Version of Record: https://www.sciencedirect.com/science/article/pii/S1567134821000605
Manuscript_ba057af0641afa3e2a0c143234bf3871

Modulation of trypanosome establishment in *Glossina palpalis palpalis* by its microbiome in the Campo sleeping sickness focus, Cameroon

Running title: Microbiome analysis among tsetse flies from Campo focus

François Sougal NGAMBIA FREITAS[a,b], Flobert NJIOKOU[b], Jean Marc TSAGMO NGOUNE[c], Guilhem SEMPERE[d,e,f], David BERTHIER[d,f], Anne GEIGER[a,b,g,*]

[a]INTERTRYP, Institut de Recherche pour le Développement, University of Montpellier, Montpellier, France.

[b]Faculty of Science, University of Yaoundé I, P.O. Box 812, Yaoundé, Cameroon.

[c]Center for Research in Infectious Diseases, Yaoundé, Cameroun

[d]CIRAD, UMR INTERTRYP, F-34398 Montpellier, France

[e]South Green Bioinformatics Platform, Biodiversity, CIRAD, INRAE, IRD, Montpellier, France

[f]INTERTRYP, Univ Montpellier, CIRAD, IRD, Montpellier, France

[g]Center for Research on Filariasis and other Tropical Diseases (CRFilMT), P.O. Box 5797, Yaoundé, Cameroon

***Corresponding author**

INTERTRYP, Institut de Recherche pour le Développement, University of Montpellier, Montpellier, France; Faculty of Science, University of Yaoundé I, P.O. Box 812, Yaoundé, Cameroon; Center for Research on Filariasis and other Tropical Diseases (CRFilMT), P.O. Box 5797, Yaoundé, Cameroon

E-mail: anne.geiger@ird.fr

Abstract

The purpose of this study was to investigate factors involved in vector competence by analyzing whether the diversity and relative abundance of the different bacterial genera inhabiting the fly's gut could be associated with its trypanosome infection status. This was investigated on 160 randomly selected *G. p. palpalis* flies - 80 trypanosome-infected, 80 uninfected - collected in 5 villages of the Campo trypanosomiasis focus in South Cameroon. Trypanosome species were identified using specific primers, and the V4 region of the 16S rRNA gene of bacteria was targeted for metabarcoding analysis in order to identify the bacteria and determine microbiome composition.

A total of 261 bacterial genera were identified of which only 114 crossed two barriers: a threshold of 0.01% relative abundance and the presence at least in 5 flies. The secondary symbiont *Sodalis glossinidius* was identified in 50% of the flies but it was not considered since its relative abundance was much lower than the 0.01% relative abundance threshold. The primary symbiont *Wigglesworthia* displayed 87% relative abundance, the remaining 13% were prominently constituted by the genera *Spiroplasma, Tediphilus, Acinetobacter* and *Pseudomonas*. Despite a large diversity in bacterial genera and in their abundance observed in micobiome composition, the statistical analyzes of the 160 tsetse flies showed an association with flies' infection status and the sampling sites. Furthermore, tsetse flies harboring *Trypanosoma congolense* Savanah type displayed a different composition of bacterial flora compared to uninfected flies. In addition, our study revealed that 36 bacterial genera were present only in uninfected flies, which could therefore suggest a possible involvement in flies' refractoriness; with the exception of *Cupriavidus*, they were however of low relative abundance. Some genera, including *Acinetobacter, Cutibacterium, Pseudomonas* and *Tepidiphilus,* although present both in infected and uninfected flies, were found to be associated with uninfected status of tsetse flies. Hence their effective role deserves to be further evaluated in order to determine whether some of them could become targets for tsetse control of fly vector competence and

consequently for the control of the disease. Finally, when comparing the bacterial genera identified in tsetse flies collected during 4 epidemiological surveys, 39 genera were found to be common to flies from at least 2 sampling campaigns.

Keywords: *Glossina palpalis palpalis*, Trypanosome, microbiome, Sleeping sickness, Nagana, Metabarcoding.

1. Introduction

Tsetse flies (*Glossina* sp.) are biological vectors of trypanosomes responsible of human African Trypanosomiasis (HAT) or sleeping sickness and of animal African Trypanosomosis (AAT) or nagana in 36 sub-Saharan African countries. They are medically and agriculturally important hematophagous dipterans. Indeed, both human and animal diseases exhibit severe health and social impacts and cause heavy economic losses in endemic countries (Shaw et al. 2013; Welburn and Maudlin, 2012). In humans, *Trypanosoma brucei rhodesiense* (Tbr) and *T. b. gambiense* (Tbg) are responsible for the acute form of HAT in East and South Africa (2% of the total HAT cases), and for the chronic form in West and Central Africa (98% of the cases; fewer than 3000 cases were reported in 2015), respectively (Büscher et al. 2017; Lindner et al. 2020). Wild and domestic animals can also be infected with *T. congolense, T. vivax* and other trypanosome subspecies such as *T. b. brucei* (Nakayima et al. 2012; Tessema et al. 2014; Kato et al. 2015).

Sixty-five million people are still at risk for HAT, though the number of new cases is currently the lowest ever recorded (WHO, 2019) thanks to entomological surveillance, medical diagnosis and drug treatment, as well as vector control. Based on this promising trend, the World Health Organization (WHO) neglected tropical diseases road map targeted elimination of HAT due to *T. b. gambiense* as a public health problem by 2020 and interruption of transmission (zero cases) for 2030 (Franco et al. 2018). Over the past century, HAT has already been close to elimination (mid-1960s), but several outbreaks were recorded in the 1970s following a relaxation of epidemiological surveillance (Brun et al. 2010). In addition to this limit, no vaccine has been developed as a consequence of the frequent switch of the variant surface glycoprotein (VSG) protein of the parasite's coat which allow it to evade the mammalian host immune system (Horn, 2014). In addition, treatments are not always carried out regularly since human populations living in the bush are far from hospitals and dispensaries. Furthermore, available drugs are responsible of many side effects, or are facing resistance of trypanosomes, thus compromising

the efficacy of the chemotherapy (Delespaux and De Koning, 2007; Geiger et al. 2011; Baker et al. 2013).

To mitigate the issues faced by current measures that had led in resurgence of the disease in the past, new alternative control strategies are needed and vector control appears as the most efficient strategy to tackle these vector-borne diseases. Its action can be directly evaluated through either the reduction of tsetse fly populations or of their vector competence, both resulting in a decrease in parasite transmission. As far as tsetse fly is concerned, understanding the mechanisms involved in its susceptibility or resistance to trypanosomes infection represent important avenues for HAT and AAT control/elimination (Geiger et al., 2018).

Tsetse fly populations had been successfully reduced using the Sterile Insect Technique (SIT) (Mogi and Teji, 1991) or insecticide and colour-treated nets, so-called tiny targets, as well as cattle dipping, odor baited trapping, and insecticide spraying, and the transmission of the disease has therefore been significantly reduced (Mahamat et al., 2017). Despite the success gathered (Kariithi et al., 2018) with this approaches, today large-scale insecticide spraying, for example, is no longer used to avoid destroying the diversity of the ecosystem's entomological fauna. The development of methods that can just interfere with the ability of tsetse flies to harbor trypanosomes appears more appropriate than their complete eradication. Indeed, growing evidence has shown that genetic engineering approach (para-transgenesis for example) (De Vooght et al., 2018) is a promising approach to block the transmission of parasites. Furthermore, the microbiome of vectors has been reported to be able to modulate/interfere with the installation, survival and the parasite growth into the fly, hence its transmission from one host to another (Abraham et al. 2017; Weiss and Aksoy, 2011). Recent studies using insect microbiota to control and prevent vector-borne diseases (Carrington et al. 2018; Hoffmann et al. 2011; Bian et al. 2010) support the need to further explore this avenue. This approach is in line with the

WHO “One Health” Concept which recognizes that the health of people and that of animals are closely linked due to their common shared environment.

Like most hematophagous insects, the tsetse fly harbors bacterial symbionts maternally transmitted to offspring, including *Wigglesworthia*, its primary symbiont which essentially contribute to the fly’s survival by producing and releasing nutrients and vitamins that are insufficiently present in the diet of the latter, and that it is unable to synthesize. In addition, tsetse flies harbor communities of indigenous bacteria acquired from the environment (Dale and Maudlin, 1999; Wang et al. 2013; Aksoy et al 2014). In most cases, exogenous microbiome comprises less than 1% of tsetse’s cumulative enteric microbiota (Aksoy et al. 2014) and their function is yet to be elucidated. Due to their close proximity (gut) these bacteria may interact with the parasite harbored by tsetse flies, for example by exchanging various molecules (Geiger et al., 2010). It thus appears important to identify bacteria naturally harbored by tsetse flies, that are associated with a susceptible or refractory status to trypanosome infection, and that could be used in the gut of the fly to block the establishment of the parasite. In this study, we are investigating whether the taxonomic composition and relative abundance of distinct bacterial taxa harbored by *Glossina palpalis palpalis* may be associated with *Trypanosoma* sp. infection status as well as with the sampling sites on the field.

2. Materials and Methods

2. 1. Sampling area

Tsetse flies were sampled in April 2018 in five villages Campo beach, Ipono, Mvas, Itdonde-Fang, Mabiogo of the Campo focus (2°20' N, 9°52' E), situated in the South region of Cameroon. The Campo focus is known to be hypo-endemic to sleeping sickness where HAT cases are still diagnosed every year (PNLTHA/Cameroon), and where AAT is still active (Figure 1). Campo is located on the Atlantic Coast side, along the Ntem River which separates Cameroon from Equatorial Guinea. In this focus, previous studies reported the presence of several tsetse fly species, namely *G. p. palpalis, G. pallicera, G. caliginea* (that could be a vector for *T. b. gambiense*)*, and G. nigrofusca* (Simo et al. 2008).

2. 2. Tsetse fly trapping and sampling

Pyramidal traps (Gouteux and Lancien, 1986) were deployed in appropriate tsetse fly biotopes (this could be equatorial forest, riversides, farmlands, cocoa lands) in each village for four consecutive days and their geographical position recorded using a global positioning system (GPS) device. The traps were put out overnight and visited twice a day (10 a.m. and 3 p.m.) and the captured flies recovered to prevent desiccation. In the field laboratory, the tsetse fly species (*G. p. palpalis, G. pallicera, G. caliginea and G. nigrofusca*) were determined using morphological taxonomic keys (Grébaut et al. 2004). Prior to dissection, all *G. p. palpalis* were sorted, using morphological criteria, into teneral - young flies that had never taken a blood meal - and non-teneral flies. Only the non-teneral flies were selected for further analysis. A total of 160 flies were randomly selected for dissection among the 540 non-teneral flies.

2. 3. Dissection of tsetse flies

Prior to dissection, the bench was sterilized using 5% sodium hypochlorite, and each fly was sterilized by immersion successively in 5% sodium hypochlorite for 10 min and twice in 70%

ethanol, each for 10 min. Dissection was carried out under sterile conditions, that is in a drop of sterile 0.9% NaCl near the flame of a Bunsen burner. The dissecting instruments were carefully cleaned after the dissection of each fly to prevent cross-contamination. After dissection, tsetse fly whole guts were observed under an optic microscope (magnification x100) to detect the presence of trypanosomes; the tsetse guts were transferred into Eppendorf tubes containing 95% ethanol and stored at room temperature on the field and at -20 °C in the laboratory until further analysis.

2. 4. DNA Extraction

Once in the laboratory, the gut samples were thawed, ethanol was removed by pipetting. The remaining material was air dried. It was disrupted using piston pellets and the DNA was extracted using the DNeasy tissue kit (Qiagen, Paris, France) according to manufacturer instructions. A total of 100 µl of elution buffer was used to recover the DNA extracted from each sample. The concentration of the DNA extracted was measured using a nanodrop (ThermoFisher Scientific, Paris). The DNA samples were stored at -80°C until use.

2. 5. Trypanosome identification by PCR

PCR amplification of parasite DNA was performed using specific primers (Table 1); it was processed as described by Herder et al. (2002): an initial denaturation step at 94°C for 5 min was followed by 40 amplification cycles. Each cycle included a denaturation step at 94°C for 30 s, an annealing step for 30 s at 55°C for *T. brucei* s.l. identification, 60°C for *T. congolense* or 62°C for *T. b. gambiense*, and an extension step at 72°C for 1 min. A final extension was performed at 72°C for 10 min. The amplified products were separated on 2% agarose gel containing ethidium bromide and visualized under UV light. Positive (reference trypanosomes DNA) and negative controls (containing PCR reagents and water instead of DNA) were included in each set of PCR amplification experiment.

2. 6. Meta-barcoding Analysis and Taxonomic assignment

PCR amplification was performed on the highly variable V4 region of the 16S rRNA gene. This region was chosen since it was previously shown that it enables the identification of more bacterial genera than the sequencing of the V3V4 region (Tsagmo Ngoune et al., 2019). More importantly, we have chosen primers allowing to detect a wider variety of bacteria than those used by Jacob et al. (2017), and Tsagmo Ngoune et al. (2019). In practice, the amplification was performed using the forward primer 5'-CAGCAGCCGCGGTAATAC-3' and the reverse primer 5'- CGCATTTCACCGCTACAC-3', according to the manufacturer (MR DNA Laboratory - http://www.mrdnalab.com/shallowater, USA) protocol. The length of the amplicons was between 100 bp to 450 bp. Then the DNA fragments (from each sample (50ng)) was prior linked to adaptors and used to prepare the individual barcode library using Nextera DNA Sample Preparation Kit (Illumina, Singapore). Experion Automated Electrophoresis Station (Bio-Rad, Hercules, USA) was used to determine library insert size. Each PCR reaction was performed in 20 µl (final volume) of a mixture containing 2µl DNA template, 5µl molecular biology-grade water, 10µl HotStarTaq Plus Master Mix, 0.5µl forward primer (10mM), 0.5µl reverse primer (10mM) and 2µl CoralLoad Concentrate. The cycling parameters were a denaturation step at 95°C for 5 min followed by 30 amplification cycles, each cycle including a denaturation step at 94°C for 30 s, an annealing step at 53°C for 40 s, and an extension step at 72°C for 1 min. A final extension was performed at 72°C for 10 min.

PCR products were checked on 2% agarose gel containing ethidium bromide and visualized under UV light to determine the success of amplification. All individual barcoded samples were pooled in equimolar proportions according on their DNA concentrations. Pooled samples were purified using calibrated Ampure XP beads and used to prepare illumina DNA libraries. Pooled library (12pM) was loaded on a 600 Cycles v3 Reagent cartridge (Illumina) and the paired-end sequencing was performed at MR DNA (www.mrdnalab.com, Shallowater, TX, USA) on the Illumina MiSeq sequencing platform according to the manufacturer's guidelines.

Before executing the metabarcoding pipeline, we generated a specific reference file for the assignment step by running CutAdapt v1.18 (Martin, 2011) with the above-mentioned primers to extract V4 reference sequences from the SILVA (Yarza et al. 2014) SSU database (release 132). CutAdapt was used again in the first stage of the bioinformatics workflow in order to filter on read quality with a threshold value of 20. Then, VSearch v2.10.2 (Rognes et al. 2016) and CutAdapt were used alternately in conducting the following tasks: (i) for each sample, merging forward and reverse reads together based on their overlapping areas; (ii) read demultiplexing to generate a single fastq file per sample; (iii) clipping barcodes and primers; (iv) removing sequences containing unknown bases; (v) computing expected error rate; and (vi) dereplicating reads at the sample-level. A single fasta file was then created by pooling remaining sequences together and passed to VSearch for global dereplication. The clustering stage was performed on the remaining unique sequences using Swarm v2.2.2 (Mahé et al. 2015). VSearch was then launched again to mark chimeric clusters.

The STAMPA pipeline (https://github.com/frederic-mahe/stampa) was then used to proceed with taxonomic assignment of the sequences representing Operational Taxonomic Units (OTU) by similarity with those in the specific reference file preliminarily obtained from SILVA SSU records, which lead to an OTU table. Finally, to retain targeted taxa at the genus level, some filtering was applied on these OTUs, that consisted in eliminating clusters with a high expected error (above 0.0002), or which were considered marginal (representing less than 3 sequences and observed in a single sample).

2. 7. Data Analysis

Statistical analyses were performed using Calypso 8.84® (Zakrewski et al. 2016) software. Rarefaction curves were first performed to check the sequencing depth and to ensure the

description and thus the identification of the quasi-totality of the OTUs present in the samples (Zakrewski et al. 2016).

Diversity estimates (α-diversity for within-groups analyses, and β-diversity for between groups comparisons) were calculated with respect to the origin of the flies as well as to their infection status. The α-diversity of the flies' microbiome, which measures the overall diversity including both abundance and evenness of the OTUs, was estimated with the Shannon index; β-diversity analysis was performed to investigate whether there are some differences in the fly microbiota according to infection status and sampling sites. Variation patterns were visualized with principal coordinate analysis (PCoA) using the Bray-Curtis index and their comparison was formally assessed by non-parametric multivariate statistical test, the permutational analysis of variance (PERMANOVA). A complementary method, the Canonical Correspondence Analysis (CCA), was applied. CCA is a multivariate method used to explore complex associations between measured variables and multiple explanatory variables (or confounding factors).

Significant differences in bacterial richness between the infected and uninfected flies, and between the different sampling sites were tested using the non-parametric Kruskal–Wallis test. In order to determine whether or not there is an association between the bacterial flora and the infection status of the fly, we have divided our sample in two groups: the first one was representing bacteria present in both uninfected and infected flies and the second included bacteria occurring only either among infected or uninfected flies. The search for potential genera significantly associated with specific group of flies (infected flies, uninfected flies, flies from each village) was performed using the Wilcoxon rank test and LEfSe (Linear discriminant analysis Effect Size) algorithm (Segata et al. 2011).

For all these analyses, the threshold for significance was set to 5%. Whatever the analysis, a bacterial taxon was considered only in case it was present in at least 5 individuals and with an abundance higher than 0.01% in each group depending on the parameter of interest (trypanosome infected vs uninfected; comparison between different villages).

2. 8. Sample contamination

Although every care systematically used in the field of microbiological research has been taken to avoid contamination of the samples, the occurrence of such a possibility cannot be entirely excluded. However given that a bacterial taxon was considered only in case it was present in at least 5 individuals and with an abundance higher than 0.01%; it can be estimated that most of the possible contaminants were excluded from the study.

3. Results

3. 1. Geographic origin of collected tsetse flies and trypanosome-infection status

As shown in Table 2 (Supplementary Table S1 provides individual data), among the 160 flies analyzed 40 % were collected in Campo-beach, 29% in Mvas, 13% in Ipono, 12% in Mabiogo, and 6% in Itdonde-Fang. Fifty percent of these sampled flies were infected with at least one trypanosome species or subspecies, among which 42.1% in Campo-beach, 38% in Ipono, 47.8% in Mvas, 70% in Itdonde-Fang and 84.2% in Mabiogo. Out of the 80 infected flies, 61 (76.25%) were *Trypanosoma congolense* Savannah-type (TcS) infected [50 (62.5%) single infected; 11 (13.75%) mixed infected]; 13 flies were infected with *Trypanosoma congolense* Forest-type (TcF) [6 (7.5%) single infections, 7 (8.75%) mixed infections]; 18 (22.5%) were infected with *Trypanosoma brucei* sensu-lato (Tbsl) [11 (13.75%) single infections; 7 (8.75%) mixed infections]; 5 flies were infected with *Trypanosoma brucei gambiense* (Tbg), all of them were mixed infected. One fly harbored the 4 trypanosome species together.

3. 2. Sequencing quality control

Raw sequencing data of the V4 region generated a total of 37,573,904 paired-end reads and 96.1% were merged, representing 36,112,664 paired-end reads. Out of these merged reads, 79.8% (29,994,337 reads) were of good quality, representing enough sequencing depth with only a few bad sequences, and therefore were retained for further analyses. The rarefaction curves (Figure 2) confirmed the sequencing quality, with the saturation of most of them between 100,000 and 250,000 reads showing that the sequencing effort was sufficient to characterize most of the OTUs.

3. 3. Global characterization of flies' microbiota

The microbiota composition of the midguts of sampled tsetse flies consisted of 14 phyla + 1 "Uncultured bacteria" corresponding to 261 taxa when assigned at the genus level. However,

only 114 bacterial genera belonging to 12 phyla + the "uncultured bacteria" (Supplementary Table S2) were considered because only these genera were recorded at least once in a sampling site (Supplementary Table S3) with a relative abundance greater than the 0.01% abundance threshold. *Proteobacteria* were present in all the 160 samples and had a mean relative abundance of 90.7% of the total microbiota, while the other phyla (as well as their respective relative abundances) were unevenly distributed across the different flies processed. The predominance of *Proteobacteria* was mainly due to the very high relative abundance of *Wigglesworthia* (87.22%), the primary symbiont of the tsetse fly which plays a crucial role in fly's survival. The phylum *Tenericutes* included only one genus (*Spiroplasma*) with a relative abundance of 6.2%; *Actinobacteria*, *Bacteroidetes* and *Firmicutes* were present with a mean relative abundance of 0.83, 0.60 and 0.86%, respectively.

By looking more closely the distribution of the bacterial community of the samples, 10 of them presented profiles displaying large abundance differences with reference to the global mean repartition of the bacterial phyla (Supplementary Table S4). So, while the *Proteobacteria* phylum is, on average, the most abundant, this is not the case, for example in samples 1, 39, 41, 57 and 120 where the most abundant phylum is *Tenericutes*. In sample 155, the relative abundance of *Bacteroidetes*, *Firmicutes, Proteobacteria* were 21.20%, 17.37% and 41.9%, respectively. In sample 20 and 44 the respective relative abundances were 50.29% and 46.82% for *Proteobacteria*, 19.71% and 13.84% for *Actinobacteria*, 11% and 9.7% for *Bacteroidetes* and 10.73% and 14.48% for *Firmicutes*. Finally, despite the prominence of *Proteobacteria* phylum, the bacterial genera it includes, as well as their respective relative abundance, are themselves very unevenly distributed as shown in Supplementary Table S5 for the 10 selected flies, where the relative abundance of *Wigglesworthia* itself varies between 4.1 and 95.4% (mean relative abundance: 87.2%), and where, in sample 155, *Tepidiphilus*, a genus including 4 extremophilic (thermophilic) bacterial species becomes prominent with 31.6% relative abundance.

3. 4. Distribution of the bacteria in flies with reference to the 5 sampling sites

Supplementary Table S3 shows the distribution of the taxa and their relative abundance according to the sampling sites. Regarding *Wigglesworthia*, its relative abundance was prominent in all sites (more than 85% in Ipono, 86% in Campo-beach and Mvas, 89% in Mabiogo and 96% in Itdonde-Fang), despite the large differences of occurrences in some individuals (for example, the *Wigglesworthia* relative abundance was lowest (4.1%) in sample 155 from Mabiogo. It was highest (95.4%) in sample 94 from Campo beach, while it was only 17.9% in sample 8 trapped in the same village) (Supplementary Table S4). In contrast, relative abundance of *Spiroplasma* varied from 1.7% in flies from Itdonde-Fang to 9.1% in those from Ipono. Similarly, *Acinetobacter* relative abundance was 0.057% in Mabiogo vs 2.65% in Mvass, and that of *Pseudomonas* was 0.05% and 1.58% in Mabiogo and Mvas, respectively. Apart such diversity regarding taxa present in flies from all foci, several other taxa with relative abundances higher than the 0.01% threshold, were only present in a given collection site. These taxa and corresponding sampling sites are presented in Supplementary Table S6. In contrast, some taxa (*Aerococcus*, *Brachybacterium*, *Kocuria*, *Nesterenkonia* and *Prevotella*) were specifically absent (abundance threshold <0.01) in Itdonde-Fang site; similarly, *Leuconostoc* was absent in Mabiogo. Finally, some taxa were not detected at all in some sites; this is the case of *Anaerospora* in flies from Mvas; *Janibacter, Dyadobacter, Phascolarctobacterium* and *Zymophilus* in Mabiogo*, Myxosarcina, Anaerospora, Zymophilus, Pseudoxanthomonas* in Itdonde-Fang*,* and *Proteiniphilum, Anaerospora* in Ipono.

3. 5. Comparative abundance of the bacteria in infected and uninfected tsetse flies

Supplementary Table S7 presents the relative abundance of the overall bacterial genera in infected and uninfected flies. The relative abundance of the genus *Wigglesworthia* was very high (87.2%), both in infected and uninfected flies, though the abundance of this symbiont was significantly higher among infected flies compared to non-infected ones (p=0.032) (Table 3).

We compared the abundance of 11 taxa that displayed highest relative abundance in infected and / or in uninfected flies. Statistically significant differences were observed (Table 3) that, in contrast to *Wigglesworthia*, associates low relative abundance of the genera *Acinetobacter* (p=0.051), *Bacillus* (p=0.008), *Cutibacterium* (p=0.00012), *Pseudomonas* (p=0.0063) and *Tepidiphilus* (p=0.025) to susceptibility of the tsetse fly to trypanosome infection. In addition, when compiling the use of Wilcoxon rank test and LEfSe score as well as relative abundances in infected or uninfected flies, it appeared that the genera *Bacteroides* (p=0.0042), *Georgenia* (p=0.0097), *Prosthecobacter* (p=0.035), even though of low abundance, were also shown to be associated with fly infection. Moreover, as shown in Supplementary Table S7, 36 bacterial genera were shown to be associated to uninfected flies only (they are listed in Table 4). Most of them were present in uninfected flies in low relative abundance (between 0.01 and 0.03%), whereas *Paenibacillus* and *Ralstonia* exhibited a relative abundance around 0.03%, two others, *Legionela and Novosphingobium,* a relative abundance around 0.04%, and *Cupriavidus* a relative abundance of 0.21%. In contrast, *Cellulosimicrobium* and *Geobacillus* were present only in infected flies with a relative abundance of 0.013 and 0.018%, respectively.

3. 6. Bacterial diversity in flies (alpha-diversity)

Globally no statistically significant differences were observed when comparing the bacterial diversity of either *Trypanosoma* infected flies to uninfected ones (p=0.11) (Figure 3A), or when comparing the flies from different villages (p=0.22) (Figure 3B). In contrast, significant differences were observed when comparing the microbiota richness of the infected flies versus uninfected (p=0.003) (Figure 3C) as well as when comparing microbiota richness among flies from different villages (p=6.2x10^{-5}) (Figure 3D). Thus, we investigated whether a given species of trypanosome could be associated to these differences. A significant difference in the composition of bacterial flora was shown in flies hosting TcS when compared to the microbiota composition of uninfected ones (p=0.047; Figure 3E), while no significant differences were

shown when comparing the microbiota composition of respectively Tbrsl (p=0.83; Figure 3F) or TcF (p=0.57; Figure 3G) infected flies with uninfected flies.

3. 7. Multivariate analysis (beta-diversity)

The Principal Coordinate Analysis (PCoA), showing the distribution of the 160 samples (flies) according to their microbiota composition (bacterial distribution) (Figure 4), has highlighted 4 groups depending on the collection site (Figure 4C) and the permutational variance analysis (PERMANOVA) based on each of these components (bacterial distribution and relative abundance) showed that there was statistically significant differences between the composition of the fly microbiota by sampling site with a p-value=0.0003.

Furthermore, when the samples were identified by their infection status, the distribution of samples from infected and non-infected flies in general or non-infected flies and flies infected by different subspecies of trypanosomes are homogeneous. This was observed whatever the 4 clusters (Figure 4A, 4B), although multivariate analysis showed statistically significant differences depending on infectious status (p=0.0003) and the subspecies of parasites (p=0.0001).

To refine the results, a Canonical Correspondence Analysis (CCA) (Figure 5), was performed which has pinpointed these differences in taxa composition between infected flies and uninfected flies (p-value=0.001). Finally a hierarchical clustering resulted in a dendrogram which revealed the existence of several clusters associated neither with the infection status of tsetse flies, nor with the sampling sites where the flies were originated from (Figure 6). All these results may indicate that beyond the bacterial composition and or sampling sites (environment differences), there are other factors that could modulate the flies' microbiome composition.

4. Discussion

The present study is part of a general project aiming at investigating the bacterial flora of field-collected tsetse flies and attempting to determine their possible involvement in the establishment of the trypanosome into the gut of the tsetse flies, hence in modulating the flies' vector competence. Previous studies conducted by Jacob et al. (2017) and Tsagmo Ngoune et al. (2019) characterized the intestinal microflora of tsetse flies from several HAT and AAT active foci in the Southern region of Cameroon, including the Campo sleeping sickness focus where the tsetse flies included in the present study were collected. In this study, the bacterial composition of the microbiome of 80 flies infected by trypanosomes, and that of 80 uninfected flies were analyzed using the taxo-genomics approach based on the sequencing of the hyper variable V4 region of the small 16S rRNA gene. It resulted in the identification of 261 bacterial genera of which only 114 exhibited relative abundance higher than the 0.01% threshold, while 83 genera were identified by Jacob et al. (2017), and only 10 by Tsagmo Ngoune et al. (2019) with similar methods. Despite these large differences in the number of bacteria, most of them belonged to the previously identified four phyla, *Actinobacteria*, *Bacteroides*, *Firmicutes* and the *Proteobacteria*, together with an additional phylum, *Tenericutes* that was only represented by the genus *Spiroplasma*. These phyla are commonly characterized, in variable abundance, in the intestine of a variety of insects (Andreotti et al. 2011; Boissière et al. 2012; Zouache et al. 2011; Terenius et al. 2012; Dillon et al. 2008; Geib et al. 2009; Gouveia et al. 2008). Compared to these insects, tsetse flies present a crucial difference in so far as the *Proteobacteria* are overabundant; on average this phylum represents around 90% relative abundance, followed by *Tenericutes* with only one species (*Spiroplasma*) showing a relative abundance around 6%, while each of the three other phyla (*Actinobacteria*, *Bacteroides* and *Firmicutes*) usually account only for about 1%.

Moreover, *Proteobacteria* overabundance is due to that of one bacterium, *Wigglesworthia,* the tsetse flies obligate primary symbiont hosted by all the flies, contributing to fly fitness (Snyder et al. 2015). The overabundance of *Wigglesworthia* as compared to that of the other tsetse fly intestinal bacteria is systematically reported in the literature (Wang et al. 2013; Aksoy et al. 2014; Snyder and Rio, 2015; Jacob et al. 2017; Tsagmo Ngoune et al. 2019), and is considered as responsible for the disruption of a proper identification of low abundant bacteria, and therefore for an underestimation of the fly's gut bacteria diversity. All these phyla are unequally distributed across the 160 samples. Supplementary Tables S3 and S4 provide an overview of the very important diversity in species richness among 10 selected samples. So, the relative abundance of the *Proteobacteria* was 22.4% in sample 39 while it was 98.9% in sample 94; that of *Tenericutes* (*Spiroplasma*) was from 0.16% in sample while it was 115 to 77.2% in sample 39. *Wigglesworthia* itself, despite its overabundance "status", is highly unevenly distributed: 4.14% of relative abundance in sample 155, 95.4% in sample 94, and 43.96 % in sample 1, which represents respectively 9.9%, 96.4% and 99.5% of the *Proteobacteria* relative abundance. Finally, large diversity was also shown regarding *Proteobacteria* other than *Wigglesworthia*. For example, the relative abundance of *Pseudomonas* was 0.05% in sample 1 while it was 5% (100 times more) in sample 8; more surprisingly, while relative abundance of *Wigglesworthia* was only 4.1% in sample 155, that of the *Tepidiphilus* (a genus that includes four species of extremophilic bacteria (Xiao-Tong et al., 2020) reached 31.6%. Where and how did tsetse fly ingest such kind of bacteria that are found in thermophilic water treatment sludge, petroleum reservoir or terrestrial hot spring? (Manaia et al. 2003; Salinas et al. 2004; Poddar et al. 2014). It was reported that the tsetse fly could ingest bacteria present on the epidermis (Poinar et al. 1979) of humans or a variety of vertebrates (Moloo et al. 1988; Simo et al. 2008) or on plants' nectar (Colman et al. 2012; Solano et al. 2015) when they feed on them.

In order to check if there would be a bacterial dynamic during this last decade, we have made a listing and a comparison of the bacteria present during the previous studies carried out in this

site. One may note that the comparison has only an "indicative" value since different technics were used in these studies for microbiome composition determination i.e. culture-dependant, and molecular with different regions amplified and different primers used (total 16S, V3V4 or V4 only). One may also remember that the molecular method allows identifying bacteria that have been ingested by the flies, but that it does not certify they are still alive and are established in the tsetse fly. The number of identified bacteria was also very different: 261 in the present investigation, 10 in 2019 and 83 in 2017 (molecular methods) and 9 in 2011 (culture method). One may also note that the very low number of bacteria reported in 2019 (Tsagmo Ngoune et al. 2019) was possibly due to the exceptionally high relative abundance of *Wigglesworthia* (98.9%) which most probably disrupted identifying low abundant bacteria. Supplementary Table S7 shows the bacteria genera identified in flies collected during at least two entomological campaigns. Fifty-six bacteria corresponded to the criteria except for *Sphingobacterium* which was identified only when using the culture method, while *Tsukamurella*, *Lactococcus*, *Staphylococcus*, *Acinetobacter*, *Pseudomonas*, *Serratia* and *Spiroplasma*, were found in flies from 3 sampling campaigns. Of interest is that some of the 56 bacteria are able to secrete antiparasitic compounds (Poinar et al., 1979; Azambuja et al. 2004). Such bacteria may impact the establishment of protozoan parasites in their insect vectors as it was reported for mosquitoes (Pumpuni et al., 1996; Beier et al., 1994; Straif et al., 1998). So, for example, bacteria belonging to the genus *Enterobacter* spp. are able to produce pigments, such as prodigiosin (Moss, 2002), which is toxic to *Plasmodium falciparum* (Lazaro et al., 2002) and to *Trypanosoma cruzi* (Azambuja et al., 2004). *S. marcescens* isolated from *Rhodnius prolixus* has been reported to be able to lyse *Trypanosoma cruzi* in vitro (Azambuja et al., 2004). Similarly regarding *Pseudomonas fluorescens* strains that were shown to be able to lyse *T. cruzi* (Mercado and Colon-Whitt, 1982). Given that similar bacteria species have been isolated from the midgut of tsetse flies, further investigations deserve to be undertaken in order to verify whether they could modulate the tsetse fly vector competence.

Regarding *Sodalis*, the tsetse flies' secondary symbiont that was previously shown to favor fly infection with the trypanosome (Farikou et al. 2010; Hamidou Soumana et al. 2014), this bacterium was identified in very low abundance, thus below the fixed threshold. It was also of low abundance in flies collected during 2 other campaigns (Tsagmo Ngoune et al., 2019; Jacob et al., 2017). Possibly differences in abundance and prevalence between previous and present investigations may be due to differences in the molecular approach, former being performed with *Sodalis* specific primers. Nevertheless, its maternal transmission from one generation to the next makes it, via para-transgenesis, a good vector for in situ expression of anti-parasitic molecules able to prevent the establishment of the parasite in the gut of tsetse flies.

The large diversity of bacteria, the very low abundance of most of them, and their uneven distribution make difficult the analysis of a possible association between the composition of the microbiome and the susceptibility / resistance of the fly to trypanosome infection. The same is true regarding a possible association with the fly sampling sites and, consequently, with a possible effect of ecological factors on the composition of the microbiome. Several statistical models have been used to analyze the data, additional ones deserve to be tested in further investigations. Despite these difficulties, original results have been obtained. So, 36 bacterial genera have been identified only in uninfected flies. The interest of this finding has however to be modulated since their relative abundance was very low. In fact, the relative abundance of 26 genera out of 36 was lesser than 0.02%, it was between 0.02% and 0.05% for 9 of them, and finally only one bacterial genus, *Cupriavidus*, had a real significant relative abundance of 0.22%. Similarly, the presence or the absence of several bacterial genera seems to be specifically associated to trypanosome infection of the tsetse flies or to the fly collection sites.

Statistical analyzes - Shannon Index, PCoA, CCA and Hierarchical clustering - provide conflicting results. However, some of them express a tendency towards an association between the microbiome, its diversity in terms of composition and individual bacteria abundance, and the infection of these flies by the trypanosome ($p=0.0003$); similarly, regarding the association

between microbiome and sampling site (p=0.0003). These results differ from those that Jacob et al. (2017) and Tsagmo Ngoune et al. (2019) recorded in the same region; the difference would come from the fact that they focused on the comparison between foci and not between the sampling sites (villages) within a given focus (here the Campo focus). It should be noted that in the five investigated villages the main anthropogenic activity is fishing. Besides, depending on the villages, agriculture, livestock, hunting and logging are carried out at different levels of intensity, which would promote differences in the environmental wild and domestic fauna thus in the diversity of bacteria that tsetse flies could possibly ingest when taking their blood meal.

The aim of this study was to investigate the association between the microbiome composition of the tsetse fly and the presence of trypanosomes, that may modulate its vector competence, even though the complexity of the interactions is not well identified at this stage (one may also remember that as for the gut identified bacteria, the molecular method allows identifying trypanosomes that have been ingested by the flies, but that it does not assess they are still alive and are establish in the tsetse fly). Tsetse flies harboring *Trypanosoma congolense* "Savanah type" exhibited a different composition of bacterial flora compared to uninfected flies. Importantly, some bacteria were strongly associated with uninfected status of tsetse flies. The culture of bacteria and their functional analysis would inform on how they can modulate tsetse flies' vector competence, likely useful to the parasite transmission by para-transgenic approach. Moreover, Mirzaei and Maurice (2017) introduced an additional contributor, bacteriophages, which might be able to modulate the composition of the human microbiome. Results reported by Hamidou Soumana et al. (2014) showed that feeding tsetse flies, whether susceptible or refractory to parasite infection, with infected blood meal resulted in a strong alteration of *Sodalis* gene expression. Most surprisingly was that genes involved in ''lysozyme activity, bacteriolytic enzyme, defense response, peptidoglycan catabolic process, cytolysis, cell death, cell wall macromolecule catabolic process'' were among the most overexpressed genes in flies refractory to trypanosome infection. This finding, together with the puzzling results described in the

literature, highlights the possible role of a bacteriophage as a major actor in tsetse fly refractoriness; this role potentially would result in both decreasing *Sodalis* population density (given that the symbiont favors tsetse fly infection by the trypanosomes), and in triggering the fly's innate immune response. The involvement of bacteriophages cannot be excluded since the analysis of *Sodalis* transcriptome revealed the presence, into the bacterial genome, of DNA sequences of viral origin. This indicates the presence of a prophage integrated into the *Sodalis* genome and means the bacterium, or some of its strains, to be lysogenic. This enlarges the field of investigation which deserves to be continued as to attempt to decipher the complex interactions between the three partners, the tsetse fly, the intestinal bacteria and the trypanosome, since the results, although partial, recorded in the tsetse flies as well as in other insect vectors of parasites, indicate that they intervene in their vector competence.

Conclusive remarks and perspectives

The molecular approach used in this study makes it possible to identify bacteria that might be of interest if alive, as they could then be isolated from the gut, cultivated and tested for their real activity as effectors able to modify some biological characteristics of their host - here the modulation of the vector competence of the tsetse fly. In previous investigation we used the culture dependent method which allowed the isolation of several bacteria from the gut of the tsetse fly (Geiger et al, 2009; 2011) and even to identify *Serratia glossinae*, a novel bacterial species (Geiger et al, 2010). Similar approach was performed in East Africa, namely by Malele et al (2018).

Aside, a comparative study of the microbiome of female and male tsetse flies could provide important information knowing that bacteria, such as *Sodalis*, can be transmitted matrilineally from one generation to the next. Analyzing the microbiome of field captured teneral flies could give a response. Such bacteria could become alternatives to *Sodalis* in the context of a para-transgenic approach.

Acknowledgements

We would like to thank Nana Djeunga Hugues Clotaire who provided language help and proof reading of this paper.

Funding

The authors thank the Labex Parafrap and the *Institut de Recherche pour le Développement* (IRD) for their financial support. FSNF is a PhD student supported by *Allocations de recherche pour une thèse au Sud* (ARTS) fellowship from the IRD/Labex Parafrap.

The collection of samples was supported by International Atomic Energy Agency (IAEA) funding.

Author contribution

AG conceived the experiments. FSNF, JMTN, and FN performed the experiments. FSNF, JMTN, GS and AG analyzed the data. AG contributed reagents/ materials/analysis tools. FSNF, JMTN, FN, DB and AG wrote the paper.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Abraham, N.M., Liu, L., Jutras, B.L., Yadav, A.K., Narasimhan, S., Gopalakrishnan, V., Ansari, J.M., Jefferson, K.K., Cava, P., Jacobs-Wagner, C., Fikrig, E., 2017. Pathogen-mediated manipulation of arthropod microbiota to promote infection. Proc Natl Acad Sci U S A. 114, E781–E90.

Aksoy, E., Telleria, E.L., Echodu, R., Wu Y., Okedi, L.M., Weiss, B.L., Aksoy, S., Caccone, A., 2014. Analysis of multiple tsetse fly populations in Uganda reveals limited diversity and species-specific gut microbiota. Appl. Environ. Microbiol. 80,4301–4312

Andreotti, R., Pérez de León, A.A., Dowd, S.E., Guerrero, F.D., Bendele, K.G., Scoles, G.A., 2011. Assessment of bacterial diversity in the cattle tick *Rhipicephalus* (*Boophilus*) *microplus* through tag-encoded pyrosequencing. BMC Microbiol. 11,6.

Azambuja, P., Feder, D., Garcia, E.S., 2004. Isolation of *Serratia marcescens* in the midgut of *Rhodnius prolixus*: impact on the establishment of the parasite *Trypanosoma cruzi* in the vector. Exp Parasitol. 107,89-96.

Baker, N., De Koning, H.P., Mäser, P., Horn, D., 2013. Drug resistance in African trypanosomiasis: the melarsoprol and pentamidine story. Trends Parasitol. 29,110-8.

Beier, M.S., Pumpuni, C.B., Bizio, J.C., Davis, J.R., 1994. Effect of paraaminobenzenoic acid, insulin, and gentamicin on *Plasmodium falciparum* development in Anopheline mosquitoes (*Diptera*: *Culicidae*). J. Med. Entomol. 31, 561–565.

Bian, G., Xu, Y., Lu, P., Xie, Y., Xi, Z., 2010. The endosymbiotic bacterium *Wolbachia* induces resistance to dengue virus in Aedes aegypti. PLoS Pathog. 6,e1000833.

Boissière, A., Tchioffo, M.T., Bachar, D., Abate, L., Marie, A., Nsango, S.E., Shahbazkia, H.R., Awono-Ambene, P.H., Levashina, E.H., Christen, R., Morlais, I., 2012. Midgut Microbiota of the Malaria Mosquito Vector *Anopheles gambiae* and Interactions with *Plasmodium falciparum* Infection. PLoSPathog. 8,e1002742

Brun, R., Blum, J., Chappuis, F., Burri, C., 2010. Human african trypanosomiasis. Lancet. 375, 148-59.

Büscher, P., Cecchi, G., Jamonneau, V., Priotto, G., 2017. Human african trypanosomiasis. Lancet. 390, 2397-409.

Carrington, L.B., Tran, B.C.N., Le, N.T.H., Luong, T.T.H., Nguyen, T.T., Nguyen, P.T., Nguyen, C.V.V., Nguyen, H.T.C., Vu, T.T., Vo, L.T., Le, D.T., Vu, N.T., Nguyen, G.T., Luu, H.Q., Dang, A.D., Hurst, T.P., O'Neill, S.L., Tran, V.T., Kien, D.T.H., Nguyen, N.M., Wolbers, M., Wills, B., Simmons, C.P., 2018. Field- and clinically derived estimates of Wolbachia-mediated blocking of dengue virus transmission potential in Aedes aegypti mosquitoes. Proc Natl Acad Sci U S A. 115,361–6.

Colman, D.R., Toolson, E. C., Takacs-Vesbach, C.D., 2012. Do diet and taxonomy influence insect gut bacterial communities? Mol Ecol. 21,5124-37.

Dale, C., Maudlin, I., 1999. *Sodalis* gen. nov. and *Sodalis glossinidius* sp. nov., a microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. Int J Syst Bacteriol. 49,267–75.

Delespaux, V., De Koning, H.P., 2007. Drugs and drug resistance in African trypanosomiasis. Drug Resist Updat. 10,30–50.

De Vooght, L., Van Keer, S., Van Den Abbeele, J., 2018. Towards improving tsetse fly paratransgenesis: stable colonization of *Glossina morsitans morsitans* with genetically modified Sodalis. BMC Microbiol.18,165.

Dillon, R.J., Webster, G., Weightman, A.J., Dillon, V.M., Blanford, S., Charnley, A.K., 2008. Composition of Acridid gut bacterial communities as revealed by 16S rRNA gene analysis. J. Invertebr. Pathol. 97, 265–272.

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.

Franco, J.R., Cecchi, G., Priotto, G., Paone, M., Diarra, A., Grout, L., Mattioli, R.C., Argaw D., 2018. Monitoring the elimination of human african trypanosomiasis: Update to 2016. PLoS Negl Trop Dis. 12, e0006890.

Geib, S.M., Jimenez-Gasco Mdel, M., Carlson, J.E., Tien, M., Hoover, K., 2009. Effect of host tree species on cellulase activity and bacterial community composition in the gut of larval Asian longhorned beetle. Environ. Entomol. 38, 686–699.

Geiger, A., Fardeau, M.-L., Falsen, E., Ollivier, B., Cuny, G., 2010. *Serratia glossinae* sp. nov., isolated from the midgut of the tsetse fly Glossina palpalis gambiensis. Int J Syst Evol Microbiol. 60,1261-5.

Geiger, A., Fardeau, M.-L., Grebaut, P., Vatunga, G., Josénando, T., Herder, S., Cuny, G., Truc, P., Ollivier, B., 2009. First isolation of *Enterobacter*, *Enterococcus*, and *Acinetobacter* spp. as inhabitants of the tsetse fly *(Glossina palpalis palpalis*) midgut. Infect Genet Evol. 9,1364-70.

Geiger, A., Hirtz, C., Bécue, T., Bellard, E., Centeno, D., Gargani, D., Rossignol, M., Cuny, G., Peltier, J.B., 2010. Exocytosis and protein secretion in *Trypanosoma*. BMC Microbiol. 10,20.

Geiger, A., Fardeau, M.-L., Njiokou, F., Joseph, M., Asonganyi, T., Ollivier, B, Cuny, G., 2011. Bacterial Diversity Associated with Populations of *Glossina* spp. from Cameroon and Distribution within the Campo Sleeping Sickness Focus. Microb. Ecol. 62,632-643.

Geiger, A., Malele, I., Abd-Alla, A.M., Njiokou, F., 2018. *Trypanosoma*-tsetse fly interactions: current and potential future research directions. BMC Microbiol. 18,162.

Gouteux, J.P., Lancien, J., 1986. Le piège pyramidal à tsé-tsé (*Diptera*: *Glossinidae*) pour la capture et la lutte. Essais comparatifs et description de nouveaux systèmes de capture. Trop. Med. Parasitol. 37, 61–66.

Gouveia, C., Asensi, M.D., Zahner, V., Rangel, E.F., Oliveira, S.M., 2008. Study on the bacterial midgut microbiota associated to different Brazilian populations of *Lutzomyia longipalpis* (Lutz & Neiva) (*Diptera*: *Psychodidae*). Neotrop. Entomol. 37, 597–601.

Grébaut, P., Mbida, J.A.M., Kondjio, C.A., Njiokou, F., Penchenier, L., Laveissiére, C., 2004. Spatial and Temporal Patterns of Human African Trypanosomosis (HAT) Transmission Risk in the Bipindi Focus, in the Forest Zone of Southern Cameroon. Vector-Borne Zoonotic Dis. 4,230–238.

Hamidou Soumana, I., Loriod, B., Ravel, S., Tchicaya, B., Simo, G., Rihet, P., Geiger, A. 2014. The transcriptional signatures of *Sodalis glossinidius* in the *Glossina palpalis gambiensis* flies negative for *Trypanosoma brucei gambiense* contrast with those of this symbiont in tsetse flies positive for the parasite: possible involvement of a *Sodalis*-hosted prophage in fly Trypanosoma refractoriness? Infect Genet Evol 24,41–56.

Herder, S., Simo, G., Nkinin, S., Njiokou, F., 2002. Identification of trypanosomes in wild animals from Southern Cameroon using the polymerase chain reaction (PCR). Parasite. 9,345–349.

Hoffmann, A.A., Montgomery, B.L., Popovici, J., Iturbe-Ormaetxe, I., Johnson, P.H., Muzzi, F., Greenfield, M., Durkan, M., Leong, Y.S., Dong, Y., Cook, H., Axford, J., Callahan, A.G., Kenny, N., Omodei, C., McGraw, E.A., Ryan, P.A., Ritchie, S.A., Turelli, M., O'Neill SL., 2011. Successful establishment of *Wolbachia* in Aedes populations to suppress dengue transmission. Nature. 476,454–7.

Horn, D., 2014. Antigenic variation in African trypanosomes. Mol Biochem Parasitol. 195,123–9.

Jacob, F., Melachio, T.T., Njitchouang, G.R., Gimonneau, G., Njiokou, F., Abate, L., Christen, R., Reveillaud, J., Geiger, A., 2017. Intestinal Bacterial Communities of Trypanosome-Infected and Uninfected *Glossina palpalis palpalis* from Three Human African Trypanomiasis Foci in Cameroon. Front. Microbiol. 8, 1464.

Kariithi, H.M., Meki, I.K., Schneider, D.I., De Vooght, L., Khamis, F.M., Geiger, A., Demirbaş-Uzel, G., Vlak, J.M., iNCE, I.A., Kelm, S., Njiokou, F., Wamwiri, F.N., Malele, I., Weiss, B.L., Abd-Alla, A.M.M., 2018. Enhancing vector refractoriness to trypanosome infection: achievements, challenges and perspectives. BMC Microbiol. 18,179.

Kato, C.D., Nanteza, A., Mugasa, C., Edyelu, A., Matovu, E., Alibu, V.P., 2015. Clinical profiles, disease outcome and co-morbidities among t. B. Rhodesiense sleeping sickness patients in uganda. PLOS ONE. 10, e0118370.

Lazaro, J.E., Nitcheu, J., Predicala, R.Z., Mangalindan, G.C., Nesslany, F., Marzin, D., Concepcion, G.P., Diquet, B., 2002. Heptyl prodigiosin, a bacterial metabolite is antimalarial in vivo and non-mutagenic in vitro. J. Nat. Toxins 11, 367– 377

Lindner, A.K., Lejon, V., Chappuis, F., Seixas, J., Kazumba, L., Barrett, M.P., Mwamba, E., Erphas, O., Akl, E.A., Villanueva, G., Bergman, H., Simarro, P., Ebeja, A.K., Priotto, G., Franco, J.R., 2020. New who guidelines for treatment of gambiense human african trypanosomiasis including fexinidazole: Substantial changes for clinical practice. Lancet Infect Dis. 20, e38-e46.

Mahamat, M.H., Peka, M., Rayaisse, J.B., Rock, K.S., Toko, M.A., Darnas, J., Brahim, G.M., Alkatib, A.B., Yoni, W., Tirados, I., Courtin, F., Brand, S.P.C., Nersy, C., Alfaroukh, I.O., Torr, S.J., Lehane, M.J., Solano, P., 2017. Adding tsetse control to medical activities contributes to decreasing transmission of sleeping sickness in the Mandoul focus (Chad). PLoS Negl Trop Dis. 11,e0005792.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., Dunthorn, M., 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. PeerJ 3,e1420.

Malele, I., Nyingilili, H., Lyaruu, E., Tauzin, M., Ollivier, B., Cayol, J.-L., Fardeau, M.-L., Geiger, A., 2018. Bacterial diversity obtained by culturable approaches in the gut of *G. pallidipes* population from a non-sleeping sickness focus in Tanzania: preliminary results. BMC Microbiol. 18,164.

Manaia, C.M., Nogales, B., Nines, O.C., 2003. *Tepidiphilus margaretifer* gen. nov. sp. nov., isolated from a thermophilic aerobic digester. Int. J. Syst. Evol. Microbiol. 53,1405-1410.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10-12

Masiga, D.K., Smyth, A.J., Ayes, P., Bromidge, T.J., Gibson, W.C., 1992. Sensitive detection of trypanosomes in tsetse flies by DNA amplification. Int. J. Parasitol. 22, 909–918.

Mercado, T.I., Colon-Whitt, A., 1982. Lysis of *Trypanosoma cruzi* by *Pseudomonas fluorescens*. Antimicrob. Agents 22, 1051–1057.

Mirzaei, M.K., Maurice, C.F., 2017. Ménage à trois in the human gut: interactions between host, bacteria and phages. Nat Rev Microbiol. 15,397-408

Mogi, M., Teji, S., 1991. Advances in Disease Vector Research. Springer, New York, pp. 47–76.

Moloo, S.K., Grootenhuis, J.G., Kar, S.K., Karstad, L., 1988. Survival and reproductive performance of female *Glossina morsitans morsitans* when maintained on the blood of different species of wild mammals. Med. Vet. Entomol. 2, 347–350.

Moser, D.R., Cook, G.A., Ochs, D.E., Bailey, C.P., McKane, M.R., Donelson, J.E., 1989. Detection of *Trypanosoma congolense* and *Trypanosoma brucei* subspecies by DNA amplification using the polymerase chain reaction. Parasitology 99, 57–66.

Moss, M., 2002. Bacterial pigments. Microbiologist 3, 10–12.

Nakayima, J., Nakao, R., Alhassan, A., Mahama C, Afakye K, Sugimoto C., 2012. Molecular epidemiological studies on animal trypanosomiases in Ghana. Parasites & Vectors. 5, 217.

Poddar, A., Lepcha, R.T., Das, S.K., 2014. Taxonomic study of the genus *Tepidiphilus*: transfer of *Petrobacter succinatimandens* to the genus *Tepidiphilus* as *Tepidiphilus succinatimandens* com. nov. emended description of the genus *Tepidiphilus* and description of *Tepidiphilus thermophiles* sp.nov. isolated from a terrestrial hot spring. Int. J. Syst. Evol. Microbiol. 64,228-235.

Poinar, G.O. Jr., Wassink, H.J., Leegwater-van der Linden, M.E., van der Geest, L. P., 1979. *Serratia marcescens* as a pathogen of tsetse flies. Acta Trop. 36, 223–227.

Pumpuni, C.B., DeMaio, J., Kent, M., Davis, J.R., Beier, J.C., 1996. Bacterial population dynamics in three anopheline species: the impact on *Plasmodium* sporogonic development. Am. J. Trop. Med. Hyg. 54, 214–218.

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4,e2584

Salinas, M.B., Fardeau, M.-L., Cayol, J.-L., Casalot, L., Patel, B.K.C., Thomas, P., Garcia, J-L., Ollivier, B., 2004. *Petrobacter succinatimandens* gen. Nov., sp. nov., a moderately thermophilic, nitrate-reducing bacterium isolated from an Australian oil weil. Int. J. Syst. Evol. Microbiol. 54,645-649.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., Huttenhower, C., 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12: R60

Shaw, A.P., Torr, S.J., Waiswa, C., Cecchi, G., Wint, G.R., Mattioli, R.C., Robinson, T.P., 2013. Estimating the costs of tsetse control options: an example for Uganda. Prev. Vet. Med. 110, 290–303.

Simo, G., Njiokou, F., Mbida Mbida, J.A., Njitchouang, G.R., Herder, S., Asonganyi, T., Cuny, G., 2008. Tsetse fly host preference from sleeping sickness foci in Cameroon: Epidemiological implications. Infect. Genet. Evol 8,34–39

Snyder, A.K., Rio, R.V.M., 2015. "*Wigglesworthia morsitans*" Folate (Vitamin B 9) Biosynthesis Contributes to Tsetse Host Fitness. Appl. Environ. Microbiol. 81,5375–5386.

Solano, P., Salou, E., Rayaisse, J.-B., Ravel, S., Gimonneau, G., Traore, I., Bouyer, J., 2015. "Do tsetse flies only feed on blood?" Infect. Genet. Evol. 36, 184–189.

Straif, S.C., Mbogo, C.N.M., Toure, A.M., Walker, E.D., Kaufman, M., Toure, Y.T., Beier, J.C., 1998. Midgut bacteria in *Anopheles gambiae* and *A. funestus* (*Diptera*: *Culicidae*) from Kenya and Mali. J. Med. Entomol. 35, 222–226

Terenius, O., Lindh, J.M., Eriksson-Gonzales, K., Bussière, L., Laugen, A.T., Bergquist, H., Titanji, K., Faye, I., 2012. Midgut bacterial dynamics in Aedes aegypti. FEMS Microbiol. Ecol. 80, 556–565.

Tessema, M.B., Girma, M., Dagnachew, S., 2014. African trypanosomes: Virulence factors, pathogenicity and host responses. Journal of Veterinary Advances. 4, 732-45.

Tsagmo Ngoune, J.M., Reveillaud, J., Sempere, G., Njiokou, F., Melachio, T.T., Abate, L., Tchioffo, M., Geiger, A., 2019. The composition and abundance of bacterial communities residing in the gut of *Glossina palpalis palpalis* captured in two sites of southern Cameroon. Parasit Vectors. 12,151.

Wang, J., Weiss, B.L., Aksoy, S., 2013. Tsetse fly microbiota: form and function. Front Cell Infect Microbiol 3.

Weiss, B., Aksoy, S., 2011. Microbiome influences on insect host vector competence. Trends Parasitol. 27,514–22.

Welburn, S.C., Maudlin, I., 2012. Priorities for the elimination of sleeping sickness. Adv Parasitol. 79,299-337.

WHO, 2019. Trypanosomiasis, human African (sleeping sickness).https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness). Accessed 11 October 2019.

Xiao-Tong, W., Jian-Jie, S., Xi-Zhe, L., Wei, L., Jian-Long, X., Dong-An, L., Yun-Zeng, Z., Lei, W., 2020. *Tepidiphilus olei* sp. nov., isolated from the production water of a water-flooded oil reservoir in PR China. Int J Syst Evol Microbiol. 70,4364-4371.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzéby, J., Amann, R., Rosselló-Móra, R., 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol. 12,635-45.

Zakrzewski, M., Proietti, C., Ellis, J.J., Hasan, S., Brion, M.-J., Berger, B., Krause, L., 2016. Calypso: a user-friendly web-server for mining and visualizing microbiome–environment interactions. Bioinformatics 33, 782–783.

Zouache, K., Raharimalala, F.N., Raquin, V., Tran-Van, V., Raveloson, L.H., Ravelonandro, P., Mavingui, P., 2011. Bacterial diversity of fieldcaught mosquitoes, Aedes albopictus and Aedes aegypti, from different geographic regions of Madagascar. FEMS Microbiol. Ecol. 75, 377–389.

Figures legends

Figure 1. Map showing the position of five villages of the Campo sleeping disease focus where tsetse flies were sampled.

Figure 2. Rarefaction curves performed for all sequenced samples (16S ribosomal subunit, V4 region). Coloured symbols designed the different individual samples.

Figure 3. Comparison of tsetse flies microbiome diversities. A, C: flies infection status; B, D: flies origin ; E, F, G : the trypanosomes species harbored by tsetse flies. TCF: *Trypanosoma congolense* 'Forest' type, TCS: *Trypanosoma congolense* 'Savanah' type, TBRSL: *Trypanosoma brucei* sensu-lato.

Figure 4. Distribution of the tsetse flies samples according to their microbiota composition based on principal coordinates analysis (PCoA) using Bray-Curtis index according to either (A) infection status, (B) *Trypanosoma* species and (C) the sampling sites. TCF: *Trypanosoma congolense* 'Forest' type, TCS: *Trypanosoma congolense* 'Savanah' type, TBRSL: *Trypanosoma brucei* sensu lato, TRBPA *Trypanosoma brucei gambiense*.

Figure 5. Differences in taxa composition between infected and uninfected tsetse flies based on Canonical Correspondence analysis (CCA) using Bray Curtis diversity index based on infection status across the 160 selected tsetse flies.

Figure 6. Relationship between different samples (represented by their identifier) according to the bacterial genera that they host using hierarchical clustering based on Bray-Curtis index values. The boxes above the sample identifiers indicate the flies' infection status.

Tables

Table 1. Primers used for PCR amplification of trypanosomes' DNA (Farikou et al. 2010)

Trypanosome species	**Primer sequences (5'-3')**	**PCR product size (base pairs)**	**References**
T. congolense forest type	GGACACGCCAGAAGGTACTT GTTCTCGCACCAAATCCAAC	350 bp	Masiga et al. (1992)
T. congolense savannah type	TCGAGCGAGAACGGGCACTTTG CGA ATTAGGGACAAACAAATCCCGC ACA	341 bp	Moser et al. (1989)
T. brucei s.l.	CGAATGAATATTAAACAATGCG CAG AGAACCATTTATTAGCTTTGTTG C	164 bp	Masiga et al. (1992)
T. b. gambiense group 1 (TRBPA)	GCGCCGACGATACCAATGC AACGGATTTCAGCGTTGCAG	239 bp	Herder et al. (2002)

Table 2. Global distribution of simple and mixed tsetse fly infections in the different villages of the Campo focus

village	Nber of flies analyzed	Nber of infected flies (%)	TcS (%)	Tbsl (%)	TcF (%)	Tc (F & S) (%)	Tbsl & TcS (%)	Tbsl & Tbg (%)	Tbsl, Tbg & TcS (%)	Tc(F & S) & Tbg & Tbsl (%)
Campo-beach	64	27 (42.1)	11	6	2	3	2	1	1	1
Ipono	21	8 (38)	0	5	2	0	0	1	0	0
Mvas	46	22 (47.8)	17	0	2	3	0	0	0	0
Itdonde-Fang	10	7 (70)	6	0	0	0	0	0	1	0
Mabiogo	19	16 (84.2)	16	0	0	0	0	0	0	0
Total	160	80 (50)	50 (31.25)	11 (6.8)	6 (3.75)	6 (3.75)	2 (1.25)	2 (1.25)	2 (1.25)	1 (0.62)

TcS, Tbsl and TcF refer to single infections. These species are also involved in mixed infection. Tbg was involved in mixed infections only.

Table 3. Comparative relative abundance, in trypanosome-infected and uninfected tsetse flies, of 11 mainly represented bacterial genera

Bacterial genera	Relative abundance (%)		p-value
	Non-infected flies (NI)	Infected flies (I)	(I vs NI)
Acinetobacter	1.65	0.24	0.051
Bacillus	0.25	0.19	0.008
Chryseobacterium	0.19	0.28	0.2
Cutibacterium	0.32	0.13	0.00012
Leuconostoc	0.2	0.21	0.48
Pseudomonas	0.98	0.29	0.0063
Spiroplasma	4.78	7.7	0.12
Tepidiphilus	1.54	1.22	0.025
Undibacterium	0.23	0.13	0.25
Wigglesworthia	86.45	87.96	0.032
Uncultured Bacterium	0.45	0.24	0.45

I: Infected; NI: Non-infected

Table 4. Bacterial genera which are present only in uninfected flies

Taxonomy	Abundance
Cupriavidus	0.2159
Legionella	0.0477
Novosphingobium	0.0412
Ralstonia	0.0354
Paenibacillus	0.0311
Luteolibacter	0.0249
Alloprevotella	0.0239
Lachnospiraceae	0.0232
Altererythrobacter	0.0210
Glutamicibacter	0.0203
Hydrogenophilus	0.0184
Gemella	0.0173
Chthoniobacter	0.0167
Comamonas	0.0160
Jeotgalicoccus	0.0156
Ruminococcus	0.0155
Turicella	0.0154
Fimbriiglobus	0.0149
Pseudarcicella	0.0146
Rikenellaceae	0.0144
Fodinicola	0.0139
Sporichthyaceae	0.0133

Macrococcus	0.0130
Streptomyces	0.0128
Alistipes	0.0121
Ruminiclostridium	0.0121
Myxosarcina	0.0119
Phascolarctobacterium	0.0118
Lysobacter	0.0113
Parabacteroides	0.0112
Hymenobacter	0.0111
Actinomyces	0.0109
Rickettsiella	0.0107
Blastopirellula	0.0103
Brevibacterium	0.0101
Schlegelella	0.0100


Atlantic ocean
Itdonde-Fang
Campo-Centre
Campo-beach
Mvas
Ipono
Rio Campo
NTEM
Mabiogo
Bongola
Akak
LEGEND
Sampling foci
HAT & AAT active foci
Road
0
2,5
5
Km

Figure 1

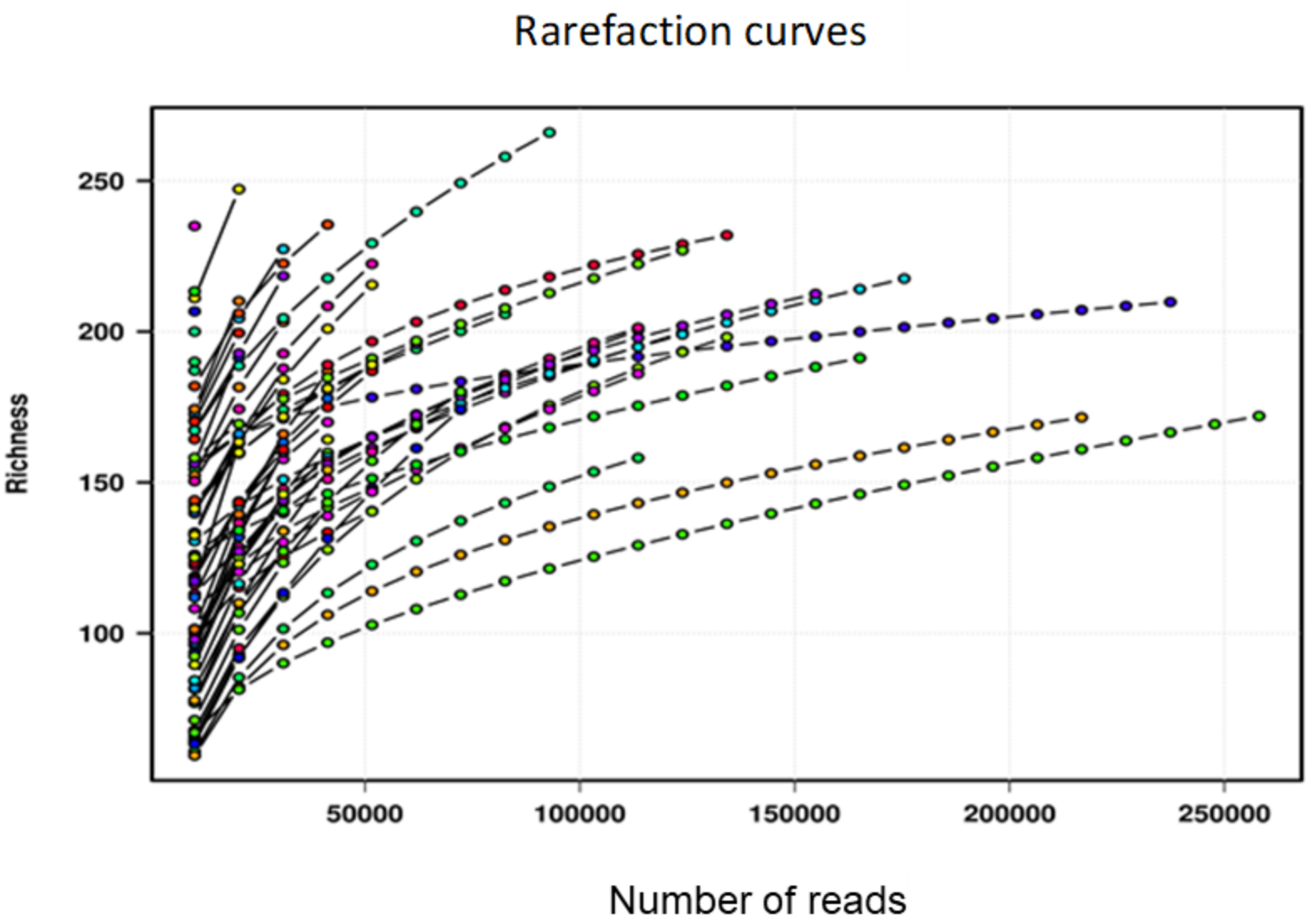
Rarefaction curves
Richness
250
200
150
100
50000
100000
150000
200000
250000
Number of reads

Figure 2

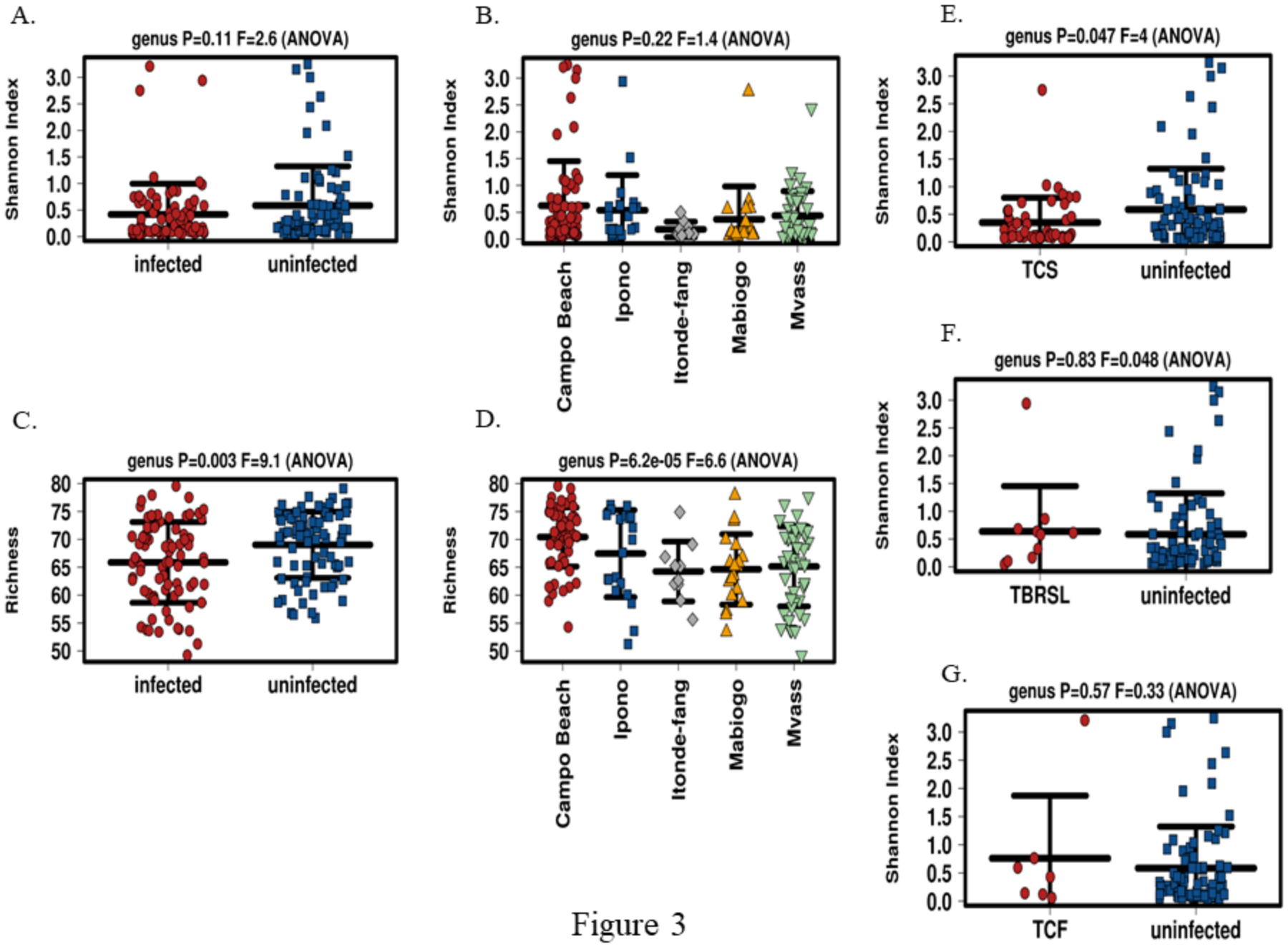
A.
genus P=0.11 F=2.6 (ANOVA)
Shannon Index
infected
uninfected
B.
genus P=0.22 F=1.4 (ANOVA)
Shannon Index
Campo Beach
Ipono
Itonde-fang
Mabiogo
Mvass
E.
genus P=0.047 F=4 (ANOVA)
Shannon Index
TCS
uninfected
C.
genus P=0.003 F=9.1 (ANOVA)
Richness
infected
uninfected
D.
genus P=6.2e-05 F=6.6 (ANOVA)
Richness
Campo Beach
Ipono
Itonde-fang
Mabiogo
Mvass
F.
genus P=0.83 F=0.048 (ANOVA)
Shannon Index
TBRSL
uninfected
G.
genus P=0.57 F=0.33 (ANOVA)
Shannon Index
TCF
uninfected

Figure 3

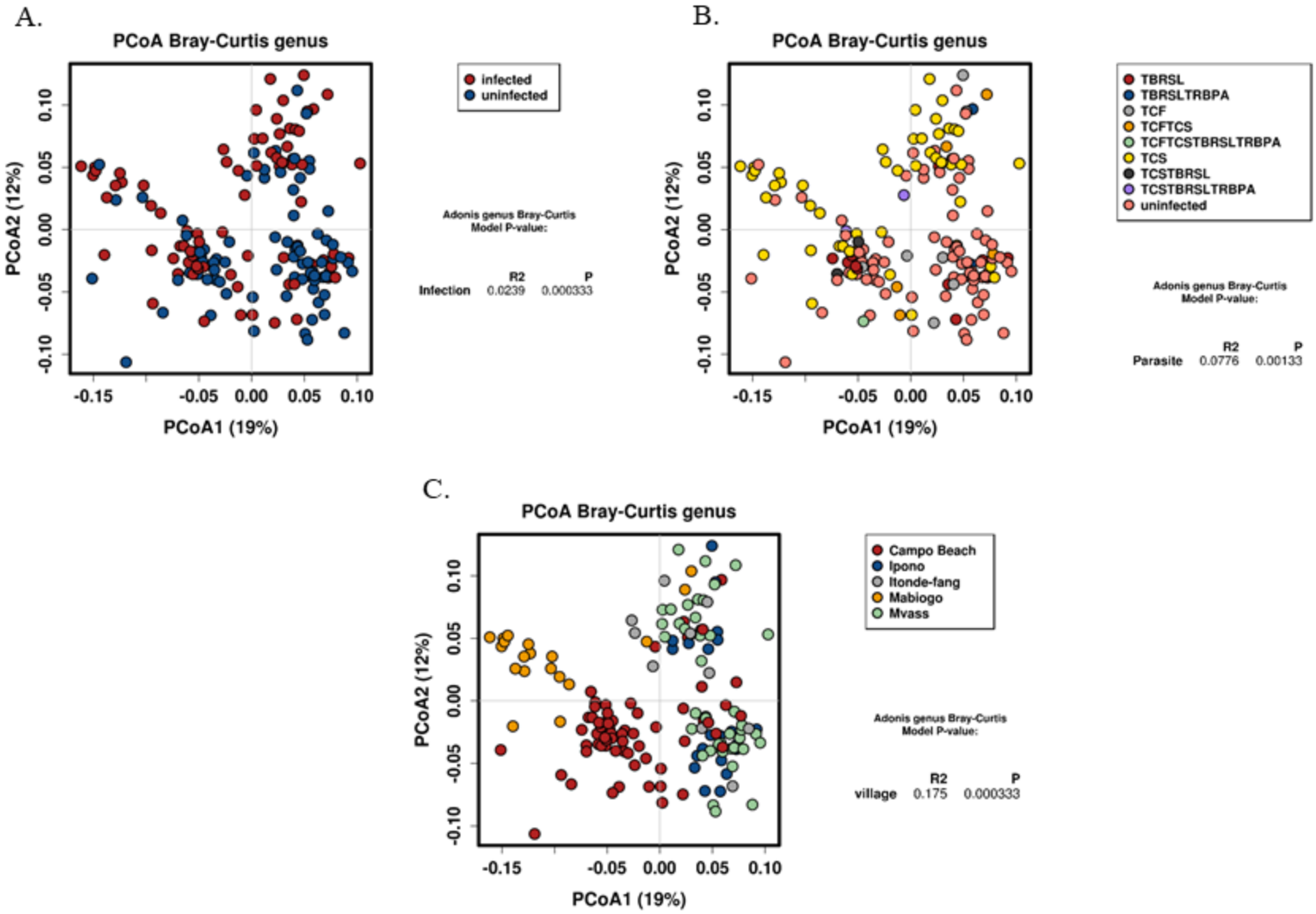
A.
PCoA Bray-Curtis genus
PCoA2 (12%)
PCoA1 (19%)
infected
uninfected
Adonis genus Bray-Curtis
Model P-value:
R2 P
Infection 0.0239 0.000333
B.
PCoA Bray-Curtis genus
PCoA2 (12%)
PCoA1 (19%)
TBRSL
TBRSLTRBPA
TCF
TCFTCS
TCFTCSTBRSLTRBPA
TCS
TCSTBRSL
TCSTBRSLTRBPA
uninfected
Adonis genus Bray-Curtis
Model P-value:
R2 P
Parasite 0.0776 0.00133
C.
PCoA Bray-Curtis genus
PCoA2 (12%)
PCoA1 (19%)
Campo Beach
Ipono
Itonde-fang
Mabiogo
Mvass
Adonis genus Bray-Curtis
Model P-value:
R2 P
village 0.175 0.000333

Figure 4

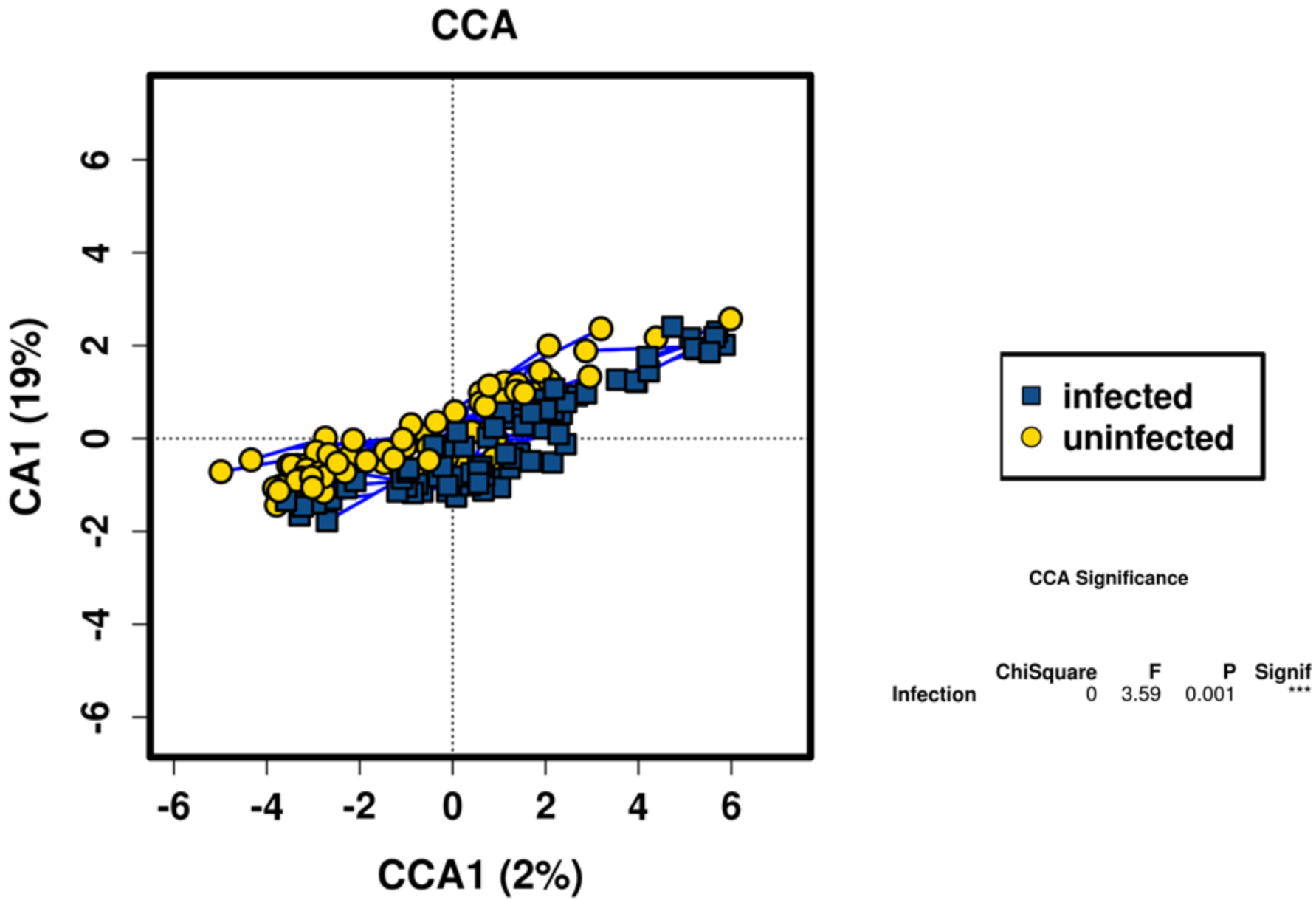
CCA
CA1 (19%)
CCA1 (2%)
-6
-4
-2
0
2
4
6
infected
uninfected
CCA Significance
ChiSquare F P Signif
Infection 0 3.59 0.001 ***

Figure 5

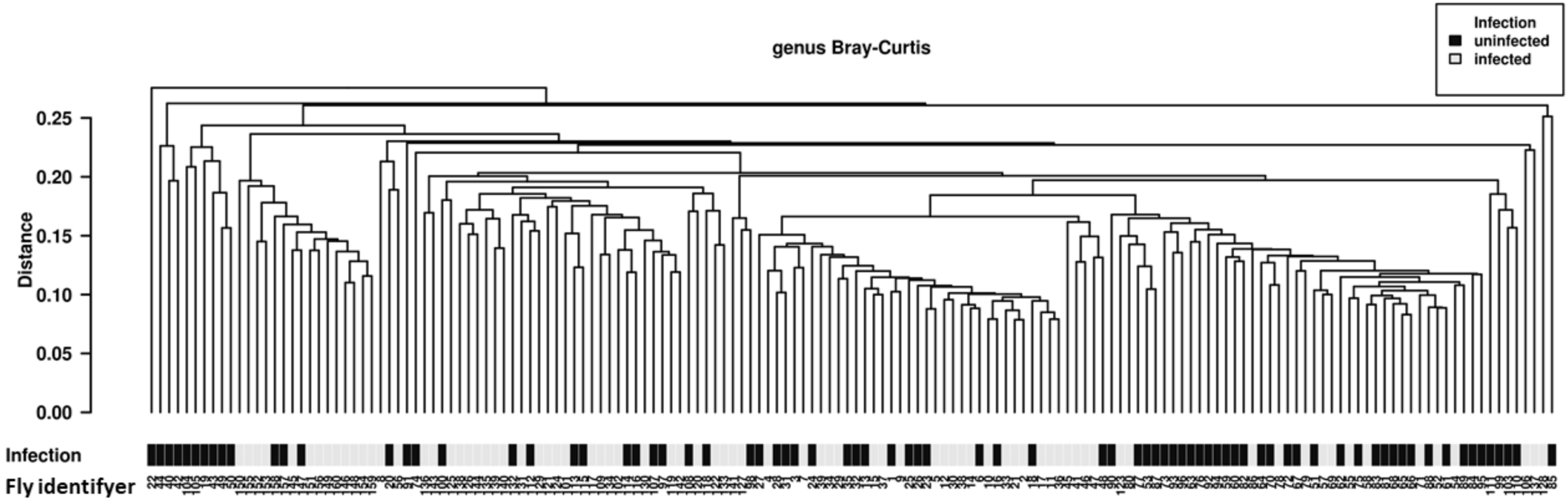
genus Bray-Curtis
Infection
uninfected
infected
Distance
0.25
0.20
0.15
0.10
0.05
0.00
Infection
Fly identifyer

Figure 6