Leishmania donovani is the only cause of visceral leishmaniasis in East Africa; previous descriptions of *L. infantum* and ''*L. archibaldi*'' from this region are a consequence of convergent evolution in the isoenzyme data

M. B. JAMJOOM[1]†, R. W. ASHFORD[1], P. A. BATES[1], M. L. CHANCE[1], S. J. KEMP[2], P. C. WATTS[2] *and* H. A. NOYES[2]*

[1] *Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK*
[2] *Animal Genomics Laboratory, School of Biological Sciences, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK*

(*Received 29 January 2004; revised 7 April 2004; accepted 8 April 2004*)

SUMMARY

Isoenzyme-based studies have identified 3 taxa/species/'phylogenetic complexes' as agents of visceral leishmaniasis in Sudan: *L. donovani*, *L. infantum* and ''*L. archibaldi*''. However, these observations remain controversial. A new chitinase gene phylogeny was constructed in which stocks of all 3 putative species isolated in Sudan formed a monophyletic clade. In order to construct a more robust classification of the *L. donovani* complex, a panel of 16 microsatellite markers was used to describe 39 stocks of these 3 species. All ''*L. donovani* complex'' stocks from Sudan were again found to form a single monophyletic clade. *L. donovani ss* stocks from India and Kenya were found to form 2 region-specific clades. The partial sequence of the glutamate oxaloacetate transaminase (GOT) gene of 17 *L. donovani* complex stocks was obtained. A single nucleotide polymorphism in the GOT gene appeared to underlie the isoenzyme classification. It was concluded that isoenzyme-based identification is unsafe for stocks isolated in *L. donovani* endemic areas and identified as *L. infantum*. It was also concluded that the name *L. archibaldi* is invalid and that only a single visceralizing species, *Leishmania donovani*, is found in East Africa.

Key words: aspartate aminotransferase, microsatellites, MLEE, kala azar, Ethiopia, chagasi.

INTRODUCTION

Visceral leishmaniasis (VL) is caused by members of the ''*Leishmania donovani* complex'' (Lainson & Shaw, 1987). The 2 main members are *L. donovani* (Laveran & Mesnil, 1903) in the Old World and *L. infantum* Nicolle, 1908 in the Old World and New World (*L. chagasi* is a junior synonym of *L. infantum* (Mauricio, Stothard & Miles, 2000)). ''*L. archibaldi*'' is a possible third member of the group that is found in Sudan (Pratlong *et al*. 2001). These 3 taxa are morphologically indistinguishable but are associated with different epidemiology, ecology and pathology. *L. infantum* and *L. donovani* have largely discrete geographical distributions; *L. infantum* is present around the Mediterranean basin and also in the New World, where it was probably introduced by European colonists (Courtenay *et al*. 2002; Mauricio *et al*. 2000). *L. donovani* is found in the Indian subcontinent and East Africa where it is associated with epidemics in adults as well as children. *L. donovani* is believed to be largely anthroponotic, although a species of mongoose has been found infected in Sudan (Elnaiem *et al*. 2001) and Dereure *et al*. (2003) reported a high (>70%) seroprevalence in dogs during 1998 and 1999 in one village in eastern Sudan. Dogs are the principal reservoir host for *L. infantum* but visceral disease in immunocompetent humans occurs only sporadically and almost exclusively in children (Ashford, 2000).

Multilocus enzyme electrophoresis-based studies (MLEE) have suggested that there may be some overlap in the range of *L. donovani* and *L. infantum*, particularly in East Africa where *L. infantum* is occasionally identified despite the predominance of *L. donovani* (Pratlong *et al*. 2001).

The classification of parasites from East Africa is further complicated by the supposed presence of ''*L. archibaldi*''. The name ''*Leishmania donovani varietas archibaldi*'' was originally applied to *Leishmania* from Sudan on dubious grounds by Castellani & Chalmers (1919). Brumpt (1936) later

* Corresponding author: Animal Genomics Laboratory, Biosciences Building, School of Biological Sciences, University of Liverpool, Crown Sreet, Liverpool L69 7ZB, UK. Tel: +44 (0)151 795 4512. Fax: +44 (0)151 795 4512. E-mail: harry@liv.ac.uk

† Current address: King Abdulaziz University, Department of Medical Parasitology, P.O. Box 80205, Jeddah 21589, Saudi Arabia.

Parasitology (2004), **129**, 399–409. © 2004 Cambridge University Press
DOI: 10.1017/S0031182004005955 Printed in the United Kingdom

included *L. donovani var. archibaldi* as a synonym of *L. donovani*. Nicoli (1963) subsequently revised the genus *Leishmania* and treated '*var. archibaldi*' as the subspecies *L. donovani archibaldi*. He applied the name to parasites from throughout East Africa and suggested, on the basis of its occurrence in rodents, that it should probably have specific status, even calling it *L. archibaldi* in the same article. The name "*L. archibaldi*" was more formally, but still provisionally, used by Lainson & Shaw (1987), who applied it to VL parasites from Sudan, suggesting that the parasite in Kenya and Somalia might be different.

With the advent of MLEE the possibility arose for a more objective taxonomic study of these parasites. Moreno *et al*. (1986) found that stocks that corresponded biologically with *L. donovani* could be distinguished from those corresponding with *L. infantum* on the basis of numerical analysis of isoenzyme data. The recent taxonomic re-descriptions of the *L. donovani* complex are, however, essentially based upon differences at a single 'diagnostic' enzyme locus since an inspection of the raw data shows that only glutamate oxaloacetate transaminase (GOT), segregates between these phenotypes. GOT_{100} was assigned to *L. infantum*, which was considered to be invariant for this allele (Moreno *et al*. 1986; Rioux *et al*. 1990). The *L. donovani* GOT enzyme variant had a mobility 13% greater than the *L. infantum* allele and consequently was designated GOT_{113}. All *L. donovani* stocks were believed to have the GOT_{113} allele. Thus the mobility of alleles of the GOT enzyme was considered sufficient to distinguish between these 2 species.

When a series of stocks became available from Sudan, it was found that both the *L. donovani* and *L. infantum* GOT character-states were represented, as well as a third state-GOT_{110}. Rioux *et al*. (1990) used the name *L. archibaldi* for the zymodeme showing this new character-state, and produced a classification of *Leishmania*, which included *L. archibaldi* as a 'phylogenetic complex', defined by its GOT mobility. A total of 3 *L. archibaldi* zymodemes have now been described, all with the GOT_{110} allele but with variations at other enzyme loci. Other Sudanese stocks were identified as *L. donovani* and *L. infantum* according to their respective GOT mobilities (Pratlong *et al*. 2001).

Thus the name *archibaldi* has been used by Rioux and co-workers' for stocks with the GOT_{110} allele, while Castellani and Chalmers applied the name to all Sudanese parasites, Nicoli used it for all east African parasites, and Lainson and Shaw applied it to all parasites from Sudan with possibly another taxon being responsible for leishmaniasis in Kenya and Somalia. The World Health Organization (Anon, 1990) included *L. archibaldi* in their list of *Leishmania* species, but did not indicate whether they used it in the sense of Lainson and Shaw, or that of Rioux *et al*., or indeed of any of the other workers.

While isoenzyme-based studies have indicated the presence of *L. infantum* and "*L. archibaldi*" in Sudan (Pratlong *et al*. 2001), this contrasts with DNA-based studies using Gp63, Internal Transcribed Spacer (ITS) and anonymous genomic sequences, that have found only a single clade of parasites in Sudan (El-Tai *et al*. 2001; Mauricio *et al*. 2001), and with sequences of anonymous genomic DNA, which failed to segregate with isoenzyme-based species (Lewin *et al*. 2002). Phylogeographical studies are usually conducted using neutral markers such as the housekeeping genes used for MLEE, however functional genes can also help reveal the effects of selective pressures. A PCR–RFLP study of the Gp63 gene (which is involved in invasion of the macrophage) of the *L. donovani* complex indicated geographical clustering that did not correlate with MLEE data (Guerbouj *et al*. 2001). In an attempt to clarify the classification of *Leishmania* from East Africa we have used the phylogeny of the chitinase gene which is involved in the interaction between the sandfly and the parasite and we have also developed a panel of 20 evolutionarily neutral microsatellite markers for *L. infantum* and *L. donovani* (Jamjoom *et al*. 2002*b*). Microsatellites are the mainstay of modern genetic studies and have largely replaced isoenzymes for many applications from individual-based studies through to shallow phylogenies (Sunnucks, 2000). Because they are PCR-based and allelic variants vary in length, microsatellite loci are well adapted for analysing large numbers of samples from small amounts of DNA and the data are, in principle, easily exchanged between laboratories. Most microsatellite loci are believed to be selectively neutral and are typically highly polymorphic. This makes it generally possible to discover more genetic diversity using microsatellites than with MLEE. We have now applied 16 of these markers to a collection of 41 stocks from Sudan, Kenya and India together with 5 *L. infantum* controls from Brazil, Spain, Algeria, and the UK (introduced). We have also sequenced parts of the chitinase and GOT genes in order to confirm our findings.

MATERIALS AND METHODS

Parasites

The stocks used in this study are listed in Table 1.

Sequencing the chitinase gene

The sequences of the *L. donovani* (AF009354) and *L. mexicana* chitinase genes were used to identify conserved regions of this gene and to design 2 pairs of primers, ChitA forward and reverse and ChitB forward and reverse (Table 2). PCR was carried out in

25 μl volumes in 0·2 ml PCR tubes with each reaction containing 2·0 mM $MgCl_2$, 75 mM Tris–HCl, 20 mM $(NH)_4SO_4$, 0·01% (v/v) Tween 20, 0·4 mM each dNTP, 25 ng genomic DNA, 40 pmol ChitA or ChitB primers and 2·5 units of *Taq* polymerase (ABgene). Thermal cycling conditions were: 94 °C for 5 min, then 30 times (94 °C for 60 sec, 58 °C for 60 sec and 72 °C for 60 sec) and finally 72 °C for 10 min. PCR products were cloned into pCR®-2.1-TOPO® (Invitrogen). The chitinase sequences were determined by cycle sequencing using CEQ DTCS chemistry (Beckman Coulter Inc.) and electrophoresis on a CEQ8000XL capillary sequencer (Beckman Coulter Inc). Sequences were aligned using MegAlign in the DNASTAR package and trees were compiled using DNAPARS and DNADIST in PHYLIP.

Microsatellite primers

A full description of the microsatellite loci and their PCR conditions is given in detail by Jamjoom *et al.* (2002*b*) except that for this work the forward primer for each locus was 5′-labelled with either D3 or D4 fluorescent dyes (PrOligo or Research Genetics) for visualization (Table 2). PCR products were pooled, according to their variation in allelic size and the fluorescent dye, along with a 400 base-pair (D1) size standard (Beckman Coulter Inc.) and separated by capillary electrophoresis through a denaturing acrylamide gel on a Ceq8000XL automated sequencer (Beckman Coulter Inc.). Microsatellite alleles were sized using the cubic model in the Ceq8000XL fragment analysis software. Alleles were converted to presence/absence data and the resultant table was used to construct phenetic trees using distance and Parsimony methods in PAUP 4.10b (Swofford, 1998). Deviation from genotypic linkage equilibrium among the entire set of locus-pair combinations was analysed using Fisher's exact test as implemented by GENEPOP v.3.3 (Raymond & Rousset, 1995).

Sequencing the GOT gene

Glutamate oxaloacetate transaminase is also known as aspartate aminotransferase. There are 2 aspartate aminotransferases in *T. brucei*, a nuclear encoded cytoplasmic enzyme and a nuclear encoded mitochondrial enzyme (Berger *et al.* 2001). The 2 genes have no significant homology as judged by BLAST search of one against the other or attempts to align them (HAN, unpublished observations). However when used as probes in a BLAST search of kinetoplastid sequences in GeneDB (http://www.genedb.org/), they both hit the same sequence in clone AC078900 from *L. major* chromosome 35, which contains the aspartate aminotransferase gene annotated as LmjF35.0820 in GeneDB. This sequence was used to design the following primers using the Primer3 program (Rozen & Skaletsky, 2000): GOT_1F 10-29, GOT_5R 1259-1277, and GOT_7R 1190-1211 (Table 2). The numbers after each primer name indicate the position of the primer in the predicted GOT gene. The primer pairs GOT_1F and GOT_5R and GOT_1F and GOT_7R both gave PCR products which were cloned into the pGEM EASY vector (Promega), cycle sequenced with the BigDyeV3.1 kit (ABI) and sequences were obtained on an ABI3100 capillary sequencer. Sequences were aligned in the SeqMan package of DNASTAR and scanned for polymorphisms by eye. A single nucleotide T→G polymorphism at position 823 was typed by direct sequencing of the PCR product of the primers GOT_SNP_F and GOT_SNP_R (Table 2).

RESULTS

Comparison of the chitinase gene sequence of the New World species *L. mexicana* with that of *L. donovani* revealed significant differences (93% identity). Therefore, this gene was tested for suitability for an intraspecific classification of the "*L. donovani* complex". Parsimony informative sites (199) were identified within the 1028 bp alignment, a classification was produced of 27 stocks (Fig. 1) and the species complexes could be identified with confidence. Both parsimony and distance methods produced trees with identical topology and only small differences (<8) in bootstrap values. There were 8 informative sites within the *L. donovani* complex and relationships within this group were less well supported. All stocks from Kenya and Sudan fell into 2 clear geographically defined groups with 56% and 58% bootstrap support, respectively. Stocks from Ethiopia and the Indian subcontinent clustered with the Sudanese stocks with 91% bootstrap support. *L. infantum* from Europe and the Mediterranean formed a clade with 85% support. "*L. infantum*" from Sudan clustered tightly with other Sudanese isolates and not with *L. infantum* from Europe and Mediterranean. Stocks labelled "*L. archibaldi*" were indistinguishable from other stocks from Sudan and did not form a distinct clade. The 8 informative sites in the alignment formed 4 haplotypes (Table 3). These haplotypes showed that a Kenyan *L. donovani* and Mediterranean *L. infantum* clade was supported by 5 positions, the *L. infantum* clade was supported by 2 positions and the Sudanese clade combining *L. donovani*, "*L. infantum*" and "*L. archibaldi*" stocks was supported by just a single position.

Although the chitinase data described above gave a classification that correlated with geography and biology better than classifications based on MLEE, the amount of the data underlying this conclusion was limited. Therefore, data from presence/absence of alleles at 16 microsatellite loci for 41 stocks of

Table 1. *Leishmania* stocks

(*Leishmania* stocks used in the study. The first column shows the putative species of the studied stocks, species names based on MLEE but reassigned in this study are shown in inverted commas. The second column presents the WHO code. This indicates host: MHOM Humans; IDUB Phlebotomus duboscqi; IMRT Phlebotomus martini; MCAN Canis familiaris; MNYC Nyctomys sp.; MRHO Rhombomys sp./Country: BD Bangladesh; BR Brazil; BZ Belize; DZ Algeria; ES Spain; ET Ethiopia; GB Great Britain; IL Lebanon; IN India; IQ Iraq; IR Iran; IT Italy; KE Kenya; NP Nepal; SD Sudan; SN Senegal; SU Soviet Union; XX Unknown/Year of isolation/Laboratory code, if 2 or more laboratories have given codes then these are separated by a semi colon. The third column is the zymodeme where known. Zymodemes are indicated as typed by the Laboratoire d'Ecologie Medicale et Pathologie Parasitaire (LEM), Montpellier (MON); n.d. indicates zymodeme is not determined. The fourth column is the clinical disease indicated by CL for cutaneous leishmaniasis and VL for visceral leishmaniasis. The fifth column shows the GOT[823] genotype as determined by direct sequencing of PCR products, the traces from these genes can be seen at www.genomics.liv.ac.uk/animal/supplementary_data/jamjoom/traces.pdf. The sixth column shows the studies completed for each stock: Ch, Chitinase gene sequence and GenBank Accession Number; Ms, microsatellite data; GOT, GOT sequencing of cloned PCR products with GenBank Accession number, * indicates a previous WHO reference strain (Chance & Walton, 1982).)

Species	WHO code	Zymodeme	Disease	GOT[823] Genotype	Analysis performed
L. major	MHOM/IL/80/Friedlin;Fv1	n.d.	CL		Ch. (AY518224), Ms.
L. major	MRHO/SU/59/LV39;Pstrain*	MON-4	CL		Ch. (AY518225), Ms.
L. major	MHOM/SU/60/LV356;LRC-L38	n.d.	CL	GG	Ch. (AY518226), Ms.
L. major	IDUB/SN/XX/LV599;DK57	n.d.	CL		Ch. (AY518227), Ms.
L. major	MHOM/SN/XX/LV622;DK72	n.d.	CL		Ch. (AY518228), Ms.
L.major	MRHO/IR/75/ER	n.d.	CL		Ch. (AY518229), Ms.
L. tropica	MHOM/SU/60/LV357;LRC-L39*	n.d.	CL	GG	Ch. (AY518230), Ms.
L. tropica	MHOM/IQ/66/LV556*	n.d.	CL		Ch.(AY518258), Ms.
L. donovani	MHOM/BD/97/LDON;BG1	MON-2	VL		Ch. (AY518231), Ms., GOT (AY529106)
L. donovani	MHOM/ET/67/HU3;LV9*	MON-18	VL	GG	Ch. (AY518233), Ms., GOT (AY529111)
L. donovani	MHOM/SD/90/D83;LEM2131	MON-18	VL	GG	Ch. (AY518234), Ms., GOT (AY529105)
L. donovani	MHOM/SD/90/D92;LEM2132	MON-18	VL	GG	Ch. (AY518235), Ms., GOT (AY529107)
L. donovani	MHOM/SD/90/2828;LEM2139	MON-18	VL	GG	Ch. (AY518236), Ms., GOT (AY529107)
L. donovani	MHOM/SD/90/D100;LEM2140	MON-18	VL	GG	Ch. (AY518237), Ms., GOT
L. donovani	MHOM/SD/92/18	MON-18	VL		Ms.
L. donovani	MHOM/IN/00/Devi;LEM138	MON-2	VL		Ms.
L. donovani	MHOM/KE/75/H9;LEM496	MON-32	VL		Ch. (AY518238), Ms.
L. donovani	MHOM/KE/55/LRC-L53;LEM707	MON-36	VL		Ms.
L. donovani	IMRT/KE/62/LRC-L57;LEM719	MON-37	VL		Ch. (AY518239), Ms.
L. donovani	MHOM/ET/84/ADDIS164;LEM980	MON-83	VL		Ch. (AY518240), Ms.
L. donovani	MHOM/KE/73/MRC74	LON-51	VL		Ms.
L. donovani	MHOM/ET/00/Hussen	LON-42	VL		Ch. (AY518241), Ms.
L. donovani	MHOM/NP/02/NEP2	n.d.	VL		Ch. (AY518242), Ms., GOT (AY529112)
L. donovani	MHOM/NP/02/NEP3	n.d.	VL		Ch. (AY518243), Ms., GOT (AY529112)
L. donovani	MHOM/NP/02/NEP5	n.d.	VL		Ch. (AY518244), Ms., GOT (AY529114)
L. donovani	MHOM/IN/75/K13;LV613	n.d.	VL		Ms.
L. donovani	MHOM/IN/77/Munni;LV614	n.d.	VL		Ms.
L. donovani	MHOM/IN/77/Agindra;LV615	n.d.	VL		Ms.
L. donovani	MHOM/IN/77/Pandey;LV616	n.d.	VL		Ms.
L. donovani	MHOM/IN/77/Shibchandra;LV617	n.d.	VL		Ms.
L. donovani	MHOM/IN/77/Devi;LV619	n.d.	VL		Ms.
L. donovani	MHOM/IN/77/Mandrika;LV621	n.d.	VL		Ms.

L. donovani	MHOM/IN/77/Rai;LV636	n.d.	VL		Ms.
L. donovani "*L. archibaldi*"	MHOM/SD/90/D75;LEM2134	MON-82	VL		Ch. (AY518245), Ms.
L. donovani "*L. archibaldi*"	MHOM/SD/90/2655;LEM2135	MON-82	VL	GT	Ch. (AY518246), Ms., GOT (AY529109)
L. donovani "*L. archibaldi*"	MHOM/SD/90/D99;LEM2137	MON-82	VL	GT	Ch. (AY518247), Ms.
L. donovani "*L. archibaldi*"	MHOM/SD/91/D1783;LEM2211	MON-82	VL	GT	Ch. (AY518248), Ms.
L. donovani "*L. archibaldi*"	MHOM/SD/92/82	MON-82	VL		Ch. (AY518232), Ms.
L. infantum	MCAN/GB/96/LV755	MON-1	VL	TT	Ch. (AY518249), Ms.
L. infantum	MHOM/BR/76/150406;M4192	MON-1	VL		Ch. (AY518250), Ms.
L. infantum	MCAN/ES/98/LEM-935;JPC;M5	MON-1	VL	TT	Ch. (AY518251), Ms., GOT (AY529110)
L. donovani "*L. infantum*"	MHOM/SD/91/D1809;LEM2213	MON-30	VL	TT	Ch. (AY518252), Ms., GOT (AY529115)
L. donovani "*L. infantum*"	MHOM/SD/97/LEM3435	MON-30	VL	TT	Ch. (AY518253), Ms., GOT (AY529116)
L. donovani "*L. infantum*"	MHOM/SD/97/LEM3431	MON-30	VL	TT	Ch. (AY518254), Ms., GOT
L. donovani "*L. infantum*"	MHOM/SD/97/LEM3441	MON-30	VL	TT	Ch. (AY518255), Ms., GOT
L. donovani "*L. infantum*"	MHOM/SD/92/30	MON-30	VL		Ch. (AY518256), Ms.
L. infantum	MHOM/ES/81/BCN1;LEM307	MON-29	VL		Ms.
L. infantum	MHOM/DZ/83/;LEM425	MON-80	VL		Ms.
L. infantum	MHOM/IT/85/ISS175;LEM1733	MON-111	VL		Ms.
L. amazonensis	MHOM/BR/97/M2269	n.d.	CL	GG	Ch. (AY518257), GOT
L. donovani	MHOM/SD/62/1-S		VL		Ch., AF009354
L. mexicana	MNYC/BZ/62/M379		CL		Ch., AY572789

L. infantum, *L. donovani* and "*L. archibaldi*" were generated and analysed by parsimony (Fig. 2) and minimum evolution using mean character differences. There were only 3 minor differences between the parsimony and minimum evolution trees, all within the large Sudanese clade, these were: (1) LEM980 from western Ethiopia clustered inside the Sudanese clade; (2) the polytomy at the root of the Sudanese clade in the parsimony tree was resolved into 2 clades in the minimum evolution tree but the grouping of taxa was consistent between the 2 trees; (3) LEM2211 could be resolved from the large polytomy that is adjacent to the parsimony tree but not in the mean distance tree. A table of all the microsatellite data is available at www.genomics.liv.ac.uk/animal/supplementary_data/jamjoom/ms_data.xls. All stocks from the Indian subcontinent, Kenya and Sudan fell into 3 clear geographically defined groups with 85–100% bootstrap support by parsimony and 98–100% support by the distance method. Classical *L. infantum* stocks from Europe, Brazil and Algeria formed a fourth discrete clade with 96–100% bootstrap support. Once again all stocks from Sudan clustered in a single clade. The "*L. infantum*" from Sudan clustered tightly with other Sudanese isolates and not with other *L. infantum* from Europe or Brazil. Stocks labelled "*L. archibaldi*" were indistinguishable from other stocks from Sudan and did not form a distinct clade. Two well-supported clades were identifiable within the 17 Sudanese stocks. Members of all 3 putative species were found in each of the clades.

Since the tree was not rooted, it is only possible to draw limited conclusions about the relationships between the geographical clades. The Indian and Kenyan clades clustered together with 77% and 89% bootstrap support. However, no rooting point is possible that would create the clade that was observed in the chitinase phylogeny composed of the Kenyan *L. donovani* and the classical *L. infantum* stocks. Consequently the two phylogenies are inconsistent in this respect.

The microsatellite dataset was scanned for loci that might be in linkage disequilibrium and hence on the same chromosome. Table 4 shows the locus pairs for which there was evidence of disequilibrium but after applying a Bonferoni correction for the large number of tests involved only LIST7-037 × LIST7-035 remained significant.

Although 16 microsatellite loci were used in this classification it is possible to identify members of the different clades with fewer microsatellite markers. Table 5 shows the microsatellite alleles from 3 microsatellite loci that segregated with specific geographical regions.

As noted in the Introduction, the MLEE-based assignment of stocks to named species within the "*L. donovani* complex" is based exclusively on the mobility of the enzyme GOT. Therefore, to examine

Table 2. Sequences of 16 primer pairs for polymorphic, dinucleotide microsatellite loci (Jamjoom *et al.* 2002*b*) and primers used to amplify and sequence regions of the chitinase and GOT genes in the *Leishmania donovani* 'complex'

(Also shown are single nucleotide polymorphism primers GOT_SNP. Dye refers to the fluorescent label used for allelic visualization on a CEQ8000XL (Beckman Coulter Inc.).)

Locus	5′Dye	Forward	Reverse
LIST7-021	D4	CCGAATACACAAGCCTCCTC	TCAGGCTTCGTCGTTTCTTT
LIST7-022	D3	GTCGCTCTGTCTCTGTGTGC	TCCGCATTTTCCTCTCCTT
LIST7-023	D3	CTTTGCGTTGCGCACTAA	GCTTGTGTTCCGTGTGTGTT
LIST7-024	D4	TAAACTGCATGGTCCCCTCT	ACAAGCACCATCATCCACAT
LIST7-025	D3	GGAGTCGTCTCTCTGTTACGC	ATCGCGTGCATGGGTATT
LIST7-027	D4	CTCTCTCGTCACCACAGCAC	AGGGGACAAGACACAGATGG
LIST7-028	D3	CACTCCACTGCGTTGGATA	CTTTGACCGCCGTTCTTT
LIST7-029	D4	GCAGAGCTTCTGCTTGGATT	GCATTGCTGTTCTCATCCAC
LIST7-031	D3	CACTGGTGGAAATAGAAAGACT	GGAGAACTAAAACGAGCAGCA
LIST7-032	D3	CTAGAGGCGTGCGGATGTA	TCGCAGTTTTCGGTCCA
LIST7-033	D4	CATTGCTGAGTGCTGCTAGTG	ATGAGCGTACTGGGCACAC
LIST7-035	D3	AAAGGTATGATACGCCTGTGG	ACCGCAAAGAACGGACAT
LIST7-036	D4	CTCTCTCGTCACCACAGCAC	TCCCTCTCGTTGGTGAAGTT
LIST7-037	D4	ATGCTGAGCCCATCAAGACT	GATGTCCCCGTTTACTCCAA
LIST7-039	D3	CTCGCACTCTTTCGCTCTTT	AGACGAGAGGAACGGAAAA
LIST7-040	D4	GCAGAGCGAGACACACAGAC	GTGCACGTTGATGTGCTTCT
ChitA		TTCTGCAGCTGGCGTGTCTTGTA	TGAGCTTGCGGCGGTGGTCCTG
ChitB		TCCTTCCCCGCTCTTTACTGTCTT	TACCCGTCATACTCGCTCAAT
GOT_1F 10-29		CCCATCACTCACGATTCACA	
GOT_5R 1259			AGCACAGATGTCCACGCA
GOT_7R 1190-1211			GAAGATTCAGGCACAAGCTCCC
GOT_SNP		CGAATCCATCACGCTCTTTAC	TGTTCTTCGACTCCGCCTAC

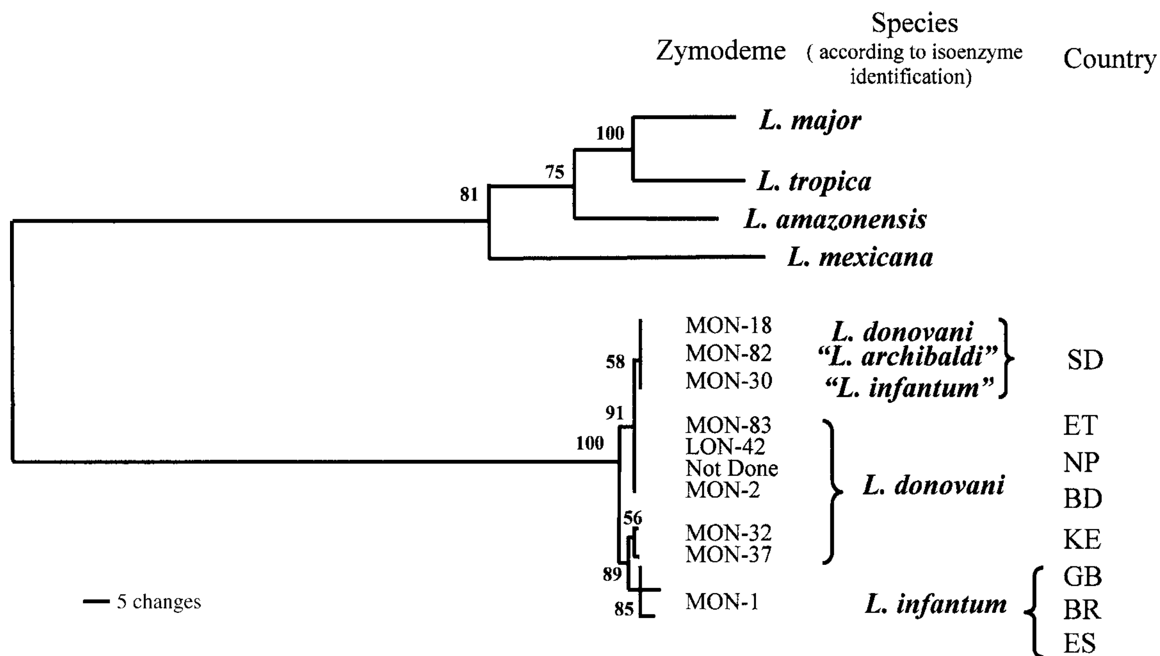

Fig. 1. Phylogeny based on partial sequencing (1028 nucleotides) of the chitinase gene. The phylogeny was compiled using the parsimony option in PAUP and bootstrap values were compiled using 100 bootstrap replicates. "*Leishmania infantum*" and "*Leishmania archibaldi*" from Sudan clearly cluster with *Leishmania donovani* and not with classical *Leishmania infantum* zymodeme MON1. The taxa included are indicated in Table 1, together with the GenBank Accession Numbers of the sequences. Country abbreviations are as in the legend to Table 1.

the validity of this distinction, the sequence of 1127 of the 1260 base pairs of this enzyme was obtained for 16 of the stocks under investigation (Table 1). Within this sequence only a single informative polymorphism was identified that give rise to an amino acid change. This T→G polymorphism at position 823 would cause a change in amino acid 275 from tyrosine in the GOT_{100} ("*L. infantum*") stocks to aspartate in

Table 3. Summary of the chitinase haplotypes within the *Leishmania donovani* 'complex'

(The number of stocks sequenced in each group is indicated in parentheses.)

	Position of the polymorphism in the chitinase gene							
Stocks	129	158	212	315	435	663	847	1034
Sudanese (14) and Ethiopian (1)	T	C	G	G	T	C	C	A
Kenya (2)	C	C	A	T	C	G	T	A
L. infantum (MON-1) (3)	C	T	A	T	C	G	T	G
Other stocks (Ethiopia, Nepal, Bangladesh) (6)	C	C	G	G	T	C	C	A
	Sudanese	*L. infantum*	Kenya and *L. infantum*					***L. infantum***

the GOT_{113} ("*L. donovani*") stocks. The additional negative charge provided by the aspartate residue in such *L. donovani* stocks would be expected to cause the GOT of these stocks to migrate faster towards the anode and consequently give rise to the faster observed mobility of GOT. All stocks that had been identified as *L. infantum* by MLEE had the 275^{tyr} allele including 2 "*L. infantum*" stocks from Sudan (LEM2213 and LEM3435). Three stocks that had been identified as "*L. archibaldi*" by MLEE had the 275^{tyr} allele whilst one had the 275^{asp}. Eighteen stocks were sequenced directly from PCR products (3 "*L. archibaldi*", 6 *L. infantum*, 6 *L. donovani*, 1 *L. tropica L. major* and 1 *L. amazonensis*). All 6 *L. infantum* as identified by MLEE had the "T" allele, all 6 *L. donovani* had the "G" allele and all 3 "*L. archibaldi*" were heterozygotes (Table 1). The other 3 species typed had the "G" allele suggesting that this is the primitive condition. Electropherograms of these data are available from www.genomics.liv.ac.uk/animal/supplementary_data/jamjoom/traces.pdf.

DISCUSSION

Validity of "L. archibaldi" *and* L. infantum *in Sudan*

The chitinase sequence-based classification and the microsatellite-based classification both show that *Leishmania* from Sudan form a single clade. This is inconsistent with the present, isoenzyme-based classification which indicates that parasites from Sudan belong to 3 distinct clades, but consistent with previous DNA-based studies (El-Tai *et al.* 2001; Mauricio *et al.* 2001). Our GOT data show that the isoenzyme-based identification of *L. donovani* and *L. infantum* is dependent on a single nucleotide polymorphism and confirms previous observations made by Mauricio and Miles at the London School of Hygiene and Tropical Medicine (Mauricio personal communication). Since both the 275^{tyr} and the 275^{asp} alleles of GOT are clearly viable, it is quite conceivable that the T→G transition that underlies the amino acid polymorphism could have occurred independently on more than one occasion. Such an independent mutation could have given rise to the "*L. infantum*" stocks found in Sudan quite separately from the same mutation that gave rise to the *L. infantum* around the Mediterranean. Consequently although GOT_{100} may be an adequate marker for identifying *L. infantum* and *L. donovani* in the Mediterranean, India and the New World, it is not suitable for this purpose in Africa.

We therefore propose that all stocks that have been identified by MLEE as *L. infantum* and that were isolated within the *L. donovani* endemic area of East Africa should be renamed, as *L. donovani*. We further suggest that the species "*L. archibaldi*" should be considered invalid, since it did not form a monophyletic clade; even if *L. archibaldi* does form a monophyletic clade within the Sudanese *L. donovani*, this would render *L. donovani* itself polyphyletic and hence invalid.

Further, if genuine *L. infantum* does exist in Sudan then it might be expected that parasites identified as *L. infantum* would have a similar epidemiology to those around the Mediterranean. A recent study of zymodemes of parasites isolated from 52 humans and 20 dogs in Sudan found no epidemiological difference in the diseases caused by the 3 putative species (Dereure *et al.* 2003). However, reanalysis of the published data shows that a significantly larger proportion of dogs were infected with *L. infantum* zymodeme parasites than *L. donovani* zymodeme parasites ($P<0{\cdot}001$, Chi Squared test). Although this would appear to support the MLEE-based hypothesis that there is *L. infantum* in East Africa, we do not believe that this observation is sufficient evidence in itself to warrant the retention of species status for *L. infantum* zymodeme parasites in East Africa.

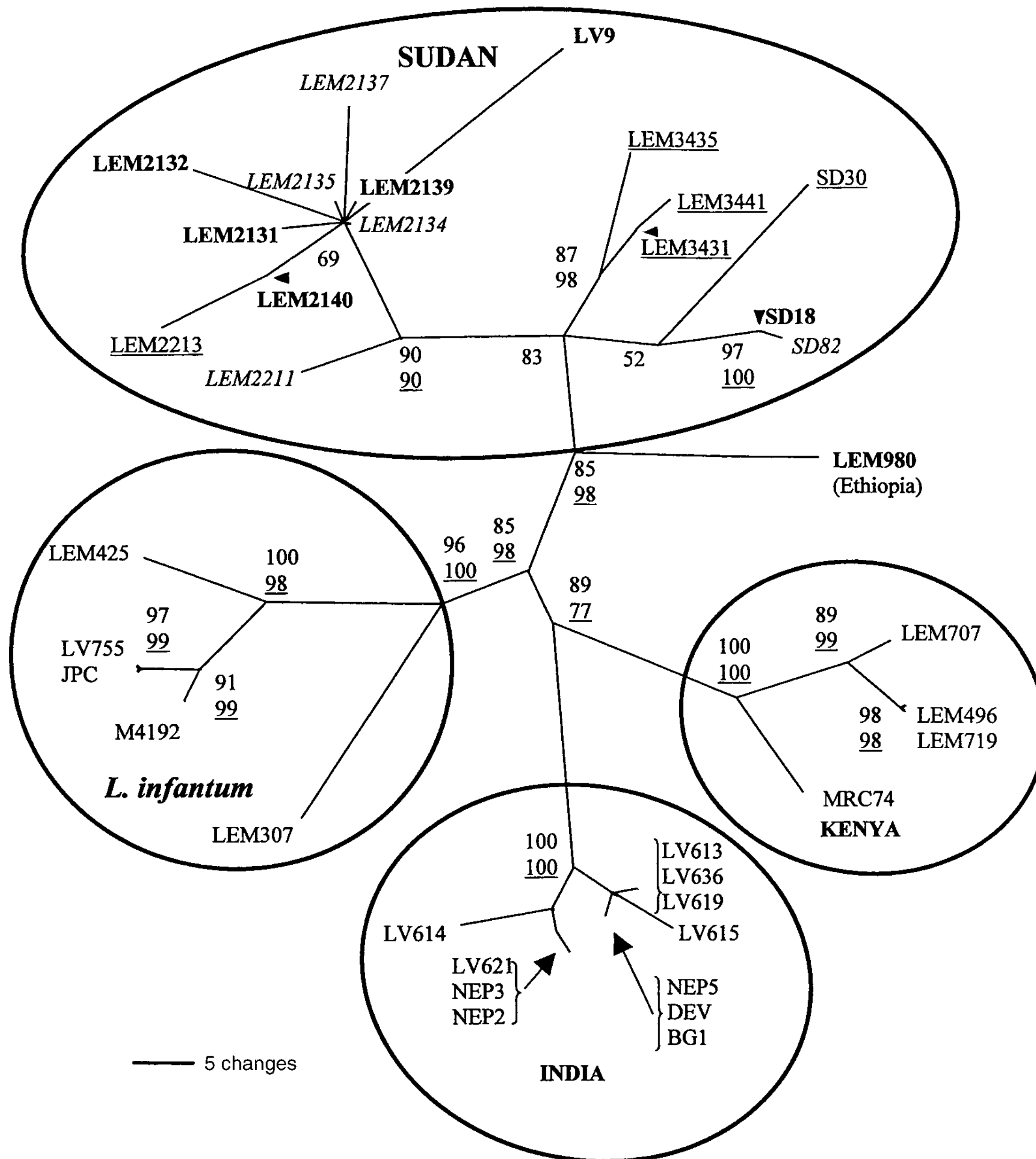

Fig. 2. Dendrogram of microsatellite data compiled using the parsimony option in PAUP. Bootstrap values were calculated using 100 bootstrap replicates. Bootstrap values from the distance tree are shown underlined underneath the bootstrap values for parsimony where the same clade was found in both trees. The MLEE-based species identification of Sudanese stocks is indicated by the typeface. *Leishmania donovani* in bold; *Leishmania infantum* underlined; *Leishmania archibaldi* in italics. These 3 putative taxa do not form distinct clades within the Sudanese clade nor does "*Leishmania infantum*" from Sudan cluster with *Leishmania infantum* from around the Mediterranean. Two distinct groups of parasites are found in Sudan but these groups do not have any correlation with MLEE identifications.

Inconsistencies with isoenzyme data

The data presented here are inconsistent with isoenzyme data in 2 respects. Firstly, we have presented molecular evidence that shows that "*L. archibaldi*" is a heterozygote for an amino acid that would cause a change in charge of the enzyme and would account for the difference in mobility of *L. donovani* and *L. infantum*. Heterozygotes at GOT have not been reported on MLEE gels by the Montpellier groups although Mauricio *et al.* (2001) refer to MON82 as a putative ASAT (= GOT) hybrid. The GOT enzyme is a homodimer and consequently the heterozygote would be expected to appear as a triplet with a strong middle band in MLEE gels. Given the small difference in relative mobility of the *L. donovani* GOT_{113} enzyme compared with the *L. infantum* GOT_{100} it is possible that a heterozygote with a strong intermediate band could be scored as a homozygote GOT_{110}. Secondly, 2 GOT loci are reported for *L. donovani* complex stocks in MLEE studies but our data only account for one of these loci. There is

Table 4. Microsatellite loci that may be in linkage disequilibrium

(Probability that 6 locus pairs are in linkage disequilibrium and hence presumed to be on the same chromosome. Note that LIST7-021, LIST7-027, LIST7-036 and possibly LIST7-022 appear to be on the same chromosome. * After applying a Bonferoni correction for the large number of tests that were made on this dataset only the LIST7-037 × LIST7-035 pair remains significant.)

Locus 1	Locus 2	Probability
LIST7-021	LIST7-027	0·975
LIST7-024	LIST7-029	0·970
LIST7-021	LIST7-036	0·964
LIST7-027	LIST7-036	0·990
LIST7-036	LIST7-022	0·980
LIST7-037	LIST7-035	1·00*

Table 5. Microsatellite loci with alleles that are associated with particular geographical regions

(The definition used to include *Leishmania infantum* stocks was parasites with a GOT^{100} allele and that had been isolated outside the *Leishmania donovani* endemic areas. The Sudan group includes samples with a GOT^{100} allele but since they were isolated within the *Leishmania donovani* endemic areas were excluded from the *Leishmania infantum* group. Further studies may show that some or all of these alleles are not restricted to particular geographical regions. These loci all appear to be independent since there was no evidence of linkage among this set of locus-pair combinations using Fisher's exact test as implemented by GENEPOP v.3.3 (Raymond & Rousset, 1995).)

	Microsatellite loci		
Region	LIST7-027	LIST7-036	LIST7-023
Sudan	187; 189; 191; 193; 195	242; 244; 247; 249	171; 173; 175
Kenya	183	238	153
India	185; 184	241	151
L. infantum	177; 179	232; 234	151

conflicting evidence from MLEE and genome sequencing projects for the number of GOT loci that generate products visible on MLEE gels. Both GOT loci are scored as having enzymes with the same relative mobility in all zymodemes in the *L. donovani* complex and in 10 out of 11 *L. major* zymodemes (Maazoun *et al.* 1986; Pratlong *et al.* 2001; Rioux *et al.* 1990). Two aspartate aminotransferases (= GOT) are annotated in GeneDB for *L. major* Friedlin strain (LmjF24.0370 and LmjF35.0820), however, these 2 genes have no sequence homology when one is used in a BLAST search against the other, and they have different predicted isoelectric points (pH 7·0 and pH 7·9) and hence would be expected to have different mobilities by MLEE. Since these 2 known GOT loci are on different chromosomes it is unlikely that their relative mobilities are so tightly linked that they change from GOT_{100} in *L. infantum* to GOT_{113} in *L. donovani* together. Consequently one of these genes is probably not being expressed at a level detectable by MLEE in promastigotes (the parasite stage usually studied by MLEE). In contrast in *L. tropica*, *L. gerbili*, *L. aethiopica* and Namibian stocks all zymodemes have 2 GOT enzymes visible on MLEE gels with substantial differences in relative mobility (15–87%) and which vary independently (Lanotte, Rioux & Serres, 1986). For these species 2 loci are clearly being detected on the MLEE gels and this is consistent with the *L. major* genome sequence data for 2 GOT loci. Gene expression studies might help to resolve these apparent anomalies. Although sequence data are presented here that would account for the MLEE data on *L. infantum* and *L. donovani*, further work will be required to confirm that the "*L. archibaldi*" heterozygotes that are predicted by the sequence data would account for the intermediate mobility of the "*L. archibaldi*" GOT isoenzyme allele. If the MLEE phenotype of "*L. archibaldi*" is shown to be a heterozygote then this could be interpreted as indicating that "*L. archibaldi*" is a hybrid between *L. infantum* and *L. donovani*. However it has been shown here that the "*L. infantum*" in Sudan is in fact *L. donovani* and probably the consequence of an independent and relatively recent mutation in the GOT gene in Sudan from local *L. donovani* stocks, an example of convergence. The microsatellite data (www.genomics.liv.ac.uk/animal/supplementary_data/jamjoom/ms_data.xls) was scanned for evidence that "*L. archibaldi*" stocks as a group may be heterozygotes of other Sudanese stocks but none was found in the 16 loci tested. Since the "*L. archibaldi*" stocks were scattered in 2 distinct clades in the Sudanese group with high bootstrap support it would appear that "*L. archibaldi*" is not a genetic entity at all and if individual stocks are hybrids then they are different hybrids that have arisen on more than one occasion. Further work will also be required to confirm that the gene sequenced here is expressed in promastigotes at a level that could be detected by isoenzymes and to discover the expression pattern of the other aspartate aminotransferase.

Geographical grouping of clades

The microsatellite classification identified strongly supported geographically based groups. However, additional stocks will have to be typed to confirm these observations since a previous study on this scale also identified regional groups but also found that some strains from Portugal, Lebanon, Iran and Italy clustered with a clade that was otherwise similar to the Sudanese clade identified here (Mauricio *et al.* 2001). Although there were clear regional groups, the relationships between the groups could not all be

resolved with confidence. This is partly because the tree was unrooted. Only 2 of the microsatellite primers isolated from *L. donovani* are informative within the potential outgroup *L. major*, and conversely only 2 microsatellites isolated from *L. major* are polymorphic in *L. donovani* (Jamjoom *et al.* 2002*a*; Jamjoom *et al.* 2002*b*). It is therefore not possible to root the microsatellite tree using the currently available loci. However, the Kenyan and Indian clades appeared to be most closely related to each other, consistent with an African rather than Asian origin for the Indian group as previously suggested by Ashford (1986).

Identification of L. infantum

New methods are now required to identify *L. infantum*. A total of 5 alleles at 3 loci were found exclusively in *L. infantum* but studies that include larger numbers of *L. infantum* stocks will be required to confirm that these alleles are diagnostic for this species. Unfortunately this means that currently there is no single marker for identifying *L. infantum*, although it is possible to classify stocks as *L. infantum* by using the microsatellites described here or Gp63 intergenic region PCR–RFLP (Mauricio *et al.* 2001).

A new diagnostic test is not required in Sudan

The confirmation that visceral leishmaniasis in Sudan is only caused by a single clade of *L. donovani* makes it possible to apply a single serological test to patients from Sudan with greater confidence. The possibility that the most widely used existing test, which is based on an *L. donovani* antigen, might be missing infections caused by other species can now be disregarded.

We are grateful to Professor Dedet and Dr Pratlong, Université de Montpellier I, France; Dr Maingon, Keele University, UK; Dr Mauricio, London School of Hygiene and Tropical Medicine, UK and Dr Javadian for the supply of parasite stocks used in this study. The study was funded in part by King Abdulaziz University, Saudi Arabia, which we gratefully acknowledge. We thank P. Ambrose and D. Reynolds for valuable technical assistance and two anonymous referees for constructive criticisms.

REFERENCES

ANON. (1990). Control of the leishmaniases. Report of a WHO Expert Committee. *World Health Organization Technical Report Series* **793**, 1–158.

ASHFORD, R. W. (1986). Speculations on the origins and evolution of Old World *Leishmania* systems. In *Leishmania : Taxonomy and Phylogeny* (ed. Rioux, J. A.), pp. 257–264. IMEEE, Montpellier.

ASHFORD, R. W. (2000). The leishmaniases as emerging and reemerging zoonoses. *International Journal for Parasitology* **30**, 1269–1281.

BERGER, L. C., WILSON, J., WOOD, P. & BERGER, B. J. (2001). Methionine regeneration and aspartate aminotransferase in parasitic protozoa. *Journal of Bacteriology* **183**, 4421–4434.

BRUMPT, E. (1936). Precis de Parasitologie. Masson et Cie, Paris.

CASTELLANI, A. & CHALMERS, A. (1919). *Manual of Tropical Medicine*, Vol. 3. Balliere, Tindall and Cox, London.

CHANCE, M. L. & WALTON, B. C. (1982). In *Proceedings of a Workshop held at the Pan American Health Organization 9–11 December, 1980*, pp. 115–121. UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, Geneva.

COURTENAY, O., QUINNELL, R. J., GARCEZ, L. M. & DYE, C. (2002). Low infectiousness of a wildlife host of *Leishmania infantum*: the crab-eating fox is not important for transmission. *Parasitology* **125**, 407–414.

DEREURE, J., EL SAFI, S. H., BUCHETON, B., BONI, M., KHEIR, M. M., DAVOUST, B., PRATLONG, F., FEUGIER, E., LAMBERT, M., DESSEIN, A. & DEDET, J. P. (2003). Visceral leishmaniasis in eastern Sudan: parasite identification in humans and dogs; host–parasite relationships. *Microbes and Infection* **5**, 1103–1108.

EL-TAI, N. O., EL-FARI, M., MAURICIO, I. L., MILES, M. A., OSKAM, L., EL-SAFI, S. H., PRESBER, W. H. & SCHONIAN, G. (2001). *Leishmania donovani*: intraspecific polymorphisms of Sudanese isolates revealed by PCR-based analyses and DNA sequencing. *Experimental Parasitology* **97**, 35–44.

ELNAIEM, D. A., HASSAN, M. M., MAINGON, R., NURELDIN, G. H., MEKAWI, A. M., MILES, M. & WARD, R. D. (2001). The Egyptian mongoose, *Herpestes ichneumon*, is a possible reservoir host of visceral leishmaniasis in eastern Sudan. *Parasitology* **122**, 531–536.

GUERBOUJ, S., VICTOIR, K., GUIZANI, I., SERIDI, N., NUWAYRI-SALTI, N., BELKAID, M., BEN-ISMAIL, R., LE-RAY, D. & DUJARDIN, J. C. (2001). Gp63 gene polymorphism and population structure of Leishmania donovani complex: influence of the host selection pressure? *Parasitology* **122**, 25–35.

JAMJOOM, M. B., ASHFORD, R. W., BATES, P. A., KEMP, S. J. & NOYES, H. A. (2002*a*). Polymorphic microsatellite repeats are not conserved between *Leishmania donovani* and Leishmania major. *Molecular Ecology Notes* **2**, 104–106.

JAMJOOM, M. B., ASHFORD, R. W., BATES, P. A., KEMP, S. J. & NOYES, H. A. (2002*b*). Towards a standard battery of microsatellite markers for the analysis of the *Leishmania donovani* complex. *Annals of Tropical Medicine and Parasitology* **96**, 265–270.

LAINSON, R. & SHAW, J. J. (1987). Evolution, classification and geographical distribution. In *The leishmaniases in biology and medicine* (ed. Peters, W. & Killick-Kendrick, R.), pp. 1–120. Academic Press, London.

LANOTTE, G., RIOUX, J. A. & SERRES, E. (1986). Approche cladistique du genre Leishmania Ross, 1903. A propos de 192 souches originaires de l'Ancien Monde, analyse numerique de 50 zymodemes identifies par 15 enzymes et 96 isoenzymes. In *Leishmania. Taxonomie et phylogenese. Applications Eco-Epidemiologiques* (ed. Rioux, J. A.), pp. 269–288. IMEEE, Montpellier.

LEWIN, S., SCHONIAN, G., EL TAI, N., OSKAM, L., BASTIEN, P. & PRESBER, W. (2002). Strain typing in *Leishmania donovani* by using sequence-confirmed amplified region analysis. *International Journal for Parasitology* **32**, 1267–1276.

MAAZOUN, R., PRATLONG, F., LANOTTE, G. & RIOUX, J. A. (1986). Le complexe *Leishmania major*. A propos de

l'analyse numerique de 35 souches identifiées par le methode enzymatique. In *Leishmania ; Taxonomy and Phylogeny ; Applications to ecology and epidemiology* (ed. Rioux, J. A.), pp. 119–128. IMEEE, Montpellier.

MAURICIO, I. L., GAUNT, M. W., STOTHARD, J. R. & MILES, M. A. (2001). Genetic typing and phylogeny of the *Leishmania donovani* complex by restriction analysis of PCR amplified gp63 intergenic regions. *Parasitology* **122**, 393–403.

MAURICIO, I. L., STOTHARD, J. R. & MILES, M. A. (2000). The strange case of *Leishmania chagasi*. *Parasitology Today* **16**, 188–189.

MORENO, G., RIOUX, J. A., LANOTTE, G., PRATLONG, F. & SERRES, E. (1986). Le complexe *Leishmania donovani s.l.* Analyse enzymatique et traitement numerique, individualisation du complexe *Leishmania infantum*, corollaires biogeographiques et phyletiques, apropos de 146 souches originaires de l'Ancien et du Nouveau Monde. In *Leishmania : taxonomie phylogenese* (ed. Rioux, J. A.), pp. 105–118. IMEEE, Montpellier.

NICOLI, R. (1963). Le genre *Leishmania* R. Ross, 1903. *Bulletin de la Societe de Pathologie Exotique* **56**, 408–416.

PRATLONG, F., DEREURE, J., BUCHETON, B., EL SAFI, S., DESSEIN, A., LANOTTE, G. & DEDET, J. P. (2001). Sudan : the possible original focus of visceral leishmaniasis. *Parasitology* **122**, 599–605.

RAYMOND, M. & ROUSSET, F. (1995). GENEPOP, Version 1·2. Population genetics software for exact tests and ecumenicisms. *Journal of Heredity* **86**, 249–249.

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**, 111–125.

ROZEN, S. & SKALETSKY, H. (2000). Primer3 on the WWW for general users and biologist programmers. In *Bioinformatics Methods and Protocols* (ed. Krawetz, S. A. & Misener, S.), pp. 365–386. Humana Press, Totowa, NJ.

SUNNOCKS, (2000). Efficient genetic markers for population biology. *Trends in Ecology and Evolution* **15**, 199–203.

SWOFFORD, D. L. (1998). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.