



A tsetse and tabanid fly survey of African great apes habitats reveals the presence of a novel trypanosome lineage but the absence of *Trypanosoma brucei*[☆]



Jan Votýpka^{a,b,*,1}, Jana Rádrová^{a,b,1}, Tomáš Skalický^{b,c,1}, Milan Jirků^{b,c}, Dagmar Jirsová^{c,d}, Andrei D. Mihalca^e, Gianluca D'Amico^e, Klára J. Petřelková^{b,f}, David Modrý^{b,d,g}, Julius Lukeš^{b,c,h,*}

^a Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic

^b Institute of Parasitology, Biology Centre, České Budějovice (Budweis), Czech Republic

^c Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), Czech Republic

^d Central European Institute of Technology, University of Veterinary and Pharmaceutical Sciences, Brno, Czech Republic

^e Department of Parasitology and Parasitic Diseases, Faculty of Veterinary Medicine, University of Agricultural Sciences and Veterinary Medicine, Cluj-Napoca, Romania

^f Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic

^g Faculty of Veterinary Medicine, University of Veterinary and Pharmaceutical Sciences, Brno, Czech Republic

^h Canadian Institute for Advanced Research, Toronto, Canada

ARTICLE INFO

Article history:

Received 5 February 2015

Received in revised form 17 June 2015

Accepted 22 June 2015

Available online 26 July 2015

Keywords:

Trypanosoma

Tsetse

Tabanids

African great apes

Gorillas

Transmission

Bloodmeal

Feeding preference

ABSTRACT

Tsetse and tabanid flies transmit several *Trypanosoma* species, some of which are human and livestock pathogens of major medical and socioeconomic impact in Africa. Recent advances in molecular techniques and phylogenetic analyses have revealed a growing diversity of previously unidentified tsetse-transmitted trypanosomes potentially pathogenic to livestock and/or other domestic animals as well as wildlife, including African great apes. To map the distribution, prevalence and co-occurrence of known and novel trypanosome species, we analyzed tsetse and tabanid flies collected in the primary forested part of the Dzanga-Sangha Protected Areas, Central African Republic, which hosts a broad spectrum of wildlife including primates and is virtually devoid of domestic animals. Altogether, 564 tsetse flies and 81 tabanid flies were individually screened for the presence of trypanosomes using 18S rRNA-specific nested PCR. Herein, we demonstrate that wildlife animals are parasitized by a surprisingly wide range of trypanosome species that in some cases may circulate via these insect vectors. While one-third of the examined tsetse flies harbored trypanosomes either from the *Trypanosoma theileri*, *Trypanosoma congolense* or *Trypanosoma simiae* complex, or one of the three new members of the genus *Trypanosoma* (strains 'Bai', 'Ngbanda' and 'Didon'), more than half of the tabanid flies exclusively carried *T. theileri*. To establish the putative vertebrate hosts of the novel trypanosome species, we further analyzed the provenance of blood meals of tsetse flies. DNA individually isolated from 1033 specimens of *Glossina* spp. and subjected to high-throughput library-based screening proved that most of the examined tsetse flies engorged on wild ruminants (buffalo, sitatunga, bongo), humans and suids. Moreover, they also fed (albeit more rarely) on other vertebrates, thus providing indirect but convincing evidence that trypanosomes can be transmitted via these vectors among a wide range of warm- and cold-blooded hosts.

© 2015 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

[☆] Note: Nucleotide sequence data reported in this paper are available in GenBank under accession numbers KP979582–4, KP979585–7, KR024686, KP307018–26, KR024687, KR024683–5, KR024688, SAMN03703425, SAMN03703430, SAMN03703434, SAMN03703438, SAMN03703509, SAMN03703516–8.

* Corresponding authors at: Department of Parasitology, Charles University, Viničná 7, 128 44 Prague, Czech Republic. Tel.: +420 221951826 (J. Votýpka). Biology Centre, Branišovská 31, 37005 České Budějovice, Czech Republic. Tel.: +420 387775416 (J. Lukeš).

E-mail addresses: vapid@natur.cuni.cz (J. Votýpka), jula@paru.cas.cz (J. Lukeš).

¹ These authors contributed equally to this work.

1. Introduction

The first scientific description of African trypanosomes was made at the end of the 19th century and the importance of the tsetse fly in transmission was discovered soon afterwards (Sharma et al., 2009). African trypanosomes are well-known causative agents of serious diseases in humans and livestock. A fatal human sleeping sickness occurs in sub-Saharan Africa and

approximately 60 million people are exposed to the risk of infection, with more than 10,000 new cases reported annually (Simarro et al., 2011). Nagana, surra, and other animal trypanosomiasis constitute such a serious problem for farming that they largely exclude productive breeds of livestock from parts of the humid and semi-humid zones of sub-Saharan African countries and annually cause approximately three million cattle deaths (Schofield and Kabayo, 2008).

The group of most significant African trypanosome species includes *Trypanosoma congolense*, *Trypanosoma vivax*, *Trypanosoma simiae* and five members of the *Trypanosoma brucei* complex. The diversity of African trypanosomes is much greater, however, and new species are continually described (Adams et al., 2006, 2010a; Hamilton et al., 2008). Nevertheless, life cycles, vertebrate hosts and vectors of many neglected and newly described trypanosome species remain unknown. It is clear that our knowledge of the diversity of African trypanosomes is limited and understanding their true diversity will require broader surveys.

Taxonomic designations of these diverse flagellates, mostly lacking species-specific morphological characters, are mostly based on their phylogenetic placements (Gibson, 2007). Moreover, trypanosome species may differ by only minor genetic differences, which reflect adaptations to particular parasitic niches. Recently, this has been exemplified by the case of *Trypanosoma evansi*, the life cycle of which is significantly influenced by just a few tiny genetic differences from *Trypanosoma brucei brucei* ancestral stock (Lai et al., 2008; Carnes et al., 2015).

Recent advances in molecular identification techniques and phylogenetic analyses have revealed previously unidentified trypanosome species in African vertebrates (Auty et al., 2012a) and tsetse flies (Adams et al., 2006, 2010a; Hamilton et al., 2008; Auty et al., 2012b). Due to the paucity of distinguishing morphological features, molecular tools comprise virtually the only way to map the high diversity of trypanosomes, yet only rarely are new forms formally described and named. Consequently, a growing number of nameless species represents an inconvenient complication, as these are largely ignored (Adams et al., 2010a).

While the mere presence of trypanosome DNA in bloodsucking flies cannot be considered proof of transmission by these vectors, it reveals their occurrence in the blood of vertebrates inhabiting the environment, in which collections occurred. New trypanosome species detected in tsetse flies could originate from wild vertebrates that are not yet known to host trypanosomes (Adams et al., 2010a) and, vice versa, undescribed trypanosomes found in wildlife could be transmitted by tsetse flies (Auty et al., 2012b). To track the entire life cycle and to reveal the bloodsucking vectors responsible for the transmission of these flagellates under natural conditions is time- and money-consuming and in some cases almost impossible. Nevertheless, the involvement of specific vertebrate hosts and invertebrate transmitters in the trypanosome life cycle could be uncovered by analyzing the blood meals and feeding patterns of possible vectors (Muturi et al., 2011). Not only do analyses of blood meals of tsetse flies trapped in their natural environments provide information about transmitted pathogens, but these hematophagous flies could also be used as “vertebrate samplers” and deliver DNA-based information on many aspects of vertebrate ecology (Calvignac-Spencer et al., 2013).

A great deal of African trypanosome biodiversity is probably still hiding in African savannahs and forests. Recent findings demonstrate that trypanosomes in and outside of the *T. brucei* complex can infect vertebrate hosts such as chimpanzees and other great apes that were previously believed to be more or less trypanosome-free (Jirků et al., 2015). To the best of our knowledge, however, no survey for trypanosomes in tsetse flies has been undertaken directly within habitats inhabited by African great

apes. In this study, tsetse and tabanid flies captured in the Dzanga-Sangha Protected Areas (DSPA), Central African Republic (CAR), were examined to determine the diversity and prevalence of harbored trypanosomes and to survey the spectrum of vertebrates from which those flies took their blood meals.

2. Materials and methods

2.1. Study site

The DSPA in the CAR include zones with various levels of protection and restricted human access, as well as multiple-use zones with controlled human activities. The human population density in the DSPA is low (ca one person per km²). The forested areas are composed of secondary and, less frequently, primary forest. Rainfall averages 1400 mm/year. Dry months typically occur between December and February, while the rest of the year has a long rainy season (albeit with low precipitation during June and July). Our sampling was carried out in the Dzanga Sector of the Dzanga-Ndoki National Park (2°50'N, 16°28'E) and its surroundings. Bai Hokou is a permanent Primate Habituation Programme research camp with local BaAka and Bantu trackers and research assistants, and up to five foreign volunteers and/or researchers. The surrounding area is covered by primary forest with several open spaces (forest glades), locally known as “bais”. Two of these, Bai-Hokou (BH; 2°51'30.6"N, 16°28'12.6"E, in the proximity of the research camp) and Bai-Gubunga (BG; 2°50'59.9"N, 16°28'05.1"E, 3 km distant from the research camp), were chosen as the main collection sites of bloodsucking insects. Both locations are frequently visited by herds of forest buffalos (*Syncerus caffer nanus*), together with bongos (*Tragelaphus eurycerus*), sitatungas (*Tragelaphus spekii*) and forest elephants (*Loxodonta cyclotis*). The surrounding primary forest is inhabited by several duiker species (e.g., *Cephalophus silvicultor*), red river hogs (*Potamochoerus porcus*), giant forest hogs (*Hylochoerus meinertzhageni*), wild chimpanzees (*Pan troglodytes*), agile mangabeys (*Cercocebus agilis*) and several groups of western lowland gorillas (*Gorilla gorilla gorilla*) at different levels of habituation (Fig. 1).

2.2. Collection of insects

Tsetse and tabanid flies were collected in the DSPA during September 2012. The insects from the BH and BG collection sites were thoroughly examined by specific PCR for trypanosome prevalence and in the case of tsetse flies for blood source. The microhabitat of each collection and the occurrence of wildlife were noted. Most bloodsucking flies were collected on forest edges using tent-like Malaise traps with a black and dark-blue striped central wall and white roof that directs insects upwards to a cylinder containing a killing jar with 70% pure ethanol. The captured insects were collected from the Malaise traps twice each day. A small number of flies were obtained by sweep netting and hand collection. Ethanol-stored specimens were transported to the laboratory, where they were sorted under a stereomicroscope according to species, sex, collection site and date of trapping. Species identification was done using an online key (Les Glossines ou mouches tsé-tsé; <http://www.cnev.fr/index.php/publications-et-outils/outils-didentification/identiciels>) and confirmed by Dr. Pascal Grébaud (IRD-CIRAD, Montpellier, France). The correctness of the determination was further verified by barcoding five selected specimens from each morpho-species. This involved sequencing of their cytochrome *c* oxidase subunit 1 (*cox1*) gene using universal primers (Hebert et al., 2003), followed by a comparison with available databases, BOLD (<http://www.boldsystems.org/>) and GenBank (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

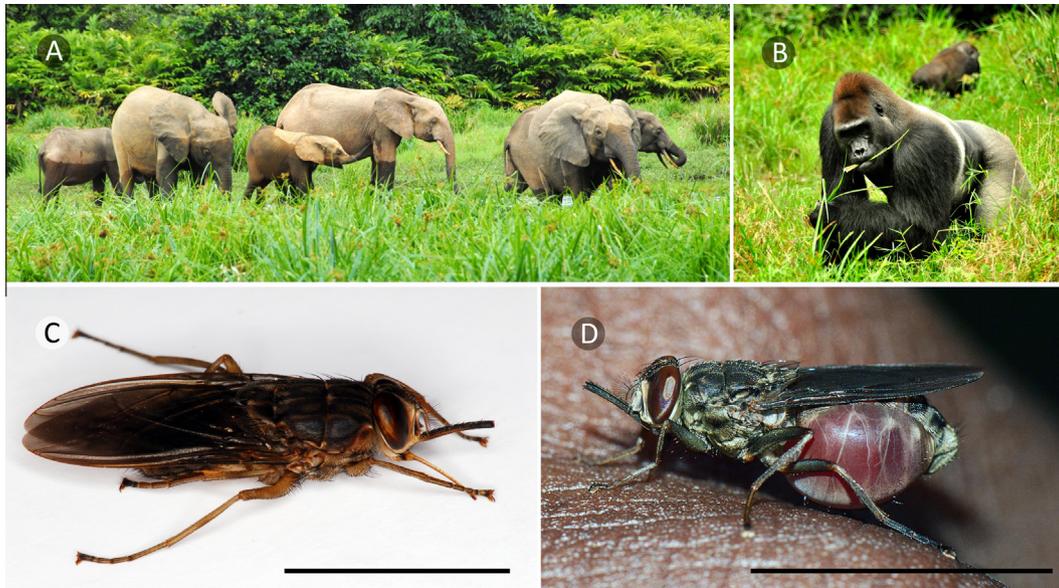


Fig. 1. Wild animals and tsetse flies in Dzanga-Sangha Protected Areas, Central African Republic. (A) Herd of forest elephants in Bai-Gubunga. (B) Makumba, a silverback gorilla from a habituated group in Bai-Hokou. (C) *Glossina (Austenia) tabaniformis*. (D) *Glossina (Nemorhina) fuscipes fuscipes*. Scale bars = 1 cm.

2.3. DNA isolation, nested PCR, sequencing and recombinant sequence check

After washing individual tsetse and tabanid flies for 1 h in distilled water, their DNA was extracted using a High Pure PCR Template Preparation Kit (Roche, Germany) according to the manufacturer's instructions. Tissue homogenates were prepared from the whole tsetse or tabanid flies. In the case of tabanid flies, some specimens were divided into three body parts (head, thorax and abdomen), in order to determine the exact location of trypanosomes within the insect, and whether it is in accordance with the expected mode of transmission. For species identification, the trypanosome 18S rRNA genes were amplified using newly designed specific nested PCR with amplicon sizes reaching up to 2.2 kb and 2.1 kb in the first and second round, respectively. The first amplification round consisting of 35 cycles was performed in a final volume of 20 μ l of PCR mix (TopBio, Czech Republic) containing 60–120 ng of total DNA isolated individually from the tsetse or tabanid fly body, and 10 pmol of each primer (S762: 5'-GACTTTTGCTTCTCTAWTG-3' and S763: 5'-CATATGCTTGTTCAGGAC-3'; Maslov et al., 1996). The annealing temperature was 55 °C for 90 s. For the next step consisting of 35 amplification cycles, 1 μ l of amplified product was used as a template with 10 pmol of newly designed nested primers (TR-F2: 5'-GARTCTGC GCATGGCTCATTACATCAGA-3' and TR-R2: 5'-CRCAGTTTGATGAGC TGCGCCT-3'). The annealing temperature was 64 °C for 90 s. Both strands of PCR products were sequenced. The search for possible recombinant sequences obtained by PCR was performed by Recombination Detection Program v.4 (RDP4) using the RDP, GENECONV, Chimaera, MaxChi, BOOTSCAN, SISCAN and 3seq methods with default settings (Martin et al., 2010).

2.4. Phylogenetic analysis

The 18S rRNA sequences obtained by nested PCR from flies were used to generate an alignment using K-align (<http://www.ebi.ac.uk/Tools/msa/kalign/>). Ambiguous positions and poorly alignable sequences were manually removed using BioEdit. The final 18S rRNA alignment included 2082 characters. An evolutionary model (TIM2+I+G) for this dataset was selected using the Akaike criterion

in Modeltest 2.1.4 and used for Maximum Likelihood inference in PhyML 3.0. A heuristic search was performed using the subtree pruning and regrafting (SPR) branch-swapping algorithm and statistical support of bipartition was assessed using bootstrap resampling (1000 replicates) as described elsewhere (Votýpka et al., 2010, 2012). The accession numbers of sequences retrieved from GenBank and used in phylogenetic reconstructions and 18S rRNA alignments are provided in figures.

2.5. Preparation and analysis of the DNA library

To establish the origins of blood meals in tsetse flies, 1033 DNAs individually isolated from captured insects were separated into pools of 10 DNA samples and each pool was used as a template for specific PCR with primers Cytbb1F (5'-CCATCMAACATYTCADC ATGATGAAA-3') and Cytbb2R (5'-GCHCTCAGAATGAYATTTGKCTCA-3'). PCR conditions were as follows: 94 °C for 5 min, then 35 cycles (94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min), and 72 °C for 7 min (Rádrová et al., 2013). Amplicons were purified using GenElute™ PCR Clean-Up (Sigma-Aldrich, Germany) and then further pooled according to tsetse fly species (*Glossina fuscipes fuscipes* versus *Glossina tabaniformis*), gender (male versus female), and capture site (BH versus BG) to form eight similarly numerous groups: (i) males of *G. f. fuscipes* from BH, (ii) females of *G. f. fuscipes* from BH, (iii) males of *G. tabaniformis* from BH, (iv) females of *G. tabaniformis* from BH, (v) males of *G. f. fuscipes* from BG, (vi) females of *G. f. fuscipes* from BG, (vii) males of *G. tabaniformis* from BG, and (viii) females of *G. tabaniformis* from BG.

DNA libraries from each group were prepared as described in the Rapid Library Preparation Method Manual (Roche) and then quantified using a TBS 380 Fluorometer (Turner, USA). Different Multiplex Identifiers (MIDs) with unique sequence tags were ligated to each group of amplicons and resulting DNA libraries were purified using Agencourt AMPure XP magnetic beads (Beckman Coulter, USA). Each DNA library was quantified using a TBS 380 Fluorometer (Turner); the number of molecules in each library was calculated and samples were diluted with Tris-EDTA buffer (10 mM TRIS, 1 mM EDTA, pH 8.0) to a concentration of 1×10^7 molecules/ μ l. Barcoded libraries were subsequently pooled and processed according to the emulsion PCR (emPCR)

Amplification Method Manual Lib-L (Roche). The library of DNA molecules was mixed with capture beads in a 1:1.7 ratio, emulsified and clonally amplified with emulsion-based PCR. Subsequently, beads carrying amplified DNA were recovered from emulsion, enriched and sequencing primers were annealed, yielding 450,000 beads. Sequencing using a GS Junior pyrosequencer (Roche) resulted in almost 47,000 raw sequences, which were analyzed and processed using the Mothur package (Schloss, 2009; www.mothur.org). Sequences were separated into eight groups according to their unique sequence tags present in MIDs. Denoising was performed by PyroNoise (minflows = 360; maxflows = 720; pdiffs = 3; bdiffs = 1) followed by sequence trimming (maxhomop = 8; minlength = 150). This resulted in 14,533 high quality sequences that were analyzed using BLASTN with an E-value cut-off set at 10–50. Statistical evaluation was performed using the Kruskal–Wallis and Mann–Whitney *U* tests (Statistica 6.1, StatSoft) and *P* values below 0.05 were considered significant. All primary sequencing data were deposited to the NCBI Sequence Read Archive under accession numbers SAMN03703425, SAMN03703430, SAMN03703434, SAMN03703438, SAMN03703509, SAMN03703516, SAMN03703517 and SAMN03703518. It should be noted that the methodology used is biased towards sequences from a recently taken blood meal where the host DNA is still highly abundant.

3. Results

3.1. Prevalence of trypanosomes and their 18S rRNA-based phylogeny

Morphological examination allowed allocation of all 1033 tsetse flies trapped in the two localities within the primary forest area of the DSPA to just two species. The subgenus *Austenia* was represented by 886 specimens of *G. tabaniformis*, while 147 specimens of *G. f. fuscipes* belonged to the subgenus *Nemorhina* (Fig. 1). All obtained *cox1* tsetse fly sequences clustered into two well-defined, internally homogenous groups (data not shown) fully corresponding with morphology-based determinations. Our *G. f. fuscipes* sequences (GenBank accession numbers KP979582–4) show 99% similarity with reference sequences. In the case of *G. tabaniformis* (KP979585–7), however, no sequences are publicly available and our sequences derived from this species constituted a separate clade among available tsetse fly sequences (data not shown).

The detection of trypanosomes by the nested PCR approach was performed on DNA samples isolated individually from a randomly selected subset of 564 tsetse flies representing 436 and 128 specimens of *G. tabaniformis* and *G. f. fuscipes*, respectively. The analyzed set was composed of 260 males and 202 females (102 specimens were homogenized prior to the determination of their sex), with 403 and 161 insects trapped at BH and BG, respectively. In total, 183 (32.4%) tsetse flies were positive for at least one trypanosome species (Table 1), with no significant correlation between infection and tsetse fly species ($\chi^2 = 0.684$, $P = 0.41$; 34.3% of *G. tabaniformis* versus 27.3% of *G. f. fuscipes*) or gender (*G. tabaniformis*: $\chi^2 = 2.073$, $P = 0.15$; 35.2% of males versus 28.1% of females; *G. f. fuscipes*: $\chi^2 = 0.406$, $P = 0.53$; 33.2% of males versus 27.7% of females) observed. The two trapping sites (BH and BG) differ in several aspects. First of all, they significantly differ in the distribution of the two tsetse fly species ($\chi^2 = 13.257$, $P < 0.0003$), with a relatively high occurrence of *G. f. fuscipes* in BH. Another notable difference is associated with species composition of vertebrates available in “bais” that serve as hosts for trypanosomes and as a source of blood for tsetse and tabanid flies. The BG locality is inhabited almost exclusively by herds of forest buffaloes, while other mammals visit both locations at approximately the same

frequency (J. Votýpka, K.J. Petrželková, D. Modrý, personal observations).

The two sites differ significantly ($\chi^2 = 72.836$, $P < 0.0001$) in overall trypanosome prevalence, with 23.1% and 55.3% of tsetse flies from BH and BG, respectively, being positive. Regarding individual trypanosome species (see below), the number of trypanosome-positive tsetse flies was statistically significant only for *Trypanosoma theileri*. In congruence with overall trypanosome prevalence, tsetse flies were more commonly infected with *T. theileri* at BG (49.8%) than they were at BH (1.9%) ($\chi^2 = 91.187$, $P < 0.0001$) which is in good correlation with the occurrence of buffaloes in the former bai.

Altogether, based on 18S rRNA, we detected eight trypanosome species inventoried as shown in Table 1 and belonging to three subgenera (Fig. 2). This substantial diversity includes six members of the subgenus *Nannomonas*, one species affiliated with the subgenus *Trypanozoon*, and one species of the subgenus *Megatrypanum*. The tsetse fly species and prevalence of all registered trypanosomes are shown in Table 1. The more abundant *G. tabaniformis* hosted all eight trypanosome species, while the less numerous *G. f. fuscipes* was infected with only the three most prevalent trypanosomes (Table 1).

In our dataset, the subgenus *Nannomonas* is represented by three members of the *T. simiae* group: (i) the typical *T. simiae* (isolates G114 and G219 from *G. tabaniformis*; GenBank accession numbers KP307021 and KR024686, respectively); (ii) the new *T. simiae* ‘Bai’ (isolate G107 from *G. tabaniformis*; accession number KP307022), which forms a new subclade within the *T. simiae* complex together with the lineage *T. simiae* ‘Tsavo’; and (iii) *Trypanosoma* sp. ‘Fly9’ (also called *T. godfreyi*-like in Adams et al., 2010a) (isolates G49 and G159 from *G. tabaniformis*; accession numbers KP307023 and KR024687, respectively). The latter trypanosome (‘Fly9’) was previously described from a tsetse fly (Malele et al., 2003) and is related to *Trypanosoma godfreyi* (Adams et al., 2010a). The species *T. congolense* is represented by the new lineage (iv) *T. congolense* ‘Dzanga-Sangha’ (isolates G22 from *G. tabaniformis*; accession number KP307024; and G34 and G39 from *G. f. fuscipes*; accession numbers KP307026 and KP307025, respectively). These three isolates form a new subclade within the species *T. congolense*, distinct from the hitherto recognized ‘Forest’, ‘Savannah’ and ‘Kilifi’ lineages. Finally, two closely related isolates within the subgenus *Nannomonas* that cluster between *T. congolense* and the *T. simiae* complex form two new species labeled (v) *Trypanosoma* sp. ‘Ngbanda’ (isolates G277 and G603 from *G. tabaniformis*; accession numbers KP307019 and KR024684, respectively) and (vi) *Trypanosoma* sp. ‘Didon’ (isolates G432 and G601 from *G. tabaniformis*; accession numbers KP307020 and KR024685, respectively).

The trypanosome labeled *Trypanosoma* sp. ‘Makumba’ (isolate G73 from *G. f. fuscipes* and isolate G433 from *G. tabaniformis*; accession numbers KP307018 and KR024683, respectively) is closely related to *Trypanosoma* sp. ‘Msubugwe’ (Hamilton et al., 2008), with which it constitutes a sister group of the subgenus *Trypanozoon* (Fig. 2). Finally, the subgenus *Megatrypanum* is represented by the *T. theileri* complex (isolate G24 from *G. f. fuscipes*; accession number KR024688), although as the sequences obtained are slightly heterogeneous, these may represent several genotypes or a complex of closely related species (data not shown). In addition to tsetse flies, we captured 81 female tabanids belonging to the genera *Chrysops* (43 specimens), *Haematopota* (6), and *Tabanus* (32), of which 48 (59.3%) were positive for *T. theileri* (Table 2). No other trypanosome species was detected. Prevalence of infection, which was confined to the thoracic and abdominal parts (Fig. 3; also see Section 4), differed significantly among genera ($\chi^2 = 10.502$, $P < 0.001$; 44.2% and 81.3% positivity for *Chrysops* and *Tabanus*, respectively). While 46.8% of tabanids

Table 1
Prevalence of trypanosomes in tsetse flies captured in the Dzanga-Sangha Protected Areas, Central African Republic.

Tsetse fly species	Examined/positive (%)	Prevalence of <i>Trypanosoma</i> spp. (%)							
		<i>T.s.</i>	<i>T.s.B</i>	Fly9	<i>T.c.</i>	Ngb.	Did.	Mak.	<i>T.t.</i>
<i>Glossina tabaniformis</i>	436/149 (34.2)	2.8	0.7	0.7	8.7	0.5	0.5	5.0	15.1
<i>Glossina fuscipes fuscipes</i>	128/35 (27.3)	0.0	0.0	0.0	7.8	0.0	0.0	2.3	17.2
Total	564/183 (32.4)	2.1	0.5	0.5	8.5	0.4	0.4	4.4	15.6

T.s., typical *Trypanosoma simiae*; *T.s.B*, *Trypanosoma simiae* 'Bai'; Fly9, *Trypanosoma* sp. 'Fly9'; *T.c.*, *Trypanosoma congolense* 'Dzanga-Sangha'; Ngb., *Trypanosoma* sp. 'Ngbanda'; Did., *Trypanosoma* sp. 'Didon'; Mak., *Trypanosoma* sp. 'Makumba'; *T.t.*, *Trypanosoma theileri*.

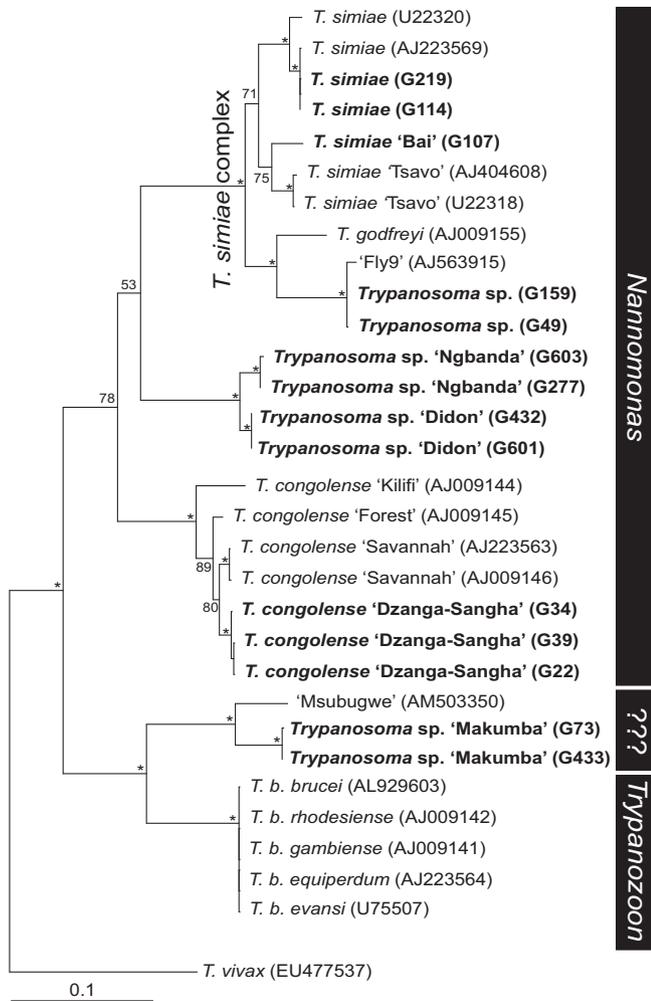


Fig. 2. 18S rRNA-based phylogeny of trypanosomes from tsetse flies. Maximum likelihood tree with bootstrap values for 1000 replicates (∗ indicates value >95). 18S rRNA sequences generated within this study are in bold. Each sequence, labeled by tsetse fly number (e.g., G219), belongs to a given species/complex/group (see Table 1). Notes on the newly described species: *Trypanosoma simiae* 'Bai' is a new member of the *T. simiae* complex; *Trypanosoma congolense* 'Dzanga-Sangha' is a new lineage of the species *T. congolense*; *Trypanosoma* sp. 'Ngbanda' and *Trypanosoma* sp. 'Didon' are two new species constituting a new clade within the subgenus *Nannomonas*; and *Trypanosoma* sp. 'Makumba' is a new member of the *Trypanosoma* sp. 'Msubugwe' group. Sequences obtained belonging to the *Trypanosoma theileri* complex of the subgenus *Megatrypanum* were not included.

were positive for trypanosomes at BH, positivity increased to 81.5% at BG ($\chi^2 = 7.717$, $P < 0.005$).

In order to identify possible recombinant sequences generated by PCR, we performed recombination detection analysis with the RDP4 program using the RDP, GENECONV, Chimaera, MaxChi, BOOTSCAN, SISCAN and 3seq detection algorithms. Application of

these algorithms failed to identify any statistically significant recombination events ($P < 0.05$).

3.2. Determination of tsetse fly blood meal

To establish upon which vertebrates the tsetse flies fed and whether this varied by collection site, tsetse fly species and/or sex, the set of 1,033 DNAs individually isolated from specimens captured at BH and BG was subdivided into eight categories. Next, an equal amount of DNA from each category was pooled and used for the preparation of eight separate Lib-L libraries (see subsection 2.5). Due to substantial differences in the number of sequences extracted from each library, only limited mutual comparison was possible. Furthermore, it should be noted that the data obtained do not represent the blood "prevalence" or the number of tsetse flies engorging on given hosts, but only the frequencies of different sequences that could to some extent be affected by such factors as blood freshness. The strength of the method used lies in its ability to detect rare sequences. Indeed, it was able to monitor a wide range of blood sources.

Eight species of warm- and two species of cold-blooded vertebrates constituted more than 99% of the 14,533 high-quality sequences obtained from the engorged tsetse flies (Table 3). The dominant source of blood at both sampled sites consisted of forest buffaloes (*Syncerus caffer nanus*; 42.9%), with humans (24.9%) representing the second most frequent source (Fig. 1). At BG, a location distant from any human settlements, human blood occurred significantly less frequently (16.0%). Three groups of large forest mammals represented an additional important source of blood: bongo and sitatunga (*Tragelaphus* spp.; 15.7%), two species of wild hogs (*Potamochoerus porcus* and *Hylochoerus meinertzhageni*; 12.7%), and duikers (*Cephalophus* spp.; 4.1%), whereas the blood of rodents (unspecified member of the order Rodentia), forest elephants (*Loxodonta cyclotis*), tortoises (Testudinidae), and crocodiles (*Osteolaemus tetraspis*) was identified only sporadically (Table 3). The sporadic presence of elephant blood is noteworthy, as these are frequent visitors to both examined sites (Fig. 1).

4. Discussion

Recently, infections by trypanosomes mainly belonging to the *T. brucei* complex were detected with unexpected frequency in tissue samples and feces of African great apes (Jirků et al., 2015). To evaluate the possibility that tsetse and/or tabanid flies take

Table 2
Prevalence of trypanosomes in tabanid flies captured in the Dzanga-Sangha Protected Areas, Central African Republic.

Tabanid genus	Examined/positive (%)
<i>Chrysops</i>	43/19 (44.2)
<i>Haematopota</i>	6/3 (50.0)
<i>Tabanus</i>	32/26 (81.3)
Total	81/48 (59.3)

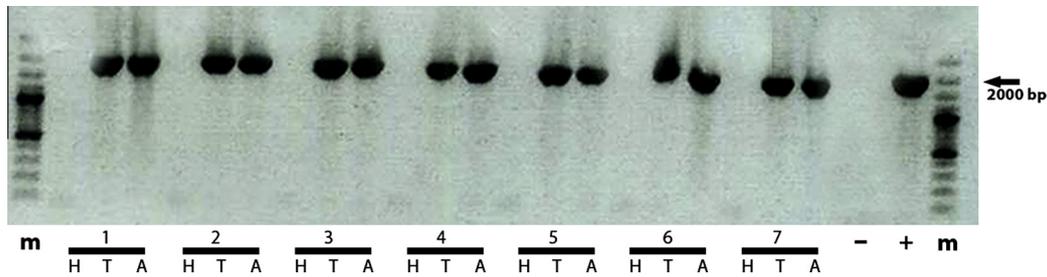


Fig. 3. Detection of *Trypanosoma theileri* in body parts of *Tabanus* spp. Each individual was dissected into three body parts: head (H), thorax (T) and abdomen (A). Trypanosomes were detected by PCR only in the thorax and abdomen. 1–7, tabanid specimen number; m, marker; –, negative and +, positive controls.

Table 3
Host cytochrome *b* gene detected in tsetse flies captured in the Dzanga-Sangha Protected Areas, Central African Republic.

Host	Collection site		Tsetse fly species		Tsetse fly sex		Seq. analysis No. of seq. (%)
	BH (No: 582)	BG (No: 451)	G.t. (No: 886)	G.f.f. (No: 147)	Male (No: 367)	Female (No: 556)	
African buffalo	40.66% ^a	68.91% ^a	50.62%	50.15%	53.26%	50.62%	2364 (42.88)
Human	24.87% ^a	16.04% ^a	23.19%	26.12%	24.03%	23.86%	1685 (30.56)
Bongo/Sitatunga	15.77%	11.46%	12.28%	19.57%	17.24%	16.38%	875 (15.87)
Bush pig/Giant hog	12.69% ^a	1.92% ^a	9.15% ^a	1.04% ^a	2.71%	3.68%	364 (6.60)
Duiker	4.40%	1.07%	3.41%	2.88%	1.75%	4.69%	173 (3.14)
Unidentified rodent	0.99%	0.42%	0.91%	0.14%	1.01%	0.26%	26 (0.47)
Forest elephant	0.30%	N.D.	0.18%	N.D.	N.D.	0.22%	10 (0.18)
Chicken	0.07%	0.05%	0.05%	0.05%	N.D.	0.05%	4 (0.07)
Unidentified tortoise	0.07%	N.D.	0.05%	N.D.	N.D.	0.00%	2 (0.04)
Dwarf crocodile	N.D.	0.03%	0.03%	N.D.	N.D.	0.03%	1 (0.02)
Total number of seq.	2495	2091	3306	2207	1763	2826	5513 (100)

BH, Bai-Hokou; BG, Bai-Gubunga; G.t., *Glossina tabaniformis*; G.f.f., *Glossina fuscipes fuscipes*; No, number of analyzed tsetse flies; seq., sequences; N.D., not detected.

^a Statistically significant difference ($P < 0.05$).

blood from apes and are involved in transmitting their trypanosomes, an extensive analysis of these potential vectors from a habitat of chimpanzees and gorillas was performed.

Upon searching the literature, we came to realize that tsetse and tabanid flies had never before been collected in a primary forest inhabited by African great apes and other wildlife. For that reason, we undertook an extensive search for trypanosomes in hundreds of flies. Employing high-throughput screening using next-generation sequencing of almost all tsetse flies enabled us to determine the spectrum of animal species upon which the flies had fed. Indeed, at both studied sites within the DSPA, the DNA of several mammalian species was detected in *G. (Austenia) tabaniformis* and *G. (Nemorhina) fuscipes fuscipes*, thereby proving the opportunistic feeding behavior of *Glossina* flies (Muturi et al., 2011). Feeding of *G. fuscipes fuscipes* and several other tsetse species on cold-blooded vertebrates such as monitor lizard, turtle and snake has been documented previously (Waiswa et al., 2006; Bouyer et al., 2007; Farikou et al., 2010) and is further corroborated by our data.

Although lowland gorillas and tsetse flies co-habit in both capture sites and the Malaise traps were applied in the bais at least twice during visits by these great apes, they do not appear to be frequent targets of the examined *Glossina* spp.. Such an observation is non-trivial and surprising, especially when the virtual absence of blood from gorillas and other apes is compared with the abundant presence of human blood. Indeed, both *Glossina* spp. were often seen to attack humans tracking the gorillas (J. Votýpka, K.J. Petrželková, D. Modrý, personal observations; Fig. 1). One possible explanation for the obtained results implies the existence of some unknown anti-tsetse fly behavior by gorillas and chimpanzees, however, other explanations must also be considered. Host odor is an important component of tsetse fly perception and these insects do not seem to be attracted nor repelled by the primate

smell. Another possibility is that the relative abundance of great apes in the sampling area is still low compared with other available hosts, and even analysis of 1000 tsetse flies is insufficient to cover the whole host spectrum and exhaustively map feeding preferences of the studied tsetse flies. Finally, because the effectiveness of black and dark-blue striped traps of various shapes differs significantly among tsetse fly species, a confounding presence of other, non-attracted and consequently undetected, *Glossina* spp. cannot be excluded.

Recently, several new trypanosome species or genotypes have been discovered in tsetse flies (Adams et al., 2006, 2010b; Hamilton et al., 2008) and large mammals (Auty et al., 2012a), gradually increasing the diversity and host range of African trypanosomes. Regrettably, these unnamed flagellates remain largely ignored, probably due to an absence of formal taxonomic description (Adams et al., 2010a). The reasons for this fact include unavailability in culture and the lack of identified vertebrate and/or insect hosts, which may well include rare species or those that have been inaccessible to investigation.

Our nested 18S rRNA-based PCR assay proved to be highly sensitive and provided sufficient information for a phylogenetic analysis that revealed an unexpectedly broad spectrum composed of at least eight trypanosome operational taxonomic units (OTUs). However, the presence of trypanosome DNA in the blood meals of tsetse flies is not proof of their capacity to transmit the parasite. Some OTUs are very closely related to species of medical and veterinary importance. At least one of these trypanosomes was detected in almost one-third of the examined tsetse flies, testifying to its active circulation in the DSPA. Two representatives of the subgenus *Nannomonas*, *T. simiae* and *T. congolense*, recently expanded into a complex of four and three lineages, respectively (Njiru et al., 2004; Hamilton et al., 2008; Adams et al., 2010a,b; Simo et al., 2012). Our findings further increase this diversity:

Trypanosoma simiae 'Bai' extends the *T. simiae* complex into five lineages (typical *T. simiae*, *T. simiae* 'Tsavo', *T. simiae* 'Bai', *T. godfreyi* and *Trypanosoma* sp. 'Fly9') and the herein newly described *T. congolense* 'Dzanga-Sangha' expands the species *T. congolense* into four lineages (*T. congolense* 'Forest', 'Savannah', 'Kilifi' and 'Dzanga-Sangha'). Finally, *Trypanosoma* sp. 'Makumba' clusters together with *Trypanosoma* sp. 'Msubugwe' (Hamilton et al., 2008; Adams et al., 2010a,b) to form an enigmatic group which may represent an evolutionary link between the subgenera *Trypanozoon* and *Nannomonas*.

Two additional novel flagellates of the subgenus *Nannomonas* first reported in this study, *Trypanosoma* sp. 'Ngbanda' and *Trypanosoma* sp. 'Didon' form, in the 18S rRNA-based tree, a new clade branching off between the *T. congolense* and *T. simiae* complexes. Such a position of this new clade raised concern about a potential hybrid molecule. Due to a very low amount of DNA, we failed to amplify additional genes (e.g. glyceraldehyde 3-phosphate dehydrogenase). However, we have obtained the same results twice and independently from different tsetse fly specimen, and further analyzed the amplified sequences using the recombinant detection RDP4 programs, which failed to identify any statistically significant recombination events. Therefore, we believe the obtained sequences are authentic.

We recently established a diagnostic ITS1-based PCR assay that allows detection of even very small amounts of trypanosome DNA in ethanol- or RNAlater-stored feces or in partially decomposed tissue samples collected from dead animals (Jirků et al., 2015). While useful for assignment to previously known species, the information obtained in the ITS1 region is insufficient for phylogenetic analyses of novel lineages which should, if possible, be based on the 18S rRNA gene.

Using the ITS1-based PCR assay, the *T. brucei* complex was frequently detected in an extensive survey of African great apes from several tropical African countries but not CAR (Jirků et al., 2015). The 18S rRNA-based nested PCR assay revealed circulation of at least eight trypanosome lineages, but we did not detect any trypanosomes belonging to the *T. brucei* complex. This result can be explained either (i) by an overall absence of this complex in the studied biotope (no information is available about *T. brucei* in humans and animals in the studied area), or (ii) by the presence of an unregistered tsetse fly species transmitting flagellates of the *T. brucei* complex.

No significant differences were observed in trypanosome prevalence among tsetse fly species or gender, although significant differences did exist between the two trapping sites in terms of overall trypanosome prevalence and, in particular, the abundance of the most common species, *T. theileri*. Since the same trend was observed in the trapped tabanid flies, we believe that these differences correspond with the higher concentration at Bai-Gubunga (BG) of forest buffaloes, of which *T. theileri* is a frequent parasite.

The extremely high positivity of tabanid flies for the *T. theileri* is likely a consequence of their increased life span compared with other insect vectors, resulting in a high number of blood meals. Furthermore, the striking differences in the prevalence of *T. theileri* can be explained by the large and highly aggressive tabanid flies of the genus *Tabanus* preferentially attacking forest buffaloes, while the relatively small members of the genus *Chrysops* seems to prefer smaller forest ungulates. The invariable restriction of *T. theileri* to the gut (Fig. 3) of these important vectors is fully compatible with the stercorarian-type development proposed by Bose and Heister (1993).

In summary, the present study confirms the feeding opportunism of tsetse flies. Indeed, an extensive analysis of their blood meal not only identified their main mammalian targets but also confirmed previous evidence for their capacity to feed on cold-blooded vertebrates. The absence of gorillas, which occur

frequently at both studied locations, in the feeding range of those tsetse flies analyzed may be due to the relatively low number of analyzed flies, the presence of unregistered tsetse fly species, or some anti-tsetse fly behavior by the apes. Finally, we have demonstrated that tsetse flies, but not tabanid flies, harbor a surprisingly broad spectrum of trypanosomes, some of which were encountered, to our knowledge, for the first time.

Acknowledgements

Research permission was obtained from the relevant ministry of the Central African Republic. Special thanks are given to A. Todd, the park staff, the Primate Habituation Programme and the World Wildlife Fund of the Dzanga-Sangha Protected Areas, Central African Republic. We would also like to thank J. Košnar and A. Chroňáková from the Institute of Soil Biology, Czech Republic, for their help with Next Generation sequencing. This study was supported by grant number M200961204 and the Praemium Academiae award from the Academy of Sciences of the Czech Republic, grant 108/2013/P from the University of South Bohemia, Czech Republic, and European Union 7th Framework Programme grants CZ.1.07/2.3.00/300032 and 316304. This research was also co-financed by the European Social Fund and state budget of the Czech Republic (CZ.1.07/2.3.00/20.0300), UEFISCDI grant PCE236/2011 and COST Action TD1303.

References

- Adams, E.R., Hamilton, P.B., Gibson, W.C., 2010a. African trypanosomes: celebrating diversity. *Trends Parasitol.* 26, 324–328.
- Adams, E.R., Hamilton, P.B., Rodrigues, A.C., Malele, I.I., Delespaux, V., Teixeira, M.M.G., Gibson, W., 2010b. New *Trypanosoma* (*Duttonella*) *vivax* genotypes from tsetse flies in East Africa. *Parasitology* 137, 641–650.
- Adams, E.R., Malele, I.I., Msangi, A.R., Gibson, W.C., 2006. Trypanosome identification in wild tsetse populations in Tanzania using generic primers to amplify the ribosomal RNA ITS-1 region. *Acta Tropica* 100, 103–109.
- Auty, H., Anderson, N.E., Picozzi, K., Lembo, T., Mubanga, J., Hoare, R., Fyumagwa, R.D., Mable, B., Hamill, L., Cleaveland, S., Welburn, S.C., 2012a. Trypanosome diversity in wildlife species from the Serengeti and Luangwa Valley ecosystems. *PLoS Negl. Trop. Dis.* 6 (10), e1828. <http://dx.doi.org/10.1371/journal.pntd.0001828>.
- Auty, H.K., Picozzi, K., Malele, I., Torr, S.J., Cleaveland, S., Welburn, S.C., 2012b. Using molecular data for epidemiological inference: assessing the prevalence of *Trypanosoma brucei rhodesiense* in tsetse in Serengeti, Tanzania. *PLoS Negl. Trop. Dis.* 6 (1), e1501. <http://dx.doi.org/10.1371/journal.pntd.0001501>.
- Bose, R., Heister, N.C., 1993. Development of *Trypanosoma* (*M.*) *theileri* in tabanids. *J. Euk. Microbiol.* 40, 788–792.
- Bouyer, J., Pruvot, M., Bengaly, Z., Guerin, P.M., Lancelot, R., 2007. Learning influences host choice in tsetse. *Biol. Lett.* 3, 113–116.
- Calvignac-Spencer, S., Leendertz, F.H., Gilbert, M.T., Schubert, G., 2013. An invertebrate stomach's view on vertebrate ecology: certain invertebrates could be used as "vertebrate samplers" and deliver DNA-based information on many aspects of vertebrate ecology. *Bioessays* 35, 1004–1013. <http://dx.doi.org/10.1002/bies.201300060>.
- Carnes, J., Anupama, A., Balmer, O., Jackson, A., Lewis, M., Brown, R., Cestari, I., Desquesnes, M., Gendrin, C., Hertz-Fowler, C., Imamura, H., Ivens, A., Kořený, L., Lai, D.-H., MacLeod, A., McDermott, S.M., Merritt, C., Monnerat, S., Moon, W., Myler, P.J., Phan, I., Ramasamy, G., Sivam, D., Lun, Z.-R., Lukeš, J., Stuart, K., Schnauffer, A., 2015. Genome and phylogenetic analyses of *Trypanosoma evansi* reveal extensive similarity to *T. brucei* and multiple independent origins for dyskinetoplasty. *PLoS Negl. Trop. Dis.* 9, e3404. <http://dx.doi.org/10.1371/journal.pntd.0003404>.
- Farikou, O., Njiokou, F., Simo, G., Asonganyi, T., Cuny, G., Geiger, A., 2010. Tsetse fly blood meal modification and trypanosome identification in two sleeping sickness foci in the forest of southern Cameroon. *Acta Trop.* 116, 81–88.
- Gibson, W., 2007. Resolution of the species problem in African trypanosomes. *Int. J. Parasitol.* 37, 829–838.
- Hamilton, P.B., Adams, E.R., Malele, I.I., Gibson, W.C., 2008. A novel, high-throughput technique for species identification reveals a new species of tsetse-transmitted trypanosome related to the *Trypanosoma brucei* subgenus, *Trypanozoon*. *Infect. Genet. Evol.* 8, 26–33.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. B Biol. Sci.* 270, 313–321.
- Jirků, M., Votýpka, J., Petřelová, K.J., Jirků-Pomajbíková, K., Kriegová, E., Vodička, R., Junglen, S., Lankaster, F., Leendertz, S.A.J., Wittig, R.M., Boesch, C., Modrý, D., Ayala, F.J., Leendertz, F.H., Lukeš, J., 2015. Wild chimpanzees are infected by *Trypanosoma brucei*. *Int. J. Parasitol. Parasit. Wildl.* 4, 277–282.

- Lai, D.H., Hashimi, H., Lun, Z.R., Ayala, F.J., Lukeš, J., 2008. Adaptations of *Trypanosoma brucei* to gradual loss of kinetoplast DNA: *Trypanosoma equiperdum* and *Trypanosoma evansi* are petite mutants of *T. brucei*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1999–2004.
- Malele, I., Craske, L., Knight, C., Ferris, V., Njiru, Z.K., Hamilton, P., Lehane, S., Lehane, M., Gibson, W., 2003. The use of specific and generic primers to identify trypanosome infections of wild tsetse flies in Tanzania by PCR. *Infect. Genet. Evol.* 3, 271–279.
- Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefevre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26, 2462–2463.
- 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.
- Muturi, C.N., Ouma, J.O., Malele, I.I., Ngure, R.M., Rutto, J.J., Mithofer, K.M., Enyaru, J., Masiga, D.K., 2011. Tracking the feeding patterns of tsetse flies (*Glossina* genus) by analysis of bloodmeals using mitochondrial cytochromes genes. *PLoS One* 6, e17284. <http://dx.doi.org/10.1371/journal.pone.0017284>.
- Njiru, Z.K., Makumi, J.N., Okoth, S., Ndungu, J.M., Gibson, W.C., 2004. Identification of trypanosomes in *Glossina pallidipes* and *G. longipennis* in Kenya. *Infect. Genet. Evol.* 4, 29–35.
- Rádrová, J., Šeblová, V., Votýpka, J., 2013. Feeding behavior and spatial distribution of *Culex* mosquitoes (Diptera: Culicidae) in wetland areas of the Czech Republic. *J. Med. Entomol.* 50, 1097–1104.
- Simarro, P.P., Diarra, A., Postigo, J.A.R., Franco, J.R., Jannin, J.G., 2011. The human African trypanosomiasis control and surveillance programme of the World Health Organization 2000–2009: the way forward. *PLoS Negl. Trop. Dis.* 5, e1007. <http://dx.doi.org/10.1371/journal.pntd.0001007>.
- Simo, G., Silatsa, B., Flobert, N., Lutumba, P., Mansinsa, P., Madinga, J., Manzambi, E., De Deken, R., Asonganyi, T., 2012. Identification of different trypanosome species in the mid-guts of tsetse flies of the Malanga (Kimpese) sleeping sickness focus of the Democratic Republic of Congo. *Parasites Vectors* 5, 201. <http://dx.doi.org/10.1186/1756-3305-5-201>.
- Schloss, P.D., 2009. A high-throughput DNA sequence aligner for microbial ecology studies. *PLoS One* 4, e8230. <http://dx.doi.org/10.1371/journal.pone.0008230>.
- Schofield, C.J., Kabayo, J.P., 2008. Trypanosomiasis vector control in Africa and Latin America. *Parasites Vectors* 1, 24. <http://dx.doi.org/10.1186/1756-3305-1-24>.
- Sharma, R., Gluenz, E., Peacock, L., Gibson, W., Gull, K., Carrington, M., 2009. The heart of darkness: growth and form of *Trypanosoma brucei* in the tsetse fly. *Trends Parasitol.* 25, 517–524.
- Votýpka, J., Maslov, D.A., Yurchenko, V., Jirků, M., Kment, P., Lun, Z.R., 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., Szabová, J., Rádrová, J., Zídková, L., Svobodová, M., 2012. *Trypanosoma culicavium* sp. nov., an avian trypanosome transmitted by *Culex* mosquitoes. *Int. J. Syst. Evol. Microbiol.* 62, 745–754.
- Waiswa, C., Picozzi, K., Katunguka-Rwakishaya, E., Oluho-Mukani, W., Musoke, R.A., Welburn, S.C., 2006. *Glossina fuscipes fuscipes* in the trypanosomiasis endemic areas of south eastern Uganda: Apparent density, trypanosome infection rates and host feeding preference. *Acta Trop.* 99, 23–29.