Infection, Genetics and Evolution 38 (2016) 101–109

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Phylogenomic reconstruction supports supercontinent origins for *Leishmania*

Kelly M. Harkins [a,b,*], Rachel S. Schwartz [c], Reed A. Cartwright [c,d], Anne C. Stone [b,e]

[a] Department of Anthropology, Human Paleogenomics Laboratory, University of California, Santa Cruz, Santa Cruz, CA, USA
[b] School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA
[c] The Biodesign Institute, Arizona State University, Tempe, AZ, USA
[d] School of Life Sciences, Arizona State University, Tempe, AZ, USA
[e] Center for Evolution and Medicine, Arizona State University, Tempe, AZ, USA

ARTICLE INFO

Article history:
Received 1 October 2015
Received in revised form 25 November 2015
Accepted 26 November 2015
Available online 18 December 2015

Keywords:
Leishmania
Phylogenomics
Evolution

ABSTRACT

Leishmania, a genus of parasites transmitted to human hosts and mammalian/reptilian reservoirs by an insect vector, is the causative agent of the human disease complex leishmaniasis. The evolutionary relationships within the genus *Leishmania* and its origins are the source of ongoing debate, reflected in conflicting phylogenetic and biogeographic reconstructions. This study employs a recently described bioinformatics method, SISRS, to identify over 200,000 informative sites across the genome from newly sequenced and publicly available *Leishmania* data. This dataset is used to reconstruct the evolutionary relationships of this genus. Additionally, we constructed a large multi-gene dataset, using it to reconstruct the phylogeny and estimate divergence dates for species. We conclude that the genus *Leishmania* evolved at least 90–100 million years ago, supporting a modified version of the Multiple Origins hypothesis that we call the Supercontinent hypothesis. According to this scenario, separate *Leishmania* clades emerged prior to, and during, the breakup of Gondwana. Additionally, we confirm that reptile-infecting *Leishmania* are derived from mammalian forms and that the species that infect porcupines and sloths form a clade long separated from other species. Finally, we firmly place the guinea-pig infecting species, *Leishmania enriettii*, the globally dispersed *Leishmania siamensis*, and the newly identified Australian species from a kangaroo, as sibling species whose distribution arises from the ancient connection between Australia, Antarctica, and South America.

1. Introduction

Leishmania is a genus of parasitic trypanosomatid protozoa responsible for the human disease complex leishmaniasis, which is estimated to cause the ninth largest disease burden among infectious diseases. *Leishmania* spp. are transmitted primarily by sandflies to human hosts and animal reservoirs in over 85 countries worldwide. Leishmaniasis is endemic to poverty-stricken countries, has risen in incidence by nearly two million cases annually, and lacks effective treatment or vaccine (Alvar et al., 2012). At least twenty *Leishmania* species cause disease in humans with three main clinical manifestations: visceral, cutaneous, and mucocutaneous. Despite extensive work since its first description in 1903, conflicting hypotheses persist regarding the evolution of the *Leishmania* genus (Cupolillo et al., 2000; Lainson, 2010; Noyes et al., 2002; Schönian et al., 2010; Van der Auwera et al., 2011). This lack of consensus regarding current species relationships is due to incongruence between molecular and non-molecular data, as well as among molecular studies (Kerr, 2000; Lysenko, 1971). Additionally, available sequence data are skewed towards human-infecting species; comparatively scarce data from wild reservoir/host populations limits evolutionary reconstructions.

The majority of molecular phylogenetic studies of *Leishmania* have relied on a few genetic loci to infer evolutionary relationships (Schönian et al., 2013). These data may not accurately reflect the true history of the genus and also provide limited information with which to estimate the age of splits between clades (Castresana, 2007; Momen and Cupolillo, 2000). The evolutionary history of this genus can be resolved with larger molecular datasets and by using these data to estimate the timing of the splits among clades within the phylogeny (Croan et al., 1997; Fraga et al., 2010; Lukes et al., 2007; Momen and Cupolillo, 2000; Noyes et al., 2000; Stevens and Rambaut, 2001).

1.1. Proposed origin hypotheses

Leishmania are parasites that require two species for a complete life cycle: an insect vector and vertebrate host. Before the development of molecular biology techniques, the genus was divided into two

* Corresponding author at: Department of Anthropology, Human Paleogenomics Laboratory, University of California, Santa Cruz, Santa Cruz, CA, USA.
E-mail address: kmharkin@ucsc.edu (K.M. Harkins).

http://dx.doi.org/10.1016/j.meegid.2015.11.030
1567-1348/

subgenera based upon where the parasite developed within the vector (midgut vs. hindgut) (Lainson and Shaw, 1972). These subgenera are *Leishmania*, consisting of all Old World species and one species complex found in the Americas, and *Viannia*, comprised exclusively of New World species. The species that infect Old World reptiles have since been placed in the subgenus *Sauroleishmania* (Saf'janova, 1982), while many other *Leishmania* species, including newly described lineages, remain unclassified.

The biogeography of vectors and distinct vertebrate hosts has led to the proposal of three hypotheses for the origins of the genus. The Palearctic hypothesis assumes an origin in Cretaceous lizards with recent migrations to the Nearctic and Neotropics across land bridges. This hypothesis suggests that *Sauroleishmania* form a clade that is sister to all other species; it is supported by non-molecular data, namely host-phylogenies, biogeography, and evidence of an ancestral parasite, *Paleoleishmania*, found fossilized in Cretaceous amber in modern-day Burma, that is similar to species restricted to New World mammals (Kerr, 2000, 2006; Kerr et al., 2000; Lysenko, 1971; Poinar, 2007).

Alternatively, *Leishmania* has been hypothesized to originate in the Neotropics. This hypothesis is supported by sequence-based phylogenies (Croan et al., 1997; Cupolillo et al., 2000; Fraga et al., 2010; Noyes et al., 2000; Stevens et al., 2001). In these phylogenies, New World species emerged 46–34 mya and are ancestral to those found in the Old World (Lukes et al., 2007). Thus, the reptile-infecting forms are derived from mammalian forms of the parasite rather than being ancestral to them, contrary to the Palearctic hypothesis. The Neotropical hypothesis, however, requires two separate intercontinental migrations of the parasites: first the ancestor of *Leishmania*/*Sauroleishmania* to the Old World approximately 24–14 mya (Lukes et al., 2007), followed by the migration of a member of the *Leishmania* subgenus back to the New World. Alternatively, there might have been two migrations into the Old World: once by an ancestor of *Sauroleishmania* and second by the ancestor of the Old World *Leishmania* species. The timing of these migrations must have coincided with appropriate land bridges.

Finally, the Multiple Origins hypothesis suggests division of *Leishmania* into two lineages on Gondwana (Momen and Cupolillo, 2000). One lineage led to the subgenera *Leishmania*, *Viannia*, and *Sauroleishmania* (termed Euleishmania), while the second diverged into all other species, which to date are restricted to New World mammals (Paraleishmania) (Cupolillo et al., 2000). The breakup of Gondwana subsequently separated the ancestors of the *Viannia* subgenus and the *Leishmania* and *Sauroleishmania* subgenera in South America and Africa, respectively. The amount of diversity observed between the *Leishmania* and *Viannia* subgenera has been cited as evidence for vicariance due to the separation of Africa from the Neotropics (Fernandes et al., 1993). *L. Leishmania* spp. (e.g. *L. Leishmania amazonensis*) were subsequently brought from Africa through Eurasia to North America. Due to the insect vector's short lifecycle and weak flying ability this process must have occurred when conditions were conducive to vector survival and the migration of vertebrate hosts (Tuon et al., 2008). The most recent date of an introduction of *Leishmania* subgenus into the Neartic via Beringia would have been the mid-Miocene when temperatures were warm enough for sandfly survival (Stevens and Rambaut, 2001).

This study represents the first phylogenomic analysis of *Leishmania*, employing over 200,000 variable sites and 49 genes from across the genome. This large dataset counters previous challenges suggesting that different substitution rates in different genes can confound estimates of relationships among *Leishmania* (Momen and Cupolillo, 2000). The size of the dataset also allows us to estimate the timing of divergence among clades, which until now has been purely speculative. Our results allow us to clarify and expand the Multiple Origins hypothesis, which proposes that *Leishmania* clades separated prior to, and during, the breakup of Gondwana. We term this revision the Supercontinent hypothesis. Additionally, we confirm that reptile-infecting *Leishmania* are derived from mammalian forms, and that the species that infect porcupines and sloths form a clade long separated from other species. We also firmly place the guinea-pig infecting species, *Leishmania enriettii*, the globally dispersed *Leishmania siamensis*, and the newly identified Australian species found in kangaroos, as sibling species whose distribution arises from the ancient connection between Australia, Antarctica, and South America.

2. Materials and methods

2.1. Whole genome shotgun sequencing

We sequenced whole genomes of twelve species of *Leishmania* and one species of *Endotrypanum* (Table 1). Organisms were grown as promastigotes (Wirth and Pratt, 1982) at 22 °C in Schneider's Medium supplemented with 20% heat inactivated FCS and 17.5 mg/mL gentamycin. The cells were pelleted and washed twice in phosphate-buffered saline (PBS). Genomic DNA was extracted according to established methods. The cells were incubated in lysis buffer (100 mM NaCL, 10 mM TrisCl pH 8.0, 0.5 mM EDTA, 0.5% SDS, 0.1 mg/mL fresh Proteinase K) for 12–18 h at 50 °C, and purified via phenol/chloroform/isopropyl alcohol. Extract was incubated in RNase A (10 mg/mL) for 2.5 h at 37 °C and dialyzed overnight at 4 °C with three changes of PBS buffer. DNA concentration was evaluated with spectrophotometry (Beckman Coulter DU730) and re-extracted if contaminated with phenol. Paired-end reads (100 bp) were sequenced on an Illumina HiSeq2000 by the University of Arizona Genome Core. The number of pairs of reads for each species ranged from 2.1 to 14.3 million, or 10–67× average coverage for the ~34 million bp genome. The raw data were deposited in the National Center for Biotechnology Information (NCBI) database and Sequence Read Archive under BioProject ID PRJNA267749.

Additional shotgun genomic data were downloaded from the European Nucleotide Archive (ENA; Table S1). These data represent all publicly available genome data per *Leishmania* taxon.

2.2. Phylogenetically informative data

2.2.1. Variable sites from across the genome

Although reference genomes are available for *Leishmania*, initial attempts to align shotgun sequencing reads to these references resulted in low alignment rates. Therefore, to extract phylogenetically informative data from the available shotgun sequences, we used the bioinformatics pipeline SISRS (Schwartz et al., 2015). This method identifies sites that are fixed for each species and variable across species to construct a multiple-species alignment. SISRS determines the nucleotide at a site in each species via strict consensus. If a site is not identical in all reads or there are no data for that site in a species, the information is considered "missing". We produced dataset VS-m6 allowing up to six taxa to have missing information per site (missing data were allowed to ensure sufficient data to determine the phylogeny). To ensure that linked sites would not bias our results, we subsampled this dataset by sampling only one site per sequence fragment, producing dataset VS-m6s. Finally, we produced an additional alignment, VS-t80, with no missing data allowed, but with a lower calling threshold of 0.8, i.e. 80% of bases for that taxon must be one allele.

2.2.2. Additional gene data

Because branch length estimates can be incorrect for variable site datasets, we developed an additional dataset of 49 genes with putative or known function. This dataset also allowed us to compare the results of two approaches to estimate the *Leishmania* phylogeny. We first obtained these genes from *Leishmania* (*Viannia*) *braziliensis* reference sequences (Table 2); the emphasis on this species was chosen on account of a related project. The sequence of each gene for all other species was then extracted from the shotgun sequencing reads using the

Table 1
Leishmania strains sequenced in this analysis. *Leishmania* isolates are from the laboratories of Diane McMahon-Pratt, Yale School of Public Health and Lucile Floeter-Winter, University of São Paulo. Raw data are curated on NCBI (BioProject ID PRJNA267749) and the Sequence Read Archive (SRA).

Species name	Isolate ID	Subgenus	Region	Location	Host of isolate	WHO Number	Lab PI	SRA Experiment Accession
Endotrypanum schaudinni	M6159	n/a	NW	Brazil	Edentata/*Choloepus*	MCHO/BR/80/M6159	McMahon-Pratt	SRX767388
L. deanei	M5088	n/a	NW	Brazil	Rodentia/*Coendou*	MCOE/BR/00/M5088	McMahon-Pratt	SRX767381
L. hertigi	M4051	n/a	NW	Panama	Rodentia/*Coendou*	MCOE/PA/00/M4051, also referred to as LV42 and C-8	McMahon-Pratt	SRX767385
L. enriettii	L88	n/a	NW	Brazil	Rodentia/*Cavia*	MCAV/BR/45/L88	McMahon-Pratt	SRX767380
L. L. tropica	LRC-L39	Leishmania	OW	Soviet Union	Primates/*Homo*	MHOM/SU/60/LRC-L39	McMahon-Pratt	SRX767378
L. L. aethiopica	LRC-L147	Leishmania	OW	Ethiopia	Primates/*Homo*	MHOM/ET/71/L100	McMahon-Pratt	SRX767387
L. L. mexicana pifanoi	Ltrod	Leishmania	NW	Venezuela	Primates/*Homo*	MHOM/VE/00/Ltrod	McMahon-Pratt	SRX767386
L. adleri	LRC-L123	"Sauroleishmania"	OW	Kenya	Reptile/*Latastia longicaudata*	RLAT/KE/57/SKINK-7	McMahon-Pratt	SRX764330
L. V. panamensis	WR120	Viannia	NW	Panama	Primates/*Homo*	MHOM/PA/74/WR120	McMahon-Pratt	SRX767384
L. V. guyanensis	M4147	Viannia	NW	Brazil	Primates/*Homo*	MHOM/BR/75/M4147	McMahon-Pratt	SRX767379
L.V. shawi	M8408	Viannia	NW	Brazil	Primates/*Cebus*	MCEB/BR/1984/M8408	Floeter-Winter	SRX764331
L.V. naiffi	M5533	Viannia	NW	Brazil	Edentata/*Dasypus*	MDAS/BR/1979/M5533	Floeter-Winter	SRX764332
L.V. lainsoni	M6426	Viannia	NW	Brazil	Primates/*Homo*	MHOM/BR/1981/M6426	Floeter-Winter	SRX764333

following pipeline: (1) Reads from each species were aligned to the reference genes using Bowtie2 (Langmead and Salzberg, 2012). (2) We used the mpileup feature of SAMTools (Li et al., 2009) to obtain the information from each read for each site. (3) The base for each site was identified based on whether at least 80% of the reads at that site contained a single base. (4) Alignments were adjusted using MAFFT (Katoh and Standley, 2013) with default settings. This pipeline is now automated as a part of SISRS.

We then added to the gene dataset all data available for two recently described *Leishmania* species: an isolate from Australian kangaroos

Table 2
Genes used in analysis according to the reference sequence.

GeneInfo identifier	Reference ID	*L. braziliensis* strain ID	Gene
>gi 154345881	XM 001568828.1	MHOM/BR/75/M2904	Tyrosine aminotransferase (TAT)
>gi 154337345	XM 001564856.1	MHOM/BR/75/M2904	RNA polymerase II (LBRM 21 2050)
>gi 389602506		MHOM/BR/75/M2904	Prostaglandin f2-alpha synthase/D-arabinose dehydrogenase (pgfs)
>gi 389603341	XM 001569011.2	MHOM/BR/75/M2904	Putative oxidoreductase (LBRM 35 4420)
>gi 112383574	gbDQ836162.1	HOM/BR/75/M2903	N-acetylglucosamine-1-phosphate transferase gene partial cds (nagt)
>gi 154346249	XM 001569012.1	MHOM/BR/75/M2904	Putative UDP-N-acetylglucosamine-dolichyl-phosphate N-acetylglucosaminephosphotransferase (LBRM 35 4430)
>gi 154336055	XM 001564214.1	MHOM/BR/75/M2904	Mitogen-activated protein kinase (MPK4)
>gi 155675719	gbEU053119.1	MHOM/PE/88/BAA2079	Mannose phosphate isomerase (MPI) gene
>gi 155675703	gbEU053111.1	MHOM/PE/88/BAA2079	Malate dehydrogenase (MDH) gene
>gi 389602969		MHOM/BR/75/M2904	HSP83 heat shock protein 83-1
>gi 9864198	gbAF291716.1		Heat shock protein 70 (hsp70) gene
>gi 322505745		MHOM/BR/75/M2904	Chromosome 36 GPI alpha-mannosyltransferase III
>gi 154341936	XM 001566870.1	MHOM/BR/75/M2904	Glyceraldehyde 3-phosphate dehydrogenase, glycosomal (LBRM 30 2950)
>gi 154332907	XM 001562666.1	MHOM/BR/75/M2904	Elongation factor-1 gamma (LBRM 09 1020)
>gi 154335073	XM 001563727.1	MHOM/BR/75/M2904	Elongation factor 1-alpha (LBRM 17 0090)
>gi 2581879	gbAF009138.1		DNA polymerase alpha gene
>ENA CAM37041.1		MHOM/BR/75/M2904	Dihydrofolate reductase-thymidylate synthase (dhfr-ts)
>ENA AB434681.1		MHOM/BR/00/LTB300	Kinetoplast pre-edited Cytb gene for cytochrome b
>gi 154345431	XM 001568603.1	MHOM/BR/75/M2904	Elongation factor 2 (LBRM 35 0270)
>gi 154342453	XM 001567125.1	MHOM/BR/75/M2904	Putative cytochrome c oxidase VIII (COX VIII) (LBRM 31 1780)
>gi 154343626	XM 001567709.1	MHOM/BR/75/M2904	Myosin XXI (LBRM 32 4110)
>gi 154340941	XM 001566374.1	MHOM/BR/75/M2904	Putative heat shock protein 90 (LBRM 29 0780)
>gi 389604019	XM 003723107.1	MHOM/BR/75/M2904	Arginine N-methyltransferase-like protein (LBRM 20 6060)
>gi 389604005	XM 003723100.1	MHOM/BR/75/M2904	Serine peptidase (LBRM 20 6000)
>gi 389603991	XM 003723093.1	MHOM/BR/75/M2904	Putative N-acyl-L-amino acid amidohydrolase (LBRM 20 5930)
>gi 389603983	XM 003723089.1	MHOM/BR/75/M2904	Mitochondrial RNA ligase 2 (LBRM 20 5890)
>gi 389603973	XM 003723084.1	MHOM/BR/75/M2904	Putative separin (LBRM 20 5840)
>gi 389603961	XM 003723078.1	MHOM/BR/75/M2904	Cell division cycle protein-like protein (LBRM 20 5780)
>gi 389603935	XM 003723065.1	MHOM/BR/75/M2904	Putative axoneme central apparatus protein (LBRM 20 5640)
>gi 389603929	XM 003723062.1	MHOM/BR/75/M2904	Putative RNA-binding regulatory protein (LBRM 20 5610)
>gi 389603919	XM 003723057.1	MHOM/BR/75/M2904	Putative coatomer beta subunit (LBRM 20 5560)
>gi 389603917	XM 003723056.1	MHOM/BR/75/M2904	Kinase-like protein (LBRM 20 5550)
>gi 389603921	XM 003723058.1	MHOM/BR/75/M2904	Putative exosome associated protein 1 (Rrp42 homologue) (LBRM 20 5570)
>gi 389603879	XM 003723037.1	MHOM/BR/75/M2904	Triosephosphate isomerase (LBRM 20 5360)
>gi 389603873	XM 003723034.1	MHOM/BR/75/M2904	Endo-1, 4-beta-xylanase z precursor-like protein (LBRM 20 5330)
>gi 389603843	XM 003723019.1	MHOM/BR/75/M2904	Putative glutaredoxin (LBRM 20 5180)
>gi 389603811	XM 003723003.1	MHOM/BR/75/M2904	Cytochrome c oxidase assembly factor-like protein (LBRM 20 5020)
>gi 389603809	XM 003723002.1	MHOM/BR/75/M2904	Phosphopantetheinyl transferase-like protein (LBRM 20 5010)
>gi 389603777	XM 003722986.1	MHOM/BR/75/M2904	Conserved hypothetical protein (LBRM 20 4850)
>gi 389603767	XM 003722981.1	MHOM/BR/75/M2904	Chaperone protein DNAJ-like protein (LBRM 20 4790)
>gi 389603765	XM 003722980.1	MHOM/BR/75/M2904	Conserved hypothetical protein (LBRM 20 4780)
>gi 154341936	XM 001566870.1	MHOM/BR/75/M2904	Glycosomal glyceraldehyde 3-phosphate dehydrogenase (LBRM 30 2950)

(strain AM-2004) (Rose et al., 2004) and *L. siamensis* (Kanjanopas et al., 2013; Leelayoova et al., 2013). These data were downloaded from GenBank. For AM-2004 we obtained partial sequences of three loci. For *L. siamensis* we obtained partial sequences of six loci, which includes two genes (Table S2).

2.3. Phylogenetic analysis

2.3.1. Concatenated variable site analysis (3 datasets)

The practice of concatenating thousands of likely unlinked sites into a single alignment (Philippe et al., 2011) can merge the disparate genealogical histories of different chromosomal regions. However, variable sites cannot be partitioned by linkage to separate the history of genes from the history of the species, and it is computationally challenging to consider each site separately for a dataset of this size. We thus treated the SISRS variable site data as a single concatenated locus (Yoder et al., 2013). For each dataset, we constructed a phylogeny using maximum likelihood (ML) in RAxML 8.0.20 (Stamatakis, 2014) with a General Time Reversible (GTR) model and substitution rates following a discrete gamma distribution with four categories for 1000 bootstrap replicates, allowing for ascertainment bias correction using the Lewis model (Lewis, 2001).

2.3.2. Gene data

For coding genes, partitioning by codon position was found to provide a better fit than by gene alone (Mueller et al., 2004). The best way to partition the data was determined by the program *partitionfinder* with default settings (Lanfear et al., 2012). Input partitions were separate codon positions for each gene. We employed the Bayesian Information Criterion for model selection to avoid overparameterization. The resulting 20 partitions were used to estimate the phylogeny in an ML framework; this analysis was implemented in RAxML using a GTR model with gamma distributed rate heterogeneity across sites, for 100 bootstrap replicates. This analysis was repeated five times with random starting seeds to identify the ML tree and avoid local optima.

2.3.3. Individual gene analysis

ML trees for each gene were constructed in RAxML with 1000 bootstrap replicates. Although concatenated multi-gene datasets are found generally to be robust in phylogenetic inference even without explicitly accounting for large variations in rates, length, and GC content, systematic biases in the dataset can lead to high support of the wrong tree (Gadagkar et al., 2005). We follow the recommendation of Gadagkar et al. (2005) to report the gene support frequency for each given partition. A majority rule (MR) consensus tree was generated in RAxML from the best scoring ML trees of individual gene trees for which we had all ingroup taxa. When necessary, the outgroup *Crithidia* was pruned using Newick Utilites (Junier and Zdobnov, 2010). All phylogenetic trees were visualized with the ape (Paradis et al., 2004) and phytools (Revell, 2012) packages in R (R Core Team, 2013).

2.3.4. Estimating divergence time

We used two approaches to estimate the timing of divergences among clades. Time estimates allow us to evaluate the plausibility of the proposed hypotheses for the origin of *Leishmania*. Due to poor fossil preservation of *Leishmania*, no secure fossil calibration dates within the genus exist. Thus, our first approach was to add two outgroup species, *Trypanosoma cruzi* and *Trypanosoma brucei*, to the concatenated gene dataset. The divergence date for these species is believed to be 100 million years ago (mya) based on the timing of the split between Africa and South America (Lukes et al., 2007; Stevens and Gibson, 1999). Using our known tree for *Leishmania* with these two additional species, divergence times were estimated using RelTime (Kumar et al., 2012; Tamura et al., 2012, 2013).

Our second approach was to use a calibration date of 40 million years for the split between *L. enriettii*, *L. siamensis* and *Leishmania* sp. Ghana, with the Australian isolate *Leishmania* sp. AM-2004, which corresponds to the breaking of the connection between South America and Australia via Antarctica. In this analysis we only used the sequences available for AM-2004 to avoid any effect of missing data on branch lengths. For this reason, we were able to include additional *Leishmania* taxa for which these few loci were also publicly available (Table S3). We constructed dated phylogenies using BEAST v1.8.2 (Drummond et al., 2012) with two unlinked partitions, one for 18s/ITS/5.8s and the other for RNA polymerase II large subunit. We estimated substitution models in jModeltest (Darriba et al., 2012); based on those results we implemented a TN93 substitution model with rates estimated from a gamma distribution with four rate categories and a relaxed lognormal molecular clock for 18,000,000 generations (high ESS values and convergences were observed in Tracer 1.5 (Rambaut et al., 2014) at this point). A normally distributed prior with a mean of 40 mya was specified for the node leading to Australia kangaroo isolate, *Leishmania* sp. AM-2004.

3. Results

3.1. Phylogenetic analyses

3.1.1. Variable sites data

The numbers of variable sites identified for each dataset were 215,644 (VS-m6), 18,312 (VS-m6s), and 2790 (VS-t80). Based on an alignment with the reference genome for *Leishmania tarentolae*, we determined that these sites are distributed across the *Leishmania* genome. GC content was over 70%, which is higher than the GC content estimated previously of 50–60% (Peacock et al., 2007). The ML phylogeny for all of these datasets supports the *Leishmania* and *Viannia* subgenera as monophyletic clades; *Sauroleishmania* is also monophyletic, although only two taxa were available for phylogenetic analysis, and it is sister to the clade comprising the *Leishmania* subgenus (Fig. 1). *L. enriettii* is supported as sister to all other Euleishmania. *Leishmania hertigi*, *Leishmania deanei*, and *Endotrypanum* (Paraleishmania) form a clade that is sister to all other *Leishmania* lineages (Euleishmania).

3.1.2. Partitioned genes

A total of 70,447 sites in 49 genes were concatenated in the alignment. The resulting majority rule consensus of the ML tree is identical to that for the variable sites data, with the exception of the placement of *L. enriettii* as sister to the clade containing the subgenera *Leishmania* and *Sauroleishmania*, rather than as sister to Euleishmania (Fig. S1). However, this placement was supported by less than 50% of individual gene trees. Interestingly, when ML trees were constructed separately from each codon position the phylogeny for each of these three trees was identical to that constructed for the variable site data, although the third codon position phylogeny had low support at several nodes. In all trees constructed using the gene datasets, *Leishmania* sp. AM-2004 and *L. siamensis* are most closely related to *L. enriettii*.

3.2. Divergence time estimates

Divergence dates were first estimated from the gene dataset for *Leishmania* and two species of *Trypanosoma* using Reltime (Kumar et al., 2012; Tamura et al., 2012, 2013) with a calibration point 100 mya at the split between two outgroup species, *T. cruzi* and *T. brucei* (Fig. 2). Because we were unable to obtain genome data for the sister genus to *Leishmania* (*Leptomonas*), we approximate the origin of *Leishmania* based on the timing of the split of the genus into two clades, and the divergence of *Leishmania* and *Crithidia*. These dates lead to an approximate origin of 90 mya.

Secondly, divergence dates were estimated using the loci for which we had available data from the Australian isolate, allowing us to include

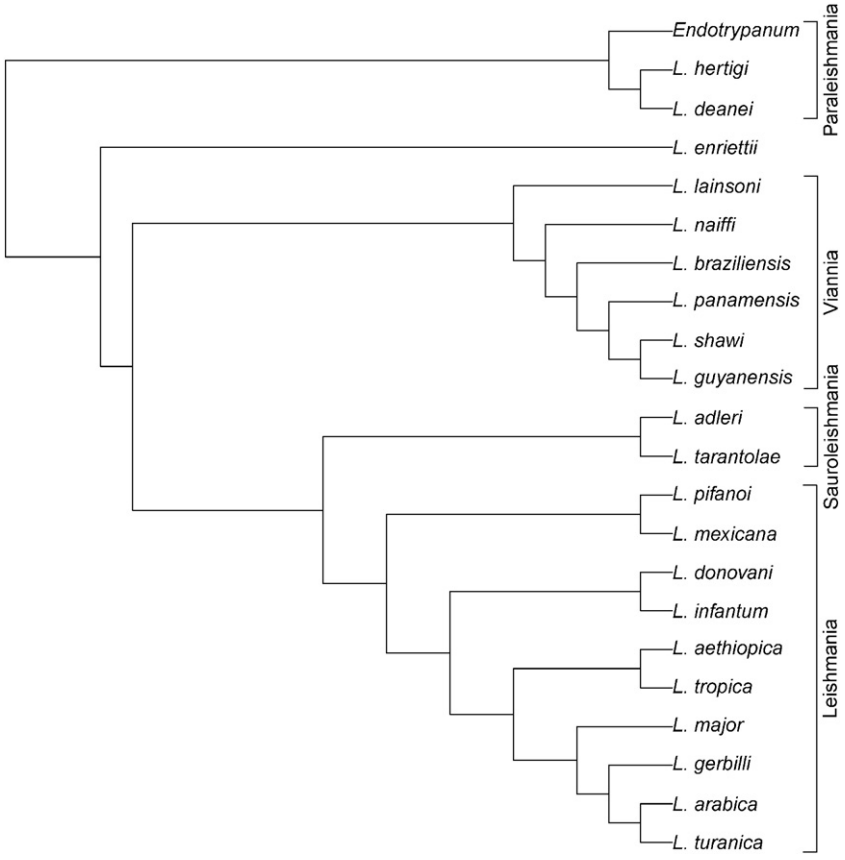

Fig. 1. Maximum likelihood tree based on 215,644 variable sites from across the genome. Data were identified using SISRS (Schwartz et al., 2015). Up to six samples were allowed to have data missing or be ambiguous for any given site. The phylogeny was constructed with RAxML 8.0.20 (Stamatakis, 2014) using the Lewis correction for exclusively variable site data (Lewis, 2001). Bootstrap support values are 100 at every node. Identical results were obtained using a dataset containing only one site per sequence fragment, and one with no missing data allowed, but with a lower calling threshold of 80%.

Leptomonas as an outgroup. This analysis was performed in BEAST v1.8.2 (Drummond et al., 2012), setting a normal prior distribution of 40 mya on the node ancestral to this isolate and other members of the "*L. enriettii* complex". These results suggest an older origin of the genus of approximately 140 mya (Figs. 3 and S2).

4. Discussion

4.1. Data

Shotgun sequencing data of thirteen *Leishmania* isolates (this study) and eleven species available publicly enabled us to mine whole genome datasets for hundreds of thousands of phylogenetically informative sites. Previous molecular datasets used for phylogenetic reconstruction ranged from 1 to 31 genes (Tschoeke et al., 2014). Rather than relying on a few known genes, our data were sampled across the genome. The SISRS method is particularly valuable for *Leishmania* spp., which like many other non-model organisms, do not have a suitable reference genome for this purpose.

4.2. Phylogeny

Our results place species of Old and New World subgenus *Leishmania* in a monophyletic clade separate from species of the exclusively New World subgenus *Viannia*. These results are consistent with other molecular-based trees, including those that place *Sauroleishmania* as sister to the *Leishmania* subgenus (Dougall et al., 2011; Fraga et al., 2010).

Our genome-wide, variable-site tree places *L. enriettii*, a parasite of the New World guinea pig, sister to all Euleishmania. This result is consistent with some prior molecular phylogenies (Croan et al., 1997; Noyes et al., 2002; Stevens et al., 2001); however, it differs from early hypotheses that suggest *L. enriettii* is a member of the New World *L. Leishmania* subgenus (Lainson, 1997; Lainson and Shaw, 1987), or later work suggesting it is sister to the entire genus (Kerr, 2006; Yurchenko et al., 2006). This tree is also supported by separate analyses of each codon positions in our gene dataset.

Although the full concatenated gene dataset produces a conflicting topology whereby *L. enriettii* is not basal to all Euleishmania but only to the *Sauroleishmania*/*Leishmania* subgenera, we believe the latter relationships are not correct. Substitutions are expected to be most rapid at the third position (Kimura, 1968). At these time scales, saturation in the third position may negatively affect the phylogenetic signal. The difference in these results is not necessarily surprising given the short branches leading to the splits of *L. enriettii* and the Euleishmania clades, which often reflects incomplete lineage sorting and would lead to different relationships estimated from different genes. Adding more loci, as we have done with the variable site dataset, is advocated in this case to increase resolution (Maddison and Knowles, 2006). It is important to note that none of our phylogenies support the current NCBI classification of *L. enriettii* as a member of the *L.* (*Leishmania*) *mexicana* species complex.

As with nearly all other analyses to date, the molecular evidence also does not support the taxonomic classification of *Endotrypanum* as a separate genus. Relative to molecular markers, there are few morphological characters used to classify unicellular organisms, calling their utility into question, when compared to molecular data (Perkins et al., 2011). Sequence data support their inclusion in the *Leishmania* genus, unless all other Paraleishmania species are misclassified.

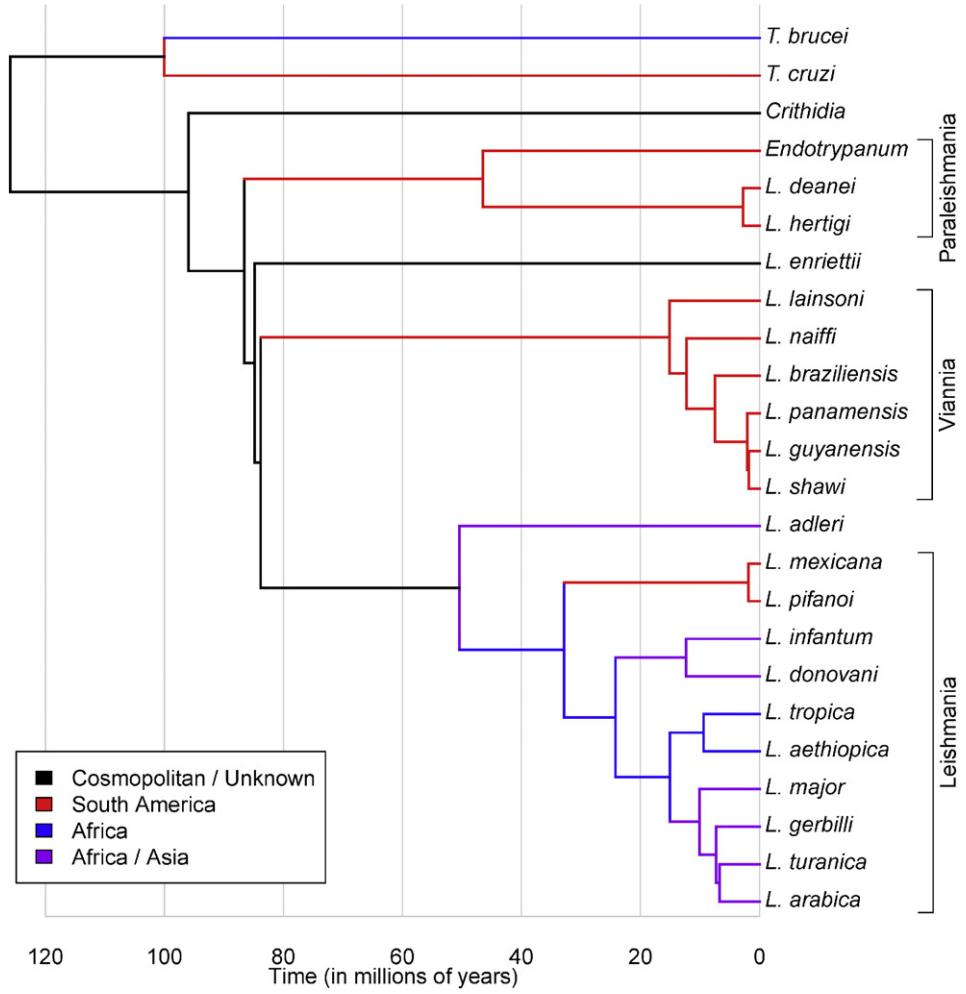

Fig. 2. Dated phylogeny based on 49 genes constructed with Reltime (Kumar et al., 2012; Tamura et al., 2012; Tamura et al., 2013) using a calibration date of 100 mya for the split of *Trypanosoma brucei* and *T. cruzi*.

4.3. Dates

The calibration of 40 mya for the isolation of the Australian isolate *Leishmania* sp. AM-2004, like other fossil dates, provides only a minimum time since two species diverged. However, estimated dates on other nodes within the tree resulting from this calibration, for example the split between *Leishmania donovani* and *Leishmania major*, are consistent with previously published dates. The date calculated for the *L. donovani–L. major* split was 24.2 mya using the large gene dataset, Reltime, and a calibration of 100 mya for the split between *T. cruzi* and *T. brucei*. The date calculated for this split based on the dataset with fewer genes, more samples, BEAST, and a calibration of 40 mya on the node ancestral to the Australia isolate was 21.1 mya. Previously published dates for this split are 24.7–14.6 mya (Lukes et al., 2007). The consistency between our results and those published previously suggests that results at other nodes may be considered reasonably valid when evaluating hypotheses for the origin of *Leishmania*.

4.4. Origins hypotheses

4.4.1. The Palearctic hypothesis

Our results allow us to reject the Palearctic origin hypothesis for *Leishmania* (Kerr, 2000, 2006). First, *Sauroleishmania* are not sister to all other *Leishmania*, as would be suggested by this hypothesis. Second, our estimated dates conflict with a Pliocene introduction of the subgenus *Viannia* to the Neotropics at ~5 mya after the reformation of the Panama isthmus. If Kerr (2000) is correct in asserting that reptiles were the first hosts of *Leishmania* in the Cretaceous, extinction events associated with the K–T boundary could have erased evidence of those lineages; however, this speculation does not affect our interpretation.

4.4.2. The Neotropical hypothesis

Our results also allow us to reject the Neotropical Origins hypothesis. Under this hypothesis, the current global distribution of *Leishmania* is a purely a result of dispersal through vector/reservoir migration and not vicariance. Lukes et al.'s (2007) version of the hypothesis proposes the ancestor of *Leishmania* evolved in the Neotropics between 46 and 36 mya. This date is derived from their hypothesis that the ancestor of the Old World species within the Leishmania subgenus migrated to the Old World via the Bering Land Bridge prior to the split of the *L. donovani* complex from other *L. Leishmania* spp. 24–14 mya. It is unclear whether this ancestor refers to the entire *Leishmania*/*Endotrypanum* clade, or some more recent ancestor; however, this hypothesis would require an additional migration of the ancestor of Sauroleishmania, which they do not discuss. Similarly, Noyes et al. (1997) and Noyes (1998) proposed that the *Leishmania*/*Endotrypanum* clade evolved in the Neotropics between 65 and 40 mya and dispersed to the Nearctic and to the Palearctic through land bridges. Regardless of which date is considered, our results suggest an origin of the clade 90–140 mya, which is not consistent with these proposed dates.

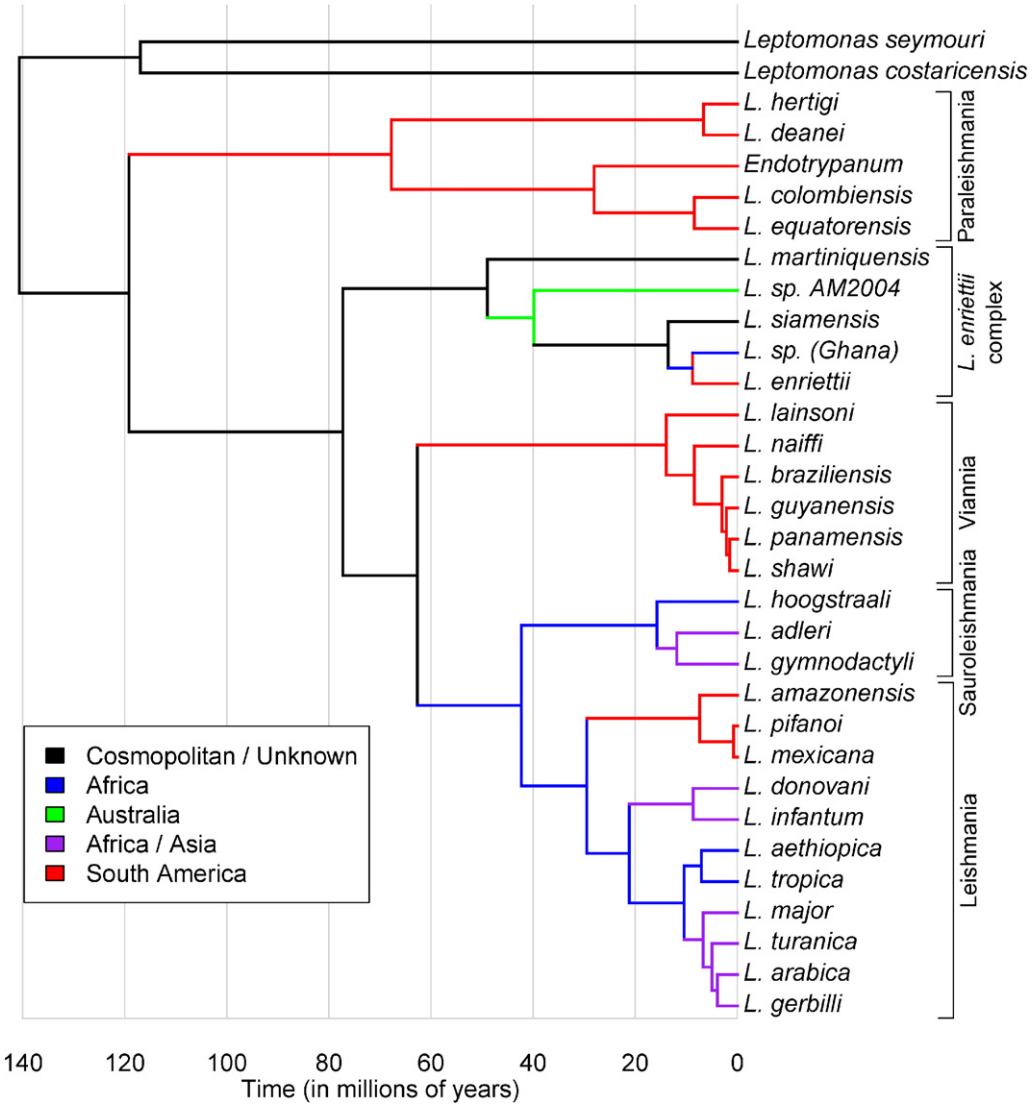

Fig. 3. Phylogeny constructed using the loci available for *L.* sp. AM-2004 in BEAST v1.8.2 (Drummond et al., 2012) with a relaxed molecular clock and a calibration date of 40 mya for the split of the AM-2004 lineage. See Fig. S2 for posterior probability and 95% HPD error bars. This phylogeny places the Australia sample as sister to *L. siamensis* and *L. enriettii* and suggests that the *Leishmania* genus emerged prior to the breakup of Gondwana.

4.4.3. The Multiple Origins hypothesis

The Multiple Origins hypothesis refers to the separation and subsequent independent evolution of the *Viannia* and *Leishmania* subgenera due to the separation of South America from Africa approximately 100 mya (Momen and Cupolillo, 2000). This is the only formalized hypothesis that includes vicariance as a mechanism for the evolution of the major *Leishmania* clades. This hypothesis also suggests an ancient separation of Paraleishmania from all other Leishmania species.

We propose a modification and expansion of the Multiple Origins hypothesis, which we now term the Supercontinent hypothesis, for the origin of the *Leishmania* genus. In this scenario, the ancestor of *Leishmania* emerged from monoxenous parasites (those found in a single host) on Gondwana (Yurchenko et al., 2006). As with the Multiple Origins hypothesis, the split between Paraleishmania and all other species occurred around the time of the separation of Gondwana ~90–100 mya. The Supercontinent proposal is in agreement with speculations by Shaw (1997), who suggested that an adaptation to mammals occurred around 90 mya when mammals began to radiate and Africa became fully isolated. Additionally, levels of genetic diversity between the *Viannia* subgenera and *Sauroleishmania*/*Leishmania* subgenera have been cited previously as a reflection of vicariance after the separation of South America and Africa (Fernandes et al., 1993). These ideas are consistent with the estimated divergence dates presented here using two methods with two types of datasets, one genome-wide and one gene-based, with separate calibration information, respectively.

Only one migration of a lineage in the *Leishmania* subgenus back to the New World is required by this hypothesis (notwithstanding the recent transfer of *Leishmania infantum* to the New World by European settlers, often termed *L. infantum chagasi* (Marcili et al., 2014; Shaw, 2006)). This migration has been proposed to have occurred via the Neartic and Beringia during the mid-Miocene when temperatures were warm enough for sandfly survival (Stevens and Rambaut, 2001). This timing is roughly consistent with our observation that New World *L. Leishmania* diverged approximately 30 mya. Furthermore, divergent *L.* (*Leishmania*) *amazonensis* isolates have been identified in central and western China, which is consistent with the clade's phylogenetic placement within Old World *Leishmania* spp. (Waki et al., 2007). The global distribution of the phlebotomine sandfly genera that almost exclusively serve as vectors for the parasite likely resulted from the

breakup of Pangaea and subsequent continental splintering (Filho and Brazil, 2003; Galati, 1995).

Our results are consistent with Early Cretaceous fossils of *Paleoleishmania proterus* found in sandflies trapped within Burmese amber ~100 mya that are reportedly evidence of the first digenetic trypanosomatids (Poinar, 2007). There is no way to confirm the genus but the organisms are morphologically similar to *Leptomonas*, the sister of *Leishmania*. Interestingly, this species is believed to be associated with reptile hosts.

The recent discovery of *L. siamensis* and other linages that fall within the "*L. enriettii* complex", a clade basal to all Euleishmania, in humans and other mammals in North America, Europe, West Africa and Asia further highlights the plausibility of an ancient global dispersal predating continental splintering. At least two of these species, *L. siamensis* and *Leishmania martiniquensis*, occur globally. Because few data are currently available for all lineages within the new *L. enriettii* clade, further work is necessary to rule out recent introduction of *Leishmania* to those regions. However, *L. martiniquensis* is believed to be endemic to Martinique Island and Thailand (Desbois et al., 2014; Liautaud et al., 2015; Pothirat et al., 2014), while autochthonous cases of *L. siamensis* infection have been found in Thailand, Switzerland, and Florida (Kanjanopas et al., 2013; Lobsiger et al., 2010; Muller et al., 2009; Reuss et al., 2012). The placement of AM-2004 as sister to lineages found in the New World (*L. enriettii*) is consistent with the biogeography of other groups separated by the breakup of Australia, Antarctica, and southern South America (e.g. the plant genus *Nothofagus*).

The Supercontinent hypothesis for *Leishmania* genus also draws parallels with a related kinetoplastid parasite, *Trypanosoma*. Stevens et al. (2001) hypothesize a southern-supercontinent origin in which *T. brucei* evolved in Africa and *T. cruzi* in the New World following the breakup of Gondwana, although a recent 'bat-seeding' hypothesis (Hamilton et al., 2012) introduces the possibility of an even more complex scenario of dispersal and vicariance for *Trypanosoma*. Divergent lineages of *Trypanosoma* have also been found in Australia (Stevens et al., 1999); therefore much like the position of the kangaroo *Leishmania* isolate is critical in our proposed origins scenario. This parallel pattern of evolution between related parasites *Trypanosoma* and *Leishmania* may suggest similar evolutionary processes.

The most ancestral lineages of *Leishmania* — those crucial to resolve the evolutionary history of the genus — are not only those typically found in wild reservoir populations, but also the samples for which fewest data exist. Animal reservoirs are critical for maintaining *Leishmania* in the wild. Reptiles and mammals, especially bats, that are currently excluded as potential reservoirs may reflect under-sampling rather than a true absence in the coevolutionary history and dissemination of *Leishmania*. Broader sampling strategies that refocus on wild host and vector populations will shed light on the deepest nodes of the phylogeny and ultimately on the processes of zoonotic transfer to humans. Until additional NGS data are available from unclassified and newly described taxa, we offer genome wide data and multiple approaches to estimate divergence dates, further contributing to our understanding of the evolution and origin of *Leishmania*.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.meegid.2015.11.030.

Acknowledgments

Genomic DNA from *Leishmania* isolates was obtained from the lab of Diane McMahon-Pratt at the Yale School of Public Health and from Lucille Floeter-Winter at the University of São Paulo. This work was supported by a National Science Foundation Doctoral Dissertation Improvement Grant [grant number BCS-1232582 to K. Harkins and A. Stone], internal funds from ASU School of Life Sciences [R. Cartwright], and a National Science Foundation Advances in Bioinformatics Grant [grant number DBI-1356548 to R. Cartwright]. Sudhir Kumar and Jay Taylor provided guidance and feedback on the development of the project.

References

Alvar, J., Velez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., den Boer, M., Team, W.H.O.L.C., 2012. Leishmaniasis worldwide and global estimates of its incidence. PLoS One 7, e35671.

Castresana, J., 2007. Topological variation in single-gene phylogenetic trees. Genome Biol. 8, 216.

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. Mol. Biochem. Parasitol. 89, 149–159.

Cupolillo, E., Medina-Acosta, E., Noyes, H., Momen, H., Grimaldi Jr., G., 2000. A revised classification for *Leishmania* and *Endotrypanum*. Parasitol. Today 16, 142–144.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Desbois, N., Pratlong, F., Quist, D., Dedet, J.P., 2014. *Leishmania* (*Leishmania*) *martiniquensis* n. sp. (Kinetoplastida: Trypanosomatidae), description of the parasite responsible for cutaneous leishmaniasis in Martinique Island (French West Indies). Parasite 21, 12.

Dougall, A.M., Alexander, B., Holt, D.C., Harris, T., Sultan, A.H., Bates, P.A., Rose, K., Walton, S.F., 2011. Evidence incriminating midges (Diptera: Ceratopogonidae) as potential vectors of *Leishmania* in Australia. Int. J. Parasitol. 41, 571–579.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Fernandes, A.P., Nelson, K., Beverley, S.M., 1993. Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. Proc. Natl. Acad. Sci. U. S. A. 90, 11608–11612.

Filho, J.D., Brazil, R.P., 2003. Relationships of new world phlebotomine sand flies (Diptera: Psychodidae) based on fossil evidence. Mem. Inst. Oswaldo Cruz 98 (Suppl. 1), 145–149.

Fraga, J., Montalvo, A.M., De Doncker, S., Dujardin, J.-C., Van der Auwera, G., 2010. Phylogeny of *Leishmania* species based on the heat-shock protein 70 gene. Infect. Genet. Evol. 10, 238–245.

Gadagkar, S.R., Rosenberg, M.S., Kumar, S., 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. J. Exp. Zool. B Mol. Dev. Evol. 304, 64–74.

Galati, E., 1995. Phylogenetic systematics of Phlebotominae (Diptera, Psychodidae) with emphasis on American groups. Bol. Dir. Malariol. San. Amb. 35, 133–142.

Hamilton, P.B., Teixeira, M.M., Stevens, J.R., 2012. The evolution of *Trypanosoma cruzi*: the 'bat seeding' hypothesis. Trends Parasitol. 28, 136–141.

Junier, T., Zdobnov, E.M., 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26, 1669–1670.

Kanjanopas, K., Siripattanapipong, S., Ninsaeng, U., Hitakarun, A., Jitkaew, S., Kaewtaphaya, P., Tan-ariya, P., Mungthin, M., Charoenwong, C., Leelayoova, S., 2013. *Sergentomyia* (*Neophlebotomus*) *gemmea*, a potential vector of *Leishmania siamensis* in southern Thailand. BMC Infect. Dis. 13, 333.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kerr, S.F., 2000. Palaearctic origin of *Leishmania*. Mem. Inst. Oswaldo Cruz 95, 75–80.

Kerr, S.F., 2006. Molecular trees of trypanosomes incongruent with fossil records of hosts. Mem. Inst. Oswaldo Cruz 101, 25–30.

Kerr, S.F., Merkelz, R., Mackinnon, C., 2000. Further support for a Palaearctic origin of *Leishmania*. Mem. Inst. Oswaldo Cruz 95, 579–581.

Kimura, M., 1968. Evolutionary rate at the molecular level. Nature 217, 624–626.

Kumar, S., Stecher, G., Peterson, D., Tamura, K., 2012. MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. Bioinformatics 28, 2685–2686.

Lainson, R., 1997. On *Leishmania enriettii* and other enigmatic *Leishmania* species of the Neotropics. Mem. Inst. Oswaldo Cruz 92, 377–387.

Lainson, R., 2010. The Neotropical *Leishmania* species: a brief historical review of their discovery, ecology and taxonomy. Rev. Pan-Amaz Saude 1, 13–32.

Lainson, R., Shaw, J.J., 1972. Leishmaniasis of the New World: taxonomic problems. Br. Med. Bull. 28, 44–48.

Lainson, R., Shaw, J.J., 1987. Evolution, classification, and geographical distribution. In: Peters, W., Killick-Kendrick, R. (Eds.), The Leishmaniases in Biology and Medicine. Academic Press, London, pp. 1–120.

Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Leelayoova, S., Siripattanapipong, S., Hitakarun, A., Kato, H., Tan-ariya, P., Siriyasatien, P., Osatakul, S., Mungthin, M., 2013. Multilocus characterization and phylogenetic analysis of *Leishmania siamensis* isolated from autochthonous visceral leishmaniasis cases, southern Thailand. BMC Microbiol. 13, 60.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Liautaud, B., Vignier, N., Miossec, C., Plumelle, Y., Kone, M., Delta, D., Ravel, C., Cabie, A., Desbois, N., 2015. First case of visceral leishmaniasis caused by *Leishmania martiniquensis*. Am.J.Trop. Med. Hyg. 92, 317–319.

Lobsiger, L., Muller, N., Schweizer, T., Frey, C.F., Wiederkehr, D., Zumkehr, B., Gottstein, B., 2010. An autochthonous case of cutaneous bovine leishmaniasis in Switzerland. Vet. Parasitol. 169, 408–414.
Lukes, J., Mauricio, I.L., Schönian, G., Dujardin, J.-C., Soteriadou, K., Dedet, J.-P., Kuhls, K., Tintaya, K.W.Q., Jirků, M., Chocholová, E., Haralambous, C., Pratlong, F., Oborniik, M., Horák, A., Ayala, F.J., Miles, M.A., 2007. Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. Proc. Natl. Acad. Sci. U. S. A. 104, 9375–9380.
Lysenko, A., 1971. Distribution of leishmaniasis in the Old World. Bull. World Health Organ. 44, 515–520.
Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55, 21–30.
Marcili, A., Speranca, M.A., da Costa, A.P., Madeira, M.D., Soares, H.S., Sanches, C.D., Acosta, I.D., Girotto, A., Minervino, A.H., Horta, M.C., Shaw, J.J., Gennari, S.M., 2014. Phylogenetic relationships of *Leishmania* species based on trypanosomatid barcode (SSU rDNA) and gGAPDH genes: taxonomic revision of *Leishmania* (*L.*) *infantum chagasi* in South America. Infect. Genet. Evol.
Momen, H., Cupolillo, E., 2000. Speculations on the origin and evolution of the genus *Leishmania*. Mem. Inst. Oswaldo Cruz 95, 583–588.
Mueller, R.L., Macey, J.R., Jaekel, M., Wake, D.B., Boore, J.L., 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. Proc. Natl. Acad. Sci. U. S. A. 101, 13820–13825.
Muller, N., Welle, M., Lobsiger, L., Stoffel, M.H., Boghenbor, K.K., Hilbe, M., Gottstein, B., Frey, C.F., Geyer, C., von Bomhard, W., 2009. Occurrence of *Leishmania* sp. in cutaneous lesions of horses in Central Europe. Vet. Parasitol. 166, 346–351.
Noyes, H.A., 1998. Implications of a Neotropical origin of the genus Leishmania. Mem. Inst. Oswaldo Cruz 93, 657–661.
Noyes, H.A., Arana, B.A., Chance, M.L., 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.
Noyes, H.A., Pratlong, F., Chance, M., Ellis, J., Lanotte, G., Dedet, J.P., 2002. A previously unclassified trypanosomatid responsible for human cutaneous lesions in Martinique (French West Indies) is the most divergent member of the genus *Leishmania* ss. Parasitology 124, 17–24.
Noyes, H.A., Morrison, D.A., Chance, M.L., Ellis, J.T., 2000. Evidence for a neotropical origin of *Leishmania*. Mem. Inst. Oswaldo Cruz 95, 575–578.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289–290.
Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M.A., 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, S.L., Faulconbridge, A., Jeffares, D., Depledge, D.P., Oyola, S.O., Hilley, J.D., Brito, L.O., Tosi, L.R., Barrell, B., Cruz, A.K., Mottram, J.C., Smith, D.F., Berriman, M., 2007. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nat. Genet. 39, 839–847.
Perkins, S.L., Martinsen, E.S., Falk, B.G., 2011. Do molecules matter more than morphology? Promises and pitfalls in parasites. Parasitology 138, 1664–1674.
Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T., Manuel, M., Worheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9, e1000602.
Poinar Jr., G., 2007. Early Cretaceous trypanosomatids associated with fossil sand fly larvae in *Burmese amber*. Mem. Inst. Oswaldo Cruz 102, 635–637.
Pothirat, T., Tantiworawit, A., Chaiwarith, R., Jariyapan, N., Wannasan, A., Siriyasatien, P., Supparatpinyo, K., Bates, M.D., Kwakye-Nuako, G., Bates, P.A., 2014. First isolation of *Leishmania* from Northern Thailand: case report, identification as *Leishmania martiniquensis* and phylogenetic position within the *Leishmania enriettii* complex. PLoS Negl. Trop. Dis. 8, e3339.
R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6.
Reuss, S.M., Dunbar, M.D., Calderwood Mays, M.B., Owen, J.L., Mallicote, M.F., Archer, L.L., Wellehan Jr., J.F., 2012. Autochthonous *Leishmania siamensis* in horse, Florida, USA. Emerg. Infect. Dis. 18, 1545–1547.
Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3, 217–223.
Rose, K., Curtis, J., Baldwin, T., Mathis, A., Kumar, B., Sakthianandeswaren, A., Spurck, T., Low Choy, J., Handman, E., 2004. Cutaneous leishmaniasis in red kangaroos: isolation and characterisation of the causative organisms. Int. J. Parasitol. 655–664 (2004/04/28 ed).
Saf'janova, V., 1982. The problem of taxonomy with *Leishmania*. Ser. Protozool. Sov. Acad Sci. Leningr. 7, 5–109.
Schönian, G., Cupolillo, E., Mauricio, I., Ponte-Sucre, A.E.A., 2013. Molecular evolution and phylogeny of *Leishmania*. Drug Resistance in *Leishmania* Parasites. Springer-Verlag, Wien, pp. 15–40.
Schönian, G., Mauricio, I., Cupolillo, E., 2010. Is it time to revise the nomenclature of *Leishmania*? Trends Parasitol. 26, 466–469.
Schwartz, R.S., Harkins, K.M., Stone, A.C., Cartwright, R.A., 2015. A composite genome approach to identify phylogenetically informative data from next-generation sequencing. BMC Bioinformatics 16, 193.
Shaw, J., 1997. Ecological and evolutionary pressures on leishmanial parasites. Braz. J. Genet. 20.
Shaw, J.J., 2006. Further thoughts on the use of the name *Leishmania* (*Leishmania*) *infantum chagasi* for the aetiological agent of American visceral leishmaniasis. Mem. Inst. Oswaldo Cruz 101, 577–579.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stevens, J., Rambaut, A., 2001. Evolutionary rate differences in trypanosomes. Infect. Genet. Evol. 1, 143–150.
Stevens, J.R., Gibson, W., 1999. The molecular evolution of trypanosomes. Parasitol. Today 15, 432–437.
Stevens, J.R., Noyes, H.A., Dover, G.A., Gibson, W.C., 1999. The ancient and divergent origins of the human pathogenic trypanosomes, *Trypanosoma brucei* and *T. cruzi*. Parasitology 118 (Pt 1), 107–116.
Stevens, J.R., Noyes, H.A., Schofield, C.J., Gibson, W., 2001. The molecular evolution of Trypanosomatidae. Adv. Parasitol. 48, 1–56.
Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., Kumar, S., 2012. Estimating divergence times in large molecular phylogenies. Proc. Natl. Acad. Sci. U. S. A. 109, 19333–19338.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Tschoeke, D.A., Nunes, G.L., Jardim, R., Lima, J., Dumaresq, A.S.R., Gomes, M.R., de Mattos Pereira, L., Loureiro, D.R., Stoco, P.H., de Matos Guedes, H.L., de Miranda, A.B., Ruiz, J., Pitaluga, A., Silva, F.P., Probst, C.M., Dickens, N.J., Mottram, J.C., Grisard, E.C., Dávila, A.M.R., 2014. The comparative genomics and phylogenomics of *Leishmania amazonensis* parasite. Evol. Bioinformatics Online 10, 131–153.
Tuon, F.F., Amato Neto, V., Sabbaga Amato, V., 2008. Leishmania: origin, evolution and future since the Precambrian. FEMS Immunol. Med. Microbiol. 54, 158–166.
Van der Auwera, G., Fraga, J., Montalvo, A.M., Dujardin, J.-C., 2011. Leishmania taxonomy up for promotion? Trends Parasitol. 27, 49–50.
Waki, K., Dutta, S., Ray, D., Kolli, B.K., Akman, L., Kawazu, S., Lin, C.P., Chang, K.P., 2007. Transmembrane molecules for phylogenetic analyses of pathogenic protists: *Leishmania*-specific informative sites in hydrophilic loops of trans-endoplasmic reticulum N-acetylglucosamine-1-phosphate transferase. Eukaryot. Cell 6, 198–210.
Wirth, D.F., Pratt, D.M., 1982. Rapid identification of *Leishmania* species by specific hybridization of kinetoplast DNA in cutaneous lesions. Proc. Natl. Acad. Sci. U. S. A. 79, 6999–7003.
Yoder, J.B., Briskine, R., Mudge, J., Farmer, A., Paape, T., Steele, K., Weiblen, G.D., Bharti, A.K., Zhou, P., May, G.D., Young, N.D., Tiffin, P., 2013. Phylogenetic signal variation in the genomes of *Medicago* (Fabaceae). Syst. Biol. 62, 424–438.
Yurchenko, V.Y., Lukes, J., Jirku, M., Zeledon, R., Maslov, D.A., 2006. *Leptomonas costaricensis* sp. n. (Kinetoplastea: Trypanosomatidae), a member of the novel phylogenetic group of insect trypanosomatids closely related to the genus *Leishmania*. Parasitology 133, 537–546.