Molecular variation of *Trypanosoma brucei* subspecies as revealed by AFLP fingerprinting

E. E. C. AGBO[1]*, P. A. O. MAJIWA[3], H. J. H. M. CLAASSEN[2] *and* M. F. W. TE PAS[1]

[1] *Division of Animal Sciences, Section for Animal Genomics, Institute for Animal Science and Health, ID-Lelystad, Edelhertweg 15, 8200 AB Lelystad, The Netherlands*
[2] *Department of Immunology, Erasmus University Rotterdam, Dr Molewaterplein 50, 3015 GE Rotterdam, The Netherlands*
[3] *International Livestock Research Institute, P.O. Box 30709, Nairobi, Kenya*

(*Received 20 September 2001 ; revised 3 December 2001 ; accepted 3 December 2001*)

SUMMARY

Genetic analysis of *Trypanosoma* spp. depends on the detection of variation between strains. We have used the amplified fragment length polymorphism (AFLP) technique to develop a convenient and reliable method for genetic characterization of *Trypanosome* (sub)species. AFLP accesses multiple independent sites within the genome and would allow a better definition of the relatedness of different *Trypanosome* (sub)species. Nine isolates (3 from each *T. brucei* subspecies) were tested with 40 AFLP primer combinations to identify the most appropriate pairs of restriction endonucleases and selective primers. Primers based on the recognition sequences of *Eco*RI and *Bgl*II were chosen and used to analyse 31 *T. brucei* isolates. Similarity levels calculated with the Pearson correlation coefficient ranged from 15 to 98%, and clusters were determined using the unweighted pair-group method using arithmetic averages (UPGMA). At the intraspecific level, AFLP fingerprints were grouped by numerical analysis in 2 main clusters, allowing a clear separation of *T. b. gambiense* (cluster I) from *T. b. brucei* and *T. b. rhodesiense* isolates (cluster II). Interspecies evaluation of this customized approach produced heterogeneous AFLP patterns, with unique genetic markers, except for *T. evansi* and *T. equiperdum*, which showed identical patterns and clustered together.

Key words: *Trypanosoma* species, *Trypanosoma brucei* subspecies, genetic analysis, AFLP fingerprinting, genetic markers.

INTRODUCTION

Trypanosoma brucei is a unicellular, protozoan parasite that is transmitted by tsetse flies. It consists of 3 subspecies – *T. b. brucei*, *T. b. gambiense* and *T. b. rhodesiense* that are indistinguishable by conventional morphological, biochemical and antigenic criteria but differ by their geographical distribution and host specificity (Gibson, Marshall & Godfrey, 1980; Gibson & Wellde, 1985). The disease, trypanosomosis (Kassai, 1988), is generally an acute to chronic, debilitating disease. *T. b. brucei*, the cause of 'nagana' in cattle does not cause disease in humans because this subspecies is lysed by normal human serum. *T. b. gambiense* and *T. b. rhodesiense* are resistant to normal human serum, although *T. b. rhodesiense* can revert to a human serum sensitive phenotype (Hawking, 1973, 1977). There is a resurgence of sleeping sickness and increasing incidence of drug resistance (Murray *et al*. 2000), thus the need for high-resolution analytical tools which will facilitate the typing of *T. brucei* isolates on a routine basis increases. Moreover, the ability to differentiate between human-infective and human non-infective isolates has important implications for studies of the transmission dynamics and anthropozoonotic potential of this parasite.

As an initial step towards an appropriate strategy for management of the disease, the genetic diversity of the parasite's populations must be assessed. Genetic markers have been used to study the diversity of trypanosome populations. These have included isoenzyme studies (Gibson *et al*. 1980; Gibson & Gashumba, 1983; Young & Godfrey, 1983; Tait, Babiker & Le Ray, 1984; Tait *et al*. 1985; Richner *et al*. 1989; Godfrey *et al*. 1990; Stevens & Godfrey, 1992), randomly amplified polymorphic DNA data (Mathieu-Daude *et al*. 1995; Kukla *et al*. 1987; Waitumbi & Murphy, 1993), microsatellite and minisatellite DNA analysis (Donelson & Artama, 1998; Morlais *et al*. 1998; MacLeod *et al*. 2000; Biteau *et al*. 2000) and restriction fragment length polymorphism, RFLP (Hide *et al*. 1994; Agbo *et al*. 2001). However, the identification of the characteristics specific to a subspecies is not yet supported by the discriminatory power of conventional diagnostic methods. Furthermore, many of the genetic typing assays have drawbacks in that they may require a relatively large amount of high-quality DNA or, as with RAPD technique, may be difficult

* Corresponding author: Division of Animal Sciences, Section for Animal Genomics, Institute for Animal Science and Health, ID-Lelystad, Edelhertweg 15, 8200 AB Lelystad, The Netherlands. Tel: +31 320 238391. Fax: +31 320 238050. E-mail: e.e.c.agbo@id.wag-ur.nl

Parasitology (2002), **124**, 349–358.
DOI: 10.1017/S003118200100138X Printed in the United Kingdom

Table 1. Trypanosome isolates used in these studies

Isolate	Origin	Year	Host
Trypanosoma brucei brucei			
1. Lump266	Kiboko, Kenya	1969	Fly, *G. pallidipes*
2. PTAG130	Daloa, Ivory Coast	1985	Pig
3. KP2, clone 7	Kouassi-Parita, Ivory Coast	1982	Fly (*G. palpalls*)
4. B8/18	Nsukka, Nigeria	1962	Pig
5. SW3/87	Democratic Rep. of Congo	1987	Pig
6. LVBG3N	Lambwe valley, Kenya	1980	Cow
7. M249	Matuga, Kenya	1981	Sheep
8. H3	Luangwa valley, Zambia	1974	Lion
9. STIB215	Serengeti, Tanzania	1971	Lion
10. 1902 (sindo)	Kenya	1971	Waterbuck
11. AnTat2/2	Nigeria	1970	Fly, *G. morsitans*
12. AnTat17/1	Democratic Rep. of Congo	1978	Sheep
13. J10	Zambia	1973	N.A.
T. b. gambiense group 1			
14. ITMAP1843	Bouenza, Congo	1975	Human
15. A006	Fontem, Cameroon	1988	Human
16. Font-1	Fontem, Cameroon	1993	Human
17. PT16	Côte d'lvoire	1992	Human
18. PT312	Côte d'lvoire	1992	Human
19. Mabia	Bouenza, Congo	1989	Human
20. NW2	Uganda	1992	Human
21. Suzena	Yambio, Sudan	1982	Human
22. 1829 (Aijo)	Badundu, D.R. Congo	1970	Human
23. 1898	Democratic Rep. of Congo	1974	Human
24. Dal972	Daloa, Ivory Coast	1978	Human
T. b. gambiense group 2			
25. TB26	Bouenza, Congo	1983	Pig
Trypanosome brucei rhodesiense			
26. TRP2320	Zambia	1983	Fly, *G. pallidipes*
27. Gambella II	Illubabor, Ethiopia	1968	Human
28. 058	Luangwa valley, Zambia	1975	Human
29. LVH56	Lambwe valley	1978	Human
30. LVH108	Lambwe valley	1980	Human
31. AnTat25.1*	Rwanda	1971	Human
T. evansi/*T. equiperdum*			
AnTat3.1 (T. ev)	South America	1969	Capybara
RoTat1.2 (T. ev)	Indonesia	1982	Water buffalo
AnTat4.1 (T. eq)	N.A.	Unknown	N.A.
STIB818 (T. eq)	Beijing, China	1979	Horse
T. congolense			
C49 (savannah)			
Gam2 (savannah)	The Gambia	1977	Cow
IL3900 (riverine/forest)	Burkina Faso	1982	Dog
ANR3 (riverine/forest)	The Gambia	1988	Fly
K45.1 (Kilifi)	Kenya	1982	Cow
WG5 (Kilifi)	Kenya	1980	Goat
T. simiae			
TsØ2	Kenya	N.A.	Bushbuck
Ken2	The Gambia	1988	Fly

* Isolate was adapted-sensitive to normal human serum.

to reproduce between laboratories (Jones *et al*. 1997; Perez, Albomoz & Dominguez, 1998). Many markers, especially isoenzymes, can evolve too fast for use as epidemiological markers, so that results can be misinterpreted (Hide, 1999). Furthermore, in the majority of these methods, only a very limited part of the genome is covered through highly specific molecular targeting of one or more repetitive DNA elements. To avoid possible bias arising from the use of such methods, the population structure should be inferred from neutral markers that distribute randomly throughout the genome.

We have used the amplified fragment length polymorphism (AFLP) technique, a PCR-based fingerprinting method, to investigate the genetic diversity of *T. brucei* isolates, and tested the technique for trypanosome genotyping in general. The technique has already been frequently used for

linkage analysis in plant genetics, and has also been applied for high resolution fingerprinting of plant, bacterial and fungal genomes (Vos *et al.* 1995; Folkerstma *et al.* 1996; Lin, Kuo & Ma, 1996). Moreover, it has been applied to the analysis of population diversity in potato cysts and root-knot nematodes (Folkerstma *et al.* 1996; Semblat *et al.* 2000). There are 2 main advantages (Dijkshoom *et al.* 1996; Janssen *et al.* 1996; Lin *et al.* 1996) of this technique in characterizing microbial populations: there is an extensive coverage of the genome under study (Janssen *et al.* 1996) and the complexity of the AFLP fingerprint can be advantageously managed by adding selective primers during PCR amplifications (Vos *et al.* 1995), and by varying the choice of restriction enzyme pairs. These features make the technique better suited for revealing higher levels of genetic variation than isoenzyme and RFLP analysis. In addition, the reproducibility is reported to be high (Vos *et al.* 1985), while the requirement for only small amounts of DNA makes it highly suitable for genotyping isolates on a large scale or for isolation of binary specific markers. A recent review by Masiga, Tait & Turner (2000) highlighted the value of this technique to parasite genetic studies. The objective of this study was to ascertain the usefulness of the AFLP technique in assessing the genetic diversity of trypanosome strains at the species or subspecies level, and as a tool to study potential relationships of various trypanosome (sub)species.

MATERIALS AND METHODS

Trypanosome materials

The 31 isolates used were directly derived from different hosts between 1968 and 1993, and have been characterized on the basis of their response to normal human serum, as well as on host and geographical origins (Gibson, W. C., personal communication). The geographical origin, year of isolation and source of strain are described in Table 1. Genomic DNA was isolated from the purified pellets using a standard method (Van der Ploeg *et al.* 1982).

AFLP reactions

AFLP markers were assayed as previously described by Vos *et al.* (1995), with the following modifications: 50–300 ng DNA was digested in a total volume of 20 μl using 10 U each of 8 combinations of enzymes (*Eco*RI/*Mse*I, *Bgl*II/*Hpa*II, *Hin*dIII/*Hpa*II, *Eco*RI/*Hin*dIII, *Eco*RI/*Bgl*II, *Eco*RI/*Bam*HI, *Bgl*II/*Hin*dIII and *Bgl*II/*Bam*HI. All restriction enzymes were from New England Biolabs or Roche Molecular Biochemicals. The digestion mixture was incubated at 37 °C for 4 h followed by the addition of 10× ligation buffer (20 mM Tris–HCl, 60 mM KCl, 1 mM EDTA, 5 mM dithiothreitol, 50% glycerol (v/v), pH 7·5 (4 °C), DNA ligase (1 U)) and 10 pmol (or 100 pmol for 4-bp recognizing endonucleases) of respective adapters (Table 2) and incubation at 15 °C for 16 h. In a series of preliminary experiments, PCR amplification conditions were optimized (data not shown). Based on these results, the final optimized reaction conditions for pre-selective PCR included 2·5 mM dNTP's, 2 mM $MgCl_2$, 1× PCR buffer (100 mM Tris–HCl, pH 8·3, 500 mM KCl, 1·5 mM $MgCl_2$, 0·01% (w/v) gelatin), 5 pM of each pre-selective primer (Table 1), 1·25 U AmpliTaq® DNA polymerase (Perkin-Elmer, Maarssen, The Netherlands) and 4 μl of 1:2-diluted ligation product. The optimized thermal cycling conditions were 2 min at 95 °C, followed by 20 cycles of 95 °C for 30 s, 56 °C for 30 s and 72 °C for 1 min. The PCR product was diluted 20-fold and 4 μl of the dilution was used for the selective amplification reaction, using primer combinations with a selective nucleotide at the 3′ end of one or both primers. The cycling conditions were as for the pre-selective amplification, except for a final extension at 60 °C for 30 min. The final products were diluted 4-fold, and 1 μl, together with a Genescan-500 internal standard (PE Applied Biosystems, Maarssen, The Netherlands), was analysed on a 7·3% denaturing sequencing gel by using an ABI 373A automated DNA sequencer. Gels were routinely prepared by using the ABI protocols and were electrophoresed for 5 h. Forty primer combinations were evaluated using all 9 possible +1/+0 or +1/+1 primer combinations for each pair of enzymes (with selective nucleotide only on 3′ end of the upstream primer, or on both primers, respectively) (Table 2). The *T. brucei* isolates were analysed using 3 selected restriction endonuclease combinations, with the set of selective primers, namely *Eco*+A/*Bgl*+0, *Eco*+A/*Bgl*+T and *Eco*+0/*Hin*d+A.

Data analysis

After electrophoresis, AFLP data were collected with the Genescan software (PE Applied Biosystems, Maarssen, The Netherlands). Densitometric values were transferred to the GelCampar v4.1 software (Applied Maths, Kortrijk, Belgium), and gels were normalized by using the internal size standard that was added to each lane. Levels of similarity between banding patterns were calculated with the Pearson correlation product-moment coefficient (Pearson, 1926). The unweighted pair group method using average linkage (UPGMA) was used to cluster the patterns (Vauterin & Vauterin, 1992).

RESULTS

Selection of AFLP enzymes and primers

AFLP genetic markers were assessed for their usefulness in characterizing molecular diversity

Table 2. Adapters and PCR primer core sequences used in this study

	Adapter	Primer core sequence
*Eco*RI	5′-CTC GTA GAC TGC GTA CC 3′-CAT CTG ACG CAT GGT TAA	5′-**6-FAM**-GAC TGC GTA CCA ATT C
*Hin*dIII	5′-CTC GTA GAC TGC GTA CC 3′-CTG ACG CAT GGT CGA	5′-**6-FAM**-GAC TGC GTA CCA GCT T
*Bgl*II	5′-CGG ACT AGA GTA CAC TGT C 3′-C TGA TCT CAT GTG ACA GCT AG	5′-**6-FAM**-GAG TAC ACT GTC GAT CT
*Mfe*I	5′-AAT TCC CAA GAG CTC TCC AGT AC 3′-G GTT CTC GAG AGG TCA TGA T	5′-**6-FAM**-GAG AGC TCT TGG AAT TG
*Xba*I	5′-GTC GTA GAC TGC GTA CG 3′-CTG ACG CAT GCG ATC	5′-**6-FAM**-GAC TGC GTA CGC TAG A
*Mse*I	5′-GAC GAT GAG TCC TGA G 3′-CTA CTC AGG ACT CAT	5′-GAT GAG TCC TGA GTA A
*Hpa*II (C/CGG)	5′-GAC GAT GAG TCC TGA T 3′-CTA CTC AGG ACT AGC	5′-GAT GAG TCC TGA TCG G

Table 3. Primer combinations that were evaluated by AFLP

(Fingerprints were scored (+ to ++++), based on relative total number of bands and uniformity of size distribution from 35–500 bp. The combinations shown in bold were selected for further analyses. 'O' is AFLP primer without any selective nucleotide; 'A' or 'T' indicate primer with the respective primer.)

		*Bam*HI			*Eco*RI			*Hin*dIII			*Hpa*II		*Mse*I				
Enzyme	Base	O	A	T	O	A	T	O	A	T	T	TC	A	G	T	C	AT
*Bgl*III	O	–	+	+	–	++++	++	–	++	+	+	+	–	–	–	–	–
	A	–	+	+	+	+++	+	+	+	+	++	+	–	–	–	–	–
	T	–	+	+	+	+++	+	+	+	+	++	+	–	–	–	–	–
*Eco*RI	O	–	+	+				–	+++	++	–	–	–*	–*	–*	–*	–*
	A	–	+	+				–	+	+	–	–	–*	–*	–*	–*	–*
	T	–	+	+				–	+	+	–	–	–*	–*	–*	–*	–*
*Hin*dIII	O	–	–	–							–	–	–	–	–	–	–
	A	–	–	–							++	+	–	–	–	–	–
	T	–	–	–							++	+	–	–	–	–	–
*Bam*HI	O				–	+	+	–	–	–	–	–	–	–	–	–	–
	A				+	+	+	–	–	–	–	–	–	–	–	–	–
	T				+	+	+	–	–	–	–	–	–	–	–	–	–

\+ Very few and weak bands.
–* Not scored because most bands are relatively very small (< 150 bp) and faint.
– Not studied.

among *T. brucei* subspecies. Nine *T. brucei* isolates were first analysed to determine the optimal AFLP conditions. Restriction digestion and adapter ligation were performed in alternate steps, followed by pre-selective and selective PCR amplification.

Nine sets of restriction enzyme combinations were tested (Table 3). The banding patterns obtained using *Eco*RI and *Mse*I restriction enzymes (with A+T-rich recognition sequences) for example, contained more bands than fingerprints generated with *Bam*HI and *Hpa*II (with G+C-rich recognition sequences). Overall, fingerprints with the pair of 6-bp/4-bp recognizing endonuclease combinations (*Eco*RI/*Mse*I, *Bgl*II/*Hpa*II and *Hin*dIII/*Hpa*II) gave unevenly distributed, non-scorable banding patterns or monomorphic profiles of mostly small-sized (< 150 bp) fragments (Table 3). Thus, we did not find the sets of 6-bp/4-bp recognizing endonuclease combinations suitable for useful AFLP fingerprinting of *T. brucei*. Two sets of restriction enzyme combinations (*Eco*RI/*Bgl*II and *Eco*RI/*Hin*dIII) were selected based on their reproducibility, even distribution of bands along the gel, number of polymorphic bands and polymorphism detected within and between subspecies. The *Eco*RI/*Bgl*II combination (with A selective nucleotide at the 3′ end of *Eco*RI primer) gave the best polymorphic, evenly distributed profiles, and was chosen for further analysis. To evaluate for the reproducibility of our results, the AFLP analyses procedures of the 9 test isolates were repeated 3 times, starting from DNA digestion.

AFLP analysis of Trypanosoma brucei

The 31 *T. brucei* stocks comprising 13 animal-infective and 18 human-infective isolates were subjected to AFLP analysis. The AFLP profiles were found to be strikingly similar, especially for the *T. b. gambiense* isolates and SW3/87 (Fig. 1A, lanes 19–31), in spite of their distant geographical origins and time of isolation. After normalization of data, the similarity levels between individual fingerprints, calculated by Pearson correlation product-moment coefficient, ranged from 60 to 96%. The fingerprint pattern showed a relatively high degree of heterogeneity of the *T. b. brucei* and *T. b. hodesiense* subspecies, while the *T. b. gambiense* isolates showed rather homogeneous patterns (Fig. 1A).

Cluster analysis

The use of Pearson coefficient (Pearson, 1926) for the calculation of inter-strain correlations allowed a rapid classification of analysed isolates into 2 main clusters, I and II (Fig. 1A). The *T. b. gambiense* isolates (Cluster I) were maximally separated from *T. b. brucei* and *T. b. rhodesiense*, both residing in Cluster II. Cluster I isolates were further subdivided into subclusters, denoted as P and Q, while cluster II has R and S subclusters (Fig. 1A). Numerical analysis of the AFLP profiles showed that the *T. b. gambiense* cluster was joined with the *T. b. brucei*/*T. b. rhodesiense* cluster at a Pearson correlation level of 60%. Similarity level among the *T. b. gambiense* isolates (Cluster I) ranged from 92 to 96%, indicating a highly identical set of isolates. Despite the overall homogeneity of this cluster, our numerical data showed a further subdivision in which *T. b. gambiense* isolates from West and Central Africa were homogeneously distributed into 2 subclusters, designated as P and Q (Fig. 1A). The first is composed of 2 Congo stocks from the same focus isolated from humans over a 14-year period, and all the 5 isolates from West Africa (Côte d'lvoire and Cameroon). The second subcluster contains 2 Democratic Republic (D.R.) of Congo stocks, isolated 14 years apart, and 1 isolate each from Congo, Sudan and Uganda. Interestingly, visual inspection of the normalized band patterns revealed 2 *T. b. gambiense*-specific marker bands, M1 and M2 (Fig. 1A). One '*T. b. brucei*' stock (SW3/87), a pig isolate from Congo, was 'outgrouped' into Cluster I. This implies that this particular isolate is closer to *T. b. gambiense* than it is to *T. b. brucei*/*T. b. rhodesiense* in Cluster II, which raises the issue of whether it is indeed a *T. b. brucei*. Given that it has the *T. b. gambiense*-specific bands M1 and M2, one is led to the conclusion that this isolate may in fact be a *T. b. gambiense*. Within Cluster II, the *T. b. brucei* and *T. b. rhodesiense* isolates were subdivided into 2 different subclusters, designated R and S. The Cluster II dendogram showed longer branch lengths of subclusters, with a similarity level of 60–80%. This indicates higher diversity among the isolates, compared to the 92–96% similarity level of Cluster I isolates. Overall, highly identical fingerprints were obtained from the 12 *T. b. gambiense* isolates, while *T. b. brucei* and *T. b. rhodesiense* isolates showed more genetic diversity than *T. b. gambiense*.

Interspecies analysis

To evaluate the applicability of the technique for interspecies analysis, AFLP data obtained from 2 independent isolates each of *T. congolense* (kilifi, riverine/forest and savannah), *T. evansi* and *T. equiperdum* were similarly analysed. The fingerprint patterns were compared to those of randomly selected *T. brucei* isolates (Fig. 1B). The dendogram of the 19 populations, based on AFLP analysis using the UPMGA program, resulted in 2 main clusters, in which the *Trypanozoon* (*Trypanosoma brucei* ssp., *T. evansi* and *T. equiperdum*) isolates were maximally separated from the *Nannomonas* (*T. congolense*) and *Pycnomonas* (*T. simiae*) at a similarity level of 15%. The similarity level between individual fingerprints of each pair of (sub)species, calculated by Pearson coefficient, ranged from 15 to 95%, indicating an overall high level of genetic heterogeneity among the subgenera (Fig. 1B). The *Trypanozoon* isolates (Cluster I) were further divided into 2 subclusters comprising (i) the *T. b. gambiense* (PT16 and Mabia), and SW3/87, a '*T. b. brucei*' which had been included to evaluate if its clustering position would also be different in an interspecies analysis. Both are distinct from other *Trypanozoon* isolates; subcluster II comprised all the other members of *Trypanozoon* subgenus (*T. b. brucei*, *T. b. rhodesiense*, *T. evansi* and *T. equiperdum*). The Kenyan *T. b. brucei* isolate from cow (LVBG3N) is out-grouped from other members of this subcluster, showing a more distant relationship. The *T. b. rhodesiense* isolates (Gambella-2 and 058) are more closely related to the *T. evansi*/*T. equiperdum* clade, both of which are separated at a similarity level of 73%. Interestingly, KP2, clone 7, a tsetse fly isolate from Côte d'lvoire is on the *same T. evansi*/*T. equiperdum* clade. Overall, the 2 subclusters show a genetic similarity level of about 68%, which is a significant level of diversity within the *Trypanozoon* subgenus.

Within Cluster II, heterogeneous fingerprint patterns were obtained, except for *T. evansi* and *T. equiperdum* which shared highly similar AFLP patterns and a high level of genetic homogeneity of 95–98% (Fig. 1B). Furthermore, in the interspecies analysis, there was a higher degree of homogeneity between the fingerprint patterns of each of the analysed pair of isolates than in the intraspecific analysis (Fig. 1B). For instance, the fingerprint patterns of IL3900 and ANR3, both *T. congolense*

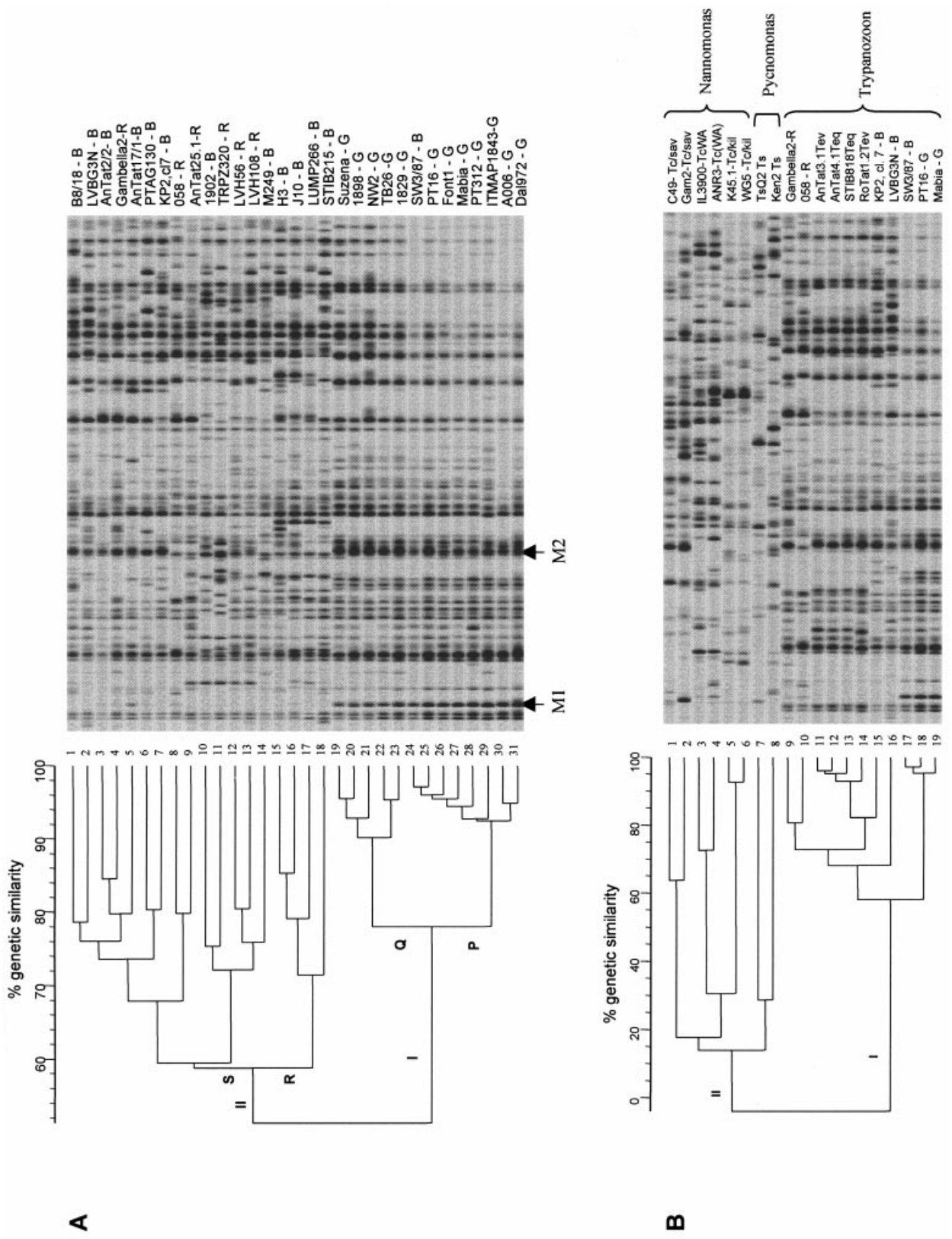

Fig. 1. For legend see facing page.

(riverine/forest) isolates, were quite similar (Fig. 1B, lanes 1 and 2). Taken together, the data indicate that there exists remarkable diversity within the genomes of *Trypanozoon* isolates.

DISCUSSION

Accurate measurement of genetic relationships requires a marker system that provides an unbiased estimate of total genomic variance and is sufficiently abundant to minimize errors due to sampling variance (Spooner *et al*. 1996). We looked for a genotyping method suitable for trypanosome (sub)species for several reasons. The limitations of previously described methods based on single genetic locus have been highlighted above. Karyotypes of *Trypanosoma brucei* isolates, for example, are known to be highly plastic (Gibson, 1991) with unusual chromosomal size polymorphisms between isolates (Melville *et al*. 1998). Therefore, any genotyping technique based on a single genetic locus may provide misleading results. The AFLP method is an extremely useful and reliable technique for detecting polymorphisms and its reproducibility is reported to be very high (Vos *et al*. 1995). The appropriate choice of restriction enzymes and the number and base composition of selective bases determine the usefulness and applicability of this technique in genetic diversity studies. In theory, each 6-bp recognizing endonuclease cleaves every 4^6 (1096) bp in the genome. Since trypanosome genomes are relatively small, the addition of only 1 selective nucleotide should yield scorable banding patterns. In our hands, the use of a pair of 6-bp recognizing endonucleases (*Eco*RI/*Bgl*II), with a selective nucleotide at the 3′ end of the *Bgl*II primer was sufficient to generate useful fingerprints for trypanosome (sub)species genotyping. We propose this combination as the standard set with which to characterize trypanosomes to the (sub)species level.

In this study, we identified optimal AFLP conditions for *Trypanosome* (sub)species fingerprinting. The restriction endonucleases that were used to genotype other species did not produce satisfactory AFLP patterns: trypanosome genomic DNA was not adequately digested with *Bam*HI, *Hha*I digestion resulted in monomorphic fragments, while *Eco*RI and *Mse*I combination resulted in too many small fragments. Restriction endonucleases *Eco*RI combined with another 6-bp-cutting endonuclease *Bgl*II resulted in reproducible, evenly distributed and the most informative band patterns. The resolution of different selective primers tested varied, but *Eco*RI + A selective nucleotide combined with *Bgl*II primer with no selective base produced the best patterns. Genetically related strains produced homologous patterns and grouped together, whereas unrelated strains separated on the dendogram.

In certain eukaryotic systems especially plants, polysaccharides are often co-precipitated with DNA (Rether, Delmas & Laoued, 1993), and extraction steps with organic solvents (such as phenol, chloroform) may fail to strip all sugars completely. With the high glycoprotein content in trypanosome surface coat (the variable surface glycoproteins, VSG) this may lead to partial or poor DNA digestion, which is an important step in the AFLP technique. To improve the chances of complete enzyme restriction, we opted for restriction digestion of genomic DNA followed by ligation in 2 separate steps. This produced better and more consistent results compared to digestion and ligation in one reaction (data not shown). To obtain reproducible fingerprints, a standardized protocol for AFLP analysis and computer-based analysis was required. Since a pre-amplification step could preferentially enrich heterosite fragments following adapter ligation (Vos *et al*. 1995), we chose the 2-step PCR amplification procedure: a pre-selective reaction (without any selective nucleotide on primer), followed by selective amplification reaction.

As in previous studies with RAPD and single-locus analyses (Mathieu-Daude *et al*. 1995; Hide *et al*. 1998; Biteau *et al*. 2000; MacLeod *et al*. 2000; Agbo *et al*. 2001), our data correlated with the close genetic relatedness among the 3 *T. brucei* subspecies. In our analysis, therefore, we aimed to group the isolates rather than distinguish them, crucial for biodiversity surveys or epidemiological research. The

Fig. 1. (A) Numerical analysis of normalized AFLP bands patterns generated from *Trypanosoma b. brucei*, *B* ($n = 13$), *T. b. gambiense*, *G* ($n = 12$) and *T. b. rhodesiense*, *R* ($n = 6$), using primer combination *Eco*RI + A/*Bgl*II + 0. The dendogram was constructed using the unweighted pair-group method using arithmetic averages, UPGMA. Correlation levels were expressed as percentage values using the Pearson coefficient. M1 and M2 denote subspecies-specific AFLP marker bands differentiating *T. b. brucei* from *T. b. gambiense*/*T. b. rhodesiense*. (B) Comparative dendogram based on AFLP band patterns from interspecies analysis of *T. congolense* – Tc/kil (kilifi), Tc/WA (riverine/forest) and Tc/sav (savannah), *T. simiae* (Ts), *T. evansi* (Tev) and *T. equiperdum* (Teq) strains. Their fingerprints were compared to the patterns generated from 2 isolates each of *T. b. gambiense* (TB26 and Mabia) and *T. b. rhodesiense* (Gambella-2 and 058), and 3 *T. b. brucei* isolates (KP2, clone 7, LVBG3N and SW3/87). Similarities between pairs were calculated by Pearson correlation coefficient. Fingerprints were generated with *Eco* + A/*Bgl* + 0 primer combinations. Two main clusters result which comprise of isolates in the *Trypanozoon* family (*T. brucei*, *T. evansi*, *T. equiperdum*) (Cluster I), and those belonging to *Nannomonas* (*T. congolense*) and *Pycnomonas* (*T. simiae*) subgenera.

Pearson correlation coefficient (Pearson, 1926) suggested that individual *T. b. brucei* and *T. b. rhodesiense* isolates from various foci may be quite distinct, while the dendogram data indicate that isolates from the same region are similar. Some of the *T. b. brucei* and *T. b. gambiense* isolates used in this study have also been examined, using RFLP, for variations in their ITS/5·8S rDNA region (Agbo *et al*. 2001), who proposed 2 distinct evolutionary lines of descent. Interestingly, in these current studies, each of the main clusters of isolates revealed by numerical analysis of AFLP variations corresponded to the respective evolutionary line of the organism. It should be noted that AFLP is independent of ITS/5·8S rDNA sequence analysis because the nucleotide positions analysed by AFLP are distributed over the whole genome. Therefore, the 2 *T. b. gambiense*-specific AFLP fragments that we detected appear to be additional reliable and specific markers for the characterization of *T. b. gambiense* stocks. Clustering analysis showed that variations were more pronounced among *T. b. brucei*/*T. b. rhodesiense* subspecies (cluster II), which suggests a slightly different level of genetic heterogeneity between the two lines.

For *T. brucei* genotyping, our data indicate that human-infective *T. b. gambiense* form a distinct homogeneous group separate from *T. b. brucei* and *T. b. rhodesiense* isolates, which cluster together. This agrees with earlier reports that *T. b. gambiense* is a homogeneous group of isolates with high levels of similarity (Biteau *et al*. 2000). The low level of molecular variation among the *T. b. gambiense* isolates may be explained by 2 hypotheses (i) the introduction of human-infective trypanosomes to West and Central Africa (the ecozone for *T. b. gambiense*) was more recent and, most probably, only a few strains were introduced, and; (ii) the cumulative effects adopted for the control of the disease. The control of Gambian sleeping sickness relies on case identification, principally using the card agglutination test (Magnus, Vervoort & Van Meirvenne, 1978), and chemotherapy of a largely asymptomatic human reservoir. On the other hand, *T. b. rhodesiense*, the cause of Rhodesian sleeping sickness is zoonotic, with a reservoir in wild animals and domestic livestock (Welburn *et al*. 2001). Combined with the acuteness of the disease, its control demands a more aggressive approach. These have included large-scale tsetse control operations, aerial spraying and aggressive medical interventions. These measures may have accumulated to induce more mutational pressure on *T. b. rhodesiense*, resulting in an increased level a genetic diversity within the subspecies.

As an interspecies yardstick with which to measure intra-specific divergence among *T. brucei* subspecies, 2 isolates each from *Nannomonas* (*T. congolense*) and *Pycnomonas* (*T. simiae*) subgenera were similarly analysed, and their fingerprints compared to *Trypanozoon* (*T. brucei*) patterns. As shown, the different isolates within each subgenus have similar but distinct fingerprint patterns, except for *T. evansi* and *T. equiperdum*, whose fingerprint patterns were closely identical. This obvious close relationship between the two 'species' is in agreement with earlier reports (Gibson *et al*. 1980; Lun, Brun & Gibson, 1992; Zhang & Baltz, 1994; Brun, Hecker & Lun, 1998; Biteau *et al*. 2000). However, a clear correlation was found between AFLP polymorphisms and trypanosome species. The specific bands present in each of the analysed species allowed us to separate the strains effectively. On the basis of the pronounced differences in their dendograms, different subtypes of *T. congolense* (kilifi, riverine/forest and savannah) could be considered as separate subspecies. Sequence information based on rRNA genes have been used to characterize and to infer the phylogenetic relationship among *Trypanozoon*, *Pycnomonas* and *Nannomonas* trypanosome parasites (Stevens *et al*. 1999; Urakawa & Majiwa, 2001; Gibson *et al*. 2001). Our data and the previous reports have corroborated the close genetic relatedness between *T. b. brucei* and *T. b. rhodesiense*. In addition, our results did show the existence of micro-heterogeneity among isolates of these subspecies. Overall, our data show that there does exist more genetic diversity within and between the subgenera than was previously reported.

The AFLP method has a higher multiplex ratio (= number of loci simultaneously analysed per experiment) than single sequence repeats (SSRs). As it simultaneously accesses multiple independent sites within the entire genome, it provides a more valuable tool for overall evaluation of the phylogenetic relatedness of trypanosome strains. Our study has established the AFLP conditions and primer combinations that permit the assessment of genetic diversity of *T. brucei* subspecies. It further demonstrated that the technique could be applied in trypanosome diversity studies, with potential use for the isolation of intra- or inter-specific genetic markers. The results also show that human infective isolates derived from different ecozones may indeed be genetically separate groups. Furthermore, they correlate with the close phenotypic relatedness between *T. b. brucei* and *T. b. rhodesiense*. In addition, our data suggest that variation between human-infective *T. brucei* subgenus is beyond only geographical origin, since the *T. b. rhodesiense* seem to be genetically more diverse, while the *T. b. gambiense* are more clonal or genetically more stable over time. By this technique, the two subspecies can be unambiguously classified by whole-genome fingerprinting. Through the evaluation of large numbers of clearly defined field samples, such AFLP fingerprinting may facilitate the identification of polymorphisms linked to parasite factors of host tropism,

and contribute to the understanding of host–parasite interactions at the molecular level. Finally, our data present some further markers for defining *T. b. gambiense*, while the detected polymorphisms of trypanosome species may be valuable tools for epidemiological studies.

We thank Drs Philippe Buscher and Magda Radwanska for their constructive comments on the manuscript. This work was funded by the Institute for Animal Science and Health (ID-Lelystad), Lelystad, The Netherlands, under internally funded iSEO project number 640-47293.

REFERENCES

AGBO, E. C., MAJIWA, P. A. O., CLAASSEN, E. H. J. M. & ROOS, M. H. (2001). Measure of molecular diversity within the *Trypanosoma brucei* subspecies – *Trypanosoma brucei brucei* and *Trypanosoma brucei gambiense* as revealed by genotypic characterisation (in the Press).

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

BRUN, R., HECKER, H. & LUN, Z. R. (1998). *Trypanosoma evansi* and *Trypanosoma equiperdum*: distribution, biology, treatment and phylogenetic relationship (a review). *Veterinary Parasitology* **79**, 95–107.

DIJKSHOORN, L., AUCKEN, H., GERNER-SMIDT, P., JANSSEN, P., KAUFMANN, M. E., GARAIZER, J., URSING, J. & PITT, T. L. (1996). Comparison of outbreak and non-outbreak *Actinobacter baumanni* strains by genotypic and phenotypic methods. *Journal of Clinical Microbiology* **34**, 1519–1525.

DONESLON, J. E. & ARTAMA, W. T. (1998). Diagnosis of *T. evansi* by polymerase chain reaction (PCR). *Journal of Protozoology Research* **8**, 204–213.

FOLKERTSMA, R. T., VAN DER VOORT, R., DE GROOT, J. N. A. M., DE GROOT, K. E., VAN ZANDVOORT, P. M., GOMMERS, F. J., HELDER, J. & BAKKER, J. (1996). Gene pool similarities of potato cyst nematode populations assessed by AFLP analysis. *Molecular Plant-Microbe Interactions* **9**, 47–54.

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

GIBSON, W. C. & GARSIDE, L. (1991). Genetic exchange in *Trypanosoma brucei*: variable location of housekeeping genes in different trypanosome stocks. *Molecular and Biochemical Parasitology* **45**, 77–90.

GIBSON, W. C. & GASHUMBA, J. K. (1983). Isoenzyme characterisation of some *Trypanozoon* stocks from a recent trypanosomiasis epidemic in Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **77**, 114–118.

GIBSON, W. C., STEVENS, J. R., MWENDIA, C. M., MAKUMI, J. N., NGOTHO, J. N. & NDUNG'U, J. M. (2001). Unravelling the phylogenetic relationships of African trypanosomes of suids. *Parasitology* **122**, 625–631.

GIBSON, W. C. & WELLDE, B. T. (1985). Characterisation of some trypanosome stocks from the South Nyanza sleeping sickness focus in Western Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **79**, 671–676.

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.

HAWKING, F. (1973). The differentiation of *Trypanosoma rhodesiense* from *T. brucei* by means of human serum. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **67**, 517–527.

HAWKING, F. (1977). The resistance to human plasma of *Trypanosoma brucei*, *T. rhodesiense* and *T. gambiense*. In *Analysis of the Composition of Trypanosome Strains*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **70**, 504–512.

HIDE, G. (1999). History of sleeping sickness in East Africa. *Clinical Microbiology Reviews* **12**, 112–125.

HIDE, G., ANGUS, S. D., HOLMES, P. H., MAUDLIN, I. & WELBURN, S. C. (1998). *Trypanosoma brucei*: comparison of circulating strains in an endemic and an epidemic area of a sleeping sickness focus. *Experimental Parasitology* **89**, 21–29.

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-infective isolates. *Parasitology* **109**, 95–111.

JANSSEN, P., COOPMAN, R., HUYS, G., SWINGS, J., BLEEKER, M., VOS, P., ZABEAU, M. & KERSTERS, K. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology* **142**, 1881–1893.

JONES, C. J., EDWARDS, K. J., CASTAGLIONE, S., WINFIELD, M. O., SALA, F., VAN DE WIEL, C., BREDEMEIJER, G., VOSMAN, B., MATTHES, M. & DALY, A. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* **3**, 381–390.

KASSAI, T. (1998). Standardized nomenclature of animal parasitic disease (SNOAPAD). *Veterinary Parasitology* **29**, 299–326.

KUKLA, B. A., MAJIWA, P. A. O., YOUNG, C. J., MOLOO, S. K. & OLE-MIOYOI, O. K. (1987). Use of species-specific DNA probes for the detection and identification of trypanosome infections in tsetse flies. *Parasitology* **95**, 1–26.

LIN, J., KUO, J. & MA, J. (1996). A PCR-based DNA fingerprinting technique: AFLP for molecular typing of bacteria. *Nucleic Acids Research* **24**, 3649–3650.

LUN, Z. R., BRUN, R. & GIBSON, W. C. (1992). Kinetoplast DNA and molecular karyotypes of *Trypanosoma evansi* and *T. equiperdum* from China. *Molecular and Biochemical Parasitology* **50**, 189–196.

MACLEOD, A., TWEEDIE, A., WELBURN, S. C., MAUDLIN, I., TURNER, C. M. R. & TAIT, A. (2000). Minisatellite marker analysis of *Trypanosoma brucei*: Reconciliation of clonal, panmictic and epidemic population genetic structures. *Proceedings of the National Academy of Sciences, USA* **97**, 13442–13447.

MAGNUS, E., VERVOORT, T. & VAN MEIRVENNE, N. (1978). A

card agglutination test with stained trypanosomes (C.A.T.T) for the serological diagnosis of *T. b. gambiense* trypanosomiasis. *Annales de la Société Belge de Medécine Tropicale* **58**, 169–176.

MAJIWA, P. A. O., MAINA, M., WAITUMBI, J. N., MIHOK, S. & ZWEYGARTH, E. (1993). *Trypanosoma* (*Nannomonas*) *congolense*: molecular characterisation of a new genotype from Tsavo, Kenya. *Parasitology* **106**, 151–162.

MASIGA, D. K., TAIT, A. & TURNER, C. M. R. (2000). Amplified restriction fragment length polymorphism in parasite genetics. *Parasitology Today* **16**, 350–353.

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.

MELVILLE, S. E., LEECH, V., GERRARD, C. S., TAIT, A. & BLACKWELL, J. M. (1998). The molecular karyotype of the housekeeping chromosomes of *Trypanosoma brucei* and the assignment of chromosome markers. *Molecular and Biochemical Parasitology* **94**, 155–173.

MORLAIS, I., GREBAUT, P., BODO, J. M. & CUNY, G. (1998). Characterisation of trypanosome infections by polymerase chain reaction (PCR) amplification in wild tsetse flies in Cameroon. *Parasitology* **116**, 547–554.

MURRAY, H. W., PEPIN, J., NUTMAN, T. B., HOFFMAN, S. L. & MAHMOUD, A. A. (2000). Tropical medicine. *British Medical Journal* **320**, 490–494.

PEARSON, K. (1926). On the coefficient of racial likeness. *Biometrika* **18**, 105–117.

PEREZ, T., ALBOMOZ, J. & DOMINGUEZ, A. (1998). An evaluation of RAPD fragment reproducibility and nature. *Molecular Ecology* **7**, 1347–1357.

RETHER, B., DELMAS, G. & LAOUED, J. (1993). Isolation of polysaccharide-free DNA from plants. *Plant Molecular Biology Reports* **11**, 333–337.

RICHNER, D., SCHWEIZER, J., BETSCHART, B. & JENNI, L. (1989). Characterisation of West African *Trypanosoma (Trypanozoon) brucei* isolates from man and animals using isoenzyme analysis and DNA hybridisation. *Parasitology Research* **76**, 80–85.

SEMBLAT, J.-P., BONGIOVANNI, M., WAJNBERG, E., DALMASSO, A., ABAD, P. & CASTAGNONE-SERENO, P. (2000). Virulence and molecular diversity of parthenogenetic root-knot nematodes, *Meloidogyne* spp. *Heredity* **84**, 81–89.

SPOONER, D. M., TIVANG, J., NIENHUIS, J., MILLER, J. T., DOUCHES, D. S. & CONTRERAS, M. A. (1996). Comparison of four molecular markers in measuring relationships among the wild potato relatives Solanum section *Etuberosum* (subgenus *Potatoe*). *Theoretical and Applied Genetics* **92**, 532–540.

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., NOYES, H. A., DOVER, G. A. & GIBSON, W. C. (1999). The ancient and divergent origins of the human pathogenic trypanosomes, *Trypanosoma brucei* and *Trypanosoma cruzi*. *Parasitology* **118**, 107–116.

TAIT, A., BABIKER, E. A. & LE RAY, D. (1984). Enzyme variation in *Trypanosoma brucei* spp. I. Evidence for sub-speciation of *T. b. gambiense*. *Parasitology* **89**, 311–326.

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

URAKAWA, T. & MAJIWA, P. A. O. (2001). Physical and transcriptional organization of RNA genes of the savannah-type *Trypanosoma congolense*. *Parasitology Research* **87**, 431–438.

VAN DER PLOEG, L. H. T., BERNARDS, A., RIJSWIJK, F. A. M. & BORST, P. (1980). Characterization of the DNA duplication-transposition that controls the expression of two genes for variable surface glycoproteins in *Trypanosoma brucei*. *Nucleic Acids Research* **10**, 593–609.

VAUTERIN, L. A. & VAUTERIN, P. (1992). Computer-aided objective comparison of electrophoresis patterns for grouping and identification of microorganisms. *European Microbiology Journal* **1**, 37–41.

VOS, P., HOGERS, M., BLEEKER, M., REIJANS, M., VAN DE LEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M. & ZABEAU, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407–4414.

WAITUMBI, J. N. & MURPHY, N. B. (1993). Inter- and intra-species differentiation of trypanosomes by genomic fingerprinting with arbitrary primers. *Molecular and Biochemical Parasitology* **58**, 501–517.

WELBURN, S. C., VRE, E. M. F., COLEMAN, P. G., ODIIT, M. & MAUDLIN, I. (2001). Sleeping sickness: a tale of two diseases. *Trends in Parasitology* **17**, 19–24.

YOUNG, C. J. & GODFREY, D. G. (1983). Enzyme polymorphism and distribution of *Trypanosoma congolense* isolates. *Annals of Tropical Medicine and Parasitology* **77**, 467–481.

ZHANG, Z. Q. & BALTZ, T. (1994). Identification of *Trypanosoma evansi*, *Trypanosoma equiperdum* and *Trypanosoma brucei brucei* using repetitive DNA probes. *Veterinary Parasitology* **53**, 197–208.