Genetic typing and phylogeny of the *Leishmania donovani* complex by restriction analysis of PCR amplified gp63 intergenic regions

I. L. MAURICIO*, M. W. GAUNT, J. R. STOTHARD† *and* M. A. MILES

Pathogen Molecular Biology and Biochemistry Unit, Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

SUMMARY

Protozoan parasites of the *Leishmania donovani* complex (*L. donovani*, *L. infantum*/*L. chagasi*) are causative agents of visceral leishmaniasis. To understand phylogeny and taxonomy within this group better we have developed 2 new polymerase chain reaction-linked restriction fragment length polymorphism (PCR–RFLP) analyses of the major surface protease (msp or gp63) intergenic (ITG) regions. We have named this approach **m**sp **i**ntergenic region **R**FLP **t**yping (MIRT). One intergenic region lies between the constitutive *msp* (*mspC*) and stationary phase *msp* (*mspS4*) genes (ITG/CS) and the other between multicopy logarithmic phase *msp* (*mspL*) genes (ITG/L). The markers generated robust and congruent phylogenies, identifying 5 genetic clusters within *L. donovani*. One cluster was synonymous with *L. infantum* (*L. chagasi*); clusters strongly correlated with isoenzyme typing and some with geographical origin. These genetic groups may be important for epidemiological and clinical studies. The congruence of the groups identified indicates suitability of these genomic targets for taxonomic studies. Furthermore, subgroups of *L. donovani* were of equivalent phylogenetic status to *L. infantum*. No evidence was found to support the existence of *L. archibaldi*. It is likely to be necessary in future to re-evaluate the taxonomic status of *L. donovani* or of *L. infantum*, as discrete species.

Key words: Leishmania donovani complex, Leishmania infantum, gp63 intergenic regions, PCR–RFLP analysis, genetic diversity, phylogenetics.

INTRODUCTION

The aetiological agents of visceral leishmaniasis (VL), or kala-azar, are protozoan parasites grouped under the *Leishmania donovani* complex: *L. donovani* (Laveran & Mesnil, 1903) Ross, 1903, *L. infantum* Nicolle, 1908; *L. chagasi* Cunha & Chagas, 1937, and *L. archibaldi* Castellani and Chalmers, 1919. Taxonomy within this complex had apparently been settled with discrimination of *L. infantum* and *L. donovani* by isoenzyme electrophoresis (Rioux *et al.* 1990), and demonstration that *L. chagasi* is synonymous with *L. infantum* (Mauricio, Stothard & Miles, 2000). However, the existence of another species in the complex (*L. archibaldi*) is still debatable. The relationships between African and Indian *L. donovani* are also not clear: there are overall similarities in epidemiology, isoenzyme characterization and molecular data, accompanied by striking differences that might reflect different human host populations. We have previously found that *L. donovani* has more genetic diversity than *L.*

* Corresponding author: Tel: +44 020 7927 2399. Fax: +44 020 7636 8739. E-mail: isabel.mauricio@lshtm.ac.uk
† Present address: Biomedical Parasitology, Department of Zoology, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

infantum and have identified 3 clades with some degree of geographical association (Mauricio *et al.* 1999).

The use of polymerase chain reaction-linked restriction fragment length polymorphism (PCR–RFLP) analysis for *Leishmania* was initiated with the development of intergenic region typing (IRT) of the ribosomal internal transcribed spacer (ITS) (Cupolillo *et al.* 1995), but several other intergenic regions are available in *Leishmania*. The genes coding for a glycoprotein of 63 Dalton (gp63), or major surface protease (msp), have been extensively studied and their intergenic regions have been described for New World *L. infantum* (Roberts *et al.* 1993). *Msp* genes have always been found in multiple copies (Maingon *et al.* 1990; Medina Acosta, Karess & Russell, 1993; Steinkraus *et al.* 1993; Victoir *et al.* 1995; Voth *et al.* 1998) but the best description of organization was for New World *L. infantum* (Roberts *et al.* 1993). Three distinct types of genes were reported to be expressed preferentially in the logarithmic or stationary growth phases, or to be constitutively expressed (Ramamoorthy *et al.* 1992) and were accordingly named as *mspL*, *mspS* and *mspC*, although in fact they are all transcribed at all stages of development Roberts *et al.* 1993). The non-coding regions and the 3′ half of the coding
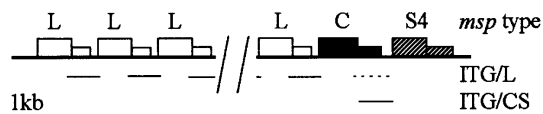
Fig. 1. Schematic representation of amplified gp63/ITG regions based on Roberts *et al.* (1993) and Ramamoorthy *et al.* (1992). C – constitutive *msp*; L – logarithmic phase *msp*; S4 – type 4 stationary phase *msp*. Larger and smaller blocks indicate, respectively, coding and non-coding regions. // represents an undetermined repeat number of *mspL* genes. ITG/L and ITG/CS PCR products are indicated by lines below, and the dashed line of the ITG/L PCR products represents the intergenic region between *mspC* and *mspS4* which is amplified simultaneously but was not used in the restriction analyses.

regions are type specific, such that specific primers can be designed for each gene type. The gene cluster has an array with an indeterminate number of *mspS*, followed by an array with an indeterminate number of *mspL* of at least 4 subtypes and then the single copy gene, *mspC*, followed by a single copy of *mspS4* (Fig. 1).

Here we describe the development of a new PCR–RFLP, based on 2 separate intergenic regions (ITG) between gp63 coding genes, and named **m**sp **i**ntergenic region **R**FLP **t**yping (MIRT). We demonstrate that this technique is useful for determining genetic diversity and for phylogenetic analysis of the *L. donovani* complex.

MATERIALS AND METHODS

Leishmania *strains*

DNA was extracted using phenol/chloroform (Kelly, 1993) from several *L. donovani* complex strains (Table 1) with various geographical origins and reflecting zymodeme diversity. Most strains had been cryopreserved and are not likely to have undergone prolonged culture. Other *Leishmania* strains were used to test PCR specificity. From the Old World: *L. aethiopica*, MHOM/ET/70/L96 and MHOM/ET/72/L100; *L. major*, MHOM/SU/73/5ASKH; *L. tropica*, MHOM/SU/74/K27. From the New World: *L. amazonensis*, MHOM/BR/73/M2269; *L. braziliensis*, MHOM/BR/84/LTB300; *L. guyanensis*, MHOM/BR/75/M4147; *L. lainsoni*, MHOM/BR/81/M6426; *L. mexicana*, MHOM/BZ/82/BEL21; *L. panamensis*, MHOM/PA/71/LS94; *L. peruviana*, MHOM/PE/84/LC39.

Amplification of intergenic regions

For amplification of ITGs within the gp63 gene array, new primers were designed with methods previously applied for coding *msp* sequences (Mauricio *et al.* 1999). Primers C9F (5′ ggc tcc cga cgt gag tta 3′, *mspC* specific, stop codon underlined) and C1R (reverse generic primer, 5′ **ggg ccc ggg** cga

cag cag cga tga ctg 3′, with *Xma*I restriction site insertion in bold) were designed specifically to amplify the ITG between *mspC* and *mspS4* genes (ITG/CS, Fig. 1), with an expected size of 1·75 kilobase pairs (kb). Primers C10F (5′**ggg aag ctt** acg tac agc gtg cag gtg 3′, *msp* generic with *Hind*III restriction site insertion in bold) and C1R were used to amplify all *msp* ITGs (Fig. 1), with 3 expected fragments of approximately 1·6, 2·0 and 4·5 kb, the smaller corresponding to ITG/L (between *mspL* genes). Primers C10F, C1R and C9F begin amplification from positions 1697, 210 and 2055, respectively, in the *mspC* GenBank M80671 (Ramamoorthy *et al.* 1992), such that each PCR product contains 96 base pairs (bp) of the 5′ end of the coding region, within the signal peptide, and 245 bp (ITG/L) or 11 bp (ITG/CS) of the 3′ end coding region, including the primer sequences. Restriction sites were added to C10F and C1R primers to differentiate between different ITG/L alleles by cloning, if necessary. PCR reactions of 100 μl included 125 ng genomic DNA, 2·5 mM MgCl$_2$, 0·2 mM dNTP, 50 μM of each primer and 1 U *Taq* polymerase (Bioline), in the buffer recommended by the manufacturer, with 5 % (ITG/L) or 10 % (ITG/CS) dimethylsulphoxide. Thermal cycling was: 1 min at 95 °C, then, for ITG/CS, 30 cycles of 1 min at 94 °C, 1 min at 65 °C and 2 min at 72 °C, and for ITG/L, 35 cycles of 1 min at 95 °C, 1 min at 70 °C and 1·5 min at 72 °C, with, in each case, a final step of 10 min at 72 °C. Identity of the amplified fragments was confirmed by partial sequence of PP75 ITG/CS (GenBank AJ009908) (Mauricio *et al.* 1999) and by hybridization of *L. infantum* IPT 1 and *L. donovani* DD8 ITG/L (positive for 2·0 kb fragment and negative for ITG/CS) with a 138 bp probe obtained by PCR amplification (data not shown) using primer C10F and C9R 5′ **ggg gga tcc** ctg gca cac ctc cac gta 3′ (with a *Bam*HI restriction site, in bold) and with the 3′ end at nucleotide 1816 of the New World *L. infantum mspC* (GenBank M80671).

The desired PCR products were separated in a 0·8 % agarose gel, purified with a QIAEX II™ gel purification Kit (Qiagen), dissolved in Tris–HCl, pH 8·5, adjusting the final concentration to be similar for all samples.

Restriction of PCR fragments

Appropriate restriction enzymes were chosen after screening published sequences for restriction sites and trial restriction experiments. Both ITG/CS and ITG/L PCR products were digested with *Alu*I, *Bst*UI, *Cfo*I, *Hae*III, *Msp*I, *Rsa*I and *Taq*I in a total volume of 10 μl, with approximately 0·5 μg of DNA and 5 U of restriction enzyme in the recommended buffer, overnight at the recommended temperature. Restriction fragments were separated at 250 V for

Table 1. *Leishmania donovani* complex strains

(Key to WHO codes: host/country of isolation/year of isolation/strain name. Hosts: insect (I), mammal (M), *Homo sapiens* (MHOM), *Canis familiaris* (MCAN), *Cerdocyon thous* (MCER), *Phlebotomus martini* (IMAR), *Arvicanthis* sp (MARV), *Meriones* sp (MMER). Countries: China (CN), Spain (ES), Ethiopia (ET), India (IN), Iran (IR), Italy (IT), Kenya (KE), Lebanon (LB), Panama (PA), Portugal (PT), Saudi Arabia (SA), Sudan (SD), former Soviet Union (SU), Tunisia (TU). 0000 is unknown year. Cryobanks: L – London School of Hygiene and Tropical Medicine (LSHTM), UK; I – Instituto de Higiene e Medicina Tropical, Lisbon, Portugal; M – Laboratoire d'Ecologie Medicale et Pathologie Parasitaire (LEMPP), Montpellier; B – Instituto Evandro Chagas, Belem, Brazil; LV – Liverpool School of Tropical Medicine and Hygiene, UK. Zymodemes are indicated as typed by the LEMPP laboratory (MON) and the LSHTM laboratory (LON). In parentheses are partial LON zymodeme typing (Evans *et al.* 1984; Le Blancq, 1986; Miles *et al.* 1980 *a,b*) in thin starch gel electrophoresis using alanine aminotransferase (ALAT, EC 2.6.1.2), aspartate aminotransferase (ASAT, EC 2.6.1.1), glucosephosphate isomerase (GPI, EC 5.3.1.9), malate dehydrogenase (MDH, EC 1.1.1.37), mannosephosphate isomerase (MPI, EC 5.3.1.8), nucleoside hydrolase using inosine as substrate (NH, EC 3.2.2.1), proline dipeptidase (PEPD, EC 3.4.13.9) and 6-phosphogluconate dehydrogenase (decarboxylating) (6PGD, EC 1.1.1.44), data not shown. Putative MON zymodemes equivalent to the LON zymodemes are also shown in parentheses.)

| Species | Code | WHO code | Zymodeme | | Cryobank of origin |
			LON	MON	
L. infantum (syn. *L. chagasi*)	I	MHOM/TU/1980/IPT 1*	49	I	L
	C	MHOM/BR/1974 /PP75*	49	1	L
	I25	IARI/PT /1989/IMT 171	(201)	24	M
	I26	IARI/PT/1989/IMT 172	(201)	24	M
	I3	MCAN/FR/1982/Pharoah	49	(1)	L
	C6	MCAN/BR/1989/Dog 124	—	—	L
	C13	MCER/BR/1981/M6445	(49)	(1)	B
	C12	MCER /BR/1989/M12084	(49)	(1)	B
	I16	MHOM/CN/1980/Strain A	—	34	L
	I17	MHOM/CY/1963/L53	(202)	—	L
	I4	MHOM/ES/1987/Lombardi	—	—	L
	C17	MHOM/HN/1988/HN122	—	—	LV
	C20	MHOM/HN/1993/HN336	—	—	LV
	I31	MHOM/MT/1985/Buck	(201)	78	L
	C2	MHOM/PA/1978/WR285	(49)	(1)	L
	I6	MHOM/PT/1982/IMT 104	—	1	L
L. donovani	D	MHOM/IN/1980/DD8*†	41	2	L
	D28	MHOM/ET/1972/Gebre 1††	50	82	L
	D21	IMAR/KE /1962/LRC L57	44	37	L
	D22	MARV/SD/1962/LRC L64	48	(30)	L
	D24	MCAN/IT/1976/Dora	50	(82)	L
	D25	MCAN/KE/0000/D2	45	—	L
	D26	MHOM/CN/0000 /Wangjie 1	(206)	35	L
	D13	MHOM/ET/0000/Hussen	42	(83/31)	L
	D12	MHOM/ET/0000/Ayele 5	52	(83/31)	L
	D27	MHOM/ET/0000/Ayele 8	56	(32)	L
	D1	MHOM/ET/1967/HU3 (LV9)*	46	18	L
	D14	MHOM/ET/1982/Bekele	42	—	L
	D15	MHOM/ET/1984/Addis 142	(56)	(32)	L
	D29	MHOM/ET/1984/Addis 164	(205)	83	L
	D4	MHOM/IN/1982/Patna 1†	(41)	(2)	L
	D6	MHOM/IN/1977/Chowd X	—	—	L
	D7	MHOM/IN/1979/STL 1–79†	41	(2)	L
	D8	MHOM/IN/1982/Nandi 1†	41	(2)	L
	D30	MHOM/KE/0000/Neal R1	56	(32)	L
	D2	MHOM/KE/1967/MRC (L)3	—	—	L
	D3	MHOM/KE/1973/MRC74	51	(2)	L
	D31	MHOM/KE/1975/Mutinga H9	56	32	L
	D16	MHOM/KE/1980/Ndandu 4A	44	(37)	L
	D17	MHOM/LB/1984/Salti 4	(46)	(18)	L
	D34	MHOM/PT/1992/IMT 180	(46)	18	I
	D32	MHOM/SA/1981/Jeddah KA	42	31	L
	D11	MHOM/SA/1987/VL6	(42)	(83/31)	L
	D9	MHOM/SA/1987/VL23	(42)	(83/31)	L
	D10	MHOM/SA/1987/VL29	(42)	(83/31)	L
	D18	MHOM/SD/0000/Khartoum	46	18	L
	D33	MHOM/SD/1982/Gilani	48	30	L
	D19	MHOM/SD/1985/A22	(48)	(30)	L
	D20	MHOM/SD/1987/UGX-marrow	(205)	(83/31)	L
	D5	MMER/IR/1996/Mesh 17	50	(82)	L

* WHO reference strain.　† *L. donovani s.s.*　†† *L. archibaldi.*
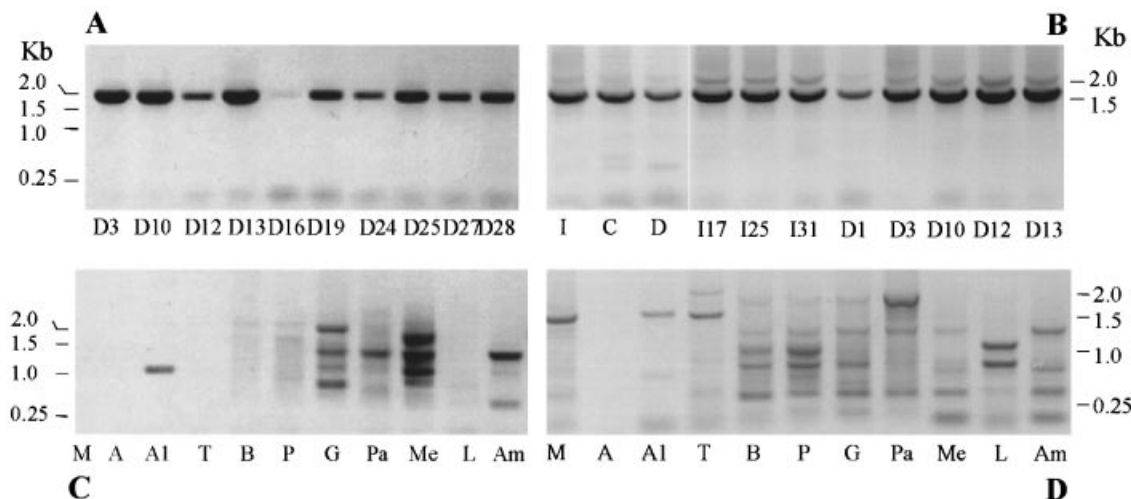
Fig. 2. Negative image of amplification products from (A) ITG/CS and (B) ITG/L within the *Leishmania donovani* complex (stronger band is ITG/L), and specificity of (C) ITG/CS and (D) ITG amplification. Only at very high DNA concentrations was a faint fragment seen, similar in size to ITG/L (1·5 Kb), but only for Old World *Leishmania* species, and stronger bands of a different size. Reference strains are: A – *L. aethiopica* (L100), A1 – *L. aethiopica* (L96), Am – *L. amazonensis*, B – *L. braziliensis*, C – American *L. infantum* (PP75); D – *L. donovani* (DD8 and others); G – *L. guyanensis*, I – *L. infantum* (IPT 1); L – *L. lainsoni*, M – *L. major*, Me – *L. mexicana*, P – *L. panamensis*, Pe – *L. peruviana*, T – *L. tropica*. *L. donovani* strains are D1 (HU3), D3 (MRC 74), D10 (VL29), D12 (Ayele 5), D13 (Hussen), D16 (Ndandu 4A), D19 (A22), D24 (Dora), D25 ('D2'), D27 (Ayele 8), D28 (Gebre 1). *L. infantum* strains are I17 (L53), I25 (IMT 171), I31 (Buck).
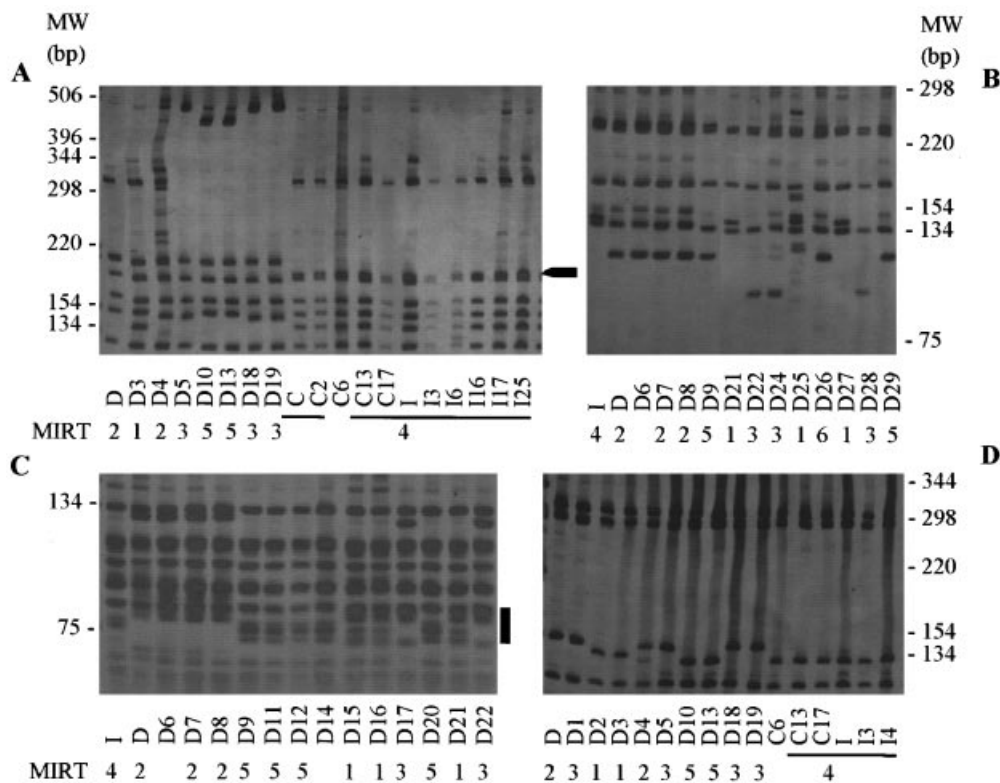
Fig. 3. Restriction profiles generated by enzymes (A) *Cfo*I and (B) *Hae*III on ITG/CS amplification products, and enzymes (C) *Cfo*I and (D) *Msp*I on ITG/L amplification products. *Cfo*I generated specific profiles for *L. infantum*/*L. chagasi* for band sizes indicated by an arrow upon ITG/CS and by a bar upon ITG/L. The MIRT types corresponding to each strain and restriction pattern are indicated.
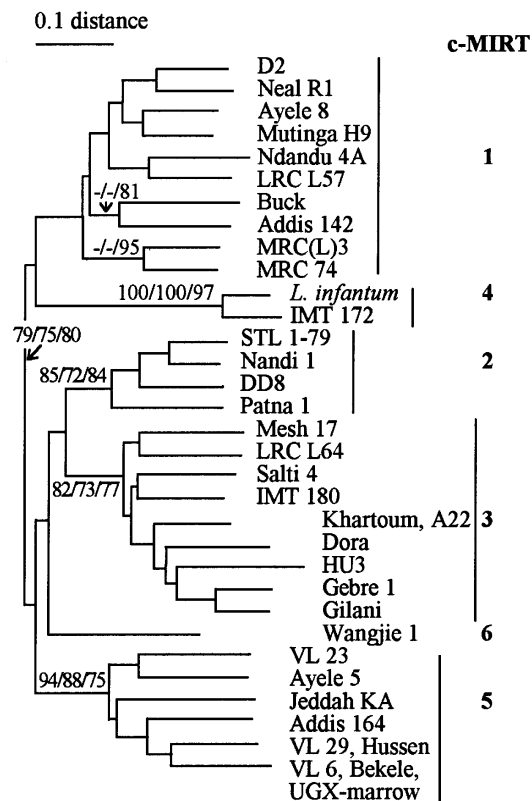
Fig. 4. Phylogenetic analysis of the *Leishmania donovani* complex using c-MIRT data. Neighbour-joining unrooted tree (PHYLIP) using Jaccard distances. Clusters are highlighted and corresponding cMIRT are shown. Bootstrap values higher than 70 % are shown for main branches for neighbour-joining/Wagner parsimony/maximum likelihood (10000 puzzling steps). Bootstrap values for UPGMA were very similar to neighbour-joining and are not shown. *L. infantum* refers to strains: IPT 1, PP75, HN122, HN336, IMT 104, L53, M12084, M6445, Pharoah, Strain A, WR285, which were all indistinguishable.

2 h in 6 % polyacrylamide gels, 3·5 % cross-linking, and silver stained (Vidigal *et al*. 1994). Reproducibility of restriction profiles was tested by using IPT 1 and DD8 replicates in each restriction batch and gel.

Phylogenetic analysis of fragment data

The RFLP fragments were scored for presence (1) or absence (0), except fragments larger than 900 bp for which resolution was poorer. Data matrices of each PCR–RFLP were analysed separately or pooled with packages SYN-TAX-pc 5.0 (Podani, 1993), PHYLIP (Felsenstein, 1993), PAUP* 4.0 (Swofford, 2000) and Puzzle 4.0.2. (Strimmer & von Haeseler, 1996). Phenograms were built from a distance matrix (Jaccard index: $\sqrt{(1-a/(a+b+c))}$) using single linkage and the unweighted pair group method with arithmetic averages (UPGMA), with SYN-TAX, and neighbour-joining (PHYLIP) clustering algorithms. Cophenetic correlation was used to verify the fit of tree derived distances (SYN-TAX)

to the observed pairwise distance matrix. Neighbour-joining and UPGMA analyses (NEIGHBOR, PHYLIP), and cladistic analysis, using Wagner or Dollo parsimony (MIX or DOLLOP, PHYLIP), were performed using taxa jumbling and global rearrangements. Maximum likelihood was obtained by quartet puzzling (Puzzle 4.0.2) incorporating an invariant rate parameter. Bootstrap analyses were performed for 100 (PHYLIP) or 1000 (PAUP*, using Upholt mean character distance for restriction sites with an heuristic search) replications, for parsimony and clustering methods, respectively, to estimate the robustness of lineage divergence.

RESULTS

Amplification of intergenic regions ITG/CS and ITG/L

ITG/CS amplification generated a single product of approximately 1·75 kb, as estimated by comparison with a 1 kb DNA ladder on agarose gel electrophoresis (Fig. 2A) which was specific for the *L. donovani* complex (Fig. 2C). ITG/L amplification generated up to 3 bands, a pattern apparently specific for the *L. donovani* complex, the smaller of which corresponded to ITG/L and was approximately 1·6 kb as estimated by comparison with a 1 kb DNA ladder on agarose gel electrophoresis (Fig. 2B). For some other *Leishmania* species it was possible to amplify a band of similar size to that of the *L. donovani* ITG/L PCR products among many non-specific bands (Fig. 2D) if a high DNA concentration was used (> 5 mg/l).

Restriction of intergenic regions ITG/CS and ITG/L

Restriction of ITG/CS (c-MIRT) and ITG/L (l-MIRT) generated rich band profiles with enough diversity to be used for strain comparisons. Enzymes *Cfo*I (Fig. 3), *Alu*I and *Rsa*I, used upon gp63 ITG/CS and *Cfo*I used upon ITG/L produced profiles specific for *L. infantum*. The restriction profiles of *L. infantum* strains were identical for both intergenic regions, except for small differences in strains IMT 172 and HN336, whilst diversity with *L. donovani* was much greater with at least 4 distinct profile types (Fig. 3). The putative *L. infantum* strain Buck had profiles resembling *L. donovani* more than *L. infantum*. In total 64 and 55 binary characters were used for c-MIRT and l-MIRT, respectively.

Phylogenetic analysis of c-MIRT

In all c-MIRT trees, 5 clusters of strains were identified, which had also been defined in the Jaccard distance matrix and were classified into 6 c-MIRT types (Fig. 4). UPGMA (cophenetic correlation coefficient 0·977) and single linkage generated

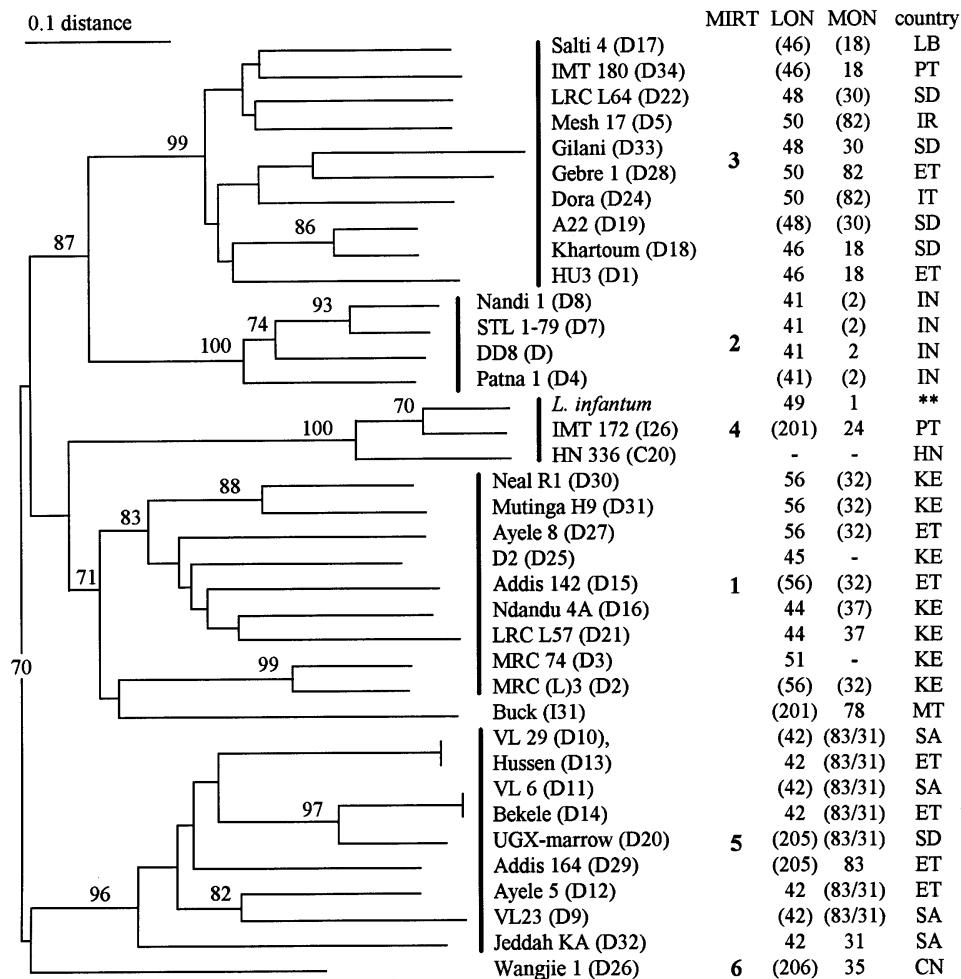Table 2. Pairwise distance (Jaccard index) among *Leishmania donovani* complex strains for c-MIRT data (lower matrix) and l-MIRT data (upper matrix)
(Values lower than 65 % of the maximum distance are shaded. VL29 is indistinguishable from Hussen; *L. infantum* refers to IPT1, PP75, WR285, M12084, NH122, Pharoah, Strain A, L53, which were all indistinguishable; VL6 is indistinguishable from Bekele and UGX-marrow for c-MIRT; Ayele 5 is indistinguishable from UGX-marrow for l-MIRT.)

	STL 1-79	Nandi 1	DD8	Patna 1	Wangjie 1	Salti 4	IMT 180	Mesh 17	Gebre 1	Gilani	LRC L64	Dora	HU3	Khartoum	A22	Ayele 8	MRC (L)3	MRC 74	Ndandu 4A	D2	Neal R1	Mutinga H9	Addis 142	LRC L57	Buck	VL23	Jeddah KA	Ayele 5	VL29	Addis 164	VL6	L. inf.	IMT 172	HN336
STL 1-79	-	0.000	0.167	0.167	0.528	0.506	0.528	0.469	0.558	0.506	0.469	0.506	0.447	0.494	0.474	0.436	0.397	0.397	0.500	0.457	0.512	0.506	0.500	0.413	0.518	0.550	0.571	0.518	0.429	0.447	0.550	0.459	0.459	0.480
Nandi 1	0.167	-	0.167	0.167	0.528	0.506	0.528	0.469	0.558	0.506	0.469	0.506	0.447	0.494	0.474	0.436	0.397	0.397	0.500	0.457	0.512	0.506	0.500	0.413	0.518	0.550	0.571	0.518	0.429	0.447	0.550	0.459	0.459	0.480
DD8	0.285	0.232	-	0.000	0.512	0.488	0.512	0.447	0.544	0.488	0.447	0.488	0.424	0.474	0.453	0.463	0.368	0.368	0.482	0.482	0.494	0.528	0.482	0.442	0.500	0.535	0.591	0.500	0.403	0.424	0.535	0.435	0.435	0.459
Patna 1	0.363	0.324	0.324	-	0.512	0.488	0.512	0.447	0.544	0.488	0.447	0.488	0.424	0.474	0.453	0.463	0.368	0.368	0.482	0.482	0.494	0.528	0.482	0.442	0.500	0.535	0.591	0.500	0.403	0.424	0.535	0.435	0.435	0.459
Wangjie 1	0.512	0.488	0.488	0.488	-	0.537	0.558	0.506	0.516	0.537	0.544	0.457	0.528	0.488	0.512	0.516	0.494	0.494	0.494	0.452	0.544	0.571	0.494	0.537	0.281	0.354	0.442	0.281	0.424	0.320	0.354	0.397	0.397	0.363
Salti 4	0.494	0.469	0.469	0.469	0.544	-	0.274	0.226	0.345	0.477	0.226	0.312	0.277	0.277	0.320	0.403	0.512	0.512	0.426	0.471	0.522	0.516	0.426	0.477	0.596	0.619	0.608	0.596	0.571	0.577	0.619	0.556	0.556	0.535
IMT 180	0.474	0.494	0.494	0.494	0.528	0.281	-	0.277	0.312	0.408	0.277	0.349	0.320	0.320	0.358	0.477	0.535	0.535	0.494	0.532	0.463	0.500	0.494	0.537	0.615	0.636	0.564	0.615	0.591	0.596	0.636	0.577	0.577	0.556
Mesh 17	0.512	0.488	0.488	0.442	0.558	0.354	0.320	-	0.349	0.436	0.229	0.226	0.164	0.164	0.232	0.457	0.474	0.474	0.477	0.477	0.488	0.522	0.477	0.522	0.571	0.596	0.584	0.571	0.541	0.550	0.596	0.524	0.524	0.500
Gebre 1	0.500	0.518	0.556	0.556	0.550	0.392	0.363	0.469	-	0.345	0.408	0.271	0.383	0.383	0.413	0.511	0.528	0.528	0.489	0.447	0.408	0.403	0.489	0.532	0.577	0.601	0.522	0.577	0.615	0.558	0.601	0.603	0.603	0.584
Gilani	0.518	0.535	0.535	0.535	0.528	0.358	0.324	0.442	0.167	-	0.436	0.378	0.413	0.463	0.442	0.565	0.512	0.512	0.546	0.511	0.383	0.431	0.546	0.552	0.564	0.590	0.506	0.564	0.603	0.506	0.590	0.591	0.591	0.571
LRC L64	0.550	0.528	0.528	0.488	0.558	0.354	0.320	0.316	0.418	0.387	-	0.316	0.164	0.281	0.232	0.457	0.474	0.474	0.477	0.516	0.488	0.522	0.477	0.482	0.571	0.596	0.584	0.571	0.541	0.550	0.596	0.524	0.524	0.500
Dora	0.518	0.494	0.535	0.535	0.528	0.358	0.392	0.442	0.285	0.324	0.442	-	0.277	0.277	0.320	0.494	0.469	0.469	0.471	0.426	0.436	0.477	0.471	0.516	0.528	0.558	0.544	0.528	0.571	0.506	0.558	0.556	0.556	0.535
HU3	0.584	0.564	0.528	0.564	0.522	0.413	0.442	0.482	0.358	0.320	0.436	0.387	-	0.232	0.167	0.482	0.453	0.453	0.500	0.500	0.469	0.506	0.500	0.506	0.556	0.584	0.571	0.556	0.524	0.535	0.584	0.506	0.506	0.480
Khartoum	0.535	0.512	0.512	0.512	0.463	0.320	0.281	0.354	0.324	0.281	0.354	0.281	0.354	-	0.167	0.482	0.500	0.500	0.500	0.500	0.512	0.544	0.500	0.544	0.556	0.584	0.571	0.556	0.524	0.535	0.584	0.506	0.506	0.480
A22	0.535	0.512	0.512	0.512	0.463	0.320	0.281	0.354	0.324	0.281	0.354	0.281	0.354	0.000	-	0.506	0.480	0.480	0.522	0.522	0.494	0.528	0.522	0.528	0.541	0.571	0.556	0.541	0.506	0.518	0.571	0.487	0.487	0.459
Ayele 8	0.626	0.636	0.636	0.608	0.537	0.670	0.636	0.655	0.652	0.636	0.655	0.655	0.661	0.629	0.619	-	0.442	0.442	0.267	0.267	0.500	0.452	0.267	0.271	0.577	0.601	0.619	0.577	0.550	0.558	0.601	0.535	0.535	0.512
MRC(L)3	0.632	0.615	0.615	0.584	0.544	0.626	0.615	0.608	0.684	0.668	0.608	0.668	0.636	0.626	0.626	0.383	-	0.000	0.413	0.413	0.424	0.469	0.413	0.358	0.429	0.474	0.541	0.429	0.435	0.453	0.474	0.408	0.408	0.435
MRC74	0.603	0.584	0.584	0.550	0.544	0.626	0.615	0.608	0.684	0.668	0.608	0.668	0.661	0.626	0.626	0.383	0.229	-	0.413	0.413	0.424	0.469	0.413	0.358	0.429	0.474	0.541	0.429	0.435	0.453	0.474	0.408	0.408	0.435
Ndandu 4A	0.615	0.596	0.626	0.596	0.590	0.661	0.652	0.645	0.692	0.700	0.645	0.677	0.670	0.661	0.661	0.408	0.413	0.413	-	0.305	0.477	0.471	0.000	0.309	0.558	0.552	0.629	0.558	0.564	0.537	0.552	0.550	0.550	0.528
D2	0.619	0.601	0.601	0.571	0.565	0.612	0.601	0.595	0.670	0.655	0.595	0.655	0.623	0.612	0.612	0.374	0.378	0.378	0.271	-	0.431	0.374	0.305	0.309	0.522	0.552	0.571	0.522	0.564	0.500	0.552	0.550	0.550	0.528
Neal R1	0.645	0.629	0.629	0.601	0.532	0.639	0.629	0.623	0.645	0.629	0.623	0.629	0.595	0.583	0.583	0.309	0.431	0.431	0.345	0.218	-	0.226	0.477	0.436	0.535	0.564	0.469	0.535	0.577	0.512	0.564	0.563	0.563	0.577
Mutinga H9	0.626	0.636	0.636	0.608	0.537	0.670	0.636	0.655	0.652	0.636	0.655	0.661	0.629	0.619	0.619	0.224	0.436	0.436	0.349	0.309	0.221	-	0.471	0.378	0.564	0.590	0.463	0.564	0.603	0.544	0.590	0.591	0.591	0.603
Addis 142	0.596	0.577	0.577	0.544	0.601	0.645	0.661	0.601	0.700	0.685	0.655	0.685	0.655	0.670	0.670	0.477	0.436	0.436	0.408	0.374	0.426	0.431	-	0.309	0.558	0.552	0.629	0.558	0.564	0.537	0.552	0.550	0.550	0.528
LRC L57	0.603	0.615	0.643	0.615	0.636	0.652	0.668	0.685	0.684	0.692	0.685	0.692	0.685	0.700	0.700	0.436	0.488	0.488	0.277	0.378	0.431	0.383	0.436	-	0.528	0.558	0.577	0.528	0.535	0.506	0.558	0.518	0.518	0.535
Buck	0.643	0.626	0.626	0.626	0.645	0.636	0.652	0.619	0.692	0.677	0.645	0.677	0.645	0.661	0.661	0.500	0.463	0.463	0.436	0.403	0.452	0.457	0.349	0.463	-	0.232	0.363	0.000	0.333	0.285	0.232	0.293	0.293	0.333
VL23	0.584	0.598	0.598	0.598	0.650	0.640	0.528	0.650	0.707	0.714	0.727	0.714	0.676	0.667	0.667	0.648	0.692	0.692	0.639	0.616	0.589	0.623	0.595	0.655	0.639	-	0.418	0.232	0.397	0.358	0.000	0.368	0.368	0.397
Jeddah KA	0.617	0.629	0.629	0.629	0.591	0.640	0.657	0.676	0.617	0.629	0.676	0.629	0.650	0.640	0.640	0.632	0.667	0.667	0.700	0.700	0.677	0.659	0.728	0.667	0.722	0.465	-	0.363	0.480	0.447	0.358	0.459	0.459	0.480
Ayele 5	0.610	0.591	0.591	0.591	0.584	0.632	0.650	0.643	0.667	0.676	0.643	0.650	0.643	0.659	0.659	0.677	0.632	0.632	0.668	0.670	0.693	0.700	0.700	0.684	0.692	0.333	0.459	-	0.333	0.285	0.397	0.293	0.293	0.333
VL29	0.676	0.659	0.659	0.659	0.528	0.692	0.684	0.677	0.676	0.684	0.677	0.659	0.652	0.643	0.643	0.636	0.643	0.643	0.677	0.678	0.655	0.661	0.728	0.714	0.721	0.480	0.429	0.363	-	0.368	0.232	0.174	0.174	0.338
Addis 164	0.659	0.643	0.643	0.643	0.544	0.677	0.692	0.685	0.632	0.643	0.661	0.643	0.608	0.652	0.652	0.645	0.652	0.652	0.685	0.686	0.663	0.677	0.714	0.700	0.707	0.453	0.397	0.324	0.281	-	0.418	0.333	0.333	0.429
VL6	0.684	0.668	0.668	0.643	0.544	0.700	0.692	0.661	0.684	0.692	0.685	0.668	0.661	0.652	0.652	0.670	0.677	0.677	0.707	0.707	0.686	0.693	0.714	0.700	0.747	0.541	0.500	0.447	0.281	0.387	-	0.368	0.368	0.397
L. infantum	0.747	0.734	0.752	0.714	0.663	0.740	0.752	0.727	0.721	0.728	0.740	0.707	0.707	0.721	0.721	0.639	0.655	0.655	0.629	0.606	0.577	0.612	0.583	0.645	0.629	0.772	0.735	0.728	0.693	0.678	0.655	-	0.000	0.297
IMT 172	0.741	0.728	0.747	0.707	0.655	0.734	0.747	0.721	0.647	0.657	0.650	0.657	0.721	0.714	0.714	0.707	0.645	0.645	0.722	0.721	0.741	0.728	0.748	0.715	0.742	0.767	0.735	0.721	0.685	0.693	0.645	0.158	-	0.200
HN336	0.747	0.734	0.752	0.714	0.663	0.740	0.752	0.727	0.721	0.728	0.740	0.707	0.707	0.721	0.721	0.639	0.655	0.655	0.629	0.606	0.577	0.612	0.583	0.645	0.629	0.772	0.735	0.728	0.693	0.678	0.655	0.000	0.158	-

	MIRT	LON	MON	country
Salti 4 (D17)		(46)	(18)	LB
IMT 180 (D34)		(46)	18	PT
LRC L64 (D22)		48	(30)	SD
Mesh 17 (D5)		50	(82)	IR
Gilani (D33)	3	48	30	SD
Gebre 1 (D28)		50	82	ET
Dora (D24)		50	(82)	IT
A22 (D19)		(48)	(30)	SD
Khartoum (D18)		46	18	SD
HU3 (D1)		46	18	ET
Nandi 1 (D8)		41	(2)	IN
STL 1-79 (D7)	2	41	(2)	IN
DD8 (D)		41	2	IN
Patna 1 (D4)		(41)	(2)	IN
L. infantum		49	1	**
IMT 172 (I26)	4	(201)	24	PT
HN 336 (C20)		-	-	HN
Neal R1 (D30)		56	(32)	KE
Mutinga H9 (D31)		56	(32)	KE
Ayele 8 (D27)		56	(32)	ET
D2 (D25)		45	-	KE
Addis 142 (D15)	1	(56)	(32)	ET
Ndandu 4A (D16)		44	(37)	KE
LRC L57 (D21)		44	37	KE
MRC 74 (D3)		51	-	KE
MRC (L)3 (D2)		(56)	(32)	KE
Buck (I31)		(201)	78	MT
VL 29 (D10),		(42)	(83/31)	SA
Hussen (D13)		42	(83/31)	ET
VL 6 (D11)		(42)	(83/31)	SA
Bekele (D14)		42	(83/31)	ET
UGX-marrow (D20)	5	(205)	(83/31)	SD
Addis 164 (D29)		(205)	83	ET
Ayele 5 (D12)		42	(83/31)	ET
VL23 (D9)		(42)	(83/31)	SA
Jeddah KA (D32)		42	31	SA
Wangjie 1 (D26)	6	(206)	35	CN

Fig. 5. Phylogenetic analysis of the *Leishmania donovani* complex from MIRT data. Neighbour-joining clustering method from Jaccard distances. Bootstrap values higher than 70 % from 1000 replicates (PAUP) are shown above main branches: values for UPGMA were very similar ($\pm 4\%$). Clusters are highlighted and corresponding MIRT are shown. LON and MON zymodeme types are shown, with partial zymodeme typing as described in Table 1 in parentheses. *L. infantum* refers to IPT 1, PP75, HN122, IMT 104, M12084, M6445, Pharoah, Strain A, L53, WR285. For country codes see legend of Table 1. ** from various Mediterranean and Latin American countries.

topologies similar to the unrooted neighbour-joining tree (Fig. 4), except for strain Wangjie 1 (c-MIRT 6) closer to c-MIRT 3. A consensus tree from 84 equally parsimonious Wagner trees (116 steps total length) and the maximum likelihood tree (log likelihood: −520·70; 6168 unresolved quartets (17·2 %); 28·1 % constant sites with pi(1) = 58·0 %) were congruent with distance methods for lineages with more than 70 % bootstrap support (Fig. 4). Dollo parsimony produced fewer alternative trees (46), but with higher length (135 steps), and identified the same clusters.

Phylogenetic analysis of l-MIRT

In the l-MIRT UPGMA dendrogram (cophenetic correlation coefficient 0·955), from Jaccard distances (Table 2) 5 groups were identified, which were also defined in the distance matrix, were classified into the following l-MIRT types. l-MIRT 1: Addis 142, Ayele 8, D2, LRC-L57, Ndandu 4A; l-MIRT 2:

DD8, Nandi 1, Patna 1, STL1-79; l-MIRT 3: subgroup 3′–A22, Dora, Gebre 1, HU3, IMT 180, Khartoum, LRC-L64, Mesh 17, Salti 4, subgroup 3″–Mutinga H9, NEAL R1, (Gilani); l-MIRT 4: MRC(L)3, MRC74, l-MIRT 5: subgroup 5′–Hussen, VL29, *L. infantum* (IPT 1, PP75, Strain A, L53, IMT 172, HN336); subgroup 5″–Addis 164, Ayele 5, Bekele, Buck, UGX-marrow, VL23, VL6, Wangjie 1.

Although the main groups remained unchanged by single linkage and neighbour-joining, the branch containing strains MRC(L)3 and MRC74 clustered separately and strains Jeddah-KA, Neal R1, Mutinga H9 and Gilani shifted positions depending on the method used. Wagner parsimony (consensus tree from 100 most parsimonious trees with a total length of 68 steps) also identified the main groups and had similar topology, but the resulting tree included several polytomies. Dollo parsimony (consensus tree from 57 trees with 78 total steps) maintained clusters l-MIRT 2, 5 and 3′, whilst
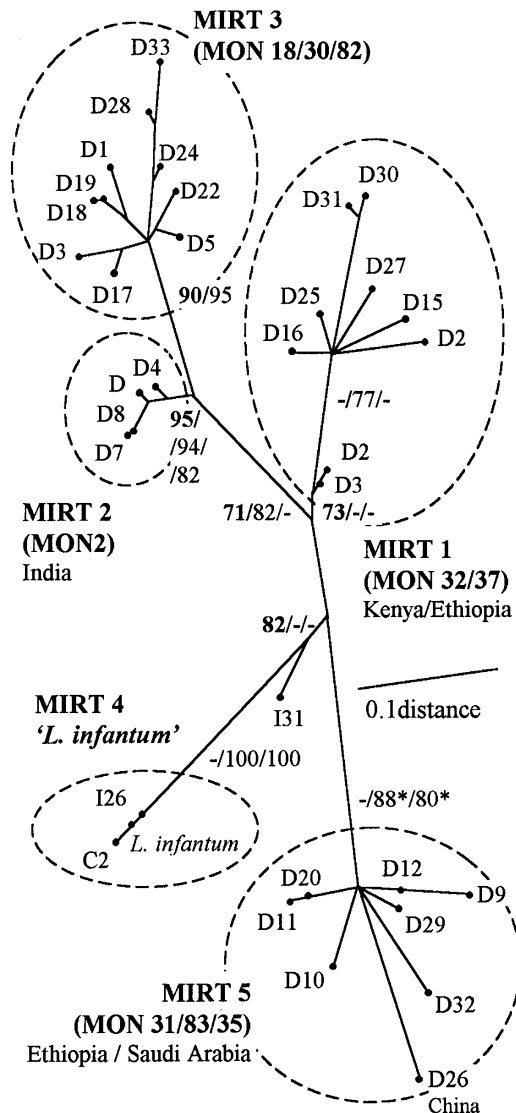
Fig. 6. Consensus (quartet puzzling) maximum likelihood tree (log L: −965·24), with 39·5% constant sites, pi(1) = 61·2% and 52360 (10·7%) unresolved quartets. Deep branch support higher than 70% (in bold) is shown for maximum likelihood (bold)/Wagner parsimony/Dollo parsimony. *Value without D26 (Wangjie 1) from Wagner parsimony. Dotted ellipses highlight genetic groups. 'Leishmania infantum' refers to strains IPT 1, PP75, HN122, IMT 104, M12084, M6445, Pharoah, Strain A, L53, WR285. Strain D11 (VL6) is indistinguishable from Bekele and D10 (VL9) is indistinguishable from Hussen. For other strain codes compare with Fig. 5 and see Table 1.

separating l-MIRT 3″ and including l-MIRT 4 in l-MIRT 1. Group 1-MIRT 3 changed dramatically by maximum likelihood (log likelihood: −362·74, with 2988 (17·0%) unresolved quartets, 52·7% constant sites, p(i) = 65·5%), but all other groups were maintained. Bootstrap support for main lineages was lower than 75% for cluster and parsimony methods, but was 88% for l-MIRT 2 in maximum likelihood.

Bootstrap and quartet puzzling analyses for l-MIRT described very few robust lineages (support

higher than 75%): Indian strains in all; Neal R1 and Mutinga H9 in UPGMA, neighbour-joining and maximum likelihood; VL6 (identical to Bekele) and UGX-marrow (identical to Ayele 5), *L. infantum* and VL29 (identical to Hussen), in UPGMA and neighbour-joining; l-MIRT 1 and 5 and Jeddah-KA, in maximum likelihood.

Phylogenetic analysis of pooled MIRT

Groups identified by l-MIRT and c-MIRT were congruent between phylogenies for robust lineages common to all methods, which, together with the proximity of the loci and assuming a low prevalence of recombination, provided the basis for combining the data sets.

Five main clusters of *L. donovani* complex strains were identified for MIRT with all phylogenetic methods tested and were classified into MIRT types (Figs 5 and 6) which were similar to c-MIRT clusters.

The main groups were identified with neighbour-joining (Fig. 5), UPGMA (cophenetic correlation coefficient 0·945) and Wagner and Dollo parsimony with high bootstrap support (Figs 5 and 6). The only exception was cluster MIRT 1 in which neighbour-joining including strain Buck. Maximum parsimony identified the same groups with higher support for MIRT 1 but lower for MIRT 5, which included strain Wangjie 1. Strains Wangjie 1 and Buck did not fit consistently into any of the groups, and strain Wangjie 1 was sufficiently different to be placed in MIRT 6.

MIRT clusters were strongly associated with zymodeme type, and geographical patterns were evident in some clusters (MIRT 2 – India – and 1 – Kenya/Ethiopia), as described in Fig. 5 and highlighted in Fig. 6. Among widespread clusters (MIRT 3 and *L. infantum*), the most remarkable example is *L. infantum* which reaches Latin America, whilst having a low degree of genetic diversity.

Congruence of genetic distances

The 5 genotypic and zymodeme clusters identified within the *L. donovani* complex were clearly detected for all pairwise genetic distances less than 65% of the maximum observed Jaccard distance (Table 2) for MIRT (0·445) and c-MIRT (0·502) but less so for l-MIRT (0·413). Congruent distance ratios were also observed for parsimony mean distances but not between c-MIRT and l-MIRT for maximum likelihood.

DISCUSSION

Typing methods

Two novel PCR–RFLP methods, as far as we are aware, are described here for phylogenetic analysis

of the *L. donovani* complex. c-MIRT (for gp63 ITG/CS) was the best method to distinguish between *L. infantum* (*L. chagasi*) and *L. donovani* strains, whilst l-MIRT (for gp63 ITG/L) complemented c-MIRT, by generating robust strain clusters when the data generated by the two methods were analysed together (MIRT).

Amplification of ITG/CS generated a single product which by restriction analysis was found to have substantial genetic diversity within the *L. donovani* complex. c-MIRT generated robust clades identified by different phylogenetic methods.

l-MIRT requires purification of the ITG/L specific band after PCR and did not generate robust clades and, most importantly, did not resolve a *L. infantum* clade. Analysis of this region is complicated by the presence of multiple copies but also because the restriction profiles were not sufficiently diverse.

Despite the differences, the two methods gave congruent lineages with bootstrap support higher than 70%, except for strain Buck, and for strain VL29 which were internal to the *L. infantum* clade by l-MIRT.

The combined MIRT analysis improved bootstrap support for the main clades. Topology of the trees was more unstable than c-MIRT across the tested tree building methods, probably due to the influence of ITG/L on the pooled data.

Phylogenetic methods

Phylogenetic analysis using several different methods verified reliability of c-MIRT and MIRT in that genetic groups within the *L. donovani* complex were consistently identified, despite the fact that RFLP data can be difficult to use for phylogenetic analysis. It is difficult to distinguish gain or loss of a restriction site from insertions or deletions among complex restriction profiles such as those generated by the large molecules, of 1·6 to 1·75 kb, used here. The diploidy of the genome and multiple copies of the genes can further complicate the analysis.

Taxa within the L. donovani *complex*

The organization of gp63 genes described for *L. chagasi* (Roberts *et al.* 1993) seems to be conserved in the *L. donovani* complex. The ITG/L fragment, in the generic gp63 ITG amplification, is much more intense than the fragment corresponding to the ITG/CS region which is compatible with a gene repeat versus a single copy gene.

MIRT confirmed that *L. chagasi* and *L. infantum* are indistinguishable (Beverley, Ismach & Pratt, 1987; Cupolillo, Grimaldi & Momen, 1994; Moreno *et al.* 1984; Schonian *et al.* 1996; van Eys *et al.* 1991) and thus synonymous (Mauricio *et al.* 2000). The genetic diversity between *L. infantum* and *L. chagasi*

strains was within the level expected for intraspecific diversity (Beverley *et al.* 1987), which explains differences between previous single strain comparisons. *L. infantum* (*chagasi*) was confirmed to be monophyletic (Mauricio *et al.* 1999, 2000), except for strain Buck which may not strictly be *L. infantum*.

Genetic groups corresponding to clusters were used instead of assigning demes. Six MIRT genetic clusters were found within the *L. donovani* complex which correlated strongly with isoenzyme typing. The existence of these groups was supported by phylogenies of *mspC* gene sequence, restriction of ITS and RAPD (Mauricio, Gaunt, Stothard and Miles, unpublished data). Identification of groups was not dependent on time of isolation, because the 2 most recent strains (Mesh 17 and IMT 180) were reliably typed as MIRT 3 and all groups encompassed strains isolated several years apart.

Cluster MIRT 4 includes only *L. infantum* strains but no other group has a specific designation. Although MON 30 strains have been considered as *L. infantum* by isoenzyme typing (Rioux *et al.* 1990), this zymodeme was identified as MIRT 3, which may have important implications for epidemiology and development of diagnostic methods.

It is not known if the 6 MIRT clades have any associated biological or clinical features, but strong correlations with zymodeme and some geographical associations were identified. MIRT 2 strains were MON 2 (LON 41) and were limited to India, MIRT 1 were MON 37/32 (LON 44/45/51/56) from Kenya/ Ethiopia, MIRT 5 strains were MON 83/31 (LON 42) from Ethiopia/Saudi Arabia, and MIRT 6 (MON 35) was from China. The reservoirs for the 3 latter groups are not clearly known, whilst MIRT 2 (India) is known to be anthroponotic and *L. infantum* (MIRT 4) is zoonotic with a principal canine reservoir.

The Indian MIRT clade is known to produce PKDL, whilst the *L. infantum* clade does not. PKDL also occurs in Africa and may depend predominantly on the host immune response, but correlation between MIRT and PKDL should be investigated. A taxonomic status for *L. infantum* and for *L. donovani* of India (MIRT 2) is valuable for diagnostics, epidemiology and disease control, although it would be better adjusted to the level of subspecies. We consider it to be inappropriate to create more taxa within the *L. donovani* complex, but reference to the *L. donovani* genetic groups identified here may be an important parameter in epidemiological and clinical studies.

We found no support for specific status for *L. archibaldi*, defined as MON 82. Zymodeme MON 82 clustered with MIRT 3, which also included zymodeme variants LON 46 and 48 (MON 18 and 30, respectively) and which is a cosmopolitan group. Rioux *et al.* (1990) placed this group at the root of the *L. donovani* complex and described 3 species, *L.*

infantum, *L. donovani* and *L. archibaldi*. However, zymodeme MON 82 (LON 50) is a putative ASAT hybrid or heterozygote of LON 46 and LON 48 and these 3 zymodemes have been found in Sudan behaving as a single population (Ashford *et al.* 1992). Thus, either the designation *L. archibaldi* should be discarded or used as a subspecies (*L. d. archibaldi*) for all members of the genetic group MIRT 3 (MON 18, 30, 82/LON 46, 48, 50). The extensive distribution of this group may have implications in epidemiology and control. One of the strains used here (IMT 180) was isolated from an AIDS patient in Portugal, in a possible needle transmission (Campino *et al.* 1994), whilst in Iran and Italy, strains of the same group were isolated from rodent and dog hosts.

Two strains did not fit consistently into any of the groups: strain Wangjie 1 (China) was thus included as the single member of MIRT 6 which might include other strains from the Far-East. Strain Buck belongs to a zymodeme which seems to be endemic to Malta (Gramiccia, Gradoni & Troiani, 1992) and did not fit well into any MIRT cluster but was closest to *L. infantum*.

The present analysis was not intended to provide a basis for speculation on the history of the *L. donovani* complex and no outgroup was used. However, it provides a framework for further analysis of the *L. donovani* complex and significant original insight into the composition of the complex. Using c-MIRT or MIRT, *L. donovani* complex strains can be easily assigned to genetic groups, or new groups could be identified. Simple genetic distance measures, such as Jaccard, can be used to assign strains to genetic groups, but profiles within the groups are sufficiently similar to allow visual identification of strains.

REFERENCES

ASHFORD, R. W., SEAMAN, J., SCHORSCHER, J. & PRATLONG, F. (1992). Epidemic visceral leishmaniasis in southern Sudan: identity and systematic position of the parasites from patients and vectors. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86**, 379–380.

BEVERLEY, S. M., ISMACH, R. B. & PRATT, D. M. (1987). Evolution of the genus *Leishmania* as revealed by comparisons of nuclear DNA restriction fragment patterns. *Proceedings of the National Academy of Sciences, USA* **84**, 484–488.

CAMPINO, L., SANTOS GOMES, G., PRATLONG, F., DEDET, J. P. & ABRANCHES, P. (1994). The isolation of *Leishmania donovani* MON-18, from an AIDS patient in Portugal: possible needle transmission. *Parasite* **1**, 391–392.

CUPOLILLO, E., GRIMALDI, G. Jr. & MOMEN, H. (1994). A general classification of New World *Leishmania* using numerical zymotaxonomy. *American Journal of Tropical Medicine and Hygiene* **50**, 296–311.

CUPOLILLO, E., GRIMALDI, G. Jr, MOMEN, H. & BEVERLEY, S. M. (1995). Intergenic region typing (IRT): a rapid molecular approach to the characterization and evolution of *Leishmania*. *Molecular and Biochemical Parasitology* **73**, 145–155.

EVANS, D. A., LANHAM, S. M., BALDWIN, C. I. & PETERS, W. (1984). The isolation and isoenzyme characterization of *Leishmania braziliensis* subsp. from patients with cutaneous leishmaniasis acquired in Belize. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **78**, 35–42.

FELSENSTEIN, J. (1993). *PHYLIP* (*Phylogeny Inference Package*) *Version 3.5c*. Distributed by the author. Department of Genetics, University of Washington, Seattle.

GRAMICCIA, M., GRADONI, L. & TROIANI, M. (1992). HIV-*Leishmania* co-infections in Italy. Isoenzyme characterization of *Leishmania* causing visceral leishmaniasis in HIV patients. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86**, 161–163.

KELLY, J. M. (1993). Isolation of RNA and DNA from *Leishmania*. In *Methods in Molecular Biology*, vol. 21 (ed. Hyde, J. E.), pp. 123–131. Humana Press, New Jersey.

LE BLANCQ, S. M. (1986). *An Epidemiological and Taxonomical Study of Old World* Leishmania *Using Isoenzymes*. Ph.D. thesis, University of London, London.

MAINGON, R. D. C., ELSON, L., HOMMEL, M., CRAMPTON, J. M. & CHANCE, M. L. (1990). Multiple distinct gp63-like genes in amastigotes of *Leishmania donovani*. In *Parasites: Molecular Biology, Drug And Vaccine Design*, (ed. Agabian N. & Cerami, A.) pp. 191–204. Willey-Liss, Inc., New York.

MAURICIO, I. L., HOWARD, M. K., STOTHARD, J. R. & MILES, M. A. (1999). Genetic diversity in the *Leishmania donovani* complex. *Parasitology* **119**, 237–246.

MAURICIO, I. L., STOTHARD, J. R. & MILES, M. A. (2000). The strange case of *Leishmania chagasi*. *Parasitology Today* **16**, 188–199.

MEDINA ACOSTA, E., KARESS, R. E. & RUSSELL, D. G. (1993). Structurally distinct genes for the surface protease of *Leishmania mexicana* are developmentally regulated. *Molecular and Biochemical Parasitology* **57**, 31–45.

MILES, M. A., LANHAM, S. M., DE SOUZA, A. A. & POVOA, M. (1980 *a*). Further enzymic characters of *Trypanosoma cruzi* and their evaluation for strain identification. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **74**, 221–237.

MILES, M. A., POVOA, M. M., DE SOUZA, A. A., LAINSON, R. & SHAW, J. J. (1980 *b*). Some methods for the enzymic

characterization of Latin-American *Leishmania* with particular reference to *Leishmania mexicana amazonensis* and subspecies of *Leishmania hertigi*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **74**, 243–252.

MORENO, G., RIOUX, J.-A., LANOTTE, G., PRATLONG, F. & SERRES, E. (1984). Le complexe *Leishmania donovani s.l.* Analyse enzymatique et traitement numerique. Individualization du complexe *Leishmania infantum*. Corollaries biogeographiques et phyletiques. Apropos de 146 souches originaires de l'Ancien et du Nouveau Monde. *Leishmania*. In *Taxonomie et Phylogenèse Applications Éco-épidémiologiques* (ed. Rioux, J. A.), pp. 105–117. IMEEE, Montpellier.

PODANI, J. (1993). *SYN-TAX pc Computer Programs for Multivariate Data Analysis in Ecology and Systematics, Version 5.0*. Sciencia Publishing, Budapest.

RAMAMOORTHY, R., DONELSON, J. E., PAETZ, K. E., MAYBODI, M., ROBERTS, S. C. & WILSON, M. E. (1992). Three distinct RNAs for the surface protease gp63 are differentially expressed during development of *Leishmania donovani chagasi* promastigotes to an infectious form. *Journal of Biological Chemistry* **267**, 1888–1895.

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.

ROBERTS, S. C., SWIHART, K. G., AGEY, M. W., RAMAMOORTHY, R., WILSON, M. E. & DONELSON, J. E. (1993). Sequence diversity and organization of the *msp* gene family encoding gp63 of *Leishmania chagasi*. *Molecular and Biochemical Parasitology* **62**, 157–171.

SCHONIAN, G., SCHWEYNOCH, C., ZLATEVA, K., OSKAM, L., KROON, N., GRASER, Y. & PRESBER, W. (1996). Identification and determination of the relationships

of species and strains within the genus *Leishmania* using single primers in the polymerase chain reaction. *Molecular and Biochemical Parasitology* **77**, 19–29.

STEINKRAUS, H. B., GREER, J. M., STEPHENSON, D. C. & LANGER, P. J. (1993). Sequence heterogeneity and polymorphic gene arrangements of the *Leishmania guyanensis* gp63 genes. *Molecular and Biochemical Parasitology* **62**, 173–185.

STRIMMER, K. & VON HAESELER A. (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution* **13**, 964–969.

SWOFFORD, D. L. (2000). *PAUP* 4.0 : Phylogenetic Analysis Using Parsimony (And Other Methods). Version 4.0 (Beta Version)*. Sinauer, Sunderland, Massachusetts (in the Press).

VAN EYS, G. J., GUIZANI, I., LIGTHART, G. S. & DELLAGI, K. (1991). A nuclear DNA probe for the identification of strains within the *Leishmania donovani* complex. *Experimental Parasitology* **72**, 459–463.

VICTOIR, K., DUJARDIN, J. C., DE DONCKER, S., BARKER, D. C., AREVALO, J., HAMERS, R. & LE RAY, D. (1995). Plasticity of gp63 gene organization in *Leishmania* (*Viannia*) *braziliensis* and *Leishmania* (*Viannia*) *peruviana*. *Parasitology* **111**, 265–273.

VIDIGAL, T. H., DIAS NETO, E., CARVALHO, O. D. S. & SIMPSON, A. J. (1994). *Biomphalaria glabrata*: extensive genetic variation in Brazilian isolates revealed by random amplified polymorphic DNA analysis. *Experimental Parasitology* **79**, 187–194.

VOTH, B. R., KELLY, B. L., JOSHI, P. B., IVENS, A. C. & MCMASTER, W. R. (1998). Differentially expressed *Leishmania major* gp63 genes encode cell surface leishmanolysin with distinct signals for glycosylphosphatidylinositol attachment. *Molecular and Biochemical Parasitology* **93**, 31–41.