Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
International Journal for Parasitology 36 (2006) 757–769
www.elsevier.com/locate/ijpara

Towards multilocus sequence typing of the *Leishmania donovani* complex: Resolving genotypes and haplotypes for five polymorphic metabolic enzymes (ASAT, GPI, NH1, NH2, PGD)☆

Isabel L. Mauricio [a,*], Matthew Yeo [a], Mehdi Baghaei [a,1], Daniela Doto [a], Francine Pratlong [b], Eva Zemanova [c], Jean-Pierre Dedet [b], Julius Lukes [c], Michael A. Miles [a]

[a] *Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK*
[b] *Laboratoire de Parasitologie and Centre National de Reference des Leishmania, Centre Hospitalier Universitaire de Montpellier, 163 rue Auguste Broussonet, 34090 Montpellier, France*
[c] *Institute of Parasitology, Czech Academy of Sciences and Faculty of Biology, University of South Bohemia, Branisovska 31, 37005 České Budejovice, Czech Republic*

Received 29 November 2005; received in revised form 6 March 2006; accepted 21 March 2006

Abstract

Multilocus enzyme electrophoresis is the gold standard for identification of *Leishmania* species and strains. Drawbacks include: only amino acid polymorphisms affecting electrophoretic mobility are detected; distinct allozymes can have coincident mobilities; few characters are available; and parasites must be cultured in bulk. So far, thousands of *Leishmania* strains have been phenotyped by multilocus enzyme electrophoresis. Here, we sequence enzyme-coding genes to provide a PCR-based higher resolution equivalent of multilocus enzyme electrophoresis, particularly for *Leishmania infantum*. Of 15 enzymes used for multilocus enzyme electrophoresis (MON typing) we have sequenced aspartate aminotransferase, glucose-6-phosphate isomerase, nucleoside hydrolase 1, nucleoside hydrolase 2 and 6-phosphogluconate dehydrogenase. Heterozygous alleles were common, with multiple heterozygous sites within a single locus for several of the genes. Haplotypes were resolved by allele-specific PCR and allele-specific sequencing. Heterozygous haplotypes conformed to the haplotypes of putative parents. One strain appeared to be hybrid across two genetic groups of the *Leishmania donovani* complex. In most cases, a single amino acid polymorphism was responsible for change in enzyme mobility. Some indistinguishable phenotypes were produced by distinct genotypes. Silent genetic polymorphisms provided enhanced discrimination over multilocus enzyme electrophoresis, for example, by subdividing the zymodeme MON-1. The PCR-based genotyping that we describe could be applied directly to clinical samples or to small volume cultures and in a multilocus sequence typing format. Furthermore, it can be used to detect recombination indirectly and for population genetics studies.

Keywords: *Leishmania donovani*; *Leishmania infantum*; Nucleoside hydrolases; Glucose-6-phosphate isomerase; Aspartate aminotransferase; 6-Phosphogluconate dehydrogenase; Multilocus sequence typing

1. Introduction

The leishmaniases are human or veterinary diseases caused by protozoan parasites of the genus *Leishmania*. Different clinical forms can be produced by different *Leishmania* species or strains: visceral leishmaniasis by the *Leishmania donovani* complex of species, mucocutaneous leishmaniasis by the *Leishmania braziliensis* complex and cutaneous leishmaniasis by most other species. Multilocus enzyme electrophoresis (MLEE; isoenzyme analysis) has been the most comprehensive method used for identification of *Leishmania*, particularly

☆ Nucleotide sequences reports in this paper are available in the GenBank database under the accessions numbers: AJ620791 to AJ620823, AM117184 to AM117191, AM157715 to AM157720, AJ620617 to AJ620646, AM11792 to AM117193, AM157721 to AM157725, AJ620688 to AJ620721, AM157726 to AM157730, AJ621066 to AJ621102, AM117504 to AM117506, AM157731 to AM157735, AM157136 to AM157174 and AM157736 to AM157739.

* Corresponding author. Tel.: +44 20 7927 2814; fax: +44 20 7636 8739.
E-mail address: isabel.mauricio@lshtm.ac.uk (I.L. Mauricio).
[1] Department of Parasitology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran.

0020-7519/$30.00 © 2006 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2006.03.006

Table 1
Leishmania donovani complex strains and WHO reference strains for other Old World species

Species WHO code	Zymodeme MON (LON)[a]	Alleles					Genotype
		asat	*gpi*	*nh1*	*nh2*	*pgd*	
Leishmania infantum							
MHOM/FR/1978/LEM75	1	1	1	1	1	1	1
MHOM/ES/1986/BCN16	1	1	1	1	1	1	1
MHOM/ES/1993/PM1	1	1	1	1	1	1	1
MHOM/FR/1995/LPN114	1	1	1	1	1	1	1
MHOM/FR/1997/LSL29	1	1	1	1	1	1	1
MHOM/PT/2000/IMT260	1	1	1	1	1	1	1
MHOM/TN/1980/IPT-1[b]	1	1	8	1	1	1	2
MHOM/ES/1991/LEM2298	183	1	2	1	1	2	3
MHOM/CN/1980/Strain A	34	1	8	1	1	2	4
MHOM/FR/1980/LEM189	11	1	1	1/2	1	2	5
MHOM/FR/1996/LEM3249	29	1	1	2	1	2	6
MHOM/MT/1985/Buck	78	1	8	2	1	1	7
MHOM/ES/1988/LLM175	198	1	2	3	1	2	8
MHOM/IT/1993/ISS800	188	1	8	2	2	2	9
MHOM/IT/1994/ISS1036	228	1	1/2	2	1	2	10
MHOM/ES/1992/LLM373	199	1	2/8	2	1	2	11
MHOM/IT/1979/Francesca	27	1/4	1/8	1/2	1/2	1	12
Leishmania donovani							
MHOM/IN/0000/Devi	2	2	3	4	3	3	13
MHOM/IN/1996/Thak35	2	2	3	4	3	3	13
MHOM/IN/1954/SC23	38	3	4	6	3	5	14[c]
MHOM/KE/1954/LRC-L53	36	3	4	6	3	5	14[c]
MHOM/KE/1973/MRC74	2 (51)	3	4	7	3	4	15
IMAR/KE/1962/LRC-L57	37	3	4	7	3	5	16[c]
MHOM/KE/1975/Mutinga H9	32	3	4	7	3	5	16[c]
MHOM/ET/1967/HU3 (LV9)[b]	18	3	5	3	3	6	17
MHOM/SD/1982/Gilani[d]	30	1	5	3	3	2/6	18
MHOM/SD/1962/3S[d]	81	1	5	3	3	6	19[c]
MHOM/SD/1997/LEM3472[d]	267	1	5	3	3	6	19[c]
MHOM/SD/1997/LEM3429[e]	257	1/3	5	4	3	6/8	20
MHOM/SD/1997/LEM3463[e]	258	1/3	5	4	4	6/9	21
MHOM/ET/1972/Gebre 1[e]	82	1/3	5	3	3	2	22
MCAN/SD/2000/LEM3946	274	3	5/6	4	4/3	2/6	23
MHOM/CN/0000/Wangjie 1	35	3	7	5	5	7	24
MHOM/SU/1984/Marz-Krim	73	3	8	4	2	2	25
MHOM/ET/0000/Hussen	31	3	6	4	4	2	26[c]
MCAN/IQ/1981/Sukkar2	–	3	6	4	4	2	26[c]
MHOM/IQ/1977/Bumm3	3	3	6	4	4	2	26[c]

[a] In parenthesis are LON zymodemes.
[b] World Health Organization (WHO) reference strains.
[c] Genotypes identical for different zymodemes.
[d] *L. infantum* strains according to Pratlong et al. (2001). Key to WHO codes: host/country/year/strain name. Hosts: IARI—*Phlebotomus ariasi*; IMAR—*Phlebotomus martini*, MCAN—*Canis familiaris*, MHOM—*Homo sapiens*, MPSM—*Psammomys* sp; MRHO—*Rhombomys opimus.* Countries: CN: China; ES: Spain; ET: Ethiopia; FR: France; IN: India; IQ—Iraq; IT: Italy; KE: Kenya; MT: Malta; PT: Portugal; SA: Saudi Arabia; SD: Sudan; SU: former Soviet Union; TN: Tunisia. 0000 is unknown year.
[e] *Leishmania archibaldi.*

the MON system, which is based on 15 enzymes (Rioux et al., 1990). Isoenzymes have also been used to generate phylogenetic trees and to provide a basis for the current taxonomy of the *L. donovani* complex, which has three designated species, *L. donovani*, *Leishmania infantum* and *Leishmania archibaldi* (Rioux et al., 1990; Pratlong et al., 2001). Multilocus enzyme electrophoresis has many drawbacks: strains with the same enzyme phenotype may in fact have distinct amino acid sequences; the degree of relationship between different phenotypes is not known; putative heterozygous phenotypes are difficult to interpret; it is time consuming and expensive, and requires large volume *Leishmania* cultures. A simpler and more discriminatory method is required.

The definition of the three named species is not consensual, as *L. archibaldi* and *L. infantum* strains in the Sudan have been found to cluster with *L. donovani* strains of the same area by a range of different molecular tools (Mauricio et al., 2001, 2004; Jamjoom et al., 2004; Zemanova et al., 2004; Kuhls et al., 2005). Here, we will thus designate *L. donovani* to zymodemes MON-30, MON-81 and MON-267, previously classified into

L. infantum and to zymodemes MON-82, MON-257 and MON-258, previously *L. archibaldi*.

Here, we amplified and sequenced the genes for aspartate aminotransferase (ASAT; or glutamate oxaloacetate transaminase, GOT1, E.C.2.6.1.1), glucose-6-phosphate isomerase (GPI, E.C.5.3.1.9), nucleoside hydrolase 1 (NH1; or nucleoside phosphorylase 1, NP1, E.C.3.2.2.2), nucleoside hydrolase 2 (NH2; or nucleoside phosphorylase 2, NP2, E.C.3.2.2.x) and 6-phosphogluconate dehydrogenase (PGD, E.C.1.1.1.44). Our aim is to investigate the deeper relationships between *L. donovani* complex zymodemes and to devise methods capable of enhancing MLEE, taking advantage of a DNA-based approach.

All five enzymes are dimeric but are encoded by a single gene, according to the available *Leishmania major* and *L. infantum* genome sequences (http://www.genedb.org/). ASAT and GPI are involved in glycolysis, PGD in the pentose phosphate pathway, and NH1 and NH2 act in nucleotide metabolism. Aspartate aminotransferase has been the character used to discriminate between the three classical species of the *L. donovani* complex (Rioux et al., 1990): *L. infantum* with phenotype 100, *L. archibaldi* with 110 and *L. donovani* with 113. Glucose-6-phosphate isomerase and NH1 are highly polymorphic, with GPI phenotype 100 and NH1 phenotypes 100 and 140 occurring in both *L. infantum* and *L. donovani*. Nucleoside hydrolase 2 and PGD are monomorphic in the *L. donovani* complex, with a single PGD variant in zymodeme MON-2 (characteristic of India).

We describe genetic polymorphisms responsible for the different observable phenotypes and others that provide further discrimination, thus creating the basis for a PCR-based enzyme typing. We show evidence of recombination in the complex, as suggested by heterozygous sequences, mosaic genotypes and a putative hybrid across different genetic groups. We determine allelic phase in heterozygotes through a new PCR-based method. We thus demonstrate that analysis of enzyme coding genes has the potential for use in genetic typing and population genetics.

2. Materials and methods

2.1. Strains and extraction of genomic DNA

Strains were chosen to be representative of all *L. donovani* complex phenotypes, including strains from different genetic groups (Table 1) and from most published zymodemes. Strains belonging to other species were also used as outgroups: *Leishmania aethiopica* MHOM/ET/1972/L100, *Leishmania arabica* MPSM/SA/1983/JISH220, *Leishmania killicki* MHOM/TN/1980/LEM163, *L. major*, MHOM/SU/1973/5-ASKH, *Leishmania tropica* MHOM/SU/1974/K27, *Leishmania turanica* MRHO/SU/1983/MARZ-051. Genomic DNA was prepared as described by Kelly (1993) from parasite mass cultures in alpha-MEM.

2.2. PCR amplification

Primers were designed from GeneDB data of the *L. major* and the *L. infantum* genome sequencing projects (http://www.genedb.org/) and for *nh1* from the *L. donovani* sequence AY033633 in GenBank/EMBL. Primers were designed if possible in the intergenic regions to amplify the entire coding regions of each gene. Different primer combinations were tested and the best amplification was achieved with the primers listed in Table 2. Each 100 µl of optimised amplification reaction contained: 125 ng genomic DNA, 50 pmol each primer, 0.2 mM dNTPs, XmM $MgCl_2$, and 5 U *Taq* polymerase (*X* is given in Table 2). Amplification conditions were: 35 cycles with 95 °C for 1 min, *Y* °C for 1 min and 72 °C for *Z* min, with a final extension at 72 °C for 10 min (*Y* and *Z* are given in Table 2). Amplification of *nh1* required 10% dimethyl sulphoxide (DMSO). The sensitivities of each PCR were tested for IPT-1 and DD8 (*L. infantum* and *L. donovani* reference strains) with serial dilutions of 25, 2.5, 0.25 and 0.025 ng of DNA in 20 µl reactions.

2.3. DNA sequencing

PCR products were purified either from a gel with QIAEX II or QIAQuick gel extraction kits (Qiagen) or directly from the PCR reaction with QIAQuick PCR extraction kit (Qiagen) if a single band was amplified. The fragments were sequenced with the PCR primers and with internal primers (Supplementary Data), to obtain forward and reverse sequences for the entire genes. We used direct sequencing with BigDye™ Terminator Cycle Sequencing V2 or V3.1 (ABI PRISM® Applied Biosystems) in ABI PRISM™ 377 or 3730 DNA Sequencers (Applied Biosystems), respectively. Partial sequences were manually assembled in BioEdit with the aid of Clustal W and a consensus sequence was generated for each strain. The consensus coding sequences were then aligned manually because no gaps were detected.

2.4. Phase determination

Allelic phase of heterozygous sequences was established for *asat*, *gpi* and *pgd* using two different methods. One method was allele-specific sequencing (Hare and Palumbi, 1999), used for more than two heterozygous sites that were located within 700 bp of each other, and the other was allele-specific PCR, which we developed here (see Section 3 for more detail). The allele-specific PCR was used for two or more sites that were at least 100 bp from each other, using combinations of single nucleotide polymorphism (SNP) specific primer pairs amplifying towards each other. Both methods use primers specific for each polymorphism at their 3′ end base and with a destabilizing mismatch at the second 3′ end position (Supplementary Data). The PCR conditions were as described above and listed in Supplementary Data. We used Diamond DNA Polymerase (Bioline) that lacks a 3′–5′ proofreading activity.

Table 2
Amplification conditions for each gene

Target	Primer name[a]	Location of 3′ end from ATG	Primer sequence[b]	*X* ($MgCl_2$) mM	*Y* (T_{ann}) °C	*Z* (ext. time) min
asat	F2new	1592	acgagcgccgtccgyaa	1	60	1.5
	R2new	−5	ttccymcatccaccaagc			
gpi	F3	12	gaatccctttttcaag**atg**agcgattat	1	58	2
	R4	1937	cccctgagaggcaatcacag			
nh1	F1	−30	cttgcttacgccgcagatac	2	65	1
	R3	1030	gaaaaaaaagacgcttcacacaagc			
nh2	F10	−100	acgtgggcgaacgcac	1	60	1.5
	R6	1137	gccatctacaccttcagtgcctcggtc			
pgd	F1	3	gaacgaatcccttattctcy**atg**	1	60	1.5
	R2	1472	ggaaccggttgagcggc			

Ext. time, extension time.
[a] F is forward and R is reverse.
[b] In bold are start in primer sequences.

2.5. DNA sequence analyses

Synonymous and non-synonymous substitution rates were calculated and a *Z*-test of selection was done using the modified Nei–Gojobori method (Nei and Gojobori, 1986; Nei and Kumar, 2000) with Jukes–Cantor distances (which corrects for multiple substitutions at the same site), assuming homogeneity between lineages and uniform rates between sites and with the observed transition/transversion ratio, using MEGA (Kumar et al., 1994). The distribution pattern of the polymorphisms was evaluated through alpha, the shape parameter of the gamma distribution, as determined through the programme PHYML (Gu and Zhang, 1997) using a GTR model (Lanave et al., 1984) and eight categories.

2.6. Protein sequence analyses

Molecular mass, overall charge and charge variation, according to pH, of the inferred protein sequence, were determined through programme ABIM: © Ixe et Ygrec, available at http://www.iut-arles.up.univ-mrs.fr/w3bb/d_abim/compo-p.html

2.7. Phylogenetics

Branches in individual gene trees had low bootstrap support due to the small number of characters, so we produced haplotype networks for individual genes using statistical parsimony (at 95% connection limit) as implemented in programme TCL (Clement et al., 2000) for a single sequence for each haplotype.

Genome trees were produced from concatenated DNA sequences of strains with one or with no heterozygous genes. The concatenated sequences were deconstructed into two separate sequences (or haplotypes), based on phase determination. Strains with two or more heterozygous genes were not included in these analyses. Phylogenetic trees from concatenated DNA sequences were produced from Kimura-2-parameter distance matrices by Neighbor-Joining (NJ) in the MEGA package (Kumar et al., 1994) and by Maximum Likelihood (ML) using programme DNAML in the PHYLIP package (Felsenstein, 1993. Phylogeny Inference Package (PHYLIP). Seattle, Distributed by the author. Department of Genetics, University of Washington). Neighbor-Joining trees were produced from Kimura-2 parameter distance matrices and from modified Nei–Gojobori (Jukes–Cantor) distance matrices for synonymous mutations only. Maximum likelihood was done with a randomized input order of sequences (10 replicates), global rearrangements, the most thorough search option, empirical transition/transversion ratios and empirical base frequencies. Bootstrap tests were done for 1000 replicates in NJ and 100 for ML. Interior branch tests of phylogeny were done for the NJ trees in Mega.

Recombination was investigated through NJ networks (NeighborNet) obtained by SplitsTree 4-Beta (Huson, 1998) using all strains and with heterozygous sequences at two or more genes coded using the standard IUPAC codes for combinations of two bases.

3. Results

3.1. Amplification of enzyme genes

All PCRs amplified the entire coding region and often part of the intergenic region of each gene in the *L. donovani* complex, yielding the expected sizes: approximately 1.3 kb for *asat*, 2.0 kb for *gpi*, 1.15 kb for *nh1*, 1.2 kb for *nh2* and 1.5 kb for *pgd*. PCR for *nh1* was specific for the *L. donovani* complex, but all other PCRs amplified the reference strains of *L. major* and *L. tropica*. Each PCR was also tested for the other Old World species *L. aethiopica*, *L. arabica*, *Leishmania gerbilli*, *L. killicki*, and *L. turanica* (gel photos available on request). A PCR product was obtained for *nh2* in all species. All species generated a PCR product for *gpi*, except *L. gerbilli*, and *asat* (with primers F2new and R2new) except *L. turanica* and *L. gerbilli*. Only *L. arabica* and *L. killicki* yielded a *pgd* PCR fragment.

PCR products were obtained from as little as 0.25 ng of extracted genomic DNA (gel photos available on request).

Table 3
Polymorphisms in the *Leishmania donovani* complex

Gene	Strain example	MON	Phenotype	Nucleotide position/nucleotide															Allele	pI	MMass
asat				216	522	**823**															
	IPT-1	1	100	C	G	**T**													1	6.78	46119,77
	DEVI	2	113	C	G	**G**													2	6.49	46071,68
	HU3 (LV9)	18	113	C	A	**G**													3	6.49	46071,68
	Francesca	27	100	TC	G	**T**													1/4		
gpi				**28**	456	477	570	687	690	708	759	796	1020	1350	**1487**	1506	1791	1815			
	LEM75	1	100	**G**	C	C	T	T	C	G	T	T	C	C	**G**	T	C	G	1	6.08	67067,35
	LEM2298	183	115	**G**	C	C	T	T	C	G	T	T	C	C	**T**	T	C	G	2	6.00	67024,32
	DEVI	2	100	**G**	T	T	T	T	T	C	C	C	T	C	**G**	T	C	G	3	6.08	67067,35
	LRC-L57	37	100	**G**	C	C	T	T	C	G	T	T	C	G	**G**	C	C	G	4	6.08	67067,35
	HU3 (LV9)	18	100	**G**	C	C	T	C	C	G	T	T	C	C	**G**	T	T	G	5	6.08	67067,35
	Bumm3	3	86	**A**	C	C	T	C	G	T	T		C	C	**G**	T	C	G	6	6.26	67066,41
	Wangjie1	35	86	**A**	C	C	T	C	G	T	T		C	C	**G**	T	C	C	7	6.26	67066,41
	IPT-1	1	100	**G**	C	C	T	T	C	G	T		C	C	**G**	T	C	A	8	6.08	67067,35
nh1				**74**	**82**	309	**431**	**574**	645	792											
	IPT-1	1	100	**A**	**G**	C	**A**	**A**	T	A									1	6.16	34237,76
	LEM3249	29	140	**A**	**G**	C	**A**	**G**	T	A									2	6.00	34238,74
	LLM175	198	140	**A**	**G**	T	**A**	**G**	T	G									3	6.00	34238,74
	Bumm3	3	140	**A**	**G**	C	**A**	**G**	T	G									4	6.00	34238,74
	Wangjie1	35	150	**G**	**G**	C	**A**	**G**	T	G									5	6.00	34211,71
	LRC-L53	36	100	**A**	**A**	C	**G**	**G**	C	G									6	6.16	34194,74
	LRC-L57	37	140	**A**	A	C	A	**G**	C	G									7	6.00	34252,77
nh2				**170**	201	306	480	546	**562**	651											
	IPT-1	1	100	**T**	C	A	A	C	**T**	C									1	4,54	39139,39
	ISS800	188	100	**T**	C	A	A	C	**C**	C									2	4.60	39113,36
	Gilani	30	100	**T**	C	A	G	C	**T**	C									3	4,54	39139,39
	Hussen	31	100	**T**	A	A	G	C	**T**	C									4	4,54	39139,39
	Wangjie1	35	100	**A**	C	G	A	T	**T**	T									5	4.54	39155,39
Pgd				**142**	194	432	678	747	**864**	890	976	1404									
	IPT-1	1	100	**G**	C	C	A	C	**C**	A	G	C							1	5.61	51944,24
	Gebre1	82	100	**G**	C	C	G	C	**C**	A	G	C							2	5.61	51944,24
	DEVI	2	93	**G**	C	C	G	C	**C**	A	A	C							3	5.80	51943,26
	MRC74	2	93	**A**	C	C	G	C	**C**	A	G	T							4	6.00	51943,30
	LRC-L53	36	100	**G**	C	C	G	C	**C**	A	G	T							5	5.61	51944,24
	3S	81	100	**G**	T	C	G	T	**C**	A	G	C							6	5.61	51972,30
	Wangjie1	35	100	**G**	C	C	G	C	**A**	A	G	C							7	5.61	51944,24
	LEM3429	257	100	**G**	**T**	*C*	G	T	C	*AG*	**G**	C							6/8		
	LEM3463	258	100	**G**	**T**	*CT*	G	T	C	*AG*	**G**	C							6/9		

Representative strains for each allele sequence and strains with heterozygous phenotypes are shown. In bold: sites that cause an amino acid polymorphism.

3.2. DNA sequences

The PCR products were directly sequenced and the sequences were deposited in GenBank/EMBL with the following numbers: *asat* AJ620791 to AJ620823, AM117184 to AM117191 and AM157715 to AM157720, *gpi* AJ620617 to AJ620646, AM11792 to AM117193 and AM157721 to AM157725, *nh1* AJ620688 to AJ620721 and AM157726 to AM157730, *nh2* AJ621066 to AJ621102, AM117504 to AM117506 and AM157731 to AM157735, and *pgd* AM157136 to AM157174 and AM157736 to AM157739. The sequences were also deposited in the *Leishmania* genotyping database (www.leishdomus.org). All strains had the expected coding region length: *asat* (1239 bp), *gpi* (1818 bp), *nh1* (945 bp), *nh2* (1059 bp) and *pgd* (1440 bp) for all the listed *L. donovani* complex strains and the other Old World species that were amplified and sequenced. No deletions or insertions were detected. The average GC content was higher than 50% for all genes: *asat* 61.7%, *gpi* 55.8%, *nh1* 60.0%, *nh2* 59.2% and *pgd* 61.4%.

Only two alternative bases were found at each polymorphic site across all strains, with the exception of site 1815 in *gpi* (Table 3). The transition/transversion ratios were mostly above 2 (*gpi* 4.0, *nh2* 2.5, *nh1* 11.7 and *pgd* 9.5) but lower in *asat* (1.1).

The number of single nucleotide polymorphic sites found within the *L. donovani* complex varied between genes, from the lowest to the highest: *asat* had 0.24% (three sites), *pgd* 0.63% (nine sites), *nh2* 0.66% (seven sites), *nh1* 0.74% (seven sites) and *gpi* 0.82% (15 sites). The nonsynonymous/synonymous rate ratio (dN/dS) also varied, but was below 1 in all genes: *nh2* 0.07, *gpi* had 0.08, *pgd* 0.22, *nh1* 0.26 and *asat* 0.36. On a *Z*-test of selection, neutrality was only rejected, against conservative selection, for *gpi*.

Polymorphic sites were distributed throughout each gene with no evident clusters, however, the gamma parameter varied. It was particularly high for *asat*, *nh1*, *nh2* and *pgd* (100), indicating very low variation of substitution rates among sites, but very low for *gpi* (0.010), meaning high variation of substitution rates.

3.3. Alleles

The polymorphisms found generated 4, 8, 7, 5 and 7 alleles for *asat*, *gpi*, *nh1*, *nh2* and *pgd*, respectively (Table 3). All recorded phenotypes could be divided into two or more alleles, with the exception of GPI-115 for which there was a single allele.

Some alleles were specific to certain genetic groups (Table 1), *asat-2* in Indian MON-2 strains; *gpi-3*, *gpi-4* and *gpi-5* in, respectively, Indian MON-2 strains, Kenyan strains and SC23, and Sudanese strains; *nh1-1* and *nh1-2* in *L. infantum* and *nh1*-5 and *nh1-6* in Kenyan strains; *nh2-1* in *L. infantum; pgd-1*, *pgd-3*, *pgd-4* and *pgd-5* in, respectively, *L. infantum*, Indian, Kenyan and Sudanese strains.

Alleles at some loci were shared across more than one genetic group. Most *L. infantum* strains and some Sudanese *L. donovani* strains shared *asat-1* and *nh1-7*. Some *L. infantum* and the *L. donovani* strain Marz-Krim shared *gpi-8* and *nh2-2*. Alleles *nh1-3* and *pgd-2* were present in different *L. donovani* groups, with the latter in *L. infantum* as well (Table 1).

3.4. Linkage phase determination strategies

We found here heterozygous gene sequences, which were detected as split peaks from direct sequencing in both directions. In these cases, it was necessary to determine allelic phase. For allelic phase determination we developed, what is as far as we know, a new procedure based on allele-specific PCR. Allele-specific primers were designed for each SNP at each site, so that they would amplify in opposite directions across each pair of sites. Amplicons were only produced for those SNPs that were present on the same contiguous DNA template, thus determining phase (Fig. 1). A faint non-specific band was occasionally detected but was normally eliminated at higher annealing temperatures (data not shown).

This new strategy was used here for two heterozygous sites if these were located at more than 100 bp from each other, such as in *asat* and *pgd*, although it is also applicable by successive PCRs to multiple pairs of adjacent sites. For *gpi* of LEM3946 we applied a combined strategy, as there was a group of three heterozygous sites within 700 bp on the 5′ end of the gene and a single heterozygous site near the 3′ end of the gene (Table 3), outside the sequencing range. In this case, we used the PCR method to determine phase between the furthermost heterozygous sites and then the sequencing strategy to determine phase between the closest sites.

3.5. Heterozygosity

We identified, for all the genes studied, strains with heterozygous sites (Table 3). Some strains had more than one heterozygous site in a single gene, in particular the Sudanese strain LEM3946 for *pgd* and *gpi* but also phenotype 110 for *asat*, LEM3463 and Gilani for *pgd* and LLM373 for *gpi.*

Heterozygous genes in some strains shared one or more alleles with alleles in homozygous strains belonging to the same genetic group, as previously defined by us (Mauricio et al., 2001, 2004). *Leishmania infantum* strains LEM189, ISS1036, LLM373 and Francesca had corresponding alleles found only in other *L. infantum* strains. Strains Gebre 1, LEM3429 and LEM3463 had corresponding alleles recorded only in other Sudanese *L. donovani*. Strain LEM3946, however, had one set of alleles shared with the Sudanese genetic group, with LEM3429 (MON-257) the most similar strain, and a second set shared with the Ethiopian/Saudi Arabian genetic group, with genotype 26 the most similar (Table 1). It should be noted, however, that strain LEM3429 has alleles that are not found in LEM3946.

Two heterozygous strains had unique alleles; LEM3429 and LEM3463 from Sudan had alleles *pgd-8* and *pgd-9*, which were not found in other strains studied here.

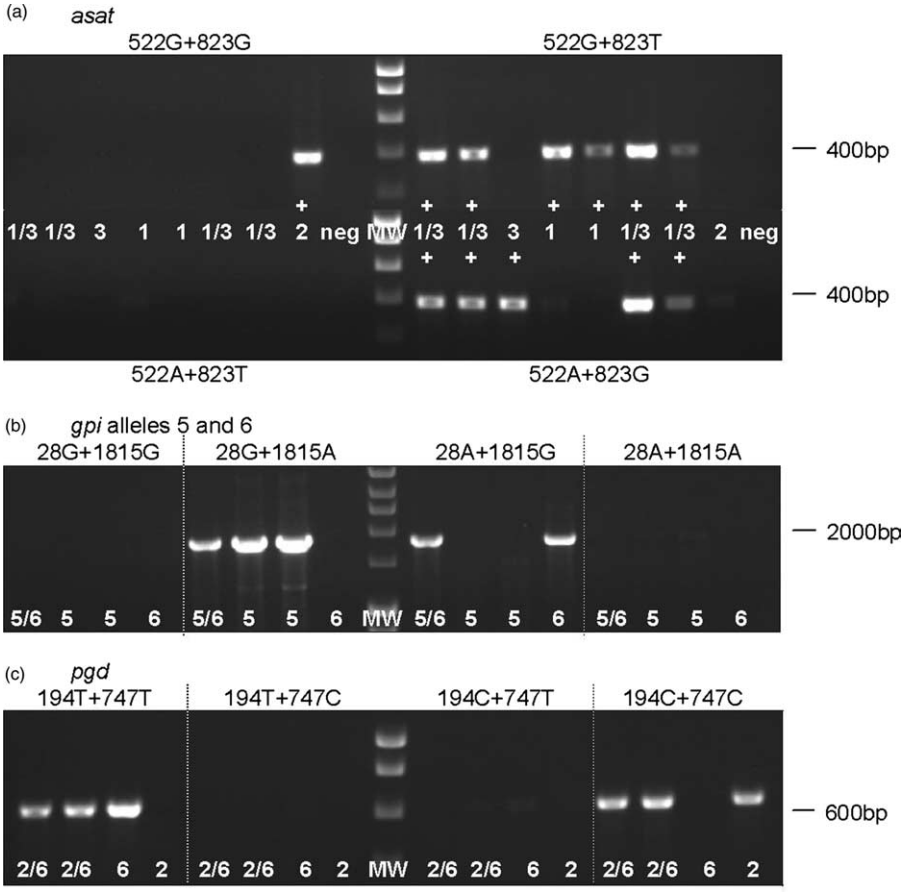

Fig. 1. Amplification products for phase determination: (a) *asat*; (b) *gpi* alleles 5 and 6; (c) *pgd*. Allele composition for each specimen is indicated. MW is molecular mass marker and sizes of bands nearest to the PCR product are shown. Haplotypes are indicated for each primer combination. Primer combinations (forward+reverse, respectively) are indicated at the top of each image (bottom also for *asat*), with the location of their 3′ end base indicated.

3.6. Relationship between derived protein sequences and expected phenotypes

In most cases, the amino acid polymorphisms found in each gene matched the expected phenotypes (Table 3). Some phenotypes were mainly due to differences in protein charge (e.g. GPI phenotypes 86 and 100) and others were due to a change in molecular mass (e.g. NH1 phenotypes 140 and 150). In two strains (3S from Sudan and Mutinga H9 from Kenya), a phenotype of 150 for NH1 was expected (Rioux et al., 1990) but they had alleles *nh1-3* and *nh1-7*, respectively, also found in strains with phenotype 140 in Sudan and Kenya, respectively.

Some strains with the same phenotype for a given enzyme had different protein sequences and the shared phenotype was caused by different amino acid polymorphisms. In one instance, the difference in charge was compensated by the change in molecular mass to give similar phenotypes (e.g. NH1 phenotype 100 for MON-1 and MON-36). PGD phenotype 93 for the two MON-2 Indian strains and MRC74 had different pI, which was expected to impart a motility change, but the electrophoresis buffer of pH 7.4 did not allow distinction between the two phenotypes, as indicated by the expected charge versus pH curve. The enzyme NH2 has only one recorded phenotype, but we identified three different amino acid sequences. In this case, the differences in charge or molecular mass either compensated each other or were not sufficient to produce distinct phenotypes (Table 3).

Some strains were heterozygous at non-synonymous sites and on isoelectric focusing produced a three-banded pattern characteristic of dimeric enzymes (data not shown), for which only the more intense middle band had previously been scored from thick-starch gel electrophoresis. Phenotype 110 of ASAT, NH1-130, GPI-95/80 and 105 were all heterozygous at a single non-synonymous site.

3.7. Combined genotypes

Twenty three (79.3%) genotypes representing a combination of the five genes were unique to a single zymodeme, so that it is possible to determine a high percentage of the known zymodemes with the DNA sequence of only these five genes rather than by the customary MLEE analysis with 15 enzymes. Zymodemes MON-81 and MON-267 had the same genotype (19), as did MON-31 and MON-3 (26) and MON-36 and MON-38 (14). The two zymodemes for which we analysed two or more strains could be subdivided by polymorphisms: MON-1 with *gpi*, into genotypes 1 and 2, and MON-2 with *pgd*, into genotypes 13 and 15.

3.8. Phylogenetic analyses

Given the low number of polymorphisms in each gene, we constructed haplotype networks (Fig. 2) for each gene using statistical parsimony, which is more appropriate to study phylogenetic relationships between sequences with little diversity and for which ancestors may still be extant. The method could not integrate outgroup sequences, as they were too divergent, so the networks could not be rooted in this way. Most adjacent haplotypes diverged by one mutation or even two, but *gpi-3* and *nh2-5* differed at, respectively, four and seven sites. In the absence of haplotype frequency, the method chose as ancestral haplotypes those with the most connections. *Leishmania infantum* haplotypes appeared for three genes as putative ancestor and *L. donovani* only for *nh1*. Strains present in the same region, such as in Sudan and in Kenya, and the *L. infantum* strains had haplotypes that were directly related but strikingly, strain SC23 (MON-38) from India consistently appeared with Kenyan haplotypes often genetically distant from the others from India, which was particularly evident for *gpi* (Fig. 2(b)). *Leishmania donovani* strain Marz-Krim had *L. donovani asat* and *nh1* alleles but also *gpi-8* and *nh2-1*, which are terminal *L. infantum* haplotypes, not directly related to *L. donovani*. Some heterozygous diplotypes were composed of two closely

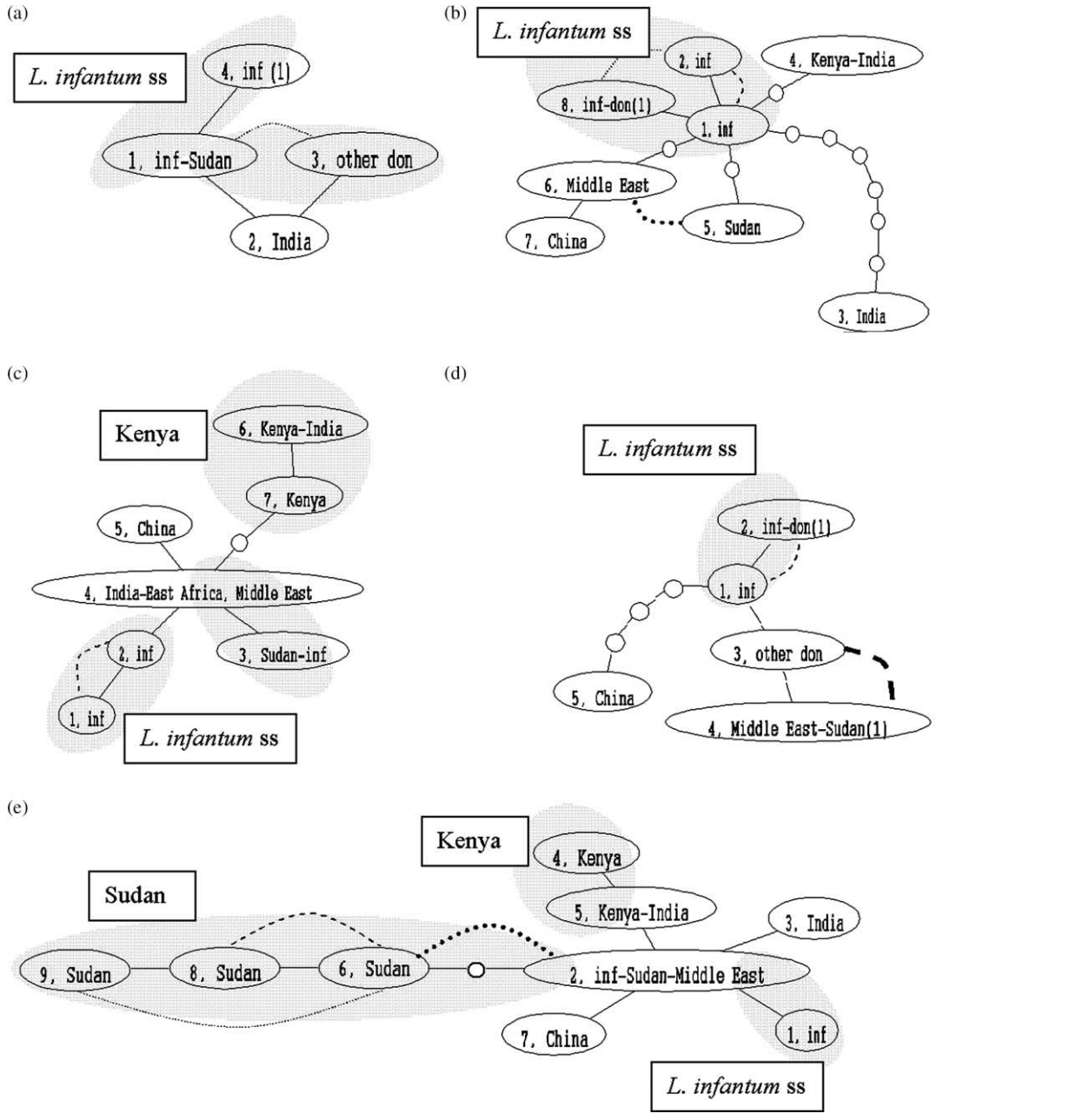

Fig. 2. Haplotype network trees for each gene; (a) *asat*, (b) *gpi*, (c) *nh1*, (d) *nh2*, (e) *pgd*. Broken curved lines represent combinations of alleles present in heterozygotes, with dashed lines connecting directly related alleles and dotted lines connecting genetically distant alleles, with thicker dashed or broken lines for the putative hybrid. Each circle represents one mutation, so that small circles represent missing intermediate alleles. Grey elipses represent populations or genetic groups with more than one sympatric allele.

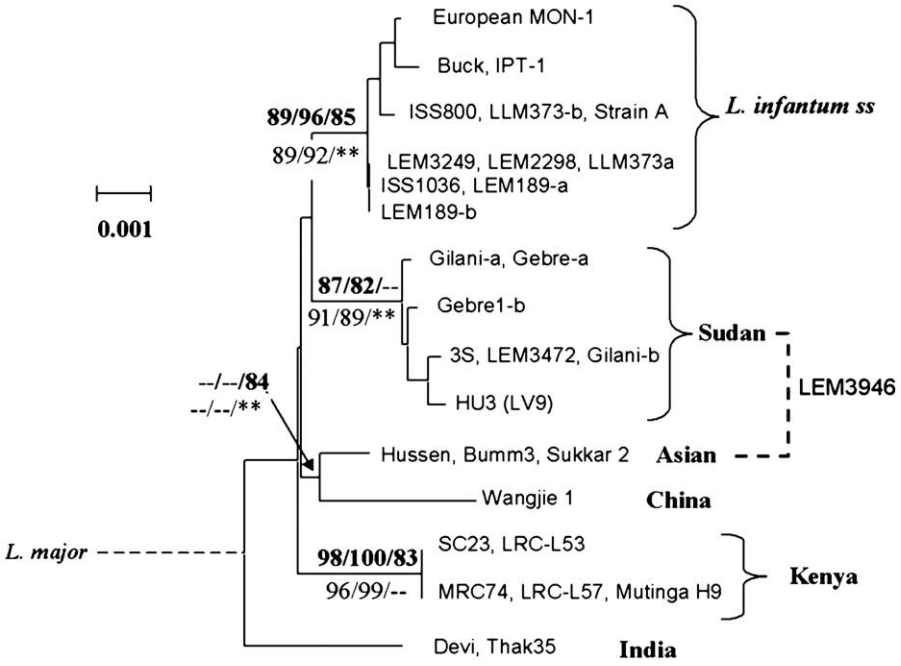

Fig. 3. Unrooted neighbour-joining (NJ) tree of combined *asat*, *gpi*, *nh1*, *nh2* and *pgd* coding regions. Distances were calculated using the modified Nei–Gojobori (Jukes–Cantor) method using synonymous mutations only with a transition/transversion ratio of 2, and a bootstrap test of 1000 replicates (values above 80% are shown in bold). Strains heterozygous at a single gene were decomposed into haplotypes (a and b). Strains were excluded for better tree definition but included in the network (Fig. 4): those heterozygous at two or more genes (Francesca, LEM3429, LEM3463 and LEM3946) or with mosaic gene composition (LLM175 and Marz–Krim). The tree is rooted by *Leishmania major* strain Friedlin genome sequences (branch not drawn to scale). The broken line represents the putative parental groups of the putative hybrid (LEM3946). Bootstrap values are shown above branches for, respectively, NJ trees using synonymous mutations, using all sites (1000 replicates) and maximum likelihood (ML) (100 replicates). Below branches are interior branch tests for the two types of NJ trees and significance in the ML tree ($P < 0.01$, **). All groups were congruent. European MON-1 includes all MON-1 strains, except IPT-1.

related haplotypes (e.g. *gpi-1/8*, *nh1-1/2* and *nh1-2/8*, *nh21/2* and *nh2-3/4*, *pgd-2/6* and *pgd-6/8*), however, others had two distantly related haplotypes (*asat-1/3*, *gpi-2/8*, *pgd-6/9*).

The concatenated data for all genes produced trees (Fig. 3) with robust groups but not with a robust inner topology. Strains with more than one heterozygous gene were not included because it was not possible to determine combined haplotypes across separate, unrelated loci, only within each locus. Strains LLM175 and Marz–Krim, which had alleles typical of different groups, were not included because they would have been pulled between two different groups and, thus, would disrupt the major groups. The groups identified were: *L. infantum sensu stricto*, 'Kenya', 'Sudan', 'India' and 'Asia'. To investigate the reasons behind the lack of support for the inner branches we used a networking method. A reticulate pattern was found mainly between the *L. infantum*, the Sudan group and strains from Saudi Arabia and the Middle East (Fig. 4), which may be caused by recombination between strains. Networking demonstrated that strains LLM175 and Marz-Krim and LEM3946 were indeed in intermediate positions between groups.

4. Discussion

Here we have amplified and sequenced the genes *asat*, *gpi*, *nh1*, *nh2* and *pgd* for five enzymes used in isoenzyme typing, or MLEE, of *Leishmania* for the *L. donovani* complex and several other Old World species (see Section 2). This is the most comprehensive study of the genetic diversity for these five genes, applied here primarily to the *L. donovani* complex. All the genes, with the exception of *gpi*, seem to evolve under neutrality. These genes are, thus, appropriate for multilocus sequence typing (MLST).

For all five enzymes we have identified polymorphisms that discriminate between indistinguishable MLEE phenotypes. Most remarkably, a silent polymorphism in *gpi* distinguishes strains of MON-1. This is an important finding with epidemiological applications, as MON-1 is the most common zymodeme of *L. infantum* in Europe.

Some alleles, such as *asat-2* for Indian MON-2, were specific to a genetic group and could thus be used as a diagnostic character. Other alleles were present in different groups, which could be due to maintenance of conserved alleles or to genetic exchange between the groups.

To avoid cloning heterozygous genes and to confirm that heterozygosity was not a result of PCR errors, we developed and described here a new method for phase determination. The faint band detected in some strains was possibly due to incomplete specificity of the primer at the PCR temperature employed or due to *Taq* polymerase slippage between templates. We have also used, if preferable, allele-specific DNA sequencing. An alternative method, sequencing of bands purified from single strand conformation polymorphism electrophoresis was not needed here. The method we describe here is suitable for rapid determination of the phase of these genes, as required in population studies.

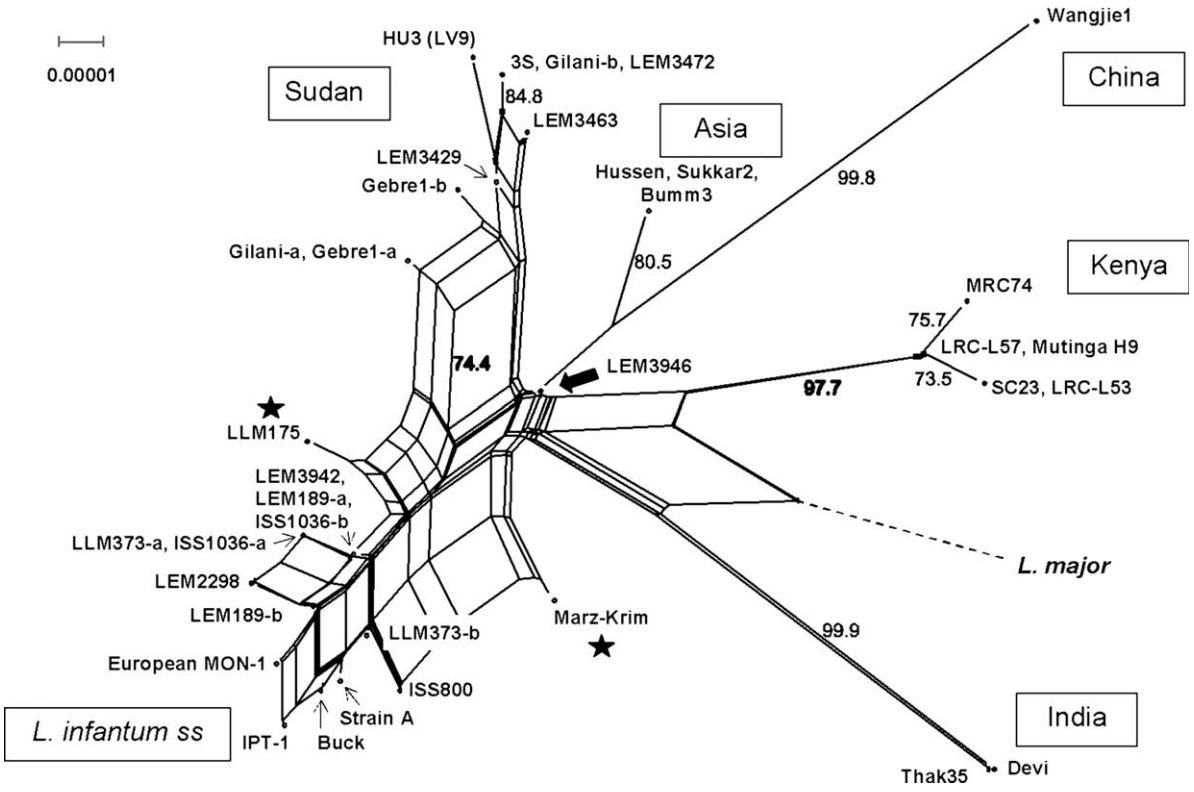

Fig. 4. Network built with Neighbor–Net using complete DNA sequences for *asat*, *gpi*, *nh1*, *nh2* and *pgd* coding regions, with 1000 bootstrap replicates. IUPAC codes for two bases were used for heterozygous sites. Distances were calculated using the Kimura-2-parameter. All strains were included and haplotypes were used where possible. The tree is rooted by *Leishmania major* Friedlin genome sequences (branch not to scale). Stars indicate strains with genome mosaics; the thick arrow indicates the putative hybrid between the Sudanese and the Middle Eastern strains. A network of diploid sequences is similar and networks without strains LEM3946, Marz–Krim and LLM175 retain the splits within each group but less between groups (trees not shown). European MON-1 includes all MON-1 strains except IPT-1.

Most isoenzyme phenotypes could be explained by the amino acid polymorphisms detected. Isoenzyme phenotype can thus be determined through DNA sequencing or by targeted typing of the sites responsible for phenotypic differences. In most cases the different electrophoretic mobilities were caused by charge differences between amino acids, in others by a change in amino acid size altering the charge/mass ratio.

Some indistinguishable phenotypes were due to distinct genotypes, which for DNA-based enzyme typing provides information not available through MLEE. An important example is strain MRC74 from Kenya, which had been typed as MON-2, the same zymodeme as most Indian strains, however, the indistinguishable PGD phenotype was due to different amino acid polymorphisms.

Unfortunately, NH1 phenotype 150 in zymodemes MON-32 and MON-81 was not matched by any amino acid polymorphisms in comparison with related genotypes. As a consequence it has not been possible here to distinguish those zymodemes from MON-37 and MON-267, respectively. It may be that NH1 in these strains suffers from post-translational modifications, that the different phenotypes from isoenzyme typing were electrophoretic artefacts, or that there is a second, not amplified (null), allele that participates in the secondary structure of the dimeric enzyme and changes its electrophoretic motility. It is also conceivable that we did not characterise the correct gene, but this is a remote possibility, given that matches were perfect for all other cases.

We have included in this analysis all different phenotypes, inclusively from different genetic groups. We have here identified a total of 26 different genotypes with these five enzyme coding genes, with which we were able to identify most zymodemes even though they are based on 15 enzymes in total. Only three pairs of zymodemes could not be distinguished. Supplementary genes are required to discriminate these zymodemes, as achieved for genotype 19 by Zemanova et al. (unpublished data; see below).

On the other hand, we were able to distinguish between strains with the same zymodeme, as for MON-1 and MON-2. More SNPs might be found within the same zymodeme by analysis of more genes and, possibly, of more strains. The enhanced discrimination requires new designations. To build on the established MLEE nomenclature, we propose adding additional numbers in lower case roman numerals and italicizing MON to signify genotype instead of phenotype. For example, the two different genotypes for zymodeme MON-2 become *MON-2i* (India, here genotype 13) and *MON-2ii* (Kenya, here genotype 15), whilst for MON-1 become *MON-1i* (LEM75, here genotype 1) and *MON-1ii* (IPT-1, here genotype 2). This nomenclature can be distinguished from that of alleles because it uses the prefix MON and the second number is a roman numeral.

Analysis of individual genes was best achieved through statistical parsimony to build haplotype networks. This method assigns ancestry to the most frequent haplotype, but this is not applicable to a non-populational sample such as ours. Furthermore, there was no outgroup sufficiently close. The programme, instead, took the haplotype with the most connections to be the ancestral one. Using these criteria, the networks disagreed on the group with the ancestral haplotypes, although, with the exception of *nh1*, they were all present in *L. infantum*. No tree had a clear division between *L. infantum* and *L. donovani* haplotypes, which disagrees with previous MLEE-based trees (Rioux et al., 1990) but agrees with molecular data (Ibrahim and Barker, 2001; Mauricio et al., 2001, 2004; Jamjoom et al., 2004; Kuhls et al., 2005).

As expected, most directly related haplotypes were found in the same geographical regions but there were some exceptions. All haplotypes in strain SC23 from India are typical of Kenya and this is probably a case of migration, whereas the presence of one allele typical of *L. infantum* in Sudan/Ethiopia and two in strain Marz-Krim suggests genetic exchange rather than migration. There were cases of heterozygotes with alleles that were directly related, whereas others had pairs of more distant alleles. The latter suggests that one of the haplotype groups is more likely to have been acquired through lateral gene transfer/genetic exchange, rather than by direct mutation. Most haplotypes were connected to others by one or two mutations at most, which would be expected, but the branches of *gpi-3* (India) and *nh2-5* (China) had many more. It is possible that intermediate haplotypes were present in lost populations or populations not included in this analysis or that the ancestors of these strains suffered rapid adaptive evolution with expansion of these extant strains.

A phylogenetic tree of the combined data for the five genes generated groups with high bootstrap values. These groups are the same as previously described using ITS, mini-exon (Mauricio et al., 2004) and *gp63* intergenic RFLP data (Mauricio et al., 2001) and have thus been confirmed with independent data. The groups are: (i) *L. infantum sensu strictu* (i.e. without MON-30, MON-81 and MON-267 from Sudan/Ethiopia), (ii) strains from Kenya (including MON-38 from India—SC23—and one Kenyan MON-2), (iii) MON-2 strains from India and (iv) strains from Sudan and the Ethiopian border (i.e. to include *L. archibaldi* and MON-30, MON-81 and MON-267). The Sudanese group was the least supported probably due to the *asat* alleles similar to *L. infantum*. By studying the gene sequences of five enzymes used in MLEE, we have thus confirmed that the current separation of the zymodemes present in the Sudanese/Ethiopian leishmaniasis focus into three different species is not appropriate and that they form a single genetic group, as has been proposed by several researchers (Ibrahim and Barker, 2001; Mauricio et al., 2001, 2004; Jamjoom et al., 2004). The name '*L. archibaldi*' is thus not validated.

Despite the robustness of the genetic groups, we found evidence of genome mosaics. Addition of strains LLM175 and Marz–Krim, which have more marked genome mosaic structure, disrupted the tree and lowered support for the *L. infantum* group. In addition, the topologies of the individual gene trees (data not shown) were not completely congruent. For example, MON-30, MON-81 and MON-267 from Sudan did cluster with *L. infantum* for *asat*, but not for other genes. Furthermore, a Neighbor–Net analysis produced a network between the *L. infantum* sensu strictu, the Sudanese group and Middle Eastern strains. This type of mosaic structure between groups suggests that genetic exchange has had at least historical importance in *Leishmania*.

Heterozygous sites were detected for a substantial number of isolates. This heterozygosity was confirmed by the detection of characteristic three banded patterns for non-synonymous sites and dimeric enzymes. A two band pattern would have been observed for mixed populations of homozygotes. Heterozygosity indicated by DNA sequencing at synonymous sites cannot be confirmed by phenotype. DNA sequencing of biological clones would confirm heterozygosity. As strains studied here were not biological clones it is conceivable that some apparently heterozygous DNA sequences might be due to polyclonal populations. However, this is extremely unlikely because: all, but one, non-synonymous sites were confirmed phenotypically to be heterozygous; split peaks and phase determinations gave no indication of the presence of minority clones; furthermore, in a parallel project with *Trypanosoma cruzi* biological clones revealed similar heterozygosity at synonymous sites (M. Yeo, unpublished data).

Many isolates had two or more heterozygous sites. Heterozygosity can be caused by mutation in one allele but it can also be caused by genetic exchange between strains with different alleles. Mutation is more likely for single heterozygous sites but recombination is a more parsimonious explanation for two or more sites. In contrast to the suggestion of Jamjoom et al. (2004), we propose that genetic exchange is a more likely explanation for the heterozygous *asat* than an independent and relatively recent mutation, given that we found the same two linked heterozygous *asat* sites in Sudanese isolates and the alleles are separated by two mutational events.

Genetic exchange is even more likely in two strains with three heterozygous genes, such as Francesca and, particularly, LEM3946, for which all heterozygous sites had precise homozygous matches in the Ethiopian/Saudi Arabian group and the Sudanese group. Heterozygosity in these strains was confirmed in at least one non-synonymous site. We thus propose that LEM3946 is a result of genetic exchange between strains from two different genetic groups. Strains of genotype 26 approximate to one of the putative parents, with strains from Ethiopia (MON-31) being the most likely candidates due to geography. The best candidate parent from the Sudanese group is strain LEM3429 (MON-257), which has all alleles that are present in the hybrid. However, this candidate parent also has other alleles at heterozygous loci not present in the hybrid, which could be due to loss of alleles upon genetic exchange or because the hybrid is not the actual F1 and the parental strain was not included in this analysis. This is the first putative hybrid described for the *L. donovani* complex, although there are several other *Leishmania* spp. hybrids described in the

literature (Belli et al., 1994; Russell et al., 1999; Ravel, Pratlong, Campino, Cortes, Lami, Serres, Dedet, 2005. Detection of three *Leishmania infantum-L. major* hybrid strains from three HIV infected patients in Portugal. Worldleish 3rd World Congress on Leishmaniosis, Palermo-Terrasini, Italy).

We did not find homozygous equivalents among other isolates for alleles in two Sudanese strains heterozygous for *pgd*. This may be due to heterozygosity resulting from point mutation, biased sampling or extinction of the predicted homozygotes.

Recombination was also suggested by the mosaic genome structure found in the *L. donovani* complex. Notably, most alleles found in heterozygous loci matched alleles found in other isolates, i.e. no intragenic recombination was observed and thus seems uncommon over short genetic distances.

Recombination between species or even genetic groups implies an historical phenomenon but within groups suggests events occurring within populations in relatively recent times. Most of the newly discovered zymodemes of *L. infantum* are heterozygous and have been found in HIV patients, with suggested syringe transmission (Cruz et al., 2002). Similarly, a high degree of heterozygosity has been found in the Sudan, where major visceral leishmaniasis epidemics have recently occurred (Desjeux, 2004). These are both cases where mixed infections are more likely to occur and thus create the opportunity for genetic exchange. In addition, recombinant parasites may have a selective advantage and thus become more common than parental strains.

So far the importance of genetic exchange in *Leishmania* has been played down (Tibayrenc and Ayala, 2002), mainly because sufficiently discriminatory markers were not available. Overall, although we did not find definitive evidence, genetic exchange emerged as a plausible and in some cases as the best, explanation for the data. More work should be done to investigate this hypothesis and housekeeping genes such as these are excellent candidates.

We have here defined the diversity for five enzyme-coding genes in the *L. donovani* complex. Polymorphisms responsible for observable phenotypes, as well as those capable of further discrimination were found. This DNA-based approach to *Leishmania* enzyme typing can be expanded immediately to several other Old World *Leishmania* species, particularly the main agents of cutaneous leishmaniasis, *L. major* and *L. tropica*, with the exception of *nh1*.

We recommend that DNA sequencing should be used for comprehensive typing, using internal primers for sequencing. This methodology can already be used by laboratories for preliminary typing of isolates, in house, by sequencing commercially or by despatching extracted DNA to be typed. Data can be compared with data from other laboratories without running reference strains in each experiment. Comparative data can be easily produced through web accessible databases. Care is required to recognise heterozygotes and/or mixed cultures as described above. This method is thus more accessible and simpler than MLEE and the sensitivity of the PCRs makes it applicable to smaller numbers of parasites in culture than MLEE, and also to clinical samples. The process of traditional *Leishmania* typing by MLEE is thus moving towards a new generation of MLST for strain typing. Here, we reported results for the genes encoding five enzymes. The results for another five enzyme genes will shortly be published (Zemanová et al., unpublished data), which, collectively, will provide enhanced discrimination compared with MLEE.

Acknowledgements

This work was mainly supported by the EC (grant QLK2-CT-2001-01810) and also by the Grant Agency of the Czech Academy of Sciences (grant Z60220518). The help of Ales Horak (Institute of Parasitology, České Budejovice) with data analysis is gratefully acknowledged.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijpara.2006.03.006.

References

Belli, A.A., Miles, M.A., Kelly, J.M., 1994. A putative *Leishmania panamensis/Leishmania braziliensis* hybrid is a causative agent of human cutaneous leishmaniasis in Nicaragua. Parasitology 109, 435–442.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659.

Cruz, I., Morales, M.A., Noguer, I., Rodriguez, A., Alvar, J., 2002. *Leishmania* in discarded syringes from intravenous drug users. The Lancet 359, 1124–1125.

Desjeux, P., 2004. Leishmaniasis: current situation and new perspectives. Comp. Immunol. Microbiol. Infect. Dis. 27, 305–318.

Gu, X., Zhang, J., 1997. A simple method for estimating the parameter of substitution rate variation among sites. Mol. Biol. Evol. 14, 1106–1113.

Hare, M.P., Palumbi, S.R., 1999. The accuracy of heterozygous base calling from diploid sequence and resolution of haplotypes using allele-specific sequencing. Mol. Ecol. 8, 1750–1752.

Huson, D., 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14, 68–73.

Ibrahim, M.E., Barker, D.C., 2001. The origin and evolution of the *Leishmania donovani* complex as inferred from a mitochondrial cytochrome oxidase II gene sequence. Infect. Gen. Evol. 1, 61–68.

Jamjoom, M.B., Ashford, R.W., Bates, P.A., Chance, M.L., Kemp, S.J., Watts, P.C., Noyes, H.A., 2004. *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. Parasitology 129, 399–409.

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

Kuhls, K., Mauricio, I.L., Pratlong, F., Presber, W., Schonian, G., 2005. Analysis of ribosomal DNA internal transcribed spacer sequences of the *Leishmania donovani* complex. Microbes Infect. 7, 1224–1234.

Kumar, S., Tamura, K., Nei, M., 1994. MEGA: molecular evolutionary genetics analysis software for microcomputers. Comput. Appl. Biosci. 10, 189–191.

Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86–93.

Mauricio, I.L., Gaunt, M.W., Stothard, J.R., Miles, M.A., 2001. Genetic typing and phylogenetic analysis 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., 2004. *Leishmania donovani* complex: genotyping with the ribosomal internal transcribed spacer and the mini-exon. Parasitology 128, 1–5.

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.

Nei, M., Kumar, S. (2000) Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Pratlong, F., Dereure, J., Bucheton, B., El-Saf, S., Dessein, A., Lanotte, G., Dedet, J.P., 2001. Sudan: the possible original focus of visceral leishmaniasis. Parasitology 122, 599–605.

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. Ann. Parasitol. Hum. Comp. 65, 111–125.

Russell, R., Iribar, M.P., Lambson, B., Brewster, S., Blackwell, J.M., Dye, C., Ajioka, J.W., 1999. Intra and inter-specific microsatellite variation in the *Leishmania* subgenus *Viannia*. Mol. Biochem. Parasitol. 103, 71–77.

Tibayrenc, M., Ayala, F.J., 2002. The clonal theory of parasitic protozoa: 12 years on. Trends. Parasitol. 18, 405–410.

Zemanova, E., Jirku, M., Mauricio, I.L., Miles, M.A., Lukes, J., 2004. Genetic polymorphism within the *Leishmania donovani* complex: correlation with geographic origin. Am. J. Trop. Med. Hyg. 70, 613–617.