J. Eukaryot. Microbiol., 51(3), 2004 pp. 283–290

Diversity of Insect Trypanosomatids Assessed from the Spliced Leader RNA and 5S rRNA Genes and Intergenic Regions[1]

SERGEI A. PODLIPAEV,[a,†] NANCY R. STURM,[b] IVAN FIALA,[a,c] OCTAVIO FERNANDES,[d] SCOTT J. WESTENBERGER,[b] MICHEL DOLLET,[e] DAVID A. CAMPBELL[b] and JULIUS LUKEŠ[a,c]

[a]*Institute of Parasitology, Czech Academy of Sciences, České Budějovice, Czech Republic, and*
[b]*Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, USA, and*
[c]*Faculty of Biology, University of South Bohemia, České Budějovice, Czech Republic, and*
[d]*Department of Tropical Medicine, Instituto Oswaldo Cruz, Rio de Janeiro, Brazil, and*
[e]*CIRAD, Campus International de Baillarguet, Montpellier, France*

ABSTRACT. We have determined the sequences of 5S rRNA and spliced leader (SL) RNA genes, and adjacent intergenic regions for representatives of all known trypanosomatid genera parasitizing insects. The genetic loci have been analyzed separately as well as by a combined approach. Several isolates, assigned by morphology to different genera (*Leptomonas* spp., *Blastocrithidia* spp.), seem to belong to a single species with an unexpectedly wide host and geographical range. An unnamed trypanosomatid isolated from rats in Egypt was found to belong to the genus *Herpetomonas*, so far associated with insect hosts only. It is closely related to *Herpetomonas ztiplika*, a parasite of a blood-sucking biting midge. Apparently several different trypanosomatid species can infect one insect species, as exemplified by *Leptomonas* sp. PL and *Wallaceina* sp. Wsd, which were isolated from different specimens of *Salda littoralis* on the same locality and day. However, since the same species of *Leptomonas* was obtained from insect hosts belonging to different genera, some insect trypanosomatids may have low host specificity. Our data revealed additional discrepancies between molecular phylogenetic data and cell morphology, rendering current trypanosomatid taxonomy unreliable.

Key Words. 5S rRNA, SL RNA, evolution, Kinetoplastida, phylogeny, Trypanosomatidae.

KINETOPLASTID flagellates are well-known for their capability to colonize eukaryotes ranging from protozoans to humans (Donelson, Gardner, and El-Sayed 1999; Vickerman 2000). The great medical and socioeconomic impact of these causative agents of severe diseases has stimulated dynamic research on trypanosomes and leishmanias of vertebrates, while the insect trypanosomatids have been studied much less. So far, about two hundred trypanosomatid species have been described from an estimated 2,000 to 2,500 insect species (Podlipaev 2000). Since there are more than a million insect species, only a tiny fraction of the predicted multitude of flagellate species has been sampled (Stevens 2001). Such an under-sampling provides only a very limited view of the diversity of insect trypanosomatids, which may be one of the most diverse groups within the Kinetoplastida (Podlipaev 2001). To date, a total of about 20 insect trypanosomatids have been included in molecular phylogenies. The addition of ten new species in the 18S ribosomal RNA (rRNA)-based trees resulted in the appearance of three new lineages (Merzlyak et al. 2001b), indicating that the biodiversity described so far may represent the tip of an iceberg.

Of the 10 currently valid trypanosomatid genera, six comprise monoxenous parasites of insects. For the remaining four genera, insect hosts serve as vectors (Vickerman 2000). The greater part of trypanosomatid biodiversity may thus be hidden in insects. Placement of the genus *Leishmania* among lineages of insect flagellates indicates that as putative ancestral parasites of insects, leishmanias colonized vertebrates, most likely during the process of blood feeding (Maslov et al. 1996; Podlipaev 2000). Such a phylogenetic position qualifies the closely-related and easy-to-cultivate insect trypanosomatids as models in the search for novel anti-leishmanial drugs. It seems a plausible hypothesis that the ancestors of phytomonads, responsible for economically important plant diseases, were of insect origin as well, since in the 18S rRNA-based trees they branch together with the insect trypanosomatids. Moreover, indirect evidence points to hemipterans, especially pentatomids, as the vectors of phytoflagellates (Camargo 1999; Dollet 2001).

For more than one hundred years (see Molyneux and Ashford 1983; Wallace 1966), taxonomy of the kinetoplastid suborder Trypanosomatina remains based on morphology observable in the light microscope and host origin of the isolate. Features, such as relative position of the nucleus and kinetoplast, cell shape and dimensions, as well as the position of the flagellum, are critical for assigning an isolate to a species and a genus. There is no doubt that the morphology-based system has proved very valuable and flexible enough to accommodate new observations. However, since a growing number of cultures from new hosts and locations has become available, and molecular methods are now being increasingly used, sequence information obtained thus far from the insect flagellates suggests that a major taxonomic revision is necessary. Phylogenetic trees do not match those using a handful of morphological criteria, and many contradictions have been identified among groups defined by sequence- and morphology-based systems. For example, in the 18S rRNA trees the symbiont-bearing members of the genera *Herpetomonas* and *Crithidia* comprise a well-supported clade, while the endosymbiont-free members of both genera are found elsewhere in the tree (Hollar, Lukeš, and Maslov 1998). At least four currently valid genera of insect trypanosomatids, namely *Crithidia*, *Blastocrithidia*, *Herpetomonas*, and *Leptomonas*, have to be considered para- or polyphyletic at present (Hollar, Lukeš, and Maslov 1998; Hughes and Piontkivska 2003; Merzlyak et al. 2001b).

''Classical'' taxonomy remains suitable to systematize the monomorphic members of the genus *Trypanosoma*, the morphology of which is always a variation on the trypomastigote form. However, morphology is insufficient to classify the trypanosomatids of insects. These flagellates, previously coined ''lower trypanosomatids'', represent a morphologically heterogeneous assembly currently ranked into six genera (i.e. *Crithidia*, *Blastocrithidia*, *Leptomonas*, *Herpetomonas*, *Rhynchoidomonas* and *Wallaceina*), which occur as pro-, choano-, opisto-, epi- and endomastigote morphotypes (Vickerman 2000). The use of these formally well-defined morphotypes has a number of limits: 1) in the host, most insect flagellates have been observed in several morphotypes with promastigotes common to almost all genera including *Phytomonas* from plants; 2) a

Corresponding Author: J. Lukeš—Telephone number: 00420-38-7775416; Fax number: 00420-38-5310388; E-mail: jula@paru.cas.cz

†Deceased. At the time of his death, Sergei Podlipaev's address: Zoological Institute, Russian Academy of Sciences, St. Petersburg, Russia

[1] The nucleotide sequences of the 5S rRNA genes have been deposited in the GenBank™ under accession numbers AY547471-AY547491, and those of the SL genes have been deposited under accession numbers AY547459-AY547470.

continuum of cell types exists, making the assignment to a specific stage highly subjective; 3) some stages (e.g. opistomastigotes) are rare in natural populations and sometimes can be observed in culture only under specific conditions.

The phylogeny of Kinetoplastida has been extensively reconstructed using both rRNA and protein-coding genes (Hannaert, Opperdoes, and Michels 1998; Lukeš et al. 1997; Simpson, Lukeš, and Roger 2002; Simpson and Roger 2004). The 5S rRNA and spliced leader (SL) RNA genes were shown to be useful to study the genetic diversity of insect and plant flagellates at and below the generic level (Campbell, Fernandes, and Sturm 1997; Dollet et al. 2000; Fernandes et al. 1997). So far, only a limited number of insect trypanosomatids was included in the 5S and SL RNA studies, whereas the most comprehensive collection of isolates from these hosts appeared in the 18S rRNA-based trees (Merzlyak et al. 2001).

In the present work we have used information contained in the sequences of both 5S and SL RNA genes and the neighboring intergenic regions to address issues, such as diversity, geographic distribution, and host specificity of insect trypanosomatids, as well as possible co-evolution with their hosts. For this, we have had at our disposition a representative set of insect trypanosomatids introduced into culture by us and others or obtained from the American Type Culture Collection (ATCC). Our interests were focused particularly on a unique set of isolates obtained in the time span of almost 20 years in Russia.

MATERIALS AND METHODS

Strain origin and cultivation conditions. The geographic origin, date of isolation, and host species are shown in Table 1. Strains of Russian origin were isolated and introduced into culture in the Podlipaev laboratory in St. Petersburg and cultivated in the BHI medium supplemented with hemin (10 mg/ml) at 26 °C. After axenization (several initial passages of primocultures were made with contaminating bacteria), the cultures were examined by light microscopy. All isolates collected in Russia are stored in the cryobank of the Institute of Parasitology, České Budějovice, Czech Republic.

Strains and DNA samples of non-Russian origin were obtained from a variety of sources. *Blastocrithidia culicis* and *Leptomonas seymouri* were obtained from ATCC and cultivated at the University of California, Los Angeles (UCLA) and Centre de Coóperation International en Récherche Agronomique pour le Développment, Montpellier (CIRAD), respectively. *Blastocrithidia leptocoridis* and *Blastocrithidia triatomae* were cultured at the Oswaldo Cruz Institute, Rio de Janeiro, Brazil. DNA samples of *Crithidia* spp. have been described previously (Fernandes et al. 1997). DNA samples of *Herpetomonas mariadeanei*, *Herpetomonas megaseliae*, and *Herpetomonas muscarum* (Landweber and Gilbert 1993) were gifts from Dr. Laura Landweber (Princeton University, USA).

DNA isolation, PCR and sequencing. DNA was isolated from exponentially growing cultures, and the 5S and SL RNA genes were amplified from about 10 ng of genomic DNA following the protocols described elsewhere (Dollet et al. 2000; Fernandes et al. 1997). The 5S rRNA gene repeat was amplified using primers 5S-L (5′-CCGTCCGATTTGTGAAGTTAAGC-3′) and 5S-R (5′-TAACTTCACAAATCGGACGGGAT-3′). The SL gene and the neighboring intergenic region were amplified using primers ME-1 (5′-TTCTGTACTTTATTGGTA-3′) and ME-2 (5′-CAATAAAGTACAGAAACTG-3′). The amplicon was isolated from 0.75% agarose gels, cloned into the pCR2.1 TOPO TA vector (Invitrogen, Carlsbad, CA), and both strands were sequenced on an automated sequencer CEQ™ 2000 using the CEQ™ DTCS sequencing kit (Beckman Coulter, Fullerton, CA). Two or more clones of 5S rRNA gene have been sequenced from eight isolates (Table 2).

Phylogenetic analysis. Using the Clustal-X program (Thomson et al. 1997), the 5S rRNA gene alone (alignment I) or along with the intergenic region (group-alignments II) were aligned with homologues from all available insect trypanosomatids, as well as representatives of the major clades within the Trypanosomatidae. Alignments of the SL RNA gene (alignment III) and SL RNA gene + intergenic region (alignment IV) were prepared in the same way. Alignments were refined by eye using the BioEdit sequence alignment editor (Hall 1999), and 5S and SL RNA genes were concatenated in BioEdit (alignment V). Two trypanosomes (*Trypanosoma brucei* and *Trypanosoma cruzi*) were used as outgroups. On the basis of alignments I, III, and V, trees were constructed using maximum parsimony (MP), maximum likelihood (ML), and distance (minimum evolution)(ME) methods, carried out using the program package PAUP*, version 4.0b8 (Swofford 2001). The MP analysis was done using heuristic search with random addition of taxa (10 replications) and gaps were treated as missing data. Three transversions/transitions (Tv/Ts) ratios were tested (1:2, 1:3 and 1:4). For the ML analysis, the likelihood ratio test (LRT) implemented in the Modeltest, version 3.06 (Posada and Crandall 1998) was used to determine the best model of evolution. ML was performed with the GTR+G model for alignment I, the JC69+G model for alignment III, and the GTR+G model for the concatenated alignment V. The distance method was executed using heuristic search with ME as the objective setting and the K2P substitution model. Clade support was assessed with bootstrapping (500 bootstrap replicates). Alignments are at (http://www.paru.cas.cz/alignments/) or are available upon request.

The following sequences retrieved from the GenBank™ were included in the analysis: 1) 5S: *Phytomonas* sp. Mang (AF243353); *Phytomonas* sp. Bergbr (AF243337); *Phytomonas* sp. Manibr (AF243354); *Phytomonas* sp. Ehisu (AF243342); *Phytomonas* sp. Hart1 (AF243345); *Phytomonas* sp. Ehigu (AF243339), *Phytomonas* sp. Ascl (AF243336) and *Phytomonas* sp. EM1 (AF243338); *Trypanosoma brucei* (M14817); *Trypanosoma cruzi* (AF545104); 2) SL: *Crithidia fasciculata* (U96170); *Crithidia acanthocephali* (U96167); *Crithidia luciliae thermophila* (U96171); *Endotrypanum schaudinni* (L05398); *Leishmania tarentolae* (AY100198); *Leptomonas collosoma* (S76723); *Phytomonas* sp. EM1 (X87138); *Phytomonas* sp. Cbe (L42377); *Phytomonas staheli* (AF121766); *Phytomonas* sp. Hart1 (AF250969); *Crithidia oncopelti* (U96172); *Trypanosoma brucei* (K02630) and *Trypanosoma cruzi* (K02632).

RESULTS

5S rRNA gene. A single repeat of the tandemly arrayed 5S rRNA gene was amplified using the protocol of Dollet et al. (2000). Size of the amplicon ranged from 0.25 kb in *Leptomonas* sp. P up to 1.6 kb in *Herpetomonas ztiplika*. The tandem arrangement of the 5S rRNA genes resulted in the amplification of two or three bands in about 20% of cases. These amplicons represented monomers, dimers, and trimers of the gene and the adjacent intergenic region (data not shown). In all cases, the sequenced region was composed of the complete 5S rRNA gene and the adjacent intergenic region. In total, the 5S rRNA gene was amplified, cloned, and sequenced from parasites originating from Northern Russia (10 isolates), the Czech Republic, and Egypt (1 isolate each), 3 isolates introduced into culture in Brazil, and 7 strains received from the ATCC (see Table 1). For most Russian isolates, the two 5S rRNA clones sequenced were either identical or differed in one to three positions (Table 2).

Table **1**. Description of insect trypanosomatid strains used in this study. Accession numbers of sequences obtained from the GenBank™ are included. (+) indicates that the gene was sequenced in frame of this study; (−) indicates that sequence of the gene is not available.

Species/Isolate name	Host	Place and time of isolation (source of culture)	5S rRNA	Spliced leader RNA
Blastocrithidia culicis (Novy et al., 1907)	*Aedes vexans* Diptera: Culicidae	ATCC 30268	+	+
Blastocrithidia gerricola Podlipaev, 1985	*Gerris lacustris* Hemiptera: Gerridae	Kaliningrad region, West Russia, 1981	−	+
Blastocrithidia leptocoridis (McCulloch, 1915)	*Leptocoris* (*Boisea*) *trivittatus* Hemiptera: Rhopallidae	ATCC 30265	+	+
Blastocrithidia miridarum Podlipaev & Frolov, 1987	*Leptopterna dolabrata* Hemiptera: Miridae	Pskov region, North-West Russia, 1987	+	+
Blastocrithidia triatoma Cerisola et al., 1971	*Triatoma infestans* Hemiptera: Reduviidae	Brazil	+	+
Crithidia acanthocephali Hanson & McGhee, 1961	*Acanthocephala femorata* Hemiptera: Coreidae	ATCC 30251	+	U96167
Crithidia fasciculata Leger, 1902	*Culex pipiens* Diptera: Culicidae	ATCC 12857	M28975	U96170
Crithidia oncopelti (Noguchi & Tilden, 1926)	*Oncopeltus fasciatus* Hemiptera: Lygaeidae	ATCC 12982	+	U96172
Crithidia luciliae (Strickland, 1911)	*Phaenicia sericata* Diptera: Calliphoridae	TCPD, USP	+	U96171 (ssp. *thermophila*)
Leptomonas collosoma Wallace et al., 1960	*Gerris dissortis* Hemiptera: Gerridae	ATCC 30261	−	S76723
Leptomonas seymouri Wallace, 1977	*Dysdercus suturellus* Hemiptera: Pyrrhocoridae	ATCC 30220	+	X07487
Leptomonas repentinus Malysheva & Frolov, 2002	*Gerris rufoscutellatus* Hemiptera: Gerridae	Leningrad region, North-West Russia, 2000	+	−
Leptomonas sp. F6	*Nabis* (*Nabicula*) *flavomarginatus* Hemiptera: Nabidae	Cape Kartesh, Tshupa bay, White Sea, North Russia, 1986	+	+
Leptomonas sp. F5	*Nabis* (*Nabicula*) *flavomarginatus* Hemiptera: Nabidae	Cape Kartesh, Tshupa bay, White Sea, North Russia, 1986	−	+
Leptomonas sp. PL	*Salda littoralis* Hemiptera: Saldidae	Cape Kartesh, Tshupa bay, White Sea, North Russia, 2001	+	−
Leptomonas sp. Sld	*Saldula pallipes* Hemiptera: Saldidae	Cape Kartesh, Tshupa bay, White Sea, North Russia, 2000	+	−
Leptomonas sp. C4	*Gerris rufoscutellatus* Hemiptera: Gerridae	Leningrad region, North-West Russia, 1987	+	+
Leptomonas sp. P	*Panorpa communis* Mecoptera: Panorpidae	Leningrad region, North-West Russia, 1988	+	+
Leptomonas sp. Cfm	*Nabis (Nabicula*) *flavomarginatus* Hemiptera: Nabidae	Pskov region, North-West Russia, 1983	+	+
Leptomonas sp. Nfm	*Nabis (Nabicula*) *flavomarginatus* Hemiptera: Nabidae	Pskov region, North-West Russia, 1991	+	+
trypanosomatid ''X'' Morsy et al., 1988	*Rattus norvegicus* Rodentia: Muridae	Alexandria, Egypt, 1983	+	+
Herpetomonas mariadeanei Yoshida et al., 1978	*Muscina stabulans* Diptera: Muscidae	ATCC 30708	+	+
Herpetomonas megaseliae Daggett et al., 1972	*Megaselia scalaris* Diptera: Phoridae	ATCC 30209	+	−
Herpetomonas muscarum (Leidy, 1856)	*Musca domestica* Diptera: Muscidae	ATCC 30260	+	−
Herpetomonas pessoai (Galvao et al., 1970)	*Zelus leucogrammus* Hemiptera: Reduviidae	ATCC 30252	X62331	X62331
Herpetomonas ztiplika Podlipaev et al., 2004	*Culicoides kibunensis* Diptera: Ceratopogonidae	Southern Moravia, Czech Republic, 1999	+	−
Wallaceina sp. Wsd	*Salda littoralis* Hemiptera: Saldidae	Cape Kartesh, Tshupa bay, White Sea, North Russia, 2001	+	−

TCPD, USP = Trypanosomatid Collection of the Parasitology Department, University of Sao Paulo, Brazil
ATCC = American Type Culture Collection, Manassas, USA

Table **2**. Number of 5S rRNA clones sequenced for each strain, with variability in the obtained sequences indicated.

	Nb. clones	5S rRNA gene sequence								
		30	58	152	186	227	240	248	280	289
Leptomonas sp. F6	2	—	—	—	A/G	—	—	—	G/C	—
Leptomonas sp. C4	2	—	—	—	—	—	—	—	—	—
Leptomonas sp. PL.	2	T/G	—	—	—	—	—	—	—	—
Wallaceina sp. Wsd.	2	—	—	A/T	—	—	—	—	—	C/T
Leptomonas sp. Cfm	2	—	—	—	—	—	A/G	C/G	T/C	—
Leptomonas repentinus	2	—	C/G	—	—	C/A	—	—	—	—
Blastocrithidia miridarum BM	2	—	—	—	—	—	—	—	—	—
Herpetomonas ztiplika	4	—	—	—	—	—	—	—	—	—

Even when using widely different PCR conditions, we were unable to amplify the 5S rRNA gene from four isolates.

Sequence comparison of the highly conserved 5S rRNA gene (alignment I; 138 characters) excluded the variable 3′ nucleotides from the analyses (13 characters). The total number of parsimony informative positions was 49. The purpose of this alignment, which contained 34 ingroup trypanosomatids and two outgroup *Trypanosoma* spp., was to map the diversity of the insect-derived species reflected in their 5S rRNA gene. While all available insect trypanosomatids have been included, we have aligned the 5S rRNA gene of a single representative of every phytomonad subgroup (Dollet, Sturm, and Campbell 2001).

Analysis of alignment I by MP and ML generated trees with generally low bootstrap values (Fig. 1A), an expected result considering the low number of informative positions. In all analyses, early branching of *Trypanosoma* spp. was strongly supported (100/99%), with *Phytomonas* spp. emerging as the earliest branches within the ingroup. With the exception of *Leptomonas* sp. Nfm, rRNA-based phylogeny revealed a high homogeneity of the North-Russian isolates (*Leptomonas* spp. F6, Sld, C4 and PL; *Leptomonas repentinus*, and *Blastocrithidia miridarum*) derived from different insect hosts and isolated over a 20-yr. time span from geographically distant localities (Table 1). The clade composed solely of these flagellates (Fig. 1A) is characterized by the GG dinucleotide in positions 18 and 19. The crown trypanosomatids are composed of this ''North-Russian'' clade and a diverse assembly of leishmanias, crithidias and leptomontads (*Leishmania tarentolae*, *Leishmania major*, *Leptomonas seymouri*, *Leptomonas* spp. P and Cfm, *Crithidia luciliae*, *Crithidia oncopelti*, *Crithidia fasciculata* and *Crithidia acanthocephali*) that form a polytomy. In alignment I the following diagnostic positions were identified for the crown trypanosomatids: A in position 2, G in position 19, and C in position 60.

A long branch leads to a clade supported by 99% bootstrap, composed of four symbiont-free *Herpetomonas* spp., one isolate of *Leptomonas* sp. Nfm, and a flagellate isolated from vertebrates—trypanosomatid ''X'' (Table 1, Fig. 1A). The following positions were characteristic for this group: A in position 1, TT in positions 8 + 9, and T in position 64 (alignment I). Another branch brings together two symbiont-free and one symbiont-bearing *Blastocrithidia* species. The relationship of *Wallaceina*

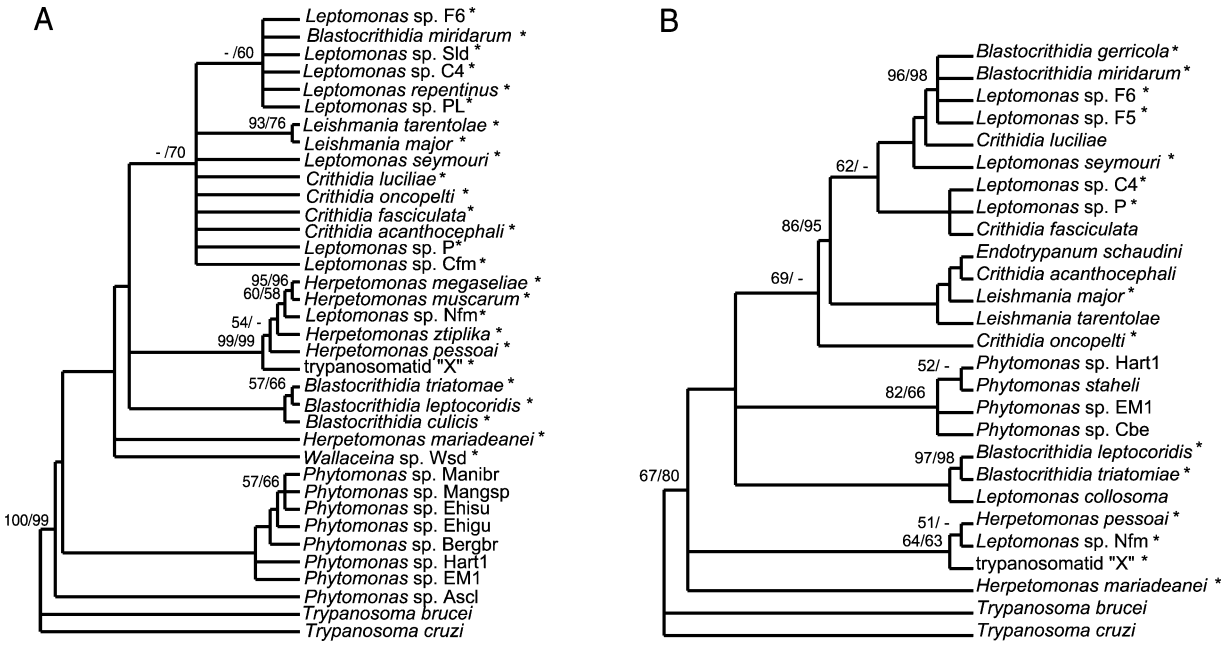

Fig. **1.** Maximum likelihood trees of the 5S rRNA gene (Fig. 1A) and spliced leader (SL) RNA gene (Fig. 1B) of trypanosomatids rooted on *Trypanosoma* spp. A strict consensus of 21 trees (− ln = 883.5840) (Fig. 1A) and 4 trees (− ln = 970.2145) (Fig. 1B) is shown. Bootstrap values (ML and MP Tv/Ts = 1:2) are indicated for the nodes gaining more than 50% support. Sequences obtained in this study are indicated with an asterisk.

sp. Wsd and *H. mariadeanei* with the other insect trypanosomatids was not resolved in the 5S RNA tree (Fig. 1A).

5S rRNA intergenic region. Cautious about the low information content in the 5S rRNA gene, we explored the potential for phylogenetic analysis of the flanking intergenic region. Since it proved impossible to create a single reliable alignment from this highly variable region, only partial alignments were analyzed (group-alignment II). This genetic locus provided further support for the existence of a highly homogeneous crown group of primarily ''North-Russian'' flagellates that differ significantly in their biology and host species (see Discussion). However, these differences are reflected neither in their 5S rRNA gene, nor in the intergenic region that contains only point mutations (group-alignments II). Multiple differences in this hypervariable region among members of the herpetomonad clade suggest radiation within this group that is not visible on the level of the 5S rRNA gene. Interestingly, the intergenic region was almost identical in *H. muscarum* and *H. megaseliae* (95% similarity). Sequence similarity in the same locus indicated that *H. pessoai* and trypanosomatid ''X'' represent a pair of closely related species (71% similarity).

SL RNA gene and intergenic region. We amplified the SL RNA gene and the flanking intergenic region for a number of new isolates (Table 1). Since we decided to sequence only one clone for each isolate, to address the issue of possible sequence heterogeneity of this multiple-copy gene in one isolate, an *in silico* analysis of multiple samplings from the SL RNA gene arrays of two model flagellates has been performed. Fifty-one and 88 SL RNA genes from the *L. major* and *T. cruzi* databases, respectively, were aligned and subjected to a phylogenetic analysis. Despite limited sequence differences, all SL RNA genes from one organism clustered with the control SL RNA sequence and were clearly distinguishable from other species (data not shown; but see http://www.paru.cas.cz/alignments/).

Numerous attempts to amplify the target gene from the remaining isolates failed for reasons that are not understood (e.g. the trypanosomatid ''X'' dual 5S-SL RNA unit was amplified by the 5S rRNA primers but not the SL RNA primers, even though the SL oligonucleotide-binding site was present). The amplicon was usually about 0.9 kb long, *Leptomonas* sp. P being an exception (see below). In all studied isolates the SL RNA exon had canonical length of 39 nucleotides and was identical with that of *L. tarentolae* (AY100198). Alignment III, which is composed of the SL RNA exon and intron, contains 52 parsimony-informative positions in a dataset of 27 species representing all trypanosomatid genera. The length of the predicted SL RNA gene varied between 85 and 101 nucleotides.

A well-supported (86/95% bootstrap) crown group of the SL RNA tree (Fig. 1B), where *T. brucei* and *T. cruzi* were considered outgroups, represented a diverse assembly composed of five *Leptomonas* spp., three *Crithidia* spp., two *Blastocrithidia* spp., *Leishmania major* and *L. tarentolae* and *Endotrypanum schaudinni* (Fig. 1B). *Crithidia oncopelti* was basal to the crown group, that was characterized by nucleotides A and T in positions 15 and 22, respectively (see alignment III). Internal branching of this large clade was poorly supported by all methods (ML, MP and ME), so the strict consensus tree failed to provide robust information about relationships of members of the six genera with the exception of the close relatedness of *Blastocrithidia gerricola* and *B. miridarum*, and *Leptomonas* spp. F5 and F6 (Fig. 1B). ME analysis returned two strict consensus trees virtually identical with the MP and ML majority rule consensus trees (data not shown).

As other members of the trichotomy, *Phytomonas* spp. appear as a monophyletic group, and *B. leptocoridis* and *B. triatominae* along with *L. collosoma* constitute a poorly supported clade (Fig. 1B). The remaining insect trypanosomatids were split into two branches of unresolved mutual relationship. *Herpetomonas pessoai* and trypanosomatid ''X'' formed a clade with *Leptomonas* sp. Nfm. *Herpetomonas mariadeanei* constituted the earliest branch of the ''non-*Trypanosoma*'' trypanosomatids, the monophyly of which was supported by 67/80% bootstrap.

The SL intron and the intergenic region in particular are highly variable segments of the SL RNA repeat, which qualifies them as regions suitable for the comparison of isolates that appear, on the basis of their 5S rRNA sequences, to be very closely related. Differences in the intron were primarily confined to variation within nucleotide stretches (e.g. 8 to 10 Ts), and to more-or-less randomly distributed point mutations. Leptomonads F5 and F6, *B. miridarum* and *B. gerricola* were virtually identical, in spite of the fact that they were isolated from unrelated hosts obtained from distant geographic localities in Russia (alignment IV). The SL intergenic region was the most polymorphic locus analyzed. It was impossible to construct an unambiguous alignment even for isolates that were closely related, as judged from the other sequenced regions (data not shown).

Concatenated analysis. Since the results of a single-gene approach are often biased due to the low number of informative positions, we have constructed a concatenated alignment of the 5S and SL RNA genes (alignment V), in which the number of parsimony-informative positions rose to 91. This reduced the number of species included, since sequences of both genes are available only for 17 insect trypanosomatids (Fig. 2). Still, the concatenated approach proved informative and confirmed several trends observed in the single-gene analyses and provided higher bootstrap support at multiple nodes.

Trypanosomatids were clearly monophyletic (100% bootstrap support in ML, MP and ME trees) with phytomonads constituting the earliest lineage. The branching order of *H. mariadeanei*, the long branch leading to *H. pessoai*, trypanosomatid ''X'' and *Leptomonas* sp. Nfm, and the *B. triatomiae* + *B. leptocoridis* branch were not resolved, but both latter branches were 100% supported by all methods. The large crown group was more robust than in the previous analyses, while internal branching within this group was resolved to a lower degree (Fig. 2). Further support for this branching order was provided by an analysis, in which three times higher weight was assigned to the highly conserved SL RNA exon when compared to the less conserved SL RNA intron (data not shown).

DISCUSSION

Within the last decade, the kinetoplastid parasites have reached prominence in molecular biology as one of the most studied groups of protists. Their phylogeny also received substantial attention, resulting in more than a hundred 18S rRNA sequences, dozens for the genus *Trypanosoma* alone. Early studies demonstrated that the 18S rRNA gene was not informative for phylogenetic analyses of closely related flagellates (Uliana et al. 1994), while genes like SL RNA and 5S rRNA seem to be more appropriate for studies of their relationships at and below the genus level (Campbell, Fernandes, and Sturm 1997; Dollet et al. 2000). Herein, we undertook an analysis of representatives of all known trypanosomatid genera parasitizing insects, using the SL RNA and 5S rRNA genes, which can be genetically linked (Drouin and de Sá 1995).

The number of discrepancies between the molecular and morphological taxonomies of these flagellates is growing, rendering five out of six genera polyphyletic in recent 18S rRNA trees (Hollar, Lukeš, and Maslov 1998; Hughes and Piontkivska 2003; Merzlyak et al. 2001b). Moreover, the almost 100-yr-old

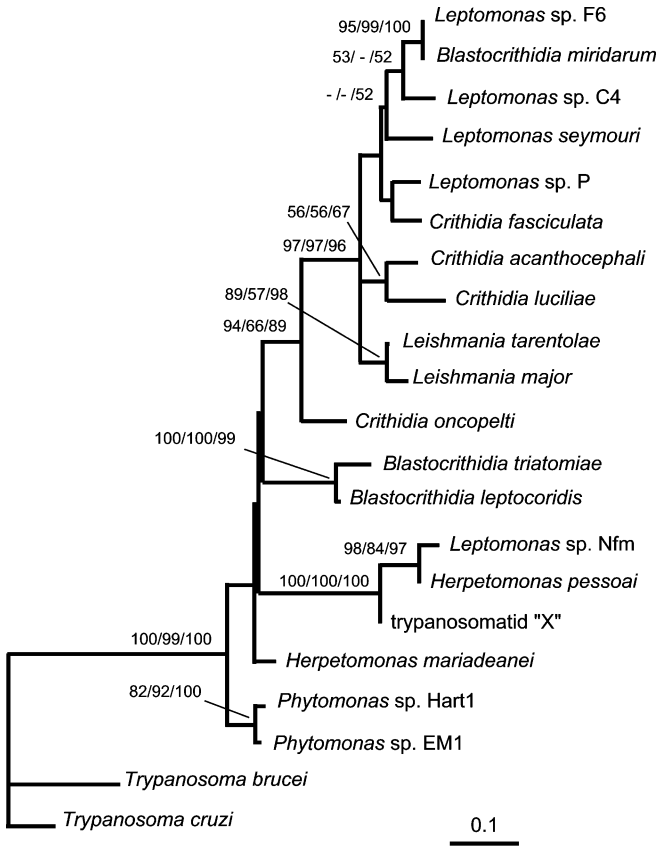

Fig. **2.** Maximum likelihood tree (− ln = 1502.6881) constructed with the concatenated sequences of the 5S rRNA and SL RNA genes of all insect trypanosomatids for which both sequences are available. The tree was rooted on *Trypanosoma* spp., bootstrap values (ML and MP Tv/Ts = 1:2) are indicated for the nodes gaining more than 50% support. The distance scale is given under the tree.

definitions of kinetoplastid genera are increasingly unsuitable, resulting in a situation, in which it is easier to describe a new species than to assign it to a genus (Dollet 2001). The problem is worsened by morphological variability of insect trypanosomatids, their low host specificity, and the high probability of mixed infections. Therefore, it seems that an investigation involving morphological description alone is now insufficient, and a study of cultured forms using molecular biology methods is needed to properly address the question ''What parasite was isolated?'' Herein, the individual genetic loci have been analyzed separately as well as by a combined approach. Although the trees were generally not highly supported, we were able to identify distinct clades within the insect flagellates and answer some questions that could not have been addressed previously.

The diversity of SL RNA sequences seems to be confined to the lower part of the tree with the crown group harboring a virtually identical version. With the exception of *C. oncopelti*, this group contains leptomonads and blastocrithidias that differ in cell morphology and were all isolated from hemipteran hosts. Based on their origin in northern Russia, they belong to a postglacial fauna that invaded the territory not earlier than 15,000 years ago (Nikonov and Shlyukov 2002). Since the divergence times of kinetoplastids have so far been calculated in hundreds of millions of years (Fernandes et al. 1993), this may be too short a period for a detectable genetic (but not morphological!) radiation of these protozoans. Based on morphology, *Leptomonas* sp. F6, F5, *Blastocrithidia gerricola* KV1, and *B. miridarum* BM4 were previously ranked into different genera. The assignment of the latter species to the genus *Blastocrithidia* was based on the structure of the epimastigote cells and nature of the cysts (''straphangers'') in natural infections, while only promastigotes were observed in culture (Podlipaev and Frolov 1987). These four flagellates correspond to the crown group with very short branches in the 18S rRNA-based trees (Merzlyak et al. 2001b), almost identical UP-PCR profiles (Bulat, Mokrousov, and Podlipaev 1999) and ND8 mRNA editing patterns (Merzlyak, Zakharova, and Kolesnikov 2001a), and very similar 5S rRNA sequences and intergenic regions (this paper). We propose that the lack of differences in the latter region may serve as a clear-cut criterion for an assignment of the different isolates into the same species. The crown isolates very likely represent the same species with a wide host range (Hemiptera: Gerridae, Nabidae, Miridae) and geographical distribution (e.g. the longest distance between localities is 1,500 km). It is notable that individual isolates were introduced into culture within a 20-yr. interval. From this observation we conclude that at least some insect trypanosomatids have low host specificity, occupy a large geographical area, and at the same time exhibit extensive morphological variability, rendering their morphology-based taxonomy unreliable.

The crown group, composed of insect trypanosomatids and the *Leishmania*/*Endotrypanum* clade, is supported by a number of molecular markers and brings together species that differ significantly by morphology, life cycles (one or two hosts), host organism (insect and/or vertebrates), and geographic origin (see Table 1). A parsimonious explanation is that the acquisition of the complex life cycle occurred once after the *Leishmania*/*Endotrypanum* branch separated from the insect trypanosomatids.

An interesting group that was unanimously highly supported in single-gene and concatenated trees is the one that embraces *Leptomonas* sp. Nfm, four symbiont-free herpetomonads, and an unnamed trypanosomatid ''X''. In the original description of this latter flagellate, isolated in 1983 from rats and stray dogs in Alexandria, Egypt, Morsy et al. (1988) noticed significant differences from an anticipated leishmania, both in morphology and enzymatic patterns. Moreover, it was predicted that this was an infectious agent ''of an unknown genus or unrecorded type'' (Morsy et al. 1988). The phylogenetic trees reported herein surprisingly affiliate the rodent/canine isolates with the genus *Herpetomonas*. To our knowledge this is the first time that a flagellate isolated from mammals has appeared within a genus, so far associated only with insects. Such an extremely wide host range for a trypanosomatid genus would be another novelty. It is tempting to speculate that this primarily insect flagellate was transferred during blood feeding and is in the process of establishing itself in warm-blooded vertebrates. This preliminary conclusion is supported by a close relatedness between trypanosomatid ''X'' and *H. ztiplika*, a herpetomonad from a blood-sucking biting midge (Podlipaev et al. 2004). Possible acquisitions of mammals as hosts by primarily insect flagellates have also been recently proposed by others (Dedet and Pratlong 2000; Jimenez et al. 1996). However, a cautionary note should be made that strain mislabeling or contamination may have occurred, and more detailed analyses including Southern hybridization, as well as isolations of more similar strains from rodents in Egypt will be needed to confirm our finding.

Herpetomonas muscarum (type species of the genus), *H. megaseliae*, and *H. pessoai* can be considered as ''classical'' members of the genus because they possess the opistomastigote stage, which is absent in *H. ztiplika* (Podlipaev et al. 2004). The ''herpetomonad'' clade, which contains species from all over the world (Table 1), does not seem to be an artifact caused by low information content of the 5S and SL RNA genes, since

the same set of species is also clearly monophyletic in the 18S rRNA phylogenies (Hollar, Lukeš, and Maslov 1998; Merzlyak et al. 2001b). The absence of a single clade that would bring together endosymbiont-containing trypanosomatids is an important feature of the 5S/SL rRNA trees that is difficult to reconcile with the 18S rRNA data (Hollar, Lukeš, and Maslov 1998; Maslov, Podlipaev, and Lukeš 2001; Merzlyak et al. 2001b) and may be caused by different evolutionary rates.

A newly isolated *Wallaceina* sp. Wsd (see below) with the characteristic endomastigote stage is interspersed between *H. mariadeanei* and the other endosymbiont-free herpetomonads. This branching order and differences in diagnostic nucleotides argue for the exclusion of *H. mariadeanei* from the genus *Herpetomonas*. Although only 18S rRNA and UP-PCR data are available for other members of the genus *Wallaceina* (Bulat, Mokrousov, and Podlipaev 1999; Merzlyak et al. 2001b), one can conclude that as typical crown species, they would appear elsewhere in the tree than *Wallaceina* sp. Wsd. Therefore, we predict that larger sampling will eventually also lead to polyphyly of this genus.

The position of *Leptomonas* sp. Nfm is noteworthy, since this isolate with most typical promastigotes could well represent a model species of the genus *Leptomonas*. However, in the 18S rRNA trees (Merzlyak et al. 2001b), *Leptomonas* sp. Nfm appeared in a poorly-supported long branch located between the ''slowly evolving'' crown and the symbiont-free herpetomonads. The contradiction between its promastigote morphology and association with the *Herpetomonas* branch makes *Leptomonas* sp. Nfm a clear example of poor applicability of cell morphotypes for trypanosomatid taxonomy. Based on its SL and 5S rRNA sequences we propose that *Leptomonas* sp. Nfm be transferred into the genus *Herpetomonas* as *Herpetomonas* sp. Nfm.

To understand the way an insect trypanosomatid circulates in nature, one has to know the extent of its host specificity. For organisms with a dearth of morphological characters this can be addressed only using molecular data. It is not easy to find an infected insect population isolated from other potential hosts that would be suitable for such a study. We have found one in the case of *Salda littoralis*, an unusual bug from the supralittoral zone of the White Sea, where it flourishes during the short summer season under a thick layer of stranded fucoids. *Salda* is strictly confined to this unique biotope, shared with only one other bug species, *Saldula pallipes*, which belongs to the same insect family. Three morphologically easily distinguishable cultures isolated from these hosts by one of us (S.A.P.) are represented by promastigotic *Leptomonas* sp. Sld and PL and endomastigotic *Wallaceina* sp. Wsd. Since *Leptomonas* sp. PL and *Wallaceina* sp. Wsd have been isolated on the same locality and day from different specimens of *S. littoralis*, we propose that the ''one host-one parasite'' paradigm should not be applied to insect trypanosomatids, since both parasites are not closely related, as judged from their 5S rRNA sequences. At the same time, leptomonads Sld and PL isolated from *S. littoralis* and *S. pallipes*, respectively, seem to belong to the same species. Since both hosts are members of different genera of the family Saldidae, and the cultures were obtained within a one-year interval, this is a strong argument for low specificity of the parasite. An alternative explanation postulates that, as a predator, *Salda* got infected while feeding on *Saldula*, its frequent prey (Merkulov and Slisarenko, pers. commun.).

The taxonomy of the insect trypanosomatids requires new criteria other than morphology and 18S rRNA-based phylogenies to re-evaluate the generic affinities of known isolates. Here we show that a combination of the SL RNA and 5S rRNA gene markers can provide relevant information to discriminate within close groups and to indicate relatedness that is not evident from inherent morphological stages. These two marker genes can also provide tools for the long-term study of kinetoplastid populations within geographically-isolated regions.

ACKNOWLEDGMENTS

We thank Lee Schnur (Hebrew University, Israel), Laura Landweber (Princeton University, USA), and Marta Teixeira and Erney Camargo (Universidade de Sao Paulo, Brazil) for kindly providing the genomic DNA from trypanosomatid ''X'' and *Herpetomonas* spp., and Helena Toma for help with some cloning. This research was supported by grants from the European Community (QLK 2-CT-2001-01810) (to J.L.) and the CIRAD, grant ATP 2002/30 (to M.D. and S.P.). S. J.W. is a trainee of the UCLA Bioinformatics Integrative Graduate Education and Research Traineeship program funded by NSF grant DGE9987641.

LITERATURE CITED

Brandao, A. A., Miranda, A., Degrave, W. M. & Sousa, M. A. 2000. The heterogeneity of choanomastigote-shaped trypanosomatids as analyzed by their kDNA minicircle size: taxonomic implications. *Parasitol. Res.,* **86**:809–812.

Bulat, S. A., Mokrousov, I. V. & Podlipaev, S. A. 1999. Classification of trypanosomatids from insects and plants by the UP-PCR (universally primed PCR) technique and cross dot blot hybridization of PCR products. *Europ. J. Protistol.,* **35**:319–327.

Camargo, E. P. 1999. *Phytomonas* and other trypanosomatid parasites of plants and fruit. *In*: Baker, J. R., Muller, R. & Rollinson, R. (ed.), Advances in Parasitology. Vol. 42. Academic Press, San Diego. p. 29–112.

Campbell, D. A., Fernandes, O. & Sturm, N. R. 1997. The mini-exon gene is a distinctive nuclear marker for distinguishing protists of the kinetoplastid/euglenoid lineage. *Mem. Inst. Oswaldo Cruz,* **92**(Suppl. I):12–13.

Dedet, J. P. & Pratlong, F. 2000. *Leishmania*, *Trypanosoma* and monoxenous trypanosomatids as emerging opportunistic agents. *J. Eukaryot. Microbiol.,* **47**:37–39.

Dollet, M. 2001. Phloem-restricted trypanosomatids form a clearly characterized monophyletic group among trypanosomatids isolated from plants. *Int. J. Parasitol.,* **31**:459–467.

Dollet, M., Sturm, N. R. & Campbell, D. A. 2001. The spliced leader RNA gene array in phloem-restricted plant trypanosomatids (*Phytomonas*) partitions into two major groups: epidemiological implications. *Parasitology,* **122**:289–297.

Dollet, M., Sturm, N. R., Sanchez-Moreno, M. & Campbell, D. A. 2000. 5S ribosomal RNA gene repeat sequences define at least eight groups of plant trypanosomatids (*Phytomonas* spp.): phloem restricted pathogens form a distinct section. *J. Eukaryot. Microbiol.,* **47**:569–574.

Donelson, J. E., Gardner, M. J. & El-Sayed, M. N. 1999. More surprises from Kinetoplastida. *Proc. Natl. Acad. Sci. USA,* **96**:2579–2581.

Drouin, G. & de Sá, M. M. 1995. The concerted evolution of 5S ribosomal genes linked to the repeat units of other multigene families. *Mol. Biol. Evol.,* **12**:481–493.

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. USA,* **90**: 11608–11612.

Fernandes, O., Teixeira, M. M. G., Sturm, N. R., Sousa, M. A., Camargo, E. P., Degrave, W. M. & Campbell, D. A. 1997. Mini-exon gene sequences define six groups within the genus *Crithidia*. *J. Eukaryot. Microbiol.,* **44**:535–539.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.,* **41**:95–98.

Hannaert, V., Opperdoes, F. R. & Michels, P. A. M. 1998. Comparison and evolutionary analysis of the glycosomal glyceraldehyde-3-phosphate dehydrogenase from different kinetoplastida. *J. Mol. Evol.,* **47**: 728–738.

Hollar, L., Lukeš, J. & Maslov, D. A. 1998. Monophyly of endosym-

biont-containing trypanosomatids: phylogeny versus taxonomy. *J. Eukaryot. Microbiol.,* **45**:293–297.

Hughes, A. L. & Piontkivska, H. 2003. Phylogeny of Trypanosomatida and Bodonida (Kinetoplastida) based on 18S rRNA: evidence for paraphyly of *Trypanosoma* and six other genera. *Mol. Biol. Evol.,* **20**:644–652.

Jimenez, M. I., Lopez-Velez, R., Molina, R., Canavate, C. & Alvar, J. 1996. HIV co-infection with a currently non-pathogenic flagellate. *Lancet,* **347**:264–265.

Landweber, L. F. & Gilbert, W. 1993. RNA editing as a source of genetic variation. *Nature,* **363**:179–182.

Lukeš, J., Jirků, M., Doležel, D., Král'ová, I., Hollar, L. & Maslov, D. A. 1997. Analysis of ribosomal RNA genes suggests that trypanosomes are monophyletic. *J. Mol. Evol.,* **44**:521–527.

Maslov, D. A., Podlipaev, S. A. & Lukeš, J. 2001. Phylogeny of the Kinetoplastida: taxonomic problems and insight into the evolution of parasitism. *Mem. Inst. Oswaldo Cruz,* **96**:394–402.

Maslov, D. A., Lukeš, J., Jirků, M. & Simpson, L. 1996. Phylogeny of trypanosomes as inferred from the small and large subunit rRNAs: implications for the evolution of parasitism in the trypanosomatid protozoa. *Mol. Biochem. Parasitol.,* **75**:197–205.

Merzlyak, E. M., Zakharova, M. Yu. & Kolesnikov, A. A. 2001a. Monogenetic trypanosomatids: comparison of the ND-8 editing gene. *Europ. J. Protozool.,* **37**:233–239.

Merzlyak, E., Yurchenko, V., Kolesnikov, A. A., Alexandrov, K., Podlipaev, S. A. & Maslov, D. A. 2001b. Diversity and phylogeny of insect trypanosomatids based on small subunit rRNA genes: polyphyly of *Leptomonas* and *Blastocrithidia*. *J. Eukaryot. Microbiol.,* **48**:161–169.

Molyneux, D. H. & Ashford, R. W. 1983. The biology of *Trypanosoma* and *Leishmania*, parasites of man and domestic animals. Taylor & Francis, London.

Morsy, T. A., Schnur, L. F., Feinsod, F. M., Michael, S. A., Saah, A., Salama, M. M. & Wahba, M. M. 1988. The discovery and preliminary characterization of a novel trypanosomatid parasite from *Rattus norvegicus* and stray dogs from Alexandria, Egypt. *Ann. Trop. Med. Parasitol.,* **82**:437–444.

Nikonov, A. A. & Shlyukov, A. I. 2002. On the time of deglaciation of Karelskii Isthmus. *Dokl. Akad. Nauk,* **387**:393–396.

Podlipaev, S. A. 2000. Insect trypanosomatids: the need to know more. *Mem. Inst. Oswaldo Cruz,* **95**:517–522.

Podlipaev, S. A. 2001. The more insect trypanosomatids under study—the more diverse Trypanosomatidae appears. *Int. J. Parasitol.,* **31**: 648–652.

Podlipaev, S. A. & Frolov, A. O. 1987. The description and laboratory cultivation of *Blastocrithidia miridarum* sp.n. (Mastigophora, Trypanosomatidae). *Parasitologija (St. Petersburg),* **21**:545–552 (in Russian, with English summary)

Podlipaev, S. A., Malysheva, M. N. & Kolesnikov, A. A. 1991. *Leptomonas rigidus* sp. n. (Trypanosomatidae)—a parasite of *Salda littoralis* L. (Hemiptera: Heteroptera). *Acta Protozool.,* **30**:121–127.

Podlipaev, S. A., Votýpka, J., Jirků, M., Svobodová, M. & Lukeš, J. 2004. *Herpetomonas ztiplika* n. sp. (Kinetoplastida: Trypanosomatidae)—parasite of the blood-sucking biting midge *Culicoides kibunensis* Tokunaga, 1937 (Diptera: Ceratopogonidae). *J. Parasitol.,* (in press).

Posada, D. & Crandall, K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics,* **14**:817–818.

Simpson, A. G. B. & Roger, A. J. 2004. Protein phylogenies robustly resolve the deep-level relationships within Euglenozoa. *Mol. Phylogenet. Evol.,* **30**:201–212.

Simpson, A. G. B., Lukeš, J. & Roger, A. J. 2002. The evolutionary history of kinetoplastids and their kinetoplasts. *Mol. Biol. Evol.,* **19**: 2071–2083.

Stevens, J. R 2001. One million insects—a lot of parasites? *Trends Parasitol.,* **17**:119–120..

Swofford, D. L. 2001. PAUP*: Phylogenetic Analysis Using Parsimony, Version 4.0b10. Sinauer Associates, Sunderland.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.,* **25**:4876–4882.

Uliana, S. R. B., Nelson, K., Beverley, S. M., Camargo, E. P. & Floeter-Winter, L. M. 1994. Discrimination amongst *Leishmania* by polymerase chain reaction and hybridization with small-subunit ribosomal DNA-derived oligonucleotides. *J. Eukaryot. Microbiol.,* **41**:324–330.

Vickerman, K. 2000. Order Kinetoplastea Honigberg, 1963. *In*: Lee, J. J., Leedale, G. F. & Bradbury, P. (ed.), An Illustrated Guide to the Protozoa. 2nd ed. American Society of Protozoologists, Lawrence, Kansas. p. 1159–1184.

Wallace, F. G. 1966. The trypanosomatid parasites of insects and arachnids. *Exp. Parasitol.,* **18**:124–193.

Received 08/28/03, 01/07/04, 02/18/04; accepted 02/19/04