Numerical taxonomy of *Trypanozoon* based on polymorphisms in a reduced range of enzymes

J. R. STEVENS *and* D. G. GODFREY

Tsetse Research Laboratory, Department of Veterinary Medicine, University of Bristol, Langford House, Langford, Bristol BS18 7DU, UK

(*Received 10 May 1991; revised 29 July 1991; accepted 29 July 1991*)

SUMMARY

Numerical analyses of *Trypanozoon* taxonomy are presented, based on the isoenzyme data of Stevens *et al.* (1992). The previous study used a reduced range of enzymes compared with earlier work; the analyses indicate the value of this rationalized system. Both recently isolated trypanosome stocks and previously studied populations were included, allowing detailed comparison with earlier studies. Relationships between zymodemes were calculated with an improved similarity coefficient program, using Jaccard's coefficient (1908), and by Nei's method (1972). Dendrograms were constructed from the matrices produced with the group-average method. The groupings produced by both numerical methods were in close agreement, and the clusters of related principal zymodemes largely matched the species, subspecies and strain groups proposed by previous workers. *Trypanozoon* biochemical taxonomy is reviewed and the groupings reinforced by this study are: the mainly East African strain groups, busoga, zambezi, kakumbi, kiboko and sindo; *T.b. gambiense* and the bouaflé strain group from West Africa, and *T. evansi*; an intermediate bouaflé/busoga group was also recognized.

Key words: *Trypanozoon*, numerical taxonomy, isoenzymes, similarity coefficient, Nei's method.

INTRODUCTION

A previous publication (Stevens *et al.* 1992) describes a simplified approach for using enzyme polymorphisms to identify subspecies and strain groups in the subgenus *Trypanozoon*. The new reduced range of enzymes recommended allow the easier allocation of a new isolate to the correct category.

This paper presents numerical analyses of the isoenzyme data from the improved range of enzymes. The principal improvements were the introduction of the enzymes superoxide dismutase and another nucleoside hydrolase. Simultaneously, several highly polymorphic enzymes and one almost invariant enzyme used previously were eliminated.

In evaluating the new range of enzymes, relationships between zymodemes are assessed by both phenotypic and genotypic numerical methods, and the groupings are compared with those previously proposed. The evaluation also presented an opportunity to test fully a new program for calculating similarities on a personal computer (Stevens & Cibulskis, 1990).

To allow comparisons with previous work, we use the terms *T.b. brucei* and *T.b. rhodesiense* throughout this paper, although a re-appraisal of this nomenclature is already suggested (Tait *et al.* 1985; Godfrey *et al.* 1990).

MATERIALS AND METHODS

Trypanosomes : isolation and preparation

The origins of the 272 primary isolates used in defining the 111 enzyme profiles included in the analyses have been given by Stevens *et al.* (1992). Full details of isolation and preparation of material are also presented.

Enzyme electrophoresis

Electrophoresis was carried out on cellulose acetate plates (CAE) or on thin-layer starch gel (TSGE), depending on which gave the best resolution.

The following 2 enzymes were examined by TSGE as described previously (Kilgour & Godfrey, 1973): EC 2.6.1.2, alanine aminotransferase (ALAT); EC 2.6.1.1, aspartate aminotransferase (ASAT).

The following 5 enzymes were examined by CAE (Lanham *et al.* 1981; Stevens *et al.* 1989); EC 1.1.1.37, malate dehydrogenase (MDH); EC 1.1.1.42, isocitrate dehydrogenase (ICD); EC 1.1.1.103, threonine dehydrogenase (TDH); EC 2.7.5.1, phosphoglucomutase (PGM); and EC 1.15.1.1, superoxide dismutase (SOD).

EC 3.2.2.1, nucleoside hydrolase, using two different substrates was also examined by CAE. The substrates were inosine (NHI) and deoxyinosine (NHD); NHI was previously labelled as NH (Gibson *et al.* 1978; Gibson, Marshall & Godfrey,

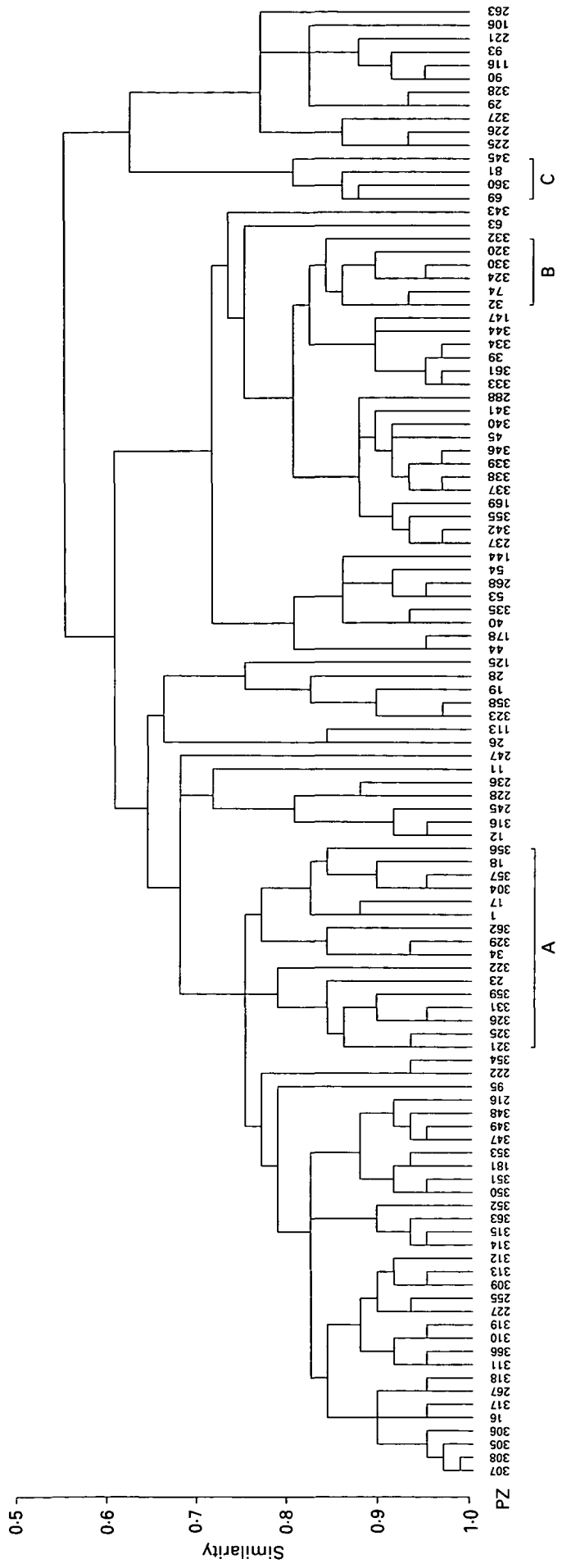

Fig. 1. Dendrogram A, constructed by UPGMA from a similarity coefficient matrix. Similarities between zymodemes were calculated by averaging the Jaccard coefficients from ten enzyme systems. PZ, Principal zymodeme.

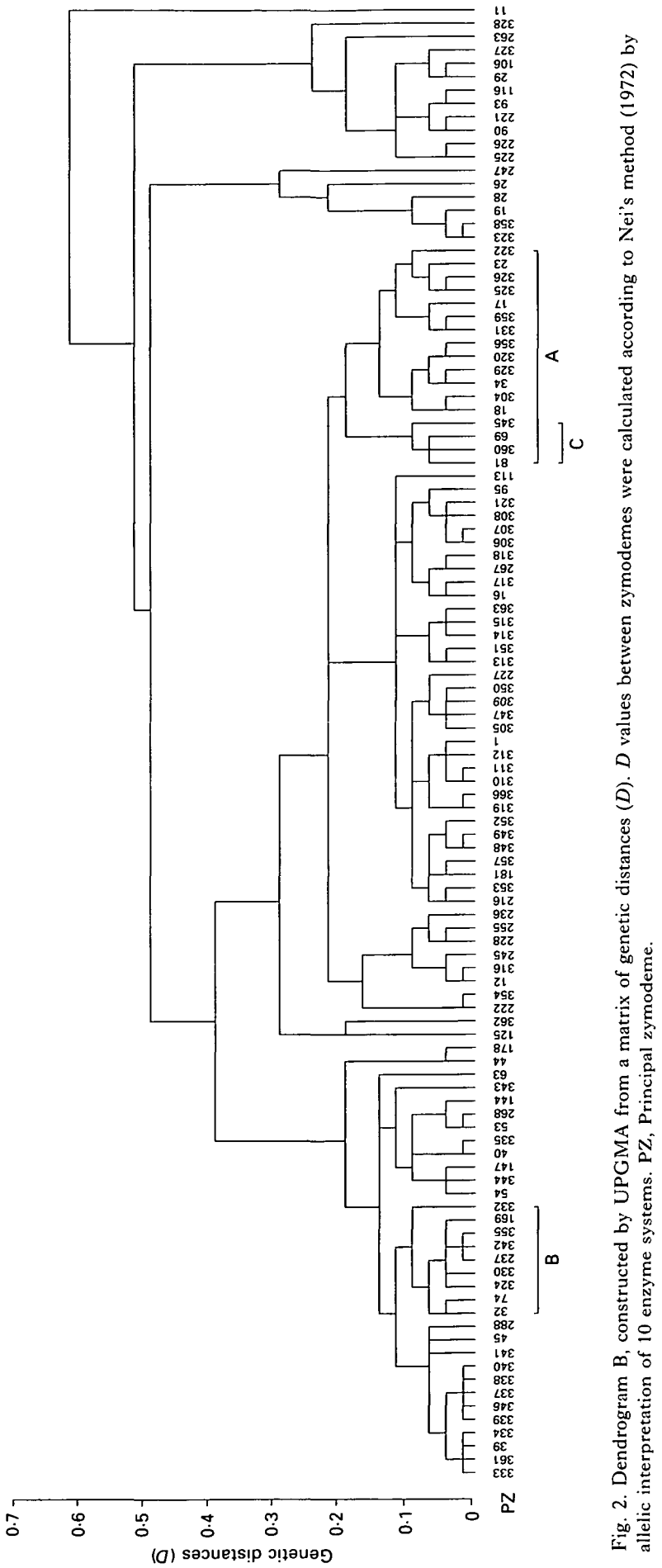

Fig. 2. Dendrogram B, constructed by UPGMA from a matrix of genetic distances (D). D values between zymodemes were calculated according to Nei's method (1972) by allelic interpretation of 10 enzyme systems. PZ, Principal zymodeme.

1980; Lanham *et al.* 1981; Stevens *et al.* 1989; Godfrey *et al.* 1990). The conditions for resolving these enzymes have been given by Stevens *et al.* (1992).

Numerical analyses

Matrices were calculated from the data of Stevens *et al.* (1992) using the isoenzyme profiles for every principal zymodeme. Each profile was regarded as an operational taxonomic unit (OTU) (Godfrey *et al.* 1990), and 111 unique OTUs were analysed. Relationships were demonstrated with dendrograms derived from distance matrices calculated by the two methods outlined below.

Similarity coefficient matrix. The similarity approach (SCM) has been widely employed for various parasitic protozoa. A notable study was that by Rioux *et al.* (1990) on the taxonomy of *Leishmania*. Most previously published work on African trypanosomes used matrices based on every band in a profile having the same weight in the final calculations; e.g. Gibson *et al.* (1980), Gashumba, Baker & Godfrey (1988), Mihok, Otieno & Darji (1990).

The program developed by Stevens & Cibulskis (1990) on a personal computer produced the similarity matrix according to Jaccard's method (1908). Similarity was calculated for every enzyme in a pair of OTUs, using all the isoenzyme bands seen in the two patterns of the same enzyme. The series of individual enzyme similarities was then used to determine an average similarity between the two OTUs. The averages for every possible pair of OTUs were entered into the similarity matrix. This approach was previously used by Le Blancq, Cibulskis & Peters (1986) for a numerical analysis of *Leishmania*. The advantages are 2-fold; firstly, equal weighting is given to all bands within an enzyme, and secondly, the 'over-weighting' of highly polymorphic enzymes in a profile is removed. The relationships are illustrated by Dendrogram A (Fig. 1).

Genetic distance matrix. Genetic interpretation by Nei's method (NM) (1972) was employed previously for *T. brucei* subspecies by Tait, Babiker & Le Ray (1984) and for *T. cruzi* by Tibayrenc *et al.* (1986), though dendrograms were not presented.

In this study, Dendrogram B (Fig. 2), was constructed from a matrix of genetic distances (D), calculated by pairwise comparison of OTUs according to NM. This necessitated genetic interpretation of enzyme patterns, based on apparent homozygous and heterozygous states in a diploid organism. Nei (1987) noted that the genetic identity (I), used for calculating D, is itself a measure of genetic similarity, ranging from 0 to 1. However, data are forced into an artificially limited range and consequently I is not widely employed as a measure of genetic similarity.

Dendrograms. Dendrograms of relationships have been constructed by the *unweighted pair-group method using arithmetic averages* (UPGMA) for a range of organisms. These have been described for *T. brucei* subspecies by Hide *et al.* (1990), Mihok *et al.* (1990) and Truc (1991).

The method is often referred to as the group-average method (Sokal & Michener, 1958, Lance & Williams, 1967). This is one of a family of techniques, collectively known as sequential, agglomerative, hierarchic, non-overlapping clustering (SAHN) methods (Sneath & Sokal, 1973). Other SAHN methods have been employed previously in studies of *Trypanozoon* taxonomy; Gibson *et al.* (1980) used the furthest neighbour method and Godfrey *et al.* (1990) used Ward's (1963) method. Both group-average and Ward's methods do well in comparison with other hierarchical clustering techniques (Dunn & Everitt, 1982). UPGMA also performs best against a range of tree-making methods (Nei, Tajima & Tateno, 1983; Sourdis & Krimbas, 1987), and was employed in this study because of the range of similarity coefficients and because it works well with linear measurements such as genetic distances (Nei, 1987). Furthermore, the method is readily available within the statistical package SPSS/PC + V3·0.

Coding of enzyme profiles

A different set of enzymes was used by Stevens *et al.* (1992) from those used in previous studies in this laboratory, by the inclusion of NHD and SOD, and the elimination of others. The reduced enzyme profile will characterize the principal zymodeme; a subordinate zymodeme will be derived from a more extensive range of enzymes. A stock previously placed in a particular zymodeme may now be allocated to a different principal zymodeme. However, to aid comparison of results, if change is unjustified, zymodeme, or profile, numbers remain as far as possible unaltered. Full details have been given by Stevens *et al.* (1992).

RESULTS

General comparison

When comparing the dendrograms produced by the two methods it should be stressed that the vertical axis scales are not directly comparable. With NM, the genetic distance D varies between 0 and infinity, and is akin to a measure of dissimilarity; with SCM, Jaccard similarity coefficients vary between 0 and 1.

The similarity coefficient method (SCM) and Nei's method (NM) showed good agreement, both

Table 1. Zymodemes of inconsistent classification

Final classification	Profile*	Zymodeme†	Godfrey *et al.* (1990)	Similarity coefficient method	Nei's method
T.b. gambiense	40	40	Unclassified	—	—
bouaflé	169	169	Unclassified	—	bou/bus
	237	237	—	—	bou/bus
	342	51	—	—	bou/bus
	355	278	—	—	bou/bus
bouaflé/busoga	32	32	busoga	—	—
	74	74	busoga	—	—
	324	25	busoga	—	—
busoga	304	—	N.D.‡	—	—
	320	37	—	bou/bus	—
	321	22	—	—	zambezi
	357	304	—	—	zambezi
	362	—	N.D.	—	Unclassified
zambezi	1	1	—	busoga	—
	95	95	kiboko	—	—
	255	255	—	—	kakumbi
kakumbi	11	11	—	—	Unclassified
	228	228	Unclassified	—	—
	236	236	Unclassified	—	—
	245	245	Unclassified	—	—
kiboko	113	113	—	—	zambezi
	125	125	—	—	Unclassified
Unclassified	44	44	bouaflé	—	—
	147	147	bouaflé	bouaflé	T.b.g.§
	178	178	bouaflé	—	—
	222	222	—	zambezi	—
	343	275	T.b.g.	—	T.b.g.
	344	285	bouaflé	bouaflé	T.b.g.
	354	223	—	zambezi	—

* Profiles are arranged in taxonomic groups. Profile number is determined from the combination of isoenzyme patterns seen in this study.

† Zymodeme number represents the combination of enzyme patterns previously observed with the 'old' range of enzymes (Godfrey *et al.* 1990). A blank entry indicates a new profile for stocks not analysed with the 'old' range of enzymes; however, such profiles are classified as far as possible according to the system of Godfrey *et al.* (1990), to allow comparison of numerical methods.

‡ N.D., not done. Although not examined before with the more extensive range of enzymes of Godfrey *et al.* (1990), these profiles clearly resembled previous busoga zymodemes.

§ T.b.g., *T.b. gambiense*.

with each other and with groupings in *Trypanozoon* previously put forward (Gibson *et al.* 1980; Gibson, Borst & Fase-Fowler, 1985; Godfrey *et al.* 1990): *T. evansi*, *T.b. gambiense*, bouaflé, kakumbi, sindo and zambezi; however, the earlier busoga strain group now appears to be subdivided. Nevertheless, confidence in the system of categories within the subgenus is enhanced because it remained broadly consistent despite using other numerical analyses on a different and reduced range of enzymes.

Godfrey *et al.* (1990) presented the modifications of earlier groupings; 74% of the zymodemes that either were, or would have been, classified by their method were similarly categorized by both SCM and NM. There was 84% agreement between SCM and NM; the agreements between Godfrey *et al.* (1990) and either SCM or NM were respectively 84% and 79%.

Table 1 presents the 29 zymodemes which were not placed in the same category by all three approaches (SCM, NM, Godfrey *et al.* 1990); included are two new profiles not recorded previously, but which would probably have been classified as busoga. All but three of the 29 zymodemes are placed in one of the range of groups (Stevens *et al.* 1992), when two of the three approaches agreed either in full or partly. Zymodemes 147, 343 and 344 were not treated in this fashion, and remain unclassified. These three fall into the area of uncertainty concerning '*non-gambiense*' infections of man, and potential animal reservoirs, in West Africa (e.g. Mehlitz *et al.* 1982; Gibson, 1986; Godfrey *et al.* 1987). Although other zymodemes of like description (e.g. Z39) are quite firmly placed in bouaflé, it is felt that any doubtful ones should at present remain unclassified.

Similarity coefficient method – Dendrogram A

The groupings formed by SCM corresponded to most subspecies and strain groups defined by Gibson *et al.* (1980, 1985) and Godfrey *et al.* (1990). These were bouaflé, *T.b. gambiense*, kakumbi, kiboko, sindo and zambezi. In contrast to the description by Godfrey *et al.* (1990), busoga appeared to be separated into three distinct groups, one related to zambezi, a second, well separated and showing affinity with sindo, and a third classed by Stevens *et al.* (1992) as bouaflé/busoga.

Nei's method – Dendrogram B

The groups defined by NM were compact compared with those defined by SCM. This may have been due to fine variation in the raw data being lost when the band patterns were genetically interpreted.

In agreement with Dendrogram A, the groups which corresponded to those proposed earlier (Gibson *et al.* 1980; Godfrey *et al.* 1990) were bouaflé, *T.b. gambiense*, kakumbi, kiboko, sindo and zambezi. The previous busoga strain group was split into two, not three as by the SCM. One was termed a bouaflé/busoga group by Stevens *et al.* (1992). The kiboko and sindo strain groups were particularly well separated from the others; *T.b. gambiense* and zambezi were tightly defined, but less well separated. Kakumbi was less well defined and was close to zambezi. Certain zymodemes previously placed in busoga (Godfrey *et al.* 1990) were only poorly differentiated from bouaflé.

Strain groups

The validity of the groupings formed is evaluated in accordance with previous terminology (Gibson *et al.* 1980, 1985; Borst, Fase-Fowler & Gibson, 1987). These earlier studies have been summarized and extended by Godfrey *et al.* (1990), and this publication forms the basis of the following evaluation.

The profiles of 80 stocks from Cameroon, Côte d'Ivoire, Uganda and Brazil are numerically analysed for the first time. Many possessed profiles described previously (Gibson *et al.* 1980; Godfrey *et al.* 1990), although several, particularly those from Uganda belonged to novel zymodemes (Stevens *et al.* 1992).

East Africa. Busoga. The SCM produced a grouping that did not correspond exactly with previous definitions of the busoga strain group. Those zymodemes that were or would have been classed as busoga, grouped into one large A, and two small, separate parts, B and C.

The 16 profiles within part A showed 75% similarity with zambezi and included isolates from both patients and livestock (Stevens *et al.* 1992). This part corresponded broadly with a large set allocated to busoga by Godfrey *et al.* (1990). The classification of part B zymodemes is discussed below in bouaflé. Part C, with 4 profiles showed 63% affinity with sindo, but was well separated from all other strain groups and the other parts of busoga; stocks were from a range of hosts in East Africa. This separation was not reproduced by NM and it bore no resemblance to any previous grouping (Godfrey *et al.* 1990), and thus, Part C remains within busoga.

By NM, zymodemes previously recognized as busoga were split into two parts. The larger part, A, with 17 zymodemes was very distinct and, as with SCM, its nearest neighbour was zambezi. It also contained the 4 zymodemes placed in part C by SCM. These zymodemes and their close relatives were previously classified as busoga; accordingly, we present parts A and C as the busoga strain group (Stevens *et al.* 1992). Part B was less well defined and mixed with bouaflé, and is discussed below.

Sixteen zymodemes in busoga were, or would have been, classified as busoga by the earlier definition; most were associated with the Lake Victoria area. Of the 56 busoga isolates 96% were from East Africa, with 75% from Uganda and 16% from Kenya. Two isolates were from Cote d'Ivoire, with one from Tanzania and another from Ethiopia; 41% of isolates were from man, 41% from cattle and 9% from pigs; one stock was from a tsetse. These percentages may be considerably biased because of a collection of new isolates mostly from patients and cattle in Uganda.

Godfrey *et al.* (1990) brought together 5 separate sets of zymodemes to form busoga. Their reasons for this were (a) the relative similarity of isoenzyme patterns between the sets and (b) their closeness in both a dendrogram and a cladogram. However, in this study, the zymodemes corresponding to this earlier concept of busoga fell into separate parts which were less readily joined and some zymodemes previously classified as busoga, were better placed in a separate bouaflé/busoga strain group (see below).

Zambezi. By SCM, the largest strain group, zambezi, with 32 zymodemes showed 75% similarity with its nearest neighbour, part A of busoga. Intragroup similarity was also high; 95% of profiles showed greater than 82% similarity. Zambezi with 33 zymodemes was also well defined by NM, and all were strongly genetically similar. By both methods, two new zymodemes (Z311, Z366) associated with man, and another, Z363 represented by a pig stock, were placed in zambezi; the new stocks were from Uganda.

A total of 29 zymodemes grouped into zambezi by both methods; 28 were, or would have been confirmed as zambezi by Godfrey *et al.* (1990). The 31 zymodemes now classed as zambezi were comprised of 73 stocks (Stevens *et al.* 1992); 85% were from man and 10% from tsetse. The remainder

came from domestic and wild animals. All isolates were from East Africa, largely from Zambia (74%) and Uganda (21%).

The constituent profiles of the zambezi strain group, using both numerical methods, remain relatively unchanged from those presented by Godfrey *et al.* (1990). The features characterizing the group as a whole were reinforced by the additional enzymes. New zymodemes containing isolates from Ugandan sleeping-sickness patients were included in zambezi by both numerical methods.

Kakumbi. Too few isolates were examined from this group to attempt to draw any conclusions other than those previously expressed by Godfrey *et al.* (1990). Notably, however, even with the reduced number of profiles with the new enzyme range, the position of kakumbi in both dendrograms was in good agreement with that proposed by Godfrey *et al.* (1990).

SCM defined the kakumbi group well; it had only 67% similarity with its nearest neighbours, zambezi and part A of busoga. However, the minimum intragroup similarity was 72%, indicating large internal enzymic differences. By NM, kakumbi showed affinity with busoga and zambezi, but did not include Z11.

All isolates were from Zambian wild animals and tsetse.

Kiboko. Kiboko was well defined by SCM and showed only 64% similarity with its nearest neighbours, the other East African strain groups. However, intragroup variability was high, with a minimum similarity of only 66%. By NM, kiboko was equally well defined, but was of slightly different composition.

Characteristically, kiboko stocks were Kenyan apart from one (Z19), which originated just across the border in Tanzania. Isolates were from a broad range of animal hosts, 50% tsetse, 50% domestic and wild animals.

Sindo. In agreement with Gibson *et al.* (1985), Godfrey *et al.* (1990) and Mihok *et al.* (1990), sindo was recognized by both numerical methods as an isolated and compact group; it is the most genetically distinct of all the groups defined by isoenzyme studies to date. Only 1 out of the 12 isolates within sindo was from man; the possibility of fleeting infections in man of otherwise animal-infective organisms has been discussed by Godfrey *et al.* (1990).

Sindo was well separated from all other *Trypanozoon* by SCM; part C of busoga, its nearest neighbour, was only 63% similar. Intragroup similarity was high, at 77% or more. By NM, sindo was separated even more distinctly, and the zymodemes included corresponded exactly in each method. The same 11 zymodemes were, or would have been, classified as sindo by earlier work.

Zymodemes were recognized by having ASAT-4, unique to this strain group. Characteristically, profiles also possessed SOD_A-3 (otherwise found only in *T.b. gambiense* or kiboko) coupled with NHD-2 or NHD-5, although the combinations were not strain group specific. All 12 isolates were from East Africa, 58% from Kenya, 42% from Zambia; 33% were from tsetse, 17% from sheep, 42% from wild animals and 1 isolate from man.

East and West Africa. Bouaflé/busoga. In the previous study by Godfrey *et al.* (1990), some zymodemes were classified as either busoga or bouaflé. Part B of busoga in each dendrogram appeared to correspond to some extent, to such a grouping. SCM and NM agreed in assigning 5 zymodemes to this bouaflé/busoga group, which included 2 zymodemes previously assigned to bouaflé. It was recognized by Stevens *et al.* (1992) as an intermediate group.

Geographical distinctions are not apparent in this group. The 5 zymodemes included 10 stocks from a range of East (30%) and West (70%) African locations. Hosts were man (30%), pigs (60%) and cattle (10%).

From the wide geographic separation of most busoga and bouaflé stocks (Stevens *et al.* 1992) it appears unlikely that populations in Z32 and Z74 are closely related to others in bouaflé/busoga, despite the enzymic similarity. However, the Lake Victoria area, from which many of the busoga stocks originate, could form the bridge between East and West Africa, so mixing of populations from the two strain groups is possible. Following on from this, the range of enzymes used in this study may be too small to adequately differentiate some proposed strain groups. Whatever, the recognition of a particular group of populations within bouaflé seems pertinent, and thus its intermediate classification as bouaflé/busoga seems justified.

The lack of SOD_A-3 in bouaflé/busoga isolates indicates that if the population from West African animals were genuine reservoirs of human sleeping sickness, they would probably be of the type labelled as '*rhodesiense*-like' (Mehlitz *et al.* 1982; Paindavoine *et al.* 1986; Hide *et al.* 1990), thus further highlighting the inadequacy of the generally accepted nomenclature.

West Africa. Bouaflé. By both methods, the majority of bouaflé zymodemes as defined by Godfrey *et al.* (1990) were contained in one main grouping. SCM defined the group tightly; zymodemes had at least 80% similarity. By NM, bouaflé zymodemes were separated by a maximum D of 0·109. However, contained within this tight group were zymodemes previously classified as busoga and *T. evansi.*

Twelve zymodemes were placed in bouaflé by SCM, NM and previous work. Another set of 4

zymodemes remained together in both dendrograms, but were related to Part B of busoga by SCM and to bouaflé by NM. Among the 16 zymodemes classed as bouaflé by Stevens *et al.* (1992), 96% of the isolates were from West Africa, of which 83% were from Côte d'Ivoire. Stocks were from a broad range of hosts, mostly animals, including 62% from domestic pigs; 7% were from tsetse and 14% from man.

The isolation of stocks with characteristics intermediate between those of bouaflé and *T.b. gambiense* is not altogether unexpected, and both Paindavoine *et al.* (1989) and Godfrey *et al.* (1990) suggested that *T.b. gambiense* may have evolved from a West African *T.b. brucei* variant. With regard to this, and the recognition of a bouaflé/busoga grouping, the suggestion by Godfrey *et al.* (1990) that the name *T.b. brucei*, generally associated with non-infectivity to man, be retained for the essentially West African bouaflé strain group seems sensible. However, as they stated, human infectivity should not be entirely excluded.

T.b. gambiense. One group of six zymodemes corresponded to *T.b. gambiense* by SCM; they showed the least intragroup variability of any strain group, with more than 86% similarity. The same six were grouped together by NM, and with the exception of Z40, they were previously similarly classified. The six are therefore listed as *T.b. gambiense* by Stevens *et al.* (1992, Table 1).

Characteristics of isolates within *T.b. gambiense* were much as those defined by Gibson *et al.* (1980) and Godfrey *et al.* (1987, 1990). NHD-1 and SOD_A-3 with SOD_B-5, appeared particularly characteristic of *T.b. gambiense*, although none of these alone was unique to the group. In agreement with Godfrey & Kilgour (1976), ALAT-1 and ASAT-2 were also characteristic, but again not diagnostic alone.

All isolates, apart from one, originated from West Africa, in particular from Côte d'Ivoire (38%) and from Cameroon (30%). Several stocks from Côte d'Ivoire, not previously characterized by isoenzymes, due to extremely low parasitaemias in rodents (Dukes *et al.* 1989), were included in Z268. The single non-West African stock was from a patient in Sudan. Host range was similarly limited, with 93% of isolates being from man and 7% from domestic animals.

In agreement with previous studies (Gibson *et al.* 1980; Paindavoine *et al.* 1986; Godfrey *et al.* 1987; Truc, 1991) *T.b. gambiense* emerged by both numerical methods as a well-defined group, though not as distinct as *T.b. gambiense* defined in the dendrogram of Godfrey *et al.* (1990). *T.b. gambiense* showed considerable affinity with some bouaflé zymodemes and *T. evansi*, and this link supports the suggestion by Paindavoine *et al.* (1989) and Godfrey *et al.* (1990), that *T.b. gambiense* evolved from bouaflé.

Pan Africa and South America. T. evansi. The reduced range of enzymes used in the present study means that previously separate zymodemes are now classified as one, Z39. This also includes some non-*T. evansi* isolates and accordingly Z39 is placed with bouaflé by all isoenzyme methods. This agrees with Gibson, Wilson & Moloo (1983) who noted that *T. evansi* resembles West African, rather than East African, *T. brucei* subspecies; in a later study Boid (1988) suggested that *T. evansi* may be evolving from *T.b. brucei* at the northern edge of the tsetse zone in Kenya and Sudan. The other *T. evansi* zymodeme Z63 is picked out as distinctly different by past and present methods.

Based on the kDNA characteristics described by Borst *et al.* (1987), Godfrey *et al.* (1990) brought together Z63 and the other *T. evansi* zymodemes. The distinctive Z63 profile (Gibson *et al.* 1983; Stevens *et al.* 1989) suggests two separate strain groups in *T. evansi*, which is confirmed by DNA probes (Masiga & Gibson, 1990). Accordingly, we now present *T. evansi* as two strain groups, mongonu (Z39) and ngurunit (Z63), named after locations where the first specimens were isolated (Gibson *et al.* 1980).

New isolates. The agreement between the position of new stocks and the existing classification system is perhaps one of the best indicators of the strength of the methods used, reaffirming the power of the classification system first suggested by Gibson *et al.* (1980).

The inclusion of the new stocks from man in zambezi, by both numerical methods was in good agreement with other epidemiological features. All were isolated from either Kigulu, in the Busoga region of Uganda, where an outbreak of rhodesian sleeping-sickness has been established since 1976 (Gibson & Gashumba, 1983), or the Tororo district, which is currently experiencing an epidemic of rhodesian sleeping-sickness (Maudlin, personal communication). Prior to inclusion of the new stocks, only four populations in three zambezi zymodemes were from Uganda. All other zambezi profiles were associated with Zambia, bar two from Tanzania and one from Rwanda. So, although somewhat unusual in origin, the new stocks conformed to the enzymic definition of zambezi (Godfrey *et al.* 1990) and are probably correctly placed. Gibson & Gashumba (1983) isolated stocks of Z5 from the Busoga epidemic, a zymodeme previously found only in Botswana, Ethiopia and Zambia and possibly introduced by military or refugee movements.

By NM, Z357, isolated from a Zebu cow was also included in zambezi, though its placement in busoga by SCM may have been more realistic, as it also met the enzymic criteria for this group of Godfrey *et al.* (1990). Indeed, all new cattle stocks met the enzymic criteria for inclusion in busoga, apart from their

possessing ALAT-10, a pattern not previously seen in Uganda. Prior to the inclusion of new material, busoga showed a range of stocks from cattle and Uganda, all with ALAT-1, -2 or -12 (Godfrey *et al.* 1990).

DISCUSSION

One purpose of this study was to evaluate the introduction of two potentially informative enzymes, NHD and SOD, into the rationalized enzyme system employed by Stevens *et al.* (1992) for identifying species, subspecies and strain groups within *Trypanozoon*. The numerical analyses based on the new range of nine enzymes produced clusters of related principal zymodemes similar to groupings previously put forward; thus the introduction of the new range was vindicated.

The credibility of any grouping system is greatly enhanced if the groupings defined by one method can be corroborated by alternative methods. The numerical methods used in this study were based on two quite different approaches, yet produced similar arrangements of the zymodemes. This was encouraging, particularly so because of the broad agreement with the phenetic and cladistic analyses previously presented (Gibson *et al.* 1980; Godfrey *et al.* 1990). Most importantly, the groupings had a sound epidemiological basis and more confidence can now be placed in the classification system first proposed by Gibson *et al.* (1980), and in the individual recognition of taxonomic groups by DNA criteria (Borst, Fase-Fowler & Gibson, 1981; Borst *et al.* 1987; Gibson *et al.* 1985; Paindavoine *et al.* 1986, 1989).

The groupings of *Trypanozoon* may be summarized as follows. *T. evansi* and *T.b. gambiense* stand as well-defined species and subspecies. On the basis of four distinct isoenzyme types, *T.b. brucei* could be subdivided into bouaflé, sindo, kiboko and kakumbi, although genetically bouaflé is quite distinct from the other three. Bouaflé represents *T.b. brucei* from West African animals, though some man-infective, potentially man-infective and East African populations are included.

In agreement with Tait *et al.* (1985) and Godfrey *et al.* (1990), we believe that *T.b. brucei*, as in the widely accepted concept of not infecting man, may not exist in East Africa. Instead we can identify at least three strain groups (sindo, kakumbi, kiboko) within *T.b. rhodesiense*, which appear to be generally confined to animal/tsetse transmission cycles. Sindo represents animal and fly stocks from East Africa and in agreement with Gibson *et al.* (1985), it has been shown to be genetically distinct from other non-man-infective *Trypanozoon*. Kakumbi and kiboko represent other isoenzymically distinct East African forms of *T.b. brucei*, with kakumbi in particular being found in flies and game; kiboko has also been shown by Gibson *et al.* (1985) to be genetically distinct from other *T. brucei* subspecies.

T.b. rhodesiense, in its classical sense of being man-infective, is subdivided into zambezi and busoga, each with characteristic isoenzyme profiles. Zambezi contains stocks isolated predominantly from man in Zambia. Busoga contains stocks also mainly from man and includes the more northerly and central African zymodemes responsible for rhodesian human sleeping sickness; busoga zymodemes are only occasionally found in West Africa. It is more enzymically variable than zambezi and shows affinity with non-man-infective strain groups, particularly bouaflé.

Broader groups are proposed by several researchers. Tait *et al.* (1985), found no enzyme variants specific to *T.b. rhodesiense* in a set of eleven stocks and concluded from this sample, that *T.b. rhodesiense* is a set of variants of *T.b. brucei*, rather than a subspecies. In contrast, using a considerably larger sample, the present study found combinations of enzyme variants which clearly defined two man-infective strain groups within East Africa, which were distinct from other strain groups in the region that appeared primarily associated with animal infectivity. However, some overlap occurred between man-infective populations from East Africa with bouaflé, the West African form of *T.b. brucei* (Stevens *et al.* 1992).

In another study of *Trypanozoon* in East Africa, Mihok *et al.* (1990) re-analysed isoenzyme results from Gibson *et al.* (1980) and Gibson & Wellde (1985), together with their own data. Their dendrogram supports recognition of the busoga and sindo strain groups. They also defined groupings approximating to zambezi. Paindavoine *et al.* (1986, 1989), using DNA techniques, were also unable to differentiate between *T.b. brucei* or *T.b. rhodesiense*, although they clearly distinguished *T.b. gambiense* from non-*gambiense* stocks.

The dendrograms produced by our two methods were similar. Nei's method (NM) requires some genetic interpretation, but allows broad comparison with a wealth of genetic and evolutionary studies, beyond those on trypanosome numerical taxonomy. The choice of method depends to a large extent on whether a quick result or a more thorough genetic investigation is required.

Genetic distances as measured by NM mean little on their own and are best understood when compared with distances found in other species. Excluding birds and some primates, genetic differences (D) between species within the same genus are generally in the range 0·05–3·00 (Nei, 1987). Thorpe (1982) concluded that populations can be regarded as separate species if $D > 0{\cdot}16$, but that conferring species status on populations separated by $D < 0{\cdot}16$ should be considered dubious. However, this criterion does not apply to birds (Avise & Aquadro,

1982) or higher primates (King & Wilson, 1975) where species are isolated by strong social and behavioural patterns, and small genetic distances between species are common, while in microbial species where reproduction is mostly asexual, large genetic distances between conspecific strains are not uncommon. This has been shown for *E. coli* (Whittam, Ochman & Selander, 1983) and for the fungus *Mucor racemosus* (Stout & Shaw, 1974).

Thus, a scale of magnitude has been developed of the genetic distance required to define a species. At one end are the primates and birds, where speciation occurs with relatively few genetic differences (Wyles, Kunkel & Wilson, 1983), while at the other are the micro-organisms where social factors do not operate and where asexual reproduction provides a basis for large genetic differences even within a species.

The high degree of interspecific variation within *Trypanozoon* fits in well with the picture for micro-organisms. However, the genetic homogeneity of *T. evansi* and bouaflé is an anomaly not easily explained. It may be that a much larger number of enzymes than the nine used needs to be studied before the true extent of the genetic separation of *T. evansi* can be realized. This apart, distinct differences in the genetic makeup of *T. evansi* and *T. brucei* subspecies exist; Borst *et al.* (1987) showed reduced mini-circle heterogeneity and an absence of maxi-circles in the kDNA of *T. evansi*, while information on host, origin, culture characteristics and tsetse transmissibility is enough to differentiate most *T. evansi* stocks from *T. brucei* subspecies.

The relatively large genetic distances between many of the strain groups within the *T. brucei* subspecies may point to the importance of asexual reproduction following limited genetic isolation in the creation of these groups (Tibayrenc *et al.* 1990). In contrast, the credibility of the groups is further enhanced by the generally small intragroup genetic distances (Avise & Aquadro, 1982). In this respect. *T. evansi*, *T.b. gambiense* and zambezi are of particular note, although bouaflé and busoga also show limited intragroup variation. Overall, the assortment of zymodemes into the strain groups, subspecies and species presented in Dendrogram B, and the genetic distances between them are in good agreement with current knowledge of genetic distances in micro-organisms.

We should like to thank numerous colleagues for help in the laboratory, and Dr R. E. Cibulskis for assistance with the program to calculate Nei's distances. We gratefully acknowledge the financial support given by the Medical Research Council and the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.

REFERENCES

AVISE, J. C. & AQUADRO, C. (1982). A comparative summary of genetic distances in the vertebrates. *Evolutionary Biology* **15**, 151–85.

LE BLANCQ, S. M., CIBULSKIS, R. E. & PETERS, W. (1986). *Leishmania* in the Old World: 5. Numerical analysis of isoenzyme data. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **80**, 517–24.

BOID, R. (1988). Isoenzyme characterisation of 15 stocks of *Trypanosoma evansi* isolated from camels in the Sudan. *Tropenmedizin und Parasitologie* **39**, 45–50.

BORST, P., FASE-FOWLER, F. & 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**, 117–31.

BORST, P., FASE-FOWLER, F. & GIBSON, W. C. (1987). Kinetoplast DNA of *Trypanosoma evansi*. *Molecular and Biochemical Parasitology* **23**, 31–8.

DUKES, P., KAUKAS, A., HUDSON, K. M., ASONGANYI, T. & GASHUMBA, J. K. (1989). A new method for isolating *Trypanosoma brucei gambiense* from sleeping sickness patients. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **83**, 636–9.

DUNN, G. & EVERITT, B. S. (1982). *An Introduction to Mathematical Taxonomy*. Cambridge: Cambridge University Press.

GASHUMBA, J. K., BAKER, R. D. & GODFREY, D. G. (1988). *Trypanosoma congolense*: the distribution of enzymic variants in East and West Africa. *Parasitology* **96**, 475–86.

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

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

GIBSON, W. C. & GASHUMBA, J. K. (1983). Isoenzyme characterization of some *Trypanozoon* stocks from a recent trypanosomiasis epidemic in Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **77**, 114–18.

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., MEHLITZ, D., LANHAM, S. M. & GODFREY, D. G. (1978). The identification of *Trypanosoma brucei gambiense* in Liberian pigs by isoenzymes and by resistance to human plasma. *Tropenmedizin und Parasitologie* **29**, 335–45.

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.

GIBSON, W. C., WILSON, A. J. & MOLOO, S. K. (1983). Characterisation of *Trypanosoma* (*Trypanozoon*) *evansi* from camels in Kenya using isoenzyme

electrophoresis. *Research in Veterinary Science* **34**, 114–18.

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. & KILGOUR, V. (1976). Enzyme electrophoresis in characterizing the causative organism of Gambian trypanosomiasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **70**, 219–24.

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.

HIDE, H., 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.

JACCARD, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise de Science Naturelle* **44**, 223–70.

KILGOUR, V. & GODFREY, D. G. (1973). Species-characteristic isoenzymes of two aminotransferases in trypanosomes. *Nature, London* **244**, 69–70.

KING, M. C. & WILSON, A. C. (1975). Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–16.

LANCE, G. N. & WILLIAMS, W. T. (1967). A general theory of classificatory sorting strategies: I. Hierarchical systems. *Computer Journal* **9**, 373–80.

LANHAM, S. M., GRENDON, J. M., MILES, M. A., POVOA, M. M. & DE SOUZA, A. A. A. (1981). A comparison of electrophoretic methods for isoenzyme characterization of trypanosomatids. 1: Standard stocks of *Trypanosoma cruzi* zymodemes from north-east Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **75**, 742–50.

MASIGA, D. K. & GIBSON, W. C. (1990). Specific probes for *Trypanosoma* (*Trypanozoon*) *evansi* based on kinetoplast DNA minicircles. *Molecular and Biochemical Parasitology* **40**, 279–84.

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. *Tropenmedizin 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.

NEI, M. (1972). Genetic distance between populations. *The American Naturalist* **106**, 283–92.

NEI, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.

NEI, M., TAJIMA, F. & TATENO, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* **19**, 153–70.

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.

RIOUX, J. A., LANOTTE, G., SERRES, E., PRATLONG, F., BASTIEN, P. & PERIERES, J. (1990). Taxonomy of *Leishmania*. Use of isoenzymes. Suggestions for a new classification. *Annales de Parasitologie Humaine et Comparée* **65**, 111–25.

SNEATH, P. H. A. & SOKAL, R. R. (1973). *Numerical Taxonomy*. San Francisco: W. H. Freeman and Company.

SOKAL, R. R. & MICHENER, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin* **38**, 1409–38.

SOURDIS, J. & KRIMBAS, C. (1987). Accuracy of phylogenetic trees estimated from DNA sequence data. *Molecular Biological Evolution* **4**, 159–66.

STEVENS, J. R. & CIBULSKIS, R. E. (1990). Analysing isoenzyme band patterns using similarity coefficients: a personal computer program. *Computer Methods and Programs in Biomedicine* **33**, 205–12.

STEVENS, J. R., LANHAM, S. M., ALLINGHAM, R. & GASHUMBA, J. K. (1992). A more efficient identification of subspecies and strain groups in *Trypanozoon* by isoenzymes. *Annals of Tropical Medicine and Parasitology* (in the Press).

STEVENS, J. R., NUNES, V. L. B., LANHAM, S. M. & OSHIRO, E. T. (1989). Isoenzyme characterization of *Trypanosoma evansi* isolated from capybaras and dogs in Brazil. *Acta Tropica* **46**, 213–22.

STOUT, D. L. & SHAW, C. R. (1974). Genetic distance among certain species of *Mucor*. *Mycologia* **66**, 969–77.

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.

THORPE, J. P. (1982). The molecular clock hypothesis: biochemical evolution, genetic differentiation and systematics. *Annual Review of Ecological Systematics* **13**, 139–68.

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.

TIBAYRENC, M., WARD, P., MOYA, A. & AYALA, F. J. (1986). Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease, have a complex multiclonal structure. *Proceedings of the National Academy of Sciences, USA* **83**, 115–19.

TRUC, P. (1991). Apport de la génétique des populations à la taxonomie de *Trypanosoma brucei* et à

l'épidémiologie de la Trypanosomiase Humaine en Afrique Centrale. Thèse, Université Montpellier II, France.

WARD, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**, 236–44.

WHITTAM, T. S., OCHMAN, H. & SELANDER, R. K. (1983). Geographic components of linkage disequilibrium in natural populations of *Escherichia coli*. *Molecular Biological Evolution* **1**, 67–83.

WYLES, J. S., KUNKEL, J. G. & WILSON, A. C. (1983). Birds, behavior and anatomical evolution. *Proceedings of the National Academy of Sciences, USA* **80**, 4394–7.