays is that the phenotype is lifecycle stage specific. The procyclic form trypanosomes in the insect gut are fully resistant to human serum. While this was believed to be because procyclics exhibit ten-fold lower rates of endocytosis then blood stream forms (Morgan *et al.*, 2001), genetically modified *T. b. brucei* with high rates of endocytosis remain human serum resistant (Natesan *et al.*, 2010). This suggests that procyclic forms of *T. brucei* are constitutively resistant to human serum and it is unclear why bloodstream forms of *T. b. brucei* are not able to utilise the same method. Why procyclic forms have developed to be resistant to human serum is unknown although it was hypothesised that because tsetse feed often, trypanosomes will be exposed to human serum several times during infection. However, procyclics quickly move to the ectoperitrophic space after gut infection and it is unknown if TLF particles can cross the peritrophic matrix. Counter-intuitively, the metacyclic stage of the life-cycle appear to be human serum sensitive (Brun & Jenni, 1987). Human infective trypanosomes such as *T. b. gambiense* and *T. b. rhodesiense* must somehow survive the initial exposure to human serum until they become bloodstream form (approximately 24 hours). This may be explained by the initial period spent by the invading

parasites in the interstitial fluid in the skin and the development of a chancre before full conversion to bloodstream forms and exposure to TLF.

1.17 *Trypanosoma brucei rhodesiense & SRA*

It was observed early in human infectivity studies that *T. b. rhodesiense* showed an inconsistent human serum resistance phenotype. Although the serum resistance phenotype was retained during animal passage, individual passages exhibit the phenotype to different degrees (Targett & Wilson, 1973; Willett & Fairbairn, 1955). Unfortunately, this also invalidated the BIIT somewhat for the original purpose of *T. b. rhodesiense* identification. It was noted that resistance appeared to be related to antigenic variation and changes to the variable antigen type (VAT) of the populations. Isogenic clones of the ETat strain of *T. b. rhodesiense* that differed only in VAT expression showed different serum resistance phenotypes (van Meirvenne *et al.*, 1976).   A particular VAT type was identified as correlating with human serum resistance- ETat 1.10. This VAT was always present in resistant lines and not in sensitive lines. Gene switching of the *VSG* did not reduce resistance to lysis, implicating an *ESAG* rather than the *VSG* itself (Rifkin *et al.*, 1994). Some VAT clones other than ETat 1.10 do show some levels of human serum resistance, although it is unclear if this is contamination by the presence of ETat 1.10 (van Meirvenne *et al.*, 1976). All of these data indicate the presence of a gene in an expression site that was differently expressed depending on the expression site being used. Although the likely location of the serum resistance element was now understood, the actual gene was unknown.

In an attempt to narrow down the options, mRNA was compared between isogenic resistant and susceptible lines. A transcript associated with resistance was identified and predicted to be for a *VSG*-like gene (De Greef *et al.*, 1992) lines possessed the gene for the protein but only resistant lines expressed it. The gene was later confirmed to be an *ESAG* in the ETat 1.10 expression site by Southern blot analysis with probes for this serum resistance associated (*SRA*) gene. Conclusive proof that human serum resistance in *T. b. rhodesiense* was due to a single gene was shown when transgenic *T. b. brucei* expressing *SRA* gained resistance to lysis (Xong *et al.*, 1998). Harking back to the earlier days of trypanosome research, the result was confirmed by inoculation into a human when a *T. b. brucei-SRA* line infected a researcher in the Pays laboratory (Gibson, 2005). Unlike the pioneering work of the Tinde experiment, this human inoculation was accidental but it clearly demonstrated that *SRA* gene was sufficient to confer human serum resistance to a *T. brucei* line.

SRA has been shown to be present in the vast majority of studied *T. b. rhodesiense* lines (Gibson, 2005) although this has become a somewhat circular argument - *SRA* is diagnostic of *T. b. rhodesiense*, so *T. b. rhodesiense* always possess *SRA*. However a recent study in Uganda has identified several human infective *T. brucei* in a *T. b. rhodesiense* focus that did not PCR amplify with *SRA* primers (Enyaru *et al.*, 2006). Additionally, in a study comparing serum sensitive and resistant isogenic lines expressing the same ETAT, the *SRA* transcript was not present in all resistant lines (De Greef *et al.*, 1989; Rifkin *et al.*, 1994). It is unknown as to whether these phenomena are due to mutations in the *SRA* gene, the expression site or an entirely new resistance trait arising in a *T. brucei* population. Structurally, *SRA* appears to be a truncated *VSG* with a large deletion of a region in the centre of the N-Terminal domain (De Greef *et al.*, 1992; De Greef & Hamers, 1994; Xong *et al.*, 1998). SRA is present on the cell surface (Milner & Hajduk, 1999), where it is internalised in the flagellar pocket during cell surface protein turnover, leading to SRA presence in the endosomes and lysosome (Oli *et al.*, 2006; Shiflett *et al.*, 2007; Vanhamme *et al.*, 2003). This correlates well with what is known about the localisation of human TLF in the lysosome before lysis.

Genetically, *SRA* is well conserved in *T. b. rhodesiense*, with less than 3% sequence variation between strains (Gibson *et al.*, 2002). This would make it seem likely that the gene arose only once (Campillo & Carrington, 2003). The slight differences present in SRA sequence segregate into northern and southern groups (Gibson *et al.*, 2002; MacLean *et al.*, 2004). This coincides with a previous distinction of *T. b. rhodesiense* into two sub-groups based on RFLP, minisatellite analysis (Gibson *et al.*, 2002; Godfrey *et al.*, 1990; MacLeod *et al.*, 2000) and disease severity (MacLean *et al.*, 2004). However, the differing genotypes and phenotypes may also be due to genetic exchange between strains resulting in the transfer of the *SRA* gene (and other traits) to *T. brucei* of different genetic backgrounds. This is supported by the fact that some populations of *T. b. rhodesiense* are more related to neighbouring *T. b. brucei* then other *T. b. rhodesiense* populations, suggesting that *T. b. rhodesiense* is a human-infective variant of *T. b. brucei* (MacLeod *et al.*, 2001). This leads to a working hypothesis that on entry into the human bloodstream, a few trypanosomes are expressing *SRA* at an active expression site. These are selected for and the population within the host become resistant. In non-human animals this selective pressure is not applied, explaining the variable nature of serum resistance when *T. b. rhodesiense* is passaged through these species.

Although *SRA* is now well described in the literature, theories of SRA mechanism are somewhat polarised. The prevailing theory is that SRA inhibits the activity of ApoL1 based on the observed co-localisation of the two proteins in parasite cells and *in vitro* (Vanhamme *et al.*, 2003). The deletion of 126 amino acids in SRA removes two surface loops normally present in a VSG protein, exposing the internal α-helices (Campillo & Carrington, 2003). Based on modelled tertiary structure it was proposed that the deletion resulted in human serum resistance as it allows an inter-chelating interaction between the exposed helices and the helical section of ApoL1, either in the lysosome or on the cell surface (Vanhamme *et al.*, 2003; Vanhamme *et al.*, 2004). The alternative hypothesis is that SRA somehow disrupts internal trafficking of TLF to the lysosome and it remains within pre-lysosome vesicles (Oli *et al.*, 2006; Shiflett *et al.*, 2007). These trafficking experiments however used the laboratory adapted *T. b. brucei* strain Lister 427 ectopically expressing *SRA* and may not necessarily be indicative of what occurs in *T. b. rhodesiense.*

Regardless of the mechanism, the discovery of *SRA* has led to advancements in diagnosis and suggested preventative action to combat sleeping sickness caused by *T. b. rhodesiense*. With the advent of specific PCR markers for *SRA*, trypanosomes with the potential to infect humans can be identified without the need for laborious human serum resistance assays such as the BIIT (Gibson *et al.*, 2002; Radwanska *et al.*, 2002; Welburn *et al.*, 2001). This has proved useful in understanding the role of wildlife and livestock in the epidemiology of the disease. For example, traditional microscopy techniques and HSR assays estimated the number of livestock cattle to be carrying human infective *T. b. rhodesiense* at approximately 1%. Modern analysis has indicated that instead this value is closer to 18% (Welburn *et al.*, 2001). This cross-sectional study may itself also be an underestimation - long term additive studies show the trypanosome burden of cattle herds could be higher by a further 20% (Maudlin 2006, personal communication).

The importance of the cattle reservoir of *T. b. rhodesiense* may become increasingly important in countries such as Uganda that are now leaving a period of long-term political unrest and population migration is increasing. In the town of Soroti in Uganda, a recent *T. b. rhodesiense* outbreak has emerged that appears to coincide with a recent movement of cattle from a traditional endemic focus of sleeping sickness in Busoga district (Fevre *et al.*, 2001). There is a significant correlation between the presence of the disease and distance from the Soroti cattle market since 1998. Cluster analysis also identified the market as a statistically significant risk factor in contracting sleeping sickness. Although the evidence is circumstantial, the emergence of

a *T. b. rhodesiense* disease focus in Soroti would seem to correlate with the recent movement of cattle into the area (Fevre *et al.*, 2001; Hutchinson *et al.*, 2003).

The knowledge of how SRA acts to inhibit the lytic component of TLF has also led to an explanation of how the serum of baboons is able to lyse *T. b. rhodesiense*. The ApoL1 present in this species possesses two amino acid substitutions in the C-Terminal helix of the protein. Normal *T. b. rhodesiense* SRA is unable to bind this protein variant and cannot inactivate the pore forming domain (Lecordier *et al.*, 2009; Perez-Morga *et al.*, 2005; Thomson *et al.*, 2009). It is proposed that the distribution of bull sperm from transgenic cattle expressing the modified baboon ApoL1 could be distributed at foci in Africa to generate herds of cattle resistant to both *T. b. brucei* and *T. b. rhodesiense*. In addition to the astronomical cost of the endeavour, the lack of widespread acceptance of transgenic animals by the African public would make the scheme appear unfeasible. The removal of *T. b. brucei*, *T. b. rhodesiense*, *T. vivax* and *T. congolense* by the application of these transgenic animals is also likely to see selection for trypanosomes that do not possess the *SRA* gene as a serum resistance mechanism, such as *T. b. gambiense* or local *T. b rhodesiense* that do not posses *SRA*. Without the competition from *T. b. brucei* or *T. b. rhodesiense* it is impossible to predict how other trypanosomes will evolve to exploit the new niche – past experience indicates that it is unlikely to be in our favour. An alternative use of baboon ApoL1 would be the creation of a stable delivery system to deliver modified ApoL1 to people infected with *T. b. rhodesiense*. This method has already been used with limited success *in vitro* on *T. b. rhodesiense* by conjugating the pore-forming domain of ApoL1 with a single domain antibody that enters the cell after binding to the VSG (Baral *et al.*, 2006).

Although the discovery of *SRA* has been a major boon to the understanding of *T. b. rhodesiense* genetics and biology, it is unable to be applied to the dominant human infective trypanosome *T. b. gambiense*. *SRA* is not present in either group of *T. b. gambiense* (De Greef *et al.*, 1992; Radwanska *et al.*, 2002; Turner *et al.*, 2004) nor in the intermediate phenotypes such as TREU927 (Turner *et al.*, 2004). An '*SRA*-like' motif has been located in the TREU927 genome but it is not involved in resistance (Vanhamme *et al.*, 2004). Elucidating the mechanism of human serum resistance in *T. b. gambiense* is vital for future disease control in Africa. The hope is to take what has been learned from the process of investigating *SRA* and *T. b. rhodesiense* and apply it to *T. b. gambiense* with more modern techniques.

1.18 *Trypanosoma brucei gambiense*

In comparison to *T. b. rhodesiense,* very little is known about the serum resistance mechanism of *T. b. gambiense*. As previously mentioned, there are two groups of *T. b. gambiense* (Gibson & Borst, 1986). Group 2 has a similar serum resistance phenotype to *T. b. rhodesiense*, with a great deal of variance in the degree of human serum resistance displayed. Group 2 *T. b. gambiense* may be in a similar situation to *T. b. rhodesiense* in that an *ESAG* has become modified to inhibit TLF. The fact that a *T. b. brucei* strain artificially selected *in vitro* to be human serum resistant shows some changes in *ESAG* expression indicates that *ESAG* modification linked to human serum resistance can occur out with *T. b. rhodesiense* (Faulkner *et al.*, 2006). However, in this particular line it has been shown that these changes to *VSG* and *ESAG* are probably unrelated to the resistance phenotype and rather it is down-regulation of the expression of the *HpHbR* gene that has conferred TLF resistance (Kieft *et al.*, 2010). There have been few studies on the human serum resistance phenotype in group 2 *T. b. gambiense*. Preliminary *in vivo* studies with the group 2 *T. b. gambiense* strain TxTat have shown a correlation between the VAT of the parasite and resistant or sensitive phenotypes, suggesting it is linked to *VSG* and ES switching (Ortiz *et al.*, 1994). It was also shown in the same line that resistant TxTat clones depleted the lytic potential of human serum more slowly than the sensitive clones (Ortiz-Ordóñez & Seed, 1995). It has however come to light that there is some confusion as to the exact identity of some TxTat lines, with some derivatives being revealed to actually be *T. b. rhodesiense* (Grab & Kennedy, 2008). Lineage data for the TxTat used in the *ex vivo* assays was unavailable so it may be possible it has been misidentified as a type 2 *T. b. gambiense*. If expression site switching is not related to resistance in group 2 *T. b. gambiense*, there could be other mechanisms involved that vary transcription in a similar manner, such as the stochastic expression that emerges from gene network interactions found in some eukaryotes (Li *et al.*, 2010). Indeed, all *T. b. brucei* exhibit expression site free, transcriptionally regulated, variable human serum resistance in that procyclic forms are always human serum resistant while bloodstream forms always sensitive, even in the same isogenic line (Natesan *et al.*, 2010). This suggests that every *T. b. brucei* has the potential to become human infective merely by modifying expression of pre-existing genes.

Very little is known about the resistance mechanisms in group 1 *T. b. gambiense* due to the difficulty of working with it under laboratory conditions. BIIT tests indicate it has a robust and invariant resistance phenotype (Gibson, 1986). This would indicate a genetic mechanism of resistance that is being constitutively expressed, completely unlike the expression linked system of *T. b. rhodesiense* or group 2 *T. b. gambiense*. In order for *T. b. gambiense* to resist lysis, the parasite must either have the ability to reduce uptake of TLF, neutralise TLF or to compensate for the effects. If resistance is due to modified uptake of TLF it could be due to modification of receptors shown to be involved in TLF uptake, such HpHbR. If the mechanism is a neutralisation of ApoL1 or its effects in the parasite cell, any number of processes involved in cell lysis could be involved and locating a specific gene or protein would be more difficult. The inability to conduct forward genetics with this particular group of the sub-species restricts the range of approaches that can potentially be used to investigate the resistance mechanism of group 1 *T. b. gambiense*.

Chapter 2: Materials & Methods

2.1 *T. b. gambiense*, *T. b. brucei* & *T. b. rhodesiense* experimental lines

STIB247 is a human serum sensitive *T. b. brucei* clone, first isolated from a Hartebeest in Serengeti in 1971 (Geigy & Kauffmann, 1973). Lister 427 is the standard laboratory *T. b. brucei* strain it is believed to have been isolated from a sheep in 1960 (Peacock *et al.*, 2008). The human serum resistant strain of Lister 427 was generated after repeated passage with increasing exposure to lytic human HDL (Faulkner *et al.*, 2006). The human serum resistant group 2 *T. b. gambiense* strain STIB386 was originally isolated in 1978 from an infected patient in Côte d'Ivoire (Felgner *et al.*, 1981). The cloned strain was adapted to grow *in vitro* in HMI9 medium (Hirumi & Hirumi, 1989) supplemented with 1.5 mM glucose, 1 mM methyl cellulose, 250 µM adenosine and 150 µM guanosine. To generate resistant and sensitive isogenic lines of STIB386, several culture adapted sub-lines were assayed for their ability to resist lysis in human serum. The most resistant STIB386 line (STIB386R) was identified and maintained in culture in the presence of 10% human serum/10% foetal bovine serum. The most sensitive line (STIB386S) was also identified and maintained without human serum. Eliane is a group 1 *T. b. gambiense* strain isolated from a human in 1952, also from Côte d'Ivoire. This isolate was maintained in modified HMI9 supplemented with 20% human serum. Human serum resistance phenotypes for these three isolates have been described previously (Turner *et al.*, 2004). Other *T. b. gambiense* group 1 strains used were isolated from infected individuals and include: MHOM/CM/75/ITMAP1789 BIM and MHOM/CM/74/ITMAP1787 MOS from Cameroon, MHOM/CG/80/ITMAP1843 PA from the Democratic Republic of Congo, MHOM/CI/83/DAL596 TOBO and MHOM/CI/83/DAL607 ISTI from Côte d'Ivoire (Mehlitz *et al.*, 1982). These isolates were grown in ICR mice and analysed *ex vivo*. The *T. b. rhodesiense* strain Baganzi was isolated from a human in South-Eastern Uganda in 1990 and maintained in the supplemented HMI9 medium with 20% foetal bovine serum. STIB386S, STIB386R and Baganzi lines were regularly assayed for human serum sensitivity to ensure their phenotypes had not varied.

2.2 Côte d'Ivoire Isolate library

The library of *T. brucei* strains used for population analysis was collected from Daloa, Côte d'Ivoire by D. Mehlitz, Freie Universität Berlin. The parasites were collected from several hosts, including human and non-human animals between 1978 and 1983 (**Table 6.1**). Samples were either provided as purified DNA from A. Tait, University of Glasgow or blood spots on Whatman FTA cards from W. Gibson, University of Bristol. DNA from the FTA cards was washed and prepared as per the manufacturer's instructions before analysis (Whatman). The collection is comprised of both group 1 and 2 *T. b. gambiense* and *T. b. brucei* isolated from the field then cloned and passaged in rodents to create stabilates. The different sub-species designations were determined using a 5 rodent BIIT test (Rickman & Robson, 1970; Targett & Wilson, 1973) and the host from which the strain was isolated. Isolates that showed 5 successful rodent infections after exposure to human serum were designated group 1 *T. b. gambiense*, 0 infections after exposure were labelled as *T. b. brucei* and any that gave an intermediate infection number were termed group 2 *T. b. gambiense*. In addition to the main sample library, the group *1 T. b. gambiense* genome strain DAL927 (DNA gifted by W. Gibson, University of Bristol), the *T. b. brucei* strain STIB247, the Côte d'Ivoire group 2 *T. b gambiense* strain STIB386 and group 1 *T. b. gambiense* strain Eliane were also genotyped.

2.3 Genotyping

Samples were genotyped for population analysis using nested PCR for seven microsatellite markers; Ch1/18, Ch1/D2/7, Ch2/PLC, Ch2/5, Ch5/JS2, Ch11/110 and Ch4/M12C12 (MacLeod *et al.*, 2005b; Morrison *et al.*, 2008) Primer sets are named with the convention that C & D are the first round combination and A & B are the second (**Table 2.1**). The Ch4/M12C12 marker proved to be monomorphic for these samples and was excluded. Markers on the same chromosome are at opposite ends of the chromosome to each other so are unlikely to be linked (MacLeod *et al.*, 2005b). All of the primers used to amplify the microsatellites have been previously identified and verified for both *T. b. brucei* and *T. b. gambiense* populations (MacLeod *et al.*, 2005b; Morrison *et al.*, 2008). The PCR conditions for both rounds for each the reaction were 95˚C 50 secs, 55˚C 50 secs and 65˚C 60 secs, for 30 cycles. PCR was performed in a total volume of 30µl with the primers at a final concentration of 10µM each and a Taq polymerase concentration of 0.25 units/µl (Thermo Scientific). To determine allele size, one of the second round primers was tagged using either AFAM or HEX fluorescent dye to be used for Genescan capillary analysis. When run on a capillary system against ROX labelled size standards (Dundee Sequencing Unit), the size of

the tagged PCR product can be accurately measured to within 2bp using the Peak Scanner® software package (Applied Biosystems). PCR products of differing sizes were concluded to be different alleles. Alleles were separated into binning groups according to the measured size. Bins were created with the methodology that differences between sizes was accurate to within 2bp, so alleles that differed by this amount or less could not be reliably distinguished. Each distinct bin was given an allele number and the two alleles for each locus was catalogued for each isolate. This allowed the creation of a multilocus genotype for each isolate (MLG). In addition to the Genescan allele sizing technique used to create the MLG, a diagnostic PCR for presence/absence for *TgsGP* was also performed. An improved set of nested *TgsGP* primers was used (**Table 2.1**) (Morrison *et al.*, 2008; Radwanska *et al.*, 2002), using PCR conditions; 95˚C 50 secs, 55˚C 50 secs and 65˚C 120 secs, for 35 cycles. PCR was performed in a total volume of 10μl with the buffers and concentrations previously described.

2.4 Identification of sub-species by PCR

To test the presence/absence of SRA and *TgsGP* in experimental lines, primer sets SRA-F with SRA-R and *TgsGP*-A with *TgsGP*-B were used to amplify from prepared genomic DNA (Primer sequences: **Table 2.1**). The SRA primer set were amplified using the following condition with Taq polymerase; 30 cycles of 95°C for 50 seconds, 55°C for 50 seconds and 65°C for 60 seconds white the *TgsGP* primers were amplified with the following; 30 cycles of 95°C for 50 seconds, 60°C for 50 seconds and 65°C for 90 seconds. PCR was performed in a total volume of 10μl with the buffers and concentrations previously described.

2.5 Sequencing

The *HpHbR* open reading frame (ORF) and the upstream and downstream sequences containing the UTRs were amplified from genomic DNA by PCR using the following conditions with PFU polymerase; 30 cycles 95°C for 50 seconds, 55°C for 50 seconds and 65°C for 120 seconds. PCR was performed in a total volume of 20μl with the buffers and concentrations previously described. The gene and UTRs were split into two halves and sequenced separately using primers HpHbR A F with HpHbR A R and HpHbR B F with HpHbR B R (**Table 2.1**).

PCR products were ligated into the TOPO PCR 2.1 plasmid (Invitrogen) and amplified using TOP10 (Invitrogen) competent cells as per the manufacturer's instructions. Plasmid was purified using a Qiagen miniprep kit and sent for DNA sequencing at DNA Sequencing & Services (MRCPPU,

University of Dundee, www.dnaseq.co.uk) using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer. As STIB386 proved to be heterozygous for *HpHbR*, 8 clones were created and sequenced. These clones aligned into two groups with differing homologues. Homologue A contained 6 individuals and homologue B contained 2. For the five group 1 *T. b. gambiense* lines, 3 clones were used for each. Both the HpHbR ORF and UTRs were identical across all 15 sequences so this gene was assumed to by homozygous for this sub-species.

The variable region of the transferrin receptor genes (*ESAG6* and *ESAG7*) were amplified from cDNA prepared as previously described. Primer sets ESAG6F & ESAG6R and ESAG7 F & ESAG7R (Table 2.1) were used under the following PCR conditions with PFU polymerase; 30 cycles 95°C for 50 seconds, 55°C for 50 seconds and 65°C for 60 seconds. PCR was performed in a total volume of 10µl with the buffers and concentrations previously described and the products were sequenced also as described previously.

2.6 Real-time Quantitative PCR

Total trypanosome RNA was extracted from approximately 5×10^7 cells using a Qiagen RNAeasy® mini kit. The RNA was subjected to three DNase I (Invitrogen) digests to remove all genomic contamination. Omniscript® RT Kit (Qiagen) was used to generate cDNA from 1µg of total RNA as per the manufacturer's instructions. Real-time-PCR was performed using cDNA from an equivalent of 50ng of total RNA, 30µM forward and reverse primers (HpHbR RT F and HpHbR RT R (Table 2.1)), 12.5µl SYBR green PCR master mix (Applied Biosystems) to a final volume of 25µl. Primer sequences are given in table 2.1. Real-time PCR conditions were: one cycle of 50°C for 2 mins, 95°C for 10 mins, followed by 40 cycles of 95°C for 15s, 60°C for 1 min. The relative amounts of specific cDNA between samples were calculated using C_T methodology (Livak & Schmittgen, 2001) calculated with the Applied Biosystems SDS v1.4 software. The endogenous control gene was GPI8 using primers GPI8F and GPI8R primers (Table 2.1). Four replicates were performed for each parasite line. All primers were designed using the Primer3 software (Rozen & Skaletsky, 2000). Mean levels of expression were compared using 1-way ANOVA and a Tukey's post hoc test.

2.7 RT-PCR

Total trypanosome RNA was extracted from approximately 5×10^7 cells and used to generate cDNA as previously described. Semi-quantitative RT PCR was performed using cDNA from an equivalent

of 50ng of total RNA amplified using the following condition with Taq polymerase; 28 cycles of 95°C for 50 seconds, 55°C for 50 seconds and 65°C for 60 seconds. PCR was performed in a total volume of 10µl with the buffers and concentrations previously described. The PCR was then run out on a 3% agarose gel for approximately 1.5 hours and visualised under ultraviolet illumination. The primer pairs used to estimate the relative transcript of each gene in the polycistron containing *HpHbR* are listed in Table 2.1. A loading control was performed using primers specific to the gene *Triose phosphate isomerase* (*TIM*).

2.8 ApoL1 ORF Plasmid

An Invitrogen Gateway®-compatible entry vector containing the ApoL1 ORF was purchased from Genecopoeia Inc, USA. To confirm the presence of a correct ApoL1 ORF, the plasmid was sent for DNA sequencing using T7 and T7Rev standard primers at DNA Sequencing & Services (MRCPPU, University of Dundee, www.dnaseq.co.uk) using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer.

2.9 Wheatgerm Cell-free Protein Expression

The Endext® Wheatgerm expression PCR kit was purchased from CellFree Sciences, Japan. This *in vitro* expression system uses crude extracts of lysed *Triticum* cells containing essential elements of the molecular machinery involved in protein translation including the ribosomes, tRNAs, aminoacyl-tRNA synthetases, initiation, elongation and termination factors. Initiation of translation in this system requires the ORF of interest be preceded by a proprietary expression promoter sequence. This can be added to the ORF via two separate rounds of PCR. The first round uses a primer consisting of ccacccaccaccacca followed by the first 15 bases of the ORF and the second adds the full promoter sequence using ccacccaccaccacca as the second reverse primer target for the proprietary second round primers. PCR conditions were supplied by the manufacturer. If the PCR method to add the initiation sequence is unsuitable, a plasmid is also available that contains the same promoter sequence preceding an insertion zone that contains sites for various enzymes. The ApoL1 ORF was removed from the Gateway® entry vector by incubating 1mg of uncut plasmid overnight with SpeI and SmaI enzymes. The cut plasmid was then run on a 1% agarose gel, visualised with ultraviolet light and the ApoL1 ORF band extracted using a Qiagen Gel Extraction Kit. This product was ligated into the Endext® wheatgerm plasmid using T4 Ligase (Invitrogen) as per the manufacturer's instructions. The plasmid was transformed into DH5α competent *E. coli* cells to be bulked up and provide enough template for the

wheatgerm expression system. After construction of either the PCR or plasmid template, it is added to the wheatgerm extract mix and the reaction proceeds as per the manufacturer's instructions. After the protein was generated, a small sample was to be removed for analysis via a Western and Coomassie gel. The rest of the reaction was to be used for protein purification under non-denaturing conditions.

2.10 Invitrogen Expressway® E. coli Cell-free Protein Expression

The Invitrogen Expressway® system is an *in vitro* cell free system that uses crude extracts of lysed *E. coli* cells. Expressway® is fully compatible with Gateway® technology and requires only a T7 promoter region upstream of the ORF in order to initiate translation. The T7 promoter is found on all Invitrogen Gateway® destination vectors so no further modification is needed once an ORF has been efficiently transferred into the destination vector. An Invitrogen compatible Gateway® entry vector containing the ApoL1 ORF (Genecoepia) was transferred into the pExp1® destination vector (Invitrogen). This plasmid was then used in an 8 hour 37°C Expressway® cell-free reaction, as per the manufacturer's instructions (Invitrogen). The reaction was fed once after two hours with further amino acids as recommended by the manufacturer. After the reaction had terminated, a small sample was removed for analysis via a Western blot and Coomassie gel. The rest of the extract was re-suspended in native binding buffer (0.5 M NaCl, 10 mM Imidazole) and used for protein purification under non-denaturing conditions.

2.11 Invitrogen Gateway® E. coli Protein Expression

The Invitrogen Gateway ® *E. coli* protein expression system is a classical live cell expression system, although with key improvements to efficiency and ease of use. An Invitrogen® compatible Gateway® entry vector containing the ApoL1 ORF (Genecoepia) was transferred into the pDest17 destination vector (Invitrogen) and transformed into BL21-AI competent *E. coli* cells. The pDest17 vector features a 6xHis-tag that precedes the ORF and can be used for purification of the protein using a nickel column. Cells were selected on 100µg/ml ampicillin agar plates overnight and colonies were checked using PCR with ApoL1 specific primers (ApoL1_F & Apol1_R (Table 2.1)) to confirm the presence of the expression vector. Transformed colonies were grown overnight in 400ml L-broth in the presence of 100µg/ml ampicillin for 8 hours and then expression was induced using 1% L-Arabidose for a further 16 hours. Cells were lysed with pH 7.8 guanidinium lysis buffer (6 M Guanidine Hydrochloride, 0.02 M Sodium Phosphate, 0.500 M NaCl) for 5 minutes and the cellular detritus removed with a 0.2 micron micro filter (Sartorius). A small

aliquot was removed for analysis with a Western blot and Coomassie stain and the remainder was used for protein purification under denaturing conditions.

2.12 Non-denaturing Protein Purification

This protein purification method used pH buffers suitable for purifying soluble protein while maintaining it in a native functional conformation. The cell-free media slurry was bound to Ni-NTA beads (Invitrogen) for one hour at pH 7.8 and then washed twice with wash buffer (0.5 M NaCl, 20 mM Imidazole). Finally, bound protein was eluted with elution buffer (0.5 M NaCl, 250 mM Imidazole). The protein was concentrated and re-suspended in a PBS buffer using 10,000MW Vivaspin columns (Sartorius). Protein purity was estimated using a Nanodrop® spectrometer (Nanodrop) and SDS-Page. A Western blot using an antibody raised against an ApoL1 peptide (Sigma-Aldrich) was used to check that the bands present were ApoL1.

2.13 Denaturing Protein Purification

This protein purification method was suitable for removing protein from insoluble inclusion bodies. Its use can create difficulties when re-folding the protein into a functional conformation but a method for re-folding ApoL1 has been published (Lecordier *et al.*, 2009). The lysed cells were bound to Ni-NTA beads (Invitrogen) for one hour at pH 7.8 and then washed twice with wash buffer (either 4M, 6M or 8 M Urea 0.02 M Sodium Phosphate, 0.5 M NaCl) at each of pH 7.8, pH 6 and pH 5.8. Finally, bound protein was eluted with wash buffer at pH 4. The protein was dialysed overnight with 0.2 M acetic acid and 0.05% Tween20 and concentrated using 10,000MW Vivaspin columns (Sartorius). Protein purity was estimated using a Nanodrop® spectrometer (Nanodrop) and SDS-Page. A Western blot using an antibody raised against an ApoL1 peptide (Sigma-Aldrich) was used to check that the bands present were ApoL1.

2.14 PBS Buffered HSR Fluorescence Death Assay

Trypanosomes were diluted to 10^7 per ml in PBS supplemented with 0.1% glucose and incubated in 25% human serum or foetal bovine serum (FBS) in a 500µl volume in a standard 24 well plate. After 4 hours the cells were re-suspended in 100µl PBS 0.1% glucose containing 0.2% Syto10 and 0.4% RedDead dye. After 15 minutes the fluorescence buffer was removed and the cells were fixed for one hour in 2.5% glutaraldehyde at 4°C. Finally, the cells were re-suspended in 50 µl PBS and the ratio of red:green emitted wavelengths measured using a fluorometer. A control group of

trypanosomes killed by immersion in 100% ethanol was used to determine the 100% death red:green fluorescence ratio and all progeny scored as a percentage of this control to indicate survival.

2.15 PBS Buffered Survival Assay

Trypanosomes were isolated from the mammalian host and either subjected to a 1 hour chilling period before transfer to the assay buffer or were re-suspended immediately. Trypanosomes were diluted to 10^6 per ml in PBS supplemented with the various factors to be investigated and incubated in 25% human serum or foetal bovine serum (FBS) in a 1ml volume in a standard 24 well plate. The numbers of cells in each well were counted with a haemocytometer for the 0 time point. The cells were incubated with 5% CO_2 at 37°C and the number of surviving trypanosomes in each well was counted by phase contrast microscopy at 4 hours to calculate percentage survival. Human serum results were normalized against the FBS controls.

2.16 HMI9 Buffered Human Serum Resistance Assay

During development of the assay, trypanosomes were isolated from a mammalian host and either subjected to a 1 hour chilling period before transfer to the assay buffer or were re-suspended immediately. Trypanosomes were diluted to 10^6 per ml in modified HMI9 and incubated in 25% human serum or foetal bovine serum (FBS) in a 1ml volume in a standard 24 well plate. The numbers of cells in each well were counted with a haemocytometer for the 0 time point. The cells were incubated with 5% CO_2 at 37°C and the number of surviving trypanosomes in each well was counted by phase contrast microscopy at several time points to calculate percentage survival. Human serum results were normalized against the FBS controls. After development of the assay, all subsequent assays on the parents and progeny were performed using a 6 hour time point.

2.17 ApoL1 Survival Assay

Trypanosomes were diluted to 10^6 per ml in modified HMI9 with 25% foetal bovine serum (FBS) in a 1ml volume in a standard 24 well plate. A dilution series of recombinant ApoL1 was formulated that would result in the desired final concentration and made up to equal volumes with protein free buffer (0.2 M acetic acid and 0.05% Tween20). A control containing equal volume of buffer was also prepared. As the concentrations of purified ApoL1 were quite high, only very small volumes of buffer and protein were needed (no more than 20µl in the 1ml reaction). It is unlikely

the Tween20 was at a physiological relevant level. Each protein volume and the control were aliquoted into different wells. The numbers of cells in each well were counted with a haemocytometer for the 0 time point. The cells were incubated with 5% CO_2 at 37°C and the number of surviving trypanosomes in each well was recorded at 24 hours to calculate percentage survival. A positive control of 25% normal human serum was ran alongside these experiments for each strain to ensure the cells reacted accordingly.

2.18 Fluorescence microscopy of TLF uptake

Trypanosomes were re-suspended in serum-free HMI9 medium at a concentration of 10^6 cells/ml and incubated in 10μg/ml of Lysotracker® (Invitrogen) and 5 μg/ml of purified human HDL that contained predominantly TLF-1 (Raper *et al.*, 1999). The presence of the lytic component of TLF-1 was confirmed with a Western blot for ApoL1. The TLF-1 was labelled with AlexaFluor®488 (Molecular Probes, Invitrogen) using the manufacturer's instructions. In control experiments, bovine HDL of a comparable density to TLF (Raper *et al.*, 1999) was also AlexaFluor tagged. The cells were incubated at 37°C for 30 minutes, 1 hour, 2 hours and 4 hours. At each time point, cells were washed once in serum-free HMI9 medium and fixed by immersion in chilled 2.5% glutaraldehyde (Sigma-Aldrich) in phosphate buffered saline for 5 minutes. The cells were re-suspended in 50% glycerol, 0.1% DAPI, 2.5% DABCO in PBS and spread onto lysine coated slides which were then protected with cover-slips sealed using ethyl acetate. The slides were imaged using a Deltavision Core system and SoftWorx package (Applied Precision) with the standard filter sets (DAPI/FITC/Texas-Red and Light transmission). Approximately 30 serial sections through each trypanosome were taken for each filter. The images were composited and the brightness, contrast and colour levels normalized between samples and exposures using the ImageJ software package (US National Institute of Health). Approximately thirty trypanosomes were imaged per time point to give an indication of the uptake in the population - these were selected at random from random fields of view. The Pearson's correlation co-efficient between the TLF and Lysotracker® fields was calculated using the Pearson-Spearman Correlation (PSC) Plug-in for ImageJ (http://www.cpib.ac.uk/~afrench/coloc.html).

2.19 Fluorescence microscopy of ApoL1 uptake

Trypanosomes were re-suspended in HMI9 medium containing 20% foetal bovine serum at a concentration of 10^6 cells/ml and incubated in 10μg/ml of Lysotracker® (Invitrogen) and several concentrations of purified recombinant ApoL1. The cells were incubated at 37°C for 4 hours. After

this period, cells were washed once in serum-free HMI9 medium and fixed by immersion in chilled 2.5% glutaraldehyde (Sigma-Aldrich) in phosphate buffered saline for 5 minutes. The cells were washed once more in chilled PBS and then re-suspended in PBS with an AlexaFluor®488 antibody raised against a 6xHis-Tag (Molecular Probes, Invitrogen) The cells were gently agitated for 1 hour and then washed twice with chilled PBS and finally re-suspended in 50% glycerol, 0.1% DAPI, 2.5% DABCO in PBS and spread onto lysine coated slides which were then protected with cover-slips sealed using ethyl acetate. The slides were imaged using a Deltavision Core system and SoftWorx package (Applied Precision) with the standard filter sets (DAPI/FITC/Texas-Red and Light transmission). Approximately 30 serial sections through each trypanosome were taken for each filter. The images were composited and the brightness, contrast and colour levels normalized between samples and exposures using the ImageJ software package (US National Institute of Health). Approximately fifteen trypanosomes were imaged per time point to give an indication of the uptake in the population - these were selected at random from random fields of view. The Pearson's correlation co-efficient between the ApoL1 and Lysotracker® fields was calculated as previously described.

2.20 *HpHbR* Phylogeny Analysis

The phylogeny of the *HpHbR* gene between isolates was assessed using the Seaview 4 interface for PHYLIP's v3.52 dnapars program (Felsenstein, 1989). A parsimony tree was built using aligned DNA sequences and a neighbour joining algorithm, bootstrapped with 100 replicates for Eliane, STIB386, the *T. b. brucei* genome reference strain TREU927 and the Tanzanian *T. b. brucei* STIB247

2.21 Quantitative trait loci (QTL) analysis

Twenty-five STIB247/STIB386 progeny created during the formulation of the group 2 *T. b. gambiense* genetic map were used for QTL analysis (Cooper *et al.*, 2008). All strains were grown initially in ICR mice and analysed *ex vivo* in a 6 hour HMI9 buffered survival assay. The inheritance of the human serum resistance trait in the progeny from the *T. b. brucei* strain STIB247 and group 2 *T. b. gambiense* strain STIB386 genetic cross was examined for co-segregation of genetic markers on the genetic map of the group 2 *T. b. gambiense* strain STIB386 (Cooper *et al.*, 2008). A quantitative trait analysis (QTL) by maximum likelihood was conducted using MapManager QTX software (Manly *et al.*, 2001) employing a significance level based on the χ^2 statistic. Genome-wide significance (P<0.001) was determined by performing 1000 permutations. Interval mapping

on each chromosome was performed using 1cM increments. QTL analysis was also repeated using linear regression by QTL express (Seaton *et al.*, 2002).

2.22 Population Analysis

The multi-locus genotype (MLG) of each isolate studied was used to create a dendrogram to infer relationship within the Côte d'Ivoire *T. brucei* population. Clustering calculator (http://www.biology.ualberta.ca/jbrzusto/cluster.php) was used to create a Phylip Drawtree string (analysed using un-weighted arithmetic average clustering and Jaccards's similarity coefficient). A bootstrap with 100 iterations was also generated. This Drawtree string was then converted into a dendrogram using Treeview software (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html). Heterozygosity, Nei's genetic distance and F_{st} were calculated using the GenAlEx software package for Microsoft Excel (http://www.anu.edu.au/BoZo/GenAlEx/). Nei's genetic distance and F_{ST} both compare the proportion of the total genetic variance contained in a subpopulation relative to the total genetic variance (Nei & Roychoudhury, 1974; Wright, 1978) These statistics indicate how similar populations are to each other, with 1 being completely different and 0 being identical. Hardy-Weinberg calculations were also performed using GenAlEx and indicate if alleles in the population are in an ideal Hardy-Weinberg state, usually as a result of completely random mating. Linkage disequilibrium calculations were performed using GenePop v4.0 software (http://genepop.curtin.edu.au/) and reveal if the same alleles in a population are usually inherited together or whether the inheritance of alleles is completely random. High levels of mating will cause alleles to be inherited in an apparently randomly assorted manner and will exhibit linkage equilibrium, while a population with little mating or strong selection will have similar alleles inherited together and exhibit linkage disequilibrium. Finally, a principal coordinate analysis (PCA) was performed on the three populations using genetic distance to generate the vectors and the GenAlEx software package for Microsoft Excel (http://www.anu.edu.au/BoZo/GenAlEx/). This analysis method performs a multi-dimensional comparison using genetic distance between individuals for both axes. The factor that causes the greatest genetic distance between individuals in the populations is used for one axis, and the feature that causes the next most distance is the other. In this way the structure of a population can be visualised.

2.23 General Statistics

Several statistical tests were performed during this thesis, all using the Minitab Software package (Minitab Solutions Ltd). One-way analysis of variance (ANOVA) tests were performed using the default settings with a Post-hoc Tukey's test to infer which sample variances statically differed from each other. General linear models (GLM) were constructed using percentage survival as the response factor and trypanosome strains and experimental conditions as the input parameters.

2.24 Solutions

Modified HMI9 Medium (500ml)
365ml Iscoves modified Dulbecco's medium + glutamax
5ml Hypoxanthine (store @ 4oC)
1.5ml Kanamycin (10mg/ml) (store@ -20oC)
5ml (store @ -20oC) Penicillin / Streptomycin (5000u pen +5000u strep)
25ml βBCPT
500mg Glucose
67mg Adenosine
71mg Guanosine
0.55g Methyl cellulose

Add above ingredients and leave overnight to dissolve on a stirrer then add heat-inactivated, sterile-filtered foetal calf serum to a concentration of 20% (v/v) to medium before use.

Phosphate Buffered Saline (PBS) (Qiagen Standard formulation)
137 mM NaCl
2.70 mM KCl
4.30 mM Na_2HPO_4
1.47 mM KH_2PO_4

The final solution was adjusted to pH 7.4 before use in lysis assays wash steps.

2.25 Primer Sequences

Table 2.1 The primer sequences used throughout the thesis

Primer	Sequence
Tb927.6.350_F	TTACATAATCGAGCATCCCGC
Tb927.6.350_R	ATGACGATTCAAGCCGTATCG
Tb927.6.360_F	TCACGAAGCCACTACATGCGT
Tb927.6.360_R	ATGCTACTTCCCCTCTACCTA
Tb927.6.370/380_F	TCAGCACACAATTGCGAGGGA
Tb927.6.370/380_R	ATGGGCACTGAGATAAGCATG
Tb927.6.400_F	TCACTGCTTATGTATCTTTGG
Tb927.6.400_R	ATGCAGGTTGATTGGGGTGCA
Tb927.6.410_F	TCAGCATCTGTAACGAACTAC
Tb927.6.410_R	ATGACACAACCAGATATATTC
Tb927.6.430_F	TATGCCTCATCCAGGGCCTGA
Tb927.6.430_R	GAAAACAATTGCGTTATCCAC
Tb927.6.440_F	CTACACCACCACCTGGAGCATAC
Tb927.6.440_R	AGAAACCGTCTTGCAGG
Tb927.6.560_F	CTACGCCGTGTTGGGTGCAAG
Tb927.6.560_R	ATGCATCTCATGGTGCCTGC
Tb927.6.740_F	CGGATCCAAAACACCAAAACA
Tb927.6.740_R	ATGAACCAGAAGCGATGCGAA
HpHbR RT F	GCCCTATGCTTATGCACATGATC
HpHbR RT R	ACCTCCGCCAGAGAAAATCTC
HpHbR A F	ACAAAGTGGCAGGTGCGTTG
HpHbR A R	ATT TTC GAT CGG GTT CCC AT
HpHbR B F	ATG GGA ACC CGA TCG AAA AT
HpHbR B R	AATCAGTTTTTTAGGGCGGC
ESAG6 F	CCGGAATTCGCTATTATTAGAACAGTTTCT
ESAG6 R	GTGTTAAAATATATC
ESAG7 F	CCGGAATTCGCTATTATTAGAACAGTTTCT
ESAG7 R	GCTCTAGACATCACTGCATTTTTTGCTTC
SRA F	GACAACAAGTACCTTGGCGC
SRA R	CAGCAACCATATTCAGAGCC
TgsGP F	TCACGGCCATCAGACGGAGA
TgsGP R	GCCATCGTGCTTGCCGCTC
JS2 F	GATTGGCGCAACAACTTTCACATACG
JS2 R	CCCTTTCTTCCTTGGCCATTGTTTTACTAT
ApoL1_F	AGGCAGATGAGCTCCGTAAA
Apol1_R	GCCAGAGGAAATGCTGAGAG
Ch1/D2/7-C	ACATTTTGGTTGTCCTGTTGG
Ch1/D2/7-D	GATCGAAGATAATAAATGCACAT
Ch1/D2/7-A	GTTGGCCGCATTATTCGATGC
Ch1/D2/7-B	GACAAGTAACACACAGGTGCACCA

Table 2.1 continued

Ch5/JS2-A	GATTGGCGCAACAACTTTCACATACG
Ch5/JS2-B	CCCTTTCTTCCTTGGCCATTGTTTTACTAT
Ch5/JS2-C	AGTAATGGGAATGAGCGTCACCAG
Ch5/JS2-D	GATCTTCGCTTACACAAGCGGTA
Ch11/110-A	GASTGCGAGGATTATGACTGTAGCG
Ch11/110-B	CCAATCTTATGCATACATGCAAGC
Ch11/110-C	TTTCTACGTCTCATTTCAACG
Ch11/110-D	ACTGCACAAACGGGTAAGGAG
CH11/51-A	AACCGATCATTCCTGTTCC
CH11/51-B	TGAGATGGTACTTGAAGAAAG
ch11/51-C	CTTACCCACAGGGCCAAA
ch11/51-D	AATGATTACGCACAAGCACA
CH1/18-A	TGTGAGAATGGTACTCACGCGCTG
CH1/18-B	ACAACGTTAGCACACAATTCCTGTG
CH1/18-C	TATAATGCGTTTGTGAGAAT
Ch1/18-D	GAAGGGAGGGAACAGAAGCAGGG
Ch2/PLC-A	TTAAGTGGACGACGAAATAACAACA
Ch2/PLC-B	CCACTGACCTTTCATTTGATCGCTTTC
Ch2/PLC-C	TTCAAACACCGTCCCCCTCAATAAT
Ch2/PLC-D	CAACGACGTTGGAAGAGTGTGAAC
TIM-F	TGCCGTTGAGTGGGTGAAGATAGC
TIM-R	CTCCCTGCTACCTGTCTTTACATC
TgsGP-A	TCACGGCCATCAGACGGAGA
TgsGP-B	GGGCTCCTGCCTCAATTGCTGCA
TgsGP-C	GTGGCAATTACTAGCAATAGCG
TgsGP-D	GCCATCGTGCTTGCCGCTC

Chapter 3: TLF-1 interactions with *T. b. gambiense* and HpHbR

3.1 Introduction

Humans possess two trypanosome lytic factors (TLF-1 & TLF-2) that can kill most trypanosome species. The human infective sleeping sickness parasites *T. b. rhodesiense* and *T. b. gambiense* are able to resist lysis by human serum and establish bloodstream infections. How these sub-species have achieved this has been researched for several years, with a particular focus on *T. b. rhodesiense* due to its ease of use in experimental procedures. One proposed mechanism for human serum resistance for *T. b. rhodesiense* stemmed from observations that a *T. b. rhodesiense* line failed to transport TLF to the lysosome and it was believed that *T. b. rhodesiense* employed a TLF avoidance strategy to resist lysis (Hager & Hajduk, 1997). However, a more widely employed resistance mechanism utilised by the majority of *T. b. rhodesiense* lines has since been described which involves the expression of a serum resistance associated (SRA) protein (De Greef & Hamers, 1994; Xong *et al.*, 1998). The *SRA* gene is found in a bloodstream expression site and its expression is therefore is closely linked to antigenic variation. When the expression site with *SRA* is activated the sub-species is able to resist lysis by human serum. However if the *SRA* containing expression site is silenced the parasite is sensitive and exhibits a lysis phenotype indistinguishable from *T. b. brucei*. This explains the observation that *T. b. rhodesiense* parasites possess a variable human serum resistance phenotype. SRA is postulated to bind to the ApoL-1 component of TLF-1 in the early endosomes as it is trafficked to the lysosome, neutralising it and preventing lysis (Oli *et al.*, 2006; Vanhamme *et al.*, 2003). *SRA* has been found to be present in the majority of *T. b. rhodesiense* isolates but absent in all strains of *T. b. brucei* (Picozzi *et al.*, 2005; Radwanska *et al.*, 2002; Welburn *et al.*, 2001). However, while it is clear that *SRA* is responsible for human infectivity in a large number of *T. b. rhodesiense* isolates; there are a small number of human serum resistant *T. b. rhodesiense* which have been shown to lack *SRA* (De Greef *et al.*, 1989; Enyaru *et al.*, 2006).

More importantly, the *SRA* gene has not been found in any isolates of the most prevalent human infective trypanosome, *T. b. gambiense* (De Greef *et al.*, 1992; De Greef *et al.*, 1989; Gibson *et al.*, 2002). This demonstrates that an alternative mechanism of human serum resistance exists in this sub-species. An additional complicating factor is that *T. b. gambiense* parasites fall into two

distinct "groups" that differ in genotype and phenotype. Isoenzyme data show that group 1 and group 2 *T. b. gambiense* populations are reliably distinguishable from each other and that group 2 is more akin to *T. b. brucei* than group 1 (Gibson *et al.*, 1980; Godfrey & Kilgour, 1976; Mehlitz *et al.*, 1982; Tait *et al.*, 1984; Zillmann *et al.*, 1984). Group 2 *T. b. gambiense* exhibit a variable human serum resistance phenotype using the BIIT in a manner superficially similar to *T. b. rhodesiense*, while group 1 *T. b. gambiense* strains are invariably resistant even after prolonged passage in laboratory rodents (Mehlitz *et al.*, 1982; Zillmann *et al.*, 1984). *In vivo* studies with the group 2 *T. b. gambiense* TxTat strain did find a correlation between changes in the VAT and human serum resistance, indicating that specific VATs were always found within resistant populations after rodent passage (Ortiz *et al.*, 1994). Additional *ex vivo* work with the same parasite line showed that the human serum sensitive group 2 *T. b. gambiense* line depleted human serum of its lytic potential faster than resistant lines derived from the same isolate (Ortiz-Ordóñez & Seed, 1995). This suggests that this resistant group 2 *T. b. gambiense* strain is depleting TLF from the media more slowly than isogenic sensitive lines. It has come to light however that there is some confusion as to the exact identity of some TxTat lines, with some derivatives revealed to be *T. b. rhodesiense* (Grab & Kennedy, 2008). Lineage data for the TxTat used in the *ex vivo* assays was unavailable so it may be possible it has been misidentified as a type 2 *T. b. gambiense*. Even less research has been performed concerning human serum resistance in the more important group 1 *T. b. gambiense*.

There are a number of possible approaches to resist lysis by human serum that *T. b. gambiense* could utilise. One such strategy involves neutralisation of the active components of the TLF particle. This is the tactic employed by most *T. b. rhodesiense* strains using SRA to inhibit ApoL-1. Another strategy would be to inhibit uptake of the lytic particles. Manipulation of the uptake of lytic particles in trypanosomes has been shown to be a viable resistance mechanism in artificially generated systems. For example, knocking-down the HpHbR receptor involved in TLF-1 uptake conferred partial human serum resistance to *T. b. brucei* (Vanhollebeke *et al.*, 2008). Parallel to this, a *T. b. brucei* line generated by *in vitro* selection with lytic human HDL showed a failure to take up TLF-1 due to down-regulation of expression for *HpHbR*, a TLF-1 receptor (Kieft *et al.*, 2010). It is possible that the resistance mechanism utilised by *T. b. gambiense* may also involve TLF avoidance rather than neutralisation, perhaps by modification of the expression or function of this gene. However, while knocking-down expression of the HpHbR receptor in *T. b. brucei* conferred full resistance to TLF-1, it had no effect on resistance to TLF-2 and little effect on

resistance to normal human serum (Vanhollebeke *et al.*, 2008). While relevant, this gene is far from the whole story.

This chapter aims to characterise features of the human serum resistance mechanism that have evolved in *T. b. gambiense* and in what ways they may differ between the two *T. b. gambiense* groups and compared with *T. b. rhodesiense*. The chapter will focus on the interactions of the predominant lytic particle TLF-1 with different parasite lines and an investigation of the HpHbR receptor. This work has been facilitated by the development of culture conditions that allow *T. b. gambiense* lines to be grown *in vitro* with stable expression of either human serum sensitivity or resistance.

3.2 Results

3.2.1 Identification of trypanosome lines

The aim of this series of studies was to investigate the nature of human serum resistance in the different *T. b. gambiense* groups in relation to other sub-species. Representative isolates from each sub-species, STIB247 (*T. b. brucei*), Eliane (group 1 *T. b. gambiense*), STIB386 (group 2 *T. b. gambiense*) and Baganzi (*T. b. rhodesiense*) were examined for the presence of the sub-species specific genes *SRA* (*T. b. rhodesiense* specific (Turner *et al.*, 2004)) and *TgsGP* (group 1 *T. b. gambiense* specific (Radwanska *et al.*, 2002)) by PCR amplification in order to verify their sub-species classification (**Figure 3.1**). The *SRA* gene was found to be present in the *T. b. rhodesiense* Baganzi strain and the *TgsGP* gene was present in the group 1 *T. b. gambiense* Elaine strain, confirming their sub-species status. The strains *T. b. brucei* STIB247 and group 2 *T. b. gambiense* STIB386 do not possess either gene (**Figure 3.1**) but can be distinguished based on human serum resistance phenotype (Turner *et al.*, 2004).
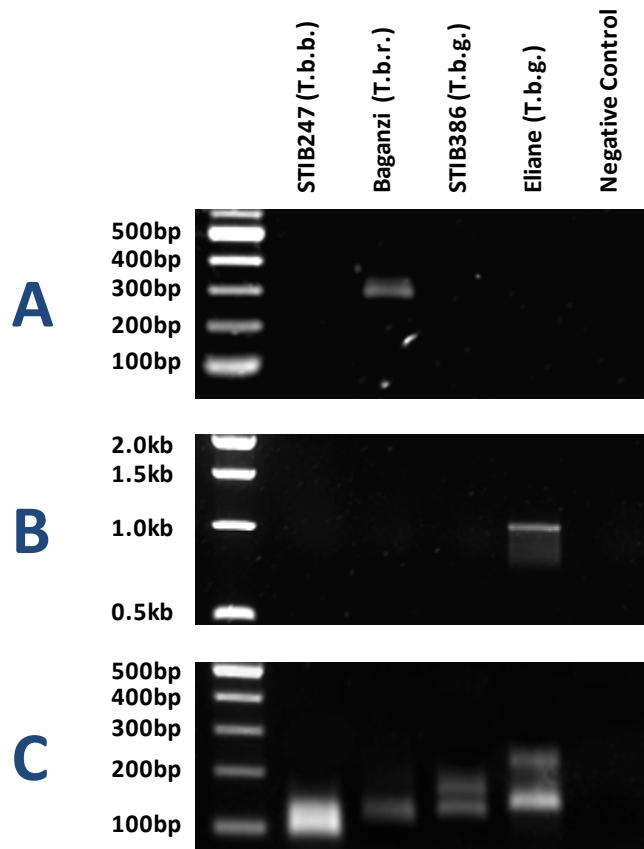
Figure 3.1 2% agarose gel showing presence/absence in various trypanosome strains of (A) *SRA*, (B) *TgsGP* and (C) microsatellite marker JS2 as a template control. A 1kb ladder (NEB) was used for *SRA* & JS2 and 100bp (NEB) used for *TgsGP*. PCR conditions & primer sequences are given in **Chapter 2: Materials & methods**.

3.2.2 Human Serum Resistance Phenotypes

Group 2 *T. b. gambiense* has a variable human serum resistance phenotype and different populations of STIB386 contain representatives of both sensitive and resistance forms in varying ratios (**Figure 3.2a**). By culturing the group 2 strain STIB386 in the presence of 10% human serum in an *in vitro* system we selected a population of human serum resistant parasites (STIB386R; **Figure 3.2b**). Several sub-lines of STIB386 gave a range of human serum resistance phenotypes of which one sensitive line was chosen that was consistently sensitive to human serum in continuous culture (STIB386S; **Figure 3.2b**). The group 1 *T. b. gambiense* strain Eliane was adapted to the culture medium. After a period of 2 weeks low growth it grew well and was consistently resistant to human serum. *T. b. brucei* STIB247 had been previously adapted to culture media and was always sensitive (**Figure 3.2**).
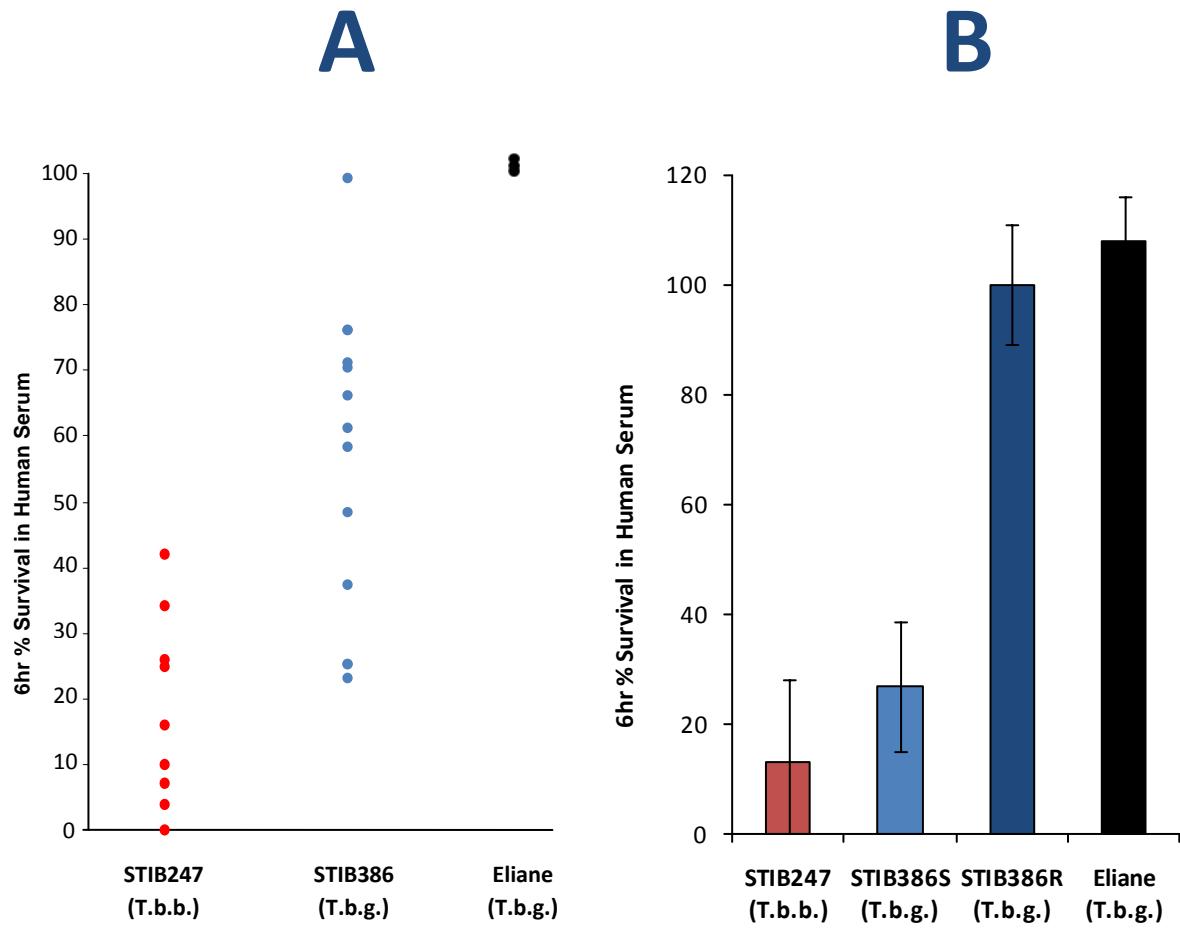
Figure 3.2a Percentage survival of several replicates of unselected *ex vivo* parasite lines exposed to human serum for 6 hours normalised against trypanosomes exposed to foetal bovine serum.

Figure 3.2b Mean percentage survival of *in vitro* culture-adapted parasite lines exposed to human serum for 6 hours, normalised against trypanosomes exposed to foetal bovine serum (n=4 ± standard error). Both sensitive and resistant clones of type 2 *T. b. gambiense* are shown, in comparison with culture adapted *T. b. brucei* (STIB247) and group 1 *T. b. gambiense* (Eliane) lines.

3.2.3 Group 2 *T. b. gambiense* bloodstream expression site analysis

Previous *in vivo* studies indicate that the switching of resistance in one group 2 *T. b. gambiense* strain correlated with a change in expression of the variable antigen type (VAT) implying that group 2 *T. b. gambiense* human serum resistance may be associated with expression site (ES) switching in a similar manner to *T. b. rhodesiense* (Ortiz *et al.*, 1994). In order to determine if a change in ES expression is associated with a switch in human serum resistance phenotype in STIB386, we examined the ES expressed in both sensitive and resistant lines. Different copies of the two genes, *ESAG6* and *ESAG7* are found in each expression site and are well-characterised components of a transferrin receptor. They contain polymorphisms that are unique to each expression site, which can be used as markers to reveal which ES is currently active at any one time point (Ligtenberg *et al.*, 1994; Salmon *et al.*, 1994; Steverding *et al.*, 1994; Steverding *et al.*, 1995). RT-PCR using primers for *ESAG6* and *ESAG7* was performed and the PCR product was cloned. Sequencing several clones for each *ESAG* revealed no sequence differences between the expressed ESAG 6 and 7 hyper-variable regions (**Figure 3.3**) or whole gene (**Appendix 1**) in STIB386S and STIB386R lines suggesting that there is no association between ES usage and the human serum resistant phenotype.

```
386R_ESAG6(BES126)  CATCACAGTTTTGCATAAGCACGGGGAAGACTGGGCCAGCAGAATACAACAACTTGCAAG
386S_ESAG6(BES126)  CATCACAGTTTTGCATAAGCACGGGGAAGACTGGGCCAGCAGAATACAACAACTTGCAAG

386R_ESAG7(BES126)  GGAGAATACGGTGATGTCAGTTCCCACGATGCGGTACGGTGGACCGAAGATCCTAGTAAA
386S_ESAG7(BES126)  GGAGAATACGGTGATGTCAGTTCCCACGATGCGGTACGGTGGACCGAAGATCCTAGTAAA
```

Figure 3.3 The sequence of the hyper variable regions (highlighted red) and surrounding bases of the expressed ESAG6 and ESAG7 in sensitive and resistant STIB386 derived from the same parent. The full gene sequences are presented in **Appendix 1**. Sequencing primers are given in **Chapter 2: Materials & methods**.

3.2.4 TLF-1 Uptake and Localisation

An important requirement of the lytic mechanism of TLF-1 is that it must be trafficked to the parasite lysosome and exposed to the low pH of the lysosome (Hager *et al.*, 1994). It has been shown in one resistant *T. b. rhodesiense* line that TLF is maintained in the late endosomes rather than being transported to the lysosome (Hager & Hajduk, 1997). A possible human serum resistance mechanism in *T. b. gambiense* therefore is that TLF-1 is not internalised or that it is not trafficked to the lysosome. We investigated the uptake and localisation of TLF-1 in our trypanosome lines using Alexafluor®488 tagged TLF-1 and the commercial dye Lysotracker®. *T. b. brucei, T. b. rhodesiense* and group 2 *T. b. gambiense* (both STIB386S and STIB386R) all showed internalisation of AlexaFluor labelled TLF-1 within one hour (**Figure 3.4**). No labelled TLF-1 was detected in the group 1 *T. b. gambiense* parasites, indicating a lack of TLF-1 uptake. This result has also been replicated independently using a FACS analysis method (Kieft *et al.*, 2010).
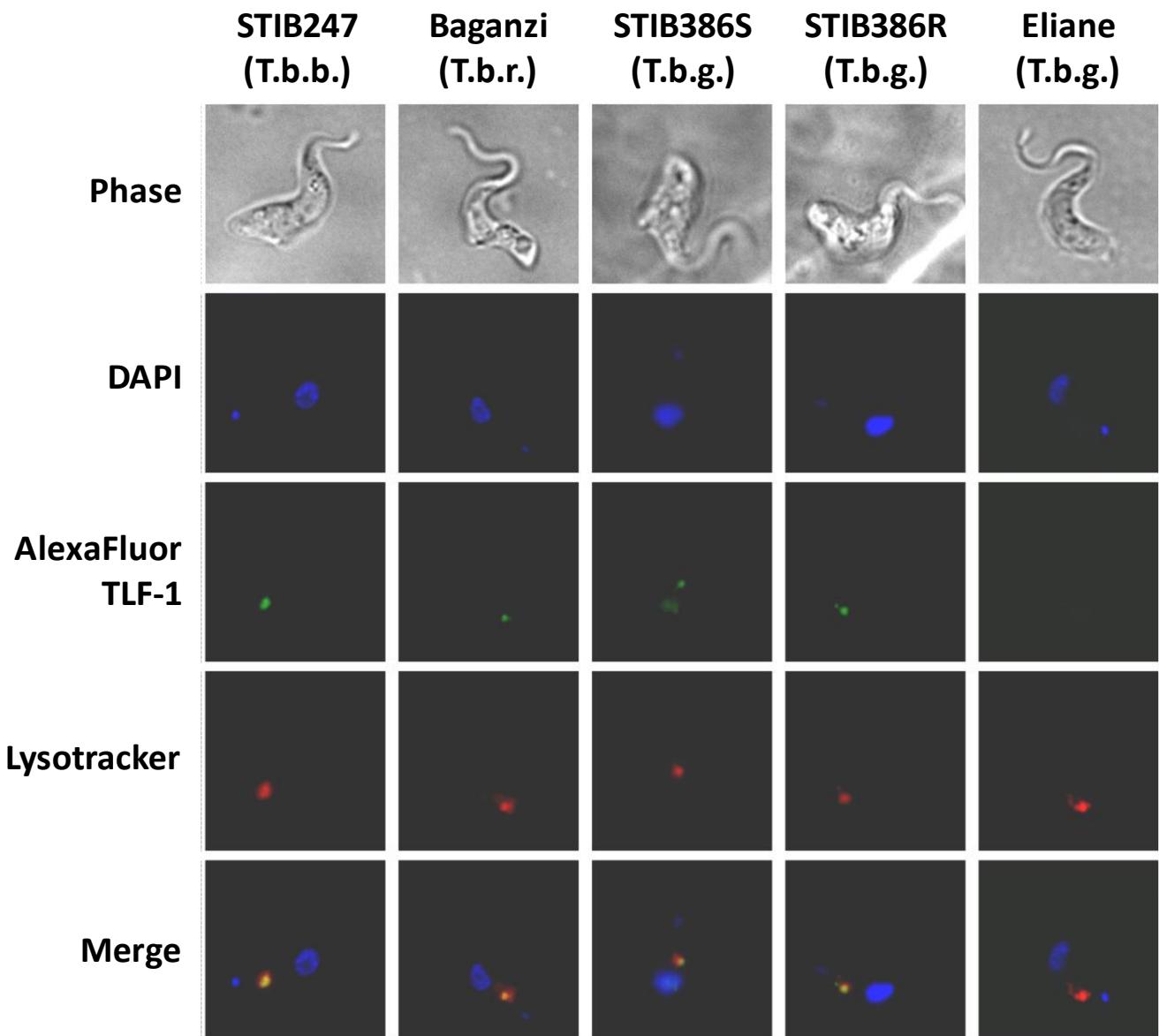
Figure 3.4 The localisation of fluorescently tagged TLF-1, Lysotracker® & DAPI in several parasite lines after one hour exposure. The mean correlation co-efficient (R) of the TLF-1 and Lysotracker® field was estimated for each parasite line: for STIB247, R=0.9 (n=14); STIB386S, R= 0.68 (n=18); STIB386R, R=0.63 (n=19); Baganzi, R=0.72 (n=28) indicating that TLF-1 was not only internalised but also trafficked to the lysosome in these strains. There was no calculated R value (n=34) for Eliane indicating no correlation between TLF-1 and the Lysotracker® dye.

3.2.5 TLF Uptake Time course

In order to investigate if there was a difference in the rate of uptake of TLF between the different strains, TLF uptake was examined over a four hour time course. Trypanosomes were counted as positive for uptake if they possessed either detectable TLF-1 within the main body of the parasite or had lysed, estimated by a percentage change in population. The majority of parasites in the *T. b. brucei*, *T. b. rhodesiense* and group 2 *T. b. gambiense* STIB386S lines all internalised TLF-1 within one hour whereas the group 2 *T. b. gambiense* STIB386R reached an asymptote of approximately 60% of the population exhibiting uptake of TLF-1 (**Figure 3.5a**). In contrast, group 1 *T. b. gambiense* showed no detectable uptake.

3.2.6 Non-lytic HDL Uptake and Localisation

The group 2 *T. b. gambiense* STIB386R strain showed a reduced degree of uptake of TLF along a time course compared to the other studied strains. In order to examine whether the reduced uptake was specific to lytic HDL suggesting changes to specific receptors, or whether the effect was due to a general down-regulation in the amount of HDL being ingested by the parasites, we analysed the uptake of bovine serum HDL of a similar size to the TLF fraction. Bovine HDL does not contain ApoL-1 or Hpr, which was confirmed using protein-specific antibodies in western blots (data not shown). All strains of *T. brucei* internalised the bovine HDL, although group 1 *T. b. gambiense* did so at a slower rate than the other sub-species (**Figure 3.5b**). This suggests that the reduced uptake of lytic HDL by the resistant group 2 *T. b. gambiense* strain is not due to a general down-regulation of uptake of all HDL but is specific to for TLF-1. In contrast, the group 1 *T. b. gambiense* strain displayed a complete lack of uptake of the lytic TLF-1 particle as well as reduced uptake of a non-lytic HDL.
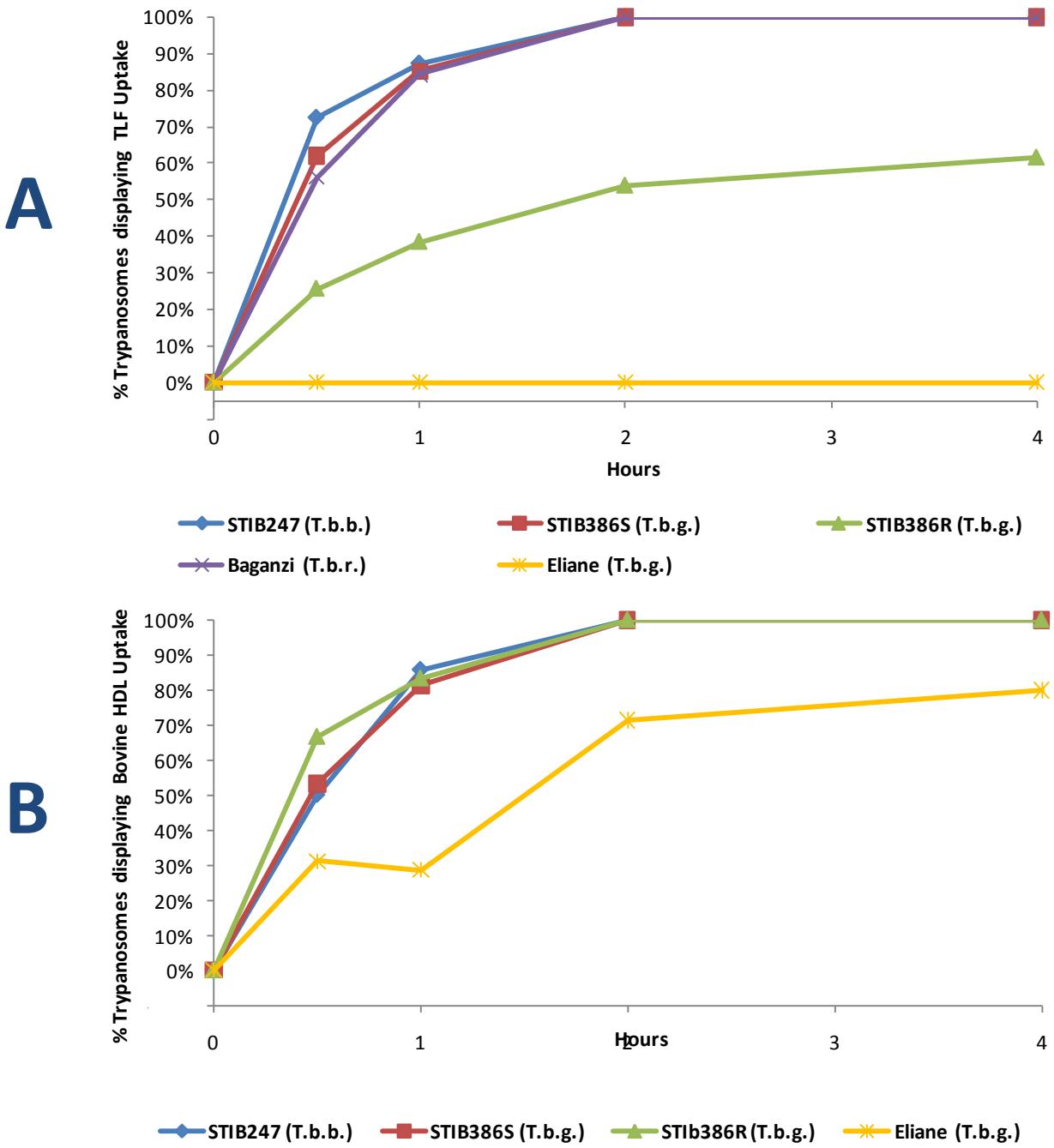
Figure 3.5a The proportion of parasites in each parasite line that exhibit uptake of TLF-1, either by the presence of visible concentrations of AlexaFluor tagged TLF-1 (Eliane, STIB386S, STIB386R and Baganzi) or a combination of percentage death and visible uptake (STIB247 and STIB386S). Cell death in these lines was attributed to the presence of TLF-1 as there was no death in cells suspended in TLF-1 free media, or B-HDL.

Figure 3.5b Uptake of bovine HDL (B-HDL) of a comparable size to TLF-1, measured by visible concentrations of AlexaFluor® tagged B-HDL in the parasite body. No death was observed in the 4 hour time course.

3.2.7 Relative HpHbR Expression

The differential uptake of TLF shown in the group 1 and group 2 *T. b. gambiense* STIB386R strains compared to *T. b. brucei* strain may be due to the variation in the levels of expression of the recently described HpHbR receptor, as a reduction of expression of this gene by RNAi or gene knockout confers some resistance to human serum in *T. b. brucei* (Vanhollebeke *et al.*, 2008). To investigate the expression of the gene in our lines the relative amount of transcript for the gene was measured using quantitative real-time PCR. STIB247, STIB386S, STIB386R and Baganzi strains all expressed similar levels of transcription for the *HpHbR* gene relative to a GPI8 endogenous control (**Figure 3.6a**). The STIB386R strain showed a slight decrease in mean level of expression of the gene compared to STIB386S but this was not statistically significant using a one way ANOVA (F_7=4.03, p = 0.091). However, the group 1 *T. b. gambiense* strain expressed a five-fold lower mean transcript level compared to the other parasite lines and comparing expression across all five strains identified significant difference (F_{18}=36.51, p < 0.01). Post-hoc Tukey's tests indicated Eliane to be different to all other lines whilst all comparisons between the other four lines were not significant.

To investigate whether the low levels of *HpHbR* expression seen in the Eliane strain was a general feature of group 1 *T. b. gambiense*, expression levels were also measured in group 1 *T. b. gambiense* isolates from the Democratic Republic of Congo, Côte d'Ivoire and Cameroon. These isolates were examined *ex vivo* and Eliane was also investigated *ex vivo* for comparison (**Figure 3.6b**). The down-regulation of transcript for the gene compared to the *T. b. brucei* STIB247 is consistent across all group 1 *T. b. gambiense* strains (F_{29}=158.97, p < 0.01). The down-regulation of transcription is more pronounced with *ex vivo* Eliane compared to the *in vitro* adapted line, however the trend is unchanged.
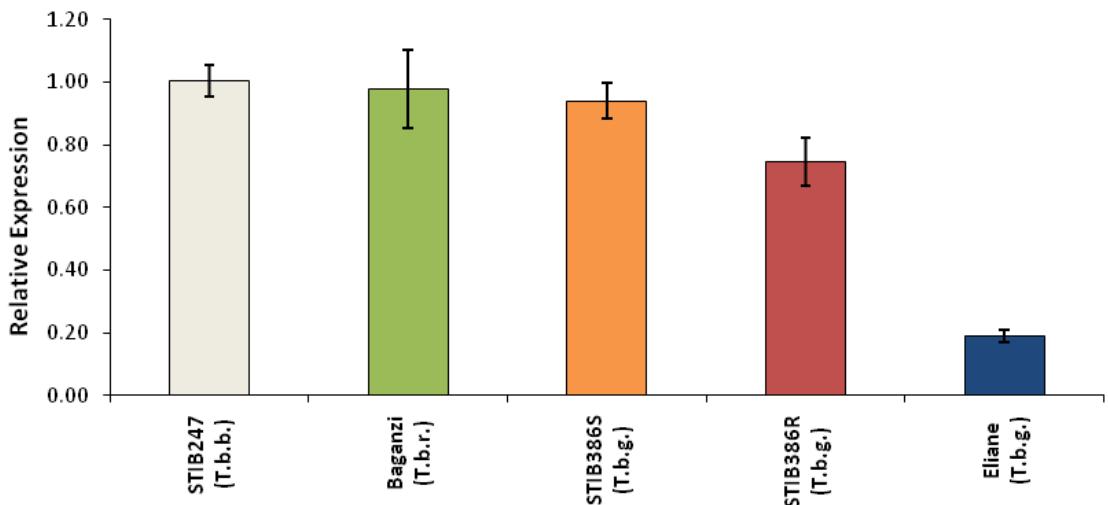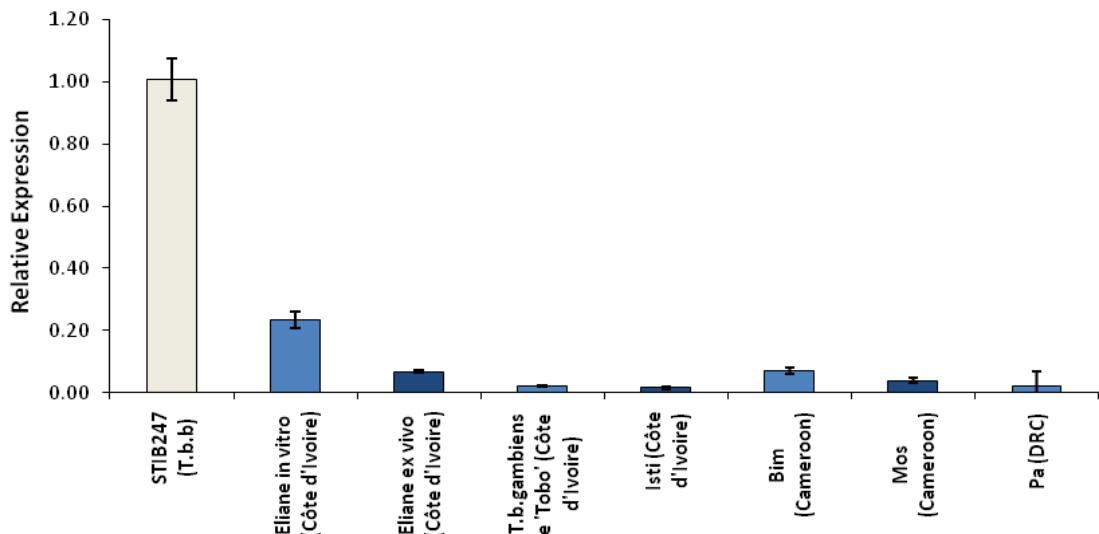
Figure 3.6a The relative abundance of transcripts for the *HpHbR* gene in each *in vitro* cell line relative to GPI8 as an endogenous control. All samples were normalised against the mean STIB247 C_T value (n = 4 for each line and standard error of the mean C_T values is shown).

Figure 3.6b The relative abundance of transcripts for the *HpHbR* gene in several group 1 *T. b. gambiense* isolates relative to GPI8 as an endogenous control. All samples were normalised against the mean *T. b. brucei* C_T value (n = 4 for each line and standard error of the mean C_T values is shown).

Real-time PCR conditions & primers are given in **Chapter 2: Materials & methods**.

3.2.8 Gene expression within the *HpHbR* polycistron

The down-regulation of *HpHbR* as a response to prolonged exposure to lytic human HDL has been demonstrated in an artificially selected human serum resistant line of *T. b. brucei* strain Lister 427 (Faulkner *et al.*, 2006). This particular line has evolved down-regulation of the expression of genes across the polycistron containing *HpHbR* and the neighbouring polycistron (Kieft 2009, unpublished). The human serum resistant line of Lister 427 therefore acts as an *HpHbR* null mutant. To investigate whether a similar chromosome regional down-regulation had occurred in a group 1 *T. b. gambiense* strain, RT-PCR was performed for genes along the polycistron containing *HpHbR* for the group 1 *T. b. gambiense* Eliane in comparison with Lister 427 and the Lister 427 *HpHbR* null mutant (**Figure 3.7**). Unlike the Lister 427 null mutant, *HpHbR* is down regulated independently of neighbouring genes in the Eliane group 1 *T. b. gambiense* strain. This implies that local modification in the 3'UTR of the *HpHbR* gene that affects the amount of transcript is more likely to be the determining factor in the low levels of *HbHbR* expression than a regional down-regulation as observed in the serum resistant Lister 427.
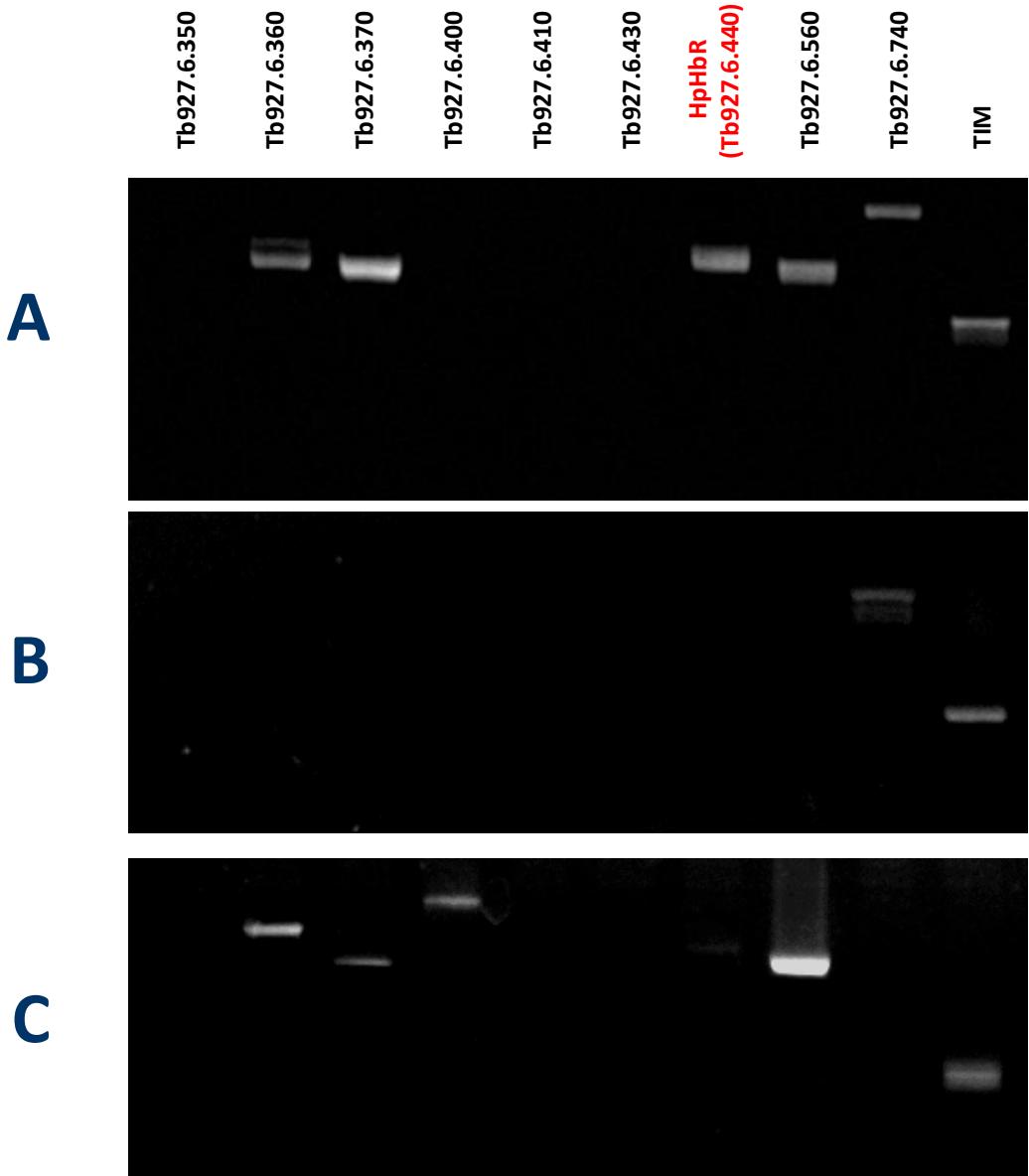
Figure 3.7 2% agarose gels showing RT PCR of the approximate relative amount of transcript for genes within the same polycistron as *HpHbR* (Tb927.6.350 - Tb927.6.430) and a neighbouring polycistron (Tb927.6.560 & Tb927.6.740) for (A) *T. b. brucei* L427, (B) an artificially selected human serum resistant L427 and (C) group 1 *T. b. gambiense* Eliane. A TIM loading control is also shown. PCR conditions and primer sequences are given in **Chapter 2: Materials & methods**.

3.2.9 *HpHbR* ORF & 3'UTR sequence

The *HpHbR* gene has been shown to be closely involved in the uptake of TLF-1 by *T. brucei* (Vanhollebeke *et al.*, 2008). We have shown that this gene is down-regulated in group 1 *T. b. gambiense*. In addition, parallel work using the *HpHbR* null mutant Lister 427 line has indicated that the group 1 *T. b. gambiense* version of the gene has lowered functionality (Kieft *et al.*, 2010). Ectopically expressing the *T. b. brucei* version of *HpHbR* in the null mutant restores full human serum sensitivity to the line. However, ectopically expressing the group 1 *T. b gambiense* version of the gene does not result in a rescue of the sensitivity phenotype (Kieft *et al.*, 2010). The open reading frame (ORF) and 3' UTR of the *HpHbR* gene were therefore sequenced to determine if there were any group 1 *T. b. gambiense* specific polymorphisms that might correlate with the differences in activity and expression. Modifications to the ORF may indicate mutations that affect function and changes in the 3' UTR may reveal features that affect gene expression due to the fact that *T. brucei* gene expression is largely regulated by mRNA stability mediated by the 3' UTR (Clayton & Shapira, 2007). The group 2 *T. b. gambiense* strain STIB386 both expresses the *HpHbR* gene and takes up TLF-1, indicating it has functional alleles of both the ORF and 3' UTR. STIB386 is heterozygous for the gene and 3' UTR, while the group 1 *T. b. gambiense* strain Eliane and *T. b. brucei* strain TREU927 are both homozygous for the ORF and 3' UTR (**Appendix 2**). The gene is very highly conserved in group 1 *T. b. gambiense*, with individuals from disease foci in Cameroon, Democratic Republic of Congo and Côte d'Ivoire possessing the same homozygous ORF and 3' UTR sequence (**Appendix 3**). Even between sub-species and groups, the open reading frame of the gene is highly conserved with only 9 non-synonymous nucleotide polymorphisms between the three different sub-species and groups. These mutations are clustered at the 3' end of the gene. There are 4 synonymous mutations clustered at the 5' end. STIB386 is heterozygous at 6 of these sites (**Figure 3.8**). There is one non-synonymous polymorphism unique to group 1 *T. b. gambiense* (Ser209Leu) that may result in the lack of gene function when the group 1 *T. b. gambiense HpHbR* is ectopically expressed in the *HpHbR* null mutant Lister 427 line (Kieft *et al.*, 2010).
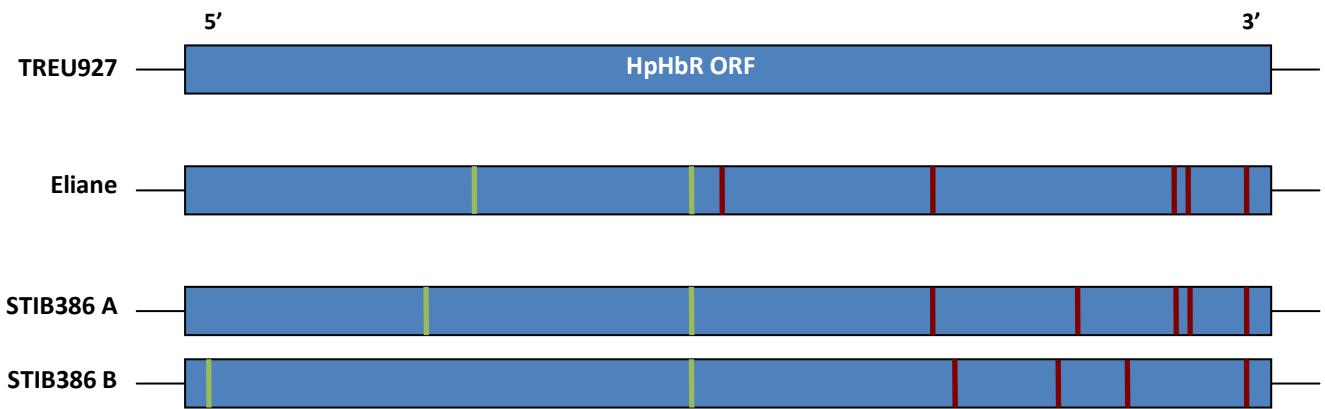
Figure 3.8 Pictorial representation of the distribution of synonymous mutations (green) and non-synonymous mutations (red) found in the *HpHbR* alleles of Eliane and STIB386 compared to the *T. b. brucei* reference strain TREU297. Full gene sequences are given in **Appendix 2**.

The polymorphisms found in the group 2 *T. b. gambiense* strain do not seem to affect the function of the receptor in either resistant or sensitive lines but one homologue of the STIB386 gene shows strongest identity with the group 1 *T. b. gambiense* lines, with only five differences between them. A parsimony tree was constructed using the nucleotide sequence of *HpHbR* to infer any possible relationships between the Eliane allele, the two STIB386 alleles and the alleles found in the *T. b. brucei* STIB247 and TREU927. This indicated that STIB386 allele A is more similar to the group 1 *T. b. gambiense* allele then either STIB386 allele B or the *T. b. brucei* alleles. It also indicates that the allele B of STIB386 and the alleles found in the *T. b. brucei* STIB247 and TREU927 are more alike (**Figure 3.9**). One hypothesis for this arrangement is that group 2 *T. b. gambiense* emerged as a hybrid between a group 1 *T. b. gambiense* and a *T. b. brucei*.
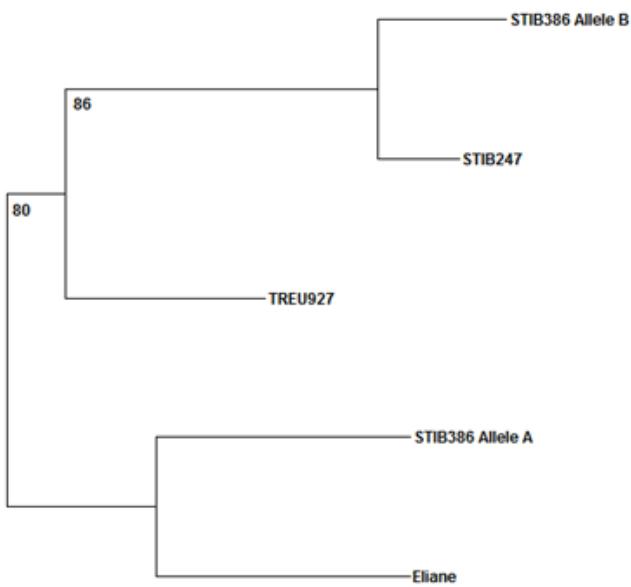
Figure 3.9 Parsimony analysis tree constructed using the nucleotide sequence of *HpHbR* from STIB247, Eliane, TREU927 and both STIB386 alleles. A neighbour joining algorithm with 100 replicates was utilised and bootstrap values above 10 are indicated.

In addition to the loss of function in the group 1 *T. b. gambiense* allele of *HpHbR*, the expression of the gene is strongly down regulated in several group 1 *T. b gambiense* foci. Gene regulation in trypanosomes is largely mediated by post-transcription initiation, controlled by mRNA stability and generally associated with the 3' UTR sequence of each gene (Clayton & Shapira, 2007). The region downstream of the gene containing the 3' UTR was therefore sequenced to investigate whether there were sequence polymorphisms (**Figure 3.10**). The length of the 3' UTR in *HpHbR* is approximately 1000bp (R. Kieft, personal communication). Like the ORF for the gene, group 1 *T. b. gambiense* and TREU927 possess a homozygous 3' UTR while STIB386 possesses a variant for each allele. There are two group 1 *T. b. gambiense* specific polymorphisms in the 3'UTR that may contribute to reduced expression of the receptor. Group 2 *T. b. gambiense* has normal expression of the *HpHbR* gene and transcripts for both 3' UTR variants have been shown to be present in approximately equal amounts in a DGE transcriptomics study (A. Macleod, personal communication) so mutations shared by both groups of *T. b. gambiense* are unlikely to contribute to the down-regulation of *HpHbR* expression observed in Eliane.
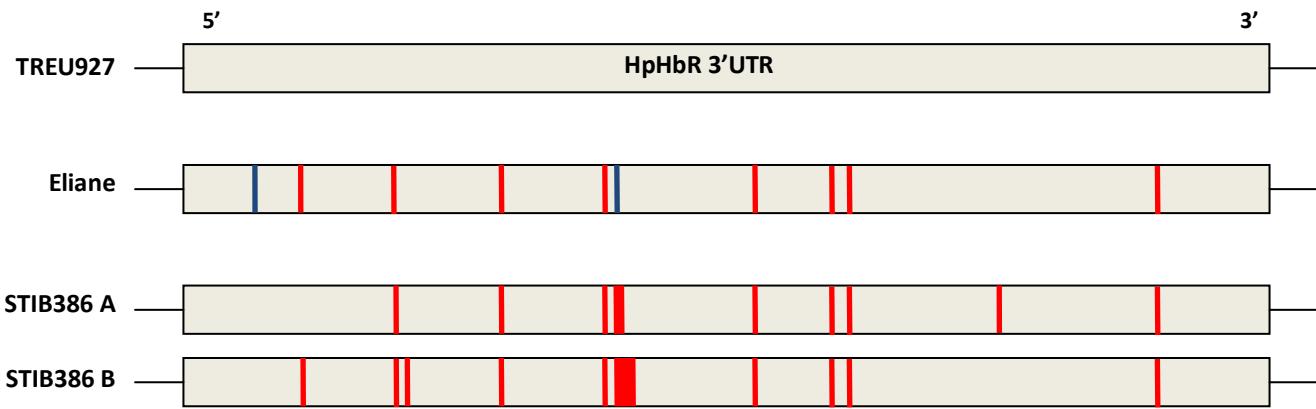
Figure 3.10 Pictorial representation of the distribution of mutations found in the 3' UTR of *HpHbR* for Eliane and STIB386 compared to the *T. b. brucei* reference strain TREU297. Eliane specific mutations are highlighted in blue. Full sequences are given in **Appendix 2**.

3.3 Discussion

These data suggest that group 1 and 2 *T. b. gambiense* have evolved different strategies to deal with lysis by human TLF-1. Group 1 *T. b. gambiense* avoid uptake of the particle but group 2 *T. b. gambiense* take up TLF-1 and then either neutralise it or compensate for its effects. Both of these mechanisms are distinct from the mechanism based on *SRA* expression that has evolved in *T. b. rhodesiense* to achieve a human serum resistant phenotype. Group 2 *T. b. gambiense* exhibits a variable phenotype as previously described (Ortiz *et al.*, 1994), but continuous exposure to a low concentration of normal human serum maintains the resistant phenotype indefinitely in culture. We have been able to show that group 2 *T. b. gambiense* internalises TLF-1 and traffics it to the lysosome like *T. b. brucei* and *T. b. rhodesiense*. In *SRA* expressing *T. b. rhodesiens*e and serum resistant group 2 *T. b. gambiense*, this localisation does not result in cell lysis. Despite the similarity of phenotype, the group 2 *T. b. gambiense* STIB386 strain used in our experiments does not possess the *SRA* gene (Turner *et al.*, 2004). While there are some differences in the rate of uptake and a possible down-regulation of the HpHbR receptor between resistant and sensitive group 2 *T. b. gambiense*, it has little impact on the overall uptake of TLF-1 compared to the dramatic lack of uptake and absence of receptor expression in group 1 *T. b. gambiense*. The lowered uptake of TLF-1 in STIB386R compared to STIB386S may explain observations that human serum incubated with *ex vivo* group 2 *T. b. gambiense* becomes less lytic over time suggesting that the trypanosomes deplete the media of lytic particles. This effect occurs faster in sensitive lines compared to resistant lines (Ortiz-Ordóñez & Seed, 1995). Unfortunately whether the reduced uptake of TLF-1 by resistant STIB386 is a true phenomena or an artefact of low biological replicates is unable to be answered due to a flaw in the design of the fluorescence microscopy experiments. The assays were time consuming and difficult to perform so a large number of trypanosomes were surveyed for each line, but only from a single experimental procedure. In

hindsight, multiple smaller experiments would have proved more effective and allowed the degree of error to be assessed. Nevertheless, we have shown that the resistant group 2 *T. b. gambiense* strain STIB386 internalises and transports TLF-1 to the lysosome, suggesting that there is either a SRA-like neutralising agent or a compensatory protective effect within the lysosome of the parasite.

Previous attempts to characterise human serum resistance/sensitivity in this sub-species using an *in vivo* system suggested that a switch between resistant and sensitive forms of group 2 *T. b. gambiense* was always accompanied by a switch in variable antigen type (VAT), suggesting that the resistance mechanism is closely related to antigenic variation (Ortiz *et al.*, 1994) with an implication that different VSGs could be expressed from different ESs in a manner similar to that seen in *T. b. rhodesiense* (Xong *et al.*, 1998). Our sequencing data show that the same *ESAG6/7* gene copies are expressed in both strains of STIB386 suggesting that, irrespective of *VSG* expression, the same dominant ES is being used in both sensitive and resistant populations and that resistance is unlikely to be related to use of a particular ES as it is in *T. b. rhodesiense.* An alternative hypothesis is that, like group 1 *T. b. gambiense* strains, STIB386 possesses a limited ESAG repertoire (Young *et al.*, 2008). However, group 2 *T. b. gambiense* have previously been shown to be more diverse than group 1 and are more similar to *T. b. brucei* (Gibson, 1986). It has also been shown that group 2 possess ESAGs more similar to those found in *T. b. brucei* and *T. b. rhodesiense* and not group 1 *T. b. gambiense*, although at present only the metacyclic ES have been investigated (Bringaud *et al.*, 2001). Indeed, the *ESAG6* and *ESAG7* being expressed by both forms of STIB386 are identical to those found in a *T. b. brucei* expression site that is absent in group 1 *T. b. gambiense* (Young *et al.*, 2008). While this suggests group 2 *T. b. gambiense* has a comparable ESAG repertoire to *T. b. brucei*, a more complete understanding of the ESAG repertoire of STIB386 is necessary.

In contrast to the group 2 *T. b. gambiense* STIB386, the group 1 *T. b. gambiense* strain Eliane does not appear to internalise and concentrate TLF-1 to any discernable degree within 4 hours and employs an avoidance strategy to counter the major lytic particle of humans. This can be explained by a substantial down-regulation of the HpHbR receptor that is known to bind TLF-1. This decrease in *HpHbR* transcription is consistent across several group 1 *T. b. gambiense* lines from the Côte d'Ivoire, Cameroon and the Democratic Republic of Congo and so would appear to be a common trait of group 1 *T. b. gambiense.* The sequence of the *HpHbR* gene and 3' UTR of isolates from Cameroon and DRC are also identical to the group 1 *T. b. gambiense* Eliane lines

used in our experiments and may serve as diagnostic marker. The similarity amongst isolates may have been expected as this sub-species has been demonstrated to be clonal (Koffi *et al.*, 2009; Morrison *et al.*, 2008). Unlike the selected human serum resistant *T. b. brucei* in which down-regulation of a genomic region around *HpHbR* has occurred, in group 1 *T. b. gambiense HpHbR* is down-regulated independently of surrounding genes. As gene expression in trypanosomes is largely regulated by transcript stability determined by the 3' UTR sequence (Clayton & Shapira, 2007), it is likely that this un-translated region is important to the stability of the *HpHbR* mRNA and consequently the expression levels. There are two group 1 *T. b. gambiense* specific polymorphisms in the 3' UTR that could potentially contribute to mRNA instability and reduced expression of HpHbR. Further investigations with 3' UTR modifications would allow this to be investigated further.

Perhaps due to the lack of expression of the gene, the ORF of *HpHbR* has accumulated polymorphisms over time which affected the phenotype and led to a decrease in function of HpHbR in group 1 *T. b. gambiense*. Knocking in and over-expressing the *T. b. brucei* version of *HpHbR* into an artificially selected L427 *HpHbR* null mutant restores full TLF-1 and human serum sensitivity. Counter to that, over-expressing the group 1 *T. b. gambiense* version of the gene does not restore the full sensitivity phenotype (Kieft *et al.*, 2010). Our group 2 *T. b. gambiense* line STIB386 is heterozygous for the gene. Interestingly, it possesses one allele more closely related to group 1 *T. b. gambiense* and a second that is more distantly related. While it is known that both alleles are expressed in both STIB386R and STIB386S (A. Macleod, personal communication), it is possible that the "group 1 like" allele exhibits the same decreased function. In the future, successful knock-ins of the two group 2 *T. b. gambiense* alleles into the *HpHbR* null *T. b. brucei* mutant may offer insight into which non-synonymous mutations are important for function of the gene. The similarity of one of the *HpHbR* alleles to the group 1 *T. b. gambiense* allele raises the prospect that group 2 are simply the result of very rare mating event of a group 1 individual and a *T. b. brucei*, although evidence to date indicates that group 1 *T. b. gambiense* does not mate (Koffi *et al.*, 2009; Morrison *et al.*, 2008). This would explain the close links between the two groups despite a different resistance phenotype. Population genetic analysis may reveal if this is the case (Chapter 6). This would be interesting as it would imply that both group 1 and 2 *T. b. gambiense* possess the same resistance mechanism, but expression of the phenotype in group 2 is not stable. Group 2 *T. b. gambiense* are amenable to forward genetic techniques (such as quantitative trait loci analysis), and comparative transcriptomics analysis making them an ideal platform to investigate the mechanism of both groups of *T. b. gambiense*.

It has been proposed that TLF particles evolved in primates to exploit the trypanosome's need for uptake of haem and lipid (Vanhollebeke *et al.*, 2008). A predicted consequence of this is that down-regulation of a receptor in response would incur a fitness cost to the trypanosome in a manner analogous to that proposed for SRA in *T. b. rhodesiense* (Coleman & Welburn, 2004). The culture-adapted strain of group 1 *T. b. gambiense* however, does not show any marked difference in growth phenotype from the *T. b. brucei* and *T. b. rhodesiense* lines. This may be due to the fact that the group 1 *T. b. gambiense* strain Eliane can internalise non-lytic particles, such as a bovine HDL of a similar density to TLF-1, although at a slower rate than other *T. brucei* sub-species. This suggests that this group 1 *T. b. gambiense* has evolved to discriminate amongst HDL particles that do or do not contain lytic components even though the bulk of proteins in these particles are identical. Although this implies that the dominant uptake mechanism for TLF-1 is the HpHbR receptor and that the predicted low-affinity receptor is less important (Drain *et al.*, 2001), it was not possible to maintain the group 1 *T. b. gambiense* strain for longer than 4 hours in serum-free media to assess whether the parasite internalises TLF-1 via this low affinity receptor over a much longer time scale. Other genes may be involved with the uptake and trafficking of different mammalian HDL fractions and group 1 *T. b. gambiense* can acquire enough metabolites using these receptors to grow unhindered. Targets for these receptors could be any of the several other major HDL proteins (Heinecke, 2009) not found in TLF-1. It is also pertinent to remember that the secondary lytic particle in humans, TLF-2, provides a second barrier that the parasite must overcome when evolving human serum resistance. Down-regulation of *HpHbR* alone does not lead to full resistance to the lytic effects of whole serum; rather it is just an decreased susceptibility to TLF-1 (Vanhollebeke *et al.*, 2008). The *T. b. rhodesiense SRA* gene neutralises the active component of both TLF particles, but whether TLF-2 resistance in *T. b. gambiense* is due to a down-regulation of other receptors or an *SRA*-like neutralising mechanism is unknown. However, as TLF-2 is believed to be internalised by interactions between the VSG coat of the parasite and TLF-2 bound IgM, it seems unlikely that there is a specific receptor involved in TLF-2 uptake (Vanhollebeke & Pays, 2010).

Chapter 4: Recombinant ApoL1 interactions with *T. b. gambiense*

4.1 Introduction to ApoL1

The general consensus in the human serum resistance field at present is that the active trypanolytic component of TLF particles is ApoL1 (Shiflett *et al.*, 2007; Vanhamme *et al.*, 2003; Vanhollebeke *et al.*, 2007). While the Hpr protein may have some toxic potential, it is primarily the ligand for the HpHbR receptor and facilitates uptake of TLF-1 and hence ApoL1 into the cell (Vanhollebeke *et al.*, 2008). Recombinant ApoL1 can lyse trypanosomes with the same changes in morphology as normal human serum, albeit at a slower rate (Lecordier *et al.*, 2009; Vanhamme *et al.*, 2003; Vanhollebeke *et al.*, 2008). The ApoL1 gene is a member of a gene family with homologues found in many multi-cellular metazoa, from fish to humans (Duchateau *et al.*, 1997). While ApoLIII appears to be the ancestral ApoL gene in mammals, it has been duplicated and then diverged to various degrees in separate lineages (Monajemi *et al.*, 2002). For example, *Homo sapiens* possess six members of the ApoL family while *Mus musculus* possess fourteen. In *Homo sapiens*, transcripts for the ApoL gene family have been found throughout the body (Duchateau *et al.*, 2001; Monajemi *et al.*, 2002), although this is likely to be due to the fact the genes are predominantly expressed by the cells of the endothelium (Vanhollebeke & Pays, 2006).

Most ApoL members are intracellular and are believed to be involved in apoptosis due to structural homology with the mammalian apoptosis protein family, Bcl-2 (Vanhollebeke & Pays, 2006). While most cell-mediated apoptosis in metazoa involves Bcl-2 proteins and permealisation of the mitochondrial membrane, certain cells of the endothelium and immune system do not appear to use this system and instead utilise a lysosomal mediated apoptosis system mediated by some ApoL proteins (Bidere *et al.*, 2003). With specific regard to ApoL1, little is known of any other functions it may possess in addition to its trypanolytic effects. Expression of ApoL1 has been implicated with involvement in tumours, replicative senescence and viral infected cells, from which it would appear to have a potential role in apoptosis in some cell types (Vanhollebeke & Pays, 2006).

Unlike the majority of ApoL family members which are intracellular, ApoL1 possesses an additional 5' signal peptide leading to its secretion by the endothelial cells in blood vessels (Vanhollebeke & Pays, 2006). The protein is usually located in human serum at a modal

concentration of approximately 6µg/ml, although this concentration varies a great deal (from 2.5µg/ml to 30µg/ml) and is closely correlated with the amount of lipid in circulation (Duchateau *et al.*, 2000). The association with lipid concentration is largely due to the fact that ApoL1 is not found free in human serum, but bound to other lipoproteins and lipid into discrete high density lipoprotein (HDL) particles (Duchateau *et al.*, 1997). ApoL1 is primarily found in two different trypanolytic particles of HDL – TLF-1 and TLF-2 (Molina-Portela Mdel *et al.*, 2005; Shiflett *et al.*, 2005). TLF-1 is a lipid rich particle internalised by trypanosomes via the HpHbR receptor, with Hpr protein acting as the ligand (Vanhollebeke *et al.*, 2008). TLF-2 contains much less lipid then TLF-1 but is considerably larger due to the binding of IgM to the particle (Raper *et al.*, 1999; Tomlinson *et al.*, 1995). TLF-2 is much less well understood due to its low stability *ex vivo*, although uptake is not Hpr mediated and is instead likely to use the weak interaction between the IgM associated with the particle and VSGs on the parasite surface (Vanhollebeke & Pays, 2010). It is believed to be internalised passively along with the VSGs as part of the cell surface recycling mechanism of the parasite.

Like other ApoL proteins, ApoL1 possesses three distinct domains - a pore-forming domain, a membrane-addressing domain and a non-functional structural domain (Perez-Morga *et al.*, 2005). The pore-forming domain shows homology to other well characterised pore forming motifs, such as bacterial colicins, diphtheria toxin and the previously mentioned Bcl-2 apoptosis protein family (Vanhollebeke & Pays, 2006). It was demonstrated that the pore forming domain of ApoL1 acts as an anionic channel when embedded in a membrane (Molina-Portela Mdel *et al.*, 2005; Perez-Morga *et al.*, 2005). These channels allow chloride ions to freely move across membranes, depolarising them and perturbing the osmotic potential of cellular compartments. In lysosomes, this leads to an uncontrolled osmotic swelling, either lysing the cell directly or indirectly by the release of digestive enzymes from damaged lysosomes (Perez-Morga *et al.*, 2005). The membrane-addressing domain is essential for efficient integration of the pore-forming domain into the plasma membrane. At a neutral pH it exposes a hydrophobic surface that binds lipid, suggesting that HDL particles can assemble passively *in situ* in the bloodstream (Perez-Morga *et al.*, 2005). At an acidic pH, such as that found in the lysosome, this domain changes conformation and releases the ApoL1 from the HDL particle. Low pH is essential to the efficiency of the lytic function of the protein (Perez-Morga *et al.*, 2005).

Considerable amounts are known about ApoL1 and its interaction with *T. b. brucei* and *T. b. rhodesiense*. ApoL1 is taken up into the parasite, usually as part of a TLF particle but also can be

internalised if it is unbound, as demonstrated by *in vitro* experiments using recombinant protein. It is then trafficked to the lysosome of both *T. b. brucei* and *T. b. rhodesiense* (Molina-Portela Mdel *et al.*, 2005; Perez-Morga *et al.*, 2005; Vanhamme *et al.*, 2003). The serum resistance associated (SRA) protein of *T. b. rhodesiense* was shown to correlate strongly with the position of ApoL1 in the lysosome (Vanhamme *et al.*, 2003). SRA protein inhibits the function of ApoL1 by strongly binding to the C-terminal helix and disrupting the pore-forming domain (Vanhamme *et al.*, 2003). As previously described, baboons and the closely related *Mandrilla* genus possess an ApoL1 protein slightly different from that found in the *Homo sapiens* (Lecordier *et al.*, 2009; Thomson *et al.*, 2009). The ApoL1 found in baboons has mutations to the C-terminal helix with which SRA interacts. Since SRA cannot bind baboon ApoL1 and inhibit the lytic affect of the protein, baboon serum can lyse *T. b. rhodesiense*. Genetic variants of human ApoL1 that have had the SRA binding motif removed to mimic baboon ApoL1 are also able to kill *T. b. rhodesiense* but not *T. b. gambiense* (Lecordier *et al.*, 2009). It is hoped that the creation of transgenic animals expressing modified or baboon *SRA* would eliminate *T. b. rhodesiense* (Thomson *et al.*, 2009) but such a strategy is probably also likely to select for *T. b. gambiense* and *SRA* negative *T. b. rhodesiense*.

Due to the fact that the SRA protein of *T. b. rhodesiense* is able to inhibit normal human ApoL1 directly, the sub-species is able to resist lysis by both TLF-1 and TLF-2 particles. Unlike *T. b. rhodesiense*, group 1 *T. b. gambiense* do not internalise TLF-1 and avoid the ApoL1 found in this particle (Chapter 3 & (Kieft *et al.*, 2010)). It achieves this by down-regulating the expression and function of the HpHbR receptor (Kieft *et al.*, 2010). This would suggest that the resistance mechanism used by *T. b. gambiense* is an ApoL1 avoidance strategy. If this is the case then how this particular sub-species group is able to resist lysis by TLF-2 is unknown. Down-regulation of expression of the HpHbR receptor in *T. b. brucei* leads to resistance to TLF-1 but not TLF-2 or normal human serum (Vanhollebeke *et al.*, 2008). In addition, it is unlikely that down-regulating a single receptor would be sufficient to avoid TLF-2 lysis due to the proposed IgM mediated uptake of the particle (Vanhollebeke & Pays, 2010). It would appear that TLF-2 has the potential to enter all *T. brucei* unhindered unless *T. b. gambiense* possesses a novel process to remove IgM from the VSG on the cell surface. If TLF-2 uptake is unavoidable then group 1 *T. b. gambiense* must also possess a method to resist the lytic effects of ApoL1 despite down-regulating the HpHbR receptor. Like other *T. brucei* strains, group 2 *T. b. gambiense* do not avoid the TLF-1 particle so must possess a mechanism to deal with ApoL1 lysis that extends to both TLF-1 and TLF-2 (Chapter 3).

Whilst other studies described in this thesis have revealed considerable insight into the interactions of TLF-1 with *T. b. gambiense*, TLF-2 presents a greater challenge. TLF-2 is notoriously difficult to work with due to low stability *ex vivo* (Raper *et al.*, 1999; Tomlinson *et al.*, 1995). It is possible however to replicate some of features of the particle if a source of recombinant ApoL1 could be generated. Like TLF-2, recombinant ApoL1 is not internalised by a receptor using Hpr as a ligand. While TLF-2 is likely internalised as part of the general membrane recycling system of trypanosomes, free ApoL1 is instead internalised via non-specific endocytosis (Lecordier *et al.*, 2009). Recombinant ApoL1 has also been shown to be trafficked to the lysosome in a manner similar to TLF particles (Perez-Morga *et al.*, 2005; Vanhamme *et al.*, 2003). Unlike recombinant Hpr, recombinant ApoL1 kills trypanosomes with morphological changes similar to those seen after exposure to normal human serum and TLF, although this occurs at a slower rate (Vanhamme *et al.*, 2003). This is probably due to a lowered efficiency of uptake compared to using the HpHbR receptor (Vanhollebeke *et al.*, 2008).

The primary aim of this thesis chapter was to design a robust and reliable protocol to express recombinant ApoL1. While several studies utilising recombinant ApoL1 have been described (Lecordier *et al.*, 2009; Perez-Morga *et al.*, 2005; Vanhamme *et al.*, 2003; Vanhollebeke *et al.*, 2007), the exact protocols used were unclear at the time this thesis was commenced. With the creation of recombinant ApoL1, several key features of the interaction between the protein and the two groups of *T. b. gambiense* can be investigated. A prime question to be answered was whether the reduced uptake of TLF-1 shown by group 1 *T. b. gambiense* was the sole resistance mechanism. If group 1 *T. b. gambiense* are inherently sensitive to ApoL1 and must avoid the protein, several therapeutic avenues present themselves to exploit this weakness. For example, a nanobody conjugated to the ApoL1 pore forming domain would be able to enter the *T. b. gambiense* cell attached to the VSG and lyse it. This system has proved somewhat successful at lysing *T. b. brucei* and *T. b. rhodesiense* but has yet to be tried with *T. b. gambiense* (S. Magez, personal communication & (Baral *et al.*, 2006). The stably sensitive and resistance lines of the group 2 *T. b. gambiense* isogenic strains were also compared to investigate the similarities and differences between the two groups of *T. b. gambiense*. Much of the examination of recombinant ApoL1 interaction with these trypanosome lines can be extended to apply to TLF-2 due to several shared features as described above.

4.2 Results

4.2.1 ApoL1 ORF

The sequencing of the purchased ApoL1 ORF revealed one non-synonymous mutation (Arg8Gly) compared to the human reference genome (**Appendix 4**). However, this region is not implicated in the trypanosome killing function of the protein (Perez-Morga *et al.*, 2005). In addition, the sequence of an independently derived ApoL1 ORF used by other research groups has been published (Lecordier *et al.*, 2009). The ApoL1 used by these researchers has been shown on several occasions to be fully functional despite also possessing the Arg8Gly mutation. This indicates that there is variation in the ApoL1 gene in humans, although this particular mutation does not impact on the lytic capability of the protein. The commercial ApoL1 ORF was therefore considered suitable to be used in further recombinant expression experiments.

4.2.2 Wheatgerm protein expression

A protocol for expression of functional recombinant ApoL1 was not available when the project commenced so a protocol had to be created. It has previously been demonstrated that ApoL1 is highly toxic to *E. coli* (Perez-Morga *et al.*, 2005). In order to avoid the ApoL1 toxicity apparent in *E. coli*, a range of cell-free *in vitro* expression systems were investigated. Initially, a wheatgerm *in vitro* expression system was utilised as, in theory, a eukaryotic model would be most likely to provide the most accurate folding and post-translation of the recombinant protein. Attempts to use the commercial Endext® system using the PCR construction of the ApoL1 ORF fused to an expression promoter sequence did not bear fruit however due to truncation of the PCR product after the reaction. Using the predictive primer affinity function found in the CLC-Bio software package, this was determined to be most likely due to the high affinity the first round forward primer had to a central region of the ApoL1 ORF (data not shown). The forward primer sequence is pre-defined by the Endext® system so using a different tag was not possible.

An alternate method was attempted that used the same Endext® wheatgerm expression system but the promoter sequence was located in an Endext® plasmid. The ApoL1 ORF successfully ligated into the correct site within the plasmid and was transformed into competent BL21-AI *E. coli* cells so that the high amount of plasmid DNA necessary could be obtained. Unfortunately, the transformed *E. coli* cells grew disappointingly and it was not feasible to grow to the volume needed to create enough plasmid DNA. Poor growth may have been due to "leaky" expression of the ORF leading to small amounts of toxic ApoL1 being produced by the bacteria. The Endext® wheatgerm *in vitro* expression system was determined to be unsuitable for our needs and available materials.

4.2.3 Invitrogen Expressway® system

A second cell-free expression system was selected for consideration - the Invitrogen Expressway® *E. coli* expression system. This system was selected due to full compatibility with the Invitrogen Gateway® vector that had been purchased and verified to be correct. It is also fully compatible with the BL21-AI *E. coli* strain available which would reduce the hypothesised leaky ORF expression believed to be present in the Endext® system. The Expressway® system proved to be able to produce recombinant soluble ApoL1 protein (**Figure 4.1**), although with a low yield (≈ 300µg per production reaction). A non-denaturing purification system was utilised to ensure the soluble protein did not require re-folding.
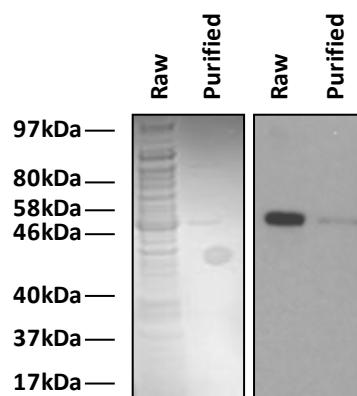
Figure 4.1 Coomassie stained gel (left) and Western blot with an Anti-ApoL1 antibody (right) performed on the raw mix from an 8 hour Expressway® protein expression and after purification under non-denaturing conditions with a nickel column. The protein band visible is at the expected size for ApoL1.

After the protein was concentrated and verified to be ApoL1, its lytic ability against *T. b. brucei* was investigated using several different concentrations of protein. Recombinant ApoL1 produced from the Expressway® system did not kill human sensitive *T. b. brucei* strain STIB247, even at relatively high concentrations compared to those found commonly *in vivo* (**Figure 4.2**). This suggests that the Expressway® system does not produced functional ApoL1, even though it is the correct size and is detected by an ApoL1 antibody. It is possible that this particular cell-free system does not provide the full suite of post-translational modification needed to create functional ApoL1.
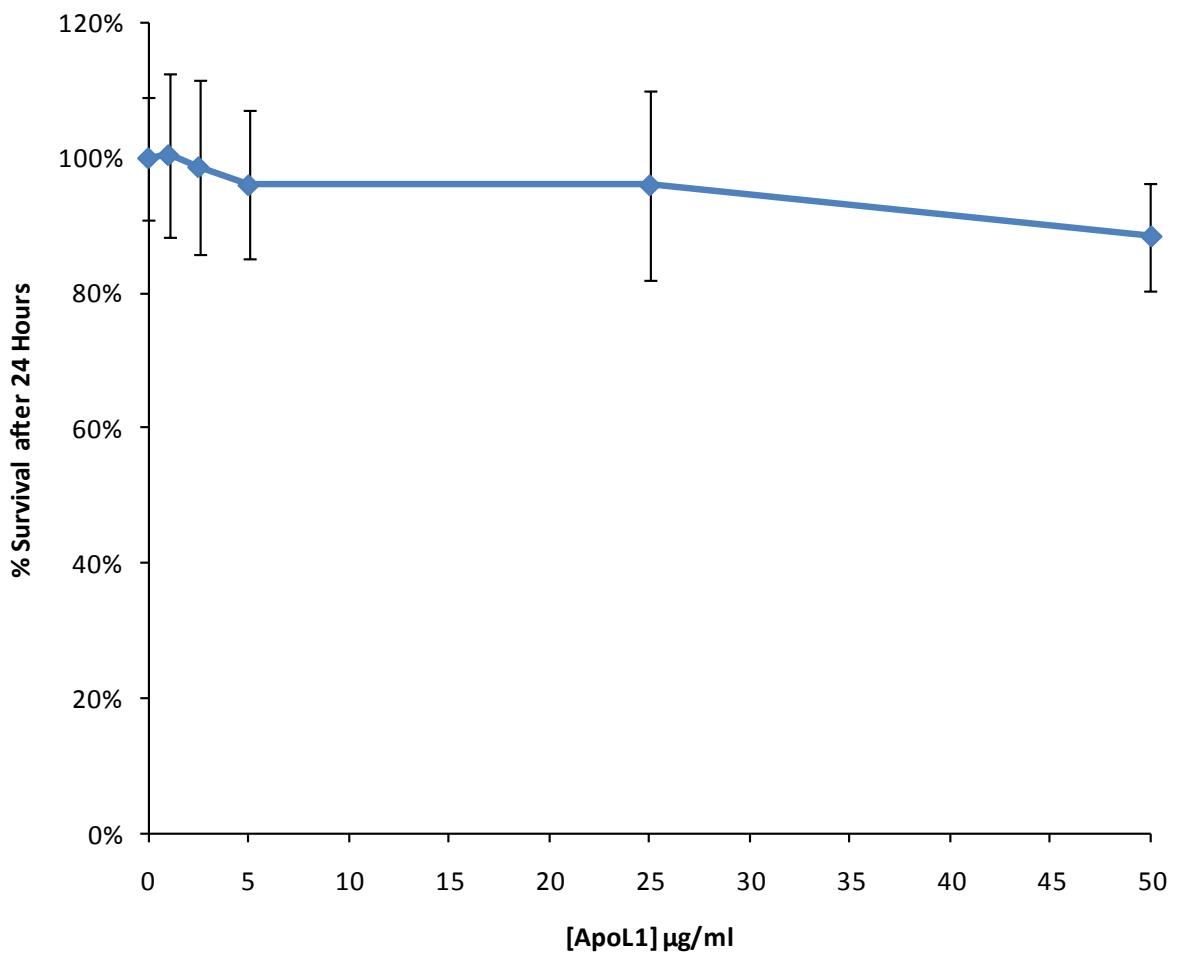
Figure 4.2 The percentage survival after 24-hours of the *T. b. brucei* strain STIB247 when exposed to various concentrations of recombinant ApoL1 obtained from the Expressway® *in vitro* protein expression system. Standard error is indicated (n=3) and samples are normalised to the 0µg/ml protein control. A positive control consisting of STIB247 exposed to 25% human serum showed 0% survival after 24 hours.

4.2.4 Gateway® *E. coli* expression system

The cell-free *in vitro* expression system could only produce low protein yield with little (if any) activity. It was decided that a classical expression system using *E. coli* would be utilised, despite the toxicity of ApoL1 to *E. coli*. The Invitrogen Gateway® expression system was employed for improved efficiency and ease of use compared to other systems. The system was also fully compatible with the Gateway® entry plasmid containing the ApoL1 ORF. The expression system proved successful, with large amounts of protein produced after induction. Unfortunately the protein was abundant but present in insoluble inclusion bodies (**Figure 4.3**).

The formation of inclusion bodies may arise due to the toxicity of ApoL1 to bacteria (Perez-Morga *et al.*, 2005). As it is difficult to purify recombinant protein from the insoluble fraction, temperature conditions were modified to provide slower inductions in the hope of improving the solubility of the protein. Solubility was not increased by lower induction temperatures, although the yield of protein decreased (**Figure 4.3**).
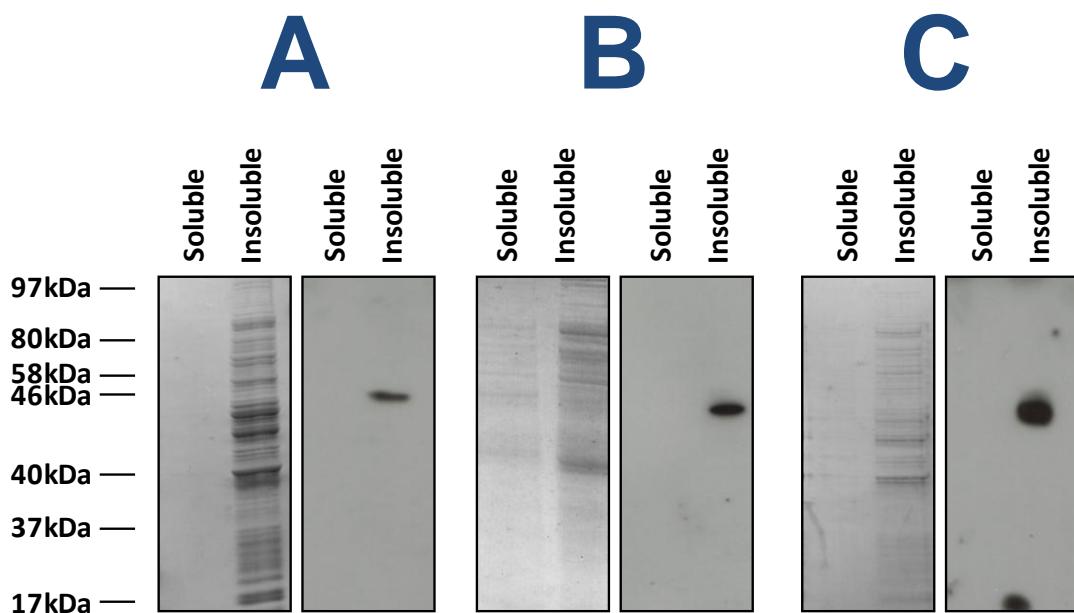
Figure 4.3 Coomassie stained gels (left) and Western blots (right) using an anti-ApoL1 antibody performed on the soluble and insoluble fractions after expression induction at (A) 16˚C (B) 26˚C and (C) 37˚C.

The toxicity of ApoL1 is such that inclusion bodies are always likely to result from an induction in *E. coli* - soluble ApoL1 within the transformed cells would quickly lyse them. Rather than avoid inclusion bodies by changing induction conditions, an alternate method would be to solubilise the protein using high concentrations of denaturing urea. This can lead to issues with proteolysis of the recombinant protein and also presents difficulties when re-folding the protein into the correct conformation. Fortuitously, data were published parallel to these experiments suggesting re-folding of recombinant ApoL1 could be performed using weak acetic acid after urea treatment solubilising the inclusion bodies (Lecordier *et al.*, 2009). 8M urea gave the best yield of protein, despite some proteolysis (**Figure 4.4**). This cell-based protein expression, in combination with a nickel column purification system gave consistently high yields of protein (≈ 200mg per protein expression).
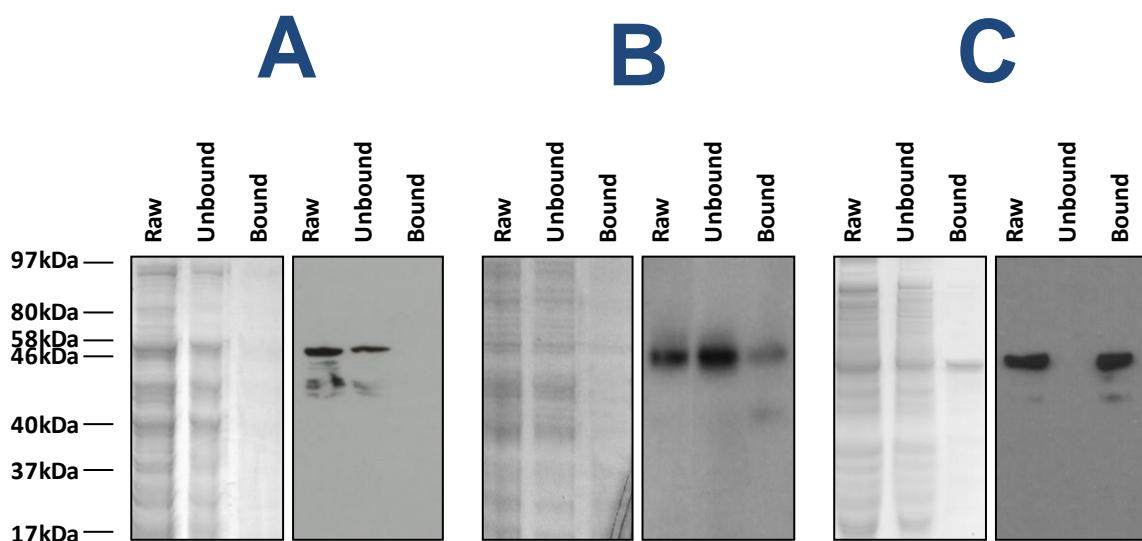
Figure 4.4 Coomassie stained gels (left) and Western blots (right) using an anti-ApoL1 antibody performed on the raw reaction mixture and soluble fraction after treatment with wash buffer containing (A) 4M Urea (B) 6M Urea and (C) 8M Urea. The predominant protein band is the correct size for ApoL1. Some protein degradation is visible.

Unlike the Expressway® *in vitro* expression system, recombinant ApoL1 from this system causes death in *T. b. brucei* at amounts close to physiological levels (**Figure 4.5**). Cells appear to die with the characteristic swollen lysosome phenotype, although cell death occurs more slowly than with normal human serum, requiring 24 hours for complete lysis of an *in vitro T. b. brucei* population (data not shown). The slower lysis rate of ApoL1 without the Hpr ligand has been

observed on numerous occasions and is likely to be due to being internalised via a less efficient non-specific endocytosis mechanism (Lecordier *et al.*, 2009).
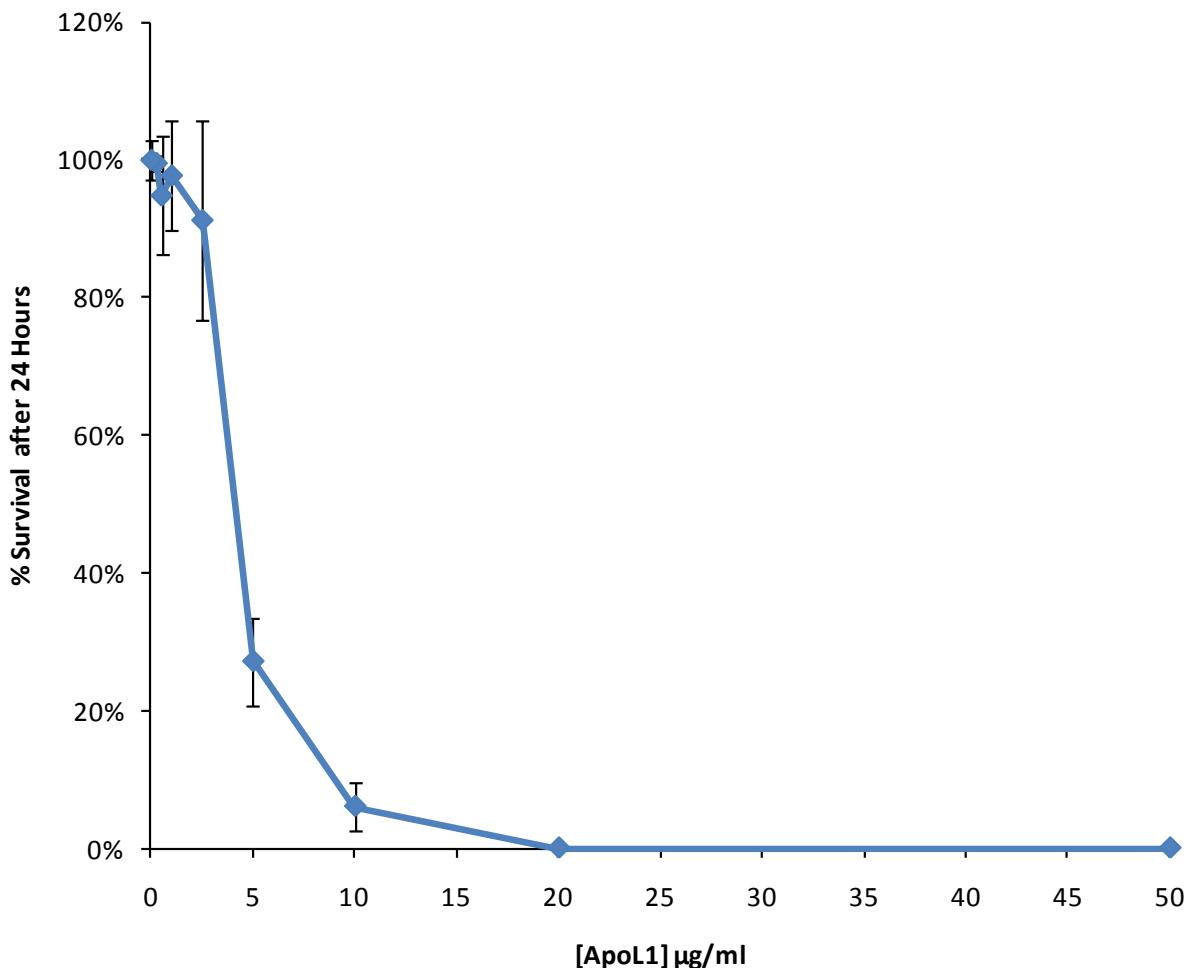
Figure 4.5 The percentage survival after 24-hours of the *T. b. brucei* strain STIB247 when exposed to various concentrations of recombinant ApoL1 obtained from the Gateway® *E. coli* protein expression system. Standard error is indicated (n=2) and samples are normalised to the 0μg/ml protein control. A positive control consisting of STIB247 exposed to 25% human serum showed 0% survival after 24 hours. A negative control consisted of a 0μg/ml concentration of ApoL1 containing solely protein free buffer.

4.2.4 Recombinant ApoL1 assays

As previously described, there are two distinct groups of *T. b. gambiense*. Group 2 show a variable phenotype with both sensitive and resistant forms, while group 1 are always resistant. The capacity for recombinant ApoL1 to lyse both sensitive and resistant forms of group 2 *T. b. gambiense*, in addition to group 1 *T. b. gambiense* and *T. b. brucei*, was examined for several concentrations of protein. Recombinant ApoL1 is able to lyse sensitive group 2 *T. b. gambiense* at levels similar to those that kill *T. b. brucei.* Lysis occurs at even low amounts of protein, approaching physiological levels. The resistant group 2 *T. b. gambiense* strain and the group 1 *T. b. gambiense* strain were unaffected by ApoL1, even at high concentrations (**Figure 4.6**).
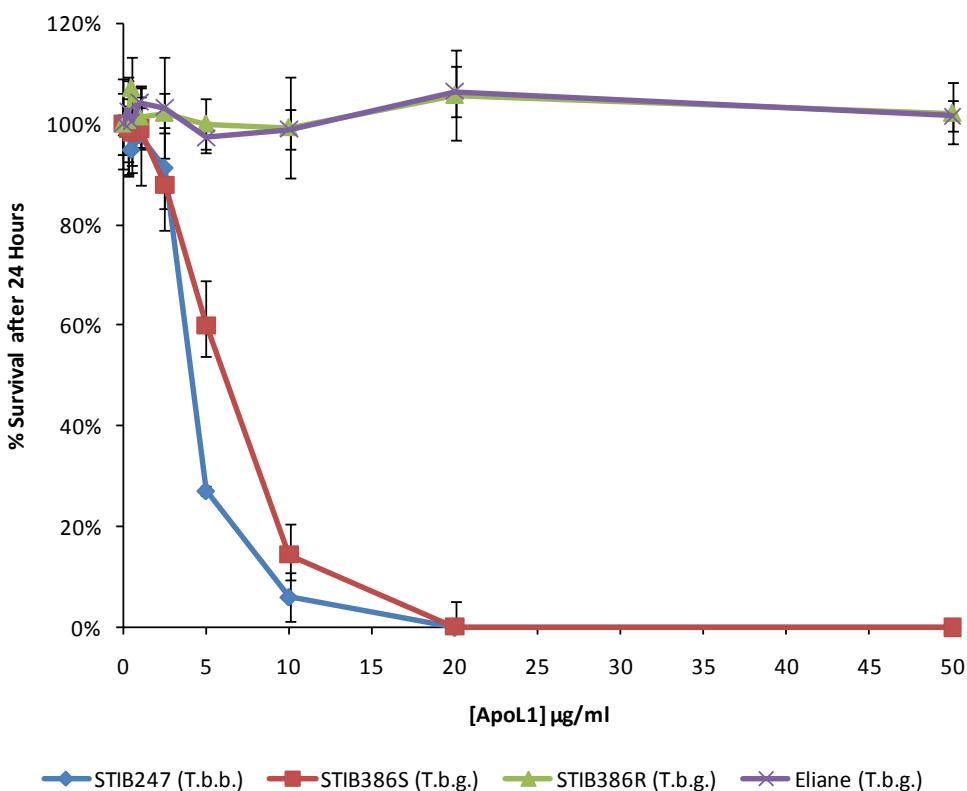
Figure 4.6 The percentage survival after 24-hours of the *T. b. brucei* strain STIB247, the resistant & sensitive isogenic lines of the group 2 *T. b. gambiense* strain STIB386 and the group 1 *T. b. gambiense* strain Eliane, when exposed to various concentrations of recombinant ApoL1 obtained from the Gateway® *E. coli* protein expression system. Values have been normalised to the 0μg/ml control well for that experiment to take into account differing growth rates. Standard error is indicated (n=2) and samples are normalised to the 0μg/ml protein control. A positive control consisted of each strain exposed to 25% human serum showed 0% survival for STIB247 and STIB386S after 24 hours, while Eliane and STIB386R survived at a level comparable to the negative control.

Interestingly, procyclic form *T. b. brucei* are fully resistant to normal human serum (Moore *et al.*, 1995). The mechanism is unknown, although it is unrelated to the decreased endocytosis displayed in this life cycle stage (Natesan *et al.*, 2010). The procyclic form of a strain of *T. b. brucei* that is sensitive to ApoL1 in bloodstream form proved able to resist the effects of recombinant ApoL1 (**Figure 4.7**). While it was not possible to show that procyclic form *T. b. brucei* internalise ApoL1, the phenotype was displayed even at high concentrations of ApoL1 that would compensate for the low levels of endocytosis. This suggests that the innate resistance mechanism of procyclic form *T. brucei* is able to overcome the lytic effects of ApoL1. This mechanism cannot involve *SRA* so there must be transcriptional regulation of genes present in the organism that can achieve the same result.
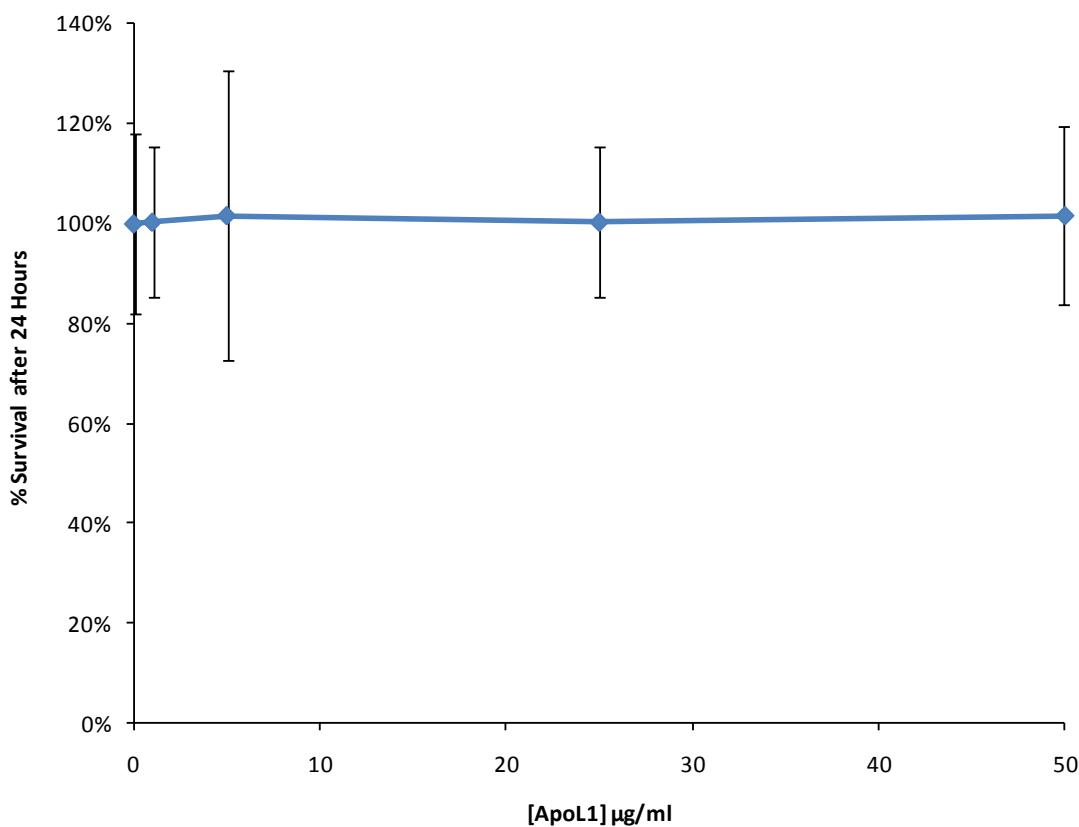

Figure 4.7 The percentage survival after 24-hours of procyclic form *T. b. brucei* strain STIB247 when exposed to various concentrations of recombinant ApoL1 obtained from the Gateway® *E. coli* protein expression system. Standard error is indicated (n=3) and samples are normalised to the 0μg/ml protein control. A positive control consisting of procyclic STIB247 exposed to 25% human serum showed no loss of survival after 24 hours.

4.2.5 Fluorescence microscopy of recombinant ApoL1

There are three possible mechanisms for resistance to the trypanolytic particles found in humans – avoidance, inhibition of lytic particles or compensating for their effects. Group 1 *T. b. gambiense* avoid the TLF-1 particle for example (**chapter 3** & (Kieft *et al.*, 2010)). While it is known that in *T. b. brucei* and *T. b. rhodesiense* recombinant ApoL1 is internalised by non-specific endocytosis, it is possible that the group 1 *T. b. gambiense* and resistant group 2 *T. b. gambiense* do not. It is also possible the recombinant protein is not being trafficked to the lysosomal compartment and so would not reach the site of action. To investigate this, the internalisation and location of ApoL1 was visualised in several *T. brucei* lines using fluorescence microscopy. All sub-species showed visible uptake of recombinant ApoL1 within four hours. The position of highest fluorescence correlates with the lysosomal dye in all strains, including the group 1 *T. b. gambiense* strain Eliane and the resistant group 2 *T. b. gambiense* strain STIB386R (**Figure 4.8**). This result would indicate that both groups of *T. b gambiense* posses the ability to resist the lytic effects of ApoL1 in the lysosome, although expression of the mechanism is variable in group 2 *T. b gambiense*.
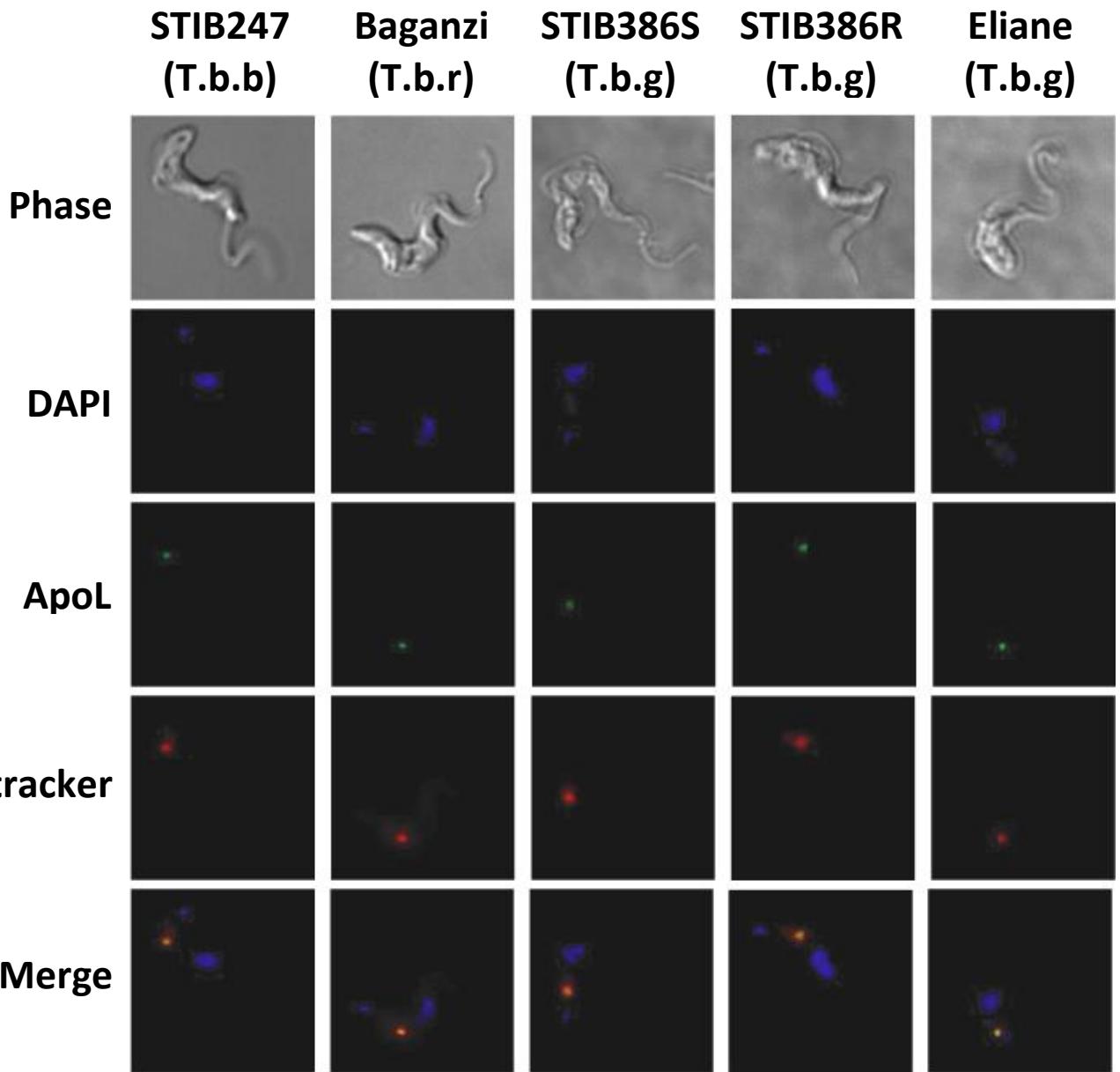
Figure 4.8 The localisation of Alexa488 labelled anti-pentaHis antibody, Lysotracker® & DAPI in several parasite lines after four hour exposure to recombinant ApoL1 featuring a 6xHis tag. The mean correlation co-efficient (R) of the TLF-1 and Lysotracker® field was estimated for each parasite line: for STIB247, R=0.87 (n=16); STIB386S, R= 0.81 (n=13); STIB386R, R=0.89 (n=15); Baganzi, R=0.76 (n=15); Eliane, R=0.85 (n=18) indicating that ApoL1 was not only internalised but also trafficked to the lysosome in these strains.

4.3 Discussion

These data show that functional recombinant ApoL1 can be made reliably and efficiently using a commercial *E. coli* expression system. The toxicity of ApoL1 to *E. coli* leads to production of the recombinant protein in insoluble inclusion bodies but these inclusion bodies can be solubilised by high urea treatment and purified using standard nickel column methods. High urea treatment disrupts the conformation of the protein but accurate re-folding can be achieved using a weak acidic buffer (Lecordier *et al.*, 2009). There is some proteolysis although a good yield of recombinant protein is still achieved. This production protocol uses common buffers and materials, making it ideal for scaling up to produce large amounts of protein. Generation of recombinant ApoL1 has allowed further investigation of fundamental questions concerning the interaction of the main lytic component of both TLF-1 and TLF-2 with *T. b. gambiense.*

As shown in other studies, recombinant ApoL1 kills *T. b. brucei* and sensitive *T. b. rhodesiense* with a similar morphology to that of normal human serum (Vanhamme *et al.*, 2003). The killing effect occurs more slowly than with normal human serum due to the fact that free ApoL1 is internalised via non-specific endocytosis (Lecordier *et al.*, 2009). The highest trypanosome killing efficiency for ApoL1 is achieved via TLF-1 with the associated Hpr protein acting as a ligand for the trypanosome HpHbR receptor (Vanhollebeke *et al.*, 2008). Inhibition of ApoL1 has been shown to be central to the human serum resistance mechanism of *T. b. rhodesiense* using the serum resistance associated (*SRA*) gene. SRA co-localises with ApoL1 within the lysosome and blocks the lytic activity of the protein by disrupting a helix in the pore forming domain (Vanhamme *et al.*, 2003). To date, the activity of recombinant ApoL1 on either group of *T. b. gambiense* has not been investigated. Our data indicate that both groups of *T. b. gambiense* internalise recombinant ApoL1 and traffic it to the lysosome, probably via non-specific endocytosis in a manner similar to that shown in *T. b. brucei* and *T. b. rhodesiense* (Vanhamme *et al.*, 2003). Recombinant ApoL1 kills the sensitive group 2 *T. b. gambiense* strain STIB386S at a similar rate to *T. b. brucei*, but not the isogenic resistant group 2 *T. b. gambiense* strain STIB386R.

This indicates that the resistance mechanism of group 2 *T. b. gambiense* is able to either neutralise or compensate the effects of ApoL1, although expression of the mechanism is variable. Pull-down assays with ApoL1 against lysosomal fractions from resistant forms of group 2 *T. b. gambiense* would reveal if this mechanism was due to the expression of an SRA-like protein that interacts with ApoL1. Unfortunately such specific pull-downs would be particularly difficult to achieve as the lysosome is an extremely sensitive organelle and centrifugation often destroys the

structure. It is likely pull-downs could only be performed on crudely separated materials. It is also possible that the resistance mechanism is not a neutralisation mechanism like *SRA* in *T. b. rhodesiense*. Instead there may be genes that compensate for the pore-forming effects of ApoL1. These could be proteins that modify lysosomal membrane integrity or fluidity to increase the capacity to resist the osmotic effects. A similar resistance effect could also be achieved via transporters that are able to counteract the influx of chloride ions. In this case, comparisons between the transcriptomes of resistant and sensitive lines of the STIB386 strain may reveal candidate genes that are differentially expressed and are involved in compensatory mechanisms. A quantitative trait locus analysis would also reveal areas of the genome implicated in involvement. Both options are available in group 2 *T. b. gambiense* due to its variable phenotype and ability to undergo mating with genes re-assorted in a Mendelian manner (Cooper *et al.*, 2008; Morrison *et al.*, 2009; Veitch *et al.*, 2010).

Group 1 *T. b. gambiense* are also able to resist the lytic affects of ApoL1, although the mechanism is constitutively expressed. There may possibly be a common mechanism between group 1 and 2 *T. b. gambiense*, although the expression of the system is variable in group 2. This may have arisen if group 2 is a hybrid generated between a group 1 *T. b. gambiense* and a *T. b. brucei*. If there is a common mechanism, this would reveal group 2 *T. b. gambiense* to be an excellent model for group 1. Resistance to ApoL1 in group 1 *T. b. gambiense* is interesting as these parasites have evolved to avoid the main ApoL1 containing particle TLF-1 by down-regulating expression and function of the HpHbR receptor (Kieft *et al* 2010, submitted manuscript). Down-regulation of the HpHbR receptor is found at several disease foci and seems to be a fundamental feature of group 1 *T. b. gambiense* (chapter 3 & (Kieft *et al.*, 2010)). Several hypotheses can be formulated for why group 1 *T. b. gambiense* avoid TLF-1 despite its inherent resistance to ApoL1; such as the idea that the TLF-1 particle evolved before TLF-2 in primates. *T. b. gambiense* may initially have evolved primate infectivity by simply avoiding the primordial TLF particle. The sub-species was then unable to use a similar mechanism to avoid uptake of TLF-2 as it is internalised by weak interactions with IgM and the surface VSG (Vanhollebeke & Pays, 2010). Instead, a second resistance mechanism to counteract ApoL1 evolved. This is pure speculation however as nothing is known about the evolution of TLF-2 or the distribution of the particle amongst primates. Another hypothesis would involve the recently described phenomena that Hpr signal peptide possesses membrane disruptive potential specific to bloodstream form trypanosomes (Harrington *et al.*, 2010). Group 1 *T. b. gambiense* may be especially susceptible to this and so the avoidance mechanism has not evolved to combat ApoL1 but rather Hpr signal peptide. While

there was no difference between the susceptibility to Hpr signal peptide between *T. brucei* sub-species, the peptide in these assays was integrating into the cell surface membrane. The site of action for the Hpr signal peptide would be the lysosomal membrane. The differences between *T. b. brucei* and *T. b. gambiense* lysosomal structure or membrane integrity have not been examined. Another hypothesis is that the *T. b. gambiense* resistance mechanism is dose dependant and by avoiding TLF-1 uptake is able to avoid the majority of the ApoL1 in the human bloodstream. This would suggest that group 1 *T. b. gambiense* may be susceptible to ApoL1 lysis if internalised at high concentrations. Using an ApoL1-nanobody conjugate to increase the uptake rate of ApoL1 above that of standard recombinant ApoL1 may reveal if this is the case (Baral *et al.*, 2006). While several theories can be invoked to explain why group 1 *T. b. gambiense* do not uptake TLF-1, it will be difficult to infer conclusions without greater knowledge of the genes involved in ApoL1 resistance.

Another interesting result revealed by these data is that sensitivity to ApoL1 mediated lysis is lifecycle stage-specific. Bloodstream form *T. b. brucei* are killed by ApoL1 while the procyclic forms of the same strain are not. This correlates with data showing that procyclic form *T. b. brucei* are fully resistant to normal human serum despite being quickly lysed when in bloodstream form (Moore *et al.*, 1995). While it was initially believed that resistance in procyclic form *T. b. brucei* was due to a 10-fold decrease in endocytosis, mutant procyclic *T. b. brucei* that exhibit bloodstream levels of endocytosis are still resistant to human serum (Natesan *et al.*, 2010). This result suggests there is a mechanism present in procyclics that is rendering them resistant to the lytic effects of ApoL1. This process may involve transcriptional regulation of genes already present. One possible gene involved is the lysosomal protein p67. This gene is up-regulated in bloodstream form *T. brucei* compared to procyclics (Kabani *et al.*, 2009; Veitch *et al.*, 2010) and artificially reducing expression of this gene does increase human serum resistance in *T. b. brucei*, although with a stark decrease in survivability (Peck *et al.*, 2008). Most of the genes identified by a transcriptional comparison between procyclic and bloodstream form *T. brucei* were hypothetical and have no known function but may contribute to the phenotype (Veitch *et al.*, 2010). Understanding the function of these hypothetical genes may reveal their importance and what impact they have on the resistance phenotype of procyclic form *T. b. brucei*. Why all bloodstream *T. brucei* cannot invoke the procyclic mechanism to resist lysis is unknown although perhaps this is the system utilised by *T. b. gambiense*. Whole transcriptome comparisons between *T. b. brucei* procyclic and bloodstream forms versus *T. b. gambiense* may reveal if this was the case.

Until recently, little was known concerning the interaction of ApoL1 and *T. b. gambiense*. The series of experiments in this chapter demonstrate that both group 1 and 2 *T. b. gambiense* internalise and transport ApoL1 to the lysosome. Group 1 and resistant group 2 *T. b. gambiense* are not lysed so both must possess a mechanism to resist the lytic effects of ApoL1. This may be an inhibitory effect, such as the binding of the SRA protein to ApoL1 in *T. b. rhodesiense* (Vanhamme *et al.*, 2003). The mechanism must be considerably different from *SRA* however as modifications to ApoL1 that allow it to kill *T. b. rhodesiense* by avoiding inhibition do not lyse *T. b. gambiense* (Lecordier *et al.*, 2009). Pull down assays investigating the binding of recombinant ApoL1 to fractions of *T. b. gambiense* cells may reveal any interactions between the protein and an inhibitor. Similar assays would also reveal a possible neutralising agent expressed by *T. b. gambiense*, such as a protease that targets ApoL1.

If however the resistance mechanism of *T. b. gambiense* is not by direct inhibition of ApoL1 but instead an ability to compensate for the effects of the protein, then association and pull down assays will not work. In group 2 *T. b. gambiense*, forward genetic and transcriptomic approaches using the isogenic resistant and sensitive STIB386 may offer insight into the processes and proteins involved (Cooper *et al.*, 2008; Veitch *et al.*, 2010). Unfortunately these systems do not appear to be transferable to group 1 *T. b. gambiense* as this group of does not seem to undergo mating in a laboratory setting so there can be no genetic crosses for QTL analysis. Additionally, it is invariably resistant so there is no opportunity to perform comparative transcriptomics with resistant and sensitive isogenic lines. It may be possible to compare the transcriptomes of *T. b. brucei* isolates to that of group 1 *T. b. gambiense* isolates, although preliminary data suggests that there are a great deal of genes that are differentially expressed between these two sub-species (N. Veitch, unpublished data). Instead the application of recombinant ApoL1 and inhibitors of cellular processes and enzymes may provide insight into possible mechanisms involved in counteracting the effects of the anionic pores formed by ApoL1. Previously, experiments such as these were hampered in that TLF-1 is not internalised by group 1 *T. b. gambiense* and TLF-2 is too unstable to be reliably used. The recombinant ApoL1 created by our protocol allows us to overcome these biological limitations and potentially investigate the resistance mechanism of group 1 *T. b. gambiense* directly.

Chapter 5: Quantitative Trait Loci (QTL) analysis of the group 2 *T. b. gambiense* human serum resistance phenotype

5.1 Introduction to quantitative trait loci analysis

Most studies investigating trypanosome phenotypes of interest have used a reverse genetics approach. Genes are manipulated, either by interfering with them *in situ* or placing them into new genetic backgrounds. Several tools exist for such analysis in trypanosomes, including RNA interference (Wang *et al.*, 2000) and stable genetic transformation (Burkard *et al.*, 2007). From changes in observed phenotype due to the manipulation of the target, the function of the gene in the whole organism can be inferred. However, most genes do not function individually and the downstream effects of the manipulation can be unpredictable or even unobservable. Another dilemma innate to these studies is that genes unique to trypanosomes are unlikely to be selected for reverse genetics approaches due to the fact most candidates for manipulation arise from homologous genes identified in other organisms. An alternate strategy to reverse genetics is the converse, forward genetics. While reverse genetics examines the phenotypes caused by specific genes, forward genetics seeks to find the genes involved with a phenotype.

A standard approach when investigating the link between phenotype and genotype starts with strongly inbred mutant parental lines that possess a phenotype of interest (Doerge, 2002). These parents are homozygous across their genomes due to continuous inbreeding. The parental line is then bred with another homozygous parent featuring a different response to the phenotype in question. This out-breeding produces heterozygous F1 progeny. Two F1 progeny are then mated and re-assortment due to crossover events causes random inheritance of genes in the F2 progeny. The process can be simplified somewhat if the parental lines used possess different phenotypes and are already heterozygous for genes associated with the phenotype (if using wild isolates for example). In this situation analysis can be performed directly on the F1 progeny. By examining the inheritance of alleles and the expression of the phenotype in the progeny it is possible to locate alleles that are more likely to be associated with the phenotypic trait of interest. This quantitative trait loci (QTL) analysis has been used to identify important genes in several parasites, including

Toxoplasma gondii, *Plasmodium falciparum* and *Trypanosoma brucei brucei* (Ferdig *et al.*, 2004; Morrison *et al.*, 2009; Su *et al.*, 2002).

Group 1 *T. b. gambiense* is the most prevalent group of this sub-species but is unsuitable for forward genetics approaches due to its inability to undergo crosses in a laboratory setting. At all disease foci studied, group 1 *T. b. gambiense* have been shown to possess a clonal population structure suggesting mating does not occur in this group (Chapter 6 & (Koffi *et al.*, 2009; Morrison *et al.*, 2008; Tait *et al.*, 1984). However, forward genetics approaches can be utilised in group 2 *T. b. gambiense* as it undergoes mating under laboratory conditions and alleles are recombined in a Mendelian manner (Cooper *et al.*, 2008; MacLeod *et al.*, 2005a). QTL analysis is an ideal technique for investigating phenotypes in this group of *T. b. gambiense* due to the publication of both the *T. brucei* genome allowing the identification of microsatellite markers (Berriman *et al.*, 2005) and the genetic map for both group 2 *T. b. gambiense* and *T. b. brucei* indicating the likelihood of recombination events between those markers in group 2 (Cooper *et al.*, 2008; MacLeod *et al.*, 2005b). The genetic map of the group 2 *T. b. gambiense* strain STIB386 comprises of 127 microsatellite markers distributed across the mega-chromosomes. The frequency of a crossover event between two markers indicates the genetic distance between them. By comparing markers on the genetic map with the inheritance of the phenotype in the progeny, markers that co-segregate with the phenotype can be identified. This allows an estimate of the likelihood of linkage to the phenotype to be made (Doerge, 2002). A statistical value is appended to each marker that indicates the logarithm of the odds (LOD) value of the likelihood that there is genetic linkage at that marker relating to the phenotype. The statistically significant LOD value used for diploid cells is usually 3, which suggests that the odds are a thousand to one in favour of genetic linkage to the phenotype (p < 0.001) (Van Ooijen, 1999).

The progeny of the cross between human serum resistant group 2 *T. b. gambiense* strain STIB386 and the human serum sensitive *T. b. brucei* strain STIB247 created to formulate the genetic map of STIB386 may prove useful to investigate the human serum resistance phenotype in group 2 *T. b. gambiense*. STIB247 is predominantly homozygous, including the majority of the 127 microsatellite markers (93%), in contrast to STIB386 which is heterozygous for all of them (MacLeod *et al.*, 2005b). While there will be an inherited contribution to the progeny phenotypes attributable to the STIB247 parent, the high homozygosity of this line means there is only a 7% chance it will segregate in the progeny. This means that differences in the phenotypes of the

progeny are more likely to be attributed to whichever of the chromosome homologues the individual progeny have inherited from the heterozygous STIB386 parent.

Using this methodology, the relationship between the phenotype and the inherited haplotype of the progeny can be ascertained. The simplest model of inheritance is one in which one parent is heterozygous for a single gene involved in a phenotype. The relevant alleles will segregate in the progeny so that the progeny will either inherit the allele or not and likewise the phenotype will either be inherited or not. By comparing the parental markers inherited by the progeny, markers more likely to be possessed by progeny displaying the phenotype suggests that genes close to those markers are involved with the phenotype. While the technique can be used to reveal phenotypes arising from single genes, such analysis is also ideal for the study of complex phenotypes that arise due to multi-gene interactions to create a quantitative trait. This is revealed as a continuous phenotypic response in the progeny, rather than the discrete presence or absence of the phenotype in question. In this case the degree to which the phenotype is displayed in each progeny is also considered in conjunction with the proportion of progeny inheriting the marker. Several mathematical and software solutions exist for quantitative trait locus analysis and have already been validated in *T. brucei*, revealing a major locus implicated in host pathogenesis (Morrison *et al.*, 2009).

The major strength of the linkage mapping approach is that it can identify genes involved in a phenotype where no obvious candidate genes exist. One such phenotype is human serum resistance. Group 2 *T. b. gambiense* can infect humans while *T. b. brucei* cannot. Fortunately a genetic cross has been performed in the laboratory between a group 2 *T. b. gambiense* and a *T. b. brucei* leading to the generation of hybrid progeny that can be used for a linkage analysis (Cooper *et al.*, 2008). An important consideration for QTL analysis with the group 2 *T. b. gambiense* progeny is that the phenotype in the parent is variable (Gibson, 1986). However, it is possible to identify loci of interest despite large amounts of variation if there is a strong heritable and therefore genetic component (Glazier *et al.*, 2002). This has been demonstrated in numerous human diseases in which there is a strong environmental or social components in addition to the genetic basis, such as type 2 diabetes mellitus (Glazier *et al.*, 2002). Preliminary QTL analysis for human serum resistance has been performed using the progeny from the group 2 *T. b. gambiense* and *T. b. brucei* cross (A. Macleod, unpublished data). The assay used to phenotype the progeny measured percentage survival after four hours incubation in human serum using a commercially

available system. This assay measures the fluorescence ratios of two different dyes that stain nuclear material (Turner *et al.*, 2004).

The two dyes used were RedDead which exhibits red fluorescence under excitation but is unable to cross an intact plasma membrane. The second dye, Syto 10, fluoresces green at the same excitation wavelength and is able to cross the plasma membrane. It was shown that progeny that were human serum sensitive would possess compromised plasma membranes after 4 hours exposure to human serum as they lysed (Turner *et al.*, 2004). A compromised plasma membrane allow both the RedDead and Syto 10 dyes to enter the cell and stain nuclear material red/green (observed as yellow) and conversely progeny that were human serum resistant would still possess intact plasma membranes and their nuclear material would only be stained green. The proportion of red and green cells can be quickly and accurately measured by fluorometry and compared to the control of 100% dead cell killed by immersion in ethanol. During the assay, cells were buffered in phosphate buffered saline (PBS) supplemented with glucose to reduce any cross-contamination in fluorescence caused by the HMI9 medium normally used to culture the trypanosomes. Forty progeny created in the formulation of the group 2 *T. b. gambiense* genetic map were assayed using this protocol (**Figure 5.1**). The phenotype segregates strongly, indicating the presence of a strong single gene effect. QTL analysis indicated strong linkage for a section of chromosome 2 that was involved in the human serum resistance phenotype measured by this assay (A. Macleod, unpublished data).
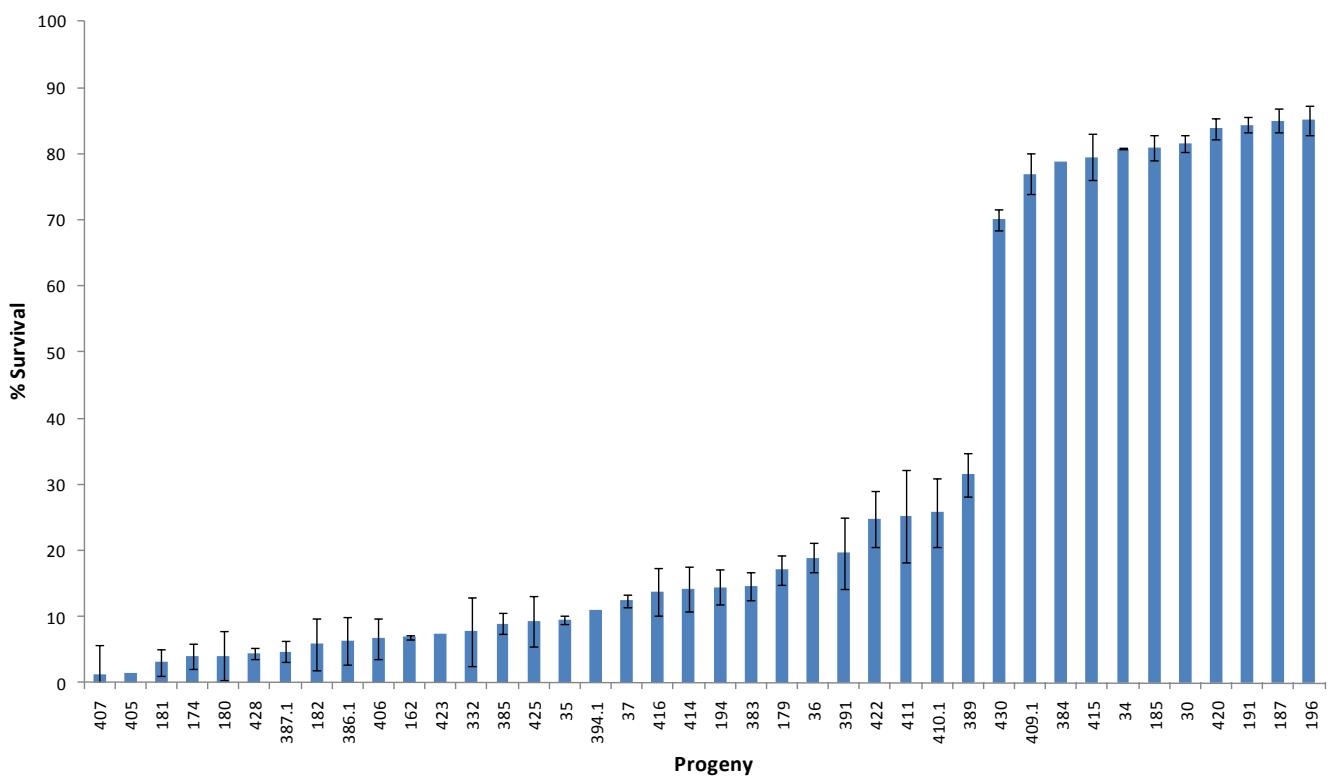
Figure 5.1 The percentage survival of different clones of progeny from the *T. b. brucei* strain STIB247 and group 2 *T. b. gambiense* strain STIB386 cross after exposure to human serum for 4 hours. Survival was assayed using the ratio of red:green fluorescence using RedDead/Syto 10 fluorescent dyes buffered in PBS supplemented with 0.1% glucose. The standard error is shown. Progeny clearly segregate into two distinct populations, one showing high survival, the other exhibiting sensitivity to the assay.

This large area containing approximately 300 genes was subsequently fine mapped to an area containing just 30 genes (A. Macleod, unpublished data). There are several candidate genes located in this area, including a family of tandemly repeated genes that encode for Invariant Surface Glycoprotein 65 (ISG65) - a receptor of unknown function with approximately 60,000 copies expressed on the cell surface (Ziegelbauer & Overath, 1993).

In completely separate parallel studies, low-temperature binding assays have indicated that there are two possible receptors for TLF-1; a high affinity receptor that has approximately 350 copies per parasite and a low affinity receptor that has 60,000 receptors per cell (Drain *et al.*, 2001). The high affinity receptor, HpHbR, has since been identified and characterised while the low affinity receptor remains elusive (Vanhollebeke *et al.*, 2008). It was hypothesised that ISG65 was the low affinity receptor for TLF-1 and the gene was taken forward for analysis as the prime candidate for

involvement with the human serum resistance phenotype in group 2 *T. b. gambiense*. Progeny that were identified as sensitive to human serum consistently possessed one particular haplotype of ISG65 genes. This haplotype was also shown to be possessed by a group 2 *T. b. gambiense* STIB386 selfer line that was homozygous for this region of chromosome 2. This STIB386 strain formed by self-fertilisation was shown to be sensitive to human serum in the same assay used to identify the hybrids (A. Macleod, unpublished data). These data heavily implicated ISG65 as a potential candidate for involvement in human serum resistance in group 2 *T. b. gambiense*.

However, shortly after starting this thesis it was noted by phase microscopy that there was significant cell death in some trypanosome lines compared to others when buffered in PBS media. The cell death was not revealed by the RedDead/Syto 10 fluorescence assay as visual cell counts were not performed during the experiment and instead comparative proportions of fluorescence were used as the assay metric and so a general decrease in cell viability occurred in both test and control cells (Turner *et al.*, 2004). This large amount of cell death may have an impact on any perceived QTL from the original data as the assay was potentially measuring an amalgam of two phenotypes rather than one - human serum resistance and PBS sensitivity. In addition, questions were raised concerning the consistency of the assay procedures, with several formulations of PBS being used at different points and the addition of variable chilling periods prior to the assay. This chapter investigates the effects of buffering human serum resistance assays in PBS compared to HMI9 medium that can sustain trypanosome growth in culture. A new assay using modified HMI9 and visual cell counting was then formulated and preliminary phenotype studies on the progeny of the *T. b. brucei* and group 2 *T. b. gambiense* cross were performed using this new method to conduct a new QTL analysis.

5.2 Results

5.2.1 Differences between assays

It is known that trypanosomes can survive for several hours in PBS supplemented with glucose and it is commonly used as a laboratory buffer. However the strain routinely used in laboratory experiments is the highly derived Lister 427 line (Peacock *et al.*, 2008). This line is heavily laboratory adapted and is able to survive in media that some trypanosome strains cannot, for example unmodified powdered HMI9 (L. Sweeny, personal communication). Often adaptation to culture of new lines requires specific and customised media formulations (Hirumi & Hirumi, 1989).

It has previously been ascertained that the strains used in this thesis grow best in culture in the presence of modified HMI9 (**Chapter 2: Materials & Methods**). It is possible that strain specific differences in response to the PBS buffer used in the assay was responsible for the original QTL rather than response to human serum. In order to ascertain how much affect the experimental buffer was potentially exerting on the assay, it was first necessary to distinguish if the parental lines differed in their response to the PBS media. If this was the case then the measured linkage to chromosome 2 is at least partly due to exposure to PBS and the presence of FBS merely reduces the killing effect of PBS. However, there was no significant difference in survivorship between STIB247 and STIB386 in either PBS or PBS supplemented with 0.1% glucose (Compared using a GLM with survival as the response factor; PBS: $F_7=1.58$ p=0.327, PBG+0.1% Glucose: $F_7=2.13$ p=0.239) (**Figure 5.2**).
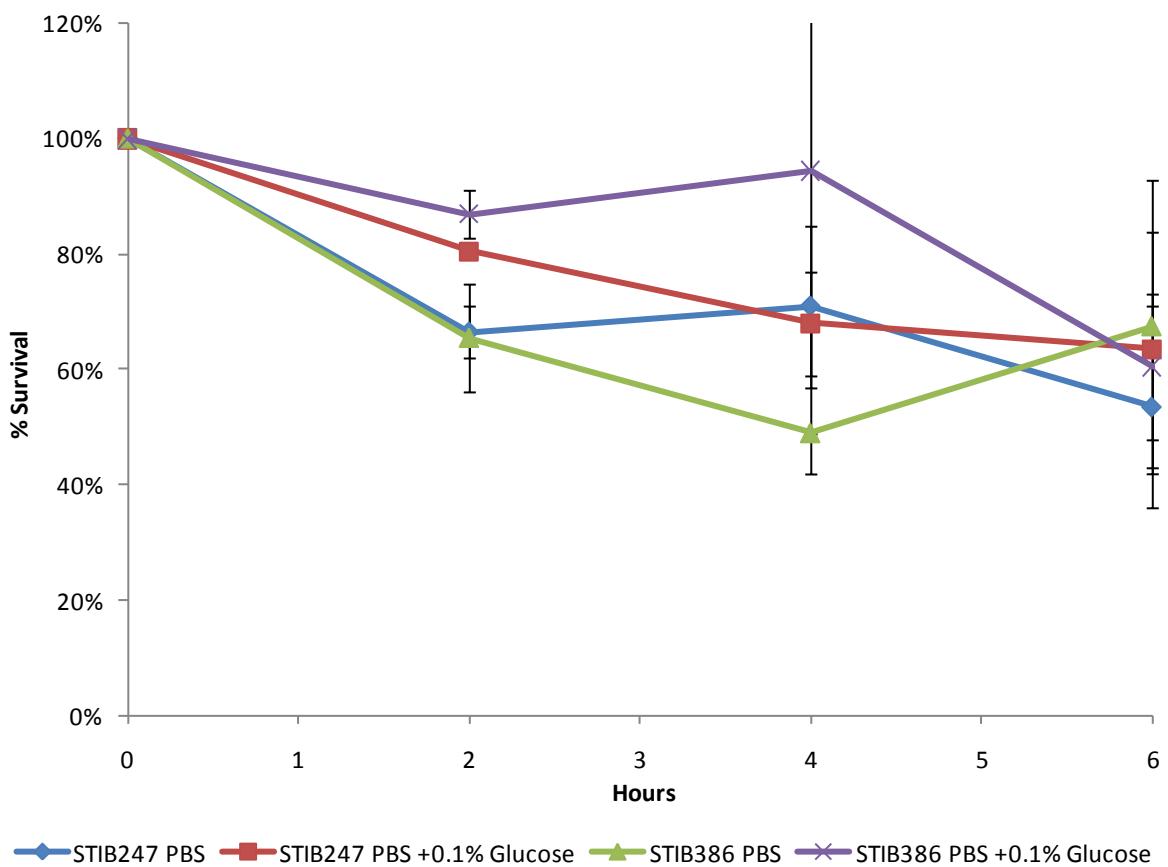
Figure 5.2 The percentage survival of the parental strains STIB247 and STIB386 over a time course in serum free PBS and PBS supplemented with 0.1%. The standard error of the mean is shown (n=2).

After discounting PBS as the sole determining variable involved in survivorship in the original assay, the response of trypanosomes in buffer with serum was investigated for several lines to quantify the differences between the PBS and modified HMI9 buffer assay. Assays were performed for the same 4 hour time course utilised in the original assay for both PBS with 0.1% glucose and modified HMI9. Trypanosomes exposed to foetal bovine serum (FBS) grew equally well in PBS with glucose and HMI9 buffer (**Figure 5.3a**). As expected, trypanosomes exposed to human serum showed varying amounts of death after 4 hours (**Figure 5.3b**). However, the survivorship seen in the PBS buffered assays in the presence of human serum was considerably lower. For example, the parental group 2 *T. b. gambiense* strain STIB386 cells derived from the same mouse passage showed approximately 45% survival after four hours exposure to human serum in modified HMI9 but less than 10% survival after exposure in the PBS with glucose buffer. This indicates that while the buffering of cells in PBS does not directly cause cell death, it does increase the effects of human serum on the strains.
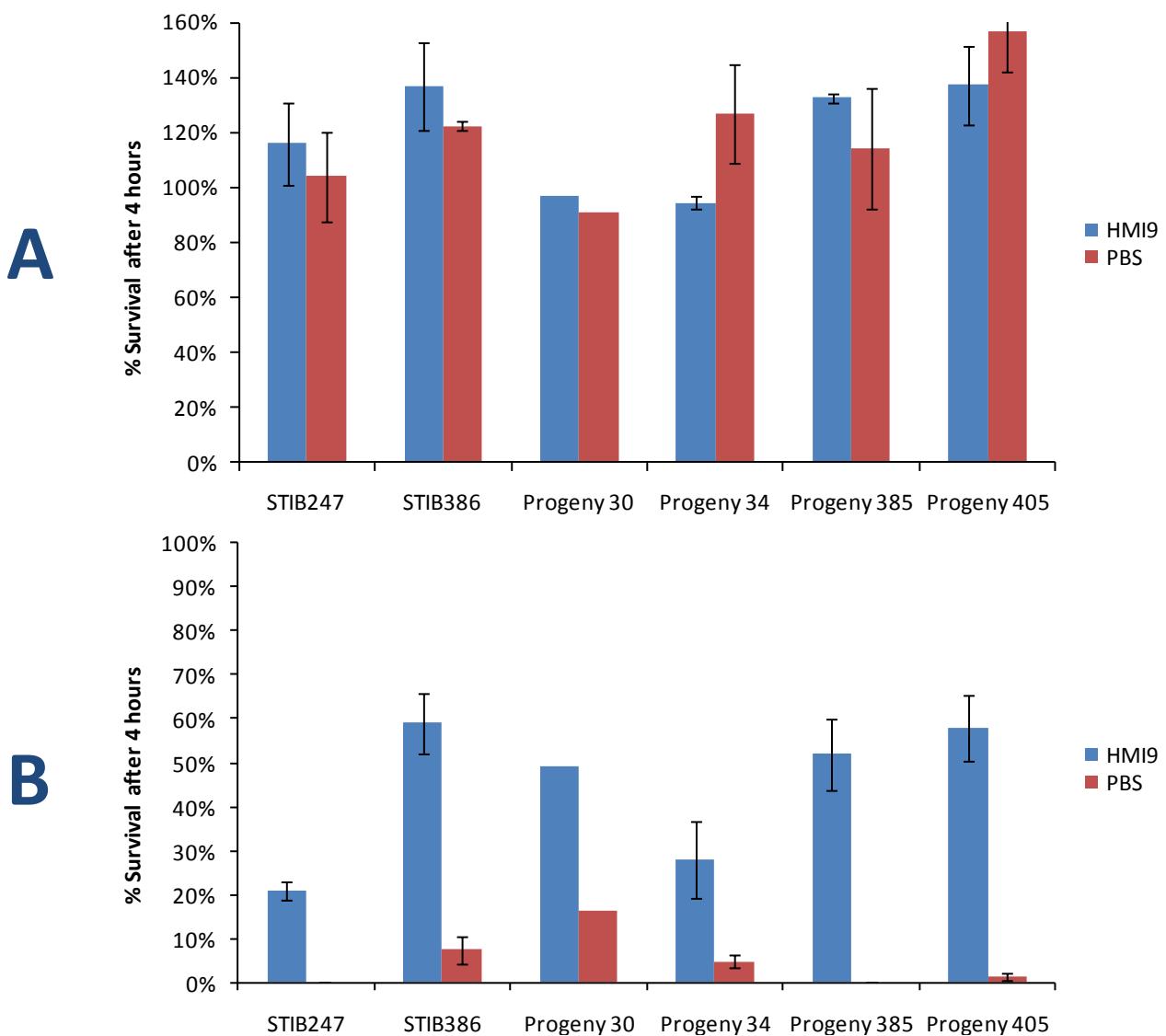
Figure 5.3a The percentage survival of the parental strains STIB247 and STIB386 and several progeny after 4 hour exposure to 25% FBS in both PBS with 0.1% glucose and modified HMI9 buffer. The standard error of the mean is shown.

Figure 5.3b The percentage survival of the parental strains STIB247 and STIB386 and several progeny after 4 hour exposure to 25% human serum in both PBS with 0.1% glucose and modified HMI9 buffer. Standard error of the mean is shown.

5.2.2 Differences between PBS & HMI9 buffers that may contribute to the phenotype

There are clear differences in the degree of lysis in the presence of human serum between PBS with glucose and the modified HMI9 buffer. Several factors were examined in order to investigate which may be contributing to the differences in phenotype response. A prime candidate that could be involved in the increased cell death by human serum between PBS and HMI9 was final glucose concentration. The PBS used in the original assays used 0.1% glucose by volume. It is known that low glucose concentration can cause death in trypanosomes with a phenotype superficially similar to human serum lysis, exhibiting gross changes in morphology and a swollen internal vacuole (Ter Kuile & Opperdoes, 1991). In the case of glucose starvation though, this large vacuole is comprised of a large vesicle at the flagellar pocket rather than swelling of the lysosome that is caused by the presence of normal human serum. Supplementing PBS with glucose greater than that present in modified HMI9 (≈0.5% by volume) resulted in the same increased sensitivity to human serum when in PBS compared to modified HMI9 buffer (**Figure 5.4**). This result suggests that lack of glucose is not the determining factor. However both the parental line STIB386 and progeny clone number 30 showed higher capacities to resist lysis compared to the other lines, suggesting that there may a heritable component that was measured by the original QTL analysis (**Figure 5.3b** & **Figure 5.4**).
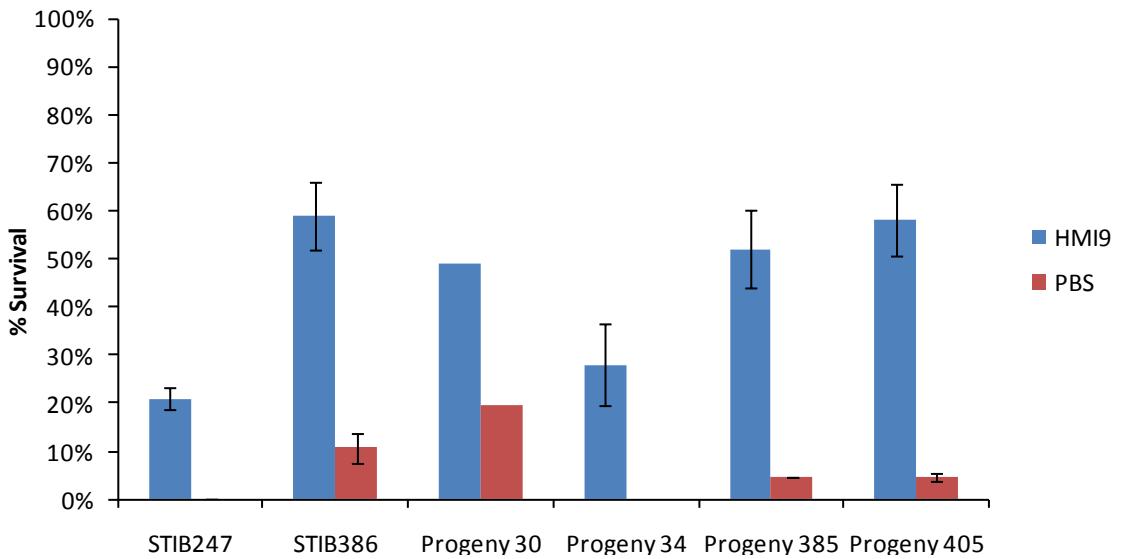
Figure 5.4 The percentage survival of the parental strains STIB247 and STIB386 and several progeny after 4 hour exposure to 25% human serum in both PBS supplemented with 1% glucose and modified HMI9 buffer. The human serum survival assays were normalised against the equivalent FBS controls and the standard error of the mean is shown (n=2, except for progeny30 for which n=1).

A second major difference between PBS and HMI9 that may be a factor is the ionic component of the buffer, specifically chloride and calcium ions. It is known that chloride ions are intimately involved in the lysis process (Pays *et al.*, 2006). HMI9 contains considerably more chloride ions in the form of $CaCl_2$ then PBS (0.165mg/ml) and this may impact on the osmotic potential of the cells. There was no difference however in lysis between normal PBS with glucose and PBS supplemented with the approximate equivalent $CaCl_2$ concentration found in modified HMI9. Trypanosomes in both buffers showed lower survival after exposure to human serum then when assayed in modified HMI9 buffer (**Figure 5.5**). This suggests that these ionic factors are not the sole determinant between PBS and modified HMI9.
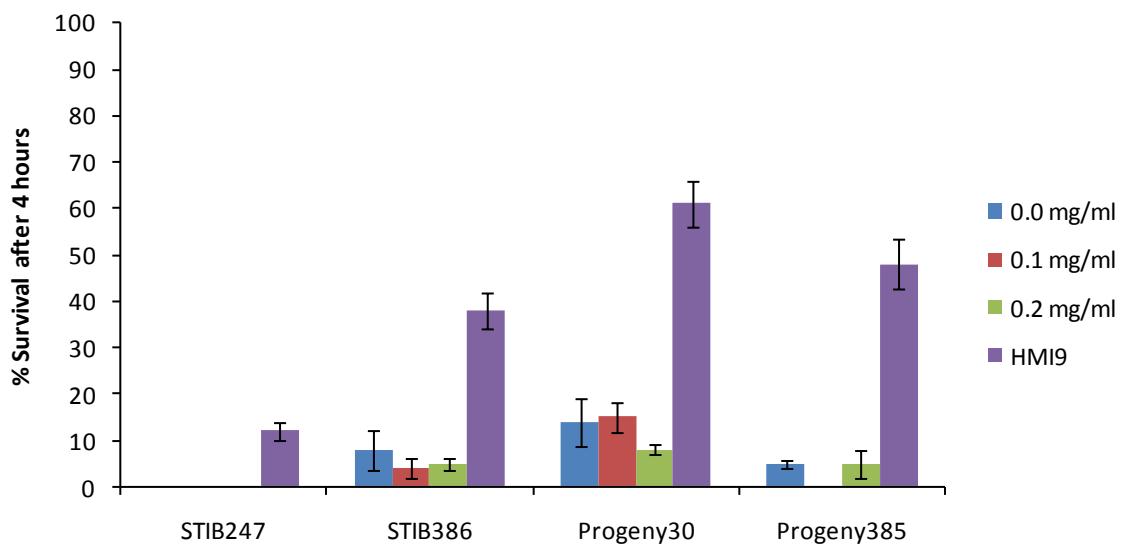
Figure 5.5 The percentage survival of the parental strains STIB247 and STIB386 and two progeny after 4 hour exposure to 25% human serum in PBS supplemented with 0.1% glucose and several concentrations of $CaCl_2$, in comparison to a modified HMI9 buffer. The human serum survival assays were normalised against the equivalent FBS controls and the standard error of the mean is shown (n=2).

The original PBS buffered assays featured a brief period of chilling before each experiment (maximally one hour). Trypanosome cells buffered in blood were kept on ice during transport between the animal host used to grow the strains and the assay site. This step was not performed in more recent assays. It is possible that the higher degree of killing exhibited by human serum in the PBS buffer may be attributable to this chilling period. One hypothesis is that lowering the temperature reduces the fluidity of the trypanosome plasma membrane. The period of chilling does not have an effect on the survival of trypanosomes exposed to human serum, although there is markedly less survival in assays buffered with PBS media rather than modified HMI9 (**Figure 5.6**).
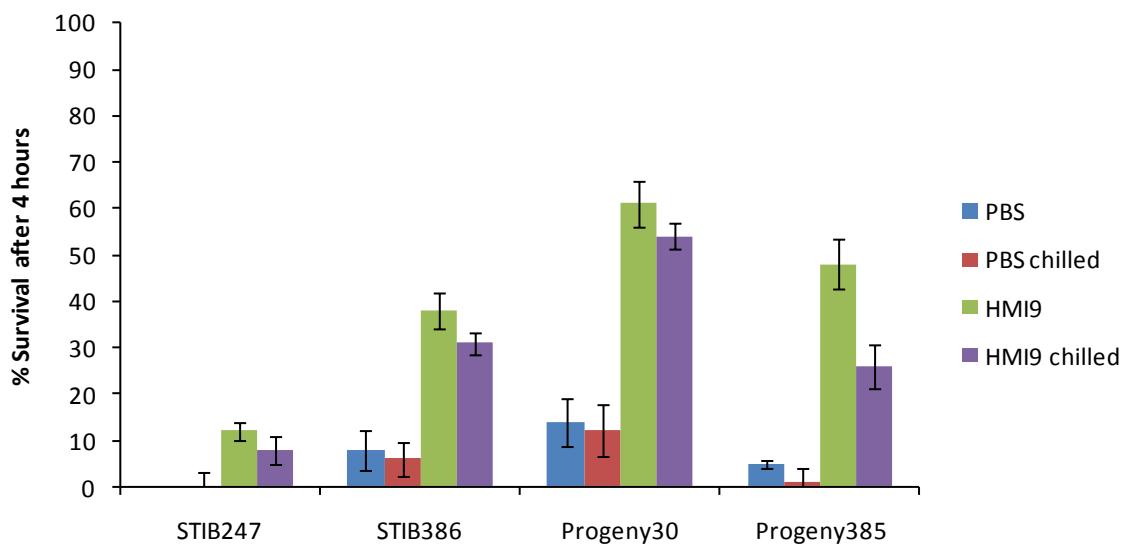
Figure 5.6 The percentage survival of the parental strains STIB247 and STIB386 and two progeny after 4 hour exposure to 25% human serum in PBS supplemented with 0.1% glucose or modified HMI9 buffer. Cells were either assayed immediately after removal from the animal host or subjected to a one hour chilling period to mimic the original assay. The human serum survival assays were normalised against the equivalent FBS controls and the standard error of the mean is shown (n =2)

In addition to the described factors, cell density and the experimental vessel used (either 1.5ml eppendorf or 24 well plate) were also examined for one trypanosome strain. Neither was found to have an effect on human serum resistance assays for that strain when buffered in modified HMI9 (data not shown). It does not appear that any single difference between the modified HMI9 used to culture the cells and the PBS used in the original QTL analysis is key to determining the difference in phenotype. However the use of PBS as a buffer for the original assay appears to magnify the lytic effects of human serum, perhaps due to a stress response initiated by a combination of various features of PBS. This suggests that the original QTL identified was unlikely to be directly involved in human serum resistance in group 2 *T. b. gambiense*, but may indicate secondary effects linked to the phenotype that can modify it, though it is unclear at this stage what these may be.

5.2.3 Development of a new assay

As the potential number of assayable differences between modified HMI9 and PBS buffer could be vast it was decided to develop a replacement assay using modified HMI9 buffer. The amount

of human serum used for exposure during the assay was maintained at 25%. In order to determine the optimal time course for the new resistance assay, *T. b. brucei* strain STIB247 and group 2 *T. b. gambiense* strain STIB386 were assayed at various time points to determine when the two parents differed most. This time course was limited to a maximum of 8 hours to minimise any potential effects of adaptation to culture. It was shown that the greatest difference between survival percentages of the two parental lines in response to human serum occurred at 6 hours (data not shown).

In conjunction with Dr. N Veitch, University of Glasgow, multiple 6 hour survival assays were then performed on the parental lines STIB247 and STIB386 *ex vivo* (**Figure 5.7**). The *T. b. brucei* STIB247 proved to be consistently sensitive to lysis caused by short term exposure to human serum. The group 2 *T. b. gambiense* exhibited the characteristic variable phenotype (Gibson, 1986). However there is a statistically significant difference between the distribution and mean phenotype response between parental lines when analysed by a one-way ANOVA (F_{16}=22.09 p < 0.01). If this difference contains a heritable component when assayed in the progeny then the contribution of different alleles can be measured by QTL analysis.
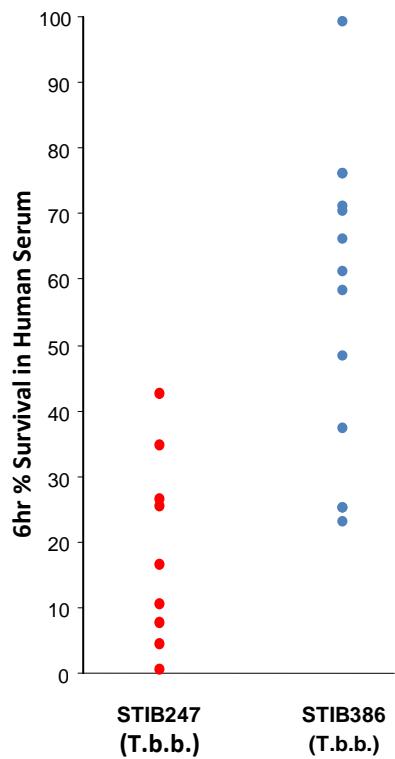
Figure 5.7 Percentage survival of several replicates of unselected parasite lines exposed to human serum for 6 hours normalised against trypanosomes exposed to foetal bovine serum. The assay was buffered in modified HMI9 media.

5.2.4 Phenotyping hybrids with new assay

To investigate if the human serum resistance phenotype segregates in the progeny of the STIB247/STIB386 cross and to test the feasibility of using a QTL approach to identify the genes involved, 25 progeny were assayed using the modified HMI9 buffered 6 hour survival assay. Approximately half of these assays were performed by Dr. N Veitch, University of Glasgow. Unlike the phenotype in the progeny previously described in the PBS buffered assays, the phenotypes measured using the modified HMI9 media buffer did not clearly segregate into two populations (**Figure 5.8**). Interestingly there are several progeny that show stronger expression of the human serum resistance phenotype then the STIB386 parent. This indicates transgressive segregation in which some progeny have possibly lost detrimental alleles during recombination and is evidence for a complicated multi-gene response (Rieseberg *et al.*, 1999).
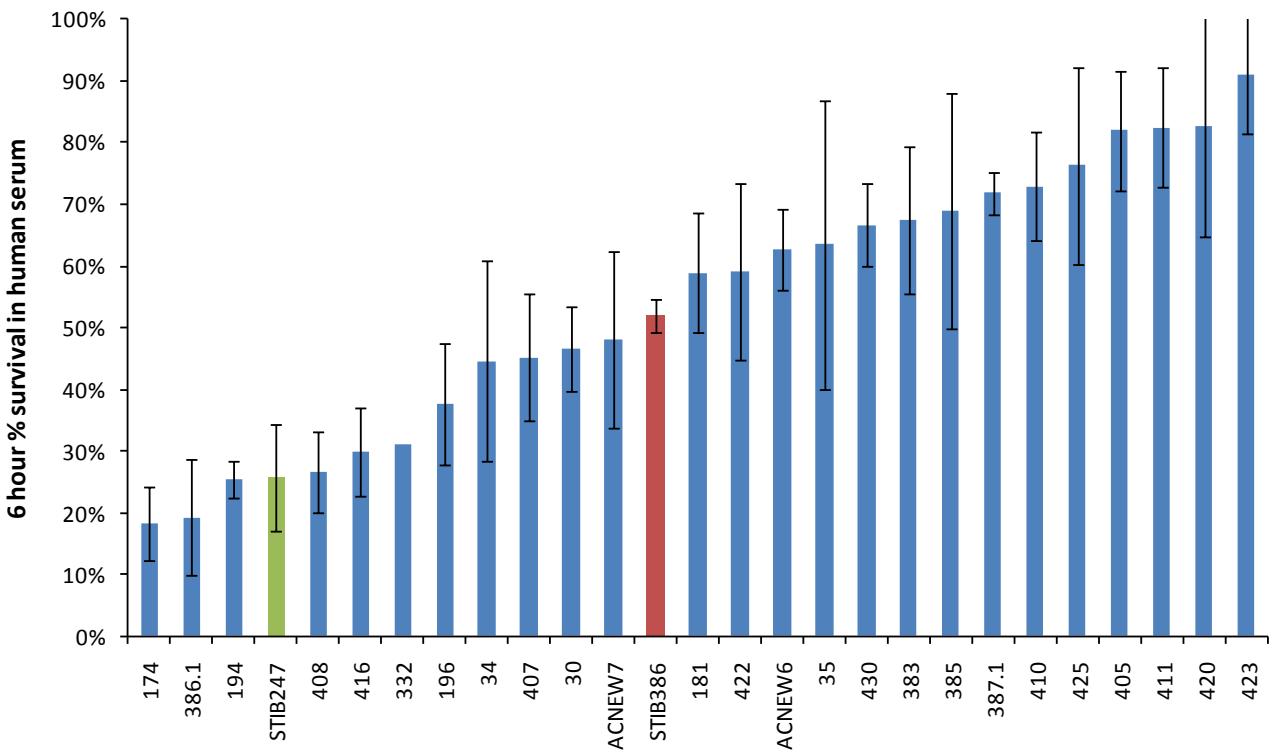
Figure 5.8 The percentage survival of different clones of progeny from the *T. b. brucei* STIB247 and group 2 *T. b. gambiense* STIB386 cross exposed to human serum (blue). The parental lines are also shown, STIB247 (green) and STIB386 (red) to highlight the transgressive segregation effect. Survival was assayed using the survival of cells visualised by microscopy. The experiments were buffered in modified HMI9 media. The standard error of the mean for each strain is shown, except for progeny 332 in which n=1 and the addition of further replicates has not been possible.

The mean percentage survival of the progeny is statistically different from each other using a one-way ANOVA (F_{65}=2.57, p=0.004) and there is greater variation between progeny than within biological replicates (R^2=56.79%). This indicates that there is a genetic component to the phenotype.

Analysis by P. Johnson, University of Glasgow has indicated that the most likely estimated heritable component for the phenotype in the progeny is 61% with a 95% confidence interval being approximately 35% to 84% (**Figure 5.9**). This evidence also strongly suggests the presence of a heritable component that would facilitate examination of the phenotype by QTL analysis.
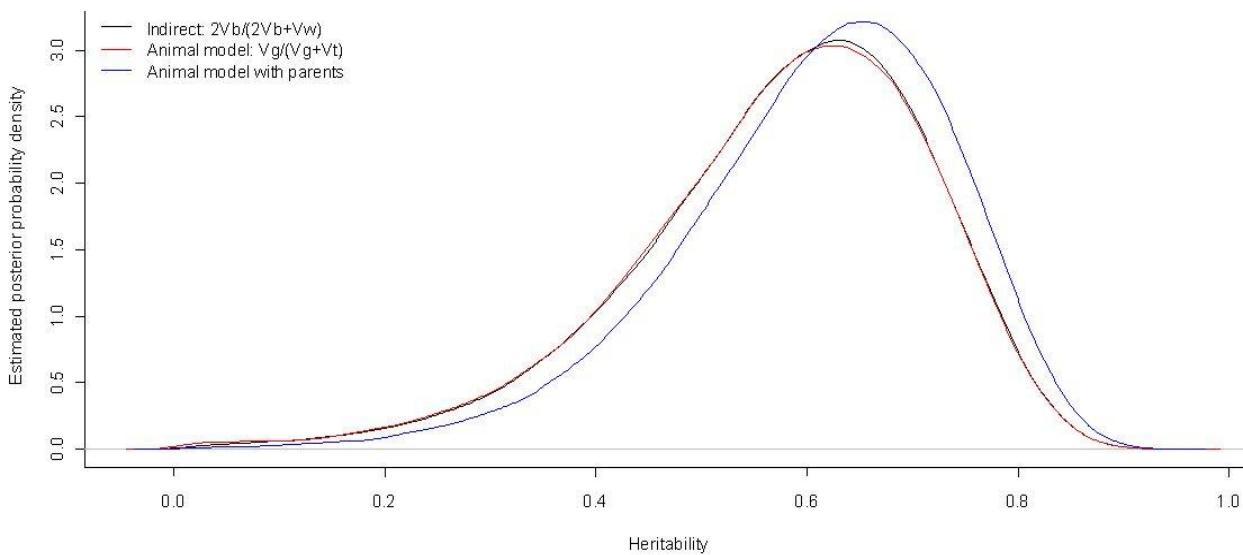
Figure 5.9 The estimated probability density of the relationship between phenotype for several values of heritability (P. Johnson, University of Glasgow). Repeat progeny assays were treated as twins and intra-clone variance was assumed to represent the environmental component of the phenotype. Different progeny clones were treated as siblings compared to each other and inter-progeny variance used to predict the genetic component. A standard inheritance model (Vg/(Vg+Vt) where Vg = genetic variance and Vt = total variance of the trait) was used to estimate the probability density that several heritability values fit the model (red). The mean value is 61% with a 95% confidence interval of approximately 35% to 84%. In addition a standard inheritance model including the parents (blue) and a non-standard model 2Vb(2Vb+Vw) (black) was also considered. The non-standard model is a custom formulation by P. Johnson, University of Glasgow in which Vb is the variance between progeny and Vw is variance within.

5.2.5 QTL Analysis

Linkage analysis was performed as a quantitative trait where no assumptions are made as to the number of loci involved using the group 2 *T. b. gambiense* genetic map (Cooper *et al.*, 2008). Genome wide linkage analysis was performed using the mean percentage survival of progeny after a 6 hour exposure to human serum. Intra-strain variance was included in the analysis to take non-genetic effects into account. Analysis of the 127 markers covering the 11 mega-chromosomes of *T. b. gambiense* identified one marker on chromosome 8 that has significant genome-wide linkage (P < 0.001) to the phenotype (**Figure 5.10**).
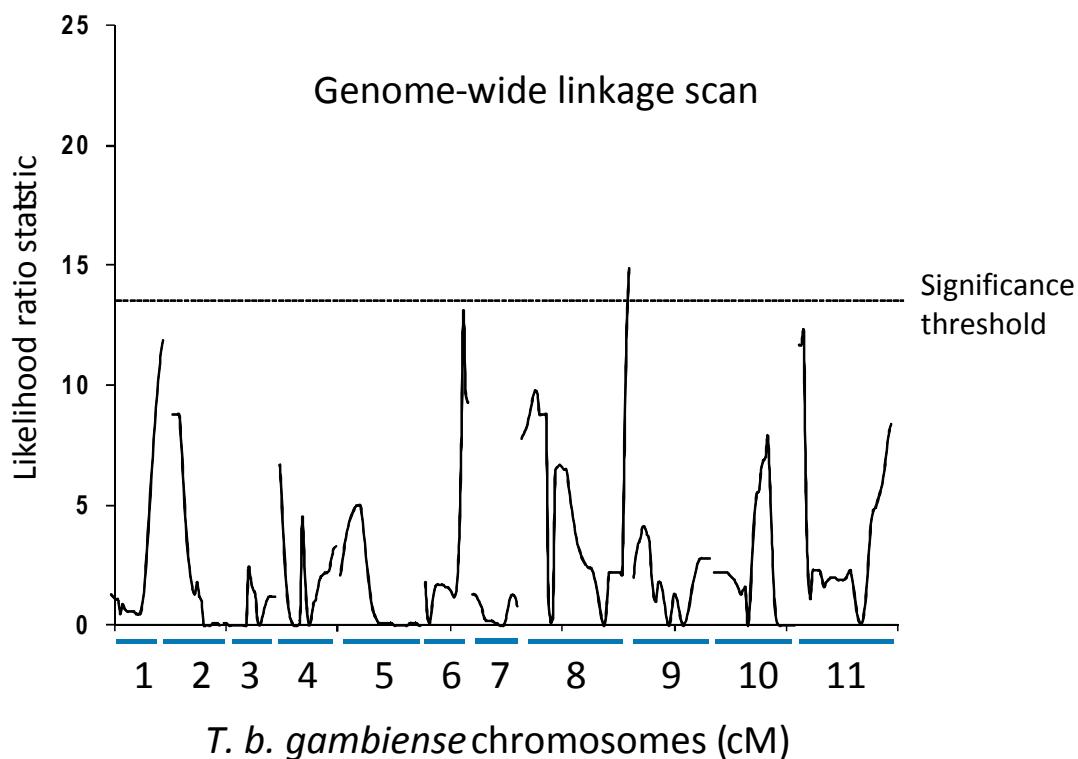
Figure 5.10 Genome-wide interval mapping analysis of human serum resistance. The likelihood ratio statistic is a measure of the significance of the linkage. The markers of the 11 chromosomal linkage groups are positioned sequentially along the horizontal axes with distances given in cM. There are between 7 and 24 markers per chromosome. The dashed line represents the threshold p = 0.001 (the probability of obtaining the LRS by chance from 1000 permutations, equivalent to LOD = 3.0).

Strongest linkage is found at microsatellite marker TB8/21 (Cooper *et al.*, 2008). The locus is estimated to account for 53% of the heritable variance seen in the human serum resistance phenotype between the progeny using MapManager QTX (Manly *et al.*, 2001). These results were confirmed using an alternative QTL approach of linear regression by QTL express (Seaton *et al.*, 2002) (data not shown). The identified locus spans 94,636bp of coding region and covers 30 genes (**Table 5.1**). The marker is the final chromosomal marker and therefore includes the telomere which has not been sequenced. This region may therefore also include non-annotated genes, ESAGs or VSGs. No other significant QTLs were identified by whole genome analysis, even when the effect on the phenotype associated with the QTL on chromosome 8 was fixed.

Table 5.1 Annotated genes found within the identified QTL on chromosome 8

Name	Description
Tb927.8.8060	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase
Tb927.8.8070	Hypothetical - expressed in procyclic form
Tb927.8.8080	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase
Tb927.8.8090	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase
Tb927.8.8100	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase
Tb927.8.8110	Hypothetical - zinc finger-like protein
Tb927.8.8120	Hypothetical - GPI signal peptide, localised to plasma membrane
Tb927.8.8130	Hypothetical
Tb927.8.8140	Small GTP-binding rab protein
Tb927.8.8150	Hypothetical -(De Maio, 1999) localised to plasma membrane
Tb927.8.8170	Hypothetical - localised to mitochondrial RNA binding complex 1
Tb927.8.8180	Hypothetical - localised to mitochondrial RNA binding complex 1
Tb927.8.8190	Hypothetical
Tb927.8.8200	Hypothetical – present in flagellum proteome
Tb927.8.8210	Hypothetical
Tb927.8.8220	Amino acid transporter
Tb927.8.8230	Amino acid transporter
Tb927.8.8240	Amino acid transporter
Tb927.8.8250	Amino acid transporter
Tb927.8.8260	Amino acid transporter
Tb927.8.8270	Hypothetical - PDEase-like
Tb927.8.8280	Hypothetical
Tb927.8.8290	Amino acid transporter AATP5
Tb927.8.8300	Amino acid transporter
Tb927.8.8310	Chaperone protein DnaJ
Tb927.8.8320	Hypothetical
Tb927.8.8330	Calpain
Tb927.8.8340	Hypothetical
Tb927.8.8350	Mitotic centromere-associated kinesin (MCAK)
Tb927.8.8360	Receptor-type adenylate cyclase GRESAG 4

5.3 Discussion

From these data it was not possible to determine any one factor that contributed to the increased human serum mediated cell lysis observed in PBS buffered assays compared to the equivalent modified HMI9 buffered assays. Such a large difference in survival may mask any possible linkage to the phenotype in question and casts considerable doubt on the conclusions of the first QTL analysis. In addition, the low survivorship was not revealed by the RedDead/Syto 10 fluorescent assay so it was determined that it was not suitable for this particular phenotype. However the linkage to chromosome 2 identified is statistically robust and several genes in the locus may be involved in the phenotype response, even if the phenotype is not directly related to lysis by human serum. There was no killing effect shown by the PBS buffer in the presence of FBS indicating that by itself PBS is not toxic to trypanosomes in the short-term. However, when trypanosomes are subjected to the presence of human serum, the added stressful effects imposed by the PBS buffer become apparent.

It would appear that a combination of the dual stresses of human serum and the PBS buffer is the phenotype being measured by the original fluorescence assay. If the response is largely mediated by stress, the most likely candidate to explain the trait is not the gene for the cell surface receptor ISG65 but the adjacent heat shock protein gene, *HSP78*. Different homologues of this gene segregate within the progeny in a manner identical to ISG65 and heat shock proteins are strongly implicated in responses to many stressful factors (De Maio, 1999; Maresca & Carratu, 1992). They are also usually transcription factors so may affect the expression of the genes responsible for the resistance mechanism found in group 2 *T. b. gambiense*. While HSP78 protein has been described as involved with a thermotolerance in yeast, its function in *T. brucei* is unknown (Krzewska *et al.*, 2001). It is also possible that any of the other 28 genes within the originally identified QTL contribute to the phenotype but this is outside of the scope of the thesis.

The possible linkage that the identified region of chromosome 2 and *HSP78* has to a buffer-mediated stress response is interesting in that it increases the sensitivity of resistant strains to human serum. If *HSP78* was validated by reverse genetic approaches to genuinely increase sensitivity to human serum then either drugs targeted to modify function of the protein or alternatively factors in the host bloodstream could be modified to initiate the HSP78 stress response and possibly allow the normal lytic factors to kill the parasite. Unfortunately heat shock proteins are a large ancestral gene family found throughout the eukaryotes and it is unlikely that a therapeutic agent would be specific enough to target just trypanosome HSP78. Eliciting the

stress response by modifying the host bloodstream is likely to also stress the patient. Finally, the process of examining the exact feature of PBS that causes the response is akin to finding a needle in the proverbial haystack. A more profitable approach is to utilise an assay that stresses trypanosomes less and allows the identification of genes involved directly in resistance to human serum in group 2 *T. b. gambiense*.

The newly derived assay using a modified HMI9 assay is hypothesised to be less stressful to trypanosomes than PBS in the short-term due to the fact that the strains used in these experiments can be quickly adapted to the media (L. Sweeny, personal communication). When survival to short-term exposure to human serum was assayed using direct observation in a modified HMI9 buffer, the survival phenotype did not segregate in the progeny in a clear manner as before. This indicates a more complicated multi-gene mediated response being displayed by the progeny than initially believed. There is a strong heritable genetic component that can be investigated using QTL analysis, although there is also a non-heritable component. The variability of the phenotype itself is an obvious cause of the non-inheritable component but it is possible that some of the phenotype may be attributed to epigenetic effects such as methylation of DNA, histone interactions along the chromosome or gene expression differences in a manner similar to *SRA*. If the QTL approach was applied to a genetic cross involving *T. b. rhodesiense* for example, there would only likely be a low heritable component. QTL analysis would be unlikely to identify the *SRA* gene directly, although it would highlight genes that respond to SRA, such as those that may cause the modification of transport of TLF within the cell shown by some strains when expressing SRA f OLI B to A (Oli *et al.*, 2006). There would also be suggestive trait linkage to the area of the genome preceding the sub-telomeric region in which *SRA* sits.

In the described investigation using group 2 *T. b. gambiense* progeny and a HMI9 buffered survival assay, a single locus on chromosome 8 was identified using QTL analysis that is associated with human serum resistance. This identified region contains 30 annotated genes, although it may also include the un-sequenced telomeric regions that might contain *ESAG*s, *VSG*s and other related genes. Fine mapping of the region at the sub-telomeric end of the locus will reveal if the telomeric region is the cause of the locus. The process to identify the 3' limits of the QTL is underway by locating extra markers that segregate in the progeny and can be used to further characterise the region. However, the DNA sequence in the sub-telomeric regions and telomeres contains many repeating structures, making it difficult to design unique PCR primers to these regions (Berriman *et al.*, 2005).

It is also necessary to reduce the number of candidate genes within the QTL to minimise wasted effort. One method to do this is to define specific properties, motifs or features of genes that might be required for the phenotype in group 2 *T. b. gambiense*. Unlike group 1 *T. b. gambiense*, the group 2 *T. b. gambiense* strain STIB386 has been shown to bind and internalise TLF (**Chapter 3**). It is possible that the proteins involved in the human serum resistance phenotype interact with one or more of the TLF components in a similar manner to SRA (Xong *et al.*, 1998). In order to perform this function, any putative inhibitory protein must be found on the cell surface or be a component of the endocytic pathway, to facilitate interaction with the serum lytic components. Of the 30 genes in this region, 12 appear to encode for proteins that have either signal peptides, transmembrane domains or GPI anchors, all of which are indicative of a cell surface location and as such are considered candidate genes. Several hypothetical proteins of unknown function may be important to the phenotype and need to be further characterised. None of the 30 genes in the region have significant homology to the *SRA* gene found *T. b. rhodesiense* although due to the fact the locus also includes the telomeric regions the presence of an SRA-like gene in the telomere cannot be ruled out and it is possible that an SRA-like gene sits in the telomere of chromosome 8. However, expression site switching has not been implicated in group 2 *T. b. gambiense* serum resistance (**Chapter 3**) suggesting that the expression of the resistance phenotype is not due to a gene in an telomeric expression site but rather due to variable expression of a gene in the non-telomeric regions of the mega-chromosomes.

The locus identified on chromosome 8 contributes to 53% of the heritable portion of the phenotype. This indicates that one or more other loci are also involved in determining the human serum resistance phenotype. Whilst these loci could lie in one of the regions of the genome not represented in the genetic map, there are also several suggestive linkages spread throughout the genome that may represent the location of loci involved to a lesser degree then the chromosome 8 locus although the power of the QTL would need to be improved with the addition of extra progeny to confirm these suggestive loci. Despite the presence of a variably expressed human serum resistance phenotype in group 2 *T. b. gambiense*, there is still a significant heritable component involved. Approximately 61% of the difference in the phenotypes of the progeny can be explained by these hereditable factors. These may be alleles or genes that need to be present to activate the variable component of resistance or alternatively alleles that are activated by the variable component. Any of these genes that are linked to the heritable component of human serum resistance in group 2 *T. b. gambiense* would make suitable targets for therapeutic opportunities.

The identification of a human serum resistance related QTL in group 2 *T. b. gambiense* has provided preliminary evidence for the existence of a genetic basis for the phenotype, although the work has been limited by time and the difficulties of formulating a new assay. The addition of more progeny to the study will improve the power of analysis and clarify the suggestive loci identified in addition to the chromosome 8 locus by the QTL analysis. A further 25 progeny have been created during the process to improve the group 2 *T. b. gambiense* genetic map (A. Cooper, unpublished data). They have been genotyped and are available for phenotyping and inclusion in QTL analysis at a future date. In addition, the loci identified on chromosome 8 by QTL analysis can be used to inform and select candidate genes received from ongoing transcriptomics studies between the isogenic sensitive and resistance group 2 *T. b. gambiense* strain STIB386. A number of genes have been shown to be differentially expressed between sensitive and resistant STIB386 using Digital Gene Expression (DGE) assays (N. Veitch, unpublished data). One candidate gene in particular, *Calpain*, lies within the identified QTL and is being taken forward for analysis by reverse genetics. Calpain is a proteolytic protein found in the lysosome of trypanosomes, although the substrate is unknown. This gene could not have been predicted to be involved in human serum resistance based on homology to other eukaryotic systems. This shows the power of utilising combinations of forward genetics approaches to identify novel genes for reverse genetics studies in a parasitic organism.

Chapter 6: The genetic relationship between the two groups of *T. b. gambiense* in Côte d'Ivoire

6.1 Introduction to the relationships between *T. b. gambiense*

Investigating questions concerning the human serum resistance phenotype in *T. b. gambiense* is complicated in that there are two distinct groups that differ in genetics and phenotype (Gibson, 1986). Group 1 *T. b. gambiense* causes chronic infections, is invariably human serum resistant and is by far the more prevalent group. It appears to be largely a disease limited to humans, although some animal reservoirs have been described (Felgner *et al.*, 1981; Gibson *et al.*, 1978; Mehlitz *et al.*, 1982). Group 2 *T. b. gambiense* is more virulent and exhibits a variable resistance mechanism similar to T*. b. rhodesiense* (Mehlitz *et al.*, 1982; Zillmann *et al.*, 1984), although it does not involve the *SRA* gene and appears not to involve expression site switching (**Chapter 3**). This group of *T. b. gambiense* has only been described in Côte d'Ivoire and always in geographical areas that overlap with that of group 1 *T. b. gambiense* (Gibson, 1986). No modern analysis using microsatellite markers of these field populations has been performed so the relationship between *T. b. brucei* and group 1 and 2 *T. b. gambiense* is currently unknown. One hypothesis for the origin of group 2 *T. b. gambiense* is that it arose due to a rare mating event between group 1 *T. b. gambiense* and *T. b. brucei*. This would explain how some group 2 *T. b. gambiense* have come to possess the diagnostic group 1 *T. b. gambiense* gene *TgsGP* (Radwanska *et al.*, 2002) and why the group 2 *T. b. gambiense* strain STIB386 possesses both a *T. b. brucei*-like *HpHbR* allele and a group 1 *T. b. gambiense*-like allele (**Chapter 3**). It would also reveal why group 2 *T. b. gambiense* foci have only arisen where there are group 1 *T. b. gambiense*. If this hypothesis is true then the population structure of a geographical focus containing *T. b. brucei* and both group 1 and 2 *T. b. gambiense* would be expected to show a degree of similarity between the three populations, with shared alleles between them. This would also mean that the human serum resistance mechanisms in both groups of *T. b. gambiense* are likely to be the same, although it is variably expressed in group 2 parasites. An alternative hypothesis is that group 2 *T. b. gambiense* is an extended host variant strain of *T. b. brucei* that is distinct from group 1 *T. b. gambiense* and possesses a novel human serum resistance mechanism. In this case there will be little overlap with alleles between the group 1 and 2 *T. b. gambiense* populations. This situation is paralleled in

T. b. rhodesiense foci in which the dominant human serum resistance mechanism is *SRA*, although several human infective but *SRA*-negative *T. brucei* isolates have been described from the same geographic region (De Greef *et al.*, 1989; Enyaru *et al.*, 2006).

Most research on *T. b. gambiense* field populations has focused on the more prevalent group 1. Previous studies with isoenzymes and AFLP have indicated that group 1 *T. b. gambiense* populations exhibit low genetic variation and are distinct from the local *T. b. brucei* population (Gibson *et al.*, 1980; Godfrey & Kilgour, 1976; Mathieu-daude & Tibayrenc, 1994; Mehlitz *et al.*, 1982; Stevens & Tibayrenc, 1996; Tait *et al.*, 1984; Zillmann *et al.*, 1984). This would appear to confirm theories that group 1 *T. b. gambiense* possesses a clonal population structure, despite the limitations inherent in the studies due to small sample number, limited diagnostic markers and little variation in those markers. More modern and comprehensive microsatellite genotyping techniques have since been utilised to confirm that group 1 *T. b. gambiense* populations are indeed clonal and distinct from local *T. b. brucei* populations (Koffi *et al.*, 2009; Morrison *et al.*, 2008). It was also shown that while group 1 *T. b. gambiense* are clonal at several disease foci, they are distinctly different between geographic locations (Morrison *et al.*, 2008). This may be because group 1 *T. b. gambiense* was more prevalent in the past but has now been reduced to genetic "islands" that have diverged due to random mutation. Alternatively the situation may have arisen as a result of the natural spread of the disease if this occurred in a stepwise manner. If transfer to new foci was limited, a founder effect would lead to diverging genotypes at each focus by random mutation. A less likely explanation is that an invariant human infectivity trait has evolved on separate occasions. Data presented in this thesis show that SNPs in the ORF and 3' UTR of the HpHbR receptor involved in TLF-1 uptake are identical between group 1 *T. b. gambiense* from different countries, which would suggest a common ancestor rather than this particular human infectivity arising on multiple occasions (**Chapter 3**).

Several diagnostic tests have been described to distinguish between group 1 *T. b. gambiense* and other *T. brucei*. Initially a blood incubation infectivity test (BIIT) was considered sufficient (Rickman & Robson, 1970; Targett & Wilson, 1973). Trypanosomes were exposed to human serum and then inoculated into a number of rodents. If all of the rodents become infected then the isolate was considered invariably resistant and deemed to be a group 1 *T. b. gambiense*. If only some rodents become infected then the resistance phenotype was considered variable and the strain a group 2 *T. b. gambiense*. If no rodents become infected then the strain was considered to be fully sensitive and a *T. b. brucei*. There are of course inherent problems with this

assay technique when trying to apply it for quick and easy identification of the different *T. brucei* sub-species in the field. A more formalised and rapid test to identify group 1 *T. b. gambiense* was the card agglutination test for trypanosomiasis (CATT). This exploits the appearance of antibodies to variable antigen type (VAT) LiTaT 1.3 that is usually found in patients infected by group 1 *T. b. gambiense* (Magnus *et al.*, 1978). However, it was subsequently found that LiTaT 1.3 is absent from some group 1 *T. b. gambiense* populations in Cameroon and so is not truly diagnostic for the sub-species (Dukes *et al.*, 1992). While the CATT has proven useful for monitoring outbreaks at some group 1 *T. b. gambiense* foci, *T. b. gambiense* infections were found in patients that were CATT negative in northern Uganda. This was blamed on the low sensitivity of the CATT and this must be considered when using it to monitor disease (Enyaru *et al.*, 1998). A second diagnostic marker proposed was PCR with primers sensitive to the *T. b. gambiense*-specific flagellar pocket glycoprotein (*TgsGP*) gene (Radwanska *et al.*, 2002). This gene was found to be conserved in all group 1 *T. b. gambiense* and appears to have evolved from *VSG* Tb10.v4.0178 (Gibson *et al.*, 2010). The function of the gene is unknown although it was initially believed to be involved in human serum resistance in *T. b. gambiense* until experiments ectopically expressing *TgsGP* in *T. b. brucei* did not confer resistance to human serum (Berberof *et al.*, 2001). The use of *TgsGP* PCR as a diagnostic for group 1 *T. b. gambiense* is hampered by the fact that some, but not all, group 2 *T. b. gambiense* are also positive for the gene (Radwanska *et al.*, 2002). This would suggest that group 1 and group 2 *T. b. gambiense* are related, although information concerning the relationship between the two groups is sparse.

There have been some field studies investigating the population genetics of group 2 *T. b. gambiense*, usually as a side effect of examining group 1 *T. b. gambiense*. The group 2 clade was proposed after it was noted that some human infective *T. b. gambiense* in a Côte d'Ivoire focus showed greater genetic variation and appeared more similar to local *T. b. brucei* (Gibson *et al.*, 1980; Godfrey & Kilgour, 1976; Mathieu-daude & Tibayrenc, 1994; Mehlitz *et al.*, 1982; Stevens & Tibayrenc, 1996; Tait *et al.*, 1984; Zillmann *et al.*, 1984). The greater variation in genetic markers and similarity to the *T. b. brucei* population suggested that group 2 *T. b. gambiense* may also be genetically competent, in contrast to the clonal group 1 *T. b. gambiense* population in which there is no evidence for genetic recombination. The possibility that different mating structures may exist in the two groups of *T. b. gambiense* has implications in evaluating the relationships between them and the evolution of traits in the population, including human infectivity. It is possible that this trait has evolved twice in separate populations of the *T. b. gambiense* or

conversely evolved once but due to mating events has become invariant in group 1 and variable in group 2 *T. b. gambiense*.

Like other protozoan parasites, *T. brucei* can undergo mating but it is not obligatory (Sternberg *et al.*, 1988; Tait, 1980; Tait & Turner, 1990). Also when mating does occur, recombinant products of self-fertilisation can be found (Tait *et al.*, 1996). These genetic features led to several population structures in the field; clonal (Mathieu-daude & Tibayrenc, 1994; Morrison *et al.*, 2008), epidemic (Hide *et al.*, 1994; MacLeod *et al.*, 2000; MacLeod *et al.*, 2001; Stevens & Tibayrenc, 1996) and panmictic (Tait, 1980). These different population structures can have an effect on human infectivity in a focus. For example, *T. b. rhodesiense* isolates in Uganda exhibit a distinct clonal population despite occurring in a background of a larger mating *T. brucei* population (MacLeod *et al.*, 2000). In a separate *T. b. rhodesiense* focus in Kenya, the *T. b. rhodesiense* and *T. b. brucei* populations were found to be genetically similar, although the *T. b. rhodesiense* isolates had less genetic variation (Tait *et al.*, 1985). These facts would suggest that *T. b. rhodesiense* is a host variant of *T. b. brucei* that undergoes clonal expansion when the *SRA* gene is transferred into an ecologically advantageous genetic background although the apparent lack of variability could be due to sampling from a restricted subset of hosts that the parasite infects. If *T. b. rhodesiense* is indeed a host variant *T. b. brucei* then this would have important implications on disease control due to evidence for mating occurring in some *T. b. brucei* populations (Hide *et al.*, 1994; MacLeod *et al.*, 2000; Stevens & Tibayrenc, 1996; Tait, 1980). Up to 18% of infected cattle in a disease focus carry *SRA* positive *T. b. rhodesiense* (Welburn *et al.*, 2001), allowing for the possibility of re-assortment of the genotypes associated with *SRA* and increasing the chance of new epidemics to emerge by allowing more virulent strains to evolve. Answering the question of whether *T. b. gambiense* also undergoes mating in the field is becoming more urgent as the geographical ranges of the two human infective sub-species, *T. b. gambiense* and *T. b. rhodesiense*, move closer together in Uganda (Fevre *et al.*, 2005; Picozzi *et al.*, 2005). If mating can occur between *T. b. gambiense* and *T. b. rhodesiense*, either directly or via a *T. b. brucei* intermediate, this could lead to the generation of a new hybrid sub-species that has multiple human serum resistance mechanisms. More importantly, the two sub-species show different tolerances to trypanosome treatment options and, if mating is possible, it will raise the possibly that any hybrid generated would be highly tolerant to several drug families (Barrett *et al.*, 2003).

This chapter uses microsatellite genotyping techniques to investigate the population structure of a collection of *T. brucei* isolates from a *T. b. gambiense* focus in Côte d'Ivoire collected by D.

Mehlitz, Freie Universität Berlin. The isolates have been characterised by BIIT and defined as group 1 *T. b. gambiense* (fully resistant), group 2 *T. b. gambiense* (intermediate resistance) or *T. b. brucei* (fully sensitive) (**Table 6.1**). In addition, the *T. b. gambiense* reference strain DAL972 was also analysed (Jackson *et al.*, 2010). All of these samples have been collected from a similar time period (1978 – 1983) and from the same geographic area (Daloa, Côte d'Ivoire). This large collection of isolates allows analysis of the population structure of the sub-species present in this geographic area at this time point and to possibly elucidate the relationships between them, especially those between group 1 and 2 *T. b. gambiense* and the implications for the evolution of the human infectivity trait.

Table 6.1 The *T. brucei* collection sorted by serum resistance phenotype, indicating isolate name, host, phenotype and the predicted sub-species & group based on a 5 rodent BIIT performed by D. Mehlitz, Freie Universität Berlin. The presence/absence of the *TgsGP* gene was confirmed by a diagnostic PCR performed alongside the genotyping process.

Isolate	Host	HSR Phenotype	Predicted sub-species from BIIT	*TgsGP*
MHOM/CI/78/TH112	Human	I	Group 2	✗
MHOM/CI/78/TH170	Human	I	Group 2	✗
MKOK/BF/80/KK1	Kob	I	Group 2	✗
MKOK/BF/80/KK17	Kob	I	Group 2	✗
MHOM/CI/78/TH149	Pig	I	Group 2	✗
MSUS/CI/78/TSW100	Pig	I	Group 2	✗
MSUS/CI/78/TSW158	Pig	I	Group 2	✓
MSUS/CI/78/TSW168	Pig	I	Group 2	✗
MSUS/CI/78/TSW182	Pig	I	Group 2	✗
MSUS/CI/78/TSW190	Pig	I	Group 2	✗
MSUS/CI/78/TSW209	Pig	I	Group 2	✓
MSUS/CI/78/TSW308	Pig	I	Group 2	✓
MSUS/CI/78/TSW77	Pig	I	Group 2	✗
MBOI/BF/80/TC125	Cattle	R	Group 1	✗
MBOI/BF/80/TC126	Cattle	R	Group 1	✗
MHOM/CI/78/TH126	Cattle	R	Group 1	✗
MHOM/CI/78/DAL069	Human	R	Group 1	✗
MHOM/CI/78/DAL072A	Human	R	Group 1	✗
VAVOUA/83/DAL542 (Adzam)	Human	R	Group 1	✓
VAVOUA/84/DAL740 (Brazo)	Human	R	Group 1	✓
GAGNEA/83/DAL595 (Kide)	Human	R	Group 1	✓
DALOA/83/DAL598 (Kosu)	Human	R	Group 1	✓
DALOA/83/DAL642 (Lisa)	Human	R	Group 1	✓
DALOA/83/DAL645 (Mago)	Human	R	Group 1	✓
DALOA/83/DAL403 (Sakon)	Human	R	Group 1	✓
DALOA/83/DAL633 (Seval)	Human	R	Group 1	✓
BOUAFLE/83/DAL587 (Yage)	Human	R	Group 1	✓
MSUS/CI/78/TSW 115	Pig	R	Group 1	✗

Table 6.1 continued

MSUS/CI/78/TSW175	Pig	**R**	Group 1	✖
MALC/BF/80/AB14	Hartebeast	**S**	*T.b.brucei*	✖
MHOM/CI/78/TH1-032	Human	**S**	Group 2	✖
MKOK/BF/80/KK14	Kob	**S**	*T.b.brucei*	✖
MKOK/BF/80/KK18	Kob	**S**	*T.b.brucei*	✖
MKOK/BF/80/KK26	Kob	**S**	*T.b.brucei*	✖
MKOK/BF/80/KK33	Kob	**S**	*T.b.brucei*	✖
MKOK/BF/80/KK7	Kob	**S**	*T.b.brucei*	✖
MSUS/CI/78/TSW065-022	Pig	**S**	*T.b.brucei*	✔
MSUS/CI/78/TSW113	Pig	**S**	*T.b.brucei*	✖
MSUS/CI/78/TSW178	Pig	**S**	*T.b.brucei*	✖
MSUS/CI/78/TSW187	Pig	**S**	*T.b.brucei*	✖
MSUS/CI/78/TSW19	Pig	**S**	*T.b.brucei*	✖
MSUS/CI/78/TSW196	Pig	**S**	*T.b.brucei*	✖
MSUS/CI/78/TSW332	Pig	**S**	*T.b.brucei*	✖
MKOD/BF/80/KD3	Waterbuck	**S**	*T.b.brucei*	✖

6.2 Results

6.2.1 Microsatellite Genotyping

There have been several studies examining the population structure of group 1 *T. b. gambiense* that have revealed foci to be largely clonal (Koffi *et al.*, 2009; Morrison *et al.*, 2008). However, very little is known about the relationship between group 1 *T. b. gambiense* and either local *T. b. brucei* or group 2 *T. b. gambiense*. The relationships between these populations may reveal if the human serum resistance phenotype in *T. b. gambiense* is likely to have evolved once or separately in each group. This study used 6 polymorphic microsatellite markers (PCR conditions and primer sequences are given in **Chapter 2: Materials & methods**) for 43 samples to investigate the relationships between the group 1 and 2 *T. b. gambiense* and *T. b. brucei* populations at a focus present in Côte d'Ivoire in the early 1980s. All 6 markers are polymorphic, featuring several alleles at each locus (**Appendix 5**). If random mating is occurring within the Côte d'Ivoire population then the number of heterozygous genotypes within the population should be a predictable proportion based on the number and frequencies of alleles present in the population. Most of the markers exhibit heterozygosity rates similar to those expected for a randomly mating population however there are markers that do not (**Table 6.2**). While a perfect match between observed and expected heterozygosity at every marker would not be expected in a field population due to the presence of some non-random mating, mutation or natural selection, this result may also suggest there is a degree of population sub-structuring at this focus that may be due to separate populations of *T. b. gambiense* and *T. b. brucei*.

Table 6.2 Polymorphisms and heterozygosity of *T. brucei* at the Côte d'Ivoire Daloa focus.

Locus	Alleles	Observed Heterozygosity	Expected Heterozygosity
Ch5/JS2	7	0.442	0.744
Ch11/110	3	0.279	0.279
Ch11/51	3	0.233	0.210
Ch1/18	8	0.837	0.706
Ch1/D2/7	8	0.372	0.672
Ch2/PLC	5	0.279	0.539

In order to further investigate the possibility of sub-structuring and possible relationships between the populations of different *T. brucei* further, a dendrogram for the focus was created using a 6 locus MLG to estimate the genetic distance between each isolate in the library (**Figure 6.1**). This revealed strong bootstrap support indicating a discrete population containing most of the isolates characterised as group 1 *T. b. gambiense*, the group 1 *T. b. gambiense* strain Eliane and the group 1 *T. b. gambiense* genome reference strain DAL972. The strains identified as *T. b. brucei* and group 2 *T. b. gambiense* form a separate population distinct from these group 1 *T. b. gambiense*. There is little bootstrap support for any of the branching nodes within the *T. b. brucei* and group 2 population, suggesting that they are largely indistinguishable from each other in the 6 locus MLG study. Additionally, six strains identified by BIIT as being group 1 *T. b. gambiense* due to an invariant human serum resistance phenotype do not cluster with the group 1 *T. b. gambiense* population and instead cluster within the *T. b. brucei* and group 2 *T. b. gambiense* population. This may indicate that they are also group 2 *T. b. gambiense*. The outlier Tanzanian *T. b. brucei* strain STIB247 shows no similarity to any of the Côte d'Ivoire isolates, as is to be expected.
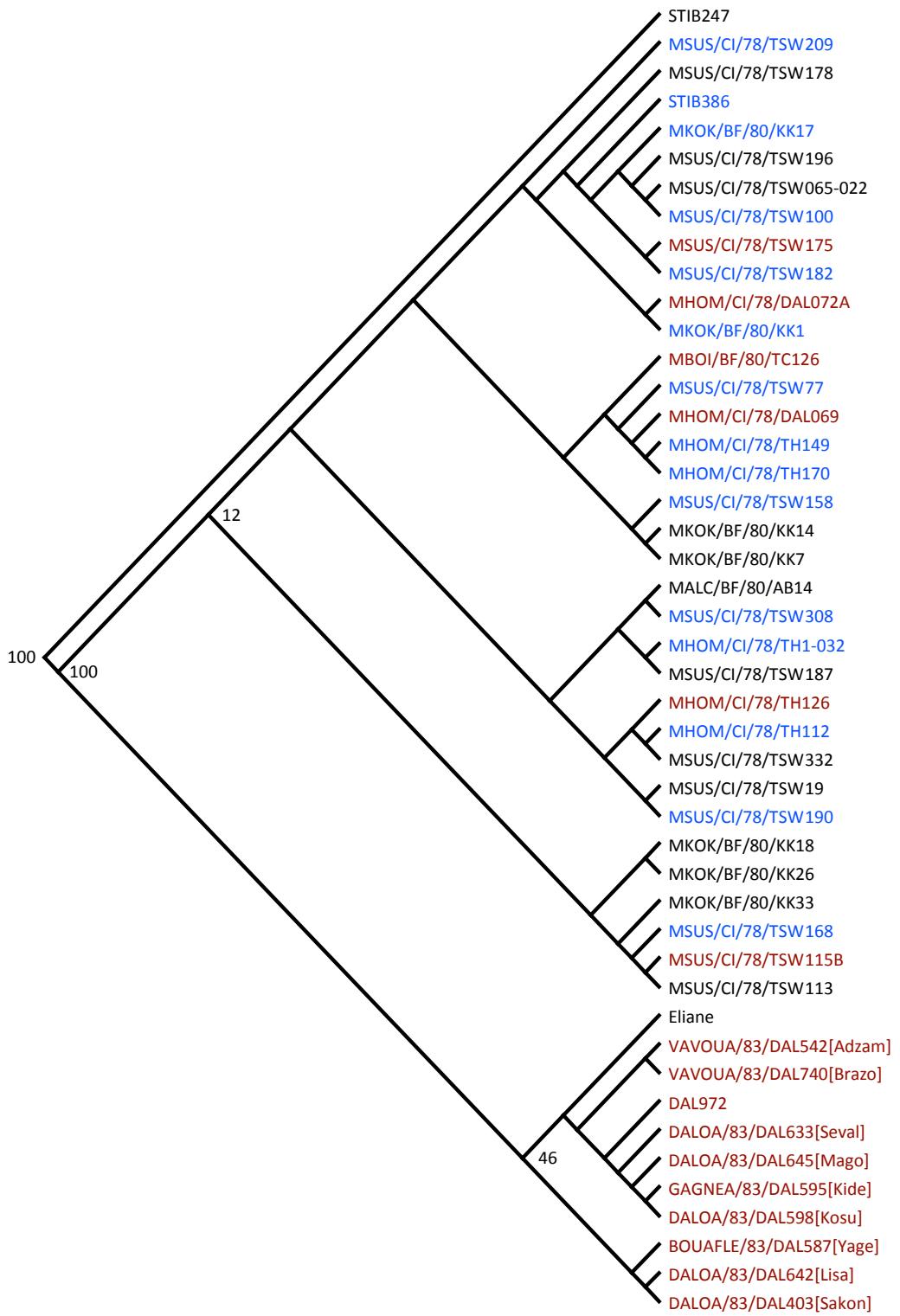
Figure 6.1 Dendrogram of multilocus genotypes (MLG) for the *T. brucei* isolates collected from Daloa, Côte d'Ivoire over the period of time 1978-1983 in addition to STIB247, STIB386, DAL972 and Eliane. Bootstrap values from 100 iterations are indicated for branch nodes with a bootstrap value above 10. Isolates exhibiting invariant human serum resistance by BIIT are indicated in red. Those expressing a variable resistance phenotype are highlighted in blue. Human serum sensitive lines are indicated in black.

6.2.2 *TgsGP*

The *TgsGP* gene is diagnostic of group 1 *T. b. gambiense* and a simple nested PCR based test can reveal the presence or absence of this gene in isolates (Radwanska *et al.*, 2002). While some group 2 *T. b. gambiense* do possess the gene, it is present in all strains of group 1 *T. b. gambiense* (Berberof *et al.*, 2001; Gibson *et al.*, 2010). In order to determine the identity of the isolates characterised as group 1 *T. b. gambiense* in the sample library by BIIT, PCR using primers for *TgsGP* was performed for the entire collection. This analysis revealed that 13 isolates amplified for the presence of *TgsGP* gene (**Table 6.1**). One of these has been characterised in a BIIT as a *T. b. brucei*. Nine of the isolates were previously characterised as group 1 *T. b. gambiense* and also cluster within the discrete group 1 *T. b. gambiense* population with Eliane and DAL972 in the 6 MLG dendrogram (**Figure 6.1**). Three of these isolates with *TgsGP* had previously been identified as group 2 *T. b gambiense* due to a variable human serum resistance phenotype in a five rodent BIIT. Several isolates that possess an invariant resistance phenotype cluster outside of the main group 1 *T. b. gambiense* population in the 6 MLG dendrogram. None of these isolates possess the group 1 diagnostic marker *TgsGP*. This suggests that these are misidentified group 2 *T. b. gambiense*, despite showing robust resistance to human serum. This could be due to the fact they do not possess the characteristic variable phenotype of group 2 *T. b. gambiense* or perhaps the isolates achieved 5 successful rodent infections in the BIIT due to chance.

As previously mentioned, one isolate identified as *T. b. brucei* also possesses *TgsGP*. This could be because even though it is a group 2 *T. b. gambiense,* it had 5 failed rodent infections in the BIIT by chance. Alternatively it could be a *T. b. brucei* strain that has inherited the *TgsGP* gene due to mating with a group 2 *T. b. gambiense*. *TgsGP* has not been shown to be present in *T. b. brucei* to date (Radwanska *et al.*, 2002) but this could be a result of testing *T. b. brucei* from areas that do not possess group 2 *T. b. gambiense* or the general lack of data for the Côte d'Ivoire focus. Similar genes to *TgsGP* that are derived from the same ancestral *VSG* are present in some *T. b. brucei* so it is also possible that cross-amplification with primers for this gene explains the perceived presence of *TgsGP* in this *T. b. brucei* (Gibson *et al.*, 2010). This indicates that the *TgsGP* PCR assay is not an ideal marker to differentiate between group 1 *T. b. gambiense* and other *T. brucei* at the Daloa, Côte d'Ivoire focus.

6.2.3 Population analysis

It appears likely from these data that there are several misidentified group 2 *T. b. gambiense* in this isolate collection and that group 2 *T. b. gambiense* were more prevalent then believed at this focus and time point. In order to perform population analysis and reveal any relationships between the two groups of *T. b. gambiense*, three populations were created based on the new data. Any *T. brucei* strain isolate in the collection that exhibited a variable human serum resistance were added to the group 2 *T. b. gambiense* population. In addition, the invariably resistant *T. b. gambiense* lines that did not possess *TgsGP* and clustered outside of the main group 1 *T. b. gambiense* population were also included in this group 2 population. The group 2 *T. b. gambiense* strain STIB386 used extensively in this thesis was also included to give a total of 21 isolates in this group. The isolates that clustered into the group 1 *T. b. gambiense* population with Eliane and the *T. b. gambiense* reference strain DAL972 were placed into the group 1 *T. b. gambiense* population to give a total of 11 isolates. Finally, all of the human serum sensitive *T. brucei* were placed into a *T. b. brucei* population, of which there were 13. The Tanzanian *T. b. brucei* strain STIB247 was removed from this analysis as it is a genetic outlier.

Table 6.3 Nei's genetic distance between the populations of group 1 and 2 *T. b. gambiense* and *T. b. brucei* at the Côte d'Ivoire Daloa focus.

	T. b. brucei	Gp 2 *T. b. gambiense*	Gp 1 *T. b. gambiense*
T. b. brucei	0.000		
Gp 2 *T. b. gambiense*	0.052	0.000	
Gp 1 *T. b. gambiense*	0.516	0.319	0.000

Table 6.4 F_{st} proportion indicating genetic distance between the populations of group 1 and 2 *T. b. gambiense* and *T. b. brucei* at the Côte d'Ivoire Daloa focus.

	T. b. brucei	Gp 2 *T. b. gambiense*	Gp 1 *T. b. gambiense*
T. b. brucei	0.000		
Gp 2 *T. b. gambiense*	0.028	0.000	
Gp 1 *T. b. gambiense*	0.223	0.144	0.000

After the establishment of the three populations of *T. brucei* present at this focus using the human serum resistance phenotype, dendrogram and *TgsGP* PCR, the relationships between these populations were further analysed using two statistical tests. Nei's genetic distance and F_{ST} both compare the proportion of the total genetic variance contained in a subpopulation relative to the total genetic variance (Nei & Roychoudhury, 1974; Wright, 1978). These statistics indicate how similar populations are to each other, with 1 being completely different and 0 being identical. Both Nei's genetic distance and the F_{st} statistic indicate that the *T. b. brucei* and group 2 *T. b. gambiense* populations are nearly indistinguishable, but both are distinct from the group 1 *T. b. gambiense* (**Tables 6.3** & **6.4**). The group 2 *T. b. gambiense* population is more similar to the group 1 *T. b. gambiense* than the *T. b. brucei*, although the overall relationship is weak and may be a result of small sample number. As they are so alike, the *T. b. brucei* and group 2 *T. b. gambiense* populations were then combined for further analysis to determine if there is evidence for mating occurring amongst strains within these two populations.

Table 6.5 Hardy-Weinberg analysis for the combined population of group 2 *T. b. gambiense* and *T. b. brucei* at the Côte d'Ivoire Daloa focus.

Locus	DF	χ^2	Probability	Significance
Ch5/JS2	15	91.034	0.000	P < 0.001
Ch11/110	3	0.525	0.913	Not Significant
Ch11/51	3	0.072	0.995	Not Significant
Ch1/18	28	40.983	0.054	Not Significant
Ch1/D2/7	15	35.948	0.002	P < 0.01
Ch2/PLC	10	14.209	0.164	Not Significant

Table 6.6 Linkage disequilibrium (statistically significant disequilibrium highlighted in bold) between pair wise polymorphic loci in the combined *T. b. brucei* and group 2 *T. b. gambiense* population at the Côte d'Ivoire focus.

Locus 1	Locus 2	Probability	Standard Error
Ch5/JS2	Ch11/110	0.177	0.013
Ch5/JS2	Ch11/51	0.096	0.009
Ch11/110	Ch11/51	0.438	0.009
Ch5/JS2	Ch1/18	0.011	0.005
Ch11/110	Ch1/18	0.144	0.017
Ch11/51	Ch1/18	0.138	0.013
Ch5/JS2	Ch1/D2/7	0.901	0.012
Ch11/110	Ch1/D2/7	0.268	0.012
Ch11/51	Ch1/D2/7	0.739	0.011
Ch1/18	Ch1/D2/7	0.152	0.023
Ch5/JS2	**Ch2/PLC**	**0.036**	**0.008**
Ch11/110	Ch2/PLC	0.975	0.003
Ch11/51	Ch2/PLC	0.293	0.010
Ch1/18	**Ch2/PLC**	**0.048**	**0.003**
Ch1/D2/7	Ch2/PLC	0.533	0.020

One method to determine the degree of mating in a population is to utilise the principles of Hardy-Weinberg. The Hardy-Weinberg hypothesis is a theoretical construct that states that the allele and genotype frequencies of a population undergoing free and random mating remain constant between generations and are in equilibrium. However, a perfect Hardy-Weinberg equilibrium is unlikely to occur in nature due to factors that disturb the equilibrium, such as non-random mating, mutation, natural selection, limited population size or random genetic drift. Despite this, the degree to which allele and genotype frequencies in a population deviate from Hardy-Weinberg equilibrium can be calculated to give an indication as to whether mating is occurring in a population. Populations closer to Hardy-Weinberg equilibrium are likely to feature some degree of mating. Analysis of the alleles in the combined *T. b. brucei* and group 2 *T. b. gambiense* population reveal that four of the six markers did not significantly deviate from Hardy-Weinberg equilibrium (Table 6.5). The two markers not in equilibrium (JS2 & D2/7) possess a large number of alleles relative to the sample size so a Hardy-Weinberg analysis will be less robust for these markers. These data are consistent with the hypothesis that some degree of mating is occurring between the group 2 *T. b. gambiense* and *T. b. brucei* populations. Hardy-Weinberg analysis of the group 1 *T. b. gambiense* is inappropriate due to the small sample size and markers being monomorphic.

A second method to assess the degree of mating within the combined group 2 *T. b. gambiense* and *T. b. brucei* population is to estimate the amount of linkage disequilibrium in allele and genotype frequencies. The presence of mating in a population will cause alleles to be inherited in an apparently randomly assorted manner and leading to linkage equilibrium between alleles, while a population with little mating or strong selection will have similar alleles inherited together and exhibit linkage disequilibrium. This study has revealed that the majority of alleles in the combined *T. b. brucei/group 2 T. b. gambiense* population show significant linkage equilibrium (Table 6.6). When combined with the Hardy-Weinberg analysis, there would appear to be some degree of re-assortment and mating both between and within the *T. b. brucei* and group 2 *T. b. gambiense* populations. The group 1 *T. b. gambiense* alleles are largely monomorphic or uninformative so a linkage disequilibrium study is not possible on this population. However the low number of alleles and limited variability present observed with studies of other foci suggest that group 1 *T. b. gambiense* possess a clonal population structure that is separate from that of the group 2 *T. b. gambiense* and *T. b. brucei* population (Koffi *et al.*, 2009; Morrison *et al.*, 2008).

Finally, principal component analysis (PCA) was performed on the isolates using two vector genetic distance comparisons (**Figure 6.2**). PCA performs a multi-dimensional comparison using genetic distance between individuals. The factor that causes the greatest genetic distance between individuals in the populations is used for one axis, and the factor that causes the next most distance is the other. In this way the structure of a population can be visualised. This PCA clearly shows that group 1 *T. b. gambiense* are distinct from the *T. b. brucei* and group 2 *T. b. gambiense* populations. These two populations however are closely related.
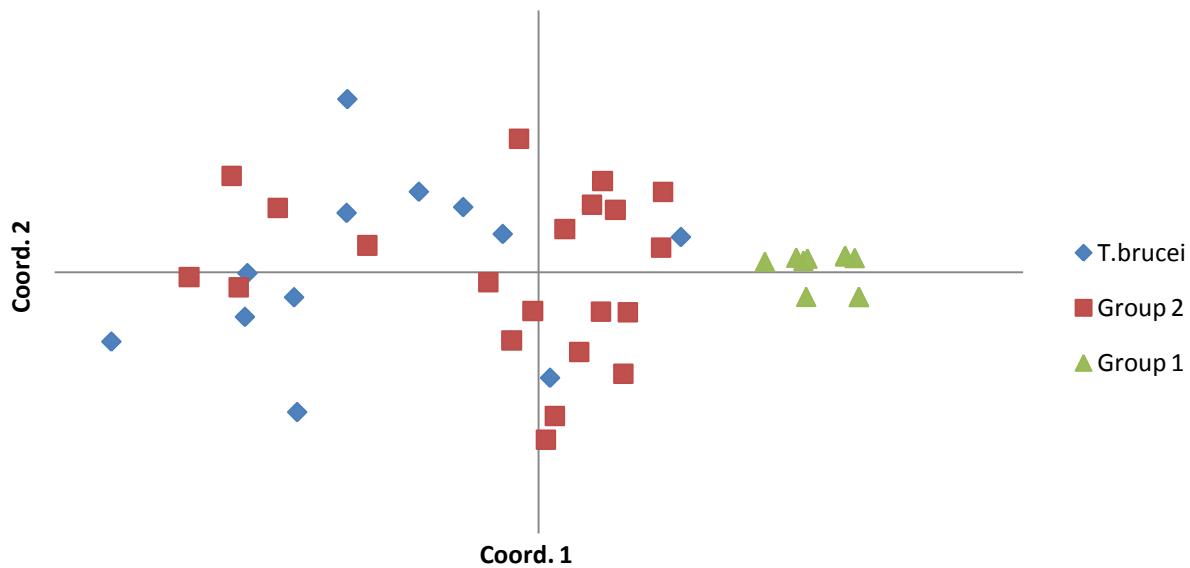
Figure 6.2 Principal component analysis (PCA) of the Côte d'Ivoire *T. brucei* isolates using a pair-wise genetic distance comparison between each isolates MLG. The x-axis explains 45.86% of the variability in the populations and the y-axis 17.96%, for a total of 63.82%. Isolates are coloured by population.

It would appear from these data that the *T. b. brucei* and group 2 *T. b. gambiense* populations exhibit some degree of mating, both within and between sub-species. The group 1 *T. b. gambiense* population is clonal and distinct from the sympatric *T. b. brucei* and group 2 *T. b gambiense* populations.

6.3 Discussion

These data confirm studies from other foci that group 1 *T. b. gambiense* is clonal (Koffi *et al.*, 2009; Morrison *et al.*, 2008). The reasons for the clonality of group 1 *T. b. gambiense* are unclear as the genes necessary for meiosis are present (Jackson *et al.*, 2010). There is an intron present in the meiotic gene SPO11 in the annotated genome of the group 1 *T. b. gambiense* strain DAL972 and it was believed this could disrupt function of the gene. It has since been revealed that the intron is an error in the sequence assembly (C. Duffy, personal communication). While group 1 *T. b. gambiense* possesses all of the necessary meiotic genes, it is unknown as to whether they can be expressed and it may be due to mis-regulation of these genes that the various group 1 *T. b. gambiense* populations have become clonal. Another reason for the low frequency of mating observed in group 1 *T. b. gambiense* is that, due to the limited number of genotypes present, the chances of a mixed genotype infection occurring in the tsetse is unlikely. This precludes any mating and the effect is compounded by a small animal reservoir and low infection prevalence and parasitaemias in humans. The low number of genotypes present could be due to a bottleneck effect that occurred to the population in the past or if group 1 *T. b. gambiense* underwent clonal expansion after it gained the human infectivity trait.   A much larger analysis of group 1 *T. b. gambiense* populations, including the vector and reservoir populations, would need to be performed to resolve whether clonality arose in this sub-species due to changes in biology or as a result of limited available genotypes.

Despite questions as to how the clonal population structure has developed, these data show that the group 1 *T. b. gambiense* genome reference strain DAL972 is clearly a good representative of group 1 strains from this area, although its similarity to isolates from other foci is unknown. Unfortunately it was not possible to compare these data to samples analysed for other group 1 *T. b. gambiense* foci due to the necessity to include new microsatellite markers. However, previous studies have shown that the MLG in the populations of various group 1 *T. b. gambiense* foci are significantly different from each other (Morrison *et al.*, 2008). In addition, examination of alleles in this study with previous studies indicates that there are few shared alleles at each marker. This suggests that this specific Daloa focus possesses a different clonal population of group 1 *T. b. gambiense* to that found in other foci. Interestingly, this study has also revealed that group 2 *T. b. gambiense* is less rare than originally thought in the 1980s. Several group 1 *T. b. gambiense* identified by the BIIT now appear to be group 2. This suggests that the BIIT is not a gold standard for differentiating between group 1 and group 2 *T. b. gambiense*. Likewise, PCR for *TgsGP* and the CATT for LiTaT 1.3 are also unsatisfactory due to either cross reaction or absence in some group 1

T. b. gambiense stocks (Dukes *et al.*, 1992; Radwanska *et al.*, 2002). It may be possible that the polymorphisms shown to be present in the *HpHbR* receptor may serve as a diagnostic test to differentiate group 1 *T. b. gambiense* from *T. b. brucei* and group 2 *T. b. gambiense* (Chapter 3). Further work must be undertaken to characterise the SNPs identified for further isolates from more foci than the three that have been examined to date.

This study has also shown that STIB386 is a good representative for Côte d'Ivoire group 2 *T. b. gambiense* in that it is distinct from the group 1 *T. b. gambiense* population and clusters with other group 2 *T. b. gambiense* strains. Previous studies have suggested that group 2 *T. b. gambiense* are more varied than group 1 and more similar to *T. b. brucei* (Gibson *et al.*, 1980; Godfrey & Kilgour, 1976; Mathieu-daude & Tibayrenc, 1994; Mehlitz *et al.*, 1982; Stevens & Tibayrenc, 1996; Tait *et al.*, 1984; Zillmann *et al.*, 1984). As group 2 *T. b. gambiense* appear highly genetically variable, it would seem prudent that further group 2 *T. b. gambiense* strains must be studied in concert with STIB386. However the data presented here show that not only is there a large and genetically varied breeding population of group 2 *T. b. gambiense*, these individuals are almost certainly breeding with local *T. b. brucei*. Indeed, the *T. b. brucei* and group 2 *T. b. gambiense* populations are largely indistinguishable and differ only in their potential to infect humans. Mating between *T. b. brucei* and group 2 *T. b. gambiense* has also been shown to occur in a laboratory setting (Cooper *et al.*, 2008). Group 2 *T. b. gambiense* are more accurately described as an extended host range variant *T. b. brucei* than a *T. b. gambiense*. Due to this, *T. b. brucei* and group 2 *T. b. gambiense* populations are very closely related and both are distinct from the group 1 *T. b. gambiense* population. Group 2 *T. b. gambiense* is more closely related to group 1 than the *T. b. brucei* are but the association is weak. This may be because group 2 *T. b. gambiense* arose as a very rare mating event between a group 1 *T. b. gambiense* and a *T. b. brucei*, or that group 1 and 2 *T. b. gambiense* arose from a similar progenitor. The association between group 1 and group 2 *T. b. gambiense* may have then been eroded over time due to group 2 *T. b. gambiense* mating with local *T. b. brucei*. Both theories would explain how some group 2 *T. b. gambiense* have come to possess the *TgsGP* gene. However, recent evidence has emerged that while *TgsGP* is a true diagnostic feature of group 1 *T. b. gambiense*, the *TgsGP* gene present in group 2 *T. b. gambiense* is actually a different derivation of the Tb10.v4.0178 *VSG* so cannot be used as evidence for relationships between group 1 and 2 *T. b. gambiense*, despite being extremely similar and amplifying with the same PCR primers (Gibson *et al.*, 2010). With regards to the human serum resistance mechanisms of the two groups, since it is not possible to

establish common heritage between the *T. b. gambiense* groups, it is prudent to assume that two different mechanisms have arisen and the two groups must be studied separately.

The degree to which mating occurs in a population is an important consideration due to it acting as a major source of variation. While the human infectivity trait present in group 2 *T. b. gambiense* may have evolved once, it is now present in a wide range of genetic backgrounds. This is in contrast to group 1 *T. b. gambiense* which has limited genetic diversity at each discrete disease focus and hence a lower capacity to evolve. The genetic backgrounds of the group 2 *T. b. gambiense* infectivity trait may be in constant flux due to evidence suggesting the presence of mating within the population. This increases the possibility that the human infectivity trait will be placed into a background that increases its effectiveness, improving the trait by natural selection. The evidence suggesting mating between group 2 *T. b. gambiense* and *T. b. brucei* also makes it feasible that the human infectivity trait can remain in the local *T. brucei* population undetected, even if the disease is removed from the human population. This is especially likely with the human infectivity trait found in group 2 *T. b. gambiense* as it can be variably expressed. Any evolutionary cost of a human infectivity trait that can be activated when needed is potentially lower than one that is constitutively expressed. This lack of a constitutively expressed mechanism may explain why type 2 *T. b. gambiense* isolates have been found in a wide variety of hosts, including humans, cattle wild deer and rodents, while all of the group 1 *T. b. gambiense* in this study were isolated from human patients. There is some debate as to the animal reservoir for group 1 *T. b. gambiense*, although domestic pigs are believed to be a possible source (Gibson, Mehlitz et al. 1978; Gibson 1986; Paindavoine, Pays et al. 1986). Several other wild fauna have also been shown to carry *T. b. gambiense* (Njiokou *et al.*, 2006) but it is important to consider how these trypanosomes were identified as group 1 – several of the group 1 *T. b. gambiense* identified by the BIIT proved to be actually group 2. Diagnostic tests created using the original BIIT determined groups would actually only differentiate between features shared by group 1 *T. b. gambiense* and misidentified group 2 *T. b. gambiense*.

An important caveat for these presented data is that the samples were collected over 27 years ago. The study has provided an insight into a snapshot of time in which group 1 and 2 *T. b. gambiense* and *T. b. brucei* existed in the same geographic range. The current situation however is largely unknown. While African sleeping sickness is still present (Kaba *et al.*, 2006; Koffi *et al.*, 2009; Koffi *et al.*, 2006), it is impossible to state how the populations of group 1 and group 2 *T. b. gambiense* have changed. Long-term temporal studies have been hampered by the fact that the

political situation in Côte d'Ivoire is in constant turmoil (Kaba *et al.*, 2006). However, the situation may improve with the formation of a peace accord between the Unity government and the FNCI rebel faction. The opportunity to sample group 1 and 2 *T. b. gambiense* with *T. b. brucei* may again present itself, allowing temporal comparisons providing important information on how group 1 and 2 *T. b. gambiense* foci develop over time.

The study presented here, in conjunction with other published data, suggests that human infectivity has arisen on at least four occasions. In Eastern Africa the evolution of *SRA* and the high levels of mating occurring there have allowed the human infectivity trait to spread through the *T. brucei* population, generating the *T. b. rhodesiense* sub-species (MacLeod *et al.*, 2000; Tait *et al.*, 1985). However, there are also human infective trypanosomes in this area that do not possess *SRA,* suggesting a second novel mechanism (De Greef *et al.*, 1989; Enyaru *et al.*, 2006). In Western Africa, group 1 *T. b. gambiense* is the dominant form of human African sleeping sickness. This group of parasites has evolved a constitutively expressed resistance mechanism that does not depend on *SRA.* After evolution of this trait, the sub-species appears to have expanded clonally to take advantage of the new niches offered by being human serum resistant. Just as in Eastern Africa, there is also a human infective *T. brucei* population separate from the dominant form that appears to have a distinct resistance mechanism. This group 2 *T. b. gambiense* population appears to be a host variant *T. b. brucei* and the human infectivity trait it possesses is freely transmitted around the local *T. b. brucei* population. The presence of mating amongst group 2 *T. b. gambiense* and *T. b. brucei* strains makes evolution of new traits more likely and may allow the human infectivity trait to be combined with new phenotypes that increase its virulence. The possibility of mating between group 1 *T. b. gambiense* and *T. b. rhodesiense* appears unlikely due to the formers clonal population structure although the population structures of Ugandan group 1 *T. b. gambiense* have not been investigated as efforts thus far have concentrated on foci in the West of the continent (Koffi *et al.*, 2006; Morrison *et al.*, 2008). Understanding whether mating occurs in the *T. b. gambiense* focus of Uganda is important due to the fact that the *T. b. gambiense* and *T. b. rhodesiense* foci are moving closer together (Fevre *et al.*, 2005; Picozzi *et al.*, 2005). In addition, *T. b. rhodesiense* mating with group 2 *T. b. gambiense* is certainly possible if the two foci meet and may generate a novel human infective sub-species. Taken together, all of these data indicate that *T. brucei* has a high zoonotic potential despite specific trypanolytic countermeasures that have been inherited by humans.

Chapter 7: Final Discussion

Very little is known about the properties or mechanisms of how the dominant human infective sub-species, *T. b. gambiense*, overcome the innate lytic factors present in humans. Matters are complicated in that there are two distinct groups of *T. b. gambiense*; group 1 *T. b. gambiense* are the more prevalent form and are invariably human serum resistant while the rarer group 2 *T. b. gambiense* exhibit a variably expressed resistance phenotype (Mehlitz *et al.*, 1982; Zillmann *et al.*, 1984). The resistance mechanism of *T. b. rhodesiense* regulated by the expression of the serum resistance associated (*SRA*) gene, is not present in either of the two groups of *T. b. gambiense* so they must possess one or more different resistance strategies (Xong *et al.*, 1998). At the start of this thesis, it was unknown as to whether the resistance mechanisms used by the two *T. b. gambiense* groups used avoidance, neutralisation (like *SRA*) or compensatory mechanisms. Indeed, it was unclear as to whether the two groups of *T. b. gambiense* use the same or different resistance mechanisms as the relationships between group 1 and 2 *T. b. gambiense* are poorly understood. It has been shown that group 1 *T. b. gambiense* populations are clonal and distinct from locally mating *T. b. brucei* populations (Koffi *et al.*, 2009; Mathieu-daude & Tibayrenc, 1994; Morrison *et al.*, 2008; Stevens & Tibayrenc, 1996). Group 2 *T. b. gambiense* show isoenzyme and AFLP profiles that are more varied and similar to *T. b. brucei*, suggesting they are undergoing some degree of mating (Gibson *et al.*, 1980; Godfrey & Kilgour, 1976; Mathieu-daude & Tibayrenc, 1994; Mehlitz *et al.*, 1982; Stevens & Tibayrenc, 1996; Tait *et al.*, 1984; Zillmann *et al.*, 1984). This suggests that the two groups of *T. b. gambiense* are distinct. This thesis has elucidated several features concerning the human serum resistance phenotype in *T. b. gambiense*; including the interactions with the major lytic HDL particle TLF-1, the ability to resist lysis by ApoL1, the identification of a genetic locus associated with the resistance trait of group 2 *T. b. gambiense* and the relationship between the two *T. b. gambiense* populations.

The group 2 *T. b. gambiense* investigated during this thesis exhibited a variably expressed resistance phenotype as expected (Chapter 3). Both isogenic resistant and sensitive *T. b. gambiense* STIB386 expressed the TLF-1 receptor HpHbR to a similar level (Chapter 3). Therefore the variability of the phenotype was not due to differential uptake of TLF-1. While there was weak evidence that the resistant form of STIB386 uptakes TLF-1 more slowly than the sensitive form, this was only from one biological replicate due to poor experimental design. In hindsight, a flow cytometry based analysis or multiple screenings from several smaller experiments would have

proved more effective in determining differences in rate of TLF-1 uptake then the single large experiment that was conducted in this thesis. In addition to TLF-1, both sensitive and resistant STIB386 internalised ApoL1 and trafficked it to the lysosome indicating that the sub-species group is variably resistant to ApoL1 and suggests that the resistance mechanism is most likely to lie within the late endosomes or lysosome (Chapter 4). While superficially resembling the variable resistance phenotype displayed by *T. b. rhodesiense*, the mechanism does not involve *SRA* (Chapter 3). Previous research has indicated that the resistance phenotype for the group 2 *T. b. gambiense* strain TxTat correlated with a switch in the variable antigen type (VAT) of the parasite (Ortiz *et al.*, 1994). This strongly suggested an expression site related mechanism that is similar to *SRA*. However, doubts have been raised as to whether the TxTat line used in this study is a group 2 *T. b. gambiense* or a *T. b. rhodesiense* (Grab & Kennedy, 2008) and sequencing of the ESAG6/7 hypervariable regions in the isogenic resistant and sensitive group 2 *T. b. gambiense* strain STIB386 did not reveal any changes in expression site switching that would point towards an ESAG/VSG associated phenotype (Chapter 3). However, this interpretation is based on the assumption that group 2 *T. b. gambiense* has a similar ESAG repertoire to that of *T. b. brucei*. This has not yet been demonstrated and it is instead possible that group 2 *T. b. gambiense* possess a restricted ESAG repertoire similar to group 1 *T. b. gambiense* which would make distinguishing between expression sites using the ESAG sequence impossible (Young *et al.*, 2008). Several different lines of evidence indicate that this is not likely; including the fact that the population structure of group 2 *T. b. gambiense* suggests that they are genetically indistinguishable from *T. b. brucei* and so are likely to possess a similar varied ESAG repertoire (Chapter 6). In addition, it has been shown that group 2 *T. b. gambiense* isolates possess more varied metacyclic ESAGs, including alleles not found in group 1 *T. b. gambiense* (Bringaud *et al.*, 2001). Finally, the *ESAG6/7* genes expressed by the resistant and sensitive group 2 *T. b. gambiense* are identical to *T. b. brucei* alleles not found in group 1 *T. b. gambiense* (Chapter 3 & (Young *et al.*, 2008)). Another assumption that was made in order to interpret the sequencing of the *ESAG6/7* genes in STIB386 is that the large ESAG repertoire that *T. b. brucei* possesses is a feature common across the sub-species and is likely to be shared with group 2 *T. b. gambiense*. So far four *T. b. brucei* lines have been investigated, and each shows a large ESAG repertoire suggesting that this is a universal feature (Hertz-Fowler *et al.*, 2008; Young *et al.*, 2008). Despite the evidence suggesting group 2 *T. b. gambiense* possesses a large ESAG repertoire, the exact variability of ESAGs in the group can only be elucidated by an intensive sequencing study of genomic ESAGs and this work is currently ongoing.

Although it appears likely that the group 2 *T. b. gambiense* resistance mechanism is unrelated to the use of particular expression sites, it still exhibits a variable phenotype (Chapter 3). Many eukaryote systems exhibit gene expression stochasticity and while this can be detrimental and is usually minimised, it is also useful in allowing organisms to adapt to an unstable and variable environment (Kaern *et al.*, 2005). This is especially applicable to bloodstream form group 2 *T. b. gambiense* in that it can find itself in a large variety of animal hosts with serum of different properties. The human host in particular offers the additional challenge of ApoL1 mediated lysis (Chapter 6). Most gene expression stochasticity emerges as a result of complex interactions between gene networks (Raser & O'Shea, 2004). Positive and negative feedback of transcription factors and other regulative genes, usually in concert with external promoters, can cause un-predictable changes to whole gene cascades and create a seemingly random expression of genes (Raser & O'Shea, 2004; Volfson *et al.*, 2006). The phenomenon of gene expression stochasticity in eukaryotic systems has largely been limited to theory formation and modelling so very little is known about the systems involved. Recent research with *Saccharomyces cerevisiae* successfully used the inherent "noise" present in the expression of transcription factors to model a large 2000 gene expression network and predict the stochastic effects (Li *et al.*, 2010). They also showed that the gene expression "noise" is a heritable trait and hence the stochastic pattern is also heritable. Nothing is known about gene expression stochasticity in trypanosomes as variable expression is intuitively assumed to always involve a change in expression sites. The fact that gene expression in trypanosomes is largely regulated by RNA stability conferred by the 3' UTR provides a large amount of expression "noise" to fuel stochastic expression due to random degradation of the RNA transcripts (Clayton & Shapira, 2007). This makes it possible that stochastic networks similar to those found in other eukaryotes could exist in *T. brucei*. Unravelling the complex network of transcription factor interactions that cause the random expression of the human serum resistance phenotype in group 2 *T. b. gambiense* poses a daunting task, but information gleaned from QTL analysis and comparative transcriptomics will offer an insight into the genes that are involved in the putative stochastic network (Chapter 5).

Interestingly procyclic form *T. brucei* are able to resist lysis by human serum, indicating that all *T. brucei* sub-species have the ability to vary transcription of genes that are already present and become human infective (Moore *et al.*, 1995). Originally this was believed to be because this life cycle stage of the parasite exhibits a 10-fold lower rate of endocytosis relative to bloodstream form (Morgan *et al.*, 2001). However, a procyclic *T. b. brucei* strain that was genetically modified to show comparable rates of endocytosis to bloodstream form remained resistant to killing by

human serum (Natesan *et al.*, 2010). This thesis has indicated that this is due the ability of procyclic form *T. brucei* to resist the effects of the trypanolytic protein ApoL1, even at extremely high concentrations of the protein (Chapter 4). While it was not possible to show that ApoL1 was internalised by procyclics, the high concentrations of ApoL1 used to overcome the lower endocytosis rate did not result in any reduction in the growth rate of the parasites. The resistance mechanism in procyclic forms is unknown but it is interesting that one of the most heavily down-regulated genes in procyclic forms relative to bloodstream forms is the lysosomal LAMP-like protein *p67* (Veitch *et al.*, 2010). When this gene is knocked-down in bloodstream form *T. b. brucei* they become human serum resistant (Peck *et al.*, 2008). The modified cells also die after 24 hours, indicating that *p67* is an essential gene to bloodstream forms and may indicate why *T. b. brucei* has not evolved to down-regulate *p67* expression in the field to be able to infect humans (Peck *et al.*, 2008). How the p67 protein invokes this human serum resistance mechanism is unknown. One potential hypothesis is that it affects lysosomal pH which would in turn stop ApoL1 achieving its pore-forming confirmation. However, knocking-down expression of *p67* had no affect on lysosomal pH (Peck *et al.*, 2008). There are size changes and other morphological features in the lysosome in *p67* knock-down *T. b. brucei*, although how these contribute to human serum resistance is again unknown. It is also possible that human serum resistance in procyclic forms is unrelated to *p67*. Further investigation into how this interesting phenotype arises may reveal a new control method against the parasites when they are in the tsetse host. Transcriptome studies comparing isogenic bloodstream form resistant and sensitive *T. b. gambiense* strain STIB386 have indicated *p67* is not significantly differentially expressed between these two lines (N. Veitch, unpublished data), so it is unlikely that this is the variable component involved in human serum resistance in group 2 *T. b. gambiense*.

Whatever the variably expressed element involved in the phenotype in group 2 *T. b. gambiense* is, there is a strong genetic component underlying the resistance mechanism revealed in the heritability of the trait (Chapter 5). Quantitative trait loci (QTL) mapping has revealed that a significant proportion of the heritable trait is located in a 30 gene region of chromosome 8 (Chapter 5). At present this QTL region extends into the telomeric region although efforts are underway to fine-map the extent of the region. The QTL may contain a key component of the trait that controls or is controlled by the variable component. Alternatively it may be a factor that needs to be present in order for the resistance trait to function. Reverse genetics on the candidate genes will allow the determination as to which theory is correct. If the QTL is found to extend into the telomeric region the candidate gene could include an ESAG or VSG related gene,

although it does not appear as if expression site switching is relevant to human serum resistance in STIB386 (Chapter 3). Fortunately, parallel to the work identifying this QTL, a comparative transcriptome comparison experiment was investigating the transcript profiles of the resistant and sensitive isogenic group 2 *T. b. gambiense* strain STIB386 allowing further refinement of the candidate genes. This transcriptome study has revealed over 500 genes differently expressed between these lines (N. Veitch, unpublished data). This is indicative of a complex multi-gene network that can cause stochastic gene expression in eukaryotes (Kaern *et al.*, 2005). However, one of the most highly differentially expressed genes, *Calpain*, lies within the chromosome 8 QTL and is now being investigated further. Trypanosome Calpain is a lysosomal protease with an unknown substrate (Liu *et al.*, 2010). It is possible that STIB386 has variations in this gene that allow it to either digest the lytic ApoL1 protein, or perhaps has functions that compensate for the pores introduced by ApoL1. This work is extremely preliminary but future options include sequencing the gene to identify polymorphisms, the utilisation of Calpain inhibitors to investigate the phenotype and also ectopically expressing the group 2 *T. b. gambiense* alleles in a *T. b. brucei* strain to investigate any impact on human serum resistance. Several other QTL of lower significance than the chromosome 8 QTL have also been highlighted (Chapter 5). These may indicate genes that are involved in the stochastic expression network and the addition of extra progeny to the QTL analysis will likely improve the power of the study and the significance of these suggestive QTLs.

The group 1 *T. b. gambiense* strain, Eliane, investigated in this thesis invariably expressed the human serum resistance phenotype, confirming previous *in vivo* research. This thesis has shown that a trait common to isolates from three different group 1 *T. b. gambiense* foci is the down-regulation and loss of function in the high-affinity receptor for TLF-1, HpHbR (Chapter 3). The group 2 *T. b. gambiense*, *T. b. brucei* and *T. b. rhodesiense* strains showed normal expression of the HpHbR receptor. Down-regulation of *HpHbR* allowed group 1 *T. b. gambiense* to avoid the major lytic particle in humans, TLF-1. While this may suggest that group 1 *T. b. gambiense* utilise a strategy to avoid lysis by ApoL1, the group 1 *T. b. gambiense* strain also displayed the ability to resist lysis by recombinant ApoL1 despite internalising and trafficking the protein to the lysosome (Chapter 4). A potential explanation for this result is that TLF-2 is likely taken into the parasite cell due to weak interactions with the cell surface VSG and the IgM normally associated with the particle (Vanhollebeke & Pays, 2010). It is difficult to envisage how group 1 *T. b. gambiense* would be able to avoid the TLF-2 particle in a manner similar to how it avoids TLF-1 so resistance

to ApoL1 is expected. Despite this, why group 1 *T. b. gambiense* strains have evolved to avoid TLF-1 while also being resistant to the lytic component of that particle is unclear.

Several hypotheses can be put forward to explain why this may be the case. One is that TLF-1 evolved in primates before TLF-2 so down-regulating *HpHbR* was a relatively straightforward way to achieve primate infectivity. However the distribution of TLF-1 and TLF-2 are unknown across the primate species so a comprehensive zoological study would need to be undertaken to determine which primate species possess each of the two particles. An alternate hypothesis is that the group 1 *T. b. gambiense* lysosome or the resistance mechanism is more sensitive to the un-cleaved and toxic Hpr Signal Peptide (Harrington *et al.*, 2010). Research into this killing mechanism is still preliminary but when trypanosomes are exposed to the Hpr signal peptide it integrates into the cell surface membrane to elicit the killing effect (Harrington *et al.*, 2010). Group 1 *T. b. gambiense* are no more sensitive to killing by cell surface peptide integration then other *T. brucei* sub-species. However, future studies investigating the effects of the Hpr Signal Peptide when it is internalised may reveal if group 1 *T. b. gambiense* is more susceptible if the peptide is in the lysosome. A final hypothesis as to why group *1 T. b. gambiense* has reduced the uptake of TLF-1 is that the resistance mechanism is not highly efficient and so must limit the amount of incoming ApoL1 to a level at which the parasite can either compensate for or neutralise the lytic protein. Exposing group 1 *T. b. gambiense* to much higher concentrations of lytic protein will make it possible gauge if there is a maximal dose that the parasite can cope with, and therefore allow testing of this hypothesis. An alternative test would be to conjugate ApoL1 to a nanobody that is easily internalised by *T. brucei* and investigate whether *T. b. gambiense* parasites taking up ApoL1 at an increased rate are more sensitive to lysis. This has been experimentally investigated with *T. b. brucei* and *T. b. rhodesiense*, although the authors do not present evidence for the effects of the nanobody conjugate on *T. b. gambiense* (Baral *et al.*, 2006). An important piece of information that emerges from the fact that group 1 *T. b. gambiense* has evolved to reduce uptake of TLF-1 is that this thesis has shown that TLF-2 is probably the more important TLF particle in elucidating the resistance mechanism of this sub-species. Most studies have previously concentrated on TLF-1 due to the fact the TLF-2 particle is extremely difficult to work with. Protocols are now being established to isolate TLF-2 and perform similar assays to those already undertaken with TLF-1.

This thesis has also shown that methods to distinguish between group 1 and group 2 *T. b. gambiense* have proved unreliable under certain conditions or in particular geographic locations.

The BIIT falsely identified some group 2 *T. b. gambiense* as group 1 (Chapter 6). Genes similar to the group 1 *T. b. gambiense* specific gene *TgsGP* are also found in some group 2 and *T. b. brucei* strains, leading to cross-amplification with *TgsGP* primers (Gibson *et al.*, 2010; Radwanska *et al.*, 2002). Finally the LiTat 1.3 based CATT is hampered due to low sensitivity and the fact that in some foci of group 1 *T. b. gambiense* LiTat 1.3 is not expressed (Dukes *et al.*, 1992). This thesis has indicated a polymorphism in the *HpHbR* gene that is unique to group 1 *T. b. gambiense* that may be useful as a diagnostic marker (Chapter 3). Three separate foci have been investigated in Côte d'Ivoire, the Democratic Republic of Congo and Cameroon but further foci will need to be added to truly define this as a diagnostic test. In addition, it may prove possible that down-regulation of the gene rather than sequence variation is the true diagnostic test. More detailed investigation into the down-regulation of the gene and the SNPs that confer this effect via the 3' UTR is needed to confirm that this is actually a common feature to the sub-species. Further elucidation of the human serum resistance mechanisms in group 1 and 2 *T. b. gambiense* will allow the refinement of diagnostic tests to aid in disease control, just as the identification of *SRA* now aids in the management of *T. b. rhodesiense* (Welburn *et al.*, 2001). The existence of *SRA*-negative *T. b. rhodesiense* and the inadequacy of the BIIT, CATT and *TgsGP* PCR tests will create issues when determining treatment options and management of livestock in areas containing overlapping sub-species. This is especially apparent in Uganda where the *T. b. gambiense* and *T. b. rhodesiense* foci are converging (Fevre *et al.*, 2005; Picozzi *et al.*, 2005).

Confirming research from other group 1 *T. b. gambiense* disease foci, the Côte d'Ivoire population has proven to be clonal (Chapter 6). The alleles present are different from other foci, indicating that they are distinct from other group 1 *T. b. gambiense*. Differences between group 1 *T. b. gambiense* foci have been shown previously and probably arose due to random mutation in the clonal population rather than sexual recombination (Morrison *et al.*, 2008). The reasons for the clonal nature of group 1 *T. b. gambiense* are unknown. The organism possesses all the necessary genes for meiosis (Jackson *et al.*, 2010) but it is possible that an inability to express one or more of these genes has led to the clonality of the group 1 *T. b. gambiense* populations. While *T. b. brucei* and group 2 *T. b. gambiense* have been shown to be able to undergo mating in a laboratory setting, this has not been possible in group 1 *T. b. gambiense*, suggesting there is a biological barrier to mating (Cooper *et al.*, 2008; Gibson *et al.*, 1995; Jenni *et al.*, 1986; MacLeod *et al.*, 2005b). An alternate hypothesis for the apparent lack of mating is that the reduced number of alleles present in the populations results in a low frequency of mixed genotype infections occurring in the tsetse vector. This low number of alleles may be due to a genetic bottleneck that

occurred in the past or the clonal expansion of a successful genotype after it acquired the human infectivity trait similar to some *T. b. rhodesiense* foci (MacLeod *et al.*, 2000). The low genetic variability exhibited by group 1 *T. b. gambiense* and the small animal reservoir should in theory make it more vulnerable to targeted treatment. While group 1 *T. b. gambiense* treatment has largely been hampered in the past due to political instability (Kaba *et al.*, 2006), a targeted drug treatment program would have a devastating effect on the parasite due to its reduced inability to evolve counter-measures.

Evidence in this thesis suggests group 2 *T. b. gambiense* parasites can undergo mating, both within the population and with local *T. b. brucei* (**Chapter 6**). In fact, the two populations are virtually indistinguishable genetically and group 2 *T. b. gambiense* is more accurately described as an extended host range variant of *T. b. brucei* than a *T. b. gambiense* (**Chapter 6**). The group 2 human infectivity trait is likely to have evolved in a *T. b. brucei* strain but, unlike group 1 *T. b. gambiense*, there was no detectable clonal expansion and the trait has remained in the *T. b. brucei* population at a low level. This may be due to the fact group *1 T. b. gambiense* were already present in the primate population so group 2 *T. b. gambiense* was not able to exploit the niche fully. Alternatively, group 2 *T. b. gambiense* could have evolved first but due to co-evolution by primates and competition from group 1 *T. b. gambiense* it has now decreased in number. Despite this, group *2 T. b. gambiense* possesses a considerable potential to evolve rapidly to overcome drug challenges, unlike group 1. The problem is compounded in that group 2 *T. b. gambiense* use an extensive animal reservoir of both domestic and wild animals (**Chapter 6**). The ability for group 2 *T. b. gambiense* to undergo mating also raises the potential of a mating event with *T. b. rhodesiense*, either directly or via a *T. b. brucei* intermediate. This would create a new sub-species with multiple human serum resistance mechanisms but, more importantly, increased tolerances to the drugs used in the treatment of the disease (Barrett *et al.*, 2003). Resistance to drug treatment options is already reported in some disease foci (de Koning, 2001; Maser *et al.*, 2003; Matovu *et al.*, 2001) and recombination and mating between sub-species may contribute to this issue. However, drug tolerances are only described for group 1 *T. b. gambiense* and *T. b. rhodesiense* so further investigation must be done with group 2 *T. b. gambiense* to determine the drug tolerances of this sub-species group.

It is hoped that exploiting the natural defence mechanisms that have evolved to combat trypanosomes infection will limit the opportunities for the parasite to evolve and overcome treatment options. For example, the increasing understanding of ApoL1 has allowed the

formation of several possible treatments for African trypanomiasis. One such proposed solution to African sleeping sickness involves the distribution of transgenic cows expressing either recombinant baboon ApoL1 (Thomson *et al.*, 2009) or truncated ApoL1 (Lecordier *et al.*, 2009). Both baboon ApoL1 and the truncated ApoL1 have key mutations at the *SRA* binding site that allow them to kill *T. b. rhodesiense* and mice expressing these genes ectopically are unable to be infected by either *T. b. brucei* or *T. b. rhodesiense* (Thomson *et al.*, 2009). Unfortunately this will have a limited effect on the human disease across Africa due to the fact that both the baboon and truncated ApoL1 used in these experiments are not able to kill *T. b. gambiense* (Lecordier *et al.*, 2009; Thomson *et al.*, 2009). *T. b. rhodesiense* causes less than 10% of human African sleeping sickness cases across the continent (WHO, 2006). Application of these transgenic cows is also likely to open new niches for non-*SRA T. b. rhodesiense* and *T. b. gambiense* due to the removal of not only *T. b. brucei* and *T. b. rhodesiense* from the cattle herd, but the more important animal trypanosomes *T. congolense* and *T. vivax*. This will need to be considered as the *T. b. gambiense* and *T. b. rhodesiense* foci in Uganda converge (Fevre *et al.*, 2005; Picozzi *et al.*, 2005). The removal of so many competitors to *T. b. gambiense* as it expands south may result in cattle herds predominantly infected not with human serum sensitive species, but with *T. b. gambiense*.

Until a variant of ApoL1 can be found that kills group 1 *T. b. gambiense* it would be prudent to advise caution in the use of transgenic cows that would greatly upset the natural populations of parasitic species. There is some tenuous evidence that there are species of baboon that possess immunity to *T. b. gambiense*, although it is unclear if this is due to ApoL1. *In vivo* studies comparing group 1 *T. b. gambiense* infections in two different baboon species showed that the parasite was able to establish an infection in *Papio hamadryas* but not *Papio papio* (Kageruka *et al.*, 1991). These experiments used a very small number of test subjects however so must be considered with caution. As previously mentioned, primate TLF-1 is unlikely to be internalised by group 1 *T. b. gambiense* due to down-regulation of expression of HpHbR (Chapter 3), so efforts to find naturally occurring variants that combat group 1 *T. b. gambiense* must concentrate on other TLF particles.

Interestingly, a mutant form of ApoL1 has recently been described in some individuals that are members of the African American ethnic group (Genovese *et al.*, 2010). This study showed that the mutant form of ApoL1 is able to kill *T. b. rhodesiense* due to mutations in the SRA-binding domain of ApoL1. Patients who were homozygous for the mutant allele exhibited acute kidney problems during mid-life but heterozygotes were asymptomatic (Genovese *et al.*, 2010). The

authors suggest that the presence of this mutant allele has been selected for in African individuals and may explain the observation that the Apol1 gene is under intense selection in Africa (Conrad *et al.*, 2010). There are several issues with this conclusion, including the fact that most African American individuals originate from Western Africa where *T. b. rhodesiense* is not present. It also seems incredulous that the sub-species of trypanosome that causes less than 10% of human disease could exert such strong evolutionary pressure that it caused the emergence of an ApoL1 allele that is deadly if homozygous. However, it is possible that the ApoL1 variant found in some humans of West African descent may also kill group 2 *T. b. gambiense*. If group 2 *T. b. gambiense* was much more prevalent in the past in Western Africa this would explain the strong selection on ApoL1 and the presence of the ApoL1 variant allele. The removal of group 2 *T. b. gambiense* due to the emergence of a new ApoL1 variant has allowed the clonal expansion of group 1 *T. b. gambiense* featuring a different resistance mechanism. An alternative hypothesis is that *T. b. rhodesiense* was much more prevalent in the past and the evolutionary pressure it exerted led to the emergence of the ApoL1 variant. The variant is predominantly found in people of Western African descent who belong to the Bantu ethnic group. This group originated in the West of Africa and spread eastwards over several millennia before stopping in what is now Southern Uganda. As the Bantu people spread eastward it would have removed *T. b. rhodeseiense* from majority of the tsetse belt. Present-day *T. b. rhodesiense* is now confined to areas populated by people of Nilotic descent (**Chapter 1**). This ethnic group migrated from the North of Africa relatively recently so are unlikely to possess the ApoL1 variant that can kill *T. b. rhodesiense*. Unfortunately this is all highly speculative and the utilisation of molecular clock techniques to determine when group 1, 2 *T. b .gambiense* or *T. b. rhodesiense* emerged from a *T. b. brucei* ancestor relative to each other is complicated by the fact that the *T. b. brucei* and group 2 *T. b. gambiense* populations are undergoing some degree of mating (**Chapter 6**). However it does introduce hope that removal of the clonal and drug vulnerable group 1 *T. b. gambiense* population will not allow the more genetically variable group 2 *T. b. gambiense* to replace it.

This thesis has showed that both group 1 and group 2 *T. b. gambiense* are able to resist lysis by the active component of human TLF, ApoL1. Group 1 *T. b. gambiense* down-regulate the HpHbR receptor to avoid the TLF-1 particle so investigations into the human serum resistance phenotype in this sub-species must concentrate on TLF-2 interactions. Group 2 *T. b. gambiense* are variably resistant to ApoL1, but there is a potentially significant genetic component involved in the phenotype located on chromosome 8. Population genetics suggests that the resistance mechanisms in *T. b. gambiense* have likely arisen on two occasions in the same geographic area. In addition, human infectivity has evolved on at least two further occasions in Western Africa, one involving the *SRA* gene (Xong *et al.*, 1998) and a second that does not (De Greef *et al.*, 1989; Enyaru *et al.*, 2006). This indicates the high potential that *T. brucei* possesses to jump species barriers and infect humans. While the divergence between *T. brucei* and *T. cruzi* occurred approximately 100 million years ago (Stevens *et al.*, 1999), the time period in which the human infective sub-species *T. b. gambiense* and *T. b. rhodesiense* emerged is unknown. Fascinatingly, several major diseases emerged at a similar point of time that correlates with the rise of Neolithic farming in Africa. For the first time in history, large static populations of humans and livestock existed together allowing increasing rates of not only zoonotic transfer, but also transfer amongst large human populations. These likely include the major diseases *Plasmodium falciparum*, measles and small pox (Brussow, 2009; Krief *et al.*, 2010; Volkman *et al.*, 2001). Indeed, the vast majority of pandemics that humans face today are caused by zoonotic transfer (Pike *et al.*, 2010).

Whilst *T. brucei* has existed for at least 100 million years (Stevens *et al.*, 1999), ApoL1 likely evolved in primates at least 25 million years ago as it was present before the divergence of the *Cercopithecidae* (baboon and mandrills) and *Hominidae* families (Monajemi *et al.*, 2002; Poelvoorde *et al.*, 2004; Seed *et al.*, 1990). ApoL1 has several possible roles in apoptosis (Vanhollebeke & Pays, 2006) and some general anti-protozoan activity (Imrie *et al.*, 2004; Samanovic *et al.*, 2009) but it is the ability to kill trypanosomes that is its defining feature. The emergence of ApoL1 would have had an immediate effect on the *T. brucei* population and offered primates with the gene a significant advantage. ApoL1 is under purifying selection in African human populations (Conrad *et al.*, 2010) and is present in many ground-dwelling primates exposed to high tsetse challenge (Lugli *et al.*, 2004; Poelvoorde *et al.*, 2004). The gene is either missing or exhibits loss of function in primates that exhibit a more arboreal lifestyle, such as the chimpanzee (Poelvoorde *et al.*, 2004). This suggests a long association between ApoL1 and primates exposed to trypanosomes. Co-evolution of the parasite with the host however has led to the evolution of *T. brucei* strains that can overcome the innate lytic properties of the serum of

these primates. The ApoL1 resistance mechanisms present in these *T. brucei* could have evolved at any point in the 25 million year time period after the emergence of ApoL1. Interestingly, both the *Cercopithecidae* and *Hominidae* families appear to have evolved convergent countermeasures to overcome the resistance mechanism of *T. b. rhodesiense,* highlighting the constant arms race between parasite and host (Genovese *et al.*, 2010; Thomson *et al.*, 2009). *T. brucei* has demonstrated a high capacity to adapt and evolve so it is unlikely that this arms race is over. Key to combating the emergence of new diseases and pandemics is the understanding of the mechanisms by which organisms are able to cross the species barrier. This allows the identification of diagnostic markers that can be used in disease detection and also opens the possibility of unique treatment options that emerge from the biology involved. This thesis has revealed features of the human serum resistant phenotype in both group 1 and 2 *T. b. gambiense* that will open new avenues for further research and possible diagnostic and treatment options in the future.

Appendices

Appendix 1 The sequence of the expressed *ESAG6* & *ESAG7* variable regions for both the isogenic stably sensitive and resistant forms of the group 2 *T. b. gambiense* strain STIB386. The hypervariable region of each gene is highlighted in red.

```
386R_ESAG6(BES126)  CATTCCAGCAGGAGTTGGAGGAAATGAGGAACGCATCCGCGTTAGCAGCAGCTGCAGCTG
386S_ESAG6(BES126)  CATTCCAGCAGGAGTTGGAGGAAATGAGGAACGCATCCGCGTTAGCAGCAGCTGCAGCTG

386R_ESAG6(BES126)  GGATAGCAGCTGGAAGACTGGAAGAGTGGATTTTTGTATTTGCACAGGCAGCAGGCAGGT
386S_ESAG6(BES126)  GGATAGCAGCTGGAAGACTGGAAGAGTGGATTTTTGTATTTGCACAGGCAGCAGGCAGGT
                                     Hyper variable region – identifies to L427 ES 126
386R_ESAG6(BES126)  CATCACAGTTTTGCATAAGCACGGGGAAGACTGGGCCAGCAGAATACAACAACTTGCAAG
386S_ESAG6(BES126)  CATCACAGTTTTGCATAAGCACGGGGAAGACTGGGCCAGCAGAATACAACAACTTGCAAG

386R_ESAG6(BES126)  AATGTTTTGATGGAACAATTGGACCTGAAACCCTTTACAAAATTGAGGATTCACGTGTAA
386S_ESAG6(BES126)  AATGTTTTGATGGAACAATTGGACCTGAAACCCTTTACAAAATTGAGGATTCACGTGTAA

386R_ESAG6(BES126)  AAGAGTCGGCGAAGACAAGATTGCTACTCCATGAAGTGTTATTATCCATTTCCTTCGGTA
386S_ESAG6(BES126)  AAGAGTCGGCGAAGACAAGATTGCTACTCCATGAAGTGTTATTATCCATTTCCTTCGGTA

386R_ESAG6(BES126)  GCTTGGGTGCGGAAAATATCCGAGGAGGAAATGGGAAGGATGGATGTAATTTGGTGAGAA
386S_ESAG6(BES126)  GCTTGGGTGCGGAAAATATCCGAGGAGGAAATGGGAAGGATGGATGTAATTTGGTGAGAA

386R_ESAG6(BES126)  CAGACAATAACGGTATATTAAAGGGGGGGGTCACCGACACGGCACAACTTAACGTGGGGTG
386S_ESAG6(BES126)  CAGACAATAACGGTATATTAAAGGGGGGGGTCACCGACACGGCACAACTTAACGTGGGGTG

386R_ESAG6(BES126)  GTGGAGTGATGAATTTTGGATCCTATCAAAACGGAAGTATGTACGTTGAGGGTGGTGAAT
386S_ESAG6(BES126)  GTGGAGTGATGAATTTTGGATCCTATCAAAACGGAAGTATGTACGTTGAGGGTGGTGAAT

386R_ESAG6(BES126)  ACGGTGATGCTACAGAGTACGGTGCAGTACGGTGGACCGAAGATCCTAGTAAAGTGAGCA
386S_ESAG6(BES126)  ACGGTGATGCTACAGAGTACGGTGCAGTACGGTGGACCGAAGATCCTAGTAAAGTGAGCA

386R_ESAG6(BES126)  TATTTAAGGATGTCATTCGCCTGTTTGCGCGGTTCCAAGAAGCAAAAAATGCAGTGATGA
386S_ESAG6(BES126)  TATTTAAGGATGTCATTCGCCTGTTTGCGCGGTTCCAAGAAGCAAAAAATGCAGTGATGA

386R_ESAG6(BES126)  AGAAAATAAAAACTACTGTGGATGAATTGACGAAATGTATAGGGCAGAAGGAGGCTGAAC
386S_ESAG6(BES126)  AGAAAATAAAAACTACTGTGGATGAATTGACGAAATGTATAGGGCAGAAGGAGGCTGAAC

386R_ESAG6(BES126)  TCACTAATGATCAGATTTACGAGGAATTTATCTGGGAGACCATAAACAGATTGGAGCTGT
386S_ESAG6(BES126)  TCACTAATGATCAGATTTACGAGGAATTTATCTGGGAGACCATAAACAGATTGGAGCTGT

386R_ESAG6(BES126)  CAAAGAGAGTGAGTGAACAA
386S_ESAG6(BES126)  CAAAGAGAGTGAGTGAACAA
```

```
386R_ESAG7(BES126)  CGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTCCGGAATTCGCTATTAT
386S_ESAG7(BES126)  CGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTCCGGAATTCGCTATTAT

386R_ESAG7(BES126)  TAGAACAGTTTCTGTACTATATTGTAACGAAGCGGAAAGTTTTAAAAGCATAGTTTAAAA
386S_ESAG7(BES126)  TAGAACAGTTTCTGTACTATATTGTAACGAAGCGGAAAGTTTTAAAAGCATAGTTTAAAA

386R_ESAG7(BES126)  AAGTGATAAGGATGAGATTTTGGTTTGTGTTGTTGGCCCTTTTGGGAAAAGAAACATATG
386S_ESAG7(BES126)  AAGTGATAAGGATGAGATTTTGGTTTGTGTTGTTGGCCCTTTTGGGAAAAGAAACATATG

386R_ESAG7(BES126)  CGTATGAAAATGAAAGGAATGCATTAAACGCAACCGCCGCTAATAAAGTGTGTGCGCTAT
386S_ESAG7(BES126)  CGTATGAAAATGAAAGGAATGCATTAAACGCAACCGCCGCTAATAAAGTGTGTGCGCTAT

386R_ESAG7(BES126)  CGACCTATCTTAAAGGAATAGCGCACAGAGTAAACAGCGAAAGTGCTGTGGTTACGGAAA
386S_ESAG7(BES126)  CGACCTATCTTAAAGGAATAGCGCACAGAGTAAACAGCGAAAGTGCTGTGGTTACGGAAA

386R_ESAG7(BES126)  AACTATCAGATTTGAAAATGAGAAGCATCCAGTTGCAGCTAACAGTAATGCGAAACAGAG
386S_ESAG7(BES126)  AACTATCAGATTTGAAAATGAGAAGCATCCAGTTGCAGCTAACAGTAATGCGAAACAGAG

386R_ESAG7(BES126)  ATCCTTCTGGCGAGAAGGATTGTAAAGACATCAGGACACTCTTGAAAACAGTATTGAGGA
386S_ESAG7(BES126)  ATCCTTCTGGCGAGAAGGATTGTAAAGACATCAGGACACTCTTGAAAACAGTATTGAGGA

386R_ESAG7(BES126)  ATGAGTTTACATTCCAGCAGGAGTTGGAGGAAATGAGGAACGCATCCGCGTTAGCAGCAG
386S_ESAG7(BES126)  ATGAGTTTACATTCCAGCAGGAGTTGGAGGAAATGAGGAACGCATCCGCGTTAGCAGCAG

386R_ESAG7(BES126)  CTGCAGCTGGGATAGCAGCCGGAAGACTGGAAGAATGGATTTTTGTATTTGCACAGGCAG
386S_ESAG7(BES126)  CTGCAGCTGGGATAGCAGCCGGAAGACTGGAAGAATGGATTTTTGTATTTGCACAGGCAG

386R_ESAG7(BES126)  CAGGCAGGTCATCACAGTTTTGCATAAGCGTGGGGAAGCATATCCCGGCTGAGCACGGCA
386S_ESAG7(BES126)  CAGGCAGGTCATCACAGTTTTGCATAAGCGTGGGGAAGCATATCCCGGCTGAGCACGGCA

386R_ESAG7(BES126)  ACTTGCAAGAATGTTTTGACGGAACAATTGGACCTGAAACCCTTTACAAAATTGAGGATT
386S_ESAG7(BES126)  ACTTGCAAGAATGTTTTGACGGAACAATTGGACCTGAAACCCTTTACAAAATTGAGGATT

386R_ESAG7(BES126)  CACGTGTAAAAGAGTCGGCAAAGACAAGCTTGCAACTCCATGAAGTGTTATCATCCATTT
386S_ESAG7(BES126)  CACGTGTAAAAGAGTCGGCAAAGACAAGCTTGCAACTCCATGAAGTGTTATCATCCATTT

386R_ESAG7(BES126)  CCTTCAATAGCTTGGGTGCGGAAAGTATTGTTGAGCAAGGAGAAAACAGAGGATGTAACC
386S_ESAG7(BES126)  CCTTCAATAGCTTGGGTGCGGAAAGTATTGTTGAGCAAGGAGAAAACAGAGGATGTAACC

386R_ESAG7(BES126)  TAATGCGGACCGCGTACGGAGGTTTGCTGGAAAGGTGTTTGTTTGAATCGCAACTTCACA
386S_ESAG7(BES126)  TAATGCGGACCGCGTACGGAGGTTTGCTGGAAAGGTGTTTGTTTGAATCGCAACTTCACA

386R_ESAG7(BES126)  TGGGGTGCCGGAGTGTTGAATTTCGGATACTGTGTGGCGGGGAACCTAAAAATAAAAGGG
386S_ESAG7(BES126)  TGGGGTGCCGGAGTGTTGAATTTCGGATACTGTGTGGCGGGGAACCTAAAAATAAAAGGG
```

Hyper variable region – identifies to L427 ES 126

```
386R_ESAG7(BES126)  GGAGAATACGGTGATGTCAGTTCCCACGATGCGGTACGGTGGACCGAAGATCCTAGTAAA
386S_ESAG7(BES126)  GGAGAATACGGTGATGTCAGTTCCCACGATGCGGTACGGTGGACCGAAGATCCTAGTAAA

386R_ESAG7(BES126)  GTGAGCATATTTAAAGGATGTCATTCGCCTGTTTGCGCGGTTCAAAGAAGCAAAAAATGC
386S_ESAG7(BES126)  GTGAGCATATTTAAAGGATGTCATTCGCCTGTTTGCGCGGTTCAAAGAAGCAAAAAATGC

386R_ESAG7(BES126)  AGTGATGTCTAGAGCAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTA
386S_ESAG7(BES126)  AGTGATGTCTAGAGCAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTA
```

Appendix 2 *HpHbR* ORF & 3' UTR sequence for several strains of *T. brucei*. The open reading frame is denoted by the blue bar. Non-synonymous polymorphisms within the gene are shown in red, synonymous in green. In the 3'UTR polymorphisms are marked solely in red. The two homologues of the *HpHbR* region in the heterozygous type 2 *T. b. gambiense* strain STIB386 are denoted A & B.

```
                        |Gene->
TREU927   (T.b.b)       ATGGAGAAACCGTCTTGCAGGGGTGCCGGTTGGGCGCAGCTTTTGTGGTGTTACGGCACC
Eliane    (Gp 1 T.b.g)  ATGGAGAAACCGTCTTGCAGGGGTGCCGGTTGGGCGCAGCTTTTGTGGTGTTACGGCACC
STIB386A  (Gp 2 T.b.g)  ATGGAGAAACCGTCTTGCAGGGGTGCCGGTTGGGCGCAGCTTTTGTGGTGTTACGGCACC
STIB386B  (Gp 2 T.b.g)  ATGGAGAAACCGTCTTGCAGGGGTGCTGGTTGGGCGCAGCTTTTGTGGTGTTACGGCACC

TREU927   (T.b.b)       TGTTGCGCTCTACTCCTCCGCCTGATAGTTGAAGCCAGTCAAGCTGCTGAGGGTTTAAAA
Eliane    (Gp 1 T.b.g)  TGTTGCGCTCTACTCCTCCGCCTGATAGTTGAAGCCAGTCAAGCTGCTGAGGGTTTAAAA
STIB386A  (Gp 2 T.b.g)  TGTTGCGCTCTACTCCTCCGCCTGATAGTTGAAGCCAGTCAAGCTGCTGAGGGTTTAAAA
STIB386B  (Gp 2 T.b.g)  TGTTGCGCTCTACTCCTCCGCCTGATAGTTGAAGCCAGTCAAGCTGCTGAGGGTTTAAAA

TREU927   (T.b.b)       ACCAAAGACGAAGTTGAGAAGGCGTGCCATCTTGCACAACAACTCAAAGAAGTTTCAATT
Eliane    (Gp 1 T.b.g)  ACCAAAGACGAAGTTGAGAAGGCGTGCCATCTTGCACAACAACTCAAAGAAGTTTCAATT
STIB386A  (Gp 2 T.b.g)  ACCAAAGACGAAGTTGAGAAGGCGTGCCATCTTGCACAACAACTCAAAGAAGTTTCAATT
STIB386B  (Gp 2 T.b.g)  ACCAAAGACGAAGTTGAGAAGGCGTGCCATCTTGCACAACAACTCAAAGAAGTTTCAATT

TREU927   (T.b.b)       ACTTTGGGAGTTATTTACCGGACCACTGAACGACACTCCGTGCAAGTTGAAGCGCATAAA
Eliane    (Gp 1 T.b.g)  ACTTTGGGAGTTATTTACCGGACCACTGAACGACACTCCGTGCAAGTTGAAGCGCATAAA
STIB386A  (Gp 2 T.b.g)  ACTTTGGGAGTTATTTACCGGACCACTGAACGACACTCCGTGCAAGTTGAAGCGCATAAA
STIB386B  (Gp 2 T.b.g)  ACTTTGGGAGTTATTTACCGGACCACTGAACGACACTCCGTGCAAGTTGAAGCGCATAAA

TREU927   (T.b.b)       ACAGCCATTGACAAACATGCGGATGCGGTGTCGCGAGCTGTGGAGGCGCTCACGAGGGTG
Eliane    (Gp 1 T.b.g)  ACAGCCATTGACAAACATGCGGATGCGGTGTCGCGAGCTGTGGAGGCGCTCACGAGGGTG
STIB386A  (Gp 2 T.b.g)  ACAGCCATTGACAAACATGCCGATGCGGTGTCGCGAGCTGTGGAGGCGCTCACGAGGGTG
STIB386B  (Gp 2 T.b.g)  ACAGCCATTGACAAACATGCGGATGCGGTGTCGCGAGCTGTGGAGGCGCTCACGAGGGTG

TREU927   (T.b.b)       GATGTGGCACTTCAGCGATTGAAAGAACTCGGGAAGGCGAATGACACGAAGGCGGTGAAA
Eliane    (Gp 1 T.b.g)  GATGTGGCACTTCAGCGATTGAAAGAACTCGGGAAGGCAATGACACGAAGGCGGTGAAA
STIB386A  (Gp 2 T.b.g)  GATGTGGCACTTCAGCGATTGAAAGAACTCGGGAAGGCGAATGACACGAAGGCGGTGAAA
STIB386B  (Gp 2 T.b.g)  GATGTGGCACTTCAGCGATTGAAAGAACTCGGGAAGGCGAATGACACGAAGGCGGTGAAA

TREU927   (T.b.b)       ATTATCGAGAACATTACCTCCGCCAGAGAAAATCTCGCTCTCTTCAATAACGAAACGCAG
Eliane    (Gp 1 T.b.g)  ATTATCGAGAACATTACCTCCGCCAGAGAAAATCTCGCTCTCTTCAATAACGAAACGCAG
STIB386A  (Gp 2 T.b.g)  ATTATCGAGAACATTACCTCCGCCAGAGAAAATCTCGCTCTCTTCAATAACGAAACGCAG
STIB386B  (Gp 2 T.b.g)  ATTATCGAGAACATTACCTCCGCCAGAGAAAATCTCGCTCTCTTCAATAACGAAACGCAG

TREU927   (T.b.b)       GCCGTACTGACGGCGAGGGATCATGTGCATAAGCATAGGGCCGCGGCATTGCAGGGGTGG
Eliane    (Gp 1 T.b.g)  GCCGTACTGACGGCGAGGGATCATGTGCATAAGCATAGGGCCGCGGCATTGCAGGGGTGG
STIB386A  (Gp 2 T.b.g)  GCCGTACTGACGGCGAGGGATCATGTGCATAAGCATAGGGCCGCGGCATTGCAGGGGTGG
STIB386B  (Gp 2 T.b.g)  GCCGTACTGACGGCGAGGGATCATGTGCATAAGCATAGGGCCGCGGCATTGCAGGGGTGG

TREU927   (T.b.b)       TCTGATGCAAAAGAAAAAGGCGATGCCGCCGCAGAGGATGTTTGGGTTCTGCTTAATGCC
Eliane    (Gp 1 T.b.g)  TCTGATGCAAAAGAAAAAGGCGATGCCGCCGCAGAGGATGTTTGGGTTCTGCTTAATGCC
STIB386A  (Gp 2 T.b.g)  TCTGATGCAAAAGAAAAAGGCGATGCCGCCGCAGAGGATGTTTGGGTTCTGCTTAATGCC
STIB386B  (Gp 2 T.b.g)  TCTGATGCAAAAGAAAAAGGCGATGCCGCCGCAGAGGATGTTTGGGTTCTGCTTAATGCC
```

```
TREU927   (T.b.b)       GCAAAAAAAGGTAATGGCAGTGCAGACGTCAAAGCAGCTGCAGAGAAATGCTCGAGATAT
Eliane    (Gp 1 T.b.g)  GCAAAAAAAGGTAATGGCAGTGCAGACGTCAAGGCAGCTGCAGAGAAATGCTCGAGATAT
STIB386A  (Gp 2 T.b.g)  GCAAAAAAAGGTAATGGCAGTGCAGACGTCAAGGCAGCTGCAGAGAAATGCTCGAGATAT
STIB386B  (Gp 2 T.b.g)  GCAAAAAAAGGTAATGGCAGTGCAGACGTCAAGGCAGCTGCAGAGAAATGCTCGAGATAT

TREU927   (T.b.b)       TCCTCAAGCAGTACTTCAGAAACTGAGTTGCAGAAAGCTATTGACGCCGCCGCTAACGTG
Eliane    (Gp 1 T.b.g)  TCCTCAAGCAGTACTTCAGAAACTGAGTCGCAGAAAGCTATTGACGCCGCCGCTAACGTG
STIB386A  (Gp 2 T.b.g)  TCCTCAAGCAGTACTTCAGAAACTGAGTTGCAGAAAGCTATTGACGCCGCCGCTAACGTG
STIB386B  (Gp 2 T.b.g)  TCCTCAAGCAGTACTTCAGAAACTGAGTTGCAGAAAGCTATTGACGCCGCCGCTAACGTG

TREU927   (T.b.b)       GGGGGTTTGTCGGCACACAAGTCGAAATATGGCGATGTGCTGAACAAGTTTAAATTGTCT
Eliane    (Gp 1 T.b.g)  GGGGGTTTGTCGGCACACAAGTCGAAATATGGCGATGTGCTGAACAAGTTTAAATTGTCT
STIB386A  (Gp 2 T.b.g)  GGGGGTTTGTCGGCACACAAGTCGAAATATGGCGATGTGCTGAACAAGTTTAAATTGTCT
STIB386B  (Gp 2 T.b.g)  GGGGGTTTGTCGGCACACAAGTCGAAATATGGCGATGTGCTGAACAAGTTTAAATTGTCT

TREU927   (T.b.b)       AATGCTTCAGTGGGAGCAGTGAGAGACACATCCGGCCGGGGCGGTAAGCATATGGAAAAG
Eliane    (Gp 1 T.b.g)  AATGCTTCAGTGGGAGCAGTGAGAGACACATCCGGCCGGGGCGGTAAGCATATGGAAAAG
STIB386A  (Gp 2 T.b.g)  AATGCTTCAGTGGGAGCAGTGAGAGACACATCCGGCCGGGGCGGTAAGCATATGGAAAAG
STIB386B  (Gp 2 T.b.g)  AATGCTTCAGTGGGAGCAGTGAGAGACACATCCGGCCGGGGCGGTAAGCATATGGAAAAG

TREU927   (T.b.b)       GTCAATAATGTGGCAAAACTTCTTAAGGATGCAGAGGTTTCTCTTGCAGCTGCAGCAGCC
Eliane    (Gp 1 T.b.g)  GTCAATAATGTGGCAAAACTTCTTAAGGATGCAGAGGTTTCTCTTGCAGCTGCAGCAGCC
STIB386A  (Gp 2 T.b.g)  GTCAATAATGTGGCAAAACTTCTTAAGGATGCAGAGGTTTCTCTTGCAGCTGCAGCAGCC
STIB386B  (Gp 2 T.b.g)  GTCAATAATGTGGCAAAACTTCTTAAGGATGCAGAGGTTTCTCTTGCAGCTGCAGCAGCC

TREU927   (T.b.b)       GAAATTGAGGAGGTTAAAAATGCACATGAAACAAAAGCACAGGAAGAGATGAAGCGCAAT
Eliane    (Gp 1 T.b.g)  GAAATTGAGGAGGTTAAAAATGCACATGAAACAAAAGTACAGGAAGAGATGAAGCGCAAT
STIB386A  (Gp 2 T.b.g)  GAAATTGAGGAGGTTAAAAATGCACATGAAACAAAAGTACAGGAAGAGATGAAGCGCAAT
STIB386B  (Gp 2 T.b.g)  GAAATTGAGGAGGTTAAAAATGCACATGAAACAAAAGCACAGGAAGAGATGAAGCGCAAT

TREU927   (T.b.b)       GGGAACCCGATCGAAAATGAATCAGAGACTAATTCAGGGGGGGAATGCGGAATCACAAGGT
Eliane    (Gp 1 T.b.g)  GGGAACCCGATCGAAAATGAATCAGAGACTAATTCAGGGGGGGAATGCGGAATCACAAGGT
STIB386A  (Gp 2 T.b.g)  GGGAACCCGATCGAAAATGAATCAGAGACTAATTCAGGGGGGGAATGCGGAATCACAAGGT
STIB386B  (Gp 2 T.b.g)  GGGAACCCGATCGAAAATGAATCAGGACTAATTCAGGGGGGGAATGCGGAATCACAAGGT

TREU927   (T.b.b)       AATGGAGATCGTGAAGATAAGAACGACGAGCAACAACAGGTCGATGAGGAGGAAACAAAG
Eliane    (Gp 1 T.b.g)  AATGGAGATCGTGAAGATAAGAACGACGAGCAACAACAGGTCGATGAGGAGGAAACAAAG
STIB386A  (Gp 2 T.b.g)  AATGGAGATTGTGAAGATAAGAACGACGAGCAACAACAGGTCGATGAGGAGGAAACAAAG
STIB386B  (Gp 2 T.b.g)  ACTGGAGATCGTGAAGATAAGAACGACGAGCAACAACAGGTCGATGAGGAGGAAACAAAG

TREU927   (T.b.b)       GTGGAAAATGGAAGCAGCGAGGAGGGGTCTTGTTGTGGAAACGAAAGTAACGGTCCCCAT
Eliane    (Gp 1 T.b.g)  GTGGAAAATGGAAGCAGCGAGGAGGGGTCTTGTTGTGGAAACGAAAGTAACGGTCCCCAT
STIB386A  (Gp 2 T.b.g)  GTGGAAAATGGAAGCAGCGAGGAGGGGTCTTGTTGTGGAAACGAAAGTAACGGTCCCCAT
STIB386B  (Gp 2 T.b.g)  GTGGAAAATGGAAGCAGCGAGGAGAGGTCTTGTTGTGGAAACGAAAGTAACGGTCCCCAT

TREU927   (T.b.b)       GTGATGAAAAACGTCATGGGGTTGAGGGACCAAGGCCCGTTGACGTGGTTAGTGGTTTC
Eliane    (Gp 1 T.b.g)  GTGATGAAAAACGTCATGGGGTTGGGGCACCAAGGCCCGTTGACGTGGTTAGTGGTTTC
STIB386A  (Gp 2 T.b.g)  GTGATGAAAAACGTCATGGGGTTGGGGCACCAAGGCCCGTTGACGTGGTTAGTGGTTTC
STIB386B  (Gp 2 T.b.g)  GTGATGAAAAACGTCATGGGGTTGAGGGACCAAGGCCCGTTGACGTGGTTAGTGGTTTC

TREU927   (T.b.b)       CGCAGTTATGCCAGTGCTTCTTTTGCTCTGCTTTCTCTTGTCCGTGTCGGTATGCTCCAG
Eliane    (Gp 1 T.b.g)  CGCAGTTATGCCAGCGCTTCTTTTGCTCTGCTTTCTCTTGTCCGTGTCGGTATTCTCCAG
STIB386A  (Gp 2 T.b.g)  CGCAGTTATGCCAGCGCTTCTTTTGCTCTGCTTTCTCTTGTCCGTGTCGGTATGCTCCAG
STIB386B  (Gp 2 T.b.g)  CGCAGTTATGCCAGCGCTTCTTTTGCTCTGCTTTCTCTTGTCCGTGTCGGTATTCTCCAG
```

```
TREU927   (T.b.b)     GTGGTGGTGTAGAAAAGACACCGTTTCTTCCAAAGACTGCGCGGGAGTTGCTCCTAAATT
Eliane    (Gp 1 T.b.g) GTGGTGGTGTAGAAAAGACACCGTTTCTTCCAAAGACTGCGCGGGAGTTGCTCCTAAATT
STIB386A  (Gp 2 T.b.g) GTGGTGGTGTAGAAAAGACACCGTTTCTTCCAAAGACTGCGCGGGAGTTGCTCCTAAATT
STIB386B  (Gp 2 T.b.g) GTGGTGGTGTAGAAAAGACACCGTTTCTTCCAAAGACTGCGCGGGAGTTGCTCCCAAATT

TREU927   (T.b.b)     GCATTTCCTTCTCTCCTCGTAACATTAAGTGGTCGTGTTGTTGTTTTTTCTTCTTTCCTC
Eliane    (Gp 1 T.b.g) GCATTTCCTTCTCTCCTCGTAACATTAAGTGGTCGTTTGTTGTTTTTTCTTCTTTCCTC
STIB386A  (Gp 2 T.b.g) GCATTTCCTTCTCTCCTCGTAACATTAAGTGGTCGTGTTGTTGTTTTTTCTTCTTTCCTC
STIB386B  (Gp 2 T.b.g) GCATTTCCTTCTCTCCTCGTAACATTAAGTGGTCGTGTTGTTGTTTTTTCTTCTTTCCTC

TREU927   (T.b.b)     TCGCGTGTGGAGGCGAGGCTGAAGACGAACTGCGGCGTCAGACTGCCAAGTGTGAGAAAG
Eliane    (Gp 1 T.b.g) TCGCGTGTGGAGGCGAGGCTGAAGACGAACTGCAGCGTCAGACTGCCAAGTGTGAGAAAG
STIB386A  (Gp 2 T.b.g) TCGCGTGTGGAGGCGAGGCTGAAGACGAACTGCGGCGTCAGACTGCCAAGTGTGAGAAAG
STIB386B  (Gp 2 T.b.g) TCGCGTGTGGAGGCGAGGCTGAAGACGAACTGCAGCGTCAGACTGCCAAGTGTGAGAAAG

TREU927   (T.b.b)     AGCGTGAAGTGGTGTGGGGATGATGCCAACGAAGATCTTTTTAATTTTTGGATGTGGGCT
Eliane    (Gp 1 T.b.g) AGCGTGAAGTGGTGTGGGGATGATGCCAACGAAGATCTTTTTAATTTTTGGATGTGGGCT
STIB386A  (Gp 2 T.b.g) AGCGTGAAGTGGTGTGGGGATGATGCCAACGAAGATCTTTTTAATTTTTGGATGTGGGCT
STIB386B  (Gp 2 T.b.g) AGCGTGAAGTGGTGTGGGGATGATGCCAACGAAGATCTTTTTAATTTTTGGATGTGGGCT

TREU927   (T.b.b)     TACGTAGTTCGCTGGTAAATATACTTAAGATCTTTCTTTTTTTTTTTT--TGTTGTTGTAT
Eliane    (Gp 1 T.b.g) TACGTAGTTCGCTGGTAAATATACTTAAGATCTTTCTTTTTTTTTTTTGTTGTTGTTGTAT
STIB386A  (Gp 2 T.b.g) TACGTAGTTCGCTGGTAAATATACTTAAGATCTTTCTTTTTTTTTTTTTTT--TGTTGTAT
STIB386B  (Gp 2 T.b.g) TACGTAGTTCGCTGGTAAATATACTTAAGATCTTTCTTTTTTTTTTTTGTTGTTGTTGTAT

TREU927   (T.b.b)     GTCTTTTTGATTTTGTTGTATTGATTGAATTATTCCCCTTGAGCTTTCTTGTCTTGTACG
Eliane    (Gp 1 T.b.g) GTCTTTTTGATTTTGTTGTATTGACTGAATTATTCCCCTTGAGCTTTCTTGTCTTGTACG
STIB386A  (Gp 2 T.b.g) GTCTTTTTGATTTTGTTGTATTGACTGAATTATTCCCCTTGAGCTTTCTTGTCTTGTACG
STIB386B  (Gp 2 T.b.g) GTCTTTTTGATTTTGTTGTATTGACTGAATTATTCCCCTTGAGCTTTCTTGTCTTGTACG

TREU927   (T.b.b)     ATACGTTGGCGAGAAGCTGGTTAAACGAAAACTTTGAAGGAAAAAAATGTGGGATGTGTT
Eliane    (Gp 1 T.b.g) ATACGTTGGCGAGAAGCTGGTTAAACGAAAACTTTGAAGGAAAAAAATGTGGGATGTGTT
STIB386A  (Gp 2 T.b.g) ATACGTTGGCGAGAAGCTGGTTAAACGAAAACTTTGAAGGAAAAAAATGTGGGATGTGTT
STIB386B  (Gp 2 T.b.g) ATACGTTGGCGAGAAGCTGGTTAAACGAAAACTTTGAAGGAAAAAAATGTGGGATGTGTT

TREU927   (T.b.b)     TGTGTCTGCTTGTGTGTGTAAACGTAATGGAGGCTTCCCCCCCC------ATTAGTGGCA
Eliane    (Gp 1 T.b.g) TGTGTCTGTTTGTGTGTGTAAACGTAATGGAGGCTTCCCCCCCCCCC---ATTAGTGGCA
STIB386A  (Gp 2 T.b.g) TGTGTCTGTTTGTGTGTGTAAACGTAATGGAGGCTTCCCCCCCCCCCCCCATTAGTGGCA
STIB386B  (Gp 2 T.b.g) TGTGTCTGTTTGTGTGTGTAAACGTAATGGAGGCTTCCCCCCCCCCCCCCATTAGTGGCA

TREU927   (T.b.b)     TGCATGCATTTCCGATGCCGACATGCGGTTTTCGGTGAGGGAACAATGCGCCTCACCGTA
Eliane    (Gp 1 T.b.g) TGGATGCATTTCCGATGCCGACATGCGGTTTTCGGTGAGGGAACAATGCGCCTCACCGTA
STIB386A  (Gp 2 T.b.g) TGCATGCATTTCCGATGCCGACATGCGGTTTTCGGTGAGGGAACAATGCGCCTCACCGTA
STIB386B  (Gp 2 T.b.g) TGCATGCATTTCCGATGCCGACATGCGGTTTTCGGTGAGGGAACAATGCGCCTCACCGTA

TREU927   (T.b.b)     TCATGTACCGTGAACTACAAGCCACAAAAATATGTGCTTGAATAGTAAGTCCTCCCAACT
Eliane    (Gp 1 T.b.g) TCATGTACCGTGAACTACAAGCCACAAAAATATGTGCTTGAATAGTAAGTCCTCCCAACT
STIB386A  (Gp 2 T.b.g) TCATGTACCGTGAACTACAAGCCACAAAAATATGTGCTTGAATAGTAAGTCCTCCCAACT
STIB386B  (Gp 2 T.b.g) TCATGTACCGTGAACTACAAGCCACAAAAATATGTGCTTGAATAGTAAGTCCTCCCAACT

TREU927   (T.b.b)     TCATTTCATGACCAGGAAATTAATCATGTCAAGCCATTGAGCAACTTTTTCTCTTCAGTG
Eliane    (Gp 1 T.b.g) TCATTTCATGACCAGGAAATTAATCATGTCAAGCCATTGAGCAACTTTTTCTCTTCAGTG
STIB386A  (Gp 2 T.b.g) TCATTTCATGACCAGGAAATTAATCATGTCAAGCCATTGAGCAACTTTTTCTCTTCAGTG
STIB386B  (Gp 2 T.b.g) TCATTTCATGACCAGGAAATTAATCATGTCAAGCCATTGAGCAACTTTTTCTCTTCAGTG
```

```
TREU927   (T.b.b)      GTATGCCTTTGTTCGTTTTCAAAAAAAAAAA--GGAAGAATAAAGAGAAAGGGAGTGGGA
Eliane    (Gp 1 T.b.g) GTATGCCTTTGTTCGTTTTCAAAAAAAAAAAAAGGAAGAATAAAGAGAAAGGGAGTGGGA
STIB386A  (Gp 2 T.b.g) GTATGCCTTTGTTCGTTTTCAAAAAAAAAAAAAGGAAGAATAAAGAGAAAGGGAGTGGGA
STIB386B  (Gp 2 T.b.g) GTATGCCTTTGTTCGTTTTCAAAAAAAAAAAAAGGAAGAATAAAGAGAAAGGGAGTGGGA

TREU927   (T.b.b)      AAACTTTTTTTTTTTTGCTACATGTGCCTTTTCTAGTGTGTGCGAGCGCCCGCTGCGAGG
Eliane    (Gp 1 T.b.g) AAACTTTTTTTTTTT--GCTATATGTGCCTTTTCTAGTGTGTGCGAGTGCCCGCTGCGAGG
STIB386A  (Gp 2 T.b.g) AAACTTTTTTTTTTTTGCTATATGCGCCTTTTCTAGTGTGTGCGAGTGCCCGCTGCGAGG
STIB386B  (Gp 2 T.b.g) AAACTTTTTTTTTTTT-GCTATATGTGCCTTTTCTAGTGTGTGCGAGTGCCCGCTGCGAGG

TREU927   (T.b.b)      AATTGAGCTTTCTCCTTTCCTTCCTTTTCCTCCTTTC
Eliane    (Gp 1 T.b.g) AATTGAGCTTTCTCCTTTCCTTCCTTTTCCTCCTTTC
STIB386A  (Gp 2 T.b.g) AATTGAGCTTTCTCCTTTCCTTCCTTTTCCTCCTTTC
STIB386B  (Gp 2 T.b.g) AATTGAGCTTTCTCCTTTCCTTCCTTTTCCTCCTTTC
```

Appendix 3 *HpHbR* ORF & 3' UTR sequence for several strains of *T. b. gambiense* from different disease foci. Eliane originated from Côte d'Ivoire, Bim from Cameroon and Pa from the Democratic Republic of Congo. The ORF is denoted by the blue bar and non-identical base positions are highlighted in red.

```
                       |Gene->
Eliane  ATGGAGAAACCGTCTTGCAGGGGTGCCGGTTGGGCGCAGCTTTTGTGGTGTTACGGCACC
Bim     ATGGAGAAACCGTCTTGCAGGGGTGCCGGTTGGGCGCAGCTTTTGTGGTGTTACGGCACC
Pa      ATGGAGAAACCGTCTTGCAGGGGTGCCGGTTGGGCGCAGCTTTTGTGGTGTTACGGCACC

Eliane  TGTTGCGCTCTACTCCTCCGCCTGATAGTTGAAGCCAGTCAAGCTGCTGAGGGTTTAAAA
Bim     TGTTGCGCTCTACTCCTCCGCCTGATAGTTGAAGCCAGTCAAGCTGCTGAGGGTTTAAAA
Pa      TGTTGCGCTCTACTCCTCCGCCTGATAGTTGAAGCCAGTCAAGCTGCTGAGGGTTTAAAA

Eliane  ACCAAAGACGAAGTTGAGAAGGCGTGCCATCTTGCACAACAACTCAAAGAAGTTTCAATT
Bim     ACCAAAGACGAAGTTGAGAAGGCGTGCCATCTTGCACAACAACTCAAAGAAGTTTCAATT
Pa      ACCAAAGACGAAGTTGAGAAGGCGTGCCATCTTGCACAACAACTCAAAGAAGTTTCAATT

Eliane  ACTTTGGGAGTTATTTACCGGACCACTGAACGACACTCCGTGCAAGTTGAAGCGCATAAA
Bim     ACTTTGGGAGTTATTTACCGGACCACTGAACGACACTCCGTGCAAGTTGAAGCGCATAAA
Pa      ACTTTGGGAGTTATTTACCGGACCACTGAACGACACTCCGTGCAAGTTGAAGCGCATAAA

Eliane  ACAGCCATTGACAAACATGCGGATGCGGTGTCGCGAGCTGTGGAGGCGCTCACGAGGGTG
Bim     ACAGCCATTGACAAACATGCGGATGCGGTGTCGCGAGCTGTGGAGGCGCTCACGAGGGTG
Pa      ACAGCCATTGACAAACATGCGGATGCGGTGTCGCGAGCTGTGGAGGCGCTCACGAGGGTG

Eliane  GATGTGGCACTTCAGCGATTGAAAGAACTCGGGAAGGCCAATGACACGAAGGCGGTGAAA
Bim     GATGTGGCACTTCAGCGATTGAAAGAACTCGGGAAGGCCAATGACACGAAGGCGGTGAAA
Pa      GATGTGGCACTTCAGCGATTGAAAGAACTCGGGAAGGCCAATGACACGAAGGCGGTGAAA

Eliane  ATTATCGAGAACATTACCTCCGCCAGAGAAAATCTCGCTCTCTTCAATAACGAAACGCAG
Bim     ATTATCGAGAACATTACCTCCGCCAGAGAAAATCTCGCTCTCTTCAATAACGAAACGCAG
Pa      ATTATCGAGAACATTACCTCCGCCAGAGAAAATCTCGCTCTCTTCAATAACGAAACGCAG

Eliane  GCCGTACTGACGGCGAGGGATCATGTGCATAAGCATAGGGCCGCGGCATTGCAGGGGTGG
Bim     GCCGTACTGACGGCGAGGGATCATGTGCATAAGCATAGGGCCGCGGCATTGCAGGGGTGG
Pa      GCCGTACTGACGGCGAGGGATCATGTGCATAAGCATAGGGCCGCGGCATTGCAGGGGTGG
```

```
Eliane   TCTGATGCAAAAGAAAAAGGCGATGCCGCCGCAGAGGATGTTTGGGTTCTGCTTAATGCC
Bim      TCTGATGCAAAAGAAAAAGGCGATGCCGCCGCAGAGGATGTTTGGGTTCTGCTTAATGCC
Pa       TCTGATGCAAAAGAAAAAGGCGATGCCGCCGCAGAGGATGTTTGGGTTCTGCTTAATGCC

Eliane   GCAAAAAAAGGTAATGGCAGTGCAGACGTCAAGGCAGCTGCAGAGAAATGCTCGAGATAT
Bim      GCAAAAAAAGGTAATGGCAGTGCAGACGTCAAGGCAGCTGCAGAGAAATGCTCGAGATAT
Pa       GCAAAAAAAGGTAATGGCAGTGCAGACGTCAAGGCAGCTGCAGAGAAATGCTCGAGATAT

Eliane   TCCTCAAGCAGTACTTCAGAAACTGAGTCGCAGAAAGCTATTGACGCCGCCGCTAACGTG
Bim      TCCTCAAGCAGTACTTCAGAAACTGAGTCGCAGAAAGCTATTGACGCCGCCGCTAACGTG
Pa       TCCTCAAGCAGTACTTCAGAAACTGAGTCGCAGAAAGCTATTGACGCCGCCGCTAACGTG

Eliane   GGGGGTTTGTCGGCACACAAGTCGAAATATGGCGATGTGCTGAACAAGTTTAAATTGTCT
Bim      GGGGGTTTGTCGGCACACAAGTCGAAATATGGCGATGTGCTGAACAAGTTTAAATTGTCT
Pa       GGGGGTTTGTCGGCACACAAGTCGAAATATGGCGATGTGCTGAACAAGTTTAAATTGTCT

Eliane   AATGCTTCAGTGGGAGCAGTGAGAGACACATCCGGCCGGGGCGGTAAGCATATGGAAAAG
Bim      AATGCTTCAGTGGGAGCAGTGAGAGACACATCCGGCCGGGGCGGTAAGCATATGGAAAAG
Pa       AATGCTTCAGTGGGAGCAGTGAGAGACACATCCGGCCGGGGCGGTAAGCATATGGAAAAG

Eliane   GTCAATAATGTGGCAAAACTTCTTAAGGATGCAGAGGTTTCTCTTGCAGCTGCAGCAGCC
Bim      GTCAATAATGTGGCAAAACTTCTTAAGGATGCAGAGGTTTCTCTTGCAGCTGCAGCAGCC
Pa       GTCAATAATGTGGCAAAACTTCTTAAGGATGCAGAGGTTTCTCTTGCAGCTGCAGCAGCC

Eliane   GAAATTGAGGAGGTTAAAAATGCACATGAAACAAAAGTACAGGAAGAGATGAAGCGCAAT
Bim      GAAATTGAGGAGGTTAAAAATGCACATGAAACAAAAGTACAGGAAGAGATGAAGCGCAAT
Pa       GAAATTGAGGAGGTTAAAAATGCACATGAAACAAAAGTACAGGAAGAGATGAAGCGCAAT

Eliane   GGGAACCCGATCGAAAATGAATCAGAGACTAATTCAGGGGGGAATGCGGAATCACAAGGT
Bim      GGGAACCCGATCGAAAATGAATCAGAGACTAATTCAGGGGGGAATGCGGAATCACAAGGT
Pa       GGGAACCCGATCGAAAATGAATCAGAGACTAATTCAGGGGGGAATGCGGAATCACAAGGT

Eliane   AATGGAGATCGTGAAGATAAGAACGACGAGCAACAACAGGTCGATGAGGAGGAAACAAAG
Bim      AATGGAGATCGTGAAGATAAGAACGACGAGCAACAACAGGTCGATGAGGAGGAAACAAAG
Pa       AATGGAGATCGTGAAGATAAGAACGACGAGCAACAACAGGTCGATGAGGAGGAAACAAAG

Eliane   GTGGAAAATGGAAGCAGCGAGGAGGGGTCTTGTTGTGGAAACGAAAGTAACGGTCCCCAT
Bim      GTGGAAAATGGAAGCAGCGAGGAGGGGTCTTGTTGTGGAAACGAAAGTAACGGTCCCCAT
Pa       GTGGAAAATGGAAGCAGCGAGGAGGGGTCTTGTTGTGGAAACGAAAGTAACGGTCCCCAT

Eliane   GTGATGAAAAAACGTCATGGGGTTGGGGCACCAAGGCCCGTTGACGTGGTTAGTGGTTTC
Bim      GTGATGAAAAAACGTCATGGGGTTGGGGCACCAAGGCCCGTTGACGTGGTTAGTGGTTTC
Pa       GTGATGAAAAAACGTCATGGGGTTGGGGCACCAAGGCCCGTTGACGTGGTTAGTGGTTTC

Eliane   CGCAGTTATGCCAGCGCTTCTTTTGCTCTGCTTTCTCTTGTCCGTGTCGGTATTCTCCAG
Bim      CGCAGTTATGCCAGCGCTTCTTTTGCTCTGCTTTCTCTTGTCCGTGTCGGTATTCTCCAG
Pa       CGCAGTTATGCCAGCGCTTCTTTTGCTCTGCTTTCTCTTGTCCGTGTCGGTATTCTCCAG
                      <-Gene|
Eliane   GTGGTGGTGTAGAAAAGACACCGTTTCTTCCAAAGACTGCGCGGGAGTTGCTCCTAAATT
Bim      GTGGTGGTGTAGAAAAGACACCGTTTCTTCCAAAGACTGCGCGGGAGTTGCTCCTAAATT
Pa       GTGGTGGTGTAGAAAAGACACCGTTTCTTCCAAAGACTGCGCGGGAGTTGCTCCTAAATT
Eliane   GCATTTCCTTCTCTCCTCGTAACATTAAGTGGTCGTTTTGTTGTTTTTTTCTTCTTTCCTC
Bim      GCATTTCCTTCTCTCCTCGTAACATTAAGTGGTCGTTTTGTTGTTTTTTTCTTCTTTCCTC
Pa       GCATTTCCTTCTCTCCTCGTAACATTAAGTGGTCGTTTTGTTGTTTTTTTCTTCTTTCCTC
```

```
Eliane     TCGCGTGTGGAGGCGAGGCTGAAGACGAACTGCAGCGTCAGACTGCCAAGTGTGAGAAAG
Bim        TCGCGTGTGGAGGCGAGGCTGAAGACGAACTGCAGCGTCAGACTGCCAAGTGTGAGAAAG
Pa         TCGCGTGTGGAGGCGAGGCTGAAGACGAACTGCAGCGTCAGACTGCCAAGTGTGAGAAAG

Eliane     AGCGTGAAGTGGTGTGGGGATGATGCCAACGAAGATCTTTTTAATTTTTGGATGTGGGCT
Bim        AGCGTGAAGTGGTGTGGGGATGATGCCAACGAAGATCTTTTTAATTTTTGGATGTGGGCT
Pa         AGCGTGAAGTGGTGTGGGGATGATGCCAACGAAGATCTTTTTAATTTTTGGATGTGGGCT

Eliane     TACGTAGTTCGCTGGTAAATATACTTAAGATCTTTCTTTTTTTTTTTTGTTGTTGTTGTAT
Bim        TACGTAGTTCGCTGGTAAATATACTTAAGATCTTTCTTTTTTTTTTTTGTTGTTGTTGTAT
Pa         TACGTAGTTCGCTGGTAAATATACTTAAGATCTTTCTTTTTTTTTTTTGTTGTTGTTGTAT

Eliane     GTCTTTTTGATTTTGTTGTATTGACTGAATTATTCCCCTTGAGCTTTCTTGTCTTGTACG
Bim        GTCTTTTTGATTTTGTTGTATTGACTGAATTATTCCCCTTGAGCTTTCTTGTCTTGTACG
Pa         GTCTTTTTGATTTTGTTGTATTGACTGAATTATTCCCCTTGAGCTTTCTTGTCTTGTACG

Eliane     ATACGTTGGCGAGAAGCTGGTTAAACGAAAACTTTGAAGGAAAAAAATGTGGGATGTGTT
Bim        ATACGTTGGCGAGAAGCTGGTTAAACGAAAACTTTGAAGGAAAAAAATGTGGGATGTGTT
Pa         ATACGTTGGCGAGAAGCTGGTTAAACGAAAACTTTGAAGGAAAAAAATGTGGGATGTGTT

Eliane     TGTGTCTGTTTGTGTGTGTAAACGTAATGGAGGCTTCCCCCCCCCCC-ATTAGTGGCATG
Bim        TGTGTCTGTTTGTGTGTGTAAACGTAATGGAGGCTTCCCCCCCCCCCCATTAGTGGCATG
Pa         TGTGTCTGTTTGTGTGTGTAAACGTAATGGAGGCTTCCCCCCCCCCC-ATTAGTGGCATG

Eliane     GATGCATTTCCGATGCCGACATGCGGTTTTCGGTGAGGGAACAATGCGCCTCACCGTATC
Bim        GATGCATTTCCGATGCCGACATGCGGTTTTCGGTGAGGGAACAATGCGCCTCACCGTATC
Pa         GATGCATTTCCGATGCCGACATGCGGTTTTCGGTGAGGGAACAATGCGCCTCACCGTATC

Eliane     ATGTACCGTGAACTACAAGCCACAAAAATATGTGCTTGAATAGTAAGTCCTCCCAACTTC
Bim        ATGTACCGTGAACTACAAGCCACAAAAATATGTGCTTGAATAGTAAGTCCTCCCAACTTC
Pa         ATGTACCGTGAACTACAAGCCACAAAAATATGTGCTTGAATAGTAAGTCCTCCCAACTTC

Eliane     ATTTCATGACCAGGAAATTAATCATGTCAAGCCATTGAGCAACTTTTTCTCTTCAGTGGT
Bim        ATTTCATGACCAGGAAATTAATCATGTCAAGCCATTGAGCAACTTTTTCTCTTCAGTGGT
Pa         ATTTCATGACCAGGAAATTAATCATGTCAAGCCATTGAGCAACTTTTTCTCTTCAGTGGT

Eliane     ATGCCTTTGTTCGTTTTCAAAAAAAAAAAAAAGGAAGAATAAAGAGAAAGGGAGTGGGAAA
Bim        ATGCCTTTGTTCGTTTTCAAAAAAAAAAAAAAGGAAGAATAAAGAGAAAGGGAGTGGGAAA
Pa         ATGCCTTTGTTCGTTTTCAAAAAAAAAAAAAAGGAAGAATAAAGAGAAAGGGAGTGGGAAA

Eliane     ACTTTTTTTTTTGCTATATGTGCCTTTTCTAGTGTGTGCGAGTGCCCGCTGCGAGGAATT
Bim        ACTTTTTTTTTTGCTATATGTGCCTTTTCTAGTGTGTGCGAGTGCCCGCTGCGAGGAATT
Pa         ACTTTTTTTTTTGCTATATGTGCCTTTTCTAGTGTGTGCGAGTGCCCGCTGCGAGGAATT

Eliane     GAGCTTTCTCCTTTCCTTCCTTTTCCTCCTTTC
Bim        GAGCTTTCTCCTTTCCTTCCTTTTCCTCCTTTC
Pa         GAGCTTTCTCCTTTCCTTCCTTTTCCTCCTTTC
```

Appendix 4 DNA sequences of the ApoL1 cDNA clone used for the generation of recombinant ApoL1 aligned to the dominant human reference ApoL1 cDNA transcript (www.ncbi.nlm.nih.gov). Non-synonymous mutations are highlighted in red, synonymous in green.

```
ApoL-1 Reference    ATGAGTGCACTTTTCCTTGGTGTGAGAGTGAGGGCAGAGGAAGCTGGAGCGAGGGTGCAA
ApoL1-cDNA          ATGAGTGCACTTTTCCTTGGTGTGGGAGTGAGGGCAGAGGAAGCTGGAGCGAGGGTGCAA

ApoL-1 Reference    CAAAACGTTCCAAGTGGGACAGATACTGGAGATCCTCAAAGTAAGCCCCTCGGTGACTGG
ApoL1-cDNA          CAAAACGTTCCAAGTGGGACAGATACTGGAGATCCTCAAAGTAAGCCCCTCGGTGACTGG

ApoL-1 Reference    GCTGCTGGCACCATGGACCCAGAGAGCAGTATCTTTATTGAGGATGCCATTAAGTATTTC
ApoL1-cDNA          GCTGCTGGCACCATGGACCCAGAGAGCAGTATCTTTATTGAGGATGCCATTAAGTATTTC

ApoL-1 Reference    AAGGAAAAAGTGAGCACACAGAATCTGCTACTCCTGCTGACTGATAATGAGGCCTGGAAC
ApoL1-cDNA          AAGGAAAAAGTGAGCACACAGAATCTGCTACTCCTGCTGACTGATAATGAGGCCTGGAAC

ApoL-1 Reference    GGATTCGTGGCTGCTGCTGAACTGCCCAGGAATGAGGCAGATGAGCTCCGTAAAGCTCTG
ApoL1-cDNA          GGATTCGTGGCTGCTGCTGAACTGCCCAGGAATGAGGCAGATGAGCTCCGTAAAGCTCTG

ApoL-1 Reference    GACAACCTTGCAAGACAAATGATCATGAAAGACAAAAACTGGCACGATAAAGGCCAGCAG
ApoL1-cDNA          GACAACCTTGCAAGACAAATGATCATGAAAGACAAAAACTGGCACGATAAAGGCCAGCAG

ApoL-1 Reference    TACAGAAACTGGTTTCTGAAAGAGTTTCCTCGGTTGAAAAGTAAGCTTGAGGATAACATA
ApoL1-cDNA          TACAGAAACTGGTTTCTGAAAGAGTTTCCTCGGTTGAAAAGTAAGCTTGAGGATAACATA

ApoL-1 Reference    AGAAGGCTCCGTGCCCTTGCAGATGGGGTTCAGAAGGTCCACAAAGGCACCACCATCGCC
ApoL1-cDNA          AGAAGGCTCCGTGCCCTTGCAGATGGGGTTCAGAAGGTCCACAAAGGCACCACCATCGCC

ApoL-1 Reference    AATGTGGTGTCTGGCTCTCTCAGCATTTCCTCTGGCATCCTGACCCTCGTCGGCATGGGT
ApoL1-cDNA          AATGTGGTGTCTGGCTCTCTCAGCATTTCCTCTGGCATCCTGACCCTCGTCGGCATGGGT

ApoL-1 Reference    CTGGCACCCTTCACAGAGGGAGGCAGCCTTGTACTCTTGGAACCTGGGATGGAGTTGGGA
ApoL1-cDNA          CTGGCACCCTTCACAGAGGGAGGCAGCCTTGTACTCTTGGAACCTGGGATGGAGTTGGGA

ApoL-1 Reference    ATCACAGCCGCTTTGACCGGGATTACCAGCAGTACCATAGACTACGGAAAGAAGTGGTGG
ApoL1-cDNA          ATCACAGCAGCTTTGACCGGGATTACCAGCAGTACCATAGACTACGGAAAGAAGTGGTGG

ApoL-1 Reference    ACACAAGCCCAAGCCCACGACCTGGTCATCAAAAAGCCTTGACAAATTGAAGGAGGTGAAG
ApoL1-cDNA          ACACAAGCCCAAGCCCACGACCTGGTCATCAAAAAGCCTTGACAAATTGAAGGAGGTGAAG

ApoL-1 Reference    GAGTTTTTGGGTGAGAACATATCCAACTTTCTTTCCTTAGCTGGCAATACTTACCAACTC
ApoL1-cDNA          GAGTTTTTGGGTGAGAACATATCCAACTTTCTTTCCTTAGCTGGCAATACTTACCAACTC

ApoL-1 Reference    ACACGAGGCATTGGGAAGGACATCCGTGCCCTCAGACGAGCCAGAGCCAATCTTCAGTCA
ApoL1-cDNA          ACACGAGGCATTGGGAAGGACATCCGTGCCCTCAGACGAGCCAGAGCCAATCTTCAGTCA

ApoL-1 Reference    GTACCGCATGCCTCAGCCTCACGCCCCCGGGTCACTGAGCCAATCTCAGCTGAAAGCGGT
ApoL1-cDNA          GTACCGCATGCCTCAGCCTCACGCCCCCGGGTCACTGAGCCAATCTCAGCTGAAAGCGGT

ApoL-1 Reference    GAACAGGTGGAGAGGGTTAATGAACCCAGCATCCTGGAAATGAGCAGAGGAGTCAAGCTC
ApoL1-cDNA          GAACAGGTGGAGAGAGTTAATGAACCCAGCATCCTGGAAATGAGCAGAGGAGTCAAGCTC

ApoL-1 Reference    ACGGATGTGGCCCCTGTAAGCTTCTTTCTTGTGCTGGATGTAGTCTACCTCGTGTACGAA
ApoL1-cDNA          ACGGATGTGGCCCCTGTAAGCTTCTTTCTTGTGCTGGATGTAGTCTACCTCGTGTACGAA

ApoL-1 Reference    TCAAAGCACTTACATGAGGGGGCAAAGTCAGAGACAGCTGAGGAGCTGAAGAAGGTGGCT
ApoL1-cDNA          TCAAAGCACTTACATGAGGGGGCAAAGTCAGAGACAGCTGAGGAGCTGAAGAAGGTGGCT

ApoL-1 Reference    CAGGAGCTGGAGGAGAAGCTAAACATTCTCAACAATAATTATAAGATTCTGCAGGCGGAC
ApoL1-cDNA          CAGGAGCTGGAGGAGAAGCTAAACATTCTCAACAATAATTATAAGATTCTGCAGGCGGAC

ApoL-1 Reference    CAAGAACTGTGA
ApoL1-cDNA          CAAGAACTGTGA
```

Appendix 5 The allele sizes determined using Genescan® analysis for 6 different microsatellite markers for several *T. brucei* isolates, predominantly from the Daloa, Côte d'Ivoire Sleeping Sickness focus. Alleles are coloured by the bin the allele was placed into and the bin ranges are also indicated.

Name	Population	JS2 A	JS2 B	11/110 A	11/110 B	11/51 A	11/51 B	18 A	18 B	D2/7 A	D2/7 B	PLC A	PLC B
STIB247	Tanzania	97	174	162	174	87	162	165	219	143	155	169	169
STIB386	Cote d'ivoire	125	158	162	174	94	105	166	193	138	163	174	217
MALC/BF/80/AB14	Cote d'ivoire	118	118	170	178	92	92	162	162	155	155	185	185
MBOI/BF/80/TC125	Cote d'ivoire	123	123	162	162	105	105	171	171	138	138	172	172
MBOI/BF/80/TC126	Cote d'ivoire	133	162	176	188	96	96	170	183	140	143	176	176
MHOM/CI/78/DAL069	Cote d'ivoire	180	180	170	170	97	97	163	171	143	143	175	175
MHOM/CI/78/DAL072 A	Cote d'ivoire	117	117	170	170	97	97	163	171	148	148	149	175
MHOM/CI/78/TH1-032	Cote d'ivoire	116	116	171	171	98	98	161	178	143	143	181	181
MHOM/CI/78/TH112	Cote d'ivoire	108	123	174	174	99	99	169	182	140	143	185	218
MHOM/CI/78/TH126	Cote d'ivoire	108	108	170	174	97	97	169	182	139	139	169	169
MHOM/CI/78/TH149	Cote d'ivoire	134	134	169	169	97	97	163	172	138	143	175	175
MHOM/CI/78/TH170	Cote d'ivoire	134	134	172	178	96	96	164	172	136	245	176	184
MKOD/BF/80/KD3	Cote d'ivoire	136	162	163	163	106	106	178	178	159	169	170	170
MKOK/BF/80/KK1	Cote d'ivoire	119	119	174	178	91	96	162	172	167	173	176	176
MKOK/BF/80/KK14	Cote d'ivoire	134	162	163	174	96	96	166	201	140	165	176	184
MKOK/BF/80/KK17	Cote d'ivoire	114	114	169	173	96	96	163	171	140	161	170	170
MKOK/BF/80/KK18	Cote d'ivoire	112	112	163	163	96	96	169	223	140	163	186	186
MKOK/BF/80/KK26	Cote d'ivoire	112	112	163	174	96	96	169	223	140	164	185	185
MKOK/BF/80/KK33	Cote d'ivoire	136	136	174	174	94	94	195	195	142	173	183	183
MKOK/BF/80/KK7	Cote d'ivoire	134	162	163	174	99	99	165	201	140	140	176	184
MSUS/CI/78/TSW 115 [CLONE B]	Cote d'ivoire	123	123	173	173	95	95	225	279	143	143	180	180
MSUS/CI/78/TSW065-022	Cote d'ivoire	127	127	163	174	92	92	162	170	143	143	173	173
MSUS/CI/78/TSW100	Cote d'ivoire	126	126	162	171	92	92	162	170	138	143	173	173
MSUS/CI/78/TSW113	Cote d'ivoire	123	138	174	174	94	94	155	225	143	143	180	199
MSUS/CI/78/TSW158	Cote d'ivoire	138	138	174	174	94	94	164	164	143	152	175	180
MSUS/CI/78/TSW168	Cote d'ivoire	122	138	176	186	99	99	181	228	140	173	185	185
MSUS/CI/78/TSW175	Cote d'ivoire	108	126	163	176	99	158	170	170	136	161	168	168
MSUS/CI/78/TSW178	Cote d'ivoire	98	98	174	174	92	92	170	170	136	161	168	168
MSUS/CI/78/TSW182	Cote d'ivoire	108	126	163	176	99	99	170	170	136	160	174	174
MSUS/CI/78/TSW187	Cote d'ivoire	127	127	176	176	95	95	162	170	143	143	184	184
MSUS/CI/78/TSW19	Cote d'ivoire	98	103	174	174	94	94	170	185	140	173	187	218
MSUS/CI/78/TSW190	Cote d'ivoire	100	140	162	176	94	94	162	180	136	140	181	181
MSUS/CI/78/TSW196	Cote d'ivoire	125	125	163	174	99	99	170	162	141	153	175	175
MSUS/CI/78/TSW209	Cote d'ivoire	109	125	174	174	94	94	166	228	138	138	148	174
MSUS/CI/78/TSW308	Cote d'ivoire	123	123	174	174	99	99	164	164	150	150	175	217
MSUS/CI/78/TSW332	Cote d'ivoire	121	127	178	178	99	99	170	183	140	143	185	185
MSUS/CI/78/TSW77	Cote d'ivoire	138	138	169	169	97	155	169	225	140	140	174	184
VAVOUA/83/DAL542 (Adzam)	Cote d'ivoire	133	214	170	170	98	105	163	172	248	248	175	175
VAVOUA/84/DAL740 (Brazo)	Cote d'ivoire	134	215	170	170	98	98	163	172	200	243	171	175
Eliane	Cote d'ivoire	137	219	170	170	96	96	164	172	242	242	174	174
GAGNEA/83/DAL595 (Kide)	Cote d'ivoire	133	213	170	170	98	104	163	171	257	257	170	170
DALOA/83/DAL598 (Kosu)	Cote d'ivoire	136	213	170	170	98	104	163	172	256	256	175	175
DALOA/83/DAL642 (Lisa)	Cote d'ivoire	183	217	169	169	96	104	163	172	248	248	171	174
DALOA/83/DAL645 (Mago)	Cote d'ivoire	134	215	170	170	98	104	163	171	255	255	171	175
DALOA/83/DAL403 (Sakon)	Cote d'ivoire	183	217	170	170	96	104	163	172	248	248	175	175
DALOA/83/DAL633 (Seval)	Cote d'ivoire	134	215	170	170	98	104	163	172	253	253	171	171
BOUAFLE/83/DAL587 (Yage)	Cote d'ivoire	181	215	170	170	96	96	164	172	242	242	171	174
DAL972	Cote d'ivoire	134	215	170	170	98	98	163	172	253	253	171	174

	JS2		11/110		11/51		18		D2/7		PLC	
1	96-102	1	161-163	1	91-99	1	162-165	1	135-142	1	148-149
2	108	2	169-177	2	103-105	2	169-172	2	160-164	2	168-176
3	112-236	3	185-187	3	153-157	3	155	3	242-248	3	180-186
4	134-138			4	162-163	4	177-184	4	173.25	4	217-218
5	179-182			5	87	5	274-278	5	149-155	5	199
6	157-161					6	201	6	252-256		
7	213-218					7	193-195	7	182		
8	173-174					8	225-228	8	158		
						9	219	9	167		
								10	199		

References

Agbo, E. C., Majiwa, P. A. O., Claassen, E. J. H. M. and Roos, M. H. (2001). Measure of molecular diversity within the *Trypanosoma brucei* subspecies *Trypanosoma brucei brucei* and *Trypanosoma brucei gambiense* as revealed by genotypic characterization. *Experimental Parasitology,* **99**(3), 123-131.

Agbo, E. E. C., Majiwa, P. A. O., Claassen, H. J. H. M. and te Pas, M. F. W. (2002). Molecular variation of *Trypanosoma brucei* subspecies as revealed by AFLP fingerprinting. *Parasitology,* **124**, 349-358.

Alsford, S., Wickstead, B., Ersfeld, K. and Gull, K. (2001). Diversity and dynamics of the minichromosomal karyotype in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology,* **113**(1), 79-88.

Balber, A. E. (1972). *Trypanosoma-Brucei* - Fluxes of Morphological Variants in Intact and X-Irradiated Mice. *Experimental Parasitology,* **31**(2), 307-&.

Baral, T. N., Magez, S., Stijlemans, B., Conrath, K., Vanhollebeke, B., Pays, E., Muyldermans, S. and De Baetselier, P. (2006). Experimental therapy of African trypanosomiasis with a nanobody-conjugated human trypanolytic factor. *Nature Medicine,* **12**(5), 580-584.

Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J. and Krishna, S. (2003). The trypanosomiases. *Lancet,* **362**(9394), 1469-1480.

Barry, J. and McCulloch, R. (2001). Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Advances in Parasitology,* **49**, 1-70.

Barry, J. D. and Turner, C. M. (1991). The dynamics of antigenic variation and growth of African trypanosomes. *Parasitology Today,* **7**(8), 207-211.

Becker, M., Aitcheson, N., Byles, E., Wickstead, B., Louis, E. and Rudenko, G. (2004). Isolation of the repertoire of VSG expression site containing telomeres of *Trypanosoma brucei* 427 using transformation-associated recombination in yeast. *Genome Research,* **14**(11), 2319-2329.

Benne, R., Van den Burg, J., Brakenhoff, J. P., Sloof, P., Van Boom, J. H. and Tromp, M. C. (1986). Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell,* **46**(6), 819-826.

Berberof, M., Perez-Morga, D. and Pays, E. (2001). A receptor-like flagellar pocket glycoprotein specific to *Trypanosoma brucei gambiense*. *Molecular and Biochemical Parasitology,* **113**(1), 127-138.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Bohme, W., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L. H., Wickstead, B., Alsmark, U. C. M., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T. J., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetter, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., MacLeod, A., Mooney, P. J., Moule, S., Martin, D. M. A., Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. A., Rabbinowitsch, E., Rajandream, M. A., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Talton, L., Turner, C. M. R., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S. L., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barret, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E. and El-Sayed, N. M. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science,* **309**(5733), 416-422.

Bidere, N., Lorenzo, H. K., Carmona, S., Laforge, M., Harper, F., Dumont, C. and Senik, A. (2003). Cathepsin D triggers bax activation, resulting in selective apoptosis-inducing factor (AIF) relocation in T lymphocytes entering the early commitment phase to apoptosis. *Journal of Biological Chemistry,* **278**(33), 31401-31411.

Bishop, J. R., Shimamura, M. and Hajduk, S. L. (2001). Insight into the mechanism of trypanosome lytic factor-1 killing of *Trypanosoma brucei brucei*. *Molecular and Biochemical Parasitology,* **118**(1), 33-40.

Bitter, W., Gerrits, H., Kieft, R. and Borst, P. (1998). The role of transferrin-receptor variation in the host range of *Trypanosoma brucei*. *Nature,* **391**(6666), 499-502.

Blanche, P. J., Gong, E. L., Forte, T. M. and Nichols, A. V. (1981). Characterization of human high-density lipoproteins by gradient gel electrophoresis. *Biochimica et Biophysica Acta,* **665**(3), 408-419.

Bonkovsky, H. L. (1991). Iron and the liver. *American Journal of Medical Science,* **301**(1), 32-43.

Bringaud, F., Biteau, N., Donelson, J. E. and Baltz, T. (2001). Conservation of metacyclic variant surface glycoprotein expression sites among different trypanosome isolates. *Molecular and Biochemical Parasitology,* **113**(1), 67-78.

Brun, R. and Jenni, L. (1987). Human serum resistance of metacyclic forms of *Trypanosoma brucei brucei*, T. *brucei rhodesiense* and T. *brucei gambiense*. *Parasitology Research,* **73**(3), 218-223.

Brussow, H. (2009). Europe, the bull and the Minotaur: the biological legacy of a Neolithic love story. *Environmental Microbiology,* **11**(11), 2778-2788.

Burkard, G., Fragoso, C. M. and Roditi, I. (2007). Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology,* **153**(2), 220-223.

Burri, C. and Brun, R. (2003). Eflornithine for the treatment of human African trypanosomiasis. *Parasitology Research,* **90 Supp 1**, S49-52.

Burri, C., Nkunhu, S., Merolle, A., Smith, T., Blum, J. and Brun, R. (2000). Efficacy of new, concise schedule for melarsoprol in treatment of sleeping sickness caused by *Trypanosoma brucei gambiense*: a randomised trial. *Lancet,* **355**(9213), 1419-1425.

Campillo, N. and Carrington, M. (2003). The origin of the serum resistance associated (SRA) gene and a model of the structure of the SRA polypeptide from *Trypanosoma brucei rhodesiense*. *Molecular and Biochemical Parasitology,* **127**(1), 79-84.

Castelli, W. P., Doyle, J. T., Gordon, T., Hames, C. G., Hjortland, M. C., Hulley, S. B., Kagan, A. and Zukel, W. J. (1977). Hdl Cholesterol and Other Lipids in Coronary Heart-Disease - Cooperative Lipoprotein Phenotyping Study. *Circulation,* **55**(5), 767-772.

Checchi, F., Filipe, J. A., Haydon, D. T., Chandramohan, D. and Chappuis, F. (2008). Estimates of the duration of the early and late stage of gambiense sleeping sickness. *British Medical Journal: Infectious Diseases,* **8**, 8-16.

Churchill, G. A. and Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics,* **138**(3), 963-971.

Clayton, C. and Shapira, M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Molecular and Biochemical Parasitology,* **156**(2), 93-101.

Coleman, P. and Welburn, S. (2004). Are fitness costs associated with resistance to human serum in *Trypanosoma brucei rhodesiense*? *Trends in Parasitology,* **20**(7), 311-315.

Connell, G. E. and Smithies, O. (1959). Human haptoglobins: estimation and purification. *Biochemistry,* **72**(1), 115-121.

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T. D., Barnes, C., Campbell, P., Fitzgerald, T., Hu, M., Ihm, C. H., Kristiansson, K., Macarthur, D. G., Macdonald, J. R., Onyiah, I., Pang, A. W., Robson, S., Stirrups, K., Valsesia, A., Walter, K., Wei, J., Tyler-Smith, C., Carter, N. P., Lee, C., Scherer, S. W. and Hurles, M. E. (2010). Origins and functional impact of copy number variation in the human genome. *Nature,* **464**(7289), 704-712.

Conway, C., Proudfoot, C., Burton, P., Barry, J. D. and McCulloch, R. (2002). Two pathways of homologous recombination in *Trypanosoma brucei*. *Molecular Microbiology,* **45**(6), 1687-1700.

Cooper, A., Tait, A., Sweeney, L., Tweedie, A., Morrison, L., Turner, C. M. and MacLeod, A. (2008). Genetic analysis of the human infective trypanosome *Trypanosoma brucei gambiense*: chromosomal segregation, crossing over, and the construction of a genetic map. *Genome Biology,* **9**(6), R103.

De Greef, C., Chimfwembe, E., Kihang'a Wabacha, J., Bajyana Songa, E. and Hamers, R. (1992). Only the serum-resistant bloodstream forms of *Trypanosoma brucei rhodesiense* express the serum resistance associated (SRA) protein. *Ann Soc Belg Med Trop,* **72 Suppl 1**, 13-21.

De Greef, C. and Hamers, R. (1994). The serum resistance-associated (*SRA*) gene of *Trypanosoma brucei rhodesiense* encodes a variant surface glycoprotein-like protein. *Molecular and Biochemical Parasitology,* **68**(2), 277-284.

De Greef, C., Imberechts, H., Matthyssens, G., Van Meirvenne, N. and Hamers, R. (1989). A gene expressed only in serum-resistant variants of *Trypanosoma brucei rhodesiense*. *Mol Biochem Parasitol,* **36**(2), 169-176.

de Koning, H. P. (2001). Transporters in African trypanosomes: role in drug action and resistance. *International Jounral of Parasitology,* **31**(5-6), 512-522.

De Maio, A. (1999). Heat shock proteins: facts, thoughts, and dreams. *Shock,* **11**(1), 1-12.

Dero, B., Zampettibosseler, F., Pays, E. and Steinert, M. (1987). The Genome and the Antigen Gene Repertoire of *Trypanosoma-brucei-gambiense* Are Smaller Than Those of *Trypanosoma-brucei-b*rucei. *Molecular and Biochemical Parasitology,* **26**(3), 247-256.

Doerge, R. W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nature Reviews Genetics,* **3**(1), 43-52.

Drain, J., Bishop, J. and Hajduk, S. (2001). Haptoglobin-related protein mediates trypanosome lytic factor binding to trypanosomes. *Journal of Biological Chemistry,* **276**(32), 30254-30260.

Duchateau, P. N., Movsesyan, I., Yamashita, S., Sakai, N., Hirano, K., Schoenhaus, S. A., O'Connor-Kearns, P. M., Spencer, S. J., Jaffe, R. B., Redberg, R. F., Ishida, B. Y., Matsuzawa, Y., Kane, J. P. and Malloy, M. J. (2000). Plasma apolipoprotein L concentrations correlate with plasma triglycerides and cholesterol levels in normolipidemic, hyperlipidemic, and diabetic subjects. *Journal of Lipid Research,* **41**(8), 1231-1236.

Duchateau, P. N., Pullinger, C. R., Cho, M. H., Eng, C. and Kane, J. P. (2001). Apolipoprotein L gene family: tissue-specific expression, splicing, promoter regions; discovery of a new gene. *Journal of Lipid Research,* **42**(4), 620-630.

Duchateau, P. N., Pullinger, C. R., Orellana, R. E., Kunitake, S. T., Naya-Vigne, J., O'Connor, P. M., Malloy, M. J. and Kane, J. P. (1997). Apolipoprotein L, a new human high density lipoprotein apolipoprotein expressed by the pancreas. Identification, cloning, characterization, and plasma distribution of apolipoprotein L. *Journal of Biological Chemistry,* **272**(41), 25576-25582.

Dukes, P., Gibson, W. C., Gashumba, J. K., Hudson, K. M., Bromidge, T. J., Kaukus, A., Asonganyi, T. and Magnus, E. (1992). Absence of the LiTat 1.3 (CATT antigen) gene in *Trypanosoma brucei gambiense* stocks from Cameroon. *Acta Trop,* **51**(2), 123-134.

Dunn, W. A., Hubbard, A. L. and Aronson, N. N., Jr. (1980). Low temperature selectively inhibits fusion between pinocytic vesicles and lysosomes during heterophagy of 125I-asialofetuin by the perfused rat liver. *Journal of Biological Chemistry,* **255**(12), 5971-5978.

Enyaru, J. C., Matovu, E., Akol, M., Sebikali, C., Kyambadde, J., Schmidt, C., Brun, R., Kaminsky, R., Ogwal, L. M. and Kansiime, F. (1998). Parasitological detection of *Trypanosoma brucei gambiense* in serologically negative sleeping-sickness suspects from north-western Uganda. *Annals of Tropical Medical Parasitology,* **92**(8), 845-850.

Enyaru, J. C., Matovu, E., Nerima, B., Akol, M. and Sebikali, C. (2006). Detection of *T.b. rhodesiense* trypanosomes in humans and domestic animals in south east Uganda by amplification of serum resistance-associated gene. *Annals of the New York Academy of Science,* **1081**, 311-319.

Faulkner, S. D., Oli, M. W., Kieft, R., Cotlin, L., Widener, J., Shiflett, A., Cipriano, M. J., Pacocha, S. E., Birkeland, S. R., Hajduk, S. L. and McArthur, A. G. (2006). In vitro generation of human high-density-lipoprotein-resistant *Trypanosoma brucei brucei*. *Eukaryote Cell,* **5**(8), 1276-1286.

Felgner, P., Brinkmann, U., Zillmann, U., Mehlitz, D. and Abu-Ishira, S. (1981). Epidemiological studies on the animal reservoir of *gambiense* sleeping sickness. Part II. Parasitological and immunodiagnostic examination of the human population. *Tropenmed Parasitol,* **32**(3), 134-140.

Felsenstein, J. (1989). PHYLIP-phylogeny inference package (version 3.2). *Cladistics,* **5**(1), 164-166.

Ferdig, M. T., Cooper, R. A., Mu, J. B., Deng, B. B., Joy, D. A., Su, X. Z. and Wellems, T. E. (2004). Dissecting the loci of low-level quinine resistance in malaria parasites. *Molecular Microbiology,* **52**(4), 985-997.

Fernandes, A. P., Nelson, K. and Beverley, S. M. (1993). Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. *Proceeds of the National Academy of Science,* **90**(24), 11608-11612.

Fevre, E. M., Coleman, P. G., Odiit, M., Magona, J. W., Welburn, S. C. and Woolhouse, M. E. (2001). The origins of a new *Trypanosoma brucei rhodesiense* sleeping sickness outbreak in eastern Uganda. *Lancet,* **358**(9282), 625-628.

Fevre, E. M., Picozzi, K., Fyfe, J., Waiswa, C., Odiit, M., Coleman, P. G. and Welburn, S. C. (2005). A burgeoning epidemic of sleeping sickness in Uganda. *Lancet,* **366**(9487), 745-747.

Forde, O. H., Thelle, D. S., Miller, N. E. and Mjos, O. D. (1977). High-Density Lipoprotein (Hdl) and Coronary Heart-Disease (Chd) - Prospective Case Control Study. *European Journal of Clinical Investigation,* **7**(3), 242-243.

Geigy, R. and Kauffmann, M. (1973). Sleeping sickness survey in the Serengeti area (Tanzania) 1971. I. Examination of large mammals for trypanosomes. *Acta Tropica,* **30**(1), 12-23.

Geigy, R., Kauffmann, M. and Jenni, L. (1973). Wild mammals as reservoirs for Rhodesian sleeping sickness in the Serengeti, 1970-71. *Transcript of the Royal Society of Tropical Medicine and Hygiene,* **67**(2), 284-286.

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., Bowden, D. W., Langefeld, C. D., Oleksyk, T. K., Knob, A. U., Bernhardy, A. J., Hicks, P. J., Nelson, G. W., Vanhollebeke, B., Winkler, C. A., Kopp, J. B., Pays, E. and Pollak, M. R. (2010). Association of Trypanolytic ApoL1 Variants with Kidney Disease in African-Americans. *Science*.

Gerrits, H., Mussmann, R., Bitter, W., Kieft, R. and Borst, P. (2002). The physiological significance of transferrin receptor variations in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology,* **119**(2), 237-247.

Gibson, W. (2009). Analysis of a genetic cross between *Trypanosoma brucei rhodesiense* and *T. b. brucei*. *Parasitology,* **99**(03), 391-402.

Gibson, W., Backhouse, T. and Griffiths, A. (2002). The human serum resistance associated gene is ubiquitous and conserved in *Trypanosoma brucei rhodesiense* throughout East Africa. *Infect Genet Evol,* **1**(3), 207-214.

Gibson, W. and Garside, L. (1991). Genetic exchange in *Trypanosoma brucei brucei*: variable chromosomal location of housekeeping genes in different trypanosome stocks. *Molecular and Biochemical Parasitology,* **45**(1), 77-89.

Gibson, W., Garside, L. and Bailey, M. (1992). Trisomy and chromosome size changes in hybrid trypanosomes from a genetic cross between *Trypanosoma brucei rhodesiense* and T. b. *brucei*. *Molecular and Biochemical Parasitology,* **51**(2), 189-199.

Gibson, W., Kanmogne, G. and Bailey, M. (1995). A successful backcross in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology,* **69**(1), 101-110.

Gibson, W., Mehlitz, D., Lanham, S. M. and Godfrey, D. G. (1978). The identification of *Trypanosoma brucei gambiense* in Liberian pigs and dogs by isoenzymes and by resistance to human plasma. *Tropenmed Parasitol,* **29**(3), 335-345.

Gibson, W., Nemetschke, L. and Ndung'u, J. (2010). Conserved sequence of the *TgsGP* gene in Group 1 *Trypanosoma brucei gambiense*. *Infect Genet Evol,* **10**(4), 453-458.

Gibson, W., Peacock, L., Ferris, V., Williams, K. and Bailey, M. (2006). Analysis of a cross between green and red fluorescent trypanosomes. *Biochemical Society Transactions,* **34**(4), 557-559.

Gibson, W. and Stevens, J. (1999). Genetic exchange in the *trypanosoma*tidae. *Advances in Parasitology,* **43**, 1-46.

Gibson, W. C. (1986). Will the real *Trypanosoma b. gambiense* please stand up. *Parasitology Today,* **2**(9), 255-257.

Gibson, W. C. (2005). The SRA gene: the key to understanding the nature of *Trypanosoma brucei rhodesiense*. *Parasitology,* **131**(Pt 2), 143-150.

Gibson, W. C. and Borst, P. (1986). Size-fractionation of the small chromosomes of Trypanozoon and Nannomonas trypanosomes by pulsed field gradient gel electrophoresis. *Molecular and Biochemical Parasitology,* **18**(2), 127-140.

Gibson, W. C., de, C. M. T. F. and Godfrey, D. G. (1980). Numerical analysis of enzyme polymorphism: a new approach to the epidemiology and taxonomy of trypanosomes of the subgenus Trypanozoon. *Advances in Parasitology,* **18**, 175-246.

Gibson, W. C., Parr, C. W., Swindlehurst, C. A. and Welch, S. G. (1978). A comparison of the isoenzymes, soluble proteins, polypeptides and free amino acids from ten isolates of *Trypanosoma evansi*. *Comparative Biochemistry and Physiology,* **60**(2), 137-142.

Gibson, W. C. and Wellde, B. T. (1985). Characterization of Trypanozoon stocks from the South Nyanza sleeping sickness focus in Western Kenya. *Transcripts of the Royal Society for Tropical Medicine and Hygiene,* **79**(5), 671-676.

Gillett, M. P. and Owen, J. S. (1991). *Trypanosoma brucei brucei*: differences in the trypanocidal activity of human plasma and its relationship to the level of high density lipoproteins. *Transcripts of the Royal Society for Tropical Medicine and Hygiene,* **85**(5), 612-616.

Glazier, A. M., Nadeau, J. H. and Aitman, T. J. (2002). Finding genes that underlie complex traits. *Science,* **298**(5602), 2345-2349.

Godfrey, D., Baker, R., Rickman, L. and Mehlitz, D. (1990). The distribution, relationships and identification of enzymic variants within the subgenus Trypanozoon. *Advances in Parasitology,* **29**, 1-74.

Godfrey, D. G. and Kilgour, V. (1976). Enzyme electrophoresis in characterizing the causative organism of Gambian trypanosomiasis. *Transcripts of the Royal Society for Tropical Medicine and Hygiene,* **70**(3), 219-224.

Godfrey, D. G., Scott, C. M., Gibson, W. C., Mehlitz, D. and Zillmann, U. (1987). Enzyme Polymorphism and the Identity of *Trypanosoma-Brucei-Gambiense*. *Parasitology,* **94**, 337-347.

Gottesdiener, K., Garcia-Anoveros, J., Lee, M. G. and Van der Ploeg, L. H. (1990). Chromosome organization of the protozoan *Trypanosoma brucei*. *Molecular and Biochemical Parasitology,* **10**(11), 6079-6083.

Grab, D. J. and Kennedy, P. G. (2008). Traversal of human and animal trypanosomes across the blood-brain barrier. *Journal of Neurovirology,* **14**(5), 344-351.

Hager, K., Pierce, M., Moore, D., Tytler, E., Esko, J. and Hajduk, S. (1994). Endocytosis of a cytotoxic human high density lipoprotein results in disruption of acidic intracellular vesicles and subsequent killing of African trypanosomes. *Journal of Cell Biology,* **126**(1), 155-167.

Hager, K. M. and Hajduk, S. L. (1997). Mechanism of resistance of African trypanosomes to cytotoxic human HDL. *Nature,* **385**(6619), 823-826.

Hajduk, S. L., Hager, K. M. and Esko, J. D. (1994). Human high density lipoprotein killing of African trypanosomes. *Annual Review of Microbiology,* **48**, 139-162.

Hajduk, S. L., Moore, D. R., Vasudevacharya, J., Siqueira, H., Torri, A. F., Tytler, E. M. and Esko, J. D. (1989). Lysis of *Trypanosoma brucei* by a toxic subspecies of human high density lipoprotein. *Journal of Biological Chemistry,* **264**(9), 5210-5217.

Hamilton, P. B., Gibson, W. C. and Stevens, J. R. (2007). Patterns of co-evolution between trypanosomes and their hosts deduced from ribosomal RNA and protein-coding gene phylogenies. *Molecular Phylogenetics and Evolution,* **44**(1), 15-25.

Hamilton, P. B., Stevens, J. R., Gaunt, M. W., Gidley, J. and Gibson, W. C. (2004). Trypanosomes are monophyletic: evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *International Journal of Parasitology,* **34**(12), 1393-1404.

Harrington, J. M., Widener, J., Stephens, N., Johnson, T., Francia, M., Capewell, P., Macleod, A. and Hajduk, S. L. (2010). The plasma membrane of bloodstream form African trypanosomes confers susceptibility and specificity to killing by hydrophobic peptides. *Journal of Biological Chemistry*.

Heinecke, J. W. (2009). The HDL proteome: a marker--and perhaps mediator--of coronary artery disease. *Journal of Lipid Research,* **50 Suppl**, S167-171.

Hertz-Fowler, C., Figueiredo, L. M., Quail, M. A., Becker, M., Jackson, A., Bason, N., Brooks, K., Churcher, C., Fahkro, S., Goodhead, I., Heath, P., Kartvelishvili, M., Mungall, K., Harris, D., Hauser, H., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Taylor, J. E., Walker, D., White, B., Young, R., Cross, G. A. M., Rudenko, G., Barry, J. D., Louis, E. J. and Berriman, M. (2008). Telomeric Expression Sites Are Highly Conserved in *Trypanosoma brucei*. *Plos One,* **3**(10), -.

Hide, G., Welburn, S. C., Tait, A. and Maudlin, I. (1994). Epidemiological relationships of *Trypanosoma brucei* stocks from south east Uganda: evidence for different population structures in human infective and non-human infective isolates. *Parasitology,* **109 (Pt 1)**, 95-111.

Hirumi, H. and Hirumi, K. (1989). Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *Journal of Parasitology,* **75**(6), 985-989.

Hope, M., MacLeod, A., Leech, V., Melville, S., Sasse, J., Tait, A. and Turner, C. M. (1999). Analysis of ploidy (in megabase chromosomes) in *Trypanosoma brucei* after genetic exchange. *Molecular and Biochemical Parasitology,* **104**(1), 1-9.

Hutchinson, O., Fevre, E., Carrington, M. and Welburn, S. (2003). Lessons learned from the emergence of a new *Trypanosoma brucei rhodesiense* sleeping sickness focus in Uganda. *Lancet: Infectious Diseases,* **3**(1), 42-45.

Imrie, H., Ferguson, D. and Day, K. (2004). Human serum haptoglobin is toxic to *Plasmodium falciparum* in vitro. *Molecular and Biochemical Parasitology,* **133**(1), 93-98.

Jackson, A. P., Sanders, M., Berry, A., McQuillan, J., Aslett, M. A., Quail, M. A., Chukualim, B., Capewell, P., MacLeod, A., Melville, S. E., Gibson, W., Barry, J. D., Berriman, M. and Hertz-Fowler, C. (2010). The Genome Sequence of *Trypanosoma brucei gambiense*, Causative Agent of Chronic Human African Trypanosomiasis. *Plos Neglected Tropical Diseases,* **4**(4), -.

Jeffreys, A. J., Neumann, R. and Wilson, V. (1990). Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell,* **60**(3), 473-485.

Jeffreys, A. J., Wilson, V. and Thein, S. L. (1985). Hypervariable Minisatellite Regions in Human DNA. *Nature,* **314**(6006), 67-73.

Jenni, L. and Brun, R. (1982). A new in vitro test for human serum resistance of *Trypanosoma (T.) brucei*. *Acta Tropica,* **39**(3), 281-284.

Jenni, L., Marti, S., Schweizer, J., Betschart, B., Le Page, R., Wells, J., Tait, A., Paindavoine, P., Pays, E. and Steinert, M. (1986). Hybrid formation between African trypanosomes during cyclical transmission. *Nature,* **322**(6075):173-5.

Kaba, D., Dje, N. N., Courtin, F., Oke, E., Koffi, M., Garcia, A., Jamonneau, V. and Solano, P. (2006). [The impact of war on the evolution of sleeping sickness in west-central Cote d'Ivoire]. *Tropical Medicine and International Health,* **11**(2), 136-143.

Kabani, S., Fenn, K., Ross, A., Ivens, A., Smith, T. K., Ghazal, P. and Matthews, K. (2009). Genome-wide expression profiling of in vivo-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in *Trypanosoma brucei*. *BMC Genomics,* **10**, 427.

Kaern, M., Elston, T. C., Blake, W. J. and Collins, J. J. (2005). Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics,* **6**(6), 451-464.

Kageruka, P., Mangus, E., Songa, E. B., Nantulya, V., Jochems, M., Hamers, R. and Mortelmans, J. (1991). Infectivity of *Trypanosoma* (Trypanozoon) *brucei gambiense* for Baboons (*Papio-hamadryas, Papio-papio*). *Annales De La Societe Belge De Medecine Tropicale,* **71**(1), 39-45.

Kieft, R., Capewell, P., Turner, C. M., Veitch, N. J., Macleod, A. and Hajduk, S. (2010). Mechanism of *Trypanosoma brucei gambiense* (group 1) resistance to human trypanosome lytic factor. *Proceedings of the National Academy of Sciences of the United States of America*.

Kilgour, V. (1980). The electrophoretic mobilities and activities of eleven enzymes of bloodstream and culture forms of *Trypanosoma brucei* compared. *Molecular and Biochemical Parasitology,* **2**(1), 51-62.

Koffi, M., De Meeus, T., Bucheton, B., Solano, P., Camara, M., Kaba, D., Cuny, G., Ayala, F. J. and Jamonneau, V. (2009). Population genetics of *Trypanosoma brucei gambiense*, the agent of sleeping sickness in Western Africa. *Proceedings of the National Academy of Sciences of the United States of America,* **106**(1), 209-214.

Koffi, M., Solano, P., Denizot, M., Courtin, D., Garcia, A., Lejon, V., Buscher, P., Cuny, G. and Jamonneau, V. (2006). Aparasitemic serological suspects in *Trypanosoma brucei gambiense* human African trypanosomiasis: a potential human reservoir of parasites? *Acta Tropica,* **98**(2), 183-188.

Kooy, R. F., Hirumi, H., Moloo, S. K., Nantulya, V. M., Dukes, P., Van der Linden, P. M., Duijndam, W. A., Janse, C. J. and Overdulve, J. P. (1989). Evidence for diploidy in metacyclic forms of African trypanosomes. *Proceedings of the National Academy of Sciences of the United States of America,* **86**(14), 5469-5472.

Krief, S., Escalante, A. A., Pacheco, M. A., Mugisha, L., Andre, C., Halbwax, M., Fischer, A., Krief, J. M., Kasenene, J. M., Crandfield, M., Cornejo, O. E., Chavatte, J. M., Lin, C., Letourneur, F., Gruner, A. C., McCutchan, T. F., Renia, L. and Snounou, G. (2010). On the diversity of malaria parasites in African apes and the origin of *Plasmodium falciparum* from Bonobos. *PLoS Pathogens,* **6**(2), e1000765.

Krzewska, J., Langer, T. and Liberek, K. (2001). Mitochondrial Hsp78, a member of the Clp/Hsp100 family in *Saccharomyces cerevisiae*, cooperates with Hsp70 in protein refolding. *FEBS Letters,* **489**(1), 92-96.

Lane, R. and Crosskey, R. (1993). *Medical insects and arachnids,* Chapman & Hall.

Laveran, A. and Mesnil, F. (1912). Trypanosomiase humaine ou maladie du sommeil. *Trypanosomes et Trypanosomiases*, 673–795.

Leach, T. M. and Roberts, C. J. (1981). Present status of chemotherapy and chemoprophylaxis of animal trypanosomiasis in the Eastern hemisphere. *Pharmacological Therapies,* **13**(1), 91-147.

Lecordier, L., Vanhollebeke, B., Poelvoorde, P., Tebabi, P., Paturiaux-Hanocq, F., Andris, F., Lins, L. and Pays, E. (2009). C-terminal mutants of apolipoprotein L-I efficiently kill both *Trypanosoma brucei brucei* and *Trypanosoma brucei rhodesiense. PLoS Pathogens,* **5**(12), e1000685.

Legros, D., Fournier, C., Gastellu, E., Maiso, F. and Szumilin, E. (1999). Therapeutic failure of melarsoprol among patients treated for late stage *Tb gambiense* human African trypanosomiasis in Uganda. *Bulletin de la Societe de pathologie exotiques,* **92**(3), 171.

Li, J. J., Min, R. Q., Vizeacoumar, F. J., Jin, K., Xin, X. F. and Zhang, Z. L. (2010). Exploiting the determinants of stochastic gene expression in *Saccharomyces cerevisiae* for genome-wide prediction of expression noise. *Proceedings of the National Academy of Sciences of the United States of America,* **107**(23), 10472-10477.

Ligtenberg, M., Bitter, W., Kieft, R., Steverding, D., Janssen, H., Calafat, J. and Borst, P. (1994). Reconstitution of a surface transferrin binding complex in insect form *Trypanosoma brucei. The EMBO Journal,* **13**(11), 2565.

Liu, W., Apagyi, K., McLeavy, L. and Ersfeld, K. (2010). Expression and cellular localisation of calpain-like proteins in *Trypanosoma brucei. Molecular and Biochemical Parasitology,* **169**(1), 20-26.

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* **25**(4), 402-408.

Lorenz, P., Barth, P., Rudin, W. and Betschart, B. (1994). Importance of acidic intracellular compartments in the lysis of *Trypanosoma brucei brucei* by normal human serum. *Transcript of the Royal Society of Tropical Medicine and Hygiene,* **88**(4), 487-488.

Lugli, E. B., Pouliot, M., Portela, M. D. M., Loomis, M. R. and Raper, J. (2004). Characterization of primate trypanosome lytic factors. *Molecular and Biochemical Parasitology,* **138**(1), 9-20.

Lumsden, W. H. R., Herbert, W. J., Gillian, J. C. and Mcneilla.Jc (1968). Nomenclature of Antigenic Types of Trypanosomes. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* **62**(1), 130-&.

Lutumba, P., Makieya, E., Shaw, A., Meheus, F. and Boelaert, M. (2007). Human African trypanosomiasis in a rural community, Democratic Republic of Congo. *Emerging Infectious Diseases,* **13**(2), 248-254.

MacLean, L., Chisi, J., Odiit, M., Gibson, W., Ferris, V., Picozzi, K. and Sternberg, J. (2004). Severity of human African trypanosomiasis in East Africa is associated with geographic location, parasite genotype, and host inflammatory cytokine response profile. *Infection and immunity, 72*(12), 7040.

MacLeod, A., Tweedie, A., McLellan, S., Taylor, S., Cooper, A., Sweeney, L., Turner, C. and Tait, A. (2005a). Allelic segregation and independent assortment in *T. brucei* crosses: proof that the genetic system is Mendelian and involves meiosis. *Molecular and Biochemical Parasitology, 143*(1), 12-19.

MacLeod, A., Tweedie, A., McLellan, S., Taylor, S., Hall, N., Berriman, M., El-Sayed, N. M., Hope, M., Turner, C. M. R. and Tait, A. (2005b). The genetic map and comparative analysis with the physical map of *Trypanosoma brucei. Nucleic Acids Research, 33*(21), 6688-6693.

MacLeod, A., Tweedie, A., Welburn, S. C., Maudlin, I., Turner, C. M. and Tait, A. (2000). Minisatellite marker analysis of *Trypanosoma brucei*: reconciliation of clonal, panmictic, and epidemic population genetic structures. *Proceedings of the National Academy of Sciences of the United States of America, 97*(24), 13442-13447.

MacLeod, A., Welburn, S., Maudlin, I., Turner, C. M. and Tait, A. (2001). Evidence for multiple origins of human infectivity in *Trypanosoma brucei* revealed by minisatellite variant repeat mapping. *Journal of Molecular Evolution, 52*(3), 290-301.

Magnus, E., Vervoort, T. and Van Meirvenne, N. (1978). A card-agglutination test with stained trypanosomes (CATT) for the serological diagnosis of TB *gambiense* trypanosomiasis. *Annales De La Societe Belge De Medecine Tropicale, 58*(3), 169.

Manly, K. F., Cudmore, R. H., Jr. and Meer, J. M. (2001). Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome, 12*(12), 930-932.

Marcello, L. and Barry, J. D. (2007). Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Research, 17*(9), 1344-1352.

Maresca, B. and Carratu, L. (1992). The biology of the heat shock response in parasites. *Parasitology Today, 8*(8), 260-266.

Maser, P., Luscher, A. and Kaminsky, R. (2003). Drug transport and drug resistance in African trypanosomes. *Drug Resistance Update, 6*(5), 281-290.

Mathieu-daude, F. and Tibayrenc, M. (1994). Isozyme Variability of *Trypanosoma-Brucei* S-L - Genetic, Taxonomic, and Epidemiologic Significance. *Experimental Parasitology, 78*(1), 1-19.

Matovu, E., Enyaru, J. C. K., Legros, D., Schmid, C., Seebeck, T. and Kaminsky, R. (2001). Melarsoprol refractory *T-b. gambiense* from Omugo, north-western Uganda. *Tropical Medicine & International Health, 6*(5), 407-411.

Matthews, K. R. (1999). Developments in the differentiation of *Trypanosoma brucei. Parasitology Today, 15*(2), 76-80.

Matthews, K. R. (2005). The developmental cell biology of *Trypanosoma brucei. Journal of Cell Science, 118*(Pt 2), 283-290.

Maynard-Smith, J. M., Smith, N. H., Orourke, M. and Spratt, B. G. (1993). How Clonal Are Bacteria. *Proceedings of the National Academy of Sciences of the United States of America, 90*(10), 4384-4388.

McCulloch, R. and Barry, J. D. (1999). A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes & Development,* **13**(21), 2875-2888.

McEvoy, R. E., Loveland, K. A. and Landry, S. H. (1988). The functions of immediate echolalia in autistic children: a developmental perspective. *Journal of Autism and Developmental Disorders,* **18**(4), 657-668.

Mehlitz, D., Zillmann, U., Scott, C. M. and Godfrey, D. G. (1982). Epidemiological studies on the animal reservoir of *Gambiense* sleeping sickness. Part III. Characterization of trypanozoon stocks by isoenzymes and sensitivity to human serum. *Tropenmed Parasitol,* **33**(2), 113-118.

Melville, S. E., Leech, V., Gerrard, C. S., Tait, A. and Blackwell, J. M. (1998). The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* and the assignment of chromosome markers. *Molecular and Biochemical Parasitology,* **94**(2), 155-173.

Miller, G. J. (1978). High-Density Lipoprotein, Low-Density Lipoprotein, and Coronary Heart-Disease. *Thorax,* **33**(2), 137-139.

Milner, J. D. and Hajduk, S. L. (1999). Expression and localization of serum resistance associated protein in *Trypanosoma brucei rhodesiense*. *Molecular and Biochemical Parasitology,* **104**(2), 271-283.

Molina-Portela Mdel, P., Lugli, E. B., Recio-Pinto, E. and Raper, J. (2005). Trypanosome lytic factor, a subclass of high-density lipoprotein, forms cation-selective pores in membranes. *Molecular and Biochemical Parasitology,* **144**(2), 218-226.

Molina-Portela, M. P., Samanovic, M. and Raper, J. (2008). Distinct roles of apolipoprotein components within the trypanosome lytic factor complex revealed in a novel transgenic mouse model. *Journal of Experimental Medicine,* **205**(8), 1721-1728.

Monajemi, H., Fontijn, R. D., Pannekoek, H. and Horrevoets, A. J. G. (2002). The apolipoprotein L gene cluster has emerged recently in evolution and is expressed in human vascular tissue. *Genomics,* **79**(4), 539-546.

Moore, D. R., Smith, A., Hager, K. M., Waldon, R., Esko, J. D. and Hajduk, S. L. (1995). Developmentally regulated sensitivity of *Trypanosoma brucei brucei* to the cytotoxic effects of human high-density lipoprotein. *Experimental Parasitology,* **81**(2), 216-226.

Morgan, G. W., Allen, C. L., Jeffries, T. R., Hollinshead, M. and Field, M. C. (2001). Developmental and morphological regulation of clathrin-mediated endocytosis in *Trypanosoma brucei*. *Journal of Cell Science,* **114**(Pt 14), 2605-2615.

Morrison, L. J., Tait, A., McCormack, G., Sweeney, L., Black, A., Truc, P., Likeufack, A. C., Turner, C. M. and MacLeod, A. (2008). *Trypanosoma brucei gambiense* Type 1 populations from human patients are clonal and display geographical genetic differentiation. *Infection, Genetics and Evolution,* **8**(6), 847-854.

Morrison, L. J., Tait, A., McLellan, S., Sweeney, L., Turner, C. M. R. and MacLeod, A. (2009). A Major Genetic Locus in *Trypanosoma brucei* Is a Determinant of Host Pathology. *Plos Neglected Tropical Diseases,* **3**(12), -.

Muranjan, M., Wang, Q., Li, Y. L., Hamilton, E., OtienoOmondi, F. P., Wang, J., vanPraagh, A., Grootenhuis, J. G. and Black, S. J. (1997). The trypanocidal Cape buffalo serum protein is xanthine oxidase. *Infection and immunity,* **65**(9), 3806-3814.

Natesan, S., Peacock, L., Leung, K., Gibson, W. and Field, M. (2010). Evidence that low endocytic activity is not directly responsible for human serum resistance in the insect form of African trypanosomes. *BMC Research Notes,* **3**(1), 63.

Navarro, M. and Gull, K. (2001). A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature,* **414**(6865), 759-763.

Nei, M. and Roychoudhury, A. K. (1974). Sampling variances of heterozygosity and genetic distance. *Genetics,* **76**(2), 379-390.

Njiokou, F., Laveissere, C., Simo, G., Nkinin, S., Grebaut, P., Cuny, G. and Herder, S. (2006). Wild fauna as a probable animal reservoir for *Trypanosoma brucei gambiense* in Cameroon. *Infection Genetics and Evolution,* **6**(2), 147-153.

Nok, A. J. (2003). Arsenicals (melarsoprol), pentamidine and suramin in the treatment of human African trypanosomiasis. *Parasitology Research,* **90**(1), 71-79.

Odiit, M., Coleman, P. G., Liu, W. C., McDermott, J. J., Fevre, E. M., Welburn, S. C. and Woolhouse, M. E. (2005). Quantifying the level of under-detection of *Trypanosoma brucei rhodesiense* sleeping sickness cases. *Tropical Medicine and International Health,* **10**(9), 840-849.

Oli, M., Cotlin, L., Shiflett, A. and Hajduk, S. (2006). Serum resistance-associated protein blocks lysosomal targeting of trypanosome lytic factor in *Trypanosoma brucei*. *Eukaryotic Cell,* **5**(1), 132.

Ortiz-Ordóñez, J. and Seed, J. (1995). The removal of trypanolytic activity from human serum by *Trypanosoma brucei gambiense* and its subsequent recovery in trypanosome lysates. *The Journal of parasitology,* **81**(4), 555-558.

Ortiz-Ordóñez, J., Sechelski, J. B. and Seed, J. (1994). Characterization of human serum-resistant and serum-sensitive clones from a single *Trypanosoma brucei gambiense* parental clone. *Journal of Parasitology,* **80**(4), 550-557.

Overath, P. and Engstler, M. (2004). Endocytosis, membrane recycling and sorting of GPI-anchored proteins: *Trypanosoma brucei* as a model system. *Molecular Microbiology,* **53**(3), 735-744.

Owen, J. S., Gillett, M. P. and Hughes, T. E. (1992). Transgenic mice expressing human apolipoprotein A-I have sera with modest trypanolytic activity in vitro but remain susceptible to infection by *Trypanosoma brucei brucei*. *Journal of Lipid Research,* **33**(11), 1639-1646.

Paindavoine, P., Pays, E., Laurent, M., Geltmeyer, Y., Le Ray, D., Mehlitz, D. and Steinert, M. (1986). The use of DNA hybridization and numerical taxonomy in determining relationships between *Trypanosoma brucei* stocks and subspecies. *Parasitology,* **92**, 31.

Paindavoine, P., Zampettibosseler, F., Coquelet, H., Pays, E. and Steinert, M. (1989). Different Allele Frequencies in *Trypanosoma-Brucei-Brucei* and *Trypanosoma-Brucei-Gambiense* Populations. *Molecular and Biochemical Parasitology,* **32**(1), 61-71.

Parker, K. R. and Mant, M. J. (1979). Effects of tsetse (*Glossina morsitans morsitans* Westw.) (Diptera: *Glossinidae*) salivary gland homogenate on coagulation and fibrinolysis. *Thrombosis and Haemostasis,* **42**(2), 743-751.

Pays, E., Lips, S., Nolan, D., Vanhamme, L. and Perez-Morga, D. (2001). The *VSG* expression sites of *Trypanosoma brucei*: multipurpose tools for the adaptation of the parasite to mammalian hosts. *Molecular and Biochemical Parasitology,* **114**(1), 1-16.

Pays, E., Tebabi, P., Pays, A., Coquelet, H., Revelard, P., Salmon, D. and Steinert, M. (1989). The genes and transcripts of an antigen gene expression site from T. *brucei. Cell,* **57**(5), 835-845.

Pays, E. and Vanhollebeke, B. (2009). Human innate immunity against African trypanosomes. *Current Opinions in Immunology,* **21**(5), 493-498.

Pays, E., Vanhollebeke, B., Vanhamme, L., Paturiaux-Hanocq, F., Nolan, D. P. and Perez-Morga, D. (2006). The trypanolytic factor of human serum. *Nature Reviews Microbiology,* **4**(6), 477-486.

Peacock, L., Ferris, V., Bailey, M. and Gibson, W. (2008). Fly transmission and mating of *Trypanosoma brucei brucei* strain 427. *Molecular and Biochemical Parasitology,* **160**(2), 100-106.

Peacock, L., Ferris, V., Sharma, R., Sunter, J., Bailey, M., Carrington, M. and Gibson, W. (2011). Identification of the meiotic life cycle stage of *Trypanosoma brucei* in the tsetse fly. *Proceedings of the National Academy of Sciences of the United States of America*.

Peck, R. F., Shiflett, A. M., Schwartz, K. J., McCann, A., Hajduk, S. L. and Bangs, J. D. (2008). The LAMP-like protein p67 plays an essential role in the lysosome of African trypanosomes. *Molecular Microbiology,* **68**(4), 933-946.

Perez-Morga, D., Vanhollebeke, B., Paturiaux-Hanocq, F., Nolan, D. P., Lins, L., Homble, F., Vanhamme, L., Tebabi, P., Pays, A., Poelvoorde, P., Jacquet, A., Brasseur, R. and Pays, E. (2005). Apolipoprotein L-1 promotes trypanosome lysis by forming pores in lysosomal membranes. *Science,* **309**(5733), 469-472.

Picozzi, K., Fevre, E. M., Odiit, M., Carrington, M., Eisler, M. C., Maudlin, I. and Welburn, S. C. (2005). Sleeping sickness in Uganda: a thin line between two fatal diseases. *British Medical Journal,* **331**(7527), 1238-1241.

Pike, B. L., Saylors, K. E., Fair, J. N., Lebreton, M., Tamoufe, U., Djoko, C. F., Rimoin, A. W. and Wolfe, N. D. (2010). The origin and prevention of pandemics. *Clinical Infectious Diseases,* **50**(12), 1636-1640.

Poelvoorde, P., Vanhamme, L., Van Den Abbeele, J., Switzer, W. and Pays, E. (2004). Distribution of apolipoprotein L-I and trypanosome lytic activity among primate sera. *Molecular and Biochemical Parasitology,* **134**(1), 155-157.

Portela, M. P. M., Raper, J. and Tomlinson, S. (2000). An investigation into the mechanism of trypanosome lysis by human serum factors. *Molecular and Biochemical Parasitology,* **110**(2), 273-282.

Radwanska, M., Claes, F., Magez, S., Magnus, E., Perez-Morga, D., Pays, E. and Buscher, P. (2002). Novel primer sequences for polymerase chain reaction-based detection of *Trypanosoma brucei gambiense. American Journal of Tropical Medicine and Hygiene,* **67**(3), 289-295.

Raper, J., Fung, R., Ghiso, J., Nussenzweig, V. and Tomlinson, S. (1999). Characterization of a novel trypanosome lytic factor from human serum. *Infection and immunity,* **67**(4), 1910-1916.

Raper, J., Portela, M. P., Lugli, E., Frevert, U. and Tomlinson, S. (2001). Trypanosome lytic factors: novel mediators of human innate immunity. *Current Opinions in Microbiology,* **4**(4), 402-408.

Raser, J. M. and O'Shea, E. K. (2004). Control of stochasticity in eukaryotic gene expression. *Science,* **304**(5678), 1811-1814.

Rickman, L. R. and Robson, J. (1970). Blood Incubation Infectivity Test - a Simple Test Which May Serve to Distinguish *Trypanosoma-brucei* from *T-rhodesiense. Bulletin of the World Health Organization,* **42**(4), 650-&.

Rieseberg, L., Archer, M. and Wayne, R. (1999). Transgressive segregation, adaptation and speciation. *Heredity,* **83**(4), 363-372.

Rifkin, M. (1978). *Trypanosoma brucei*: some properties of the cytotoxic reaction induced by normal human serum. *Experimental Parasitology,* **46**(2), 189-206.

Rifkin, M. R. (1984). *Trypanosoma brucei*: biochemical and morphological studies of cytotoxicity caused by normal human serum. *Experimental Parasitology,* **58**(1), 81-93.

Rifkin, M. R., De Greef, C., Jiwa, A., Landsberger, F. R. and Shapiro, S. Z. (1994). Human serum-sensitive *Trypanosoma brucei rhodesiense*: a comparison with serologically identical human serum-resistant clones. *Molecular and Biochemical Parasitology,* **66**(2), 211-220.

Roditi, I. and Clayton, C. (1999). An unambiguous nomenclature for the major surface glycoproteins of the procyclic form of *Trypanosoma brucei. Molecular and Biochemical Parasitology,* **103**(1), 99-100.

Rogers, D. and Robinson, T. (2004). Tsetse distribution. *The trypanosomiases*, 139–179.

Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Molecular Biology,* **132**(3), 365-386.

Salmon, D., Geuskens, M., Hanocq, F., Hanocq-Quertier, J., Nolan, D., Ruben, L. and Pays, E. (1994). A novel heterodimeric transferrin receptor encoded by a pair of *VSG* expression site-associated genes in T. *brucei. Cell,* **78**(1), 75-86.

Samanovic, M., Molina-Portela, M. P., Chessler, A. D., Burleigh, B. A. and Raper, J. (2009). Trypanosome lytic factor, an antimicrobial high-density lipoprotein, ameliorates Leishmania infection. *PLoS Pathogens,* **5**(1), e1000276.

Schweizer, J., Pospichal, H., Hide, G., Buchanan, N., Tait, A. and Jenni, L. (1994). Analysis of a new genetic cross between two East African *Trypanosoma brucei* clones. *Parasitology,* **109**, 83.

Seaton, G., Haley, C. S., Knott, S. A., Kearsey, M. and Visscher, P. M. (2002). QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics,* **18**(2), 339-340.

Seed, J., Sechelski, J. and Loomis, M. (1990). A survey for a trypanocidal factor in primate sera. *Journal of Protozoology,* **37**(5), 393-400.

Seyfang, A., Mecke, D. and Duszenko, M. (1990). Degradation, Recycling, and Shedding of *Trypanosoma-Brucei* Variant Surface Glycoprotein. *Journal of Protozoology,* **37**(6), 546-552.

Shiflett, A. M., Bishop, J. R., Pahwa, A. and Hajduk, S. L. (2005). Human high density lipoproteins are platforms for the assembly of multi-component innate immune complexes. *Journal of Biological Chemistry,* **280**(38), 32578-32585.

Shiflett, A. M., Faulkner, S. D., Cotlin, L. F., Widener, J., Stephens, N. and Hajduk, S. L. (2007). African trypanosomes: intracellular trafficking of host defense molecules. *Journal of Eukaryote Microbiology,* **54**(1), 18-21.

Shimamura, M., Hager, K. and Hajduk, S. (2001). The lysosomal targeting and intracellular metabolism of trypanosome lytic factor by *Trypanosoma brucei brucei. Molecular and Biochemical Parasitology,* **115**(2), 227-237.

Sibley, C. H. and Hunt, S. Y. (2003). Drug resistance in parasites: can we stay ahead of the evolutionary curve? *Trends Parasitology,* **19**(11), 532-537.

Simpson, L. (1972). The kinetoplast of the hemoflagellates. *American Journal of Tropical Medicine and Hygiene,* **32**, 139–207.

Smith, A. and Hajduk, S. (1995). Identification of haptoglobin as a natural inhibitor of trypanocidal activity in human serum. *Proceedings of the National Academy of Sciences of the United States of America,* **92**(22), 10262-10266.

Sternberg, J., Tait, A., Haley, S., Wells, J. M., Le Page, R. W., Schweizer, J. and Jenni, L. (1988). Gene exchange in African trypanosomes: characterisation of a new hybrid genotype. *Molecular and Biochemical Parasitology,* **27**(2-3), 191-200.

Stevens, J., Noyes, H., Schofield, C. and Gibson, W. (2001). The molecular evolution of *Trypanosomatidae*. *Advances in Parasitology*, 1-53.

Stevens, J. R., Noyes, H. A., Dover, G. A. and Gibson, W. C. (1999). The ancient and divergent origins of the human pathogenic trypanosomes, *Trypanosoma brucei* and T. cruzi. *Parasitology,* **118 (Pt 1)**, 107-116.

Stevens, J. R. and Tibayrenc, M. (1996). *Trypanosoma brucei* sl: Evolution, linkage and the clonality debate. *Parasitology,* **112**, 481-488.

Steverding, D., Stierhof, Y., Chaudhri, M., Ligtenberg, M., Schell, D., Beck-Sickinger, A. and Overath, P. (1994). ESAG 6 and 7 products of *Trypanosoma brucei* form a transferrin binding protein complex. *European journal of cell biology,* **64**(1), 78.

Steverding, D., Stierhof, Y. D., Fuchs, H., Tauber, R. and Overath, P. (1995). Transferrin-binding protein complex is the receptor for transferrin uptake in *Trypanosoma brucei*. *Journal of Cell Biology,* **131**(5), 1173-1182.

Stich, A., Abel, P. M. and Krishna, S. (2002). Human African trypanosomiasis. *British Medical Journal,* **325**(7357), 203-206.

Stockdale, C., Swiderski, M. R., Barry, J. D. and McCulloch, R. (2008). Antigenic variation in *Trypanosoma brucei*: Joining the DOTs. *Plos Biology,* **6**(7), 1386-1391.

Su, C. L., Howe, D. K., Dubey, J. P., Ajioka, J. W. and Sibley, L. D. (2002). Identification of quantitative trait loci controlling acute virulence in Toxoplasma gondii. *Proceedings of the National Academy of Sciences of the United States of America,* **99**(16), 10753-10758.

Tait, A. (1980). Evidence for diploidy and mating in trypanosomes. *Nature,* **287**(5782), 536-538.

Tait, A. (1983). Sexual processes in the *kinetoplastida*. *Parasitology,* **86 (Pt 4)**, 29-57.

Tait, A., Babiker, E. A. and Le Ray, D. (1984). Enzyme variation in *Trypanosoma brucei* spp. I. Evidence for the sub-speciation of *Trypanosoma brucei gambiense*. *Parasitology,* **89 (Pt 2)**, 311-326.

Tait, A., Barry, J., Wink, R., Sanderson, A. and Crowe, J. (1985). Enzyme variation in T. *brucei* ssp. II. Evidence for T. b. *rhodesiense* being a set of variants of T. b. *brucei*. *Parasitology,* **90**(01), 89-100.

Tait, A., Buchanan, N., Hide, G. and Turner, C. M. (1996). Self-fertilisation in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology,* **76**(1-2), 31-42.

Tait, A., Macleod, A., Tweedie, A., Masiga, D. and Turner, C. M. (2007). Genetic exchange in *Trypanosoma brucei*: evidence for mating prior to metacyclic stage development. *Molecular and Biochemical Parasitology,* **151**(1), 133-136.

Tait, A., Turner, C., Le Page, R. and Wells, J. (1989). Genetic evidence that metacyclic forms of *Trypanosoma brucei* are diploid. *Molecular and Biochemical Parasitology,* **37**(2), 247-255.

Tait, A. and Turner, C. M. (1990). Genetic exchange in *Trypanosoma brucei*. *Parasitology Today,* **6**(3), 70-75.

Targett, G. A. and Wilson, V. C. (1973). The blood incubation infectivity test as a means of distinguishing between *Trypanosoma brucei brucei* and *T. brucei rhodesiense*. *International Jounral of Parasitology,* **3**(1), 5-11.

Taylor, S., Barragan, A., Su, C., Fux, B., Fentress, S. J., Tang, K., Beatty, W. L., El Hajj, H., Jerome, M., Behnke, M. S., White, M., Wootton, J. C. and Sibley, L. D. (2006). A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen *Toxoplasma gondii. Science,* **314**(5806), 1776-1780.

Ter Kuile, B. H. and Opperdoes, F. R. (1991). Glucose uptake by *Trypanosoma brucei*. Rate-limiting steps in glycolysis and regulation of the glycolytic flux. *Journal of Biological Chemistry,* **266**(2), 857-862.

Thomson, R., Molina-Portela, P., Mott, H., Carrington, M. and Raper, J. (2009). Hydrodynamic gene delivery of baboon trypanosome lytic factor eliminates both animal and human-infective African trypanosomes. *Proceedings of the National Academy of Sciences of the United States of America,* **106**(46), 19509-19514.

Tibayrenc, M., Kjellberg, F. and Ayala, F. J. (1990). A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. *Proceedings of the National Academy of Sciences of the United States of America,* **87**(7), 2414-2418.

Tomlinson, S., Jansen, A. M., Koudinov, A., Ghiso, J. A., Choi-Miura, N. H., Rifkin, M. R., Ohtaki, S. and Nussenzweig, V. (1995). High-density-lipoprotein-independent killing of *Trypanosoma brucei* by human serum. *Mol Biochem Parasitology,* **70**(1-2), 131-138.

Tomlinson, S., Muranjan, M., Nussenzweig, V. and Raper, J. (1997). Haptoglobin-related protein and apolipoprotein AI are components of the two trypanolytic factors in human serum. *Molecular and Biochemical Parasitology,* **86**(1), 117-120.

Truc, P. and Tibayrenc, M. (1993). Population-Genetics of *Trypanosoma-Brucei* in Central Africa - Taxonomic and Epidemiologic Significance. *Parasitology,* **106**, 137-149.

Turner, C., Sternberg, J., Buchanan, N., Smith, E., Hide, G. and Tait, A. (1990). Evidence that the mechanism of gene exchange in *Trypanosoma brucei* involves meiosis and syngamy. *Parasitology,* **101**(03), 377-386.

Turner, C. M., McLellan, S., Lindergard, L. A., Bisoni, L., Tait, A. and MacLeod, A. (2004). Human infectivity trait in *Trypanosoma brucei*: stability, heritability and relationship to sra expression. *Parasitology,* **129**(Pt 4), 445-454.

Tytler, E. M., Moore, D. R., Pierce, M. A., Hager, K. M., Esko, J. D. and Hajduk, S. L. (1995). Reconstitution of the trypanolytic factor from components of a subspecies of human high-density lipoproteins. *Molecular and Biochemical Parasitology,* **69**(1), 9-17.

van Luenen, H. G. A. M., Kieft, R., Mussmann, R., Engstler, M., ter Riet, B. and Borst, P. (2005). Trypanosomes change their transferrin receptor expression to allow effective uptake of host transferrin. *Molecular Microbiology,* **58**(1), 151-165.

van Meirvenne, N., Maginus, E. and Janssens, P. G. (1976). The effect of normal human serum on trypanosomes of distinct antigenic type (ETat 1 to 12) isolated from a strain of *Trypanosoma brucei rhodesiense*. *Annales De La Societe Belge De Medecine Tropicale,* **56**(1), 55-63.

Van Ooijen, J. W. (1999). LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity,* **83**, 613-624.

Vanderploeg, L. H. T., Schwartz, D. C., Cantor, C. R. and Borst, P. (1984). Antigenic Variation in *Trypanosoma Brucei* Analyzed by Electrophoretic Separation of Chromosome-Sized DNA-Molecules. *Cell,* **37**(1), 77-84.

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D. P., Lins, L., Van Den Abbeele, J., Pays, A., Tebabi, P., Van Xong, H., Jacquet, A., Moguilevsky, N., Dieu, M., Kane, J. P., De Baetselier, P., Brasseur, R. and Pays, E. (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature,* **422**(6927), 83-87.

Vanhamme, L. and Pays, E. (2004). The trypanosome lytic factor of human serum and the molecular basis of sleeping sickness. *International Journal of Parasitology,* **34**(8), 887-898.

Vanhamme, L., Renauld, H., Lecordier, L., Poelvoorde, P., Van Den Abbeele, J. and Pays, E. (2004). The *Trypanosoma brucei* reference strain TREU927/4 contains T. *brucei rhodesiense*-specific SRA sequences, but displays a distinct phenotype of relative resistance to human serum. *Molecular and Biochemical Parasitology,* **135**(1), 39-47.

Vanhollebeke, B., De Muylder, G., Nielsen, M., Pays, A., Tebabi, P., Dieu, M., Raes, M., Moestrup, S. and Pays, E. (2008). A haptoglobin-hemoglobin receptor conveys innate immunity to *Trypanosoma brucei* in humans. *Science,* **320**(5876), 677-681.

Vanhollebeke, B., Lecordier, L., Perez-Morga, D., Amiguet-Vercher, A. and Pays, E. (2007a). Human serum lyses *Trypanosoma brucei* by triggering uncontrolled swelling of the parasite lysosome. *Journal of Eukaryote Microbiology,* **54**(5), 448-451.

Vanhollebeke, B., Nielsen, M. J., Watanabe, Y., Truc, P., Vanhamme, L., Nakajima, K., Moestrup, S. K. and Pays, E. (2007b). Distinct roles of haptoglobin-related protein and apolipoprotein L-I in trypanolysis by human serum. *Proceedings of the National Academy of Sciences of the United States of America,* **104**(10), 4118-4123.

Vanhollebeke, B. and Pays, E. (2006). The function of apolipoproteins L. *Cellular and Molecular Life Sciences,* **63**(17), 1937-1944.

Vanhollebeke, B. and Pays, E. (2010). The trypanolytic factor of human serum: many ways to enter the parasite, a single way to kill. *Molecular Microbiology*, 76(4):806-814.

Veitch, N. J., Johnson, P. C., Trivedi, U., Terry, S., Wildridge, D. and MacLeod, A. (2010). Digital gene expression analysis of two life cycle stages of the human-infective parasite, *Trypanosoma brucei gambiense* reveals differentially expressed clusters of co-regulated genes. *BMC Genomics,* **11**, 124.

Vickerman, K. (1969). On Surface Coat and Flagellar Adhesion in Trypanosomes. *Journal of Cell Science,* **5**(1), 163-&.

Volfson, D., Marciniak, J., Blake, W. J., Ostroff, N., Tsimring, L. S. and Hasty, J. (2006). Origins of extrinsic variability in eukaryotic gene expression. *Nature,* **439**(7078), 861-864.

Volkman, S. K., Barry, A. E., Lyons, E. J., Nielsen, K. M., Thomas, S. M., Choi, M., Thakore, S. S., Day, K. P., Wirth, D. F. and Hartl, D. L. (2001). Recent origin of *Plasmodium falciparum* from a single progenitor. *Science,* **293**(5529), 482-484.

Wang, Z. F., Morris, J. C., Drew, M. E. and Englund, P. T. (2000). Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. *Journal of Biological Chemistry,* **275**(51), 40174-40179.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. and Lathrop, M. (1992). A second-generation linkage map of the human genome. *Nature,* **359**(6398), 794-801.

Welburn, S. C., Picozzi, K., Fevre, E. M., Coleman, P. G., Odiit, M., Carrington, M. and Maudlin, I. (2001). Identification of human-infective trypanosomes in animal reservoir of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. *Lancet,* **358**(9298), 2017-2019.

WHO (2006). Human African trypanosomiasis (sleeping sickness): epidemiological update. *Weekly Epidemiol Record,* **81**(8), 71-80.

Willett, K. C. and Fairbairn, H. (1955). The Tinde experiment: a study of *Trypanosoma rhodesiense* during eighteen years of cyclical transmission. *Annals of Tropical Medical Parasitology,* **49**(3), 278-292.

Wright, S. (1978). Evolution and the genetics of populations, Vol. 4. *Variability within and among natural populations. University of Chicago Press, Chicago, Illinois*.

Xong, H. V., Vanhamme, L., Chamekh, M., Chimfwembe, C. E., Van Den Abbeele, J., Pays, A., Van Meirvenne, N., Hamers, R., De Baetselier, P. and Pays, E. (1998). A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense. Cell,* **95**(6), 839-846.

Young, R., Taylor, J. E., Kurioka, A., Becker, M., Louis, E. J. and Rudenko, G. (2008). Isolation and analysis of the genetic diversity of repertoires of VSG expression site containing telomeres from *Trypanosoma brucei gambiense*, *T. b. brucei* and *T. equiperdum*. *BMC Genomics,* **9**, 385.

Ziegelbauer, K. and Overath, P. (1993). Organization of two invariant surface glycoproteins in the surface coat of *Trypanosoma brucei*. *Infection and immunity,* **61**(11), 4540-4545.

Zillmann, U., Mehlitz, D. and Sachs, R. (1984). Identity of Trypanozoon stocks isolated from man and a domestic dog in Liberia. *Tropenmed Parasitol,* **35**(2), 105-108.