Molecular epidemiology of African sleeping sickness

Hide, G and Tait, A

http://dx.doi.org/10.1017/S0031182009990333

Molecular Epidemiology of African Sleeping Sickness

GEOFF HIDE[a] AND ANDREW TAIT[b]

[a]*Centre for Parasitology and Disease, Biomedical Sciences Research Institute, School of Environment and Life Sciences, University of Salford, Salford, UK, M5 4WT*

[b]*Faculty of Veterinary Medicine, Division of Infection and Immunity, University of Glasgow, Bearsden Road, Glasgow, G61 1QH, Scotland, U.K.*

Corresponding Author:

Geoff Hide

Centre for Parasitology and Disease, Biomedical Sciences Research Institute, School of Environment and Life Sciences, University of Salford, Salford, UK, M5 4WT

Telephone No: 0161 295 3371

Fascimile No: 0161 295 5015

E-mail No: g.hide@salford.ac.uk

Running title: Molecular epidemiology of African sleeping sickness

Keywords: African trypanosomiasis; *Trypanosoma brucei*; sleeping sickness; molecular epidemiology; epidemics; microsatellites; genetics; population structure

SUMMARY

Human sleeping sickness in Africa, caused by *Trypanosoma brucei* spp. raises a number of questions. Despite the widespread distribution of the tsetse vectors and animal trypanosomiasis, human disease is only found in discrete foci which periodically give rise to epidemics followed by periods of endemicity A key to unravelling this puzzle is a detailed knowledge of the etiological agents responsible for different patterns of disease – knowledge that is difficult to achieve using traditional microscopy. The science of molecular epidemiology has developed a range of tools which have enabled us to accurately identify taxonomic groups at all levels (species, subspecies, populations, strains and isolates). Using these tools, we can now investigate the genetic interactions within and between populations of *Trypanosoma brucei* and gain an understanding of the distinction between human- and nonhuman-infective sub species. In this review, we discuss the development of these tools, their advantages and disadvantages and describe how they have been used to understand parasite genetic diversity, the origin of epidemics, the role of reservoir hosts and the population structure. Using the specific case of *T.b.rhodesiense* in Uganda, we illustrate how molecular epidemiology has enabled us to construct a more detailed understanding of the origins, generation and dynamics of sleeping sickness epidemics.

INTRODUCTION

Human sleeping sickness in Africa is a puzzling disease. It is caused by the parasite *Trypanosoma brucei* (see table 1). Despite the widespread distribution of the tsetse vectors and animal trypanosomiasis, human disease is found in very specific foci and appears as epidemics followed by periods of endemicity (reviewed in Hide, 1999). The locations of these foci are very stable over time but occasionally new epidemics emerge in new localities. For example, the Busoga focus of Uganda has had epidemics recorded from 1900 to the present day while a new epidemic has emerged in the nearby Soroti region – an area free of sleeping sickness prior to 1998 (Welburn *et al.,* 2001). An understanding of the factors which cause these epidemics could clearly lead to an understanding of how to eradicate the disease. Sleeping sickness was one of the earliest examples of the application of molecular epidemiology to a parasitic disease and, as will be seen, some important insights have been made into the origins and dynamics of these epidemics. In this review, we will trace the development of molecular epidemiological approaches with a particular emphasis on African sleeping sickness. As the journal “Parasitology” has played a significant role in publishing these developments it is an appropriate topic for the centenary issue.

MOLECULAR EPIDEMIOLOGY

The molecular epidemiology of parasites describes the application of DNA based techniques to the understanding of parasite epidemiology and diversity (Hide & Tait, 1991). The approach was prompted by an emerging understanding that parasites could not be easily packaged into groups of organisms that produced defined diseases in defined hosts. Instead, there were complex epidemiological cycles involving multiple hosts, sometimes multiple

diseases, reservoir hosts, specific strains, emerging drug resistance, genetic exchange between strains and even the apparent emergence of new parasite species. Thus tools were needed which could be used to gain a higher level of resolution than morphology in defining the organisms involved. In fact, the discipline of molecular epidemiology strictly emerged from several earlier strands of research. Principally this was the use of phenotypic approaches such as isoenzyme (or allozyme) analysis as a tool to study the epidemiology of parasitic infections. One of the key epidemiological questions in parasitology is determining the correct identity (diagnosis) and taxomomic level of a given parasite (eg. species, subspecies, strains, variants). This has often proved challenging especially in the case of microscopic parasites (e.g. parasitic protozoa) where morphological distinctions are not easily identified. Two linked problems exist – defining the scope of a taxonomic unit and establishing unambiguous markers for each unit. The scope of a taxonomic unit could be relatively easily established in the absence of any form of gene exchange between units such as occurs in clearly distinct species or clonally reproducing parasites. However in earlier studies the precise identity of parasitic species or the role of gene exchange was often not known. An inability to clearly define taxonomic units therefore makes it difficult to identify markers that can be used diagnostically to distinguish that taxonomic unit. There is unfortunately a degree of circularity in the problem. There are two levels of identification needed: (1) diagnosis – generally referring to diagnostic tools capable of discriminating to the species level and defining a particular association between causative agent and disease and (2) epidemiological tools – generally referring to markers that can be used at the sub-species or strain level to investigate detailed interactions within parasite transmission cycles. This problem of circularity in developing appropriate markers can be overcome by using independent markers that are not linked to the, often very subjective, medical criteria by

which parasites were traditionally classified. DNA based molecular methods offered a potential solution to these problems.

FROM BIOCHEMICAL TO MOLECULAR EPIDEMIOLOGY

The advent of protein based technologies such as gel electrophoretic separation of isoenzymes (allozymes) enabled distinction of morphologically identical organisms and could be used as a tool for genetic studies. These approaches were initially used with the free-living protozoa both as markers for studying genetic exchange and for defining species (previously defined as syngens) as illustrated by work with *Paramecium aurelia* (Tait, 1970a; Tait, 1970b). Key centres in Edinburgh, London and subsequently Bristol then applied these approaches to investigating the genetics and taxonomy of the medically important protozoa, *Plasmodium* and *Trypanosoma*. For example these approaches were used for identifying different levels of taxonomic units (eg species, subspecies and strains) in *Plasmodium* (Carter & Walliker, 1975) and *Trypanosoma* spp (Kilgour and Godfrey, 1973; Miles *et al.,* 1977; Godfrey ,1978; Gibson *et al.,* 1978; Gibson *et al.,* 1980; Tait *et al.,* 1984, 1985). Furthermore, using these techniques, genetic exchange was shown to occur in *Plasmodium* spp (Walliker *et al.,* 1971, 1973) and *Trypanosoma spp* (Tait,1980; Gibson *et al.,* 1980; Jenni *et al.,* 1986; Gibson *et al.,* 1989).

When used to analyse the African trypanosomes causing human sleeping sickness, these approaches revealed a complex taxonomic situation. The classical taxonomy described three subspecies, *Trypanosoma brucei brucei*, (non-human infective) *Trypanosoma brucei rhodesiense* (East African, human infective) and *Trypanosoma brucei gambiense* (West African, human infective). Multilocus enzyme electrophoresis (MLEE) studies revealed that

there were clearly subgroups or strains within both *T.b.gambiense* and *T.b.rhodesiense* as well as a high degree of variation between isolates of *T.b.brucei* (Gibson *et al.,* 1980; Tait *et al.,* 1984, 1985). Furthermore, these studies showed that different levels of taxonomy appeared to be operating – the genetic relationships showed that *T.b.gambiense* appeared to behave like a separate subspecies while *T.b.rhodesiense* appeared to be a host range variant of *T.b.brucei*.

It was becoming apparent that the question of what defines a species in parasitic protozoa (and other micro-pathogens) was a complex one to address. Traditionally, the species concept is based on presence or absence of genetic exchange between taxonomic units. In the case of *T.brucei* sspp, the extent of the occurrence of genetic exchange in natural populations was unclear. Tibayrenc *et al.,* 1990 conducted a reanalysis of MLEE data from a wide range of pathogens including *T.brucei* and *Plasmodium* and proposed a clonal theory, based on a range of measures such as high levels of heterozygosity, linkage disequilibrium and over representation of particular genotypes. In the case of *T.b.rhodesiense* there was evidence for clonality but the results with *T.b.brucei* were less convincing. Further analysis of a large collection of isolates from across Africa (Mathieu-Daude & Tibayrenc, 1994) was interpreted as showing limited genetic exchange although geographical sub-structuring could be an explanation given the diversity of the origins of the strains. These conclusions conflicted with published data showing the existence of genetic exchange (eg Tait and Turner, 1990). Analysis of a large collection of isolates from the Lambwe Valley focus, in kenya, provided a partial resolution of the debate by showing that the population had an epidemic population structure (Maynard-Smith *et al.,* 1993) i.e. that is there was underlying genetic exchange masked by the local expansion of a small number of genotypes. Further analysis of populations from Cote d'Ivoire, Zambia and Busoga, Uganda, (Stevens and

Tibayrenc, 1995) showed a diversity of population structures with evidence for clonality in some populations but an epidemic structure in others (Uganda). This diversity was also illustrated by MLEE analysis of sympatric stocks collected during a sleeping sickness epidemic in Tororo, Uganda (Figure 1). In this study, it was demonstrated that the population structure of the *T.b.brucei* stocks appeared to conform to a random mating population structure while the population structure of *T.b.rhodesiense* was recognized as epidemic (Hide *et al.,* 1994). The debate over the role of genetic exchange in *T.brucei* was not fully resolved using MLEE.

MLEE is a highly informative technique and enabled epidemiological analyses to be based on a genetic interpretation, with the essential conclusions for African trypanosomes remaining unchallenged today. However, these methods require relatively large quantities of pure parasite material and so are cumbersome for large scale epidemiological analysis. Initially, blood from patients and animals had to be isolated, followed by several passages through mice (or in culture) and further mouse amplification to produce enough material for analysis. In some cases, the Kit for In Vitro Isolation (KIVI) of trypanosomes has been used to isolate trypanosomes (Truc *et al.,* 1992). During these processes, cloning from individual trypanosomes was also often necessary with several years work required to analyse an appropriately sized sample set. An additional problem associated with these methods is the potential of the methods to select for particular types of trypanosomes, such as those better adapted to growth in mice, and thereby introduce bias into the distribution of parasites recovered (Koffi *et al*., 2007).

Initially, DNA technology was focussed on developing diagnostic tools to identify *Trypanosoma* species. As it is the subject of another review in this issue (Gibson, 2009), it will not be considered further here and we will consider the development of DNA based tools

for measuring intraspecific variation (Table 2). Initial DNA studies on intraspecific variation in *T.brucei* were based on analysis of restriction enzyme site variation in the kinetoplast DNA (Borst *et al.,* 1981). However with limited stocks available and a relatively small number of restriction enzyme sites covered, it was not possible to discriminate between the three subspecies. A larger scale analysis (Gibson *et al.,* 1985) however demonstrated that a difference could be detected in stocks originating from East and West Africa. Restriction Fragment Length Polymorphism (RFLP) of VSG antigen genes and numerical taxonomy, based on those RFLPs, revealed that *T.b.gambiense* was quite distinct from *T.b.rhodesiense* or *T.b.brucei* but that the latter two subspecies could not be distinguished (Paindavoine *et al.,* 1986). However, it was discovered that two forms of *T.b.gambiense* existed (Paindavoine *et al.,* 1986; 1989) one of which was probably the classical *T.b.gambiense* (Type 1) and one which has been called either non-gambiense or Type 2 or “Bouafle” (Gibson, 1986; Paindavoine *et al.,* 1989; Gibson, 2007). This latter group is relatively rare and may constitute a variant of *T.b.brucei* that has acquired human infectivity (Gibson, 2007). RFLP methods were developed using repetitive DNA sequences including the ribosomal RNA genes to investigate RFLP variation using multiple loci (Hide *et al.,* 1990). Although these data could not be interpreted genetically and could not detect mixtures of trypanosomes, they offered a highly discriminatory approach to “fingerprinting” trypanosome isolates. The RFLP data combined with numerical taxonomy clearly distinguished the two types of *T.b.gambiense* and showed that the Type 2 appeared to be a subset of West African *T.b.brucei* strains (Hide *et al.,* 1990). Furthermore, comparison of two *T.b.rhodesiense* foci in Zambia and Uganda showed that there were two genetically distinct groups in the two foci (Hide *et al.,* 1991). A detailed analysis of the 1988 sleeping sickness epidemic in Tororo, Uganda, showed that the human isolates (ie *T.b.rhodesiense*) were genetically homogeneous and could be clearly

distinguished from non-human isolates (ie *T.b.brucei*) by numerical taxonomy (Hide *et al.,* 1994) using RFLP data. Furthermore, the *T.b.rhodesiense* stocks were found to be frequent in cattle and were closely related to isolates from previous epidemics since the 1960s in that region. Combining molecular and traditional epidemiological data it was shown that humans were five times more likely to acquire sleeping sickness via cattle than from other humans (Hide *et al.,* 1996), thus demonstrating the importance of cattle as a reservoir for sleeping sickness during an epidemic (Hide, 1999). Comparison of epidemic and endemic areas of sleeping sickness (Hide *et al.,* 1998) and areas free of sleeping sickness (Hide *et al.,* 2000) revealed the presence of the same human infective strains in patients, cattle and tsetse suggesting that the presence of a human infective strain was not the only factor determining whether an epidemic developed.

It was observed that *T.b.rhodesiense* isolates circulating in the Tororo region of Uganda were genetically homogeneous whilst the sympatric *T.b.brucei* isolates were much more diverse (Hide *et al.,* 1994). This raised the question as to the contribution of genetic exchange in "field" populations of trypanosomes. High levels of genetic exchange would facilitate the "spread" of phenotypes, such as human infectivity and drug resistance, potentially into new populations and could possibly explain the emergence of new epidemics. This collection of stocks offered an opportunity to investigate this question since they were isolated at the same place and time. However, as suitable DNA based techniques were not available, it was necessary to resort to MLEE analysis to show that genetic exchange was frequent in *T.b.brucei* but epidemic in *T.b.rhodesiense* (Hide *et al.,* 1994).

This suggested that genetic exchange could contribute to the genetic variation in trypanosome isolates in the field and importantly could contribute to the epidemiology of the disease. Two problems beset further analysis of these questions. Firstly, the inability to

genetically interpret multilocus RFLP markers and, secondly, despite DNA based methods providing higher levels of resolution, the requirement for sufficient pure parasite material for analysis.

PCR BASED MOLECULAR EPIDEMIOLOGICAL TOOLS

The advent of PCR to amplify and identify DNA at source or from small amounts of parasite material – such as a single trypanosome (Macleod *et al.,* 1997; Cox *et al.,* 2005) - offered tremendous advantages particularly in the area of diagnosis. In terms of epidemiological analysis, where multilocus analysis was clearly important, this was more difficult. Some single locus typing such as PCR-RFLP were found to be of limited use (Tilley & Hide, 2001). Random Amplification of Polymorphic DNA (RAPD) analysis was one of the initial solutions to PCR amplification of multiple loci (Mathieu-Daude *et al.,* 1995; Stevens *et al.,* 1995). This technique relied on the use of a short single PCR primer which randomly annealed to the DNA sample, a banding pattern was produced for each stock and numerical taxonomy used to assess the relatedness of stocks with each other. This was a good multilocus approach, useful with small amounts of DNA, which probably sampled the full extent of the genome rather than being restricted to certain genes or parts of the genome. However, this method is highly susceptible to contaminating host or other DNA – bands will be generated from any DNA template, it is not possible to interpret genetically and is unable to detect mixed infections of trypanosomes. A similar approach, Amplified Fragment Length Polymorphism (AFLP), uses PCR primers based on restriction enzyme sites to look for variation between trypanosome stocks (Agbo *et al.,* 2002) and has the same advantages and disadvantages of RAPD analysis. To overcome the problem of contaminating host DNA, a technique termed MGE-PCR (Mobile Genetic Element PCR) was developed utilising the positional variation of

mobile genetic elements to detect genetic variation between parasite isolates (Hide & Tilley, 2001; Terry *et al.,* 2001; Tilley *et al.,* 2003). Using a single primer designed to be specific for the repeat regions of the RIME mobile genetic element of *T.brucei*, amplification of bands between adjacent RIME elements generates a banding pattern which varies from stock to stock based on differences in the positions of RIME elements and so can be used to identify groups of similar isolates by numerical taxonomy. This approach has the advantage that it is PCR-based, specific to *T.brucei* DNA, multilocus, representative of a significant portion of the genome and generates a single "fingerprint" for each isolate or stock. The disadvantages are that it cannot detect mixed infections and it cannot be interpreted genetically.

The markers that offer the greatest prospects are the mini- and micro-satellite markers (eg Macleod *et al.,* 1997; Biteau *et al.,* 2000). These short repeated sequences vary in length or sequence between trypanosome strains and can be amplified and analysed by population genetic methods (MacLeod *et al.,* 2001a). Such markers can be used to detect mixed infections (MacLeod *et al.,* 1999; Koffi *et al*., 2007), are highly specific and sensitive tools for detection of *T.brucei ssp* and can be used without the need to amplify parasites in rodents (Morrison *et al.,* 2007). The development of PCR based methods for detailed analysis of trypanosome stocks has opened up many avenues of research. In the remainder of this review we will focus on the contributions of these techniques to our understanding of the diversity of trypanosome isolates in the field, the role of genetic exchange in the epidemiology of sleeping sickness, the nature of human infective strains and the generation of sleeping sickness epidemics.

Many studies have examined the diversity of *T.brucei* stocks obtained from field isolates. What is clear is that the initial biochemical and molecular studies were correct in suggesting substructuring within each of the subspecies *T.b.gambiense* and *T.b.rhodesiense*. Analyses

of stocks of *T.b.gambiense* Type 1 clearly show this group to be genetically highly homogeneous within a focus and distinguishable from *T.b.brucei* or *T.b.rhodesiense* by RAPD (Jamonneau *et al.,* 2002) microsatellite analysis (Biteau *et al.,* 2000; Jamonneau *et al.,* 2002), minisatellite analysis (MacLeod *et al.,* 2001a; MacLeod 2001b; MacLeod 2001c), MGE-PCR (Simo *et al.,* 2005) and AFLP (Simo *et al.,* 2008). In most cases, Type 1 *T.b.gambiense* has been shown to be genetically homogeneous despite differing symptoms in humans being associated with identical parasite genotypes (Jamonneau *et al.,* 2002). However, more detailed studies have revealed significant differences between isolates from different foci within the Type 1 group (Truc *et al.,* 2002; Morrison *et al.,* 2008; Koffi *et al.,* 2009). Interestingly, a new group of *T.b.gambiense* was identified by micro- and mini-satellites that was associated with asymptomatic disease in humans (Jamonneau *et al.,* 2004a). It is clear from the microsatellite studies that *T.b.gambiense* Type 1 is clonal and very limited, if any, genetic exchange occurs in the field (Morrison *et al.,* 2008; Koffi *et al.,* 2009).

The clonal nature of *T.b.gambiense* has allowed it to be clearly identified and shown to be zoonotic. Early isoenzyme studies demonstrated its presence in dogs, pigs, bovines and wild game (Mehlitz *et al.,* 1982) suggesting a possible animal reservoir. Recent research using PCR based tools have shown that pigs (Njiokou *et al.,* 2006; Simo *et al.,* 2006) and wild fauna (Njiokou *et al.,* 2006) have the potential to be animal reservoirs. Although in another study, using microsatellite analysis, pigs were not considered to be an active reservoir (Jamoneau *et al.,* 2004b). It is also possible that untreated, parasitologically unconfirmed, seropositive individuals could also act as a human reservoir (Garcia *et al*., 2006; Checchi *et al*., 2008). In a clonal pathogen, such as *T.b.gambiense*, the lack of genetic exchange is

helpful since it ensures a reasonable level of stability and reproducibility in genetic markers for use in epidemiological studies.

As with *T.b.gambiense*, minisatellite analysis has been used to investigate whether *T.b.rhodesiense* is clonal (MacLeod *et al.,* 2000). When the sympatric stocks from Tororo, Uganda (Hide *et al.,* 1994), were examined *T.b.rhodesiense* was found to be clonal suggesting that little or no genetic exchange is occurring in this population (MacLeod *et al*., 2000). This is in contrast with earlier MLEE studies, using the same stocks, that suggested it was epidemic (Hide *et al.,* 1994) – that is, genetic exchange masked by the amplification of a small number of genotypes. This discrepancy might be explained by the more detailed and sophisticated analyses available using minisatellites.

However, as predicted by the earlier MLEE and RFLP studies, genetic differences between different geographical populations can be clearly detected by minisatellite DNA analysis such as between Uganda and Zambia (MacLeod *et al.,* 2001b) and using polymorphism in the SRA (Serum Resistant Associated) gene between populations in Uganda and Malawi or Zambia (Gibson *et al.,* 2002; McLean *et al.,* 2004).

The SRA gene (De Greef *et al.,* 1989) has been shown to be capable of transforming a non human-infective *T.brucei* to a human-infective phenotype (Xong *et al.,* 1998). This gene is present in all human-infective field samples from East Africa and absent in animal isolates and has consequently been used as a diagnostic tool for *T.b.rhodesiense* (Welburn *et al.,* 2001; Gibson *et al.,* 2002). Although, this gene looks promising as the candidate gene determining human infectivity in *T.brucei*, it is clearly absent from both *T.b.gambiense* Type 1 and 2 (Turner *et al.,* 2004) showing that there is more than one mechanism that determines human infectivity. There is growing evidence to suggest that in at least some epidemics, such

as the 1998 epidemic in Soroti in Uganda, that the spread of trypanosomes expressing the SRA gene may be responsible for the generation of epidemics (Welburn *et al.,* 2001).

MOLECULAR EPIDEMIOLOGY: THE ORIGINS AND DYNAMICS OF *TRYPANOSOMA BRUCEI RHODESIENSE* SLEEPING SICKNESS EPIDEMICS

The 1988 Tororo sleeping sickness epidemic in Uganda resulted in a legacy of trypanosome stocks isolated during an epidemic (Hide *et al.,* 1994). In addition to analysis by RFLP, they have been analysed using minisatellite markers (MacLeod *et al.,* 2000), MGE-PCR (Tilley *et al.,* 2003) and the SRA gene (Welburn *et al.,* 2001). A clear distinction has emerged between the human infective stocks and non human-infective stocks in this region. The finding that cattle are an important reservoir for human infective stocks (Hide *et al.,* 1996) has been confirmed by more recent studies (Welburn *et al*., 2001). Isolates collected from humans ten years on from this epidemic have been shown by MGE-PCR to contain some identical genotypes thus confirming the stability of genotypes within this focus (Tilley *et al.,* 2003). A new epidemic started in December 1998 in Soroti, Uganda (Fevre *et al.,* 2001), and is currently ongoing and spreading north and west (Fevre *et al.,* 2005). Based on the evidence that cattle are an important reservoir for sleeping sickness, it has been postulated that cattle infected with human-infective *T.brucei* were imported from the Tororo District into Soroti (Fevre *et al.,* 2001). These cattle were then fed upon by local tsetse, passed onto the local people and the epidemic started. Statistical analysis of the location of sleeping sickness cases in relation to the Soroti cattle market (Fevre *et al.,* 2001) and the presence of SRA positive *T.brucei* isolated from the cattle at this market has supported this view (Welburn *et al.,* 2001). Other studies conducted on *T.brucei* isolates collected from Soroti suggest that they have a different clinical disease profile (eg more chronic presentation) to isolates

collected in Tororo and also the corresponding microsatellite profiles were found to be different (McLean *et al.,* 2007). This supports the view that, although some of the isolates have probably originated in Tororo, there may be other origins of these *T.brucei* isolates. These results clearly show that the importation of cattle from a high risk area may have been responsible for the generation of this new epidemic. It also shows the importance of tracking cattle movements particularly from high risk areas, such as Tororo, and the identification of the geographical locations of high risk areas.

Based on the extensive molecular epidemiological data from Uganda, it is possible to formulate a model of how sleeping sickness foci are generated and how epidemics develop. It is clear from the data that foci, such as the Busoga focus, must have the presence of a human infective *T.brucei* strain ("*T.b.rhodesiense"*). Molecular epidemiological studies suggest that this may be a single predominant clonal strain which remains stable over time (although multiple clonal strains could also be present). New epidemics can be generated by the influx of a human strain into an area previously occupied by only animal strains. This could be by the importation of a reservoir species, such as cattle, infected with a human strain. Local socio-political customs or events may determine the amplitude and duration of the epidemic. Despite having a wide antigenic repertoire, clonality will ensure that any given *T.b.rhodesiense* strain will have reduced genetic variability. Consequently, with time, a local human population will become immune/less susceptible or subjected to health interventions causing the epidemic to decline. On entering a naïve population, such as Soroti, clonal human strains would spread quickly. It is clear that different foci appear to have genetically different human strains (*T.b.rhodesiense* or *T.b.gambiense*) which could be generated either by the SRA gene being introduced into a new genetic background or by the introduction of one of the other mechanisms of human infectivity emerging in a previously non human-

infective strain. This mechanism could involve genetic exchange between a non human- and a human-infective strain. With time and selection through humans, emergent human strains might become established and generate new foci. The data suggest that in the field genetic exchange occurs but is masked by the epidemic spread of a small number of strains. The relatively small numbers of human infective strains, the relatively small number of foci and the long term stability of foci might suggest that there is a minimal effect of genetic exchange in generating new human strains and foci. However, in practice, the generation of even a single new focus creates a significant health issue. Important questions arise from this model. Are there low frequencies of genetic exchange generating new human-infective strains or is the generation of new human-infective strains by genetic exchange frequent but fixation of those strains within the population (to produce a focus) infrequent? Also, given the relatively limited data from foci in Zambia, Tanzania, Malawi and Mozambique, are the conclusions from the Ugandan studies typical for *T.b.rhodesiense*? As can be seen from this review, the molecular tools are now available to address such questions.

As with any model, this one generates more questions than it answers. We have come along way to understanding the generation of sleeping sickness epidemics but further questions need to be addressed. The future of molecular epidemiology will be based on our increasing ability to be able to conduct large scale epidemiological studies which will involve carefully controlled sampling strategies, application of high throughput whole genome technologies for both host and parasite, development of Single Nucleotide Polymorphism (SNP) chips, development of bioinformatics tools and integration with mathematical models of disease transmission. These approaches and technologies provide opportunities for the future and offer the prospect of addressing some of the questions raised in this review.

ACKNOWLEDGEMENTS

We would like to acknowledge all of our colleagues who have contributed to the studies we have reviewed here. While we have tried to provide appropriate coverage in this review, we acknowledge that a great deal of important work could not be included here for reasons of space.

REFERENCES

AGBO, E.E.C., MAJIWA, P.A.O., CLAASSEN, H.J. and TE PAS, M.F.W. (2002) Molecular variation of *Trypanosoma brucei* subspecies as revealed by AFLP fingerprinting. *Parasitology* **124**, 249-358.

BITEAU, N., BRINGAUD, F., GIBSON, W.C., TRUC, P. and BALTZ, T. (2000) Characterization of *Trypanozoon* isolates using a repeated coding sequence and microsatellite markers
Molecular and Biochemical Parasitology **105**,185-201.

BORST, P., FASE-FOWLER, F., and GIBSON W.C. (1981). Quantitation of genetic differences between *Trypanosoma brucei gambiense*, *rhodesiense* and *brucei* by restriction enzyme analysis of kinetoplast DNA. *Molecular and Biochemical Parasitology* **3**:117-131.
.

CARTER, R. and WALLIKER, D. (1975). New observations on the malaria parasites of rodents of the Central African Republic - *Plasmodium vinckei petteri* subsp. nov. and *Plasmodium chabaudi* Landau, 1965. *Annals of Tropical Medicine and Parasitology*. **69**,187-96.

CHECCHI, F., FILIPE, J.A., BARRETT, M.P. and CHANDRAMOHAN, D. (2008). The natural progression of gambiense sleeping sickness: what is the evidence? *PLoS Neglected Tropical Diseases.* **2** 303.

COX, A., TILLEY, A., MCODIMBA, F., FYFE, J., HIDE, G. and WELBURN, S.C. (2005). A PCR based assay for detection and differentiation of African trypanosome species in blood. *Experimental Parasitology,* **111**, 24-29.

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*. *Molecular and Biochemical Parasitology*, **36**,169-176.

FÈVRE, E.M., COLEMAN, P.G., ODIIT, M.D., MAGONA, J., WELBURN, S.C. and WOOLHOUSE, M.E.J. (2001). The origins of a new sleeping sickness outbreak (caused by *Trypanosoma brucei* infection) in eastern Uganda. *The Lancet* **358**, 625-628.

FÈVRE, 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. *The Lancet*. **366**, 745-747.

GARCIA, A., COURTIN, D., SOLANO, P., KOFFI, M. and JAMONNEAU, V. (2006). Human African trypanosomiasis: connecting parasite and host genetics. *Trends in Parasitology* **22**, 405-409.

GIBSON, W.C. (1986). Will the real *Trypansoma brucei gambiense* please stand up? *Parasitology Today* **2**, 255-257.

GIBSON, W.C. (1989). Analysis of a genetic cross between *Trypanosoma brucei rhodesiense* and *T. b. brucei*. *Parasitology*, **99**, 391-402.

GIBSON, W.C. (2009) Species-specific probes for the identification of the African tsetse-transmitted trypanosomes. *Parasitology*. In Press (current volume).

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. *Tropenmedezin und Parasitologie*, **29**, 335-45

GIBSON, W.C., MARSHALL, T.F. DE C. and GODFREY, D.G. (1980) Numerical analysis of enzyme polymorphism. A new approach to the epidemiology and taxonomy of trypanosomes of the genus *Trypanozoon*. *Advances in Parasitology* **18**, 175-245.

GIBSON, W., BORST, P. and FASE-FOWLER, F. (1985). Further analysis of intraspecific variation in Trypanosoma brucei using restriction site polymorphisms in the maxi-circle of kinetoplast DNA. *Molecular and Biochemical Parasitology*, **15,** 21-36

GIBSON, W.C., BACKHOUSE, T. and GRIFFITHS, A. (2002) The human serum resistance associated gene is ubiquitous and conserved in *Trypanosoma brucei rhodesiense* throughout East Africa. *Infection, Genetics and Evolution* **1**, 207–214.

GIBSON, W.C. (2007) Resolution of the species problem in African trypanosomes. *International Journal for Parasitology*, **37**, 829 - 838.

GODFREY, D.G. (1978). Identification of economically important parasites. *Nature*, 273, 600-604.

HIDE, G. (1999) The history of sleeping sickness in East Africa. *Clinical Microbiological Reviews* **12**, 112-125.

HIDE, G., CATTAND, P., LE RAY, D., BARRY, J.D. and TAIT, A. (1990). The identification of *T. brucei* subspecies using repetitive DNA sequences. *Molecular and Biochemical Parasitology*, **39,** 213-226.

HIDE, G. and TAIT, A. (1991) The molecular epidemiology of parasites. *Experientia* **47**, 128-142.

HIDE, G., BUCHANAN, N., WELBURN, S.C., MAUDLIN, I., BARRY, J.D. and TAIT, A. (1991). *Trypanosoma brucei rhodesiense*: Characterization of stocks from Zambia, Kenya and Uganda using repetitive DNA probes. *Experimental Parasitology,* **72**, 430-439.

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 and non human trypanosomes. *Parasitology*,**109**, 95-111.

HIDE, G., TAIT, A., MAUDLIN, I. and WELBURN, S.C. (1996). The origins, dynamics and generation of *Trypanosoma brucei rhodesiense* epidemics in East Africa. *Parasitology Today***,** **12**, 50-55.

HIDE G., ANGUS, S., HOLMES, P.H., MAUDLIN, I. and WELBURN, S.C. (1998). Comparison of *Trypanosoma brucei* strains circulating in an endemic and an epidemic sleeping sickness focus. *Experimental Parasitology*. **89**, 21-29.

HIDE, G., TILLEY, A., WELBURN, S., MAUDLIN, I. and TAIT, A. (2000). *Trypanosoma brucei*: identification of trypanosomes with genotypic similarity to human infective isolates in tsetse isolated from a region free of human sleeping sickness. *Experimental Parasitology*. **96**, 67 – 74.

HIDE, G. and TILLEY, A. (2001) Use of mobile genetic elements as tools for molecular epidemiology. *International Journal of Parasitology,* **31,** 599-602.

JAMONNEAU, V., GARCIA, A., RAVEL, S., CUNY, G., OURY, B., SOLANO, P., N'GUESSAN, P., N'DRI, L., SANON, R., FRÉZIL, J.L. and TRUC, P. (2002). Genetic characterization of *Trypanosoma brucei gambiense* and clinical evolution of human African trypanosomiasis in Côte d'Ivoire. *Tropical Medicine and International Health,* **7,** 610-621.

JAMONNEAU, V., RAVEL, S., GARCIA, A., KOFFI, M., TRUC, P., LAVEISSIÈRE, C., HERDER, S., GRÉBAUT, P., CUNY, G. and SOLANO, P. (2004a). Characterization of *Trypanosoma brucei s.l.* infecting asymptomatic sleeping-sickness patients in Côte d'Ivoire: a new genetic group? *Annals of Tropical Medicine and Parasitology*, **98,** 329-37.

JAMONNEAU, V., RAVEL, S., KOFFI, M., KABA, D., ZEZE, D.G., NDRI, L., SANE, B., COULIBALY, B., CUNY, G. and SOLANO, P, (2004b). Mixed infections of trypanosomes in tsetse and pigs and their epidemiological significance in a sleeping sickness focus of Côte d'Ivoire. *Parasitology*. **129**, 693-702.

JENNI, L., MARTI, S., SCHWEIZER, J., BETSCHART, B., LE PAGE, R.W.F., WELLS, J.M., TAIT, A., PAINDAVOINE, P., PAYS, E. and STEINERT, M. (1986). Hybrid formation between African trypanosomes during cyclical transmission. *Nature* **322**, 173-175.

KILGOUR, V. and GODFREY, D.G. (1973). Isoenzymes of alanine aminotransferase as possible specific characters of trypanosomes. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **67**, 11.

KOFFI, M., SOLANO, P., BARNABÉ, C., DE MEEÛS, T., BUCHETON, B., CUNY. G. and JAMONNEAU, V. (2007). Genetic characterisation of Trypanosoma brucei s.l. using microsatellite typing: new perspectives for the molecular epidemiology of human African trypanosomiasis. *Infection Genetics and Evolution*. **7**, 675-84.

KOFFI. M., DE MEEÛS, 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, USA.* **106**, 209-214.

MACLEAN, L., CHISI, J.E., ODIIT, M., GIBSON, W.C., FERRIS, V., PICOZZI, K. and STERNBERG, J.M. (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**, 7040-7044.

MACLEAN, L., ODIIT, M., MACLEOD, A., MORRISON, L., SWEENEY, L., COOPER, A., KENNEDY, P.G. and STERNBERG, J.M. (2007). Spatially and genetically distinct African trypanosome virulence variants defined by host interferon-gamma response. *Journal of Infectious Diseases*, **196**,1620-1628.

MACLEOD, A., TURNER, C.M.R., and TAIT, A. (1997). Detection of single copy gene sequences from single trypanosomes. *Molecular and Biochemical Parasitology*, **84**, 267-270.

MACLEOD, A., TURNER, C.M.R. and TAIT, A. (1999). A high level of mixed *Trypanosoma brucei* infections in tsetse flies detected by three hypervariable minisatellites. *Molecular and Biochemical Parasitology,* **102**, 237-248.

MACLEOD, A., TWEEDIE, A., WELBURN, S.C., MAUDLIN, I., TURNER, M.J. and TAIT, A. (2000). Minisatellite marker analysis of *Trypanosoma brucei*: reconciliation of clonal, panmictic and epidemic population genetic theories. *Proceedings of the National Academy of Sciences* **97**, 13442-13447

MACLEOD, A., TURNER, C.M.R. and TAIT, A. (2001a). The detection of geographical substructuring of *Trypanosoma brucei* populations by the analysis of minisatellite polymorphisms. *Parasitology* **123**, 475-482.

MACLEOD, A., TAIT, A. and TURNER, C.M. (2001b). The population genetics of *Trypanosoma brucei* and the origin of human infectivity. *Philosophical Transactions of the Royal Society of London B,* **356**, 1035-1044.

MACLEOD, A., WELBURN, S., MAUDLIN, I., TURNER, C.M. and TAIT, A. (2001c). Evidence for multiple origins of human infectivity in *Trypanosoma brucei* revealed by minisatellite variant repeat mapping. *Journal of Molecular Evolution* **52**, 290-301.

MATHIEU-DAUDÉ, F. and TIBAYRENC, M. (1994). Isozyme variability of Trypanosoma brucei s.l.: genetic, taxonomic, and epidemiological significance. *Experimental Parasitology*, **78**, 1-19.

MATHIEU-DAUDE, F., STEVENS, J., WELSH, J., TIBAYRENC, M. and MCLELLAND, M. (1995). Genetic diversity and population structure of *Trypanosoma brucei*: clonality versus sexuality. *Molecular and Biochemical Parasitology* **72,** 89-101.

MAYNARD SMITH, J., SMITH, N.H., O'ROURKE, M. and SPRATT, B.G. (1993). How clonal are bacteria? *Proceedings of the National Academy of Sciences* **90,** 4384-4388.

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. *Tropenmedezin und Parasitologie*, **33**, 113-118.

MILES, M.A., TOYE, P.J., OSWALD, SC. and GODFREY, D.G. (1977). The identification by isoenzyme patterns of two distinct strain-groups of *Trypanosoma cruzi,* circulating independently in a rural area of Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **71**, 217-225.

MORRISON, L.J., MCCORMACK, G., SWEENEY, L., LIKEUFACK, A.C., TRUC, P., TURNER, C.M., TAIT, A. and MACLEOD, A. (2007). Use of multiple displacement amplification to increase the detection and genotyping of trypanosoma species samples immobilized on FTA filters. *American Journal of Tropical Medicine and Hygiene,* **76,**1132-1137.

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,** 847-54.

NJIOKOU, F., LAVEISSIÈRE, C., SIMO, G., NKININ, S., GRÉBAUT, 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,** 147-53.

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**, 31-50.

PAINDAVOINE, P., ZAMPETTI-BOSSELER, 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**, 61-71.

SIMO, G., HERDER, S., NJIOKOU, F., ASONGANYI, T., TILLEY, A. and CUNY, G. (2005). *Trypanosoma brucei s.l.*: characterisation of stocks from Central Africa by PCR analysis of mobile genetic elements. *Experimental Parasitology*, **110**, 353-62.

SIMO, G., ASONGANYI, T., NKININ, S.W., NJIOKOU, F. and HERDER, S. (2006). High prevalence of *Trypanosoma brucei gambiense* group 1 in pigs from the Fontem sleeping sickness focus in Cameroon. *Veterinary Parasitology,* **139**, 57-66.

SIMO, G., CUNY, G., DEMONCHY, R. AND HERDER, S. (2008). *Trypanosoma brucei gambiense*: study of population genetic structure of Central African stocks using amplified fragment length polymorphism (AFLP). *Experimental Parasitology*, **118**, 172-80.

STEVENS, J.R. and TIBAYRENC, M. (1995). Detection of linkage disequilibrium in *Trypanosoma brucei* isolated from tsetse flies and characterized by RAPD analysis and isoenzymes. *Parasitology*, **110**, 181-186.

TAIT, A. (1970a). Genetics of NADP isocitrate dehydrogenase in *Paramecium aurelia. Nature*, **225**, 181-182.

TAIT, A. (1970b). Enzyme variation between syngens in *Paramecium aurelia*. *Biochemical Genetics*, **4**, 461-70.

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

TAIT, A., BABIKER, E.A. and LE RAY, D. (1984). Enzyme variation in *Trypanosoma brucei* ssp. I. Evidence for the subspeciation of *Trypanosoma brucei gambiense*. *Parasitology* **89**, 311-326.

TAIT, A., BARRY, J.D., WINK, R., SANDERSON, A. and CROWE, J.S. (1985). Enzyme variation in *Trypanosoma brucei* ssp. II. Evidence for *T. b. rhodesiense* being a subset of variants of *T. b. brucei*. *Parasitology*. **90**, 89-100.

TAIT, A. and TURNER, C.M.R. (1990). Genetic exchange in *Trypanosoma brucei*. *Parasitology Today* **6**, 70-75.

TERRY R.S, SMITH, J.E., DUNCANSON, P. and HIDE, G. (2001). MGE-PCR: a novel approach to the analysis of *Toxoplasma gondii* strain differentiation using mobile genetic elements. *International Journal of Parasitology*, **31**, 155-161.

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 taxonomic consequences. *Proceedings of the National Academy of Sciences*, **87**, 2414-2418.

TILLEY, A. and HIDE, G. (2001). Characterization of *Trypanosoma brucei* stocks by Intergenic Region Typing. *Annals of Tropical Medicine and Parasitology*, **95**, 617-621.

TILLEY, A, WELBURN, S.C., FEVRE, E., FEIL, E.J. and HIDE, G. (2003). *Trypanosoma brucei:* Trypanosome strain typing using PCR analysis of mobile genetic elements (MGE-PCR). *Experimental Parasitology,* **104**, 26 –32.

TRUC, P., AERTS, D., MCNAMARA, J. J., CLAES, Y., ALLINGHAM, R. LE RAY, D. and GODFREY, D. G. (1992). Direct isolation *in vitro* of *Typanosoma brucei* from man and other animals, and its potential value for the diagnosis of gambian trypanosomiasis. *Transactions of the Royol Society of Tropical Medicine and Hygiene.* **86**, 627-629.

TRUC, P., RAVEL, S., JAMONNEAU, V., N'GUESSAN, P. AND CUNY, G. (2002). Genetic variability within *Trypanosoma brucei gambiense*: evidence for the circulation of different

genotypes in human African trypanosomiasis patients in Côte d'Ivoire. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **96**, 52-55.

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**, 445-454

WALLIKER, D., CARTER, R. and MORGAN, S. (1971). Genetic recombination in malaria parasites. *Nature*, **232**, 561-562.

WALLIKER, D., CARTER, R. and MORGAN, S. (1973). Genetic recombination in *Plasmodium berghei*. *Parasitology*. **66**, 309-320.

WELBURN, S.C., PICOZZI, K., FÈVRE, E.M., COLEMAN, P.G., ODIIT, M., CARRINGTON, M. and MAUDLIN, I. (2001). Identification of human infective trypanosomes in animal reservoirs of sleeping sickness in Uganda by means of serum-resistance-associated (*SRA*) gene. *The Lancet*, **358**, 2017-2019.

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**, 839-346.

TABLE LEGENDS

Table 1. Characteristics of the subspecies of *Trypanosoma brucei*

Table 2 Molecular epidemiological methods for analysing African trypanosomes. The table shows some of the characteristics of each method: 1. Subspecies discrimination, ability to discriminate between the three subspecies of *T.brucei*. 2. Identification of populations or groups, the ability to define populations or groups within the subspecies. 3. Population genetics, the ability to be interpreted genetically to enable population genetics to be carried out. 4. Detection of mixed infections, able to detect mixed infections or not to be confounded by the presence of mixed infections. 5. PCR based amplification of DNA, can be used directly on blood or tissue samples without the need to amplify parasites in experimental animals or *in vitro*. 6. Requires pure parasites, parasites need to be amplified and purified. MLEE, Multilocus Enzyme Electrophoresis; RFLP, Restriction Fragment Length Polymorphism; PCR-RFLP, PCR amplified RFLP; RAPD, Randomly Amplified Polymorphic DNA; AFLP, Amplified Fragment Length Polymorphism; MGE-PCR, Mobile Genetic Element – PCR; Minisatellites and Microsatellites, analysis of length polymorphism in highly repeated sequences; SRA, detection of the Serum Resistance Associated gene which is found in human infective *T.b.rhodesiense*.

FIGURE LEGEND

Figure 1 A map of the Busoga focus of *T.b.rhodesiense* sleeping sickness in Uganda. Recent epidemics occurred in Tororo (1988 to date) and Soroti (1998 to date).