# *Trypanosoma brucei* s.l.: evolution, linkage and the clonality debate

J. R. STEVENS[1] *and* M. TIBAYRENC[2]

[1] *School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK*
[2] *UMR CNRS/ORSTOM 9926 : Génétique Moléculaire des Parasites et des Insectes Vecteurs, ORSTOM, BP 5045, 34032, Montpellier, France*

(*Received 6 September 1995; revised 13 December 1995; accepted 13 December 1995*)

SUMMARY

The Index of Association ($I_A$) has been proposed by Maynard Smith *et al.* (1993) as a general method for characterizing the population structures of microorganisms as either: clonal, epidemic, cryptic species or panmictic. With reference to the current debate surrounding the mode of reproduction in parasitic protozoa, this study explores (i) the suitability and limitations of the $I_A$ for characterizing populations of *Trypanosoma brucei* s.l., and (ii) the idea that the significance of genetic differences between populations may be better understood if the evolution, spread and temporal stability of certain parasite genotypes are also considered. Four populations of *T. brucei* from Côte d'Ivoire, Uganda and Zambia are analysed using the $I_A$ and a complementary test for linkage disequilibrium, test *f* of Tibayrenc, Kjellberg & Ayala (1990). The two populations from Uganda are characterized as epidemic, while the others appear more or less clonal; the merits of the two methods are compared. The implications of the various population classifications are discussed with reference to genotype longevity in each region; the evolutionary and biomedical consequences of the genetic non-homogeneity of *T. brucei* are reviewed.

Key words: *Trypanosoma brucei*, Index of Association, linkage disequilibrium, clonality.

## INTRODUCTION

Since the publication of the clonal theory by Tibayrenc, Kjellberg & Ayala in 1990 debate surrounding the mode of reproduction in natural populations of parasitic protozoa has intensified. Within the framework of this theory, discussion of the reproductive process of *Trypanosoma brucei* s.l., the agent of African sleeping sickness, which a number of studies (Tait, 1990; Tibayrenc *et al.* 1990, 1991; Maynard Smith *et al.* 1993) indicate to be intermediary on a scale of sexuality/clonality (Tibayrenc, 1995), has been especially intense. Until recently, methods available for broad-scale comparisons of levels of recombination were limited; however, methods developed by Tibayrenc *et al.* (1990) and more recently by Maynard Smith *et al.* (1993) provide indices suitable for comparing populations using data from biochemical characterization studies. This paper explores the idea that the significance of genetic differences observed between populations of *T. brucei* may be better understood if the evolution and recent historical spread of the parasite are also considered. Evolutionary studies to date appear to fall into either of two categories: early, largely descriptive works (e.g. Ormerod, 1961) and, more recently, numerical/phylogenetic studies (e.g. Godfrey *et al.* 1990). It is the advent of such numerical approaches which has enabled more rigorous, quantitative studies, less prone to observer bias.

In this study the index of association ($I_A$) of Maynard Smith *et al.* (1993) and a test of linkage disequilibrium (test *f*; Tibayrenc *et al.* 1990) are used to compare levels of recombination (defined by the degree of linkage detected with each statistical method) and to investigate the population structure of 4 populations of *T. brucei* from across Africa. Data sets were selected to include only samples originating from a single area and a discrete time period; data from Côte d'Ivoire were further subdivided as there is considerable evidence to suggest that *T. b. gambiense* is a genetically discrete entity (reviewed by Gibson, 1986). Such criteria were vital to ensure that isolates came from a population which all other things being equal *could* have been truly panmictic (Tibayrenc & Ayala, 1991).

## MATERIALS AND METHODS

### *Origins of data*

*Bouaflé, Côte d'Ivoire.* Data are taken from Mehlitz *et al.* (1982). The 73 trypanosome isolates were collected from various hosts (humans, pigs, cattle, antelopes) from villages and towns in central Côte d'Ivoire, an area bounded by the towns of Bouaflé, Daloa and Vavoua; collections were made at sites within a 60 km radius during September and November 1978. Sleeping sickness is endemic to the region, primarily in rural areas (Mehlitz, Brinkmann & Haller, 1981). Isoenzyme characterization was

 

performed with 7 enzyme systems. Stocks possessing ALAT pattern 1 were classified as *T. b. gambiense* (10%); all these stocks, from humans and a pig, were highly resistant to human serum. All other stocks had ALAT pattern 2 and were classified as human infective non-*gambiense* (11%) or, if animal derived, *T. b. brucei* (79%) (Gibson, Marshall & Godfrey, 1980); human serum resistance test results for these stocks were highly variable (Mehlitz *et al.* 1982). Isolates collected by Mehlitz *et al.* (1982) from Upper Volta (now Burkina Faso) were not included in the analysis due to their spatial separation from the Ivoirian isolates.

*Luangwa Valley, Zambia.* Data are taken from Godfrey *et al.* (1990), with additional isoenzyme characterization by Stevens *et al.* (1992). The 67 isolates were collected from the Luangwa Valley in eastern Zambia over a 5-year period from 1980 to 1984; 72% of stocks were from humans, with a further 7% being isolated from animals and 21% from tsetse. The valley is approximately 400 km long and includes several wildlife reserves. Sleeping sickness is endemic to the area and the somewhat chronic nature of the disease is well documented (Buyst, 1974; Godfrey *et al.* 1990).

Characterization was performed with 9 enzyme systems interpreted as 10 putative loci (Stevens *et al.* 1992); all except 5 stocks were characterized as *T. b. rhodesiense* (Stevens & Godfrey, 1992), with the majority (75%) being classified as strain group *zambezi* (Gibson *et al.* 1980). The 5 unclassified stocks included 3 isolates from humans and were, by numerical analysis, close to *T. b. rhodesiense* of one form or another (Stevens & Godfrey, 1992).

*Busoga Region (including Tororo District), Uganda.* Data are taken from Enyaru *et al.* (1993). The 142 isolates analysed in this study were collected from humans, domestic animals and tsetse in the districts of Busoga and Tororo in southeast Uganda over a 4-year period from 1988 to 1991; the area borders Lake Victoria and is approximately 110 km in diameter. Sleeping sickness is endemic to the area, although the disease is characterized by epidemic outbreaks, the most recent of which began in 1976 (Abaru, 1985; Mbulamberi, 1989). Isoenzyme characterization was performed with 8 enzyme systems; 97% of stocks were classified as *T. b. rhodesiense* (strain group *busoga* 51%, *zambezi* 42%, *kiboko* 3·5%), 3% of stocks were classified as intermediate between *T. b. rhodesiense* and *T. b. brucei.*

*Tororo District, Uganda.* Data are taken from Maudlin *et al.* (1990). The 44 isolates from Tororo District were from humans and domestic animals; they are independent of those included in the Busoga analysis, but approximate to a more localized subset for comparative purposes. All samples were collected in November 1988, from villages within a 25 km radius and from patients attending the nearby sleeping sickness hospital in Tororo town. The epidemiological and historical characteristics of the area are as described above (Busoga Region). Isoenzyme characterization was performed with 9 enzyme systems interpreted as 10 putative loci (Stevens & Welburn, 1993); all stocks were classified as *T. b. rhodesiense*, strain groups *busoga* (84%) or *zambezi* (16%) (Stevens *et al.* 1992).

### *Isoenzymes*

Isoenzyme characterization for the data sets included in this study was undertaken by, or in cooperation with, the MRC Trypanosomiasis Group at the Tsetse Research Laboratory, Bristol. This ensured that the techniques used and the levels of variation identified are comparable between studies.

### *Linkage disequilibrium statistics*

The index of association ($I_A$) employed is that presented by Maynard Smith *et al.* (1993) and as discussed by those authors, primarily due to the difficulty of applying an accurate allelic interpretation to many of the published isoenzyme patterns included, each diploid genotype at a locus was treated as an 'allele' in the current study. The standard error of $I_A$ is calculated for the null hypothesis (J. Maynard Smith, personal communication), i.e. assuming no barriers to panmixia. Any departure (values of $I_A > 0$) thus represents the departure due to linkage disequilibrium (LD). The benefits of the $I_A$ value derive primarily from the simple, but statistically measurable, definition of the existence, or otherwise, of LD within a given data set. Beyond this, it is the use of parallel analyses, i.e. the degree of LD detected when calculating $I_A$ for all individuals compared to the value for $I_A$ obtained when analysing only the electrophoretic types (ETs; in this case zymodemes) after removing the over-represented clones, which allows the underlying population structure to be defined. According to Maynard Smith *et al.* (1993) 4 possible types of population structure can be considered. (i) 'True' clonality: organisms undergo clonal evolution, with the clones being stable in space and time, even on an evolutionary scale. Examples given by Maynard Smith *et al.* (1993): bacteria – *Salmonella* sp.; parasitic protozoa – *T. cruzi.* (ii) Epidemic clonality: organisms undergo occasional bouts of clonal propagation, while the species is basically sexual. Based largely on empirical observations on the life-time of *Neisseria meningitidis* ETs, Maynard Smith (personal communication) suggested that the clones generated in this way are ephemeral and their life-time is at best a few years. Examples given by Maynard Smith *et al.* (1993): bacteria – *N. meningitidis*; parasitic

Table 1. Index of Association test for linkage disequilibrium (Maynard Smith *et al.* 1993)

| | |
|---|---|
| Bouaflé, Côte d'Ivoire: 73 isolates, 32 zymodemes, 7 loci | |
| All isolates: | $I_A = 0{\cdot}53 \pm 0{\cdot}0029$ |
| ETs: | $I_A = 0{\cdot}14 \pm 0{\cdot}0060$ |
| Luangwa Valley, Zambia: 67 isolates, 39 zymodemes, 10 loci | |
| All isolates: | $I_A = 1{\cdot}14 \pm 0{\cdot}0022$ |
| ETs: | $I_A = 0{\cdot}75 \pm 0{\cdot}0028$ |
| Tororo District, Uganda: 44 isolates, 12 zymodemes, 10 loci | |
| All isolates: | $I_A = 0{\cdot}80 \pm 0{\cdot}0087$ |
| ETs: | $I_A = 0{\cdot}01 \pm 0{\cdot}0206$ |
| Busoga Region (including Tororo District), Uganda: 142 isolates, 50 zymodemes, 8 loci | |
| All isolates: | $I_A = 0{\cdot}30 \pm 0{\cdot}0028$ |
| ETs: | $I_A = -0{\cdot}06 \pm 0{\cdot}0079$ |

protozoa – *T. brucei*. (iii) Cryptic speciation: the species is actually subdivided into 2 or more discrete phylogenetic lineages, each one being panmictic. Example given by Maynard Smith *et al.* (1993): bacteria – *Rhizobium meliloti*; no example given for parasitic protozoa. (iv) Panmixia: genetic exchange is frequent enough to prevent the propagation of any discrete genetic lineage and no detectable departures from random mating are observed. Example given by Maynard Smith *et al.* (1993): bacteria – *Neisseria gonorrhoeae*; parasitic protozoa – *Plasmodium falciparum*.

To distinguish 'true' clonality from epidemic clonality Maynard Smith *et al.* (1993) proposed linkage disequilibrium analyses of the genotypes (ETs) rather than the individuals (stocks). If the LD disappears, it is taken that the population is epidemic rather than clonal. Such an approach relies on the supposition that identical genotypes are the result of recent clonal propagation.

Test *f* of Tibayrenc *et al.* (1990), as for the $I_A$, provides a measure of linkage disequilibrium and can be used to analyse either all isolates or ETs only. The test (see Tibayrenc *et al.* (1990) for full details) is a Monte-Carlo based method, in which a probability value is assigned to the observed level of LD on the basis of the level of LD in random reassortments of the observed genotypes, sufficient random reassortments being performed (e.g. 10000) to allow construction of a frequency distribution with which to compare the observed LD. In this study test *f* was used to provide an alternative but complementary analysis of the 4 data sets, providing additional support for results obtained with both tests.

## RESULTS

Values for $I_A$ and associated standard errors are presented in Table 1. Data from Côte d'Ivoire and Zambia gave $I_A$ values significantly different from zero for the analyses of all isolates and the ETs only, indicating significant association within the data and suggesting a potentially clonal population structure for these trypanosomes. Subdivision of the Côte d'Ivoire data set provided further evidence of the clonal nature of *T. b. gambiense*, with all 7 *T. b. gambiense* isolates being classified in 2 zymodemes; $I_A$ analysis was not performed due to the small sample size. The non-*gambiense* Côte d'Ivoire stocks, both human and animal derived, were characterized as having an epidemic population structure. (All 66 isolates, $I_A = 0{\cdot}16 \pm 0{\cdot}07$; 30 ETs only, $I_A = -0{\cdot}04 \pm 0{\cdot}08$). The 2 Ugandan data sets gave $I_A$ values significantly different from zero for the analyses of all isolates, while analysis of ETs only did not evidence significant association. According to the proposals of Maynard Smith *et al.* (1993), this suggests a basically sexual population structure which is prone to significant epidemic bursts of clonal reproduction leading to temporary disequilibrium.

Comparative *f*-test statistics are presented in Table 2. Overall, the pattern of results and the levels of LD obtained with test *f* are in good agreement with those obtained using the $I_A$. The Zambian data set showed significant linkage disequilibrium at the 0·0001 level for the analyses of both all isolates and ETs only, suggesting significant departures from panmixia in this population of *T. brucei*. For the Côte d'Ivoire population as a whole, significant LD was detected in the analysis of all isolates, but not in the analysis of ETs only ($P = 0{\cdot}1555$) indicating an epidemic population structure; this non-significant result with test *f* did not accord with that of the $I_A$ analysis, the only comparison of linkage in this study not to do so. The Ugandan populations exhibited significant LD in the analyses of all isolates, but not in the analyses of ETs only, once again suggesting an epidemic population structure for these trypanosomes.

Summary statistics on the duration of genotypes in the regions from which the sample data sets derive are presented in Table 3. Zymodeme duration data are taken from Stevens (1991) based on the zymodeme definitions of Gibson *et al.* (1980) and Godfrey *et al.* (1990). The data set includes isolates of *T. brucei* s.l. collected and characterized by the MRC Trypanosomiasis Research Group and associates over a 31-year period from 1959 to 1989.

## DISCUSSION

Results from both analyses show that the population structure varies between populations of trypanosomes across the continent, such that it appears impossible to generalize results from single populations to *T. brucei* as a whole. Moreover, it appears possible that such differences may be connected with the evolution of African trypanosomes with respect

Table 2. Test for linkage disequilibrium (Test-*f*, Tibayrenc *et al.* 1990)

(All tests were performed at the 0·01 significance level, with 10000 random reassortments.)

| | |
|---|---|
| Bouaflé, Côte d'Ivoire | |
| ETs only: 32 genotypes, 7 loci | $P = 0{\cdot}1555$, i.e. non-significant |
| All isolates: 73 isolates (32 genotypes), 7 loci | $P < 0{\cdot}0001$, i.e. significant |
| Luangwa Valley, Zambia | |
| ETs only: 39 genotypes, 10 loci | $P < 0{\cdot}0001$, i.e. significant |
| All isolates: 67 isolates (39 genotypes), 10 loci | $P < 0{\cdot}0001$, i.e. significant |
| Tororo District, Uganda | |
| ETs only: 12 genotypes, 10 loci | $P = 0{\cdot}5204$, i.e. non-significant |
| All isolates: 44 isolates (12 genotypes), 10 loci | $P < 0{\cdot}0001$, i.e. significant |
| Busoga Region (including Tororo District), Uganda | |
| ETs only: 50 genotypes, 8 loci | $P = 0{\cdot}2377$, i.e. non-significant |
| All isolates: 142 isolates (50 genotypes), 8 loci | $P < 0{\cdot}0001$, i.e. significant |

Table 3. Zymodeme duration

(Data are from Stevens (1991), sub-specific groupings are defined primarily on the basis of zymodeme; see Godfrey *et al.* (1990) and Stevens & Godfrey (1992) for full discussion of definitions.)

| |
|---|
| *T. b. gambiense*: 6 zymodemes, 38 stocks |
| Mean zymodeme duration: 8·3 years |
| Standard error: 3·1 years |
| Maximum duration: 20 years |
| *T. brucei* s.l., strain group *zambezi*: 31 zymodemes, 73 stocks |
| Mean zymodeme duration: 4·1 years |
| Standard error: 1·4 years |
| Maximum duration: 29 years |
| *T. brucei* s.l., strain group *busoga*: 20 zymodemes, 56 stocks |
| Mean zymodeme duration: 2·4 years |
| Standard error: 1·5 years |
| Maximum duration: 23 years |

to (a) recent historically documented movements of trypanosomiasis, and (b) the long-term evolution of *Trypanozoon*.

A lack of markers suitable for parasite characterization and population studies meant that research into evolutionary aspects of African trypanosomiasis was practically impossible (Godfrey, 1984) until the early 1980s. However, due to the range of characterization studies now completed (e.g. Gibson *et al.* 1980; Tait, 1980; Tait, Babiker & Le Ray, 1984; Gibson & Wellde, 1985; Tait *et al.* 1985; Paindavoine *et al.* 1986; Godfrey *et al.* 1987, 1990; Hide *et al.* 1990, 1991, 1994; Stevens & Godfrey, 1992; Enyaru *et al.* 1993; Truc & Tibayrenc, 1993; Mathieu-Daudé *et al.* 1995) sufficient data are available to begin to explore both the patterns of movement and the evolution of *T. brucei* s.l.

Early studies (e.g. Mackichan, 1944; Ormerod, 1961) document the spread of *T. b. rhodesiense* from the Zambezi basin (Zambia and Zimbabwe) around 1908, northwards through modern-day Malawi and Tanzania, to Kenya and Uganda by 1942 (reviewed by Koerner, de Raadt & Maudlin, 1995). Interestingly, during this time the course of the disease supposedly also changed from being sporadic and chronic, as it usually occurred in the Zambezi basin, to being epidemic and acute in Kenya and Uganda. This theory, however, is based largely on the reports of early colonial medical staff and in the light of more recent biochemical studies, it seems probable that the apparent movement of the disease may have been due more to the movements of the early observers than to the trypanosomes or their vectors.

Notably, a study by Godfrey *et al.* (1990) which presented the first major cladistic analysis (Distance Wagner Method, Farris (1972); R. D. Baker, personal communication) of *T. brucei* trypanosomes reached a different conclusion from the early researchers, namely that the strain groups and subspecies of *Trypanozoon* observed today (Gibson *et al.* 1980; Tait *et al.* 1984, 1985; Godfrey *et al.* 1990) are the product of a radiative movement from an ancestral group – *busoga* – which is now, for the most part, confined to central East Africa. For *zambezi* strain group and *T. b. gambiense* stocks it is possible, therefore, that such predominantly man-infective parasites evolved from a type, such as *busoga*, which was (and still is) closely associated with an animal reservoir (Godfrey *et al.* 1990; Hide *et al.* 1994); the ancestral type may have originated around Lake Victoria, and then spread and adapted. Thus, it appears that the scenario described by Godfrey *et al.* (1990) provides the most parsimonious explanation of the distribution of *T. brucei* groupings seen today, and while the resolution of such a problem is not an absolute requirement of a taxonomic or epidemiological study, the consequences of one or other theory being correct will have profound implications for phylogenetic relationships between strain groups, particularly those described for East African *Trypanozoon*.

Methods of detecting linkage disequilibrium provide an extension to phylogenetic methods of studying evolution. In particular, the index of association (Maynard Smith *et al.* 1993) allows the

structure of parasite populations to be defined, in turn permitting comparison with information on their evolutionary position, and a number of such studies have now been completed. An analysis by Maynard Smith *et al.* (1993) of *T. brucei* from the Lambwe Valley, Kenya (Gibson & Wellde, 1985; Mihok, Otieno & Darji, 1990) indicated an epidemic population structure which, together with the 2 Ugandan data sets presented in the current paper, would confirm the epidemic status of the *T. brucei* trypanosomes from this central region of its range. More recently, a major study by Hide *et al.* (1994) also used the index of association to demonstrate an epidemic population structure for a group (defined by RFLPs) of man-infective trypanosomes isolated from southeast Uganda.

The epidemic structures defined for populations from Uganda and Kenya thus provide an indication of the underlying degree of genetic exchange occurring in the geographical centre of the range of *T. brucei* (Gibson, 1990, 1995). In contrast, the clonal population structures evidenced for populations originating away from this central region (i.e. from Côte d'Ivoire and Zambia) indicate a lack of, or significant barriers to, recombination in these areas. Again, the phylogenetic analysis of Godfrey *et al.* (1990) indicates that such populations, from the extreme geographic ends of the range of *T. brucei*, also correspond to evolutionary end-points.

Furthermore, despite there being no direct links between them, the populations classified in the current study as clonal (i.e. *T. b. gambiense* from Côte d'Ivoire and the population from Zambia) share several major characteristics: the chronic nature of the disease in the regions where these strains were isolated (Buyst, 1974; Godfrey *et al.* 1987), the extent of their geographical removal from the possible ancestral point in central East Africa, and their relatively large phylogenetic separation from the hypothetical ancestral type (Godfrey *et al.* 1990). Unfortunately, to date, and certainly for the parasites included in the data sets used in this study, no well-documented information relating the clinical nature of the disease caused by a particular isolate to its strain type is available. However, a range of circumstantial evidence indicates that trypanosomes from strain groups which, in this study have been shown to exhibit clonal population structures, appear to be associated with regions where more chronic and sometimes asymptomatic forms of sleeping sickness are reported e.g. Zambia, Ghana, Togo (reviewed by Godfrey *et al.* 1990).

In the current study 2 populations are classified by $I_A$ analysis as clonal,* while 2 appear to be epidemic. A similar pattern of results is also obtained with test *f* (Table 2) with one exception, the analysis of ETs only from Bouaflé, Côte d'Ivoire, which did not show significant LD. Interpreted directly, such a result indicates the Bouaflé population to be epidemic rather than clonal; however, while this may be the case, a number of factors remain to be considered for this data set. First, the inability of isoenzymes to detect fine levels of genetic variability (Stevens & Tibayrenc, 1995) may be important – 5 of the 12 enzymes used by Mehlitz *et al.* (1982) were completely invariant. The difficulties in evidencing linkage in such situations are well known (Tibayrenc, 1993) and have been described for a number of other parasitic protozoa (e.g. *Giardia* sp., Andrews *et al.* 1989; *Leishmania* sp., Guerrini, 1993). Secondly, after subdivision of the Bouaflé data into *T. b. gambiense* and non-*gambiense* (both human and animal derived), $I_A$ analysis indicates that while *T. b. gambiense* is probably clonal, non-*gambiense T. brucei* have an epidemic population structure. Thus, an important part of the linkage disequilibrium evidenced in the analysis of the undivided Bouaflé data set appears to have been due to the substructuring present in the combined data. Such a result illustrates the danger of inadvertently combining inappropriate data. It should also be noted that of the 3 populations which did not show significant LD with test *f*, the Bouaflé population was the nearest to being significant ($P = 0{\cdot}1555$), again emphasizing the importance of considering such results as part of a scale of linkage rather than as an absolute categorization of a given population. Such findings highlight further the differences in sensitivity between the $I_A$ and *f* tests used in this work and, more generally, the importance of the choice of method in any study of this type. In particular, the $I_A$, which in its basic form treats all genotypes as equally different, irrespective of the similarities between particular genotypes, appears less suitable to detect subtle differences in levels of relatedness between genotypes. Of course, in studies where it is possible to reliably interpret genotypes in terms of individual alleles an improved level of definition is to be expected (G. Hide, personal communication). Nonetheless, we propose that test *f* is somewhat less coarse as all the information *within* each genotype is considered in its calculation, free from assumptions connected with allelic interpretation; thus, we suggest, it is more able to take account of the richness of the data.

The Ugandan data sets analysed in this study were not subdivided, as isoenzyme characterization indicated them to be 100% *T. b. rhodesiense* (Tororo District) or more than 97% *T. b. rhodesiense* (Busoga Region). However, as the study by Hide *et al.* (1994) shows, distinct genetic subdivisions within *T. brucei* from this area can be evidenced using additional markers (e.g. RFLP, human serum resistance tests). While such additional characterization is not possible when re-analysing published data, subsequent

* See discussion regarding the subdivision of the Bouaflé data into *T. b. gambiense* and non-*gambiense*.

analyses should consider the significance of such subdivisions in trypanosome populations from this area of East Africa.

Having defined a population structure for a given data set, what are the merits of such a system of classification and to what do such definitions, particularly the term 'epidemic', relate? An epidemic population can be defined (Maynard Smith *et al.* 1993; Tibayrenc, 1995) as a population which exhibits explosive clonal bursts from an underlying sexual base, such bursts leading to temporary disequilibrium which nevertheless endure long enough to be of considerable biomedical interest. Furthermore, in addition to their immediate capacity to cause a disease epidemic, such populations, by virtue of their underlying sexual base, may harbour an immense reservoir of genetic diversity capable of throwing up a range of genotypes, the consequences of which, for example in terms of drug resistance, may be important.

The time-scale over which epidemics occur also has important implications for the parasite population genetics and the manner in which definitions are applied. From long-term data (summarized by Stevens, 1991) collected in the same regions as the data sets analysed using the LD measures it can be seen that there is considerable variation in the duration of zymodemes (genotypes) isolated from each of the regions; such information can provide a valuable insight into the long-term stability of genotypes in nature. In Zambia and Côte d'Ivoire the longest enduring zymodemes lasted 20 and 29 years, while the average duration of genotypes was 4 and 8 years, respectively. However, in the central region of Africa, while the longest surviving zymodeme was sampled over 23 years, the average duration was less than 2·5 years; thus, on average, genotypes in 'clonal' populations exist apparently unrecombined for at least twice as long as genotypes in 'epidemic' populations. The exceptions to this seen in Uganda (where 1 zymodeme was sampled over a period of 23 years) may highlight the existence of some genetically isolated clonal lineages even within a predominantly sexual population. Such genotypes may be able to endure due to the range of host cycles/niches available in a given habitat. For example, Hide *et al.* (1994) isolated a genetically distinct group of man-infective genotypes from the Tororo area of Uganda which appeared to have existed in East Africa unaltered for over 30 years. Additionally, the effect of sample size should also be considered. The *busoga* data set comprised of 56 stocks, while the *T. b. gambiense* sample included only 38 stocks. Primarily, however, smaller sample size would be expected to affect the detection of the more rare short-term genotypes rather than the more stable longer-term zymodemes (Cibulskis, 1992).

From a biomedical viewpoint, the consequences of the long-term stability of certain zymodemes and, in particular, the continued existence of certain genotypes associated with epidemic outbreaks of sleeping sickness remain unknown. Until such time as a population can be unequivocally defined as either clonal, sexual or otherwise, then the value of identifying genotypes (and zymodemes) remains considerable for its potential to aid medical, epidemiological and phylogenetic studies. Alternatively, populations conforming strictly to one or other definition may not exist in nature, instead the variation observed may best be viewed as part of a continuum of clonality/sexuality (Tibayrenc, 1995); in this case, population definitions should be regarded only as convenient working headings and, for example, even epidemic populations could vary between being more, or less, clonal/sexual. The importance of recognizing such a fact will depend on the objectives of individual studies.

This paper has attempted to provide an answer to the long-running debate as to whether the population structure of the species *T. brucei* is clonal or sexual. Obviously, the inflexibility of using existing data sets has numerous drawbacks; however, by taking a 'snapshot' of a number of populations it has been possible to demonstrate that for *T. brucei* both states can exist, perhaps even sympatrically. Such a conclusion, corroborated by the results from 2 independent numerical methods, highlights the fact that *T. brucei* populations are no longer freely mixing, but are subdivided across their range. Furthermore, it appears that the type of population structure observed may be associated with host and the biomedical properties of the population under study.

Thus, much of the past debate concerning the population structure of *T. brucei* may in part have been due to attempts to generalize conclusions for *T. brucei* as a whole using information gained from diverse individual population examples. However, as many characterization studies have now shown *T. brucei* is not a genetically homogeneous species (e.g. *T. b. gambiense* – Paindavoine *et al.* 1989; *T. b. rhodesiense* – Gibson *et al.* 1980; Hide *et al.* 1991) and the many varied populations and subspecies of *T. brucei* cannot be treated simply as one panmictic group. Indeed, several recent studies (Hide *et al.* 1994; Stevens *et al.* 1994) indicate that the underlying level of genetic variation in *T. brucei* may be considerably greater than previously recognized. Thus, while the change in LD observed between the two approaches to calculating linkage (*all isolates* and *ETs only*) with the $I_A$ and *f*-test statistics provides a measure of the evenness of spread of isolates within ETs, and therefore an indication of the importance of the over-representation of particular epidemic zymodemes which do not appear to be undergoing recombination, additional characterization of such zymodemes (e.g. using RAPDs) appears likely to reveal further genetic

variation. For example, in a recent study by these authors (Stevens & Tibayrenc, 1995) 15 stocks of *T. b. gambiense* were classified into 8 zymodemes using 17 loci; however, by RAPD analysis 13 RAPD ETs were identified in the same stocks using only 8 random primers. Moreover, while the limited genetic diversity identified by isoenzymes significantly reduced the power of LD analyses, the increased variation detected by RAPD analysis allowed LD tests to be adequately performed and significant linkage to be demonstrated. The importance of such a result on the findings of the current study is unknown, but it seems probable that a considerable amount of genetic diversity has remained undetected, corresponding to a reduced capability to detect linkage. Perhaps not surprisingly then, given the different ecological conditions (particularly fly species and number, and thus opportunities for transmission and recombination) in the regions from where they were isolated, differences between populations exist. In turn, such differences have allowed differential parasite evolution over time and, consequently, generalizations about the species become useless and may in fact become misleading as this study, one of the first to consider discrete data sets from a range of different sources, demonstrates.

We thank J. Maynard Smith for helpful discussions and information regarding the calculation of the standard error of $I_A$. For additional advice and information we thank R. D. Baker (cladistic analysis), G. Hide ($I_A$ analysis), S. Noel (*f*-tests) and I. Maudlin for useful discussions on some historical aspects of trypanosomiasis and for providing access to unpublished material. The financial support of the Ministère Français de l'Education Nationale, the EC ('Science' grant number SCI*CT92-0761) and a Royal Society Travel Grant (J.R.S.) to visit the Ugandan Trypanosomiasis Research Organization are gratefully acknowledged. J.R.S. is currently supported by the Wellcome Trust.

## REFERENCES

ABARU, D. E. (1985). Sleeping sickness in Busoga, Uganda, 1976–1983. *Tropenmedizin und Parasitologie* **36**, 72–6.

ANDREWS, R. H., ADAMS, M., BOREHAM, P. F. L., MAYRHOFER, G. & MELONI, B. P. (1989). *Giardia intestinalis*: electrophoretic evidence for a species complex. *International Journal for Parasitology* **19**, 183–90.

BUYST, H. (1974). The epidemiology, clinical features, treatment and history of sleeping sickness on the northern edge of the Luangwa fly belt. *Medical Journal of Zambia* **8**, 2–12.

CIBULSKIS, R. E. (1992). Genetic variation in *Trypanosoma brucei* and the epidemiology of sleeping sickness in the Lambwe Valley, Kenya. *Parasitology* **104**, 99–109.

ENYARU, J. C. K., STEVENS, J. R., ODIIT, M., OKUNA, N. M. & CARASCO, J. F. (1993). Isoenzyme comparison of *Trypanozoon* isolates from two sleeping sickness areas of south-eastern Uganda. *Acta Tropica* **55**, 97–115.

FARRIS, J. S. (1972). Estimating phylogenetic trees from distance matrices. *American Naturalist* **106**, 645–68.

GIBSON, W. C. (1986). Will the real *Trypanosoma b. gambiense* please stand up. *Parasitology Today* **2**, 255–7.

GIBSON, W. C. (1990). Trypanosome diversity in Lambwe Valley, Kenya – sex or selection? *Parasitology Today* **6**, 342–3.

GIBSON, W. C. (1995). The significance of genetic exchange in trypanosomes. *Parasitology Today* **11**, 465–8.

GIBSON, W. C., MARSHALL, T. F. DE C. & GODFREY, D. G. (1980). Numerical analysis of enzyme polymorphism: a new approach to the epidemiology of trypanosomes of the subgenus *Trypanozoon*. *Advances in Parasitology* **18**, 175–246.

GIBSON, W. C. & WELLDE, B. T. (1985). Characterization of *Trypanozoon* stocks from the South Nyanza sleeping sickness focus in Western Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **79**, 671–6.

GODFREY, D. G. (1984). Molecular biochemical characterization of human parasites. In *Recent Advances in Tropical Medicine*, vol. 1 (ed. Gilles, H. M.) pp. 289–319. Edinburgh: Churchill Livingstone.

GODFREY, D. G., BAKER, R. D., RICKMAN, L. R. & MEHLITZ, D. (1990). The distribution, relationships and identification of enzymic variants within the subgenus *Trypanozoon*. *Advances in Parasitology* **29**, 1–74.

GODFREY, D. G., SCOTT, C. M., GIBSON, W. C., MEHLITZ, D. & ZILLMANN, U. (1987). Enzyme polymorphism and the identity of *Trypanosoma brucei gambiense*. *Parasitology* **94**, 337–47.

GUERRINI, F. (1993). Génétique des populations et phylogénie des *Leishmania* du Nouveau-Monde. Thèse ès Sciences, Université des Sciences et Techniques du Languedoc, Montpellier II.

HIDE, G., BUCHANAN, N., WELBURN, S., MAUDLIN, I., BARRY, J. D. & TAIT, A. (1991). *Trypanosoma brucei rhodesiense*: Characterisation of stocks from Zambia, Kenya and Uganda using repetitive DNA probes. *Experimental Parasitology* **72**, 430–9.

HIDE, G., CATTAND, P., LE RAY, D., BARRY, J. D. & TAIT, A. (1990). The identification of *Trypanosoma brucei* subspecies using repetitive DNA sequences. *Molecular and Biochemical Parasitology* **39**, 213–26.

HIDE, G., WELBURN, S. C., TAIT, A. & 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**, 95–111.

KOERNER, T., DE RAADT, P. & MAUDLIN, I. (1995). The 1901 Uganda sleeping sickness epidemic revisited: A case of mistaken identity? *Parasitology Today* **11**, 303–6.

MACKICHAN, I. W. (1944). Rhodesian sleeping sickness in Eastern Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **38**, 49–60.

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

MAUDLIN, I., WELBURN, S. C., GASHUMBA, J. K., OKUNA, N. & KALUNDA, M. (1990). The role of cattle in the epidemiology of sleeping sickness in Uganda. *Proceedings of the VII International Congress of Parasitology, Paris, France.*

MAYNARD SMITH, J., SMITH, N. H., O'ROURKE, M. & SPRATT, B. G. (1993). How clonal are bacteria? *Proceedings of the National Academy of Sciences, USA* **90**, 4384–8.

MBULAMBERI, D. B. (1989). Possible causes leading to an epidemic outbreak of sleeping sickness. Facts and hypotheses. *Annales de la Société belge de Médecine tropical* **69**, 173–9.

MEHLITZ, D., BRINKMANN, U. & HALLER, L. (1981). Epidemiological studies in the animal reservoir of gambiense sleeping sickness. Part I. Review of literature and description of the study areas. *Tropenmedizin und Parasitologie* **32**, 129–33.

MEHLITZ, D., ZILLMANN, U., SCOTT, C. M. & GODFREY, D. G. (1982). Epidemiological studies in the animal reservoir of gambiense sleeping sickness. Part III. Characterization of *Trypanozoon* stocks by isoenzymes and sensitivity to human serum. *Tropennmedizin und Parasitologie* **33**, 113–18.

MIHOK, S., OTIENO, L. H. & DARJI, N. (1990). Population genetics of *Trypanosoma brucei* and the epidemiology of human sleeping sickness in the Lambwe Valley, Kenya. *Parasitology* **100**, 219–33.

ORMEROD, W. E. (1961). The epidemic spread of Rhodesian Sleeping Sickness 1908–1960. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **55**, 525–38.

PAINDAVOINE, P., PAYS, E., LAURENT, M., GELTMEYER, Y., LE RAY, D., MEHLITZ, D. & 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. & STEINERT, M. (1989). Different allele frequencies in *Trypanosoma brucei brucei* and *Trypanosoma brucei gambiense* populations. *Molecular and Biochemical Parasitology* **32**, 61–72.

STEVENS, J. R. (1991). Numerical studies on the biochemical characteristics of trypanosomes in the subgenus *Trypanozoon*. Ph.D. thesis, University of Bristol.

STEVENS, J. R. & GODFREY, D. G. (1992). Numerical taxonomy of *Trypanozoon* based on polymorphisms in a reduced range of enzymes. *Parasitology* **104**, 75–86.

STEVENS, J. R., LANHAM, S. M., ALLINGHAM, R. & GASHUMBA, J. K. (1992). A simplified method for identifying subspecies and strain groups in *Trypanozoon* by isoenzymes. *Annals of Tropical Medicine and Parasitology* **86**, 9–28.

STEVENS, J. R., MATHIEU-DAUDE, F., MCNAMARA, J. J., MIZEN, V. H. & NZILA, A. (1994). Mixed populations of *Trypanosoma brucei* in wild *Glossina palpalis palpalis*. *Tropical Medicine and Parasitology* **45**, 313–18.

STEVENS, J. R. & TIBAYRENC, M. (1995). Detection of linkage disequilibrium in *Trypanosoma brucei* isolated from tsetse flies and characterized by RAPD analysis and isoenzymes. *Parasitology* **110**, 181–6.

STEVENS, J. R. & WELBURN, S. C. (1993). Genetic processes within an epidemic of sleeping sickness in Uganda. *Parasitology Research* **79**, 421–7.

TAIT, A. (1980). Evidence for diploidy and mating in trypanosomes. *Nature, London* **287**, 536–8.

TAIT, A. (1990). Genetic exchange and evolutionary relationships in protozoan and helminth parasites. *Parasitology* (Suppl.) **100**, S75–S87.

TAIT, A., BABIKER, E. A. & LE RAY, D. (1984). Enzyme variation in *Trypanosoma brucei* spp. I. Evidence for the subspeciation of *Trypanosoma brucei gambiense*. *Parasitology* **89**, 311–26.

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

TIBAYRENC, M. (1993). *Entamoeba, Giardia* and *Toxoplasma*: clones or cryptic species? *Parasitology Today* **9**, 102–5.

TIBAYRENC, M. (1995). Population genetics of parasitic protozoa and other microorganisms. *Advances in Parasitology* **36**, 47–115.

TIBAYRENC, M. & AYALA, F. J. (1991). Towards a population genetics of microorganisms: the clonal theory of parasitic protozoa. *Parasitology Today* **7**, 228–32.

TIBAYRENC, M., KJELLBERG, F., ARNAUD, J., OURY, B., BRENIERE, S. F., DARDE, M.-L. & AYALA, F. J. (1991). Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. *Proceedings of the National Academy of Sciences, USA* **88**, 5129–33.

TIBAYRENC, M., KJELLBERG, F. & 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, USA* **87**, 2414–18.

TRUC, P. & TIBAYRENC, M. (1993). Population genetics of *Trypanosoma brucei* in Central Africa: taxonomic and epidemiological significance. *Parasitology* **106**, 137–49.