Sequence variation of the Cytochrome *b* gene of various human infecting members of the genus *Leishmania* and their phylogeny

G. E. LUYO-ACERO[1], H. UEZATO[1]*, M. OSHIRO[2], K. TAKEI[2], K. KARIYA[2], K. KATAKURA[3], E. GOMEZ-LANDIRES[4], Y. HASHIGUCHI[5] *and* S. NONAKA[1]

[1] *Department of Dermatology, Faculty of Medicine, University of the Ryukyus, 207 Uehara, Nishihara-cho, Okinawa 903-0125, Japan*
[2] *Division of Cell Biology, Graduate School of Medicine, University of the Ryukyus, 207 Uehara, Nishihara-cho, Okinawa 903-0125, Japan*
[3] *Department of Parasitology, Gunma University School of Medicine 3-39-22 Showa-machi, Maebashi 371-8511, Japan*
[4] *Departmento de Medicina Tropical, Facultad de Medicina, Universidad Catolica Santiago de Guayaquil 10833, Guayaquil, Ecuador*
[5] *Department of Parasitology, Kochi Medical School, Nankoku city, Kochi 783-8505, Japan*

(*Received 5 February 2003; revised 20 October 2003; accepted 23 October 2003*)

SUMMARY

The Cytochrome *b* (*Cyt b*) gene has proved to be useful for identification and classification of many mammals and plants. In order to evaluate the utility of this gene for discrimination of *Leishmania* parasites as well as for exploring their phylogenetic relationships, we determined the nucleotide sequences of the *Cyt b* gene from 13 human-infecting *Leishmania* species (14 strains) from the New and Old Worlds. The *Cyt b* genes, approximately 1080 base pairs, were found to be A/T rich, and their 5′ terminal-editing regions were highly conserved. The nucleotide sequence variation among them was enough to discriminate parasite species; 245 nucleotide positions were polymorphic and 190 positions were parsimony informative. The phylogenetic relationships based on this gene, showed good agreement with the classification of Lainson & Shaw (1987) except for the inclusion of *L.* (*L.*) *major* in the *L.* (*L.*) *tropica* complex and the placement of *L. tarentolae* in another genus. These data show that the *Cyt b* gene is useful for phylogenetic study of *Leishmania* parasites.

Key words: *Leishmania*, Cytochrome *b*, phylogeny.

INTRODUCTION

The genus *Leishmania* comprises 30 members of which 21 species are pathogenic to mammals including humans, causing a variety of diseases called leishmaniasis, which affect 88 countries worldwide, mostly developing countries (WHO, 2000).

Although biochemical (isoenzyme electrophoresis) and immunological (monoclonal antibodies) techniques have been utilized to identify species and subspecies of *Leishmania* parasites, these methods are time consuming (Kreutzer & Christensen, 1980; Grimaldi, David & McMahon-Pratt, 1987).

Since analysis of DNA gives the most specific and stable identification criteria, molecular approaches have also been used. In order to increase the sensitivity of their methodologies, most researchers have targeted genes and regions of genes that are present in high copy number (Uliana *et al.* 1991; Van Eys *et al.* 1992; Meredith *et al.* 1993; Fernandes *et al.* 1994; Cupolillo *et al.* 1995). However, the discriminatory power of these targets depends on the level of inter- and intra-species variability of their sequences (Van Eys *et al.* 1992; Van Eys & Meredith, 1996).

In kinetoplastid protozoa, the maxicircle component of the kinetoplast is homologous to the mitochondrial genome of other organisms and has about 50 copies (Stuart, 1983). The cytochrome *b* (*Cyt b*) gene is contained in the mitochondrial genome of a wide variety of living forms and encodes the central catalytic subunit of an enzyme present in the respiratory chain of mitochondria. The gene has been widely used for phylogenetic studies and identification of animals and plants (Irwin, Kocher & Wilson, 1991; Degli Esposti *et al.* 1993).

The *Cyt b* gene of *L. tarentolae* consists of two regions, the edited region (the most 5′ region of 23 bp) that undergoes RNA editing, and the non-edited region (the 3′ region of 1056 bp). The RNA editing process has been described in mitochondrial genes of kinetoplastid protozoa (Benne, 1994). In the RNA editing process, uridylate (U) residues are inserted or deleted to repair a frameshift present in the genomic sequence and to create the AUG codon for

* Corresponding author: Department of Dermatology, Faculty of Medicine, University of the Ryukyus, 207 Uehara, Nishihara-cho, Okinawa 903-0125, Japan. Tel: +81 98 895 1153. Fax: +81 98 895 1417. E-mail: huezato@med.u-ryukyu.ac.jp

Parasitology (2004), **128**, 483–491. © 2004 Cambridge University Press
DOI: 10.1017/S0031182004004792 Printed in the United Kingdom

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press

Table 1. *Leishmania* strains used in this study, gene size, A + T, G + C content and frequencies of the cytochrome *b* gene

(E. R., edited region; Non-E. R., non-edited region; F, frequency.)

Species	International code	*Cyt b* (bp)	E. R.	Non-E. R.	E. R.		Non-E. R.	
					A + T content (F)	G + C content (F)	A + T content (F)	G + C content (F)
L. (*L.*) *donovani*	MHOM/SD/62/2S-25M-C2	1079	23	1056	11 + 1 (0·48 + 0·04)	9 + 2 (0·39 + 0·09)	285 + 526 (0·27 + 0·50)	169 + 76 (0·16 + 0·07)
L. (*L.*) *infantum**	MHOM/TN/80/IPT1	1079	23	1056	11 + 1 (0·48 + 0·04)	9 + 2 (0·39 + 0·09)	284 + 528 (0·27 + 0·50)	170 + 74 (0·16 + 0·07)
L. (*L.*) *chagasi**	MHOM/BR/74/PP75	1080	24	1056	11 + 2 (0·46 + 0·08)	9 + 2 (0·38 + 0·08)	280 + 529 (0·27 + 0·50)	171 + 76 (0·16 + 0·07)
L. (*L.*) *tropica**	MHOM/SU/58/Strain OD	1080	24	1056	11 + 2 (0·46 + 0·08)	9 + 2 (0·38 + 0·08)	273 + 535 (0·26 + 0·51)	174 + 74 (0·16 + 0·07)
L. (*L.*) *major**	MHOM/SU/73/5ASKH	1080	24	1056	11 + 2 (0·46 + 0·08)	9 + 2 (0·38 + 0·08)	264 + 538 (0·25 + 0·51)	180 + 74 (0·17 + 0·07)
L. (*L.*) *aethiopica**	MHOM/ET/72/L100	1080	24	1056	11 + 2 (0·46 + 0·08)	9 + 2 (0·38 + 0·08)	277 + 529 (0·26 + 0·50)	171 + 79 (0·16 + 0·08)
L. (*L.*) *mexicana**	MHYC/BZ/62/M379	1079	23	1056	10 + 1 (0·43 + 0·04)	10 + 2 (0·43 + 0·09)	274 + 545 (0·26 + 0·52)	165 + 72 (0·16 + 0·07)
L. (*L.*) *amazonensis**	MHOM/BR/73/M2269	1078	22	1056	11 + 0 (0·50 + 0)	9 + 2 (0·41 + 0·09)	274 + 551 (0·26 + 0·52)	164 + 67 (0·16 + 0·06)
L. (*L.*) *garnhami**	MHOM/VE/76/JAP78	1079	23	1056	11 + 1 (0·48 + 0·04)	9 + 2 (0·39 + 0·09)	275 + 551 (0·26 + 0·52)	163 + 67 (0·15 + 0·06)
L. (*V.*) *braziliensis**	MHOM/BR/75/M2904	1078	22	1056	12 + 0 (0·55 + 0)	8 + 2 (0·36 + 0·09)	287 + 548 (0·27 + 0·52)	158 + 63 (0·15 + 0·06)
L. (*V.*) *braziliensis**	MHOM/EC/88/INH-03	1078	22	1056	12 + 0 (0·55 + 0)	8 + 2 (0·36 + 0·09)	286 + 548 (0·27 + 0·52)	159 + 63 (0·15 + 0·06)
L. (*V.*) *panamensis**	MHOM/BR/71/LS94	1078	22	1056	12 + 0 (0·55 + 0)	8 + 2 (0·36 + 0·09)	286 + 546 (0·27 + 0·52)	160 + 64 (0·15 + 0·06)
L. (*V.*) *guyanensis**	MHOM/BR/75/M4147	1078	22	1056	12 + 0 (0·55 + 0)	8 + 2 (0·36 + 0·09)	291 + 544 (0·28 + 0·51)	157 + 64 (0·15 + 0·06)
L. (*L.*) *major*-like	MHOM/EC/88/PT-115	1080	24	1056	11 + 2 (0·46 + 0·08)	9 + 2 (0·38 + 0·08)	264 + 539 (0·25 + 0·51)	180 + 73 (0·17 + 0·07)
L. tarentolae†		1079	23	1056	12 + 1 (0·52 + 0·04)	8 + 2 (0·35 + 0·09)	287 + 519 (0·27 + 0·49)	166 + 84 (0·16 + 0·08)

* WHO Reference strain.

† Nucleotide sequence as deposited in GenBank accession no. M10126.

Table 2. Primer designation and consensus sequences designed from species belonging to the order Kinetoplastida

Name of primer	Primer sequence	Species‡	Accession no.	nt positions†	Forward/reverse
COIIIF	taatacgactcactataGTTTATATTGACATTTTGTWGATT*	*L. tarentolae*[a]	M10126	5271-5294	Forward
		T. brucei[b]	M20379	819-842	Forward
		B. culicis[c]	U05814	998-1021	Forward
LCBF1	GGTGTAGGTTTTAGTTTAGG	*L. tarentolae*[a]	M10126	5456-5475	Forward
		T. brucei[b]	M17998	100-119	Forward
		T. borreli[d]	U11684	91-110	Forward
LCBR2	CTACAATAAACAAATCATAATATACAATT	*L. tarentolae*[a]	M10126	6321-6293	Reverse
		T. brucei[b]	M17998	965-937	Reverse
		T. borreli[d]	U11684	959-930	Reverse
MURF4R	gggttttcccagtcacgacgAATCTCTCTCTCCCTT*	*L. tarentolae*[a]	M10126	6609-6594	Reverse
		L. (L.) mexicana[e]	AF118655	86-71	Reverse
		L. (V.) panamensis[f]	AF118653	90-75	Reverse
		L. (L.) donovani[g]	AF118654	85-70	Reverse

* Lower case bases do not form part of the consensus sequences.
† Nucleotide positions as given in their respective Accession numbers.
‡ [a]*Leishmania tarentolae*, [b]*Trypanosoma brucei*, [c]*Blastocrithidia culicis*, [d]*Trypanoplasma borreli*, [e]*Leishmania (L.) mexicana*, [f]*Leishmania (V.) panamensis*, [g]*Leishmania (L.) donovani*.

translation initiation. In *L. tarentolae*, the transcript of the *Cyt b* gene undergoes insertion of 39 U residues in 15 sites within the edited region. Consequently, (1) a total of 20 codons are created including the AUG codon (Met) for translation initiation by insertion of one U residue between positions 2 and 3 in the edited region, and (2) the ATT codon (Leu) becomes the first amino acid in the non-edited region (Feagin *et al*. 1988).

In the present study, in an attempt to evaluate the utility of the *Cyt b* gene for discriminating *Leishmania* species and for reconstructing their phylogenetic relationships, we determined the coding sequence of this gene from 13 human-pathogenic species (14 strains). These species represent the most common causative agents of leishmaniasis in the New and Old Worlds.

MATERIALS AND METHODS

Parasites

The *Leishmania* strains used in this study are listed in Table 1. *L.* (*L.*) *chagasi* PP75 was kindly provided by Dr M. Hide (IRD de Montpellier, Laboratory CEPM UMR CNRS/IRD 9926, Cedex 5, France).

Cell culture and DNA extraction

Promastigotes were grown at 26 °C in RPMI 1640 medium (Sigma, USA) supplemented with 10% heat-inactivated foetal bovine serum (Bio Whittaker, USA), 50 U/ml penicillin and 50 μg/ml streptomycin. Promastigotes were harvested at the stationary phase of growth by centrifugation at 2000 ***g*** for 10 min. Genomic DNA was extracted from the promastigote pellets by using GenomicPrep™ Cell and Tissue DNA Isolation Kit (Amersham Pharmacia Biotech, USA) following the manufacturer's instructions.

Design of PCR primers and PCR conditions

For PCR amplification of the whole *Cyt b* gene from *Leishmania* parasites, oligonucleotide primers were designed on the basis of consensus sequences found in *Cyt b* and adjacent genes: *COIII* and *MURF4* of parasites belonging to the order kinetoplastida (Table 2). Positions and directions of the primers are indicated in Fig. 1. PCR reactions were done in a total volume of 50 μl. Each reaction mixture contained 400 ng of DNA template, 100 pM of each primer, 0·2 mM of each dNTP, 1·25 units of Ex *Taq* polymerase, 10 mM Tris–HCl, pH 8·3, 50 mM KCl and 1·5 mM $MgCl_2$ (TaKaRa, Japan). PCR conditions were as follows: initial denaturation at 94 °C for 1 min followed by 40 cycles of denaturation at 94 °C for 1 min, annealing at 50 °C for 1 min and extension at 72 °C for 1 min and a final extension step at 74 °C for 5 min.

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press

Table 3. *Leishmania* cytochrome *b* gene internal primers used for sequencing

Name of primer	Primer sequence	Forward/reverse
LCBF2	GTTACCATGTACAATGATGTC	Forward
LCBF3	TAATACGACTCACTATAGATAARTTTACTGGWTTA*	Forward
LCBF3′	GCWGTRCCWGATAARTTTACTGG*	Forward
LCBF4	TGTTATTGAATATGAGGTAGTG	Forward
LCBF5	GTACCTGATAAATTTACTGGT	Forward
LCBR1	TCTGGTAAAATTTTATCAGATGTTTT	Reverse
LCBR3	GGGTTTTCCCAGTCACGACGTACAWATAAAACAACTAAAA*	Reverse
LCBR3′	ACAAAATAYCARTTWGTACAWATAAAACAAC*	Reverse
LCBR4	GAACTCATAAAATAATGTAAACAAAA	Reverse
LCBR5	TTTGTACAAATAAAACAACTAAAAAA	Reverse

* The degenerate primers contain the degenerate bases R = G or A,W = A or T, Y = T or C.

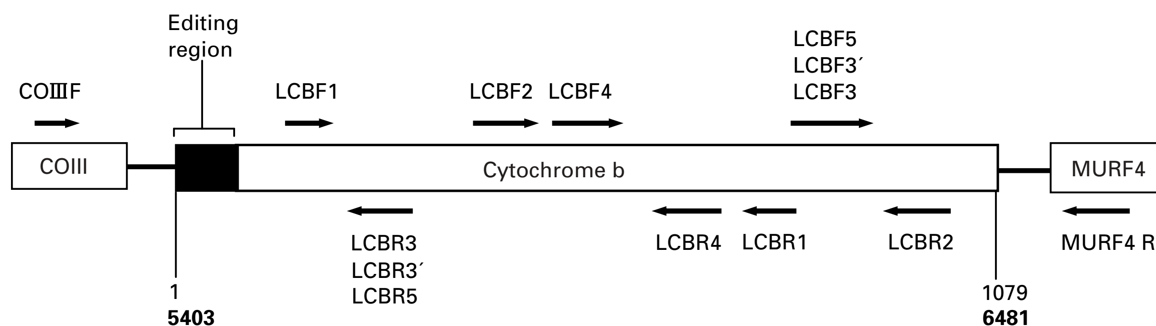


Fig. 1. Schematic representation of the *Cyt b* gene in the maxicircle component of kinetoplastid parasites, indicating the positions of the oligonucleotide primers. COIII and MURF4 represent the genes cytochrome *c* oxidase subunit III and maxicircle unidentified reading frame 4 respectively. LCBF and LCBR represent forward and reverse primers. Bold numbers, 5403 and 6481, indicate the first and the last nucleotide positions of the pre-edited *Cyt b* gene of *L. tarentolae* (Accession no. M10126) corresponding to the nucleotide positions 1 and 1079 in our analysis. (■) Editing region (23 bp). (—) Intergenic region. (□) Non-edited region of *Cyt b* gene.

Subcloning and DNA sequencing

The PCR products were separated by electrophoresis on agarose gels, purified by using the GFX PCR DNA and Gel Band purification kit (Amersham Pharmacia Biotech, USA) and ligated to pT7 blue T-vector (Novagen, USA) by using TaKaRa DNA ligation solution version 1 (TaKaRa, Japan). At least 3 clones for each PCR product were sequenced on both strands. DNA sequencing was carried out on an ABI PRISM 310 automated sequencer (Applied Biosystem, USA) by using the Big Dye terminator cycle sequencing ready reaction kit (Applied Biosystem, USA). DNA sequencing was performed by using the 2 vector-specific primers, the 4 gene-specific primers for subcloning and 10 internal primers (Table 3). The composite sequences containing the *Cyt b* genes flanked by the 5′ and 3′ intergenic sequences were assembled.

Phylogenetic analysis and trees

Sequences were assembled and edited with the program Genetyx Mac 11.0.0 (Software Development Co. Ltd, Japan). Phylogenetic analysis and trees: neighbour joining (NJ) and maximum parsimony (MP) were performed by using the MEGA 2.1 program (Kumar *et al.* 2001) available online at http://www.megasoftware.net. The nucleotide sequence data reported in this paper have been submitted to the DDBJ/EMBL/GenBank nucleotide sequence databases with the Accession numbers AB095957–AB095970.

RESULTS

Structure of the Cyt b *genes of* Leishmania *species*

The first and the last nucleotide positions of the *Cyt b* gene of each tested species/strains were determined based on comparison with the *Cyt b* gene sequence of *L. tarentolae*.

For each species/strain, the first nucleotide position of the edited region was designated as the first nucleotide position of the *Cyt b* gene (position 1), and the third nucleotide of the translation termination codon within the non-edited region was designated as the last nucleotide position of the *Cyt b* gene (positions 1078, 1079 or 1080). Table 1 shows the length, total nucleotide content as well as percentage of A + T and G + C contents of the *Cyt b* gene in each species. The *Cyt b* genes of *Leishmania* parasites were revealed to be A + T rich.

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press

Table 4. Amino acid and nucleotide sequences percentage identity between different species/strains of *Leishmania*

(The numbers above the diagonal are amino acid sequence identity excluding the 5′ end edited region of *L. tarentolae* and the numbers below the diagonal are nucleotide sequence percentage identity. Underlined numbers represent amino acid and nucleotide sequence percentage identity of species belonging to the same complex.)

	L. donovani complex			*L. tropica* complex			*L. mexicana* complex			*L. braziliensis* complex					
Species	*L. don*	*L. inf*	*L. cha*	*L. tro*	*L. maj*	*L. aeth*	*L. mex*	*L. ama*	*L. garn*	*L. b* M2904	*L. b* INH03	*L. pan*	*L. guy*	*L. m*-like	*L. tarent**
L. don		100	100	99·7	99·4	100	99·1	99·1	99·1	96·3	96·6	96·6	96·3	99·4	98·6
L. inf	99·5		100	99·7	99·4	100	99·1	99·1	99·1	96·3	96·6	96·6	96·3	99·4	98·6
L. cha	98·9	98·8		99·7	99·4	100	99·1	99·1	99·1	96·3	96·6	96·6	96·3	99·4	98·6
L. tro	92·8	92·9	93·1		99·1	99·7	98·9	98·9	98·9	96·0	96·3	96·3	96·0	99·1	98·9
L. maj	92·6	92·9	93·0	94·1		99·4	98·6	98·6	98·6	96·3	96·6	96·6	96·3	100	98·0
L. aeth	92·7	92·8	92·8	96·0	93·1		99·1	99·1	99·1	96·3	96·6	96·6	96·3	99·4	98·6
L. mex	91·3	91·4	91·2	91·4	90·8	91·8		100	100	95·4	95·7	95·7	95·4	98·6	98·3
L. ama	91·3	91·4	91·2	91·5	90·9	91·3	97·6		100	95·4	95·7	95·7	95·4	98·6	98·3
L. garn	91·4	91·3	91·3	91·7	91·0	91·5	97·7	99·7		95·4	95·7	95·7	95·4	98·6	98·3
L. bra M2904	90·3	90·1	90·1	89·4	89·2	89·3	89·4	89·8	89·9		99·7	99·7	99·4	96·3	96·3
L. bra INH-03	90·4	90·2	90·2	89·5	89·3	89·4	89·5	89·9	90·0	99·8		100	99·7	96·6	96·6
L. pan	90·1	89·9	89·9	89·5	89·0	89·2	89·4	89·8	89·9	98·7	98·9		99·7	96·6	96·6
L. guy	90·2	90·0	90·0	89·5	89·0	89·2	89·4	89·8	89·9	98·8	99·0	98·8		96·3	96·3
L. maj-like	92·7	93·1	93·1	94·2	99·9	93·0	91·3	91·0	91·1	89·3	89·4	89·1	89·1		98·0
*L. tarentolae**	89·7	89·5	89·5	89·2	88·6	89·3	90·3	90·4	90·6	90·6	90·8	91·0	91·0	88·7	

* Nucleotide sequence as deposited in GenBank Accession no. M10126.

No. of U residues inserted: 1 1 6 1 2 5 2 8 2 3 3 1 2 1 1 ▼

Species																			
L. tarentolae	A	A	G	C G	G	A G A	G	A	A	A	A G A A A	A	G	G	C	T	T T A	A C T	29
L. donovani	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	.	.	. . .	. . .	29
L. infantum	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	.	.	. . .	. . .	29
L. mexicana	.	.	.	. .	.	. . .	.	G	.	G	. . . . .	.	.	.	.	.	. . .	. . .	29
L. garnhami	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	.	.	. . .	. . .	29

No. of U residues inserted: 1 1 6 1 2 5 2 8 2 3 3 1 2 1 0 ▼

Species																		
*L. tarentolae**	A	A	G	C G	G	A G A	G	A	A	A	A G A A A	A	G	G	C - T	T T A	A C T	29
L. chagasi	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	. T .	. . .	. . .	30
L. tropica	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	. T .	. . .	. . .	30
L. aethiopica	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	. T .	. . .	. . .	30
L. major	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	. T .	. . .	. . .	30
L. major-like	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	. T .	. . .	. . .	30

No. of U residues inserted: 1 1 6 1 2 5 2 8 2 3 3 1 2 1 2 ▼

Species																		
*L. tarentolae**	A	A	G	C G	G	A G A	G	A	A	A	A G A A A	A	G	G	C - T	T T A	A C T	29
L. amazonensis	.	.	.	. .	.	. . .	.	G	.	.	. . . . .	.	.	.	. - -	. . .	. . .	28
L.V. guyanensis	.	.	.	. .	.	. . .	.	.	.	.	. . . . .	.	.	.	. - -	. . .	. . .	28
L.V. panamensis	.	.	.	. .	.	. . .	.	.	.	.	. . . . .	.	.	.	. - -	. . .	. . .	28
L.V. brazil. INH-O3	.	.	.	. .	.	. . .	.	.	.	.	. . . . .	.	.	.	. - -	. . .	. . .	28
L.V. brazil. M2904	.	.	.	. .	.	. . .	.	.	.	.	. . . . .	.	.	.	. - -	. . .	. . .	28

Fig. 2. Alignment of the 5′ edited region of the *Cyt b* gene of all *Leishmania* species/strains analysed in this study compared to *L. tarentolae*. Identical bases are marked as dots (.) and gaps as dashes (-). Numbers and positions of U residues inserted to the *L. tarentolae* edited region are shown. The boundary between the edited and non-edited regions is indicated by an arrowhead. The created initiation codon in the edited region and the first two codons of the non-edited region are underlined.

Nucleotide sequence comparison of the Cyt b *gene among* Leishmania *species*

The nucleotide sequences of the *Cyt b* gene obtained in this study together with the sequence from *L. tarentolae* were compared to each other. The pairwise nucleotide sequence comparison revealed that the inter-species degree of identity ranged from 88·7% to 99·5%. When we compared 2 strains of *L.* (*V.*) *braziliensis*: M2904 and INH-03, their sequences showed 99·8% identity. We also observed that sequences of *L. major*-like isolated in Ecuador and the WHO reference strain *L.* (*L.*) *major* isolated in Turkestan (former USSR) had 99·9% identity as shown in Table 4.

Amino acid sequence comparison of the Cyt b *gene product among* Leishmania *species*

Amino acid sequences of the *Cyt b* gene products, downstream of the edited region, were compared. The pairwise amino acid sequence comparison revealed no variation within species belonging to *L.* (*L.*) *donovani* and *L.* (*L.*) *mexicana* complexes. However, amino acid variation was observed within species belonging to *L.* (*L.*) *tropica* (99·1–99·7%) and *L.* (*V.*) *braziliensis* complexes (99·4–100%) as seen in Table 4.

Multiple alignments of nucleotide and amino acid sequences of the Cyt b *gene*

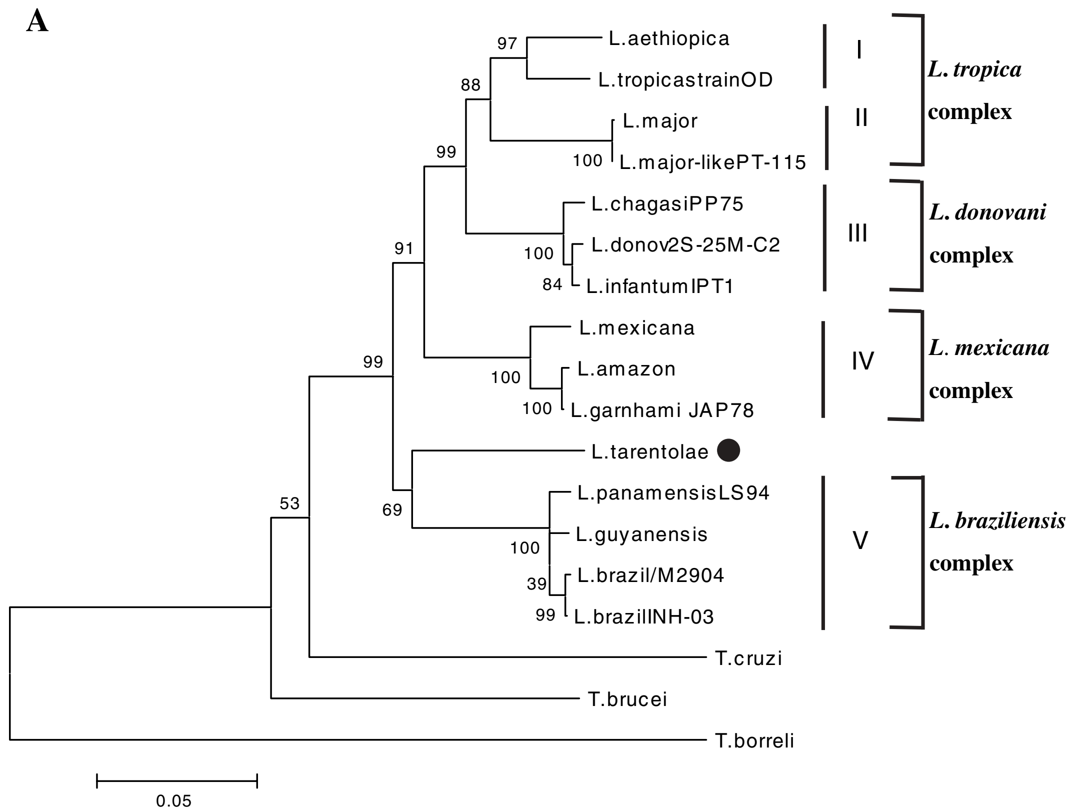
All *Cyt b* gene sequences were aligned. The multiple alignments revealed that 245 nucleotide positions were polymorphic, of which 55 positions were singletons and 190 positions were informative under parsimony criteria, 2 nucleotide positions involved insertion or deletions. These insertions or deletions were observed in the boundaries between the edited (22–24 bp) and the non-edited regions (1056 bp) of the *Cyt b* gene coding sequence (Fig. 2). Alignment of the amino acid sequence corresponding to the non-edited region revealed amino acid substitutions at 19 positions, of which 17 sites were parsimony informative: 3 Leu–Ile substitutions, 11 Val–Ile substitutions (1 singleton), 1 Cys–Ser substitution (singleton), 1 Phe–Tyr substitution, 1 Phe–Leu substitution, 1 Thr–Val substitution and 1 Ala–Val–Ile amino acid substitution. Most of these substitutions (17) were located within 6 out of the 8 transmembrane α helices, helices B, C, D, E, G and H, of the *Cyt b* protein structural model (Howell, 1989).

Comparison of nucleotide frequency and nucleotide substitution with other trypanosomatids

The nucleotide frequencies of all 4 bases were determined. Frequency of base T ranged from 0·49 to 0·52, base A ranged from 0·25 to 0·27, base G ranged from 0·15 to 0·17 and base C ranged from 0·06 to 0·08. Similar nucleotide frequencies for bases T and C have been observed in the partial non-edited coding region of the *Cyt b* gene (516 bp) of trypanosome species belonging to the subgenus *Schizotrypanum* (Barnabe, Brisse & Tibayrenc, 2003).

All kinds of nucleotide substitutions were found in the *Cyt b* gene. It was revealed that the overall transitions (5226) representing 56% were higher than the overall transversions (4074) representing 44%. While transition A⇆G was the most frequent substitution (3135) representing 60% of all transitions, transversion A⇆T was the most frequent (3090) representing 76% of all transversions. An excess of transition A⇆G was also observed by Barnabe *et al.* (2003).

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press



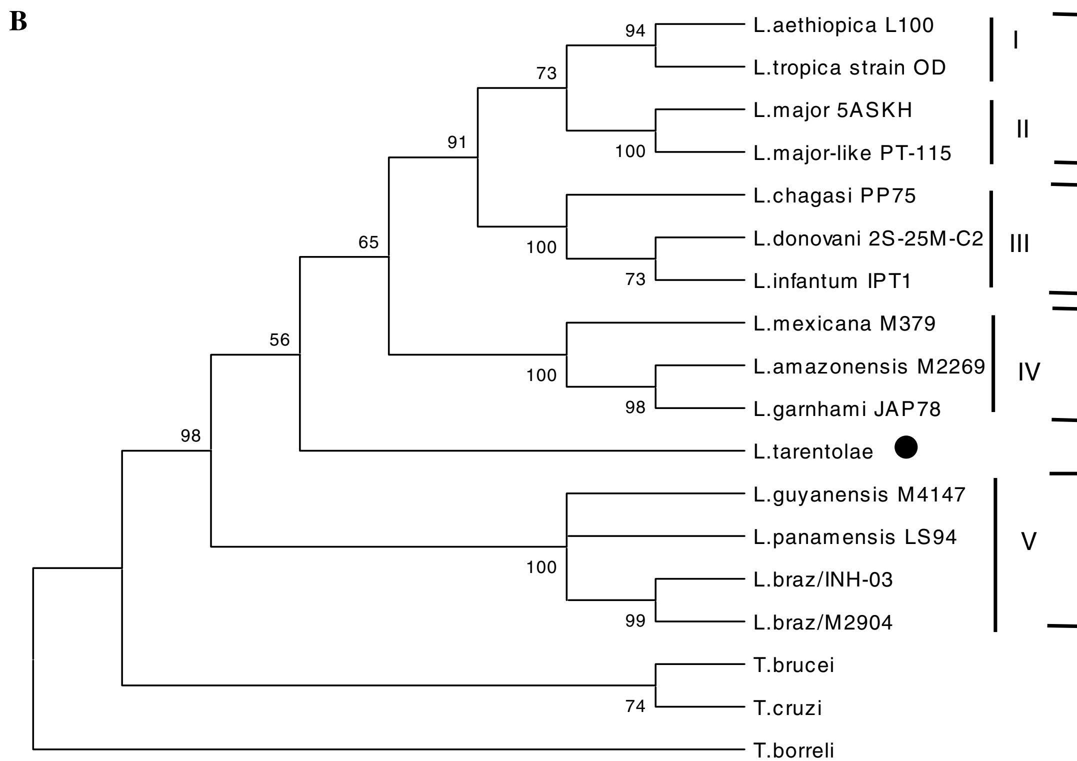


Fig. 3. Phylogenetic relationships of various members of the genus *Leishmania* based on the nucleotide non-edited coding sequences of the *Cyt b* gene. The numbers in the branches correspond to the bootstrap values based on 1000 replicates. (●) Circle indicates the position of *Sauroleishmania* (*L. tarentolae*) species in their respective phylogenetic tree. Bar indicates a clade formed based on the *Cyt b* gene. Bracket indicates *Leishmania* complexes as described by Lainson & Shaw (1987). (A) NJ tree constructed with the Tamura and Nei distance (1993), outgroup = *Trypanosoma cruzi* (U43567), *Trypanosoma brucei* (M17998) and *Trypanoplasma borreli* (U11684). (B) consensus MP tree constructed by using the complete deletion and branch and bound options.

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press

Phylogenetic trees inferred from Cyt b *gene sequences*

All *Leishmania* species analysed in this study were monophyletic. In the NJ and MP trees, the mammalian *Leishmania* clustered into 5 clades. Clade I was represented by *L.* (*L.*) *aethiopica*/*L.* (*L.*) *tropica* and clade II was represented by *L.* (*L.*) *major* (Fig. 3A and B). Clades III, IV and V corresponded to *L.* (*L.*) *donovani*, *L.* (*L.*) *mexicana* and *L.* (*V.*) *braziliensis* complexes respectively, as described by Lainson & Shaw (1987).

DISCUSSION

Taking as a model the *Cyt b* gene of *L. tarentolae* and its RNA editing process, each tested species/strains had a 22–24 bp region very similar to the edited region of *L. tarentolae*, suggesting that similar RNA editing to that of *L. tarentolae* may take place in these *Leishmania* parasites. These edited regions were followed by non-edited regions of 1056 bp starting with the TTA (Leu) codon as in *L. tarentolae*. At their 3′ ends, the edited regions of 22 bp had a deletion of one T residue, while those of 24 bp had an insertion of one T residue, relative to that of *L. tarentolae* (23 bp). We assume that these deletions or insertions may be corrected by the RNA editing process, possibly by adding zero, one, or two U residues just downstream of the position 22 bp in the edited region. This assumption is supported by the observations that (1) non-edited regions of all tested species/strains encoded a highly homologous stretch of 351 amino acids, (2) most amino acid substitutions are biochemically conservative, and (3) conserved amino acids in a wide variety of species (Degli Esposti *et al.* 1993) were also conserved in all the species/strains analysed in this study. The fact that most of the amino acid substitutions were found in the transmembrane regions is congruent with previous studies of the *Cyt b* gene of different species. (Degli Esposti *et al.* 1993; Farias *et al.* 2001).

The DNA sequence variation of the *Cyt b* gene from each of the human-infecting *Leishmania* species/subspecies determined in this study was enough to distinguish each one of them, and allowed us to explore their phylogenetic relationships. In contrast, not all parasites could be distinguished at the amino acid level, and the amino acid-based phylogenetic trees showed low bootstrap values (data not shown), probably due to the lack of informative sites.

The phylogenetic relationships, based on *Cyt b* gene nucleotide sequences, appear to agree with the classification of *Leishmania* proposed by Lainson & Shaw (1987) with two exceptions: (1) the inclusion of *L.* (*L.*) *major* within the *L.* (*L.*) *tropica* complex, because of its notable earlier divergence from the *L.* (*L.*) *tropica*/*L.* (*L.*) *aethiopica* clade and (2) the placement of *L. tarentolae*. Their proposed classification places *L. tarentolae* in another genera (*Sauroleishmania*) on the basis that *L. tarentolae* infects lizards, not mammals. However, the phylogenetic trees reported in this study showed that *Sauroleishmania* and mammalian *Leishmania* formed a well-supported monophyletic group. Our result is consistent with previous studies, using nuclear and mitochondrial molecular markers, that place *Sauroleishmania* within the genus *Leishmania* (Croan, Morrison & Ellis, 1997; Brewster & Barker, 1999).

The NJ and MP trees had similar topologies with a small difference. In the NJ tree *L. tarentolae* and the *L.* (*V.*) *braziliensis* clade formed a monophyletic group that was not highly supported (69% bootstrap value), while in the MP tree they appeared to be paraphyletic. The ambiguous position of *L. tarentolae* in this study may reflect the discrepancy existing in the literature on its position. *L. tarentolae* clustered together with species belonging to subgenus *Viannia* when minicircle conserved regions were used for phylogenetic analysis (Yurchenko, Kolesnikov & Lukes, 2000). In contrast, this species clustered together with species belonging to subgenus *Leishmania*, when the ATPase 6 gene was used (Brewster & Barker, 1999). One explanation for this discrepancy is the possibility of horizontal transfer of maxicircle genes among *Leishmania* species as described in *Trypanosoma cruzi* (Machado & Ayala, 2001). This explanation is not inconsistent with the clonal population structure of *Leishmania* proposed by Tibayrenc, Kjellberg & Ayala (1990), since this theory accepts rare genetic exchange events.

Evolution of RNA editing has been described at inter-genus level in the order kinetoplastida (Landweber & Gilbert, 1994; Maslov *et al.* 1994). In the present study, we observed that the edited region of the *Cyt b* gene was highly conserved among *Leishmania* species. Although the edited regions were too short to draw any conclusion, our observation is consistent with a previous report regarding the first edited region of the ATPase 6 gene (Brewster *et al.* 1999), suggesting that the RNA editing process within the genus *Leishmania* may be evolving at a very slow rate.

We sincerely thank Dr Mallorie Hide for providing the *L.* (*L.*) *chagasi* WHO reference strain, Dr Carlos Machado for his helpful comments on the phylogenetics analyses and Dr Masato Umikawa for his support during the initial phases of this study. This study was supported by grants No. 14256002 from the Japanese Ministry of Education, Culture, Sports, Science and Technology and No. 15590371 from the Japanese Society for the Promotion of Science.

REFERENCES

BARNABE, C., BRISSE, S. & TIBAYRENC, M. (2003). Phylogenetic diversity of bat trypanosomes of subgenus *Schizotrypanum* based on multilocus enzyme electrophoresis, random amplified polymorphic DNA,

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press

and cytochrome *b* nucleotide sequences analyses. *Infection, Genetics and Evolution* **2**, 201–208.

BENNE, R. (1994). RNA editing in trypanosomes. *European Journal of Biochemistry* **221**, 9–23.

BREWSTER, S. & BARKER, D. C. (1999). The ATPase subunit 6 gene sequence predicts that RNA editing is conserved between lizard- and human-infecting *Leishmania*. *Gene* **235**, 77–84.

CROAN, D. G., MORRISON, D. A. & ELLIS, J. T. (1997). Evolution of the genus *Leishmania* revealed by comparison of DNA and RNA polymerase gene sequences. *Molecular and Biochemical Parasitology* **89**, 149–159.

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.

DEGLI ESPOSTI, M., DE VRIES, S., CRIMI, M., GHELLY, A., PATARNELLO, T. & MEYER, A. (1993). Mitochondrial cytochrome *b*: evolution and structure of the protein. *Biochimica et Biophysica Acta* **1143**, 243–271.

FARIAS, I. P., ORTI, G., SAMPAIO, I., SCHNEIDER, H. & MEYER, A. (2001). The cytochrome *b* gene as a phylogenetic marker: The limits of resolution for analyzing relationships among cichlid fishes. *Journal of Molecular Evolution* **53**, 89–103.

FEAGIN, J. E., SHAW, J. M., SIMPSON, L. & STUART, K. (1988). Creation of AUG initiation codons by addition of uridines within cytochrome *b* transcripts of kinetoplastids. *Proceedings of the National Academy of Sciences, USA* **85**, 539–543.

FERNANDES, O., MURTHY, V. K., KURATH, U., DEGRAVE, W. M. & CAMPBELL, D. A. (1994). Mini-exon gene variation in human pathogenic *Leishmania* species. *Molecular and Biochemical Parasitology* **66**, 261–271.

GRIMALDI, G. Jr., DAVID, J. R. & McMAHON-PRATT, D. (1987). Identification and distribution of new world *Leishmania* species characterized by serodeme analysis using monoclonal antibodies. *American Journal of Tropical Medicine and Hygiene* **36**, 270–287.

HOWELL, N. (1989). Evolutionary conservation of protein regions in the proton-motive cytochrome *b* and their possible roles in redox catalysis. *Journal of Molecular Evolution* **29**, 157–169.

IRWIN, D. M., KOCHER, T. D. & WILSON, A. C. (1991). Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution* **32**, 128–144.

KREUTZER, R. D. & CHRISTENSEN, H. A. (1980). Characterization of *Leishmania* spp. by isozyme electrophoresis. *American Journal of Tropical Medicine and Hygiene* **29**, 199–208.

KUMAR, S., TAMURA, K., JAKOBSEN, I. B. & NEI, M. (2001). MEGA2: Molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.

LAINSON, R. & SHAW, J. J. (1987). Evolution, classification and geographical distribution. In *Leishmaniasis in Biology and Medicine* (ed. Peters, W. & Killick-Kendrik, R.), pp. 1–120. Academic Press, London.

LANDWEBER, L. F. & GILBERT, W. (1994). Phylogenetic analysis of RNA editing: A primitive genetic phenomenon. *Proceedings of the National Academy of Sciences, USA* **91**, 918–921.

MACHADO, C. A. & AYALA, F. J. (2001). Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. *Proceedings of the National Academy of Sciences, USA* **98**, 7396–7401.

MASLOV, D. A., AVILA, H. A., LAKE, J. A. & SIMPSON, L. (1994). Evolution of RNA editing in kinetoplastid protozoa. *Nature, London* **368**, 345–347.

MEREDITH, S. E. O., ZIJLSTRA, E. E., SCHOONE, G. J., KROON, C. C. M., VAN EYS, G. J. J. M., SCHAEFFER, K. U., EL-HASSAN, A. M. & LAWYER, P. G. (1993). Development and application of the polymerase chain reaction for the detection and identification of *Leishmania* parasites in clinical material. *Archives de l'Institut Pasteur de Tunis* **70**, 419–431.

STUART, K. (1983). Kinetoplast DNA, mitochondrial DNA with a difference. *Molecular and Biochemical Parasitology* **9**, 93–110.

TAMURA, K. & NEI, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**, 512–526.

TIBAYRENC, M., KJELLBERG, F. & AYALA, F. (1990). A clonal theory of parasitic protozoa: The population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Nigeria*, *Plasmodium*, *Trichomonas* and *Trypanosoma* and their medical and taxonomical consequences. *Proceedings of the National Academy of Sciences, USA* **87**, 2414–2418.

ULIANA, S. R. B., AFFONSO, M. H. T., CAMARGO, E. P. & FLOETER-WINTER, L. M. (1991). *Leishmania*: Genus identification based on a specific sequence of the 18S ribosomal RNA sequence. *Experimental Parasitology* **72**, 157–163.

VAN EYS, G. J. J. M. & MEREDITH, S. E. O. (1996). Detection and characterization of *Leishmania* parasites by DNA-based methods. In *Methods in Molecular Biology* (ed. Clapp, J. P.), pp. 227–242. Humana Press Inc, NJ.

VAN EYS, G. J. J. M., SCHOONE, G. J., KROON, N. C. M. & EBELING, S. B. (1992). Sequence analysis of small subunit ribosomal RNA genes and its use for detection and identification of *Leishmania* parasites. *Molecular and Biochemical Parasitology* **51**, 133–142.

WORLD HEALTH ORGANIZATION (2000). *WHO Information Fact Sheet N-116*. Available from URL: http://www.who.int//inf-fs/en/fact116.html.

YURCHENKO, V., KOLESNIKOV, A. & LUKES J. (2000). Phylogenetic analysis of Trypanosomatina (Protozoa: Kinetoplastida) based on minicircle conserved regions. *Folia Parasitologica* **47**, 1–5.

https://doi.org/10.1017/S0031182004004792 Published online by Cambridge University Press