Protist, Vol. 166, 551–568, November 2015
http://www.elsevier.de/protis
Published online date 21 October 2015

Protist

ORIGINAL PAPER

Host-specificity of Monoxenous Trypanosomatids: Statistical Analysis of the Distribution and Transmission Patterns of the Parasites from Neotropical Heteroptera

Eugene Kozminsky[a], Natalya Kraeva[b], Aygul Ishemgulova[b], Eva Dobáková[c,1], Julius Lukeš[c,d,e], Petr Kment[f], Vyacheslav Yurchenko[b,d], Jan Votýpka[d,g], and Dmitri A. Maslov[h,2]

[a]Zoological Institute, Russian Academy of Sciences, St.-Petersburg, 199034, Russia
[b]Life Science Research Centre, University of Ostrava, 70200 Ostrava, Czech Republic
[c]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, 37005 České Budějovice (Budweis), Czech Republic
[d]Faculty of Science, University of South Bohemia, 37005 České Budějovice (Budweis), Czech Republic
[e]Canadian Institute for Advanced Research, Toronto, ON M5G 1Z8, Canada
[f]Department of Entomology, National Museum, 19300 Prague, Czech Republic
[g]Department of Parasitology, Faculty of Science, Charles University, 12843 Prague, Czech Republic
[h]Department of Biology, University of California - Riverside, Riverside, CA 91521, USA

Submitted May 28, 2015; Accepted August 18, 2015
Monitoring Editor: Michael Melkonian

Host-parasite relationships and parasite biodiversity have been the center of attention for many years; however the primary data obtained from large-scale studies remain scarce. Our long term investigations of trypanosomatid (Euglenozoa: Kinetoplastea) biodiversity from Neotropical Heteroptera have yielded almost one hundred typing units (TU) of trypanosomatids from one hundred twenty host species. Half of the parasites' TUs were documented in a single host species only but the rest were found parasitizing two to nine species of hosts, with logarithmic distribution best describing the observed distribution of parasites among hosts. Different host superfamilies did not show significant differences in numbers of trypanosomatid TUs they carry, with exception of Pyrrhocoroidea which showed higher parasite richness than any other group tested. Predatory reduviids shared significantly larger numbers of parasite TUs with phytophagous mirids and coreids than the numbers shared between any other

[1]Permanent address: Faculty of Natural Sciences, Comenius University, Bratislava, Slovakia
[2]Corresponding author; fax +1 951 827 4286
e-mail maslov@ucr.edu (D.A. Maslov).

http://dx.doi.org/10.1016/j.protis.2015.08.004

groups. These results show that the specificity of trypanosomatid-heteropteran associations is not very strict: parasites seem to be transmissible between different host groups within the same niche and predatory hosts may acquire parasites from their prey.

Key words: Trypanosomatids; Heteroptera; host-parasite specificity; biodiversity; Spliced Leader RNA.

Introduction

For nearly a century the "one host – one parasite" paradigm has been the main criterion used to define species of insect trypanosomatids (Podlipaev 1990; Vickerman 1976; Wallace 1966). Historically, the notion of a strict specificity of the parasite–host systems has its roots in investigations of multicellular parasites which undergo complex life cycles and, therefore, have evolved intricate adaptations to particular host species. In trypanosomatid protists (Euglenozoa: Kinetoplastea), strict host specificity is supported by data from dixenous insect-transmitted parasites of vertebrates (Hoare 1972; Lainson and Shaw 1987). Thus, in *Leishmania*, a narrow range of competent vectors is defined by the specific and complex interactions between the parasite and the sand fly accomplished by the parasite's surface lipophosphoglycan molecules (Sacks 2001; Turco and Sacks 2003). On the contrary, monoxenous trypanosomatids appear to have a relaxed stringency of associations with their hosts increasing the frequency of non-specific infections. These speculations are supported by the apparent lack of parasite-host co-evolutionary patterns in the analyses of flagellates in several host taxa (Bulat et al. 1999; Merzlyak et al. 2001; Votýpka et al. 2010). In the latter cases, the investigated hosts were found to bear rather diverse sets of parasites (Chandler and James 2013; Maslov et al. 2007; Westenberger et al. 2004; Wilfert et al. 2011). Such observations led to the departure from the venerable and easy-to-grasp "one host – one monoxenous parasite" philosophy replaced with the concept of the flexible "broad" specificity of these organisms (Kostygov et al. 2014; Podlipaev 2003).

Nonetheless, there is evidence in favor of strong and specific interactions of at least some monoxenous trypanosomatids and their hosts. In some of such cases, the parasites are found lining the host's intestinal walls as a dense carpet, or can be found in just one particular compartment, e.g in the midgut, hindgut, Malpighian tubules or hemolymph (J.V, D.A.M and J.L., unpubl. data) (Schwarz et al. 2015). The flagellates' attachment is not random: thus, *Blastocrithidia triatomae* colonizes the midgut epithelium and hindgut walls of its host, *Triatoma infestans*, and in the hindgut it is preferentially found attached in the areas different from those used by *Trypanosoma cruzi* (Schaub et al. 1992; Schaub and Boker 1986). In other cases, e.g. *Leptomonas wallacei* infecting *Oncopeltus fasciatus*, the protists are found in the gut lumen in association with perimicrovillar membranes as well as attached to the hindgut walls. Noticeably, the progression from one gut section to another was accompanied by a change of the parasite's morphology, as well as the surface carbohydrates suggesting specific and dynamic interactions on the molecular level (Romeiro et al. 2003a, b). Surface carbohydrates have also been implicated in the process of colonizing salivary glands of the mosquito *Aedes aegypti* by a monoxenous trypanosomatid *Blastocrithidia culicis* (Nascimento et al. 2010). These interactions resemble those which are observed between dixenous parasites and their vectors. However, while the development of dixenous trypanosomatids in a vector results in producing an infective (e.g. metacyclic) stage, in monoxenous species its main purpose may be to maintain a life-long infection of the host which would also increase the chance of transmission.

Demonstration of the complex and specific interactions of monoxenous parasites and their hosts has been supplemented by the recent biodiversity surveys which revealed associations of certain large groups (orders) of hosts with respective major clades of parasites (Maslov et al. 2013; Teixeira et al. 2011; Týč et al. 2013; Votýpka et al. 2012a, 2013). The smaller groups of trypanosomatids with apparent host specificity at the family or genus levels have also been discovered (Maslov et al. 2007; Votýpka et al. 2012b). At the same time, there is a growing number of examples suggesting non-specific or transient infections, e.g. in insectivorous insects which can acquire parasites from prey or in dipterans which feed on nutrient-rich substrates and can sustain a long-term survival of parasites in their gut lumen (Carvalho and Deane 1974; Chandler and James 2013; Maslov et al. 2007; Týč et al. 2013; Yurchenko et al. 2006b). Therefore, the emerging view is that neither the

broad nor the strict specificity paradigms are applicable to the entire group which turned out to be more diverse than was anticipated even a few years ago (Maslov et al. 2013). This diversity includes not only genetic make-ups of the organisms but also various specificity levels which in turn are defined by the differences in interactions between hosts and parasites. The current study aims at further testing the variable specificity hypothesis and at estimating the relative occurrence of parasites with the narrow and/or broad host range. The multi-year study presented herein concentrates on parasites in Neotropical Heteroptera, so far the most extensively sampled host order and biogeographic region.

Results

Current Status of the Survey

The recent survey of trypanosomatid parasites in two Neotropical regions is summarized in Table 1. Out of 1,210 insects that were dissected and analyzed for trypanosomatids using light microscopy, 250 insects were found to contain parasites in the intestinal tract. DNA from the intestinal samples was preserved and subsequently used for amplification of a trypanosomatid-specific genetic marker, Spliced Leader (SL) RNA gene repeat. An entire panel of the amplified products for the Ecuador 2008 collection is shown in Figure 1, as illustration. The repeat size varies in a species-specific manner (Podlipaev et al. 2004; Ramos et al. 1996), therefore the heterogeneity of amplicons in the collection was the first indication of the diversity among the encountered trypanosomatids. Accordingly, most amplicons were cloned and sequenced. In those cases when a set of similar-size amplicons was obtained from a host population (e.g. samples 190-194 or 223-228 in Fig. 1), some samples were skipped as potentially representing identical trypanosomatids. The obtained SL repeat sequences were first analyzed by a multiple alignment procedure which served to identify clusters of most related sequences, and sequences within clusters were subsequently analyzed by pair-wise comparisons. A threshold of 90% identity was applied to delineate individual typing units (TUs) which currently serve as proxies for "molecular species" of trypanosomatids (Maslov et al. 2007). After combining this analysis with the earlier surveys (Maslov et al. 2007; Westenberger et al. 2004), the tally contains 95 trypanosomatid TUs originated from 120 species of Neotropical Heteroptera. Our current Neotropical SL repeat database includes ~450 entries including 193 SL RNA repeat sequences determined in this work (Table 2; see also table 1 of Maslov et al. 2007, and table 1 of Westenberger et al. 2004).

General Characterization of Host-parasite Associations

Considering the multitude of parasites and hosts, we began asking questions about patterns in host-parasite associations. As a pre-requisite to this analysis, the SL sequences were used to derive a neighbor-joining dendrogram which reflects phylogenetic relationships among trypanosomatids (Fig. 2). The repeat's most conserved regions (encompassing ~150 bp from the -100 position in relation to the transcription start to the transcription termination sites by the oligo(T)-block at the 3′ end of the SL RNA gene) are alignable across the entire family and were used for this purpose. The dendrogram is consistent with the known phylogeny of the group that was established earlier and shows that the Trypanosomatidae family is subdivided into several major clades (summarized in Maslov et al. 2013). In accordance with previous analyses (Maslov et al. 2007, 2010; Votýpka et al. 2012b), most trypanosomatid flagellates of Heteroptera fell into two major clades: Leishmaniinae and *Blastocrithidia*. The SL tree also shows that some flagellates belong to the '*jaculum*' and *Phytomonas* clades, while remaining TUs formed individual branches. Identities of some of these branches have been determined by analysis of the SSU rRNA genes, while some other branches appear to represent novel clades (data not shown). However, in either case these branches represent minority groups. Thus, trypanosomatids of Heteroptera predominantly belong to four major clades: Leishmaniinae, *Blastocrithidia*, '*jaculum*' and *Phytomonas*; with the last three clades so far being limited to this group of insect hosts.

Considering the distribution of the major clades or subclades among the higher order taxonomic units (superfamilies, families, groups of more closely related genera) of hosts, no preferential associations were found at this level (data not shown). However, on the low taxonomic level, there are several cases which appear to represent preferential parasite-host relations. Thus, a group of related trypanosomatids (TU's *6/7*, *130*, *162-166*) is mostly associated with several species of the family Alydidae. TU*1* is seen in association with *Dysdercus* (Pyrrhocoridae) (although in other geographic regions this widespread TU is also found

Table 1. Summary of the field work[a].

Region:	Ecuador 2008	Costa Rica
Dissected hosts (species)	624 (~140)	586 (~130)
Infected hosts (species)	125 (69)	125 (54)
Samples with successful PCR	117 (93.6%)	74 (59.2%)
Samples sequenced	75	72
No. of TUs found	38	35
Increase in the total TU count	27[b]	17[c]

[a]The presented data extend the previously published 2004-2007 work
[b]In addition to the 2004-2007 count
[c]In addition to the combined 2004-2007 and Ecuador 2008 counts

in other pyrrhocorid genera [Votýpka et al. 2012b; Westenberger et al. 2004]), TU*17* occurs in several genera of Miridae, TU*12* in three species of *Largus* (Largidae), TU*25* in four species of *Leptoscelis* (Coreidae), the related TUs *132* and *133* in two genera of Rhopalidae, etc. It must be mentioned that in the latter cases we cannot exclude that there are additional hosts outside of those groups. The preferential species-specific associations do not represent the prevailing type in these parasite-host systems, e.g. TU*166* is found in five host families, and the cases when a single TU is found in two or three families are common.

Quantitative Characterization of Host-parasite Associations

Since a single parasite TU is often capable of infecting more than one host species, we then attempted to quantify the observed distribution of parasites among host species. A numerical parameter that defines this distribution can be used as a measure of host specificity of parasites. The histogram presented in Figure 3 ("*Observed*") shows the numbers of individual TUs parasitizing respective numbers of hosts (***x***, which varies from 1 to 9). About half of the TUs (49 out of 95) were found in a single host

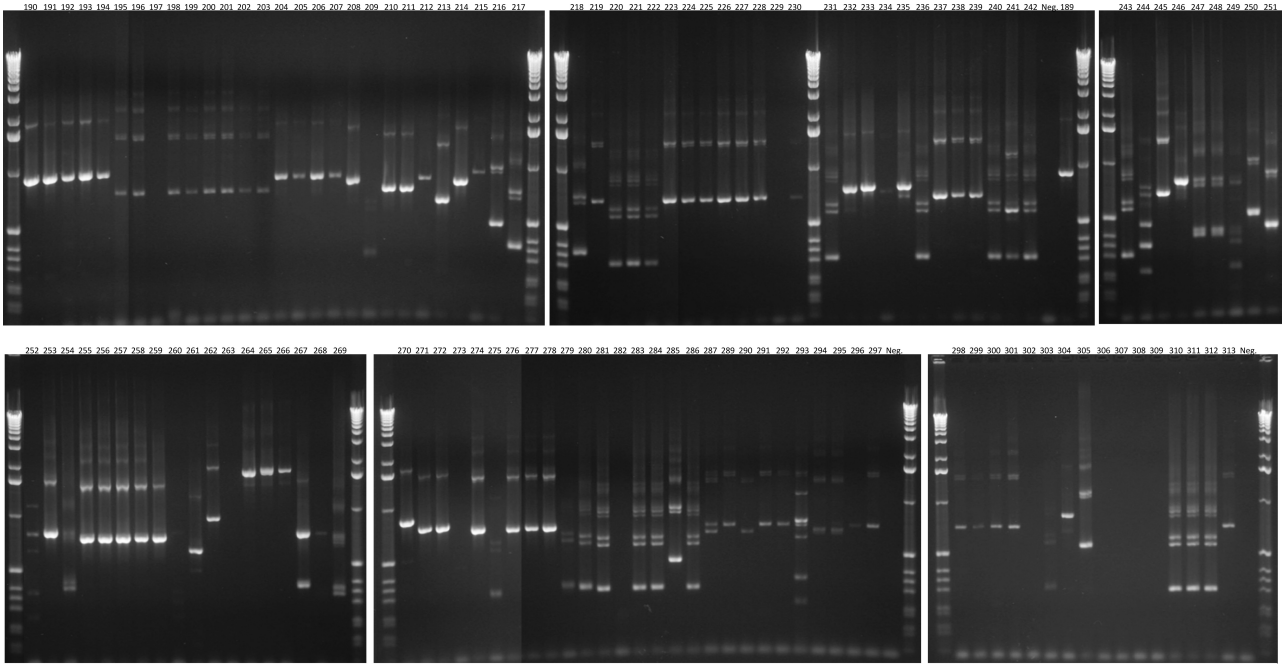


Figure 1. A panel of the PCR amplified SL RNA gene repeat fragments from the set of intestinal DNA samples collected in Ecuador in 2008. The amplified DNA was fractionated in 1% agarose gel. The figure represents a composite image of several gels. 1 kb DNA ladder (Life Technologies) was used as size standards. The numbers atop of each panel are consistent with the DNA sample IDs given in Table 2 (e.g. lane 190 represents sample 190MD etc).

Table 2. Summary of SL RNA sequences from Ecuador (samples 189MD to 313OT) and Costa Rica (samples 314AR to 438AR) obtained during the 2008-2012 studies.

DNA Sample	GenBank™ No.	Typing Unit	Collection Locale	Host species (Family)	SL Repeat Size (bp)
189MD	KP717764	92	Mindo	*Fortunacoris castaneus* (Miridae)	872
190MD	KP717765	92	Mindo	*Fortunacoris castaneus* (Miridae)	869
191MD	KP717766	92	Mindo	*Fortunacoris castaneus* (Miridae)	869, 870
191MD	KP717767, KP717768	17	Mindo	*Fortunacoris castaneus* (Miridae)	648
192MD	KP717769	54	Mindo	*Calocorisca altiplana* (Miridae)	864
193MD	KP717770	92	Mindo	*Fortunacoris castaneus* (Miridae)	871
196MD	KP717771, KP717772	94	Mindo	*Neofurius* sp. (Miridae)	678
197MD	KP717773	140	Mindo	*Neofurius* sp. (Miridae)	773
200MD	KP717774, KP717775	94	Mindo	*Neofurius* sp. (Miridae)	677, 678
204MD	KP717776	95	Mindo	*Fulvius* cf. *breddini* (Miridae)	803
208AL	KP717777, KP717778	19/20	Napo-1	*Ricolla quadrispinosa* (Reduviidae)	754, 759
209AL	KP717779	142	Napo-1	*Montina* sp. (Reduviidae)	257
210AL	KP717780, KP717781	138	Napo-1	*Hypselonotus* cf. *subterpunctatus* (Coreidae)	686
211AL	KP717782, KP717783	138	Napo-1	*Hypselonotus* cf. *subterpunctatus* (Coreidae)	686, 687
212AL	KP717784	28	Napo-1	*Zoreva lacerna* (Coreidae)	794
213AL	KP717785-KP717787	136	Napo-1	*Neopamera* sp. (Rhyparochromidae)	597, 602
214BN	KP717788 - KP717790	140	Baños	*Piezogaster rubropictus* (Coreidae)	772, 773
215BN	KP717791	54	Baños	*Stenodema andina* (Miridae)	864
216BN	KP717792 - KP717794	93	Baños	*Jadera parapectoralis* (Rhopalidae)	447
217LO	KP717795	139	Loja	*Eurygerris kahli* (Gerridae)	686
218LO	KP717796	96	Loja	*Eurygerris kahli* (Gerridae)	319
219VB	KP717797 - KP717799	41	Vilcabamba	*Collaria oleosa* (Miridae)	269
221VB	KP717800	17	Vilcabamba	*Collaria oleosa* (Miridae)	649
223VB	KP717801, KP717802	17	Vilcabamba	*Neotropicomiris nordicus* (Miridae)	651
224VB	KP717803, KP717804	41	Vilcabamba	*Neotropicomiris nordicus* (Miridae)	269
224VB	KP717805	17	Vilcabamba	*Neotropicomiris nordicus* (Miridae)	648
226VB	KP717806	17	Vilcabamba	*Neotropicomiris nordicus* (Miridae)	648
227VB	KP717807, KP717808	17	Vilcabamba	*Neotropicomiris nordicus* (Miridae)	648, 651
228VB	KP717809 - KP717812	17	Vilcabamba	*Neotropicomiris nordicus* (Miridae)	648, 649, 651
230VB	KP717813	44/90	Vilcabamba	Eccritotarsini gen. sp. 1 (Miridae)	731
231VB	KP717814, KP717815	41	Vilcabamba	*Collaria oleosa* (Miridae)	268, 269
232VB	KP717816,KP717819-KP717822	99	Vilcabamba	*Repipta* sp. (Reduviidae)	717, 723, 724
232VB	KP717817, KP717818	137	Vilcabamba	*Repipta* sp. (Reduviidae)	749
233VB	KP717823, KP717824	137	Vilcabamba	*Heza* cf. *ephippium* (Reduviidae)	749

Table 2 (*Continued*)

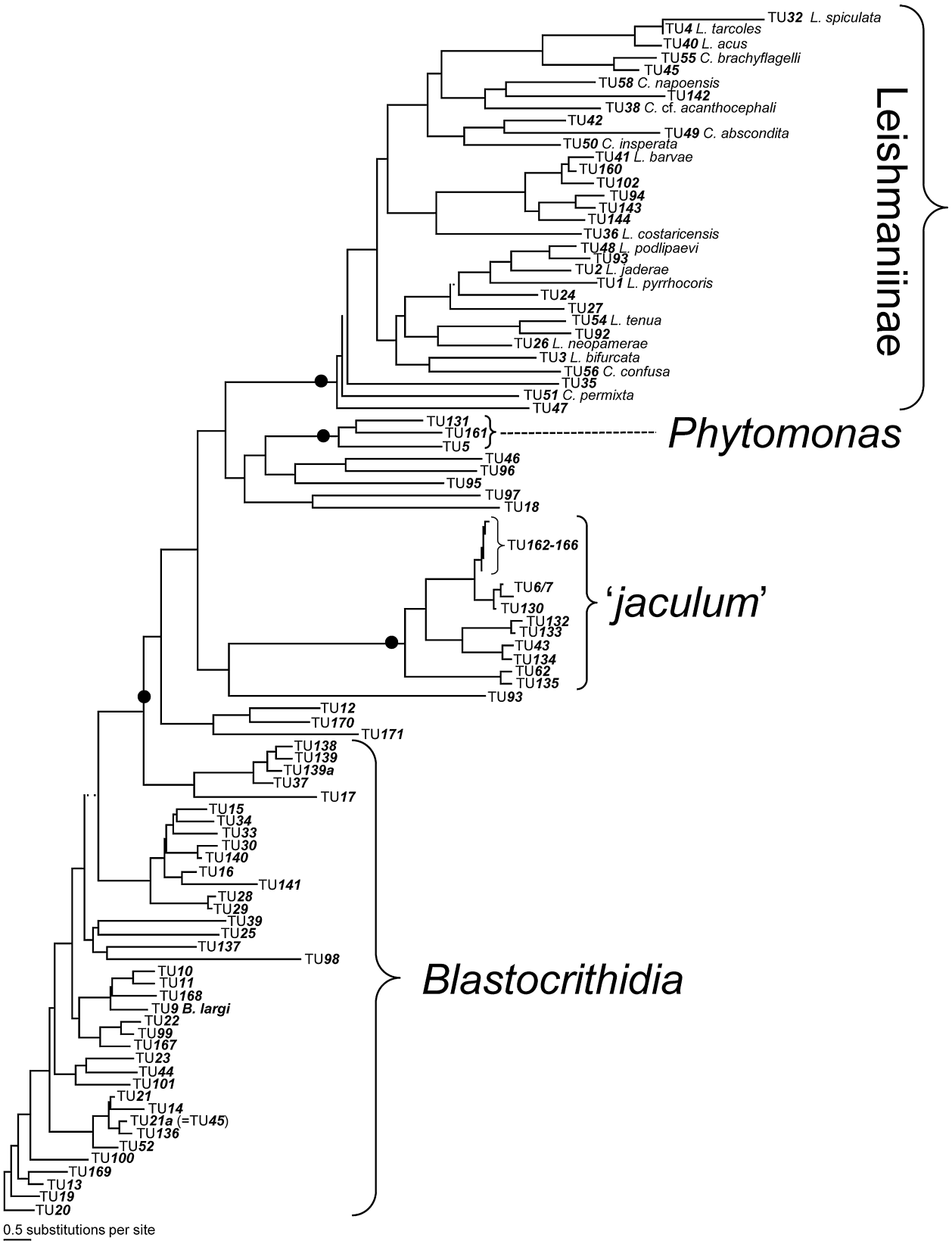
DNA Sample	GenBank™ No.	Typing Unit	Collection Locale	Host species (Family)	SL Repeat Size (bp)
235VB	KP717825, KP717826	137	Vilcabamba	*Repipta* sp. (Reduviidae)	750
236VB	KP717827 - KP717831	41	Vilcabamba	*Collaria oleosa* (Miridae)	269
237VB	KP717832, KP717833	92	Vilcabamba	*Neotropicomiris nordicus* (Miridae)	870, 871
240VB	KP717834 - KP717836	41	Vilcabamba	*Collaria oleosa* (Miridae)	269
242VB	KP717837	41	Vilcabamba	*Collaria oleosa* (Miridae)	269
243VB	KP717838	41	Vilcabamba	*Collaria oleosa* (Miridae)	269
244VB	KP717839 - KP717842	17	Vilcabamba	*Proba* sp. (Miridae)	649, 651
245VB	KP717843, KP717844	17	Vilcabamba	*Proba* sp. (Miridae)	651, 657
246VB	KP717845	33/34	Vilcabamba	*Chlorocoris* cf. *complanatus* (Pentatomidae)	777
247VB	KP717846	133	Vilcabamba	*Niesthrea vincentii* (Rhopalidae)	356
248VB	KP717847	133	Vilcabamba	*Niesthrea vincentii* (Rhopalidae)	354
249VB	KP717848	133	Vilcabamba	*Harmostes serratus* (Rhopalidae)	354
250VB	KP717849, KP717850	135	Vilcabamba	*Zicca taeniola* (Coreidae)	462, 468
251VB	KP717851	132	Vilcabamba	*Harmostes prolixus* (Rhopalidae)	391
252VB	KP717852	14	Vilcabamba	*Hypselonotus andinus* (Coreidae)	586
253VB	KP717853, KP717854	101	Vilcabamba	*Dysdercus mimus distanti* (Pyrrhocoridae)	750, 754
254VB	KP717855	129	Vilcabamba	Undet. species (Alydidae)	360
255LM	KP717856	139	Limon	*Hypselonotus andinus* (Coreidae)	686
261LM	KP717857	98	Limon	*Limnogonus aduncus* (Gerridae)	579
262AT	KP717858	42	Atacames	*Niesthrea vincentii* (Rhopalidae)	886
264AL	KP717859	134	Napo-1	*Leptoscelis pallida* (Coreidae)	1650
266AL	KP717860	134	Napo-1	*Leptoscelis pallida* (Coreidae)	1701
267AL	KP717861	25	Napo-1	*Leptoscelis serrata* (Coreidae)	714
268AL	KP717862	130	Napo-1	*Dysdercus ruficeps* (Pyrrhocoridae)	304
269AL	KP717863 - KP717866	130	Napo-1	*Vilga westwoodi* (Coreidae)	304, 305, 306, 331
270OT	KP717867	97	Otongachi	*Ricolla pallidinervis* (Reduviidae)	816
271OT	KP717868	100	Otongachi	*Proxys victor* (Pentatomidae)	731
272OT	KP717869	100	Otongachi	*Proxys victor* (Pentatomidae)	731
275OT	KP717870	139	Otongachi	Undet. species (Coreidae))	686
279OT	KP717871, KR056218	35	Otongachi	*Ricolla pallidinervis* (Reduviidae)	277, 285
280OT	KP717872, KP717874	44/90	Otongachi	*Ricolla pallidinervis* (Reduviidae)	735, 738
280OT	KP717873	145	Otongachi	*Ricolla pallidinervis* (Reduviidae)	275
283OT	KP717875 - KP717880	41	Otongachi	*Collaria oleosa* (Miridae)	269
283OT	KP717881	44/90	Otongachi	*Collaria oleosa* (Miridae)	735
285OT	KP717882	102	Otongachi	*Collaria oleosa* (Miridae)	423

287OT	KP717883	144	Otongachi	Eccritotarsini gen. sp. 1 (Miridae)	627
289OT	KP717884	44/90	Otongachi	Eccritotarsini gen. sp. 1 (Miridae)	731
290OT	KP717885	144	Otongachi	*Horciasisca* cf. ecuadorensis (Miridae)	627
293OT	KP717886	44/90	Otongachi	Eccritotarsini gen. sp. 1 (Miridae)	735
294OT	KP717887	44/90	Otongachi	Eccritotarsini gen. sp. 1 (Miridae)	739
297OT	KP717888	44/90	Otongachi	Eccritotarsini gen. sp. 1 (Miridae)	730
303OT	KP717889	145	Otongachi	*Ricolla pallidinervis* (Reduviidae)	273
304OT	KP717890	95	Otongachi	*Ricolla pallidinervis* (Reduviidae)	807
305OT	KP717891	8	Otongachi	*Zicca commaculata* (Coreidae)	523
311OT	KP717892	41	Otongachi	*Collaria oleosa* (Miridae)	269
313OT	KP717893	143	Otongachi	Eccritotarsini gen. sp. 1 (Miridae)	689
314AR	KR056219	131	Arenal	*Hypselonotus atratus hilaris* (Coreidae)	435
316AR	JF734884	19/20	Arenal	*Largus maculatus* (Largidae)	770
316AR (c)	JF734885, KR056220	50	Arenal	*Largus maculatus* (Largidae)	299, 301
317AR	KR056221	19/20	Arenal	*Ricolla simillima* (Reduviidae)	754
318AR	JF734886	12	Arenal	*Largus maculatus* (Largidae)	476
320AR	JF734888	19/20	Arenal	*Largus maculatus* (Largidae)	754
320AR (c)	JF734887	56	Arenal	*Largus maculatus* (Largidae)	428-431, 433
321AR	KR056222	19/20	Arenal	*Ricolla simillima* (Reduviidae)	778
324RV	JF937069, JF937068	1	Guanacaste-3	*Dysdercus obscuratus flavipenuis* (Pyrrhocoridae)	1068, 1069
325RV (i, c)	JF937070, JF937071	1	Guanacaste-3	*Dysdercus obscuratus flavipenuis* (Pyrrhocoridae)	1067, 1069
326RV (i, c)	JF937072, JF937073	1	Guanacaste-3	*Dysdercus obscuratus flavipenuis* (Pyrrhocoridae)	1069, 1070
327MV	KR056223	139	Monteverde	*Hypselonotus atratus hilaris* (Coreidae)	674
329MV (i-B, c)	JF937074, JF937075	1	Monteverde	*Dysdercus mimulus* (Pyrrhocoridae)	1065, 1068
329M (i-A)	KR056224	23	Monteverde	*Dysdercus mimulus* (Pyrrhocoridae)	769
330MV	KR056225	33/34	Monteverde	*Dysdercus bimaculatus* (Pyrrhocoridae)	776
331MV (i, c)	JF734889, JF734890	32	Monteverde	*Proba sallei* (Miridae)	239
332MV (i, c)	JF734891, JF734892	32	Monteverde	*Proba sallei* (Miridae)	239
333MV (c)	JF734893	32	Monteverde	*Proba sallei* (Miridae)	239
334MV (c)	JF734894	32	Monteverde	*Proba sallei* (Miridae)	239
335VL_A	KR056226	55	Tarcoles	*Piezogaster odiosus* (Coreidae)	288
335VL_B	KR056227	9	Tarcoles	*Piezogaster odiosus* (Coreidae)	704

Table 2 (*Continued*)

DNA Sample	GenBank™ No.	Typing Unit	Collection Locale	Host species (Family)	SL Repeat Size (bp)
336VL	KR056228	13	Tarcoles	*Piezogaster odiosus* (Coreidae)	768
337VL (i, c)	JF734895 - JF734897	54	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	863, 864
338VL_A	JF734998	13	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	769
338VL_B	JF734999	54	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	866
339VL (c)	JF734900	54	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	862
340VL (cult)	JF734901, JF734902	55	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	288
341VL (i, c)	JF734903, JF734904	54	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	862, 864
342VL (i, c)	JF734905, JF734906	55	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	288
343VL (i, c)	JF734907 - JF734909	55	Tarcoles	*Prepops* cf. *accinctus* (Miridae)	288
345VL	KR056229	141	Tarcoles	*Camptischium clavipes* (Coreidae)	776
346VL_A	KR056230	134	Tarcoles	*Leptoscelis tricolor* (Coreidae)	1655
346VL_B	KR056231	55	Tarcoles	*Leptoscelis tricolor* (Coreidae)	288
349DR	KR056232	171	Guanacaste-2	*Oncopeltus aulicus* (Lygaeidae)	652
351SR	KR056233	167	Guanacaste-4	*Brachystethus rubromaculatus* (Pentatomidae)	774
352MV	KR056234	169	Monteverde	*Mormidea integella* (Pentatomidae)	797
355MV	KR056235	170	Monteverde	*Stenomacra marginella* (Largidae)	485
356MV	KR056236	170	Monteverde	*Stenomacra marginella* (Largidae)	486
357VL-A	KR056237	13	Tarcoles	*Repipta taurus* (Reduviidae)	764
357VL-B	KR056238	6/7	Tarcoles	*Repipta taurus* (Reduviidae)	278
360VL	KR056239	9	Tarcoles	*Largus maculatus* x *cinctus* (Largidae)	703
370VL-A	KR056240	162	Tarcoles	*Stenocoris* sp. (Alydidae)	335
370VL-B	KR056241	164	Tarcoles	*Stenocoris* sp. (Alydidae)	407
371VL	KR056242	164	Tarcoles	*Stenocoris* sp. (Alydidae)	383
372VL-A	KR056243	166	Tarcoles	*Stenocoris* sp. (Alydidae)	359
372VL-B	KR056244	163	Tarcoles	*Stenocoris* sp. (Alydidae)	311
372VL-C	KR056245	6/7	Tarcoles	*Stenocoris* sp. (Alydidae)	332
373VL	KR056246	9	Tarcoles	*Pselliopus punctipes* (Reduviidae)	705
374VL	KR056247	9	Tarcoles	*Repipta* sp. (Reduviidae)	707
378VL	KR056248	166	Tarcoles	*Camptischium clavipes* (Coreidae)	359
379VL	KR056249	2	Tarcoles	*Camptischium clavipes* (Coreidae)	513
380SC-A	KR056250	138	Guanacaste-5	*Hypselonotus lineatus* (Coreidae)	665
380SC-B	KR056251	171	Guanacaste-5	*Hypselonotus lineatus* (Coreidae)	650
386SR	KR056252	168	Guanacaste-4	*Apiomerus* cf. *longispinnis* (Reduviidae)	720

388RV	KR056253	168	Guanacaste-3	*Apiomerus* sp.1 (Reduviidae)	685
390RV	KR056254	165	Guanacaste-3	*Hyalymenus subinermis* (Alydidae)	551
391RV	KR056255	171	Guanacaste-3	*Oncopeltus fasciatus* (Lygaeidae)	650
393RV	KR056256	33/34	Guanacaste-3	*Apiomerus* sp.1 (Reduviidae)	785
396LB	KR056257	166	Las Brisas	*Hyalymenus tarsatus* (Alydidae)	359
398LB	KR056258	166	Las Brisas	*Hyalymenus pulcher* (Alydidae)	359
399LB	KR056259	161	Las Brisas	*Catorhintha* sp. (Coreidae)	233
400LB	KR056260	48	Las Brisas	*Jadera aeola aeola* (Rhopalidae)	480
403LB	KR056261	2	Las Brisas	*Jadera obscura* (Rhopalidae)	513
404LB	KR056262	2	Las Brisas	*Jadera obscura* (Rhopalidae)	513
405LB	KR056263	2	Las Brisas	*Jadera obscura* (Rhopalidae)	513
411LB	KR056264	164	Las Brisas	*Hyalymenus tarsatus* (Alydidae)	384
412LB	KR056265	164	Las Brisas	*Hyalymenus pulcher* (Alydidae)	407
413LB	KR056266	166	Las Brisas	*Hyalymenus pulcher* (Alydidae)	359
415LB	KR056267	25	Las Brisas	*Leptoscelis tricolor* (Coreidae)	725
416LB-A	KR056268	166	Las Brisas	*Acanthocephala* cf. *bicoloripes* (Coreidae)	359
423LB	KR056269	142	Las Brisas	*Harpactorinae* sp. (Reduviidae)	258
424LB	KR056270	166	Las Brisas	*Leogorrus litura* (Reduviidae)	359
425LB	KR056271	41	Las Brisas	*Collaria oleosa* (Miridae)	269
426LB	KR056272	166	Las Brisas	*Collaria oleosa* (Miridae)	359
431AR	KR056273	12	Arenal	*Largus* cf. *cinctus* (Largidae)	475
432AR	KR056274	163	Arenal	*Hyalymenus tarsatus* (Alydidae)	311
433AR	KR056275	166	Arenal	*Anochrostomus formosus* (Lygaeidae)	359
434AR-A	KR056276	166	Arenal	*Hyalymenus* sp. (Alydidae)	359
434AR-B	KR056277	41	Arenal	*Hyalymenus* sp. (Alydidae)	269
435AR-A	KR056278	160	Arenal	*Collaria oleosa* (Miridae)	539
435AR-B	KR056279	41	Arenal	*Collaria oleosa* (Miridae)	269
436AR	KR056280	171	Arenal	*Collaria oleosa* (Miridae)	650
438AR	KR056281	41	Arenal	*Collaria oleosa* (Miridae)	269


TU32 L. spiculata
TU4 L. tarcoles
TU40 L. acus
TU55 C. brachyflagelli
TU45
TU58 C. napoensis
TU142
TU38 C. cf. acanthocephali
TU42
TU49 C. abscondita
TU50 C. insperata
TU41 L. barvae
TU160
TU102
TU94
TU143
TU144
TU36 L. costaricensis
TU48 L. podlipaevi
TU93
TU2 L. jaderae
TU1 L. pyrrhocoris
TU24
TU27
TU54 L. tenua
TU92
TU26 L. neopamerae
TU3 L. bifurcata
TU56 C. confusa
TU35
TU51 C. permixta
TU47
Leishmaniinae
TU131
TU161
TU5
Phytomonas
TU46
TU96
TU95
TU97
TU18
TU162-166
TU6/7
TU130
TU132
TU133
TU43
TU134
TU62
TU135
'jaculum'
TU93
TU12
TU170
TU171
TU138
TU139
TU139a
TU37
TU17
TU15
TU34
TU33
TU30
TU140
TU16
TU141
TU28
TU29
TU39
TU25
TU137
TU98
TU10
TU11
TU168
TU9 B. largi
TU22
TU99
TU167
TU23
TU44
TU101
TU21
TU14
TU21a (=TU45)
TU136
TU52
TU100
TU169
TU13
TU19
TU20
Blastocrithidia
0.5 substitutions per site

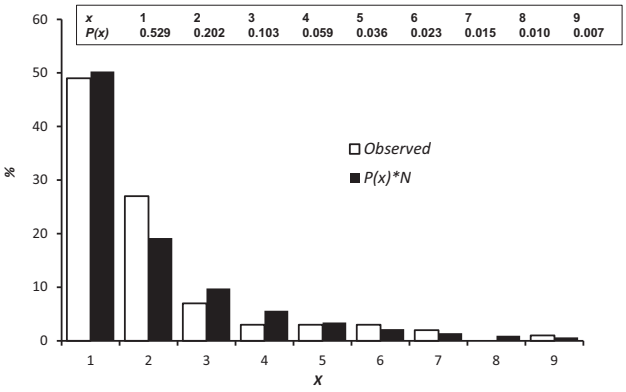


Figure 3. The distribution of trypanosomatids according to the number of parasitized host species. The X axis shows number of host species in which a single randomly selected TU from our collection can be found (***x***). The Y axis shows the actual percentage of typing units per each host category (*Observed*) or the respective percentage-based probability (***P*(*x*)**) calculated for the logarithmic distribution with $p = 0.76$ and the total number of typing units (***N***) of 95. Inset shows the calculated values of ***P*(*x*)** for each value of ***x***.

species, while the rest of parasites inhabited two or more host species including one TU that was found in nine host species.

The observed distribution of parasites was found to represent the logarithmic distribution described with the equation $P(x) = (-1)^* p^x / [x^* \ln(1-p)]$, where *P(x)* is a probability of finding a single randomly selected TU from our collection in *x* host species (*x* varies from 1 to 9), and *p* being the distribution parameter. The exact value of *p* (0.76) was determined by the method of sequential iterations (Korolyuk et al. 1985). The calculated values of ***P(x)*** for each value of ***x*** are in Figure 3 (inset). Based on these probabilities and the total number of TUs (***N*** = 95), the predicted numbers of TUs for each category of ***x*** are shown by the graph in Figure 3 as ***P(x)**N*** along with the observed numbers. The differences between the predicted and the empirical distributions are not statistically significant ($\chi^2 = 6.93$; ***P*** = 0.544; $\nu = 8$). As the parameter *p* unequivocally defines the distribution, which illustrates the degree of specificity of individual parasite TUs to their hosts, the parameter itself can serve as a numerical measure of this specificity as further discussed below.

Variations in Specificity of Host-parasite Associations among the Superfamilies of Heteroptera

The results presented above indicate that a given parasite TU is able to colonize a variety of hosts quite often, which must be due to a widespread biochemical or physiological compatibility with the hosts. We hypothesized that under the conditions when many parasites and hosts can form associations, the distribution of the former among the hosts is shaped up mainly by ecological factors, for instance those that facilitate transmission. Therefore, we looked for potential links between the hosts' feeding habits or some aspects of social behavior (e.g. propensity for aggregation) and the loads of parasite they carry. The hosts were categorized according to the major taxonomic subdivisions (superfamilies) which also allow for separating groups with different ecological characteristics: 1) Reduvioidea - arthropod-feeding (except for haematophagous Triatominae) predators (family Reduviidae); 2) Miroidea - mostly feeding on plant juices, often specific for certain host plants, some groups being phytozoophagous, zoophagous, or mycophagous (family Miridae); 3) Coreoidea - phytophagous (relatively large-size members of families Coreidae, Rhopalidae and Alydidae); 4) Lygaeoidea - mostly feeding on plants with a few taxa being predatory or haematophagous (families Lygaeidae and Rhyparochromidae); 5) Pentatomoidea - mostly plant feeding (family Pentatomidae); 6) Pyrrhocoroidea - omnivorous members of families Pyrrhocoridae and Largidae (Schuh and Slater 1995). (A few semi-aquatic gerrids, also present in our collection, were omitted from the analysis.)

For each superfamily we analyzed numbers of host species parasitized by a single TU (Table 3) and numbers of TUs found in a single host species (Table 4). In each case the data followed a logarithmic distribution, with 'goodness of fit' between the empirical and theoretical distributions evaluated by χ^2 test. Considering the host range parasitized

Figure 2. A neighbor-joining dendrogram derived from the multiple alignment of the trimmed SL RNA gene sequences (from -100 position to the 3′-end of the T-block). The tree was constructed using K2 distances and heuristic approach. The tree includes only monoxenous trypanosomatids discovered during the survey. Each TU is represented by a single sequence. Species names are shown when available. The dots indicate ancestral nodes of the marked clades. The tree is unrooted.

Table 3. Host richness analysis: numbers of trypanosomatid TUs associated with variable numbers of host species per TU.

	No. infected host species per TU									Logarithmic distribution parameters[a]				
	1	2	3	4	5	6	7	8	9	p	$\chi^2(\nu)$	χ^2_{crit}	μ	σ^2
No. TUs in each category for:														
All hosts	49	27	7	3	3	3	2	0	1	0.76	9.92 (11.5)	20.34	2.24	4.47
Reduvioidea	14	6	0	0	1					0.58	17.08 (15.0)	24.94	1.59	1.24
Miroidea	16	6	1	1	1					0.62	2.14 (6.4)	13.24	1.68	1.57
Pyrrhocoroidea	12	2	3	0	0	0	1			0.58	15.13 (12.73)	21.99	1.60	1.29
Coreoidea	26	14	0	2	0	1				0.66	14.86 (8.9)	16.72	1.78	1.99
Lygaeoidea	8	1								0.22	0.03 (1.9)	5.80	1.14	0.17
Pentatomoidea	8									-	-	-	-	-

[a]Meaning of the mathematical symbols: p - logarithmic distribution parameter; χ^2 - calculated and χ^2_{crit} - critical (5%) values for 'goodness of fit' chi-square test; μ - mathematical expectation; σ^2 - variance, ν - degrees of freedom

Table 4. Parasite richness analysis: numbers of host species associated with variable numbers of parasite TUs per host species.

	No. parasite TUs per host species								Logarithmic distribution parameters[a]				
	1	2	3	4	5	6	7	8	p	$\chi^2(\nu)$	χ^2_{crit}	μ	σ^2
No. hosts in each category for:													
All hosts	79	26	8	2	3	1	0	1	0.59	10.25 (15.11)	25.13	1.62	1.36
Reduvioidea	11	4	1	1	1				0.66	1.86 (6.80)	13.76	1.81	2.09
Miroidea	14	4	3	0	0	0	0	1	0.63	50.27 (55.32)	72.53	1.70	1.66
Pyrrhocoroidea	8	3	1	1	1	1			0.76	3.28 (9.21)	17.20	2.24	4.48
Coreoidea	28	13	3	0	1				0.64	5.33 (5.69)	12.13	1.74	1.80
Lygaeoidea	10								-	-	-	-	-
Pentatomoidea	6	1							0.28	0.05 (1.86)	5.75	1.19	0.25

[a]Mathematical designations are same as in Table 3.

by a single TU ('host diversity' analysis, Table 3) among all hosts, most trypanosomatids were found in a single host species (49 cases), whereas the remaining TUs were found to parasitize a larger number (up to 9 host species from several superfamilies for one particular TU). On average, a single trypanosomatid TU was found parasitizing 2.24 host species. Within each host group all the calculated parameters, including average numbers of parasitized hosts were similar (1.59 - 1.78), except for Lygaeoidea (1.14) but this group was among the two least sampled (with pentatomids being the other). We conclude that the parasites found within each superfamily show a similar level of specificity towards their hosts.

We next looked at the potential differences among the hosts with respect to the parasite loads those hosts carry ('parasite diversity' analysis) (Table 4). The table shows that in most cases (79 out of 120) a single host species carries a single parasite TU, but two TUs per host were not uncommon (seen in 26 host species), and in one host species there were as many as eight TUs. On average, among all groups, a single infected host carries 1.62 TUs. Among the individual superfamilies, the omnivorous hosts (Pyrrhocoroidea) had the highest diversity of parasites ($\mu = 2.24$), followed by predatory reduviids ($\mu = 1.81$), with smaller numbers of parasites found in phytophagous hosts. However, according to Student's *t*-test, the pairwise differences were not significant in any of those cases (Table 5), although the differences for Pyrrhocoroidea – Coreoidea and Pentatomoidea – Coreoidea ($\alpha = 0.285$ and $\alpha = 0.290$, respectively) were the closest to the 0.05 threshold value for significance. These results indicate that the parasite load of pyrrhocorids is more diverse compared to the remaining analyzed groups of heteropterans. Validity of these tentative conclusions may benefit from increasing the sampling size.

Table 5. Values of Student's *t*-test criterion for comparison of mathematical expectations of the number of trypanosomatid TUs per host species between the superfamilies of Heteroptera.

	Reduvioidea	Miroidea	Pyrrhocoroidea	Coreoidea
Miroidea	0.24 (38)[a]			
Pyrrhocoroidea	0.70 (31)	0.97 (35)		
Coreoidea	0.18 (61)	0.11 (65)	1.08 (58)	
Pentatomoidea	1.10 (23)	1.02 (27)	1.29 (20)	1.07 (50)

[a]Number of degrees of freedom is shown in parentheses.

Differences in the Shared Parasite TUs among the Host Groups

Next, we compared the composition of parasite loads among the groups. The aim of this analysis was to find out if there is (are) a pair(s) of superfamilies that would share a relatively high number of parasites, a situation which would suggest special interactions between such groups. In particular, we intended to verify the hypothesis that predatory hosts (reduviids) acquire some of their parasites from their prey, namely phytophagous mirids and/or other insects. To this end, the sets of TUs found in the host groups were pairwise compared using the Jaccard index (JI) (Levandowski and Winter 1971). This numerical index represents a fraction of shared TU's out of a total number ***m*** of TUs found in two host groups under comparison. It can vary from 0 (no overlap in sets of TUs in two hosts) to 1 (a complete overlap). The results of the pairwise comparisons are presented in Table 6. Three pairs involving reduviids stood out by the level of similarity in the faunas of their parasites: that with Miroidea (Red-Mir, JI = 0.12), Pyrrhocoroidea (Red-Pyrr, 0.15), and Coreoidea (Red-Cor, 0.12). The fourth such pair was Coreiodea - Miroidea (Cor-Mir, 0.16). Since we expected the JI values to be influenced by sampling size (***m***), a regression analysis was used to investigate this relationship (data not shown). Confidence intervals were then calculated for three levels of significance: 0.05 (Fig. 4), 0.01 and 0.001 (data not shown). The graph reflects a general trend toward increasing JI values with an increase of the total number of parasite TUs in two groups under comparison (Fig. 4). While most of the JI values did fall within the confidence intervals of regression, those for Red-Mir and Red-Pyrr were found outside. The respective JI values were, therefore, significantly higher (99.9% probability) than those anticipated for the respective sampling size. The Cor-Mir pair was found to be borderline for the 95% confidence band. Interestingly, two pairs (Mir-Pyrr and Cor-Pent) showed a lower than expected number of shared TUs.

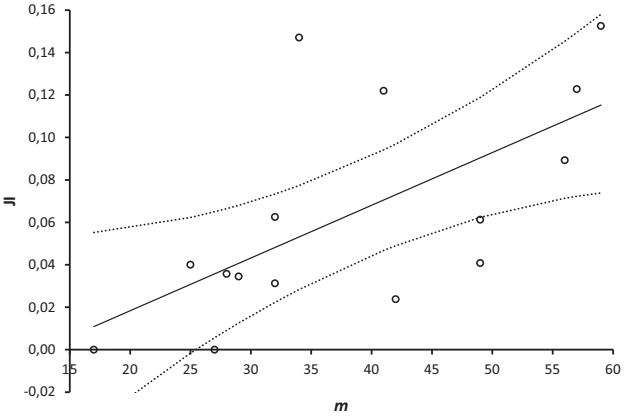


Figure 4. Linear regression analysis of Jaccard similarity coefficients for the trypanosomatid typing units shared by pairs of Heteroptera superfamilies. Total number (***m***) of TUs in a pair of compared host taxa is shown along the X axis. Jaccard index (JI) is plotted along the Y axis. Expected values of JI for a given *m* value are represented by a solid line. The dotted lines show the confidence interval boundaries (95% significance level).

Discussion

The Currently Known Segment of Trypanosomatid Diversity

This study analyzes data assembled in frame of a multi-year investigation of the trypanosomatid diversity in Heteroptera from several localities in the Neotropics. To the 51 TUs found earlier (Maslov et al. 2007; Westenberger et al. 2004; Yurchenko et al. 2006a, 2008, 2009) the analysis presented here adds 27 new TUs collected in Ecuador in 2008 and 17 new TUs discovered later in Costa Rica (Table 1). The decreasing increments reflect the fact that many previously found TUs were encountered repeatedly. The total number of host species examined for trypanosomatids was ~350 with 124 (nearly 30%) of them being positive for trypanosomatid infection. This number is certainly an underestimate of the real proportion

Table 6. Jaccard similarity index (JI) for the parasitofaunas from the superfamilies of Heteroptera.

	Reduvioidea	Miroidea	Pyrrhocoroidea	Coreoidea	Lygaeoidea
Miroidea	0.12	-			
Pyrrhocoroidea	0.15	0.02	-		
Coreoidea	0.12	0.15	0.09	-	
Lygaeoidea	0.03	0.06	0.00	0.06	-
Pentatomoidea	0.04	0.03	0.04	0.04	0.00

of competent host species due to low prevalence of some infections. Most of the analyzed insects were from common species that were collected by sweep netting or by attraction to light traps. Therefore, hosts with different ecology, seasonality and behavior may likely have eluded the survey. The species richness of Neotropical Heteroptera is understudied and the species numbers inferred from the published sources (600 to 700 species in Costa Rica (Champion 1897; Distant 1880) and 584 species in Ecuador (Froescher 1981)) are underestimated. For example, there are at least 143 species belonging to the family Reduviidae alone in Costa Rica (C. Weirauch, pers. commun.) and out of these, we sampled only 35 species, finding ten species infected with at least eight trypanosomatid TUs. Considering the number of new species described every year from Neotropics we may estimate that there are more than 1000 species of Heteroptera in Costa Rica, and more than two thousands in Ecuador. These numbers indicate that there is still a large potential for the discoveries of new trypanosomatids. The total number of trypanosomatid TUs so far discovered from those regions is 95 from 120 host species (in a few infected hosts the parasites were not genotyped). Out of the found TUs just a few were cultivable; several new trypanosomatid species have been described based on those cultures (Jirků et al. 2012; Maslov et al. 2010; Votýpka et al. 2012b; Yurchenko et al. 2006b, 2008, 2009). The question remains open how many of the new trypanosomatid TUs discovered represent separate species. SL sequences only serve as initial guides on the path of new species discovery, while a broader range of molecular, morphological and other characters has to be evaluated for validation of a TU as a new species (Maslov et al 2013).

Not only in Reduviidae but across the entire spectrum of samples the number of trypanosomatid TUs discovered so far (95) has been lagging behind the number of respective host species (120), although not by far. As discussed below, there are cases when a single TU was shared by several species, but those are balanced by several TUs encountered in a single host species. Using the logarithmic distribution to describe the parasite-host associations, on average there were 2.24 host species per single TU and 1.62 TUs per infected host species.

Large-scale Dispersal Patterns of Trypanosomatids

A substantial number (21) of trypanosomatid TUs were found in both countries indicating a potentially broad dispersal in the Neotropics and probably beyond (at least in some cases). Thus, TU*48* (= *Leptomonas podlipaevi*) is also found in California (Yurchenko et al. 2006a), and TU1 (= *Leptomonas pyrrhocoris*) has been shown to occur globally (Votýpka et al. 2010, 2012a, b). The still fragmentary nature of the data suggests that the number of broadly distributed TUs may be relatively high. For example, although TU*28* was found only in Ecuador, it was encountered in four different biogeographic regions (the Amazonian lowland, the Andean foothills and the cloud forests on the western and eastern Andean slopes) and its hosts included members of three heteropteran families (Coreidae, Pentatomidae and Miridae). The occurrence of TU*28* in such a diverse range of ecological settings indicates that this trypanosomatid is indeed broadly distributed in the Neotropics.

On the other side, endemism can be tentatively inferred in those cases when the geographic distribution or the host range (or both) of a trypanosomatid TU are limited. Thus, TU*13* has been repeatedly found in three different hosts (Largidae, Miridae and Reduviidae) but only in one locality in Costa Rica. However, the question whether or not monoxenous trypanosomatid species can be truly endemic still remains open until a more thorough investigation of this issue.

On the Specificity of Host–parasite Associations

The apparent departure of host-parasite relationships from the ‘one host – one parasite’ scenario dictates the necessity of finding the appropriate

means of describing these relationships. We found that the observed distribution of parasites among the host species can be well approximated by logarithmic distribution (Fig. 3). It shows the probability ***P(x)*** of finding a single trypanosomatid TU in ***x*** number of host species. In our case the probability of finding a parasite in one host only was 0.529 and in two hosts 0.202, decreasing further for larger numbers of hosts. The shape of this distribution is mathematically defined by parameter *p* and, from the biological perspective, reflects the specificity of parasites for their hosts. For example, a very steep decline of the curve would be observed if most or all parasites are found in one host species reflecting very high specificity. Such a distribution would be defined by *p* parameter values approaching 0. If, on the other hand, trypanosomatids were rather unspecific with respect to their hosts, then the distribution would be more broad and uniform, and the *p* value would approach 1. Thus, a value of the *p* parameter can be regarded as a numerical measure of the specificity of parasites to hosts. This is based on the assumption that the observed distribution, approximated with the respective logarithmic distribution, accurately reflects the real occurrence of parasites in their hosts. The insufficient sampling size and/or low prevalence of parasites do not directly affect *p* values, but can do this indirectly by causing distorted distributions when the existing host-parasites associations become overlooked.

The observed distribution and, therefore, the parameter *p* of the best approximating distribution represent an integration of all factors which can affect the ability of trypanosomatids to colonize hosts and successfully spread in populations. These include physiological and/or biochemical compatibility between the sets of hosts and parasites which inhabit the area, and also those factors that affect transmission from one infected host to another. Moreover, they also include factors related to hosts' feeding habits, their cohabitation with other species, propensity to aggregate in certain areas, etc. As any or all of these would change in a different area, the parameter *p* would also be different. Moreover, in a second major host group – Diptera, the parameter *p* would also be different even for the same area.

A disproportionally high number of shared parasites between certain host groups (such as reduviids and various phytophagous species, or between coreids and mirids) was revealed in this study. Such sharing might suggest the existence of special conditions which favor the exchange or transfer of parasites. For the pair including reduviids these conditions likely involve predation. However, the high JI value for the coreids-mirids pair indicates that predation is not an exclusive pre-requisite for sharing parasites. A broader hypothesis such as the one based on sharing the exact ecological niche and involving particular aspects of the hosts' natural history, e.g. feeding habits, common food sources etc., might be better suited to explain this phenomenon.

There is a growing understanding in parasitology that the concept of host specificity is not restricted to just number of parasitized host species (basic specificity) but also takes into account differences in prevalence of infection among the hosts, their phylogenetic relationships and variations in host range and prevalence across the geographic space (Krasnov et al. 2011; Poulin et al. 2011). The work presented herein should be regarded as one of the first steps towards providing an objective measure to host specificity for protist parasites using monoxenous trypanosomatids as a model system. Although only the most basic specificity concept was used throughout the analyses, we approached the host specificity problem from a new angle. In this work we applied a statistical approach to search for the numerical parameter that would characterize an entire group of parasites rather than properties of individual species. In the future it should be possible to use variations of this parameter (*p* of logarithmic distribution) to compare different parasite-host associations. Furthermore, the analysis of variations in parasite richness per host species or the reciprocal analysis of host richness per parasite TU has provided the first-ever estimates of the host specificity variation across the major host taxa and that can also be extended to additional hosts and geographic areas. In the future, with accumulation of additional data on host-parasite associations, supplemented with analyses of the host phylogeny, it should become possible to address the other facets of the trypanosomatid host specificity as well.

Methods

Collecting of insects: Insect have been collected in Ecuador in June-July of 2008 and Costa Rica in September 2009, June 2010 and April 2012. A summary of the collection data is given in Tables 1 and 2. Several of the collection locales were the same as specified in our previous reports (Maslov et al. 2007; Westenberger et al. 2004); those in Ecuador were: Mindo and Napo-1; and in Costa Rica were: Arenal, Monteverde, and Tarcoles (previously designated Carara-1). Additional collection locales in Ecuador were: Atacames (province Esmeraldas, near the city of Atacames, S 00° 52′ 31″, W 79° 50′ 32″), Baños (province Tungurahua, near the city of Baños, S 01° 24′ 14″, W 078° 25′ 56″), Limón (province Morona Santiago, near the

city of Limón, S 02° 58′ 14″, W 78° 25′ 21″), Loja (province Loja, Loja Botanical Garden, S 04° 02′ 05″, W 079° 11′ 56″), Otongachi (province Santo Domingo, 33 km W of the city of Santo Domingo, Otongachi reserve, S 00° 19′ 00″, W 78° 56′ 52″), Vilcabamba (province Loja, near the town of Vilcabamba, Rumi Wilco reserve, S 04° 15′ 22″, W 079° 13′ 06″); and in Costa Rica were: Guanacaste-3 (province Guanacaste, near the southern boundary of PN Rincon de la Vieja, N 10° 45′ 15′, W 85° 20′ 58″), Guanacaste-4 (province Guanacaste, PN Santa Rosa, N 10° 50′ 06″, W 85° 37′ 31″), Guanacaste-5 (province Guanacaste, near the town of Santa Cecilia, N 11° 02′ 19″, W 85° 23′ 59″), Las Brisas (province Limon, 5 km S of the town of Pocora, Las Brisas reserve, N 10° 08′ 03″, W 83° 36′ 18″). Most insects were collecting by sweep-netting, some were collected by attraction to light or hand-picking. Insects were kept alive in individual vials and usually dissected on the same day.

Field treatment of samples: Insect were surface-sterilized with ethanol and dissected; gut content was suspended in 1 x SSC (0.15 M NaCl, 0.015 M sodium citrate). Live material in the smears was inspected using light microscopy with 400-fold magnification and phase contrast. Post-dissection remains (xenotypes) and intact specimens from same populations were preserved for subsequent identification and final deposition in the collection of National Museum in Prague (Czech Republic). If trypanosomatids were detected in the intestinal smears, the material was transferred into a vial containing the buffer solution (1 mL) for preservation of DNA (1% SDS, 100 mM EDTA, pH 8.0) (Maslov et al. 2007; Westenberger et al. 2004). A fraction of the gut smear material was inoculated in 1 mL of cultivation medium (below) to establish primary cultures. Samples and primary cultures were stored at the ambient temperature for the duration of a trip (1-3 weeks) prior to transfer to the laboratory.

Treatment of cultures: If primary cultures showed evidence of trypanosomatid propagation the subsequent cultivation was attempted in various media including Brain Heart Infusion medium and M199 medium (both from Life Technologies, Carlsbad, USA) supplemented with 10 μg/ml hemin, 10% heat-inactivated fetal bovine serum (Atlanta Biologicals, Lawrenceville, USA) and a cocktail of antibiotics (Maslov et al. 2010; Westenberger et al. 2004). Axenic cultures were established as described previously (Yurchenko et al. 2009).

DNA amplification and analysis: DNA contained in the intestinal lysate or from cells in cultures was purified using PureLinkTM Genomic DNA kit (Life Technologies). PCR amplification and sequencing of Spliced Leader (SL) RNA gene repeats was conducted as described previously (Jirků et al. 2012; Maslov et al. 2007, 2010; Westenberger et al. 2004). The genotyping was performed using multiple alignment by Clustal-X, ver. 2.0 (Larkin et al. 2007) and neighbor-joining cluster analysis with K2P distances by PAUP* 4.0, beta version (Swofford 1998). The threshold of 90% sequence identity was used to separate typing units (Maslov et al. 2007). The GenBank™ accession numbers of the new sequences determined in the course of this study (KP717764 - KP717893, KR056219- KR056281), along with the previously determined sequences, are given in Table 2.

Statistical analysis: Logarithmic distribution was used for approximating the empirical distributions for the number of host species parasitized by a single trypanosomatid TU and the number of TUs parasitizing a single host species. Logarithmic distribution parameter p was estimated by sequential iterations. The goodness of fit between the empirical distribution and theoretical expectation was determined using χ^2 -test corrected for small expectation values (Nass 1959). The expectation μ and variance σ^2 for the logarithmic distribution were determined analytically using the equations correlating those and the parameter p (Korolyuk et al. 1985). Pairwise comparison of mathematical expectations was performed using the Student's t-test. Numbers of shared TUs among host superfamilies were evaluated using Jaccard similarity index (JI) which represents a fraction of the shared TUs out of the total number of TUs found in two host groups analyzed (Levandowski and Winter 1971). Linear regression analysis was used to calculate an expected JI value for each sampling size. Confidence intervals (CI) of linear regression were calculated for three levels of significance (95%, 99% and 99.9%) using STATISTICA, ver. 7.0 (StatSoft). Values of JI found outside of CI were considered significantly different from those expected for respective sampling sizes.

Acknowledgements

We are indebted to the colleagues in Costa Rica and Ecuador who made the field work possible by helping with permits, providing access to collection sites and helping with logistics in many other ways: G. Onore, R. Zeledón, E. Berlin, R. Vargas, R. Blanco, F. Campos. We also thank H. Brailovsky, P.-P. Chen, F. Chérot, D. Chłond, L. R. Davranoglou, D. Forero, N. Nieser, M. Schwartz, J. L. Stehlík and Ch. Weirauch for help with host identifications, W. Suen for help with the sequencing, and T. Yurchenko and T. Portnaya for help with statistical terms. This work was supported by UCR Academic Senate grants to D.A.M., the grant "Dynamics and conservation of gene pools" of the Fundamental Research Program of the Presidium of the Russian Academy of Sciences to E.K., by Praemium Academiae award to J.L., by Ministry of Culture of the Czech Republic grant DKRVO 2014/13 to P.K., by the Czech Science Foundation grants P305/14/23986S to J. L. and P506/13/24983S to V.Y. who is also supported by the funds of the Moravskoslezský Kraj research initiative.

References

Bulat SA, Mokrousov IV, Podlipaev SA (1999) Classification of trypanosomatids from insects and plants by the UP-PCR (universally-primed PCR) technique and cross dot blot hybridization of PCR products. Eur J Protistol **35**:319–326

Carvalho ALM, Deane MP (1974) Trypanosomatidae isolated from *Zelus leucogrammus* (Perty, 1834) (Hemiptera, Reduviidae), with a discussion on flagellates of isectivorous bugs. J Protozool **21**:5–8

Champion GC (1897-1901) Insecta. Rhynchota. Hemiptera-Heteroptera. In Godman FD, Salvin O (eds) Biologia Centrali-Americana. Volume II. Published for the editors by R. H. Porter, London.

Chandler JA, James PM (2013) Discovery of trypanosomatid parasites in globally distributed *Drosophila* species. PLoS ONE **8**:e61937

Distant WL (1880-1893) Insecta. Rhynchota. Hemiptera-Heteroptera. In Godman FD, Salvin O (eds) Biologia Centrali-Americana. Volume I. Published for the editors by Dulau & Co., London.

Froescher RC (1981) Heteroptera or True Bugs of Ecuador: a Partial Catalog. Smithsonian Contributions to Zoology, no. 322. Smithsonian Institute Press, Washington

Hoare CA (1972) The Trypanosomes of Mammals. A Zoological Monograph. Blackwell, Oxford, 750 p

Jirků M, Yurchenko VY, Lukeš J, Maslov DA (2012) New species of insect trypanosomatids from Costa Rica and the proposal for a new subfamily within the Trypanosomatidae. J Eukaryot Microbiol **59**:537–547

Korolyuk VS, Portenko NI, Skorokhod AV, Turbin AF (1985) Handbook on Probability Theory and Mathematical Statistics (In Russian). Nauka, Moscow, 640 p.

Kostygov AY, Grybchuk-Ieremenko A, Malysheva MN, Frolov AO, Yurchenko V (2014) Molecular revision of the genus *Wallaceina*. Protist **165**:594–604

Krasnov BR, Mouillot D, Shenbrot GI, Khokhlova IS, Poulin R (2011) Beta-specificity: the turnover of host species in space and another way to measure host specificity. Int J Parasitol **41**:33–41

Lainson R, Shaw JJ (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

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2. 0. Bioinformatics **23**:2947–2948

Levandowski M, Winter D (1971) Distance between sets. Nature **234**:34–35

Maslov DA, Votýpka J, Yurchenko V, Lukeš J (2013) Diversity and phylogeny of insect trypanosomatids: all that is hidden shall be revealed. Trends Parasitol **29**:43–52

Maslov DA, Yurchenko VY, Jirků M, Lukeš J (2010) Two new species of trypanosomatid parasites isolated from Heteroptera in Costa Rica. J Eukaryot Microbiol **57**:177–188

Maslov DA, Westenberger SJ, Xu X, Campbell DA, Sturm NR (2007) Discovery and barcoding by analysis of Spliced Leader RNA gene sequences of new isolates of Trypanosomatidae from Heteroptera in Costa Rica and Ecuador. J Eukaryot Microbiol **54**:57–65

Merzlyak E, Yurchenko V, Kolesnikov AA, Alexandrov K, Podlipaev SA, Maslov DA (2001) Diversity and phylogeny of insect trypanosomatids based on small subunit rRNA genes: Polyphyly of *Leptomonas* and *Blastocrithidia*. J Eukaryot Microbiol **48**:161–169

Nascimento MT, Garcia MC, da Silva KP, Pinto-da-Silva LH, Atella GC, Motta MC, Saraiva EM (2010) Interaction of the monoxenic trypanosomatid *Blastocrithidia culicis* with the *Aedes aegypti* salivary gland. Acta Trop **113**: 269–278

Nass CAG (1959) The χ^2 test for small expectations in contingency tables, with special reference to accidents and absenteeism. Biometrika **46**:365–385

Podlipaev SA (1990) Catalogue of the world fauna of the family Trypanosomatidae (Protozoa). Proc Zool Inst USSR Acad Sci (In Russian) **144**:1–178

Podlipaev SA (2003) Host specificity of homoxenic trypanosomatids. Parasitologia (In Russian) **37**:3–17

Podlipaev SA, Sturm NR, Fiala I, Fernandes O, Westenberger SJ, Dollet M, Campbell DA, Lukeš J (2004) Diversity of insect trypanosomatids assessed from the spliced leader RNA and 5S rRNA genes and intergenic regions. J Eukaryot Microbiol **51**:283–290

Poulin R, Krasnov BR, Mouillot D (2011) Host specificity in phylogenetic and geographic space. Trends Parasitol **27**:355–361

Ramos A, Maslov DA, Fernandes O, Campbell DA, Simpson L (1996) Detection and identification of human pathogenic *Leishmania* and *Trypanosoma* species by hybridization of PCR-amplified mini-exon repeats. Exp Parasitol **82**:242–250

Romeiro A, Monteiro Leal LH, de SW, Attias M (2003a) Interaction of *Leptomonas wallacei* with the intestinal tract of its natural host *Oncopeltus fasciatus* (Hemiptera: Lygaeidae). J Invertebr Pathol **82**:41–49

Romeiro A, Saraiva EM, de SW, Attias M (2003b) *Leptomonas wallacei* shows distinct morphology and surface carbohydrates composition along the intestinal tract of its host *Oncopeltus fasciatus* (Hemiptera: Lygaeidae) and in axenic culture. J Eukaryot Microbiol **50**:409–416

Sacks DL (2001) *Leishmania*-sand fly interactions controlling species-specific vector competence. Cell Microbiol **3**:189–196

Schaub GA, Boker CA (1986) Scanning electron microscopic studies of *Blastocrithidia triatomae* (Trypanosomatidae) in the rectum of *Triatoma infestans* (Reduviidae). J Protozool **33**:266–270

Schaub GA, Neukirchen K, Golecki I (1992) Attachment of *Blastocrithidia triatomae* (Trypanosomatidae) by flagellum and cell body in the midgut of the reduviid bug *Triatoma infestans*. Eur J Protistol **28**:322–328

Schwarz RS, Bauchan GR, Murphy CA, Ravoet J, de Graaf DC, Evans JD (2015) Characterization of two species of Trypanosomatidae from the honey bee *Apis mellifera*: *Crithidia mellificae* Langridge and McGhee, and *Lotmaria passim* n. gen., n. sp. J Eukaryot Microbiol **62**:567–583

Schuh RT, Slater JA (1995) True Bugs of the World (Hemiptera: Heteroptera). Classification and Natural History. Comstock Publishing Associates, Cornell University Press, Ithaca and London, 336 p

Swofford DL (1998) PAUP* 4. 0: Phylogenetic Analysis Using Parsimony (and Other Methods), beta version. Sinauer Associates, Inc., Sunderland, MA

Teixeira MMG, Borghesan TC, Ferreira RC, Santos MA, Takata CS, Campaner M, Nunes VL, Milder RV, de SW, Camargo EP (2011) Phylogenetic validation of the genera *Angomonas* and *Strigomonas* of trypanosomatids harboring bacterial endosymbionts with the description of new species of trypanosomatids and of proteobacterial symbionts. Protist **162**:503–524

Turco SJ, Sacks DL (2003) Control of *Leishmania*-sand fly interactions by polymorphisms in lipophosphoglycan structure. Methods Enzymol **363**:377–381

Týč J, Votýpka J, Klepetková H, Suláková H, Jirků M, Lukeš J (2013) Growing diversity of trypanosomatid parasites of flies (Diptera: Brachycera): frequent cosmopolitism and moderate host specificity. Mol Phylogenet Evol **69**:255–264

Vickerman K (1976) The Diversity of the Kinetoplastid Flagellates. In Lumsden WHR, Evans DA (eds) Biology of the Kinetoplastida. Academic Press, London, pp 1–34

Votýpka J, Klepetková H, Jirků M, Kment P, Lukeš J (2012a) Phylogenetic relationships of trypanosomatids parasitising true bugs (Insecta: Heteroptera) in sub-Saharan Africa. Int J Parasitol **42**:489–500

Votýpka J, Klepetková H, Yurchenko VY, Horák A, Lukeš J, Maslov DA (2012b) Cosmopolitan distribution of a trypanosomatid *Leptomonas pyrrhocoris*. Protist **163**:616–631

Votýpka J, Maslov DA, Yurchenko V, Jirků M, Kment P, Lun ZR, Lukeš J (2010) Probing into the diversity of trypanosomatid flagellates parasitizing insect hosts in South-West China reveals both endemism and global dispersal. Mol Phylogenet Evol **54**:243–253

Votýpka J, Suková E, Kraeva N, Ishemgulova A, Duží I, Lukeš J, Yurchenko V (2013) Diversity of trypanosomatids (Kinetoplastea: Trypanosomatidae) parasitizing fleas (Insecta: Siphonaptera) and description of a new genus *Blechomonas* gen. n. Protist **164**:763–781

Wallace FG (1966) The trypanosomatid parasites of insects and arachnids. Exptl Parasitol **18**:124–193

Westenberger SJ, Sturm NR, Yanega D, Podlipaev SA, Zeledon R, Campbell DA, Maslov DA (2004) Trypanosomatid biodiversity in Costa Rica: genotyping of parasites from Heteroptera using the Spliced Leader RNA gene. Parasitology **129**:537–547

Wilfert L, Longdon B, Ferreira AG, Bayer F, Jiggins FM (2011) Trypanosomatids are common and diverse parasites of *Drosophila*. Parasitology **138**:858–865

Yurchenko V, Lukeš J, Xu X, Maslov DA (2006a) An integrated morphological and molecular approach to a new species description in the Trypanosomatidae: the case of *Leptomonas podlipaevi* n. sp., a parasite of *Boisea rubrolineata* (Hemiptera: Rhopalidae). J Eukaryot Microbiol **53**:103–111

Yurchenko VY, Lukeš J, Jirků M, Maslov DA (2009) Selective recovery of the cultivation-prone components from mixed trypanosomatid infections: a case of several novel species isolated from Neotropical Heteroptera. Int J Syst Evol Microbiol **59**:893–909

Yurchenko VY, Lukeš J, Jirků M, Zeledón R, Maslov DA (2006b) *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

Yurchenko VY, Lukeš J, Tesařová M, Jirků M, Maslov DA (2008) Morphological discordance of the new trypanosomatid species phylogenetically associated with the genus *Crithidia*. Protist **159**:99–114