Parasitol Res (2010) 107:1049–1065
DOI 10.1007/s00436-010-1969-9

ORIGINAL PAPER

Species discrimination and phylogenetic inference of 17 Chinese *Leishmania* isolates based on internal transcribed spacer 1 (ITS1) sequences

Bin-Bin Yang · Xian-Guang Guo · Xiao-Su Hu · Jian-Guo Zhang · Lin Liao · Da-Li Chen · Jian-Ping Chen

Received: 24 April 2010 / Accepted: 18 June 2010 / Published online: 9 July 2010

Abstract Leishmaniasis is a geographically widespread disease, caused by protozoan flagellates of the genus *Leishmania*. This disease still remains endemic in China, especially in the west and northwest frontier regions. To date, the phylogenetic relationships among Chinese *Leishmania* isolates are still unclear, and the possible taxonomic diversity remains to be established. In this study, the ITS1–5.8S fragments of ten isolates collected from different foci in China were determined. To infer the phylogenetic relationships among them, seven sequences of Chinese *Leishmania* isolates retrieved from GenBank were also included. Both parsimony and Bayesian analyses reveal an unexpected but strongly supported clade comprising eight newly determined isolates, which is sister to other members of subgenus *Leishmania*. In combination with genetic distance analysis, this provides evidence of the occurrence of an undescribed species of *Leishmania*. Our results also suggest that (1) the isolate IPHL/CN/77/XJ771 from Bachu County, Xinjiang Uygur Autonomous Region is not *Leishmania infantum* but *Leishmania donovani*; (2) the status referring to an isolate MRHO/CN/88/KXG-2 from a great gerbil in Karamay as *Leishmania turanica*, formerly based on multilocus enzyme electrophoresis, is recognized; (3) an earlier finding demonstrating the *L. donovani* identity of isolate MHOM/CN/80/801 from Kashi city is corroborated; (4) the three isolates from eastern Jiashi County, Xinjiang Uygur Autonomous Region, causing desert type of zoonotic visceral leishmaniasis (see Wang et al., Parasitol Int (in press), 2010), belong to *L. donovani* instead of *L. infantum*. In addition, the results of this study make an important contribution to understanding the heterogeneity and relationships of Chinese *Leishmania* isolates, further indicating that the isolates from China may have had a more complex evolutionary history than expected.

The authors wish it to be known that, in their opinion, the first two authors, Bin-Bin Yang and Xian-Guang Guo, should be regarded as joint first authors.

B.-B. Yang · X.-S. Hu · J.-G. Zhang · L. Liao · D.-L. Chen · J.-P. Chen
Department of Parasitology,
West China School of Preclinical and Forensic Medicine,
Sichuan University,
Chengdu 610041 Sichuan, China

B.-B. Yang · X.-S. Hu · J.-G. Zhang · L. Liao · D.-L. Chen (✉) · J.-P. Chen (✉)
Animal Disease Prevention and Food Safety Key Laboratory of Sichuan Province, Sichuan University,
Chengdu 610064 Sichuan, China
e-mail: cdl1978119@sina.com
e-mail: jpchen007@163.com

X.-G. Guo
Chengdu Institute of Biology, Chinese Academy of Sciences,
Chengdu 610041 Sichuan, China

Introduction

Leishmaniasis is a vector-borne disease transmitted by sand flies, caused by a protozoan flagellate of the genus *Leishmania*. It is endemic in 88 countries on four continents and occurs in at least four major forms: cutaneous, diffuse cutaneous, mucocutaneous, and visceral (Desjeux 2004). The current global estimate of 350 million people at risk of leishmaniasis and of an incidence per year of 500,000 for visceral leishmaniasis (VL) and 1.5 million for cutaneous leishmaniasis (CL) belies higher burdens of disease in endemic foci. In the current context of worldwide (re) emergence and spreading of leishmaniasis, the relevance of species identification further gains importance.

As with many parasites, the taxonomy of the genus *Leishmania* is very complex because species definitions and boundaries are hard to define. The links between clinical disease, geographic origin, and nomenclature of *Leishmania* species have long been controversial and confusing. It has been recognized that the gold standard for taxonomy and strain typing of *Leishmania* based on multilocus enzyme electrophoresis (MLEE, also known as zymodeme typing) has several disadvantages: It requires large culture volumes and specialized equipment, it is costly and laborious, and it has relatively poor discriminatory power (Kuhls et al. 2005; Bañuls et al. 2007). To overcome these problems, DNA-based methods have been widely used for *Leishmania* spp. identification and/or phylogenetic reconstruction with a variety of targets such as protein-coding genes, non-coding segments, microsatellites, and restriction fragment length polymorphisms (e.g., Piarroux et al. 1995; Noyes et al. 1997; El Tai et al. 2001; Ibrahim and Barker 2001; Kuhls et al. 2005; Asato et al. 2009; Fraga et al. 2010; Montalvo et al. 2010). Notably, Lukeš et al. (2007) went a long way to resolving the issues for the *Leishmania donovani* complex. They proposed that this species complex consists of only two species, *L. donovani* sensu stricto (with the synonym *Leishmania archibaldi*) and *L. infantum* (with the synonym *Leishmania chagasi*), by using a combination of widely used DNA analysis techniques and further demonstrated that geographic origin of a strain is a more important predictor of genetic relatedness than the type of disease caused (visceral versus cutaneous leishmaniasis). Few isolates from China, however, were included in these studies, and no work has yet focused on the phylogenetic relationships among different isolates from different epidemic areas in China.

Leishmaniasis remains endemic in China, especially in the west and northwest frontier regions. Both human VL and CL occur in China, with most VL cases, rare CL cases (Guan et al. 1992a; Zheng et al. 2009; Wang et al. 2010). According to different geographical origin, infective agent, and clinical evidences, epidemic foci of VL in China were classified into three types, i.e., plain foci, hill foci, and desert foci (Lu et al. 1994). Different foci and types of leishmaniasis in China have brought forth the difficulty of identifying the strains of *Leishmania*, as the etiological pathogens cannot be distinguished easily on the basis of morphological characteristics. On the basis of MLEE of six representative isolates from the plain, mountainous, and desert regions, Xu et al. (1984) first identified the causative agents responsible for VL as *L. donovani* sensu lato and *L. infantum*. Xu et al. (1984) also provided support for the specific status of *Leishmania gerbilli*, a nonpathogenic parasite of the great gerbil *Rhombomys opimus* in the desert areas of Gansu Province and Xinjiang Uygur Autonomous Region in N.W. China (Wang et al. 1964a, b). Interestingly, Xu et al. (1989) further found that two isolates from kala azar patients in Kashi city of Xinjiang could not be designated definitely as *L. infantum* or *L. donovani* by isoenzyme electrophoresis. Subsequently, on the basis of zymodeme typing, Guan et al. (1995) first reported that *L. turanica* had been found in great gerbils from Xiaoguai Farm in Karamay, Xinjiang Uygur Autonomous Region. Although in recent years considerable work has been done to determine the heterogeneity of different isolates from three types of foci in China by analyzing kDNA and/or nDNA (e.g., Lu et al. 1994; Lu et al. 1997, 1998; Lu et al. 2001; Hu et al. 2002), the phylogenetic and taxonomical relationships among isolates covering China are still unclear. On the basis of kDNA and nDNA heterogeneity, Lu et al. (1994) classified 19 *Leishmania* isolates from epidemiologically different foci in China into five genotypes (groups I–V). They found that members of group II, tentatively designated as *L. infantum* sensu lato, displayed much heterogeneity in both kDNA and nDNAs. They further inferred that the isolates in China were more heterogeneous than previously thought, necessitating the reassignment of some isolates into different groups. Such a perspective is therefore still far from clearly promoting us to pursue further studies of the molecular phylogenetics of *Leishmania*.

Sequence data of the ribosomal RNA (rRNA) gene, in particular the two highly variable internal transcribed spacer regions (ITS1 and ITS2), have been successfully used to resolve taxonomic questions and to determine phylogenetic affinities among closely related *Leishmania* species (e.g., Dávila and Memen 2000; Schönian et al. 2000; El Tai et al. 2001; Kuhls et al. 2005; Parvizi et al. 2008) and other parasites (e.g., Chen et al. 2007; Lin et al. 2007; Al-Kandari and Al-Bustan 2010). Thus, the starting point of our work is to obtain a better understanding of the status of different isolates from different epidemic areas in China by DNA-based taxonomy. The ITS1–5.8S fragments were sequenced to analyze the phylogenetic relationships of ten isolates collected from different foci in China, in conjunction with seven Chinese isolates retrieved from GenBank. The aims of the present study are (1) to report a set of original ITS1–5.8S sequences for ten *Leishmania* isolates from different foci in China, (2) to determine the level of variation among ITS1–5.8S sequences among them, and (3) to infer the phylogenetic relationships among isolates in China and explore the taxonomic implications.

Materials and methods

DNA extraction, amplification, and sequencing protocols

The ten *Leishmania* isolates used in this study were listed in Table 1. Promastigotes were cultivated in medium 199 supplemented with 15% heat-inactivated fetal bovine serum

Table 1 List of *Leishmania* strains, origin, and database accession numbers, including sequences of *L. donovani* complex retrieved from GenBank

Sequence type (sequence numbers)	Sequence length (bp)	GenBank accession numbers	MLEE-based species assignment	WHO code	Origin	Reference
–	313	HM130599	n.d.	MHOM/CN/?/GS6	Gansu, China	This study
–	313	HM130600	n.d.	MCAN/CN/60/GS1	Gansu, China	This study
–	313	HM130601	n.d.	MHOM/CN/90/SC10H2	Sichuan, China	This study
–	313	HM130602	n.d.	MHOM/CN/?/GS5	Gansu, China	This study
–	313	HM130603	n.d.	MHOM/CN/83/GS2	Gansu, China	This study
–	313	HM130604	n.d.	MHOM/CN/84/SD1	Shandong, China	This study
–	311	HM130605	n.d.	MHOM/CN/84/JS1	Jiangsu, China	This study
–	313	HM130606	n.d.	MCAN/CN/?/SC11	Sichuan, China	This study
–	320	HM130607	*L. turinica*[a]	MRHO/CN/88/KXG-2	Karamay, China	This study
H1 (6)	297	AM901450	n.d.	MHOM/IN/1961/L13	India	Alam et al. 2009a
		AJ634360	*L. donovani*	MHOM/ET/00/HUSSEN	Ethiopia	Kuhls et al. 2005
		AJ634359	*L. archibaldi*[b]	MHOM/SD/97/LEM3463	Sudan	Kuhls et al. 2005
		AJ634358	*L. archibaldi*[b]	MHOM/SD/97/LEM3429	Sudan	Kuhls et al. 2005
		AJ634357	*L. archibaldi*[b]	MHOM/SD/93/GE	Sudan	Kuhls et al. 2005
		AJ634356	*L. donovani*	MCAN/SD/00/LEM3946	Sudan	Kuhls et al. 2005
H2 (5)	297	HM130608	n.d.	IPHL/CN/77/XJ771	Bachu, China	This study
		GQ367487	n.d.	MHOM/CN/08/JIASHI-2	Jiashi, China	Wang et al. 2010
		GQ367488	n.d.	MHOM/CN/08/JIASHI-5	Jiashi, China	Wang et al. 2010
		GQ367486	n.d.	MHOM/CN/08/JIASHI-1	Jiashi, China	Wang et al. 2010
		AJ000294	*L. donovani*	MHOM/CN/00/Wangjie1	China	Kuhls et al. 2005
H3 (31)	298	GQ444144	*L. infantum*	MHOM/IR/04/IPI-UN10	Iran	Mahmoudzadeh-Niknam et al. (unpublished data)
		FM164420	*L. infantum*	MHOM/UZ/2007/KU	Uzbekistan	Alam et al. 2009b
		FM164419	*L. infantum*	MHOM/UZ/2007/OBA	Uzbekistan	Alam et al. 2009b
		FM164418	*L. infantum*	MHOM/UZ/2007/ERD	Uzbekistan	Alam et al. 2009b
		FM164417	*L. infantum*	MHOM/UZ/2007/KOM	Uzbekistan	Alam et al. 2009b
		FM164416	*L. infantum*	MHOM/UZ/2007/MUA	Uzbekistan	Alam et al. 2009b
		EU326227	–	MHOM/BR/74/PP75	Brazil	Thakur et al. (unpublished data)
		AM502245	*L. infantum*	MCAN/ES/98/LLM-877	Spain	Peacock et al. 2007
		AJ634355	*L. infantum*	MCAN/ES/86/LEM935	Spain	Kuhls et al. 2005
		AJ634354	*L. infantum*	MHOM/IT/93/ISS800	Italy	Kuhls et al. 2005
		AJ634353	*L. infantum*	MHOM/IT/94/ISS1036	Italy	Kuhls et al. 2005
		AJ634352	*L. infantum*	MHOM/ES/92/LLM373	Spain	Kuhls et al. 2005
		AJ634351	*L. infantum*	MHOM/FR/80/LEM189	France	Kuhls et al. 2005
		AJ634350	*L. infantum*	MHOM/MT/85/BUCK	Malta	Kuhls et al. 2005
		AJ634349	*L. infantum*	MHOM/ES/91/LEM2298	Spain	Kuhls et al. 2005
		AJ634348	*L. infantum*	MHOM/FR/96/LEM3249	France	Kuhls et al. 2005
		AJ634347	*L. infantum*	MHOM/ES/88/LLM175	Spain	Kuhls et al. 2005
		AJ634346	*L. infantum*	MCAN/FR/87/RM1	France	Kuhls et al. 2005
		AJ634345	*L. infantum*	MHOM/CN/54/Peking	Shannxi, China	Kuhls et al. 2005
		AJ634344	*L. infantum*	MHOM/PT/00/IMT260	Portugal	Kuhls et al. 2005
		AJ634343	*L. infantum*	MHOM/ES/86/BCN16	Spain	Kuhls et al. 2005
		AJ634342	*L. infantum*	MHOM/FR/97/LSL29	France	Kuhls et al. 2005
		AJ634341	*L. infantum*	MHOM/ES/93/PM1	Spain	Kuhls et al. 2005
		AJ634340	*L. infantum*	MHOM/FR/95/LPN114	France	Kuhls et al. 2005
		AJ634339	*L. infantum*	MHOM/FR/78/LEM75	France	Kuhls et al. 2005

Table 1 (continued)

Sequence type (sequence numbers)	Sequence length (bp)	GenBank accession numbers	MLEE-based species assignment	WHO code	Origin	Reference
		AJ000306	*L. chagasi*[c]	MHOM/BR/85/M9702	Brazil	Kuhls et al. 2005
		AJ000304	*L. chagasi*[c]	MHOM/BR/74/PP75	Brazil	Kuhls et al. 2005
		AJ000303	*L. infantum*	MHOM/CN/78/D2	Xinjiang, China	Kuhls et al. 2005
		AJ000295	–	MHOM/ES/87/Lombardi	Spain	Schöenian et al. (unpublished data)
		AJ000289	*L. infantum*	MHOM/TN/80/IPT1	Tunisia	Kuhls et al. 2005
		AJ000288	*L. infantum*	MHOM/FR/62/LRC-L47	France	Kuhls et al. 2005
H4 (1)	304	AM901452	*L. donovani*	MHOM/IQ/1981/SUKKAR2	Iraq	Alam et al. 2009a, 2009b
H5 (1)	299	AJ276260	*L. donovani*	–	Sudan	El Tai et al. 2001
H6 (21)	300	GQ367489	n.d.	MHOM/CN/80/801	Kashi, China	Wang et al. 2010
		AM901448	n.d.	MHOM/LK/2002/L60b	Sri Lanka	Alam et al. 2009a
		AM901447	n.d.	MHOM/LK/2002/L60c	Sri Lanka	Alam et al. 2009a
		EU753232	–	NICD/IN/30/A	India	Thakur (unpublished data)
		EU753231	–	NICD/IN/28/A	India	Thakur (unpublished data)
		EU753230	–	NICD/IN/24/A	India	Thakur (unpublished data)
		EU753229	–	NICD/IN/23/A	India	Thakur (unpublished data)
		EU753228	–	NICD/IN/19/A	India	Thakur (unpublished data)
		AJ634378	n.d.	MHOM/IN/01/BHU20140	India	Kuhls et al. 2005
		AJ634377	*L. donovani*	MHOM/IN/96/THAK35	India	Kuhls et al. 2005
		AJ634376	*L. donovani*	MHOM/IN/00/DEVI	India	Kuhls et al. 2005
		AJ634375	*L. donovani*	MHOM/IN/54/SC23	India	Kuhls et al. 2005
		AJ634374	*L. donovani*	MHOM/KE/83/NLB189	Kenya	Kuhls et al. 2005
		AJ000297	*L. donovani*	MHOM/KE/85/NLB323	Kenya	Kuhls et al. 2005
		AJ000296	n.d.	MHOM/KE/84/NLB218	Kenya	Kuhls et al. 2005
		AJ000292	*L. donovani*	MHOM/IN/80/DD8	India	Kuhls et al. 2005
		AJ000290	n.d.	MHOM/IN/71/LRC-L51a	India	Kuhls et al. 2005
H7 (5)	301	AJ634377	*L. donovani*	MHOM/IN/96/THAK35	India	Kuhls et al. 2005
		AJ634373	*L. donovani*	MHOM/ET/67/HU3	Ethiopia	Kuhls et al. 2005
		AJ634372	*L. donovani*	MHOM/SD/93/9S	Sudan	Kuhls et al. 2005
		AJ634371	*L. infantum*[d]	MHOM/SD/93/452BM	Sudan	Kuhls et al. 2005
		AJ634370	*L. infantum*	MHOM/SD/97/LEM3472	Sudan	Kuhls et al. 2005
H8 (1)	300	AJ249616	*L. donovani*	–	Sudan	El Tai et al. 2000
H9 (2)	301	AJ249615	*L. donovani*	–	Sudan	El Tai et al. 2000
		AJ276259	*L. donovani*	–	Sudan	El Tai et al. 2001
H10 (1)	303	AM901449	*L. donovani*	MHOM/IN/1983/CHANDIGARH	India	Alam et al. 2009a
H11 (1)	298	AJ249621	*L. donovani*	–	Sudan	El Tai et al. 2000
H12 (14)	303	AJ276258	*L. donovani*	–	Sudan	El Tai et al. 2001
		AM901453	*L. donovani*	MCAN/MA/2002/AD3	Morocco	Alam et al. 2009a
		AJ634369	*L. infantum*[d]	MHOM/SD/82/GILANI	Sudan	Kuhls et al. 2005
		AJ634368	*L. donovani*	MHOM/SD/93/338	Sudan	Kuhls et al. 2005
		AJ634367	*L. archibaldi*[b]	MHOM/ET/72/GEBRE1	Ethiopia	Kuhls et al. 2005
		AJ634366	*L. archibaldi*[b]	MHOM/SD/93/35-band	Sudan	Kuhls et al. 2005
		AJ634365	n.d.	MHOM/SD/62/LRC-L61	Sudan	Kuhls et al. 2005
		AJ634364	*L. infantum*[d]	MHOM/SD/93/597-2	Sudan	Kuhls et al. 2005
		AJ634363	*L. infantum*[d]	MHOM/SD/93/762L	Sudan	Kuhls et al. 2005
		AJ634362	*L. infantum*[d]	MHOM/SD/93/45-UMK	Sudan	Kuhls et al. 2005
		AJ634361	*L. infantum*	MHOM/SD/62/3S	Sudan	Kuhls et al. 2005
		AJ249612	*L. donovani*	–	Sudan	El Tai et al. 2000

Table 1 (continued)

Sequence type (sequence numbers)	Sequence length (bp)	GenBank accession numbers	MLEE-based species assignment	WHO code	Origin	Reference
		AJ000293	n.d.	MHOM/SD/68/1S	Sudan	Kuhls et al. 2005
		AJ000291	n.d.	MHOM/SD/75/LV139	Sudan	Kuhls et al. 2005
H13 (1)	301	AJ249614	*L. donovani*	–	Sudan	El Tai et al. 2000

[a] Named previously on the basis of zymodeme analysis (Guan et al. 1995)

[b] As synonym of *L. donovani* according to Lukeš et al. (2007)

[c] As synonym of *L. infantum* according to Lukeš et al. (2007)

[d] Identified as *L. infantum* according to the zymodeme, MON30, whereas recent analyses have shown that it is *L. donovani* (Jamjoom et al. 2004; Zemanova et al. 2004)

at 25°C. Approximately $1–5\times10^9$ promastigotes were collected at room temperature by centrifugation at 4,000 rpm for 10 min and washed with distilled water. Total genomic DNA was extracted from the promastigotes using a standard sodium dodecyl sulfate-proteinase K procedure, as described by Sambrook and Russell (2001). The primers of LITSR (5′-CTGGATCATTTTCCGATG-3′) and L5.8S (5′-TGATACCACTTATCGCACTT-3′; El Tai et al. 2000) were used to amplify ITS1–5.8S segments. The PCR protocols were 94°C for 3 min followed by 30 cycles of 94°C for 30 s, 49°C for 30 s, 72°C for 1 min, and then a final elongation step at 72°C for 10 min. The PCR products were purified on a 2.0% agarose gel stained with ethidium bromide, using a commercial DNA purification kit following the manufacturer's protocol. Sequencing was performed using the same PCR primers with ABI Big Dye Terminator chemistry on an ABI 3730 automated sequencer. The sequences have been deposited in GenBank under accession numbers HM130599–HM130608 (Table 1).

Sequence alignment and analyses

A set of ITS1–5.8S sequences of *Leishmania* were retrieved from GenBank, including nine species of subgenus *Leishmania* and two species of subgenus *Viannia* (see Tables 1 and 2). The sequences were first aligned using Clustal X 1.83 (Thompson et al. 1997) with a gap-opening penalty of 5 and gap-extension penalty of 1, following the recommendation (use of small gap costs) of Hickson et al. (2000). The aligned matrix from this procedure was checked by eye, and minor adjustments were made manually with SeaView v.4.2.5 (Gouy et al. 2010). The data matrices are available from the corresponding author.

Compositional heterogeneity was evaluated using Chi-square (χ^2) tests implemented in PAUP* 4.0b10 (Swofford 2002) and assessed using the software SeqVis v.1.3 (Ho et al. 2006) to visualize and to conduct matched-pairs tests of symmetry (Ababneh et al. 2006). Evidence of evolution under conditions more complex than that assumed by commonly applied models (i.e., stationary, reversible, and homogeneous conditions) was inferred if the scatter of dots in the tetrahedral plots was widely dispersed and if x% of the matched-pairs tests of symmetry produced p values greater than or equal to x; this procedure is consistent with that advocated by Jermiin et al. (2008). Substitution saturation was tested by inspecting a new entropy-based index as implemented in DAMBE (Xia and Xie 2001). For this approach, if I_{ss} (i.e., index of substitution saturation) is not smaller than $I_{ss.c}$ (i.e., critical I_{ss}), then we can conclude that the sequences have experienced severe substitution saturation (Xia et al. 2003; Xia and Lemey 2009). The K80 +G distance matrices (Kimura, 1980) were computed with MEGA v. 4.1 (Tamura et al. 2007), with the gamma shape of 0.5780.

Phylogenetic analyses

Phylogenetic hypotheses of *Leishmania* were generated with ITS1–5.8S rRNA segments using two types of commonly applied phylogenetic method: heuristic searches using equally weighted maximum parsimony (MP) analyses performed with the program PAUP* and Bayesian inference (BI) with the program MrBayes v.3.2 (Ronquist and Huelsenbeck 2003). In both MP and BI analyses, gaps were treated as missing data.

For heuristic searches under parsimony, invariant characters were removed from the dataset, and all remaining characters were treated as equally weighted. Each search involved ten random addition replicates, one tree held at each step, TBR branch swapping, steepest descent on, and a maximum of 10,000 saved trees; all other search settings were left at default values. Non-parametric bootstrapping was used to generate phylogeny confidence values (Felsensten 1985), with 1,000 pseudoreplicates using a heuristic tree search for each pseudoreplicate. *Leishmania panamensis* (FJ948422) was used to root the trees. Because intraspecific gene evolution cannot always be represented

Table 2 List of the other strains, origin, and database accession numbers retrieved from GenBank

Species	GenBank accession number	WHO code	Origin	Sequence length (bp)	Reference
L. mexicana	AF466383	MNYC/BZ/62/M379	Brazil	320	Berzunza-Cruz et al. 2002
	FJ948436	–	–	320	de Almeida et al. (unpublished data)
	FJ948433	–	–	321	de Almeida et al. (unpublished data)
	FJ948435	–	–	318	de Almeida et al. (unpublished data)
L. amazonensis	AF339753	–	–	314	Berzunza-Cruz et al. 2002
	DQ182536	MHOM/BR/73/M2269	Brazil	315	Rotureau et al. 2006
	FJ753371	–	–	316	de Almeida et al. (unpublished data)
L. tropica	FJ948459	–	India	305	de Almeida et al. (unpublished data)
	FJ948461	–	India	306	de Almeida et al. (unpublished data)
	FJ948460	–	India	305	de Almeida et al. (unpublished data)
	FJ948464	–	India	304	de Almeida et al. (unpublished data)
	FJ948465	–	India	301	de Almeida et al. (unpublished data)
	HM004586	–	Isfahan	309	Mahmoudzadeh-Niknam (unpublished data)
	AJ000302	IROS/NA/76/ROSSI-II	–	303	Schöenian et al. (unpublished data)
	AJ300485	MHOM/TN/88/TAT3	Tunisia	308	Schöenian et al. (unpublished data)
	AJ000301	MHOM/KE/84/NLB297	Kenya	306	Schöenian et al. (unpublished data)
	FJ948452	–	–	310	de Almeida et al. (unpublished data)
	FJ948456	–	–	310	de Almeida et al. (unpublished data)
	FJ948451	–	–	313	de Almeida et al. (unpublished data)
	GQ913688	MHOM/AF/88/KK27	Afghanistan	303	de Almeida et al. (unpublished data)
	FJ948458	–	India	302	de Almeida et al. (unpublished data)
	FJ460459	MHOM/EG/06/RTC-67	India	305	Shehata et al. 2009
	FJ948457	–	India	301	de Almeida et al. (unpublished data)
L. major	AY260965	MHOM/Ir/02/PIICC1	Iran	320	Tashakori et al. (unpublished data)
	AJ300482	MTAT/KE/??/NLB089A	Kenya	320	Schöenian et al. (unpublished data)
	DQ295824	IPAP/EG/89/SI-177	–	322	Fryauff et al. 2006
	FJ753395	–	–	319	Schöenian et al. (unpublished data)
	AJ272383	–	Turkmenistan	319	Chendrik et al. (unpublished data)
	AJ300481	MHOM/SD/90/SUDAN3	Sudan	320	Schöenian et al. (unpublished data)
	GQ471900	MRHO/IR/75/ER	Iran	322	Mahmoudzadeh-Niknam et al. (unpublished data)
L. turanica[a]	EF413079	–	Iran	320	Parvizi and Ready 2008
	AJ272378	–	Uzbekistan	320	Chendrik et al. (unpublished data)
	AJ272379	–	Turkmenistan	320	Chendrik et al. (unpublished data)
	AJ272380	–	Turkmenistan	320	Chendrik et al. (unpublished data)
	AJ272381	–	Turkmenistan	320	Chendrik et al. (unpublished data)
	AJ272382	–	Kazakhstan	320	Chendrik et al. (unpublished data)
L. gerbilli	AJ300486	–	Uzbekistan	319	Schöenian et al. (unpublished data)
L. aethiopica	GQ920677	–	–	301	de Almeida et al. (unpublished data)
	GQ920674	–	–	321	de Almeida et al. (unpublished data)
	GQ920673	–	–	322	de Almeida et al. (unpublished data)
	GQ920675	–	–	324	de Almeida et al. (unpublished data)
L. panamensis	FJ948442	–	–	286	de Almeida et al. (unpublished data)
L. braziliensis	DQ182537	MHOM/BR/84/LTB300	Brazil	279	Rotureau et al. 2006

[a] The six strains share a common allele of ITS1–5.8S with the isolate MRHO/CN/88/KXG-2, as shown in Table 1

by a bifurcating tree, haplotype networks may more effectively portray the relationships among haplotypes within species (reviewed by Posada and Crandal (2001)). Therefore, we constructed unrooted parsimony networks of haplotypes for *L. donovani* complex and *Leishmania* sp. (see below) using TCS v.1.21 (Clement et al. 2000), with gap treated as a fifth state.

Prior to Bayesian analyses, the best-fit model of evolution, K80 + G, was selected using jModeltest v. 0.1.1 (Posda, 2008) under the Bayesian information criterion (Schwarz 1978), following recent recommendations (Posada and Buckley 2004). We estimated posterior probability distributions by allowing four incrementally heated Markov chains (default heating values) to proceed for four million generations, with samples taken every 200 generations. Analyses were repeated beginning with different starting trees to ensure that our analyses were not restricted from the global optimum (Huelsenbeck et al. 2002). Convergence was first tested by examining the average deviation of the split frequencies of the two runs, in order to determine whether the two runs had converged. MCMC convergence was also explored by examining the potential scale reduction factor (PSRF) convergence diagnostics for all parameters in the model (provided by the *sump* and *sumt* commands) and graphically using the cumulative, compare, and absolute difference options of the program AWTY (Nylander et al. 2008). The first one million generations, before this chain reached apparent stationarity, were discarded, and the remaining samples from the independent runs were pooled to obtain the final approximation of the posterior distribution of trees. To yield a single hypothesis of phylogeny, the posterior distribution was summarized as a 50% majority-rule consensus.

In addition, as gap (or "indel") characters have been widely recognized as a valuable source of data for phylogenetic inference across the tree of life (e.g., Dessimoz and Gil 2010), phylogenetic information from indel events of ITS1 was also included in MP and BI by coding indel events into a separate data matrix with the program SeqState (Müller 2005) using the simple indel coding method (Simmons and Ochoterena 2000). In the latter, all indels are scored as binary characters regardless of their length. In BI, a discrete model employing identical rates of forward and backward transitions (Lewis 2001) was applied to the indel matrix.

Bayesian hypothesis testing

We used Bayes factors to compare our preferred Bayesian tree topology (see below) to Bayesian trees with constraint. This method differs from traditional hypothesis testing because it does not offer a criterion for absolute rejection of a null hypothesis but instead an evaluation of the evidence in favor of the null hypothesis (Kass and Raftery 1995). The phylogeny inferred from the ITS1–5.8S data set was constrained to alternative hypotheses. Constraint analyses were conducted in MrBayes v.3.2 using the command *prset topologypr= constraint*. All analyses consisted of two simultaneous runs each with an abbreviated three MCMC chains run for four million generations or more (as necessary). The Bayes factor was determined by calculating the marginal likelihood for both unconstrained and constraint analyses using Tracer v.1.5 (Rambaut and Drummond 2009). The difference in these ln-transformed marginal likelihoods was compared to the table provided by Jeffreys (1935, 1961) and further modified by Raftery (1996). Based on these tables, we consider a 2ln Bayes factor ≥10 as significant evidence for a hypothesis (Kass and Raftery 1995).

Results

Base composition and nucleotide substitution patterns

The newly determined ITS1–5.8S fragments ranged in size from 297 bp for isolate IPHL/CN/77/XJ771, 311 bp for isolate MHOM/CN/84/JS1, and 313 bp for the remaining isolates with exception of isolate MRHO/CN/88/KXG-2 (320 bp). Specifically, we found that the isolate MRHO/CN/88/KXG-2 shared the same ITS1 sequence with six strains of *Leishmania turanica* retrieved from GenBank, as listed in Table 2. The 5.8S rRNA segment was 69 bp in length. The alignment of the *Leishmania* taxa required accommodation of 85–112 gaps in the ITS1 region per sequence. Indels (insertion/deletion events) represented between 20.7% and 27.4% of the aligned sequence length. Most indels were 3–4 bp in length, and the maximum indel length was 19 bp. Of the 409 aligned characters, 111 were

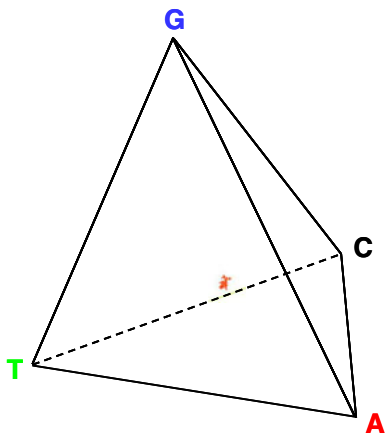

Fig. 1 Tetrahedral plots for ITS1–5.8S dataset, which were obtained using the Select Sites command from the View menu in the program SeqVis (Ho et al. 2006)

Table 3 Summary of results from matched-pairs tests of symmetry

Threshold (p value[a])	ITS1–5.8S	
	Number[b]	Proportion
0.05	0	0
0.01	0	0
0.005	0	0
0.001	0	0
0.0005	0	0
0.0001	0	0
0.00005	0	0

[a] The smallest p value is 0.1290

[b] The number of times that the matched-pairs test of symmetry resulted in a p value below the threshold (number of tests is 1,653)

variable, with 89 parsimony-informative. Percentage base compositions were A, 31.76; C, 20.01; G, 22.54; T, 25.69. The average maximum likelihood estimated Ti/Tv ratio was 1.45.

A base stationarity test showed insignificant differences among taxa in base composition bias in the data (χ^2=22.34, df=171, p=1.00). Figure 1 presented the tetrahedral plot from the ITS1–5.8S rRNA. Clearly, there was no conspicuous compositional heterogeneity in the alignment. The implication of this plot was that these sites were likely to have evolved under the same stationary, reversible, and homogeneous conditions. To corroborate whether this was the case, the matched-pairs test of symmetry was used in conjunction with the alignment. Table 3 summarized the distribution of p values. The distribution of p values clearly showed that the evolutionary process was likely to have been stationary, reversible, and homogeneous, implying that it would be wise to analyze this data using a phylogenetic approach that assumes a stationary, reversible, and homogeneous evolutionary process. The observed I_{ss} value of 0.764 was not significantly different from the $I_{ss.c}$ value of 0.692 for a symmetrical topology (p=0.4757, two-tailed test) and was significantly greater than the $I_{ss.c}$ value of 0.362 for an asymmetrical topology (p=0.0001, two-tailed test), suggesting that the ITS1–5.8S might have experienced substitution saturation.

K80 distances among the *Leishmania* species except *Leishmania* sp. ranged from near zero (between *Leishmania braziliensis* and *L. panamensis*) to 0.224 (between *Leishmania mexicana* and *L. panamensis*). Most pairwise comparisons mentioned above had divergence values of less than 0.224, with 0.102 on average. Meanwhile, the divergence between *Leishmania* sp. and other species ranged from 0.104 (*Leishmania* sp. versus *Leishmania aethiopica*) to 0.231 (*Leishmania* sp. vs. *L. panamensis*), with an average of 0.147 (Table 4).

Phylogenetic relationships

The heuristic search of the ITS1–5.8S matrix resulted in 10,000 equally parsimonious trees of 149 steps, with high values of CI (0.8456) and RI (0.9599). In the strict consensus phylogram (Fig. 2), eight isolates in China formed a strongly supported clade (clade A; *Leishmania* sp.; BP=100%) that was sister to the remaining members of subgenus *Leishmania* (BP=100%). *Leishmania amazonensis* and *L. mexicana* formed a robust clade (BP=100%) that was basal to all remaining subgenus *Leishmania* species (BP=100%). Within the other members of subgenus *Leishmania*, *L. donovani* complex clustered with *Leishmania tropica* (BP=66%), next joined by *L. turanica* plus *L.*

Table 4 Pairwise genetic distances for ITS1–5.8S segments among *Leishmania* species in this study

		1	2	3	4	5	6	7	8	9	10	11
1	*L. donovani* complex	–										
2	*L. tropica*	0.039	–									
3	*L. turinica*	0.043	0.038	–								
4	*L. gerbilli*	0.065	0.060	0.023	–							
5	*L. major*	0.054	0.047	0.051	0.063	–						
6	*L. aethiopica*	0.041	0.023	0.052	0.074	0.055	–					
7	*L. amazonensis*	0.047	0.074	0.087	0.099	0.085	0.080	–				
8	*L. mexicana*	0.053	0.070	0.083	0.094	0.081	0.078	0.008	–			
9	*L. braziliensis*	0.158	0.179	0.161	0.168	0.175	0.178	0.211	0.212	–		
10	*L. panamensis*	0.163	0.185	0.166	0.173	0.181	0.184	0.223	0.224	0	–	
11	*Leishmania* sp.	0.114	0.108	0.131	0.152	0.143	0.104	0.131	0.129	0.225	0.231	–

The substitution model, K80 + G, with gamma shape of 0.5780, was selected using jModeltest v. 0.1.1 (Posda, 2008) under the Bayesian information criterion (BIC; Schwarz 1978)

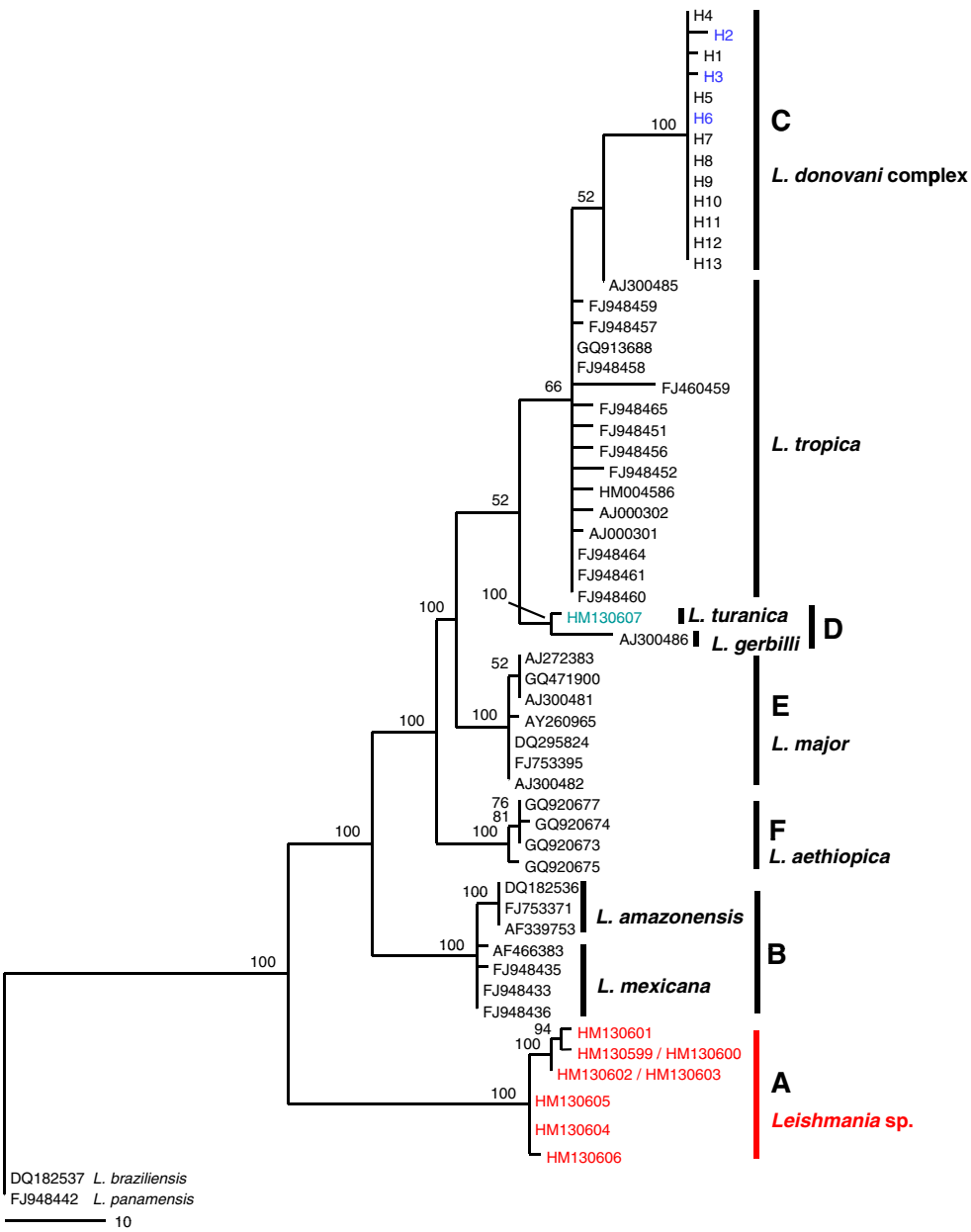

Fig. 2 Maximum parsimony consensus tree from 1,000 bootstrap replicates of ITS1–5.8S dataset by using PAUP*. *Numbers above the branch* represent percent recovery in bootstrap analysis (1,000 pseudoreplicates). `Tree length=149, CI=0.8456, RI=0.9599

gerbilli (BP=52%), *L. major* (BP=100%), and finally by *L. aethiopica* (BP=100%). Nevertheless, the monophyly of *L. tropica* was not supported. Unexpectedly, the isolate MRHO/CN/88/KXG-2, identified as *L. turanica* by MLEE, did not cluster where expected, which appeared in the *Leishmania* sp. branch. When indels of the ITS1 were treated as additional characters, the heuristic search yielded 10,000 equally parsimonious trees of 298 steps, with high values of CI (0.7651) and RI (0.9435). As shown in Fig. 3, the consensus tree was similar to Fig. 2 with respect to the placements of clades A and B. The placements of other species were incongruent with those in Fig. 2.

For the BI analyses, the likelihood value of the 50% majority consensus tree (Fig. 4) was ln L=−1,541.09. The average PSRF was 1.001. Overall, as well with maximum parsimony analyses, *Leishmania* sp., consisting of eight isolates from China, was sister to the subgenus species (PP=0.76). Similarly, *L. amazonensis* and *L. mexicana* formed a robust clade (PP=1.00) that was sister to all remaining subgenus *Leishmania* species (PP=0.96). The relationships within the remaining species were similar to those of Fig. 3 except for recognizing the monophyly of *L. tropica* (PP=0.56) instead of *L. donovani* complex. When the ITS1 indels were incorporated as additional characters, the resultant 50% majority consensus tree was shown as Fig. 5, with ln L of −2,167.78 and the average PSRF of 1.001. In this context, the topology is similar to Fig. 4 except that the monophyly of *L. donovani* complex was

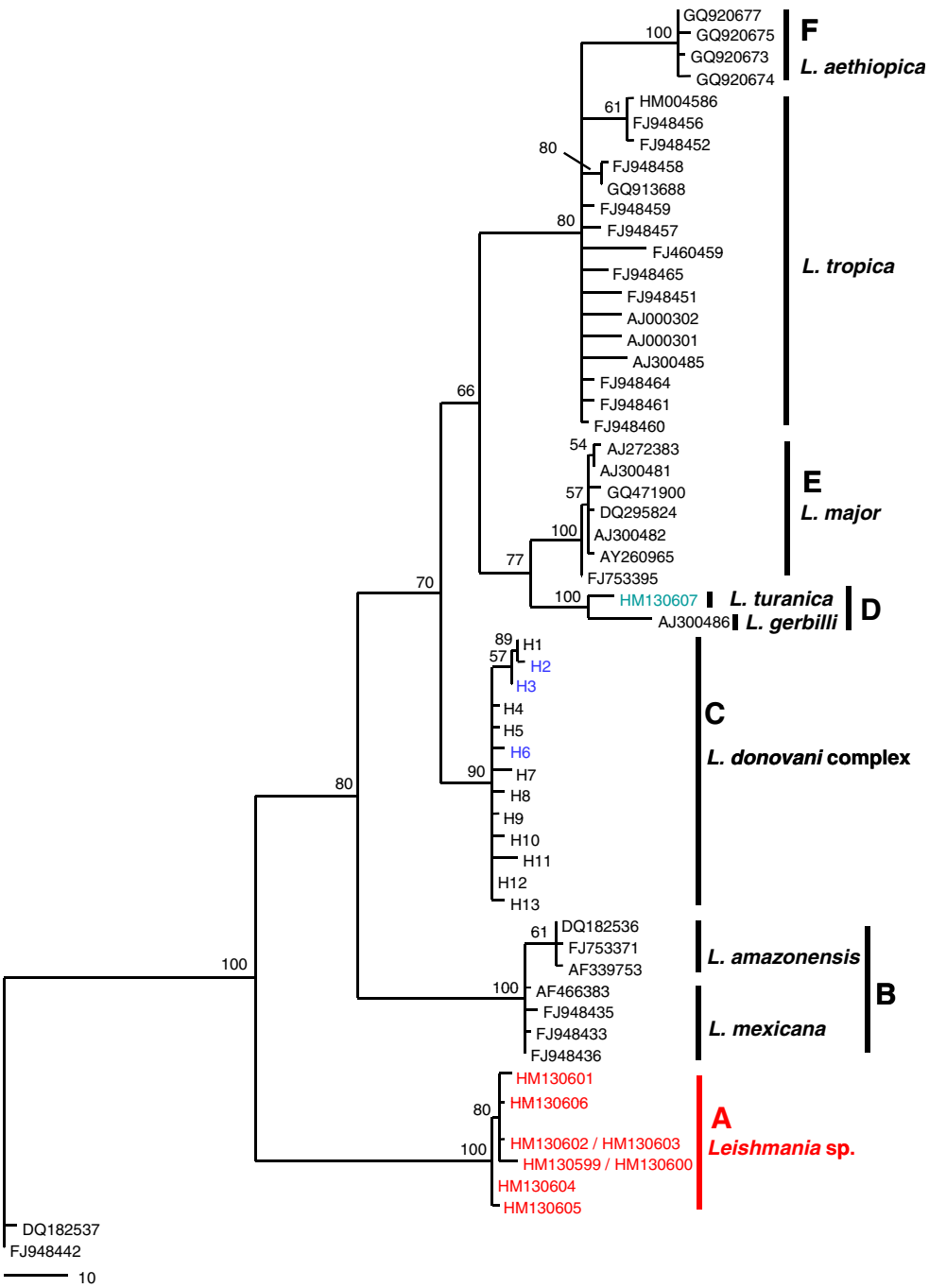

Fig. 3 Maximum parsimony consensus tree from 1,000 bootstrap replicates of ITS1–5.8S with indel coding by using PAUP*. *Numbers above the branch* represent percent recovery in bootstrap analysis (1,000 pseudoreplicates). Tree length=298, CI=0.7651, RI=0.9435

recovered with moderate posterior probability (PP=0.88) instead of *L. tropica*.

To get additional insight into the relationships among the *L. donovani* complex strains, we analyzed our data set, using the coalescent-based statistical parsimony network approach. The network of 13 haplotypes was shown as Fig. 6. H9 and H12 seemed to be central haplotypes, and the haplotype diversity was highest in Sudan. *L. donovani* revealed much more polymorphism than *L. infantum* despite a wider geographical distribution for the latter (see H3 in Fig. 6). *L. infantum* (H3) was most closely related to the H1 of *L. donovani*, with one mutational step. H2, shared by five strains from China, was also most closely related to H1. Having an advantage over the bifurcating tree in detail at the intraspecific level, the haplotype network could intuitively reflect the genetically greater distances between the singleton (H11) and one central haplotype (H12; five mutational steps, see Fig. 6). As shown in Fig. 7, the haplotype shared by GSH2 and GS2 (i.e., HM130602 and HM130603) was the interior

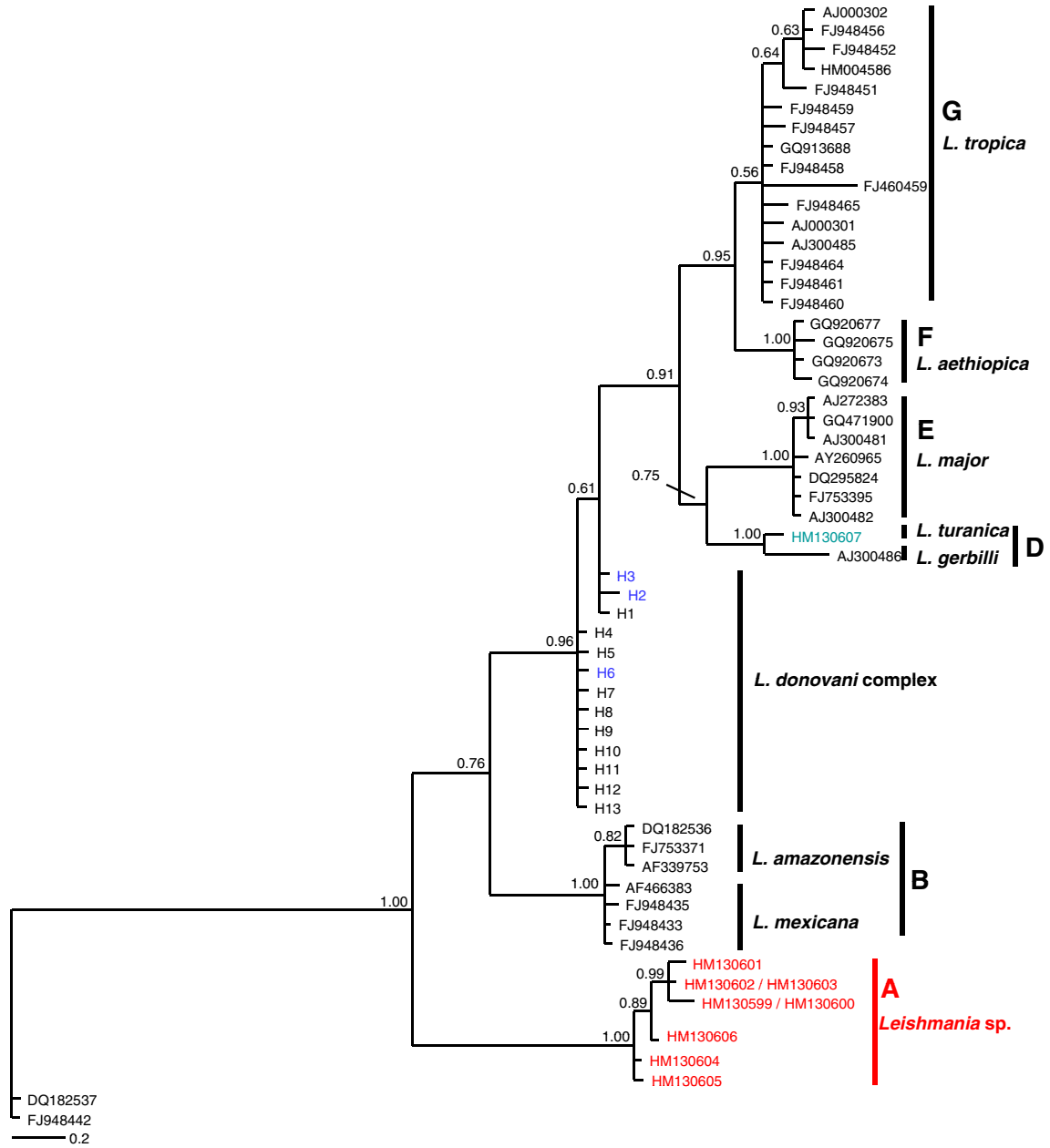

Fig. 4 The 50% majority-rule consensus tree inferred from Bayesian inference of ITS1–5.8S dataset by using MrBayes v. 3.2. *Numbers at nodes* represent Bayesian posterior probabilities

haplotype of *Leishmania* sp. and may be older than any other haplotypes. There were three tip haplotypes harbored by GS6/GS1, SC10H2, and JS1, respectively. Similarly, this network reflected a greater distance between JS1 and the interior haplotype as four mutational steps.

Bayesian hypothesis testing

Bayes factor comparisons were summarized in Table 5. The analyses conducted reflect our primary interests of evaluating the inclusion of *Leishmania* sp. in *L. donovani* complex. As mentioned above, analyses of the ITS1 data resulted in a Chinese *Leishmania* clade that excluded *L. turanica* and *L. donovani* complex. Bayes factor analyses of the ITS1–5.8S incorporating indels coding were conducted to compare topologies with constraints to the optimal tree topology. In all cases, there was very strong (2ln Bayes factor >10) evidence against the constrained topologies.

Discussion

Probabilistic methods, namely maximum likelihood and BI, have progressively supplanted the MP method for inferring

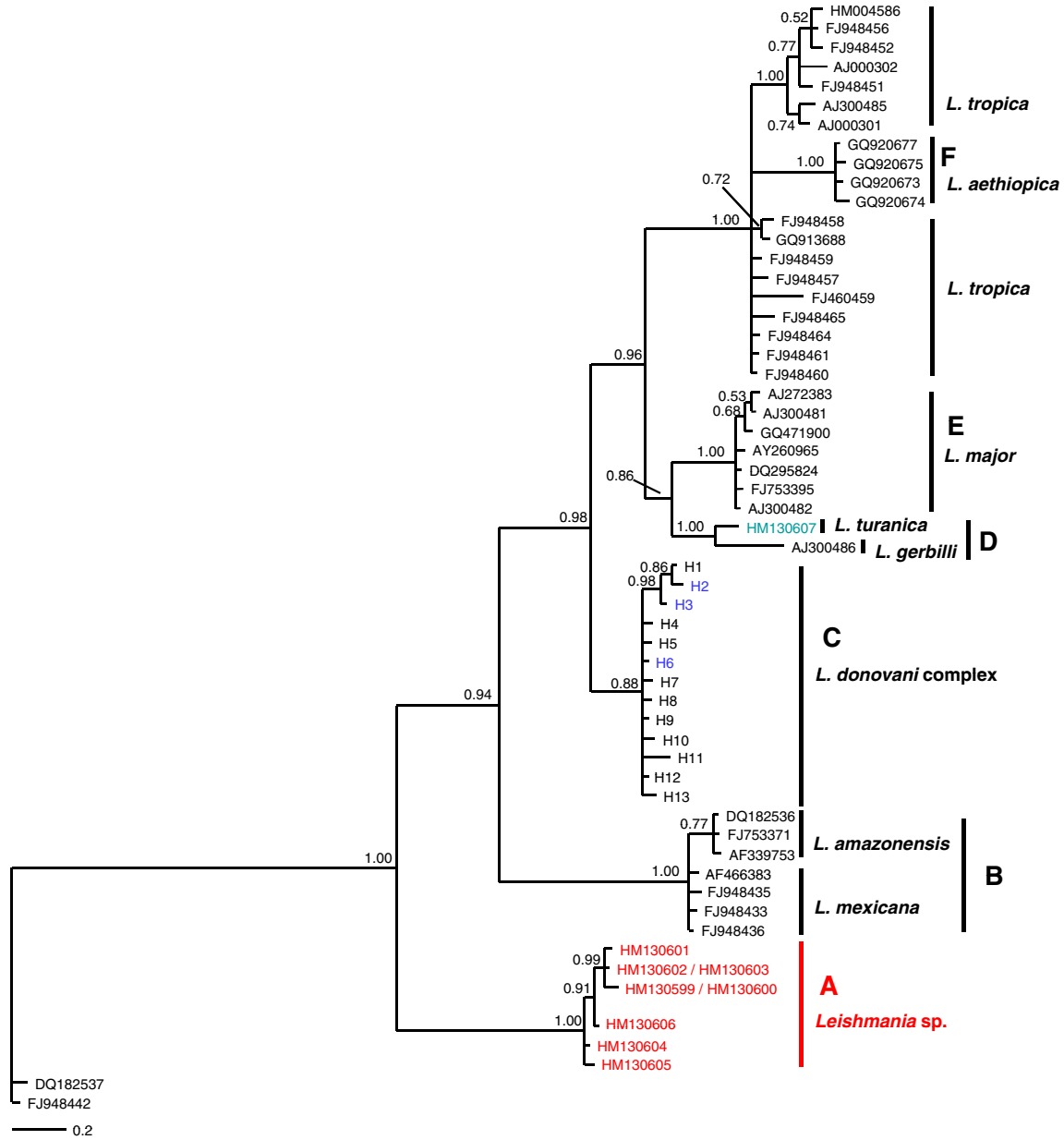

Fig. 5 The 50% majority-rule consensus tree inferred from Bayesian inference of ITS1–5.8S plus indel coding by using MrBayes v. 3.2, with indels treated in a manner similar to the simple gap coding outlined by Simmons and Ochoterena (2000). *Numbers at nodes* represent Bayesian posterior probabilities

phylogenetic trees. One of the major reasons for this shift is that MP is much more sensitive to the Long Branch Attraction artifact than are probabilistic methods. Based on simulation studies, Philippe et al. (2005) found that MP can be affected by heterotachy and that it is much less efficient than probabilistic methods in dealing with all other evolutionary heterogeneities. Thus, in support of several recent studies (e.g., Gadagkar and Kumar 2005; Gaucher and Miyamoto 2005; Spencer et al. 2005), Philippe et al. (2005) strongly urged the continued preference of probabilistic methods for inferring phylogenies from real sequences. In our study, there is no significant compositional heterogeneity in ITS1 sequences, which are likely to have evolved under the same stationary, reversible, and homogeneous conditions. The MP analysis incorporating indels coding resulted in a topology that is congruent with the Bayesian trees. However, Bayesian support values in the BI tree (Fig. 5) were found to be

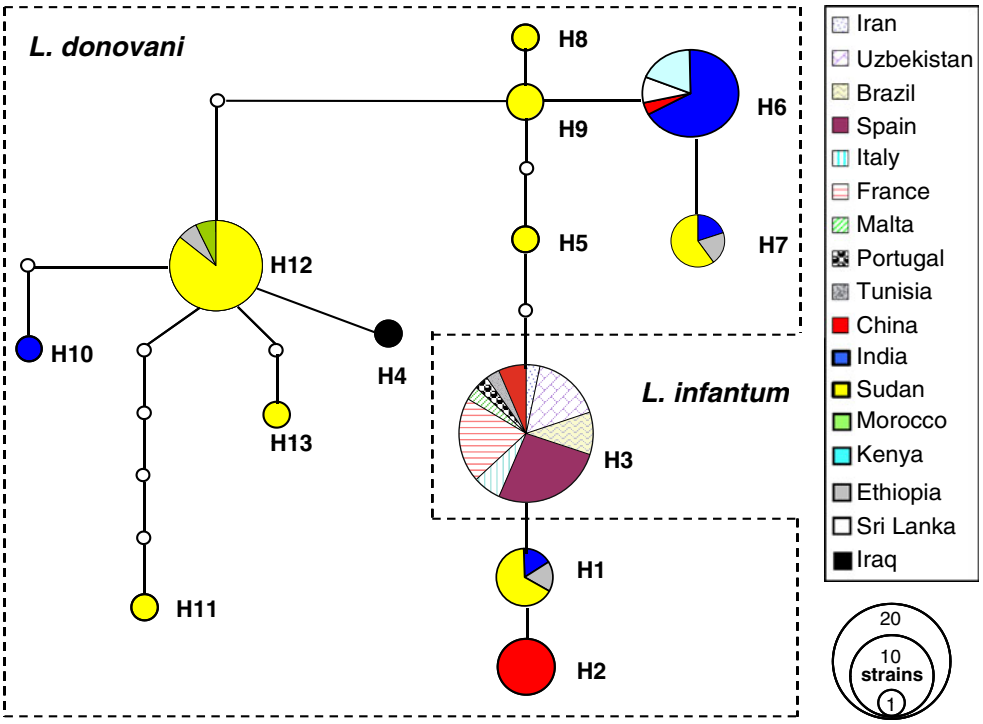

Fig. 6 Statistical parsimony network showing genetic relationships and distance among 13 haplotypes of *L. donovani* complex from different countries. Numbers of haplotypes correspond to Table 1. In the network, *solid circles* indicate sampled haplotypes; *small hollow circles* indicate unsampled or extinct haplotypes. Each mutation step is shown as either a *short or longer line* connecting neighboring haplotypes (including observed and unobserved one). The *size of the solid circles* roughly represents the numbers of strains carrying the haplotype, with the scale given beside the network; *different filled patterns* represent the corresponding geographical origin from which the haplotype was sampled

comparatively higher than bootstrap values for the clades in that MP tree (Fig. 3), suggesting that the Bayesian inference could be properly applied to the phylogenetic analyses of subgenus *Leishmania*. Considering congruent with the recent studies on the interrelationships of subgenus *Leishmania* (Asato et al. 2009; Fraga et al. 2010), we tentatively support the relationships inferred from BI of the dataset incorporating indels coding (Fig. 5) as the preferred phylogeny.

As expected, the isolate MRHO/CN/88/KXG-2, previously identified as *L. turanica* by MLEE (Guan et al. 1995), clusters with *L. gerbilli*. This result is in congruent with the taxonomic scheme published by the World Health Organization (WHO 1990). Interestingly, a common allele was shared by the isolate MRHO/CN/88/KXG-2 with six other strains of *L. turanica* from Central Asia, as shown in Table 2. This further lends support that the MRHO/CN/88/KXG-2 belongs to *L. turanica*. On the other hand, as highlighted by Guan et al. (1992b), *Leishmania* parasite of the Karamay great gerbils (including isolate MRHO/CN/88/KXG-2) was pathogenic to monkey and man, causing cutaneous leishmaniasis. This medical characteristics, however, is different from that of *L. turanica*, being nonpathogenic to humans, as described by Strelkova et al. (1990). Thus, more isolates of *L. turanica* from different geographical areas and multiple loci are required for phylogeographic studies in order to clarify the intra-species genetic diversity and complex phylogeographic pattern.

The species concept has long been a matter of debate which is far more resolved (De Meeûs et al. 2003; de Queiroz 2007). As noted by Bañuls et al. (2002), any new species of *Leishmania* should be based on the clearly distinct phylogenetic approach of Tibayrenc's discrete typing unit (Tibayrenc 1998). Eight new sequences reported in this work formed a robust clade (clade A;

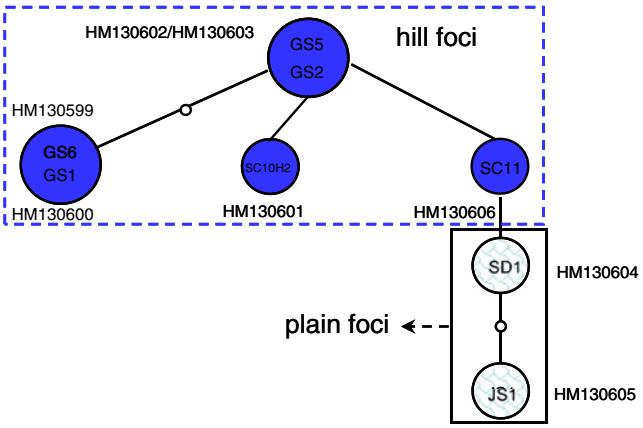

Fig. 7 Statistical parsimony network showing genetic relationships and distance among six haplotypes of *Leishmania* sp. from different sites in China. In the network, *solid circles* indicate sampled haplotypes; *small hollow circles* indicate unsampled or extinct haplotypes. Each mutation step is shown as either a *short or longer line* connecting neighboring haplotypes (including observed and unobserved one)

Leishmania sp.) that is sister to the remaining members of subgenus *Leishmania*. When we constrained all the isolates in China to form a monophyletic group, the tree obtained differs significantly from the BI tree based on 2ln Bayes factor comparison (55.856>10; see Table 5). Thus, we reject the monophyletic origin of Chinese *Leishmania* isolates and exclude the possibility that all the isolates in this study only belong to *L. donovani* complex. In addition, the mean genetic divergence (K80 distance) between *Leishmania* sp. and other species is 0.147, which is higher than that among several other species (mean 0.102). We further confirm that *Leishmania* sp. is an undescribed pathogenic species endemic in China, comprising isolates from hill foci, desert foci, and plain foci (see Fig. 7).

Judging from the kinetoplast and nuclear DNA heterogeneity, the isolate IPHL/CN/77/XJ771 was tentatively designated as *L. infantum* sensu lato by Lu et al. (1994). This hypothesis, however, is challenged by our molecular data. There is complete identity of the ITS1 sequence of the isolate IPHL/CN/77/XJ771 and that of a strain of *L. donovani* (MHOM/CN/00/Wangjie1). They share a common haplotype H2 with three isolates from eastern Jiashi County. As shown in the haplotype network (Fig. 6), H2 is most closely related to H1, which is shared by four strains from Sudan, one from India, and one from Ethiopia. Consequently, we argue that the isolate IPHL/CN/77/XJ771 is *L. donovani* instead of *L. infantum*.

The three isolates from eastern Jiashi County, causing desert type of zoonotic visceral leishmaniasis, were designated as *L. infantum* based on the genetic analysis of the ITS1 sequence (Wang et al. 2010). This conclusion, however, should be interpreted with caution, since it was not deduced from a robust phylogenetic tree, and the MP tree was misleading (see Fig. 3 in Wang et al. 2010). In contrast, as mentioned above, the coalescent-based statistical parsimony network approach provides additional insight into the relationships among the *L. donovani* complex. We infer that the three isolates should belong to *L. donovani*. As can be seen from Fig. 6, the isolate MHOM/CN/80/801 from Kashi city shares H6 with several isolates from India, Kenya, and Sri Lanka, and there is only one mutational step between H6 and a central haplotype H9 from Sudan. In combination with the results of Lukeš et al. (2007), the demonstration that the isolate MHOM/CN/80/801 belongs to *L. donovani* is corroborated, in accord with Wang et al. (2010).

In conclusion, phylogenetic analyses suggested that Chinese *Leishmania* isolates do not form a monophyletic group, but among which eight newly determined isolates form a monophyletic group, being sister to other members of subgenus *Leishmania*. The genetic distance analysis further provides evidence of the occurrence of an undescribed species of *Leishmania*. Our results also suggest that the isolate IPHL/CN/77/XJ771 is *L. donovani*; the three isolates from eastern Jiashi County, Xinjiang Uygur Autonomous Region belong to *L. donovani* instead of *Leishmania infantum*. In addition, the results of this study make an important contribution to understanding the heterogeneity and relationships of Chinese *Leishmania* isolates, further indicating that the isolates from China may have had a more complex evolutionary history than expected. However, more samples from different geographical areas and multiple independent evolving loci are required for phylogenetic studies in order to clarify the evolutionary history among Chinese *Leishmania* isolates. It further might be useful in understanding the links between clinical disease, geographic origin, and nomenclature of *Leishmania* species.

Table 5 Summary of 2ln Bayes factor comparisons of alternative phylogenetic hypotheses

Constraint (H_0)	ln marginal likelihood		Evidence against H_0	
Alternative phylogenetic hypotheses	ln *L*: unconstrained (Fig. 5)	ln *L*: constrained	2ln Bayes factor ($2\ln(B_{10})$)	
Monophyly constraint of *L. donovani* complex[a]	−2187.027	−2268.638	163.222	Very strong
Monophyly constraint of isolates from China with exception of MRHO/CN/88/KXG-2[b]	−2187.027	−2214.955	55.856	Very strong

Marginal likelihoods were calculated using the method of Suchard et al. (2001) using Tracer 1.5 (Rambaut and Drummond 2009)

2ln Bayes factors ≥10 are considered very strongly different (Kass and Raftery 1995), indicating evidence against alternative hypotheses

[a] Constrained tree with *Leishmania* sp. except MRHO/CN/88/KXG-2 embedded within *L. donovani* complex (H1–H13)

[b] Constrained tree with *Leishmania* sp. clustering with H2, H3, and H6

Acknowledgments This work was supported by the National Natural Science Foundations of China (30771883, 30800094) and the National Project of Important Infectious Diseases (2008-ZX10004-011). X-G Guo was supported by the National Natural Science Foundation of China (30700062). We thank Dianmei Lu and Zhibiao Xu for kinkly help with collecting some important references.

References

Ababneh F, Jermiin LS, Ma C, Robinson J (2006) Matched-pairs tests of homogeneity with applications to homologous nucleotide sequences. Bioinformatics 22:1225–1231

Alam MZ, Kuhls K, Schweynoch C, Sundar S, Rijal S, Shamsuzzaman AK, Raju BV, Salotra P, Dujardin JC, Schönian G (2009a) Multilocus microsatellite typing (MLMT) reveals genetic homogeneity of *Leishmania donovani* strains in the Indian subcontinent. Infect Genet Evol 9:24–31

Alam MZ, Kovalenko DA, Kuhls K, Nasyrova RM, Ponomareva VI, Fatullaeva AA, Razakov SA, Schnur LF, Schönian G (2009b) Identification of the agent causing visceral leishmaniasis in Uzbeki and Tajiki foci by analysing parasite DNA extracted from patients' Giemsa-stained tissue preparations. Parasitology 136:981–986

Al-Kandari WY, Al-Bustan SA (2010) Molecular identification of *Probolocoryphe uca* (Sarkisian, 1957; Digenea: Microphallidae) from Kuwait Bay using ITS1 and ITS2 sequences. Parasitol Res 106:1189–1195

Asato Y, Oshiro M, Myint CK, Yamamoto Y, Kato H, Marco JD, Mimori T, Gomez EA, HashiguchiY UH (2009) Phylogenic analysis of the genus *Leishmania* by cytochrome *b* gene sequencing. Exp Parasitol 121:352–361

Bañuls A-L, Hide M, Tibayrenc M (2002) Evolutionary genetics and molecular diagnosis of *Leishmania* species. Trans R Soc Trop Med Hyg 96(Suppl 1):S9–S13

Bañuls A-L, Hide M, Prugnolle F (2007) *Leishmania* and the leishmaniases: a parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. Adv Parasitol 64:1–109

Berzunza-Cruz M, Cabrera N, Crippa-Rossi M, Cabrera TS, Pérez-Montfort R, Becker I (2002) Polymorphism analysis of the internal transcribed spacer and small subunit of ribosomal RNA genes of *Leishmania mexicana*. Parasitol Res 88:918–925

Chen D-L, Wang G-T, Yao W-J, Nie P (2007) Utility of ITS1–5.8S–ITS2 sequences for species discrimination and phylogenetic inference of two closely related bucephalid digeneans (Digenea: Bucephalidae): *Dollfustrema vaneyi* and *Dollfustrema hefeiensis*. Parasitol Res 101:791–800

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1660

Dávila AMR, Memen H (2000) Internal-transcribed-spacer (ITS) sequences used to explore phylogenetic relationships within *Leishmania*. Ann Trop Med Parasitol 94:651–654

De Meeûs T, Durand P, Renaud F (2003) Species concepts: what for? Trend Parasitology 19:425–427

De Queiroz K (2007) Species concepts and species delimitation. Syst Biol 56:879–886

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

Dessimoz C, Gil M (2010) Phylogenetic assessment of alignments reveals neglected tree signal in gaps. Genome Biol 11: R37

El Tai NO, Osman OF, El Fari M, Presber W, Schönian G (2000) Genetic heterogeneity of ribosomal internal transcribed spacer (ITS) in clinical samples of *Leishmania donovani* spotted on filter papers as revealed by single-strand conformation polymorphisms (SSCP) and sequencing. Trans R Soc Trop Med Hyg 94:575–579

El Tai NO, El Fari M, Mauricio I, Miles MA, Oskam L, El Safi SH, Presber WH, Schönian G (2001) *Leishmania donovani*: intraspecific polymorphisms of Sudanese isolates revealed by PCR-based analyses and DNA sequencing. Exp Parasitol 97:35–44

Felsensten JP (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Fraga J, Montalvo AM, de Doncker S, Dujardin J-C, der Auwera GV (2010) Phylogeny of *Leishmania* species based on the heat-shock protein 70 gene. Infect Genet Evol 10:238–245

Fryauff DJ, Hanafi HA, Klena JD, Hoel DF, Appawu M, Rogers W, Puplampu N, Odoom S, Kweku M, Koram K, Wilson MD, Raczniak G, Boakye D (2006) ITS-1 DNA sequence confirmation of *Leishmania major* as a cause of cutaneous leishmaniasis from an outbreak focus in the Ho district southeastern Ghana. Am J Trop Med Hyg 75:502–504

Gadagkar SR, Kumar S (2005) Maximum likelihood outperforms maximum parsimony even when evolutionary rates are heterotachous. Mol Biol Evol 22:2139–2141

Gaucher EA, Miyamoto MM (2005) A call for likelihood phylogenetics even when the process of sequence evolution is heterogeneous. Mol Phylogenet Evol 35:624–636

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27:221–224

Guan L-R, Yang Y-Q, Ren H-Y, Chai J-J (1992a) Eight cutaneous *Leishmania*sis cases detected in Karamay, Xinjiang Uygur Autonomous Region, China. Southeast Asian J Trop Med Public Health 23:803–804

Guan L-R, Yang Y-Q, Xu Y-X, Wu J-T (1992b) Leishmaniasis in Karamay XI. The development of cutaneous leishmaniasis in monkey and man experimentally infected with Leishmania from Karamay big gerbil. Endemic Diseases Bulletin 10:263–266 (in Chinese with English abstract)

Guan L-R, Yang Y-Q, Ren H-Y, Shen W-X (1995) Discovery and study of *Leishmania turanica* for the first time in China. Bull World Health Organ 69:595–601

Hickson RE, Simon C, Perrey SW (2000) The performance of several multiple-sequence alignment programs in relation to secondary-structure features for an rRNA sequence. Mol Biol Evol 17:530–539

Ho JWK, Adams CE, Lew JB, Maatthews TJ, Ng CC, Shahabi-Sirjan A, Tan LH, Zhao Y, Easteal S, Wilson SR, Jermin LS (2006) SeqVis: visualization of compositional heterogeneity in large alignments of nucleotides. Bioinformatics 22:2162–2163

Hu X-S, Bu L, Ma Y, Wang Y, Jing B, Yi T (2002) Difference in DNA sequences in SSU rDNA variable regions among pathogens isolated from different epidemic foci of visceral leishmaniasis in China. Chin Med J (Engl) 115:1457–1459

Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. Syst Biol 51:673–688

Ibrahim ME, Barker DC (2001) The origin and evolution of the *Leishmania donovani* complex as inferred from a mitochondrial cytochrome oxidase II gene sequence. Infect Genet Evol 1:61–68

Jamjoom MB, Ashford RW, Bates PA, Chance ML, Kemp SJ, Watts PC, Noyes HA (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

Jeffreys H (1935) Some tests of significance, treated by the theory of probability. Proc Cambridge Philos Soc 31:201–222

Jeffreys H (1961) Theory of probability. Oxford University Press, London

Jermiin LS, Jayaswal V, Ababneh F, Robinson J (2008) Phylogenetic model evaluation. In: Keith J (ed) Methods in molecular biology: bioinformatics. Humana Press, Totowa

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90:773–795

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kuhls K, Mauricio IL, Pratlong F, Presber W, Schönian G (2005) Analysis of ribosomal DNA internal transcribed spacer sequences of the *Leishmania donovani* complex. Microbes Infect 7:1224–1234

Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. Syst Biol 50:913–925

Lin R-Q, Dong S-J, Nie K, Wang C-R, Song H-Q, Li A-X, Huang W-Y, Zhu X-Q (2007) Sequence analysis of the first internal transcribed spacer of rDNA supports the existence of the intermediate *Fasciola* between *F. hepatica* and *F. gigantica* in mainland China. Parasitol Res 101:813–817

Lu H-G, Zhong L, Guan L-R, Qu J-Q, Hu X-S, Chai J-C, Xu Z-B, Wang C-T, Chang K-P (1994) Separation of Chinese *Leishmania* isolates into five genotypes by kinetoplast and chromosomal DNA heterogeneity. Am J Trop Med Hyg 50:763–770

Lu F-L, Hu X-S, Jing B-Q, Luo P, Lin F-Q (1997) Analysis of kDNA of *Leishmania* isolates from hill and plain foci of China. Chin J Parasitol Parasit Dis 5:101–103 (in Chinese with English abstract)

Lu F-L, Hu X-S, Jing B-Q, Ma Y (1998) Analysis of nuclear DNA gene types of *Leishmania* isolates from hilly and plain foci of China. Chin J Parasitol Parasit Dis 16:432–435 (in Chinese with English abstract)

Lu D-M, Hu X-S, Qiao Z-D (2001) Analysis of *Leishmania* species and strains from China by RAPD technique. Chin J Parasitol Parasit Dis 19:290–293 (in Chinese with English abstract)

Lukeš J, Mauricio IL, Schönian G, Dujardin J-C, Soteriadou K, Dedet J-P, Kuhls K, Tintaya KWQ, Jirků M, Chocholová E, Haralambous C, Pratlong F, Oborník M, Horák A, Ayala FJ, Miles MA (2007) Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. Proc Natl Acad Sci USA 104:9375–9380

Montalvo AM, Fraga J, Monzote L, Montano I, de Doncker S, Dujardin JC, der Auwera GV (2010) Heat-shock protein 70 PCR-RFLP: a universal simple tool for *Leishmania* species discrimination in the New and Old World. Parasitology. doi:10.1017/S0031182010000089

Müller K (2005) SeqState: primer design and sequence statistics for phylogenetic DNA data sets. Appl Bioinform 4:65–69

Noyes HA, Arana BA, Chance ML, Maingon R (1997) The *Leishmania hertigi* (Kinetoplastida; Trypanosomatidae) complex and the lizard *Leishmania*: their classification and evidence for a neotropical origin of the *Leishmania-Endotrypanum* clade. J Eukaryot Microbiol 44:511–517

Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24:581–583

Parvizi P, Ready PD (2008) Nested PCRs and sequencing of nuclear ITS-rDNA fragments detect three Leishmania species of gerbils in sandflies from Iranian foci of zoonotic cutaneous leishmaniasis. Trop Med Int Health 13:1159–1171

Parvizi P, Moradi G, Akbari G, Farahmand M, Ready PD, Piazak N, Assmar M, Amirkhani A (2008) PCR detection and sequencing of parasite ITS-rDNA gene from reservoirs host of zoonotic cutaneous leishmaniasis in central Iran. Parasitol Res 103:1273–1278

Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M, Kerhornou A, Ivens A, Fraser A, Rajandream MA, Carver T, Norbertczak H, Chillingworth T, Hance Z, Jagels K, Moule S, Ormond D, Rutter S, Squares R, Whitehead S, Rabbinowitsch E, Arrowsmith C, White B, Thurston S, Bringaud F, Baldauf SL, Faulconbridge A, Jeffares D, Depledge DP, Oyola SO, Hilley JD, Brito LO, Tosi LR, Barrell B, Cruz AK, Mottram JC, Smith DF, Berriman M (2007) Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. Nat Genet 39:839–847

Philippe H, Zhou Y, Brinkmann H, Rodrigue N, Delsuc F (2005) Heterotachy and long-branch attraction in phylogenetics. BMC Evol Biol 5:50

Piarroux R, Fontes M, Perasso R, Gambarelli F, Joblet C, Dumon H, Auilici M (1995) Phylogenetic relationships between Old World *Leishmania* strains revealed by analysis of a repetitive DNA sequence. Mol Biochem Parasitol 73:249–252

Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of akaike information criterion and Bayesian approach. Syst Biol 53:793–808

Posada D, Crandal KA (2001) Intraspecific gene genealogies: trees grafting into networks. Trends Ecol Evol 16:37–45

Posda D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256

Raftery AE (1996) Hypothesis testing and model selection. In: Gilks WR, Spiegelhalter DJ, Richardson S (eds) Markov chain Monte Carlo in practice. Chapman and Hall, London, pp 163–188

Rambaut A, Drummond AJ (2009) Tracer v1.5, available from http://beast.bio.ed.ac.uk/Tracer

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

Rotureau B, Ravel C, Couppie P, Pratlong F, Nacher M, Dedet JP, Carme B (2006) Use of PCR-restriction fragment length polymorphism analysis to identify the main New World *Leishmania* species and analyze their taxonomic properties and polymorphism by application of the assay to clinical samples. J Clin Microbiol 44:459–467

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, New York

Schönian G, Akuffo H, Lewin S, Maasho K, Nylén S, Pratlong F, Eisenberger CL, Schnur LF, Presber W (2000) Genetic variability within the species *Leishmania aethiopica* does not correlate with clinical variations of cutaneous Leishmaniasis. Mol Biochem Parasitol 106:239–248

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464

Shehata MG, Samy AM, Doha S, Fahmy AR, Kaldas RM, Furman BD, Villinski JT (2009) First report of *Leishmania tropica* from a classical focus of *L. major* in North-Sinai Egypt. Am J Trop Med Hyg 81:213–218

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analysis. Syst Biol 49:369–381

Spencer M, Susko E, Roger AJ (2005) Likelihood, parsimony, and heterogeneous evolution. Mol Biol Evol 22:1161–1164

Strelkova MV, Shurkhal AV, Kellina OI, Eliseev LN, Evans DA, Peters W, Chapman CJ, Le Blancq SM, van Eys GJ (1990) A new species of *Leishmania* isolated from the great gerbil *Rhombomys opimus*. Parasitology 101:327–335

Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous time Markov chain evolutionary models. Mol Biol Evol 18:1001–1013

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer, Sunderland, MA

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Tibayrenc M (1998) Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. Int J Parasitol 28:85–104

Wang J, Qu J-Q, Guan L-R (1964a) A study on the *Leishmania* parasites of the great gerbil in northwest China. Acta Parasitologica Sinica 1:105–117 (in Chinese with English abstract)

Wang J, Xiong G-H, Hu Y-D, Lui P-Z, Niu H-F, Sai S-Y (1964b) The *Leishmania* infection found in *Rhombomys opimus* its transmission and relation to man. Acta Parasitologica Sinica 1:17–23 (in Chinese with English abstract)

Wang J-Y, Gao C-H, Yang Y-T, Chen H-T, Zhu X-H, Lv S, Chen S-B, Tong S-X, Steinmann P, Ziegelbauer K, Zhou X-N (2010) An outbreak of the desert sub-type of zoonotic visceral leishmaniasis in Jiashi, Xinjiang Uygur Autonomous Region, People's Republic of China. Parasitol Int. doi:10.1016/j.parint.2010.04.002

WHO (1990) Control of leishmaniases. World Health Organization, Geneva

Xia X, Lemey P (2009) Assessing substitution saturation with DAMBE. In: Lemey P (ed) The phylogenetic handbook. Cambridge University Press, Cambridge, pp 611–626

Xia X, Xie Z (2001) DAMBE: data analysis in molecular biology and evolution. J Hered 92:371–373

Xia X, Xie Z, Salemi M, Chen L, Wang Y (2003) An index of substitution saturation and its application. Mol Phylogenet Evol 26:1–7

Xu Z-B, Le Blancq S, Evans DA, Peters W (1984) The characterization by isoenzyme electrophoresis of *Leishmania* isolated in the People's Republic of China. Trans R Soc Trop Med Hyg 78:689–693

Xu Z-B, Liu Z-T, Long J-Y, Chai J-J, Chen W-K (1989) Further characterization of Chinese *Leishmania* isolates by isoenzyme electrophoresis. Chin Med J (Engl) 102:679–685

Zemanova E, Jirku M, Mauricio IL, Miles MA, Lukeš J (2004) Genetic polymorphism within the *Leishmania donovani* complex: correlation with geographic origin. Am J Trop Med Hyg 70:613–617

Zheng C-J, Wang L-Y, Xu X, Zhu X-H, Wu W-P (2009) Visceral Leishmaniasis in China during 2004–2007. Chin J Parasitol Parasit Dis 27:344–346 (in Chinese with English abstract)