



HAL
open science

iDNA screening: Disease vectors as vertebrate samplers

Arthur Kocher, Benoit de Thoisy, François Catzeflis, Sophie Valière,
Anne-Laure Bañuls, Jérôme Murienne

► To cite this version:

Arthur Kocher, Benoit de Thoisy, François Catzeflis, Sophie Valière, Anne-Laure Bañuls, et al.. iDNA screening: Disease vectors as vertebrate samplers. *Molecular Ecology*, 2017, 26 (22), pp.6478-6486. 10.1111/mec.14362 . hal-04917760

HAL Id: hal-04917760

<https://hal.science/hal-04917760v1>

Submitted on 28 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **iDNA screening: disease vectors as vertebrate samplers**

2 Arthur Kocher^{a,b,*}, Benoit De Thoisy^{c,d}, François Catzeflis^e, Sophie Valiere^f, Anne-Laure Bañuls^b and Jérôme
3 Murienne^a.

4 ^aCNRS, Université Toulouse III Paul Sabatier, ENFA; UMR5174 EDB (Laboratoire Evolution et Diversité
5 Biologique), Toulouse, France; ^bUMR MIVEGEC (IRD 224 – CNRS 5290 – Université de Montpellier), 911 Avenue
6 Agropolis, F34394 Montpellier, France; ^cInstitut Pasteur de la Guyane, 23 avenue Pasteur, 97300 Cayenne,
7 French Guiana; ^dAssociation Kwata, 16 avenue Pasteur, 97300 Cayenne, French Guiana; ^eInstitut des Sciences
8 de l'Evolution, Case Courrier 064, CNRS UMR-5554, Université Montpellier-2, Place E. Bataillon, F-34095
9 Montpellier, France; ^fGeT–PlaGe, Genotoul, INRA Auzeville, 31326 Castanet-Tolosan, France.

10 *Keywords:* sand fly; mosquito; blood meal; dipteran; insect; feeding preference;

11 *Running title:* Disease vectors as vertebrate samplers

12 **Corresponding author:* arthur.kocher@gmail.com

13

14

15 **Abstract**

16 In the current context of global change and human-induced biodiversity decline, there is an urgent need for
17 developing sampling approaches able to accurately describe the state of biodiversity. Traditional surveys of
18 vertebrate fauna involve time-consuming and skill-demanding field methods. Recently, the use of DNA derived
19 from invertebrate parasites (leeches and blowflies) was suggested as a new tool for vertebrate diversity
20 assessment. Blood meal analyses of arthropod disease vectors have long been performed to describe their
21 feeding behaviour, for epidemiological purposes. On the other hand, this existing expertise has not yet been
22 applied to investigate vertebrate fauna *per se*. Here, we evaluate the usefulness of hematophagous dipterans
23 as vertebrate samplers. Blood-fed sand flies and mosquitoes were collected in Amazonian forest sites, and
24 analysed using high-throughput sequencing of short mitochondrial markers. Blood meal identifications
25 highlighted contrasting ecological features and feeding behaviour among dipteran species, which allowed
26 unveiling arboreal and terrestrial mammals of various body size, as well as birds, lizards and amphibians.
27 Additionally, lower vertebrate diversity was found in sites undergoing higher levels of human-induced
28 perturbation. These results suggest that, in addition to providing precious information on disease vector host
29 use, dipteran blood meal analyses may represent a useful tool in the study of vertebrate communities.
30 Although further effort is required to validate the approach and consider its application to large scale studies,
31 this first work opens up promising perspectives for biodiversity monitoring and eco-epidemiology.

32

33 **Introduction**

34 The current rate of species disappearance is sometimes considered a sixth "mass extinction" crisis (Ceballos *et al.* 2015). Rapid modifications of biological communities, at both global and local scales, raise concerns as to
35 *al.* 2015). Rapid modifications of biological communities, at both global and local scales, raise concerns as to
36 the concomitant alteration of essential ecosystem functions and services (Cardinale *et al.* 2012). Developing an
37 indicator system to monitor the state and dynamics of biodiversity through time is a major challenge to reach
38 the 2020 Aichi Targets defined by Parties to the United Nations Convention on Biological Diversity. A set of
39 Essential Biodiversity Variables has been recently defined (Pereira *et al.* 2013), among which taxonomic
40 diversity of communities is relevant to several Aichi Targets. However, there are still uncertainties regarding
41 the economical and practical feasibility of developing monitoring systems at global and local scales.

42 Vertebrate community composition is frequently used as an indicator for the state of entire biota, because it is
43 relatively well known, occupies a wide variety of ecological niches, provides key ecosystem functions, and
44 contains species that are highly sensitive to human-induced disturbances (Ceballos & Ehrlich 2002; Collen *et al.*
45 2009; de Thoisy *et al.* 2010; Ceballos *et al.* 2015). Vertebrate surveys have been relying on time-consuming and
46 expertise-demanding observational methods such as line transects, track counts or live trapping (Smallwood &
47 Fitzhugh 1995; Voss & Emmons 1996; Vieira & Monteiro-Filho 2003; de Thoisy *et al.* 2008). Technological
48 progress has allowed the development of alternative approaches, including remote detection through camera
49 trapping (Rovero *et al.* 2014) or the sequencing of environmental DNA (Andersen *et al.* 2012; Taberlet *et al.*
50 2012; Valentini *et al.* 2016). Recently, the use of DNA derived from invertebrate parasites (iDNA) was suggested
51 as a new tool for vertebrate diversity assessment. Preliminary studies conducted on leeches (Schnell *et al.*
52 2012) and carrion flies (Calvignac-Spencer *et al.* 2013b; Lee *et al.* 2016) gave promising results but the method
53 has not yet been applied to large datasets.

54 Host identification from blood meals of disease-transmitting arthropods has long been performed for
55 estimating their feeding preference (Tesh *et al.* 1972; Tempelis 1975; Humair *et al.* 2007) because it is an
56 important epidemiological parameter for vector-borne diseases (Kent 2009; Hamer *et al.* 2011; Simpson *et al.*
57 2012). On the other hand, despite their worldwide distribution, the numerous methods available for their

58 collection and analyses, and the good quality of iDNA they provide (i.e. blood; Calvignac-Spencer *et al.* 2013a),
59 disease vectors have not been used to investigate the state of vertebrate diversity *per se*.
60 In this study, we first investigate the usefulness of dipteran blood meal analyses for vertebrate diversity
61 assessment. In particular, we evaluate the applicability of this approach for between-site comparisons. The
62 study was conducted in the Amazonian territory of French Guiana. This region is more than 80% covered by
63 primary lowland rainforests and as such harbours one of the largest continuous rainforest block that benefits
64 from a relatively favourable conservation status (Hammond 2005). It thus represents one of the few remaining
65 ‘natural laboratories’ where undisturbed biodiversity can be studied over large spatial scales. However,
66 demographic pressure and illegal gold mining activities constitute rising threats to biodiversity in part of the
67 territory (de Thoisy *et al.* 2005; Hammond *et al.* 2007; de Thoisy *et al.* 2010). Blood-fed sand flies (Psychodidae:
68 Phlebotominae) and mosquitoes (Culicidae) were collected in forest sites undergoing contrasting levels of
69 anthropogenic pressure. Dipteran and blood meal identifications were performed by PCR amplification and
70 high-throughput sequencing of short mitochondrial markers. The diversity of vertebrates observed through
71 blood meals was then compared between sites after rarefying samples down to equal size and dipteran species
72 composition.

73 **Materials and methods**

74 *Sampling*

75 Sampling was performed in the area of Saint-Georges de l'Oyapock in March 2016. Three study sites were
76 selected in order to constitute a gradient of anthropogenic pressure, as defined by the Human FootPrint index
77 (HFP index) which has been shown to be a good predictor of vertebrate diversity (de Thoisy *et al.* 2010, 2016).
78 In particular, hunting pressure was expected to increase with proximity to Saint-Georges town. In order to limit
79 environmental variability, the sites were chosen within the same geomorphological landscape harbouring a
80 highly-dissected plateau (Guitet *et al.* 2013) and the same forest type (Guitet *et al.* 2015). Site “SG1”, the most
81 impacted site, was located along the road leading to Brazil, close to the Oyapock river, at 1.2 km from Saint-
82 Georges and exhibited an HFP index of 32. Sites “SG6” and “SG30” were located along the national road 2

83 (constructed 15 years ago) at 6.8 km and 30.3 km from Saint-Georges and presented an HFP index of 26 and 21
84 respectively (Figure 1).

85 Nine Center for Disease Control (CDC) light-traps were set in each site, during four consecutive nights (from
86 6pm to 6am) for a total of 108 trap-nights. After each night, insects contained in the traps were killed by
87 freezing. Blood-fed mosquito and sand fly females were sorted and kept in individual microcentrifuge tubes
88 with 95% ethanol.

89 *DNA amplification and sequencing*

90 We extracted DNA from each engorged female using a without-boiling Chelex protocol (Casquet *et al.* 2012).

91 Two PCR were then performed: the first to amplify the Ins16S_1 marker (Ins16S_1-F:

92 TRRGACGAGAAGACCCTATA; Ins16S_1-R: TCTTAATCCAACATCGAGGTC; Clarke *et al.* 2014), which allows species-

93 level identifications of mosquitoes and sand flies (Kocher *et al.* 2017; Talaga *et al.*, 2017a); the second to

94 amplify the 12S-V5 marker (12S-V5-F: TAGAACAGGCTCCTCTAG; 12S-V5-R: TTAGATACCCCACTATGC; Riaz *et al.*

95 2011), that was shown to provide accurate identifications of mammals (Kocher *et al.* in press). The 12S-V5

96 marker was designed for metabarcoding of vertebrates. It thus presents several interesting properties for

97 blood-meal analyses: (i) wide taxonomic coverage (sand flies and mosquitoes may feed on a large variety of

98 vertebrates), (ii) good taxonomic resolution and (iii) very short size (*c.* 100 bp), which allows the detection of

99 highly degraded DNA (Taberlet *et al.* 2012).

100 The PCR protocol was the same for both fragments. Amplification was performed in 25 μ L mixtures containing

101 3 μ L of DNA template, 12 μ L of AmpliTaq Gold PCR Master Mix[®] (5U. μ L⁻¹; Applied Biosystems, Foster City, CA,

102 USA), 2.5 μ L of each primer (5 μ M), and nuclease-free water (Promega, Madison, WI, USA). The PCR mixture

103 was denatured at 95°C (10 min) and followed by 35 cycles of 30s at 95 °C, 30s at 50°C and 30s at 72 °C,

104 completed at 72 °C for 10 mins. Tags of eight base pairs with at least five differences between them were

105 added at the 5' end of each primer to enable the sequencing of the multiple PCR products in a single

106 sequencing run (Binladen *et al.* 2007).

107 PCR products were pooled and sent for library construction and sequencing to the GeT-PlaGe core facilities of
108 Genotoul (Toulouse, France). Samples were diluted in ultrapure water. A volume of 130 μ L containing 3 μ g of
109 DNA was purified using the HighPrep PCR system (Magbio Genomics, Gaithersburg, MD, USA) and used for
110 library construction with the Illumina NEXTflex PCR-Free DNA sequencing kit following the instructions of the
111 supplier (Bioo Scientific corp., Austin, TX, USA). Purified fragments were end-repaired, A-tailed and ligated to
112 sequencing indexed adapters. The quality of the library was controlled using the Fragment Analyzer (Advanced
113 Analytical, Ames, IA, USA) and quantified by qPCR with the Library Quantification Kit - Illumina Genome
114 Analyzer-SYBR Fast Universal (CliniSciences, Nanterre, France). The library was pooled with that of other
115 projects and loaded onto the Illumina MiSeq cartridge according to the manufacturer instructions. The quality
116 of the run was checked internally using PhiX. Quality filtering was performed by the Consensus Assessment of
117 Sequence and Variation (CASAVA) pipeline. The sequencing data was stored on the NG6 platform (Mariette *et*
118 *al.* 2012) and all computations were performed on the computer cluster of the Genotoul bioinformatic
119 platform (Toulouse, France).

120 *Bioinformatic treatment and taxonomic assignment*

121 The sequencing data was analysed using the `OBITOOLS` package (Boyer *et al.* 2016), as described previously
122 (Kocher *et al.* 2017). Paired-end reads were aligned, merged and then assigned to their corresponding sample
123 based on the tagged primer sequences with two mismatches allowed. Low quality reads were removed and
124 reads were then dereplicated. Taxonomic assignment of sequences was performed with `ecotag` (part of the
125 `OBITOOLS` package). We used reference libraries for Amazonian sand flies (Kocher *et al.* 2017), mammals
126 (Kocher *et al.* in press) and mosquitoes (Dataset S3) that we further complemented with sequences from
127 GenBank using `ecoPCR` (Riaz *et al.* 2011). When a sequence had less than 97% identity with its best match in
128 the reference database, we considered the taxonomic assignment at the generic rank or above. Since human
129 presence on the sampling sites was likely scarce and not observed during the sampling period, human
130 sequences identified in blood meals were regarded as laboratory contamination and discarded as a
131 conservative measure. In each sample, only the most abundant sequence was considered (*i.e.* we did not
132 consider potential mixed blood-meals; see *Results and discussion* for further explanations).

133 When dipterans or vertebrates could not be identified at the species level, we defined molecular taxonomic
134 units (MOTUs) within taxa. We used the Poisson Tree Process (Zhang *et al.* 2013) as implemented in `mPTP`
135 (Kapli *et al.* 2016). The method seeks to classify the branches of a phylogenetic tree into two processes: within
136 species (corresponding to a coalescence process) and between species (corresponding to a speciation process).
137 Because the method uses a phylogenetic tree, we first performed a phylogenetic analysis. Sequences were
138 aligned using `muscle` (Edgar 2004) and a Maximum Likelihood analysis was performed in `RAxML v. 8`
139 (Stamatakis 2014), with the GTR+ Gamma substitution model.

140 *Blood meals diversity analysis*

141 We had to consider two analytical issues when comparing the diversity of dipteran blood meals found between
142 sites. First, the number of blood-fed individuals differed among sites, leading to the well-known "sampling"
143 problem of diversity analysis, referring to the fact that the observed number of species is sensitive to sample
144 size (Gotelli & Colwell 2011). Secondly, each hematophagous species may exhibit distinct feeding preferences,
145 and their proportions in the samples varied between sites. Hence, the diversity of vertebrates observed
146 through blood meals is prone to site-specific bias, precluding the use of classical rarefaction analysis or the
147 direct comparison of diversity estimators.

148 We thus simulated rarefied blood meals pools by randomly drawing, without replacement, a number of
149 individuals from each dipteran species (or MOTU) corresponding to the lowest occurrence of the given species
150 among sites. Using this procedure, the obtained subsamples have identical dipteran composition and can be
151 compared. For each site, we simulated 10 000 rarefied communities and computed mean Hill numbers of
152 diversity (Chao *et al.* 2014) as well as the range of values containing the middle 95% of their distribution. Sand
153 flies of the same genus appeared to exhibit similar feeding preferences (*i.e.* *Nyssomyia* feeding mostly on
154 arboreal vertebrates and especially on xenarthrans, *Psychodopygus* feeding mainly on armadillos and
155 *Evandromyia* feeding on terrestrial rodents). Therefore, we grouped sand flies at the generic level in order to
156 increase the maximum base size of rarefied samples. These analyses can be reproduced using data and scripts
157 available in the Supporting Information.

158 **Results and discussion**

159 *Dipteran sampling and identifications*

160 In total, 209 blood-fed dipterans were collected, including 99 mosquitoes and 110 sand flies. 77, 76 and 56
161 specimens were collected in SG30, SG6 and SG1 respectively. The insect Ins16S_1 marker (Clarke *et al.* 2014)
162 was successfully amplified and sequenced for 94.9% of mosquitoes and 84.5% of sand flies, allowing to identify
163 187 specimens (average sequencing depth=6465x). Eight mosquitoes and eighteen sand fly species (or MOTUs)
164 were found (Table 1, Supporting Information).

165 *Blood meal analyses*

166 The vertebrate 12S-V5 marker (Riaz *et al.* 2011) was successfully amplified and sequenced for 94.9% of
167 mosquitoes and 66.4% of sand flies, allowing taxonomic assignment of 167 blood meals (average sequencing
168 depth=2166x). The lower success rate obtained with sand flies is likely explained by the smaller volume of their
169 blood meals, which is an important factor for successful iDNA amplification (Calvignac-Spencer *et al.* 2013a). A
170 total of 22 vertebrate MOTUs were found, of which a majority (nineteen) was identified at the species level
171 (Table 1, Supporting Information). These encompassed terrestrial and arboreal mammals of various size, as
172 well as birds, lizards and amphibians (Figure 2). Therefore, on the whole, mosquitoes and sand flies seemed to
173 be efficient vertebrate samplers, allowing the detection of a wide range of species.

174 The largely dominant mosquito MOTU (*Culex* MOTU 1) appeared to exhibit an opportunistic behaviour, since it
175 had fed on ten mammal species belonging to four different orders, as well as on two bird species (Table 1).

176 *Culex* MOTU 4 was found to feed on two lizards and a frog. All other mosquito species had fed on birds, but
177 were represented by few specimens. For some sand flies, our observations were consistent with existing
178 knowledge: the canopy-dwelling *Nyssomyia umbratilis* feeding mainly on arboreal mammals including a large
179 proportion of xenarthrans (Le Pont & Pajot 1980; Christensen *et al.* 1982) and the ground-dwelling
180 *Bichromomyia flaviscutellata* feeding on small rodents (Lainson & Shaw 1968; Tesh *et al.* 1971). In particular,
181 these known vectors of *Leishmania guyanensis* and *L. amazonensis*, respectively, were found to have fed on
182 the main known respective *Leishmania* reservoirs in the region (*Choloepus didactylus* and *Proechimys cuvieri*;
183 Rotureau 2006). For other species, this study provides the first blood meal data. Five *Psychodopygus* species

184 (*Ps. amazonensis*, *Ps. ayrozai*, *Ps. hirsuta*, *Ps. s. maripaensis* and *Ps.* MOTU 1) fed almost exclusively on
185 armadillos. The nine-banded armadillo (*Da. novemcinctus*) is the known reservoir of *L. naiffi* (Lainson & Shaw
186 1988), and *Ps. ayrozay* and *Ps. s. maripaensis* are suspected vectors of the same parasite (De Freitas *et al.* 1985;
187 Fouque *et al.* 2007). The association between sand flies and armadillos has been known for a long time, as it
188 can be attested by the Amerindian (*Tupi*) name for sand fly, "tatuquira" (literally armadillo-fly; Von Ihering
189 1968). We here report the first direct evidence of this vector-host interaction. The strong observed feeding
190 preference of *Psychodopygus* sand flies for armadillos may explain the relatively low number of human cases
191 caused by *L. naiffi*, as previously suspected, *Ps. ayrozai* being considered as a non-anthropophilic species
192 (Lainson & Shaw 2010). Finally, sand flies of the genus *Evandromyia*, for which no diet data was available, were
193 observed to feed mainly on terrestrial rodents. Of anecdotal but surprising mention is the detection of the
194 rodent *Isothrix sinnamariensis* (the brush-tailed rat, weighing ca. 200 g) in a sand fly blood meal. This rare
195 echimyid was previously know from only three localities in French Guiana, one locality in Guyana, and one in
196 Suriname (Lim & Catzeflis 2014).

197 Several studies have reported the detection of multiple blood meal sources in single hematophagous dipterans
198 (González *et al.* 2015; Sales *et al.* 2015), or the possibility to amplify vertebrate host DNA from specimens that
199 are not visibly blood-fed (Moreno *et al.* 2017). In our case, several vertebrate species were identified in 68% of
200 the specimens, with up to 6 and for a mean of 1.95 species. However, only considering the majority sequence
201 from visibly blood-fed specimens appeared to us as a necessary conservative measure. PCR-based methods are
202 extremely sensitive to laboratory contaminations, and the use of high-throughput sequencing leads to frequent
203 inter-sample "leakage" due to tag-switching events (Esling *et al.* 2015; Schnell *et al.* 2015). Of important note,
204 although considering the possibility for mixed blood meal sources would have nearly double the size of our
205 dataset, it would have led to the detection of only one additional vertebrate species (*Potos flavus*, without
206 considering obvious laboratory contaminants that may not be present in the sampling region). In our opinion,
207 this is strong evidence that most of vertebrate identifications we could have considered to originate from
208 mixed blood meals were rather laboratory or sequencing artefacts. We argue that results obtained from not
209 visibly blood-fed specimens, or the detection of multiple blood meal sources should be taken with caution

210 when using high-throughput sequencing of multiplexed PCR products. Furthermore, proper negative controls
211 (e.g. strictly phytophagous dipteran males collected in the same traps) should be used when reporting the
212 detection of frequent field and laboratory contaminants such as human DNA.

213 *Diversity comparisons*

214 The percentage of blood meal pairwise identity was higher within trap-nights than within sites, for both
215 mosquitoes (87.6% versus 32.4%; $\chi^2=347.68$; $p<10^{-15}$) and sand flies (53.5% versus 31.6%; $\chi^2=7.91$; $p<0.005$),
216 indicating that insects collected in the same trap and during the same night were likely to have fed on the same
217 individual or group of conspecifics. Therefore, blood meal data was converted into occurrences by trap-night
218 for each dipteran species (or genus for sand flies, see Material and Methods) before diversity analysis.

219 Decreasing biodiversity is a typical signature of human-induced disturbance (Turner 1996; Brashares *et al.*
220 2004) and was expected along our sampling sites. Taking raw blood meal data, we observed a marked decrease
221 in vertebrate species richness with increasing HFP and proximity to town (seventeen, eleven and seven species
222 in SG30, SG6 and SG1 respectively, Figure 3).

223 However, since different numbers of specimens were collected in each site, with varying proportions of
224 dipteran species characterized by specific feeding habits, direct comparisons of observed blood meal richness
225 are inevitably biased. Rarefaction is a commonly employed procedure to account for sampling size in diversity
226 analyses (Gotelli & Colwell 2011). Here, we rarefied our sample in order to estimate Hill numbers of diversity
227 (Chao *et al.* 2014) based on blood meals originating from identical pools of dipterans.

228 Rarefied samples included 13 randomly drawn specimens: four *Culex* MOTU 1, one *Evandromyia*, three
229 *Nyssomyia* and five *Psychodopygus*. Resulting estimated Hill numbers of order zero, one and two (*i.e.* species
230 richness, exponential of the Shannon-Wiener index and Simpson index) only slightly differed between SG30
231 and SG7, but still markedly decreased in SG1 (Figure 3). Of note, some species that are typically rare or absent
232 in highly disturbed habitat were found only in site SG30 or SG6 (e.g. *Alouatta maconnelli*, *Metachirus*
233 *nudicaudatus*, *Pecari tajacu*, *Tamandua tetradactyla*; *Tinamus major*; de Thoisy *et al.* 2008, 2010). This suggests
234 that dipteran blood meals may provide a useful tool for vertebrate diversity comparisons, even with moderate

235 sampling effort. However, the main limitation of this study is the lack of results simultaneously obtained with
236 classical methods and further work is needed to validate such an approach for vertebrate diversity assessment.
237 Of note, other studies are currently being conducted in the same area in the framework of the “Observatoire
238 Hommes-Milieux Oyapock” (<http://ohm-oyapock.in2p3.fr/>). Those involve classical camera-trapping, measures
239 of seed dispersal, bioacoustic monitoring and dung beetles sampling. This should bring complementary
240 information on vertebrate diversity along the same anthropogenic pressure gradient and allow interesting
241 comparisons with our results in a near future.

242 Dipteran species collected in this study exhibited a variety of ecological features and complementarity feeding
243 behaviour. Therefore, the analyses of the whole dipteran community, rather than specific species, allowed the
244 detection of a broad spectrum of vertebrate hosts. We argue that future studies should develop integrated
245 sampling protocols that are able to maximise the diversity of iDNA-carrying arthropods collected on study sites.
246 However, such a community-based approach also comes with drawbacks, since between-site comparisons are
247 prone to biases related to arthropod community compositions. In this study, we adopted a conservative
248 approach in this regard, but the development of specific analytical tools is another main challenge to be
249 addressed in future iDNA studies.

250 *Conclusions*

251 Although of preliminary nature, this work demonstrates that dipteran blood meal analyses can be used as a
252 single source of information to screen for the presence of a wide range of vertebrates. Our study also suggests
253 that the approach may be suitable for highlighting variation in vertebrate community composition between
254 sampling sites. Knowing the large expertise existing for the study of disease vectors, these results set the stage
255 for promising perspectives. Nevertheless, further studies involving comparisons of dipteran blood meal analysis
256 and classical vertebrate surveys are required to better evaluate the efficiency of the method.

257 Arthropods themselves are already frequently used as biodiversity indicators (Dedieu et al. 2015; Talaga et al.
258 2017b) and blood meal identifications could easily be complemented with insect community analyses in order
259 to extend the scope of such studies. Furthermore, our results provide precious contributions to the knowledge
260 of disease vector feeding habits. In recent years, questions have been raised about the impact of biodiversity

261 changes on the transmission risk of zoonotic infectious diseases (Ostfeld & Keesing 2012). In combination with
262 DNA-based pathogen detection, the present approach could open up new avenues for both biodiversity
263 assessment and eco-epidemiology.

264 **Acknowledgments**

265 This work was supported by “Investissement d’Avenir” grants managed by Agence Nationale de la Recherche
266 (LABEX CEBA: ANR-10-LABX-25-01; LABEX TULIP: ANR-10-LABX-41, ANR-11-IDEX-0002-02), Observatoire
267 Homme-Milieu Oyapock (USR 3456, within the frame of LABEX DRIIHM) as well as the Inkermann fund hosted
268 at the Fondation de France. We would like to thank Pierre-Michel Forget and François Feer for their precious
269 help in the design of this study; Antoine Fouquet (LEEISA CNRS), Romain Girod and Pascal Gaborit (Institut
270 Pasteur de la Guyane) for logistical support; Stanislas Talaga for providing us the mosquito reference database
271 and Fabian Fischer for his kind revision of the English language.

272 **References**

- 273 Andersen K, Bird KL, Rasmussen M *et al.* (2012) Meta-barcoding of “dirt” DNA from soil reflects vertebrate
274 biodiversity. *Molecular Ecology*, **21**, 1966–1979.
- 275 Binladen J, Gilbert MTP, Bollback JP *et al.* (2007) The use of coded PCR primers enables high-throughput
276 sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLOS ONE*, **2**, e197.
- 277 Boyer F, Mercier C, Bonin A *et al.* (2016) obitools: a unix-inspired software package for DNA metabarcoding.
278 *Molecular ecology resources*, **16**, 176–182.
- 279 Brashares JS, Arcese P, Sam MK *et al.* (2004) Bushmeat Hunting, Wildlife Declines, and Fish Supply in West
280 Africa. *Science*, **306**, 1180–1183.
- 281 Calvignac-Spencer S, Leendertz FH, Gilbert MTP, Schubert G (2013a) An invertebrate stomach’s view on
282 vertebrate ecology: Certain invertebrates could be used as “vertebrate samplers” and deliver DNA-
283 based information on many aspects of vertebrate ecology. *BioEssays*, **35**, 1004–1013.
- 284 Calvignac-Spencer S, Merkel K, Kutzner N *et al.* (2013b) Carrion fly-derived DNA as a tool for comprehensive
285 and cost-effective assessment of mammalian biodiversity. *Molecular Ecology*, **22**, 915–924.
- 286 Cardinale BJ, Duffy JE, Gonzalez A *et al.* (2012) Biodiversity loss and its impact on humanity. *Nature*, **486**, 59–
287 67.
- 288 Casquet J, Thebaud C, Gillespie RG (2012) Chelex without boiling, a rapid and easy technique to obtain stable
289 amplifiable DNA from small amounts of ethanol-stored spiders. *Molecular Ecology Resources*, **12**, 136–
290 141.
- 291 Ceballos G, Ehrlich PR (2002) Mammal Population Losses and the Extinction Crisis. *Science*, **296**, 904–907.
- 292 Ceballos G, Ehrlich PR, Barnosky AD *et al.* (2015) Accelerated modern human-induced species losses: Entering
293 the sixth mass extinction. *Science Advances*, **1**, e1400253.

- 294 Chao A, Gotelli NJ, Hsieh TC *et al.* (2014) Rarefaction and extrapolation with Hill numbers: a framework for
295 sampling and estimation in species diversity studies. *Ecological Monographs*, **84**, 45–67.
- 296 Christensen HA, Arias, JR A, de Vasquez, AM de V, de Freitas, RA de F (1982) Hosts of sandfly vectors of
297 *Leishmania braziliensis guyanensis* in the central Amazon of Brazil. *The American journal of tropical*
298 *medicine and hygiene*, **31**, 239–242.
- 299 Clarke LJ, Soubrier J, Weyrich LS, Cooper A (2014) Environmental metabarcodes for insects: *in silico* PCR reveals
300 potential for taxonomic bias. *Molecular Ecology Resources*, **14**, 1160–1170.
- 301 Collen B, Loh J, Whitmee S *et al.* (2009) Monitoring Change in Vertebrate Abundance: the Living Planet Index.
302 *Conservation Biology*, **23**, 317–327.
- 303 De Freitas R, Biancardi C, Castellon E (1985) Flagellate infections of brazilian sand flies (diptera: psychodidae):
304 isolation in vitro and biochemical identification of *Endotrypanum* and *Leishmania*. *Am. J. Trop. Med.*
305 *Hyg*, **34**, 1098–1108.
- 306 Dedieu N, Vigouroux R, Cerdan P, Céréghino R (2015) Invertebrate communities delineate hydro-ecoregions
307 and respond to anthropogenic disturbance in East-Amazonian streams. *Hydrobiologia*, **742**, 95–105.
- 308 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids*
309 *research*, **32**, 1792–1797.
- 310 Esling P, Lejzerowicz F, Pawlowski J (2015) Accurate multiplexing and filtering for high-throughput amplicon-
311 sequencing. *Nucleic Acids Research*, **43**, 2513–2524.
- 312 Fouque F, Gaborit P, Issaly J *et al.* (2007) Phlebotomine sand flies (Diptera: Psychodidae) associated with
313 changing patterns in the transmission of the human cutaneous leishmaniasis in French Guiana.
314 *Memórias do Instituto Oswaldo Cruz*, **102**, 35–40.
- 315 González E, Gállego M, Molina R *et al.* (2015) Identification of blood meals in field captured sand flies by a PCR-
316 RFLP approach based on cytochrome b gene. *Acta Tropica*, **152**, 96–102.
- 317 Gotelli NJ, Colwell RK (2011) Estimating species richness. *Biological diversity: frontiers in measurement and*
318 *assessment*, **12**, 39–54.
- 319 Guitet S, Brunaux O, de Granville J, Richard-Hansen C (2015) *Catalogue des habitats forestiers de Guyane*. ONF
320 and DEAL, Cayenne.
- 321 Guitet S, Cornu J-F, Brunaux O *et al.* (2013) Landform and landscape mapping, French Guiana (South America).
322 *Journal of Maps*, **9**, 325–335.
- 323 Hamer GL, Chaves LF, Anderson TK *et al.* (2011) Fine-scale variation in vector host use and force of infection
324 drive localized patterns of west nile virus transmission (RE Paul, Ed.). *PLOS ONE*, **6**, e23767.
- 325 Hammond DS (2005) *Tropical Forests of the Guiana Shield: Ancient Forests in a Modern World*. CABI.
- 326 Hammond DS, Gond V, de Thoisy B, Forget P-M, DeDijn BPE (2007) Causes and Consequences of a Tropical
327 Forest Gold Rush in the Guiana Shield, South America. *AMBIO: A Journal of the Human Environment*,
328 **36**, 661–670.
- 329 Humair P-F, Douet V, Cadenas FM *et al.* (2007) Molecular identification of bloodmeal source in ixodes ricinus
330 ticks using 12S rDNA as a genetic marker. *Journal of Medical Entomology*, **44**, 869–880.
- 331 Kapli P, Lutteropp S, Zhang J *et al.* (2016) Multi-rate Poisson tree processes for single-locus species delimitation
332 under maximum likelihood and Markov chain Monte Carlo. *bioRxiv*.
- 333 Kent RJ (2009) Molecular methods for arthropod bloodmeal identification and applications to ecological and
334 vector-borne disease studies. *Molecular Ecology Resources*, **9**, 4–18.
- 335 Kocher A, Gantier J-C, Gaborit P *et al.* (2017) Vector soup: high-throughput identification of Neotropical
336 phlebotomine sand flies using metabarcoding. *Molecular Ecology Resources*, **17**, 172–182.

- 337 Kocher A, Thoisy BD, Catzefflis F *et al.* (in press) Evaluation of short mitochondrial metabarcodes for the
338 identification of Amazonian mammals. *Methods in Ecology and Evolution*.
- 339 Lainson R, Shaw JJ (1968) Leishmaniasis in Brazil: I. Observations on enzootic rodent leishmaniasis—
340 incrimination of *Lutzomyia flaviscutellata* as the vector in the lower amazonian basin. *Transactions of*
341 *the Royal Society of Tropical Medicine and Hygiene*, **62**, 396–405.
- 342 Lainson R, Shaw J (1988) *Leishmania (Viannia) naiffi* sp. n., a parasite of the armadillo, *Dasypus novemcinctus*
343 (L.) in Amazonian Brazil. *Annales de parasitologie humaine et comparee*, **64**, 3–9.
- 344 Lainson R, Shaw JJ (2010) New world leishmaniasis. *Topley and Wilson's microbiology and microbial infections*.
- 345 Le Pont F, Pajot F-X (1980) Leishmaniose en Guyane Francaise. 1. Etude de l'écologie et du taux d'infection
346 naturelle du vecteur *Lutzomyia (Nyssomyia) umbratilis* Ward et Fraiha, 1977 en saison sèche.
347 Considérations épidémiologiques. *Cahiers ORSTOM: entomologie médicale et parasitologie*, **18**, 359–
348 382.
- 349 Lee P-S, Gan HM, Clements GR, Wilson J-J (2016) Field calibration of blowfly-derived DNA against traditional
350 methods for assessing mammal diversity in tropical forests. *Genome*, **59**, 1008–1022.
- 351 Lim BK, Catzefflis F (2014) History of mammalogy in the Guianas. In: *Historia de la mastozoologia en*
352 *Latinoamerica, las Guayanas y el Caribe* (eds Ortega J, Luis Martinez L, Tirira DG), pp. 265–292.
353 Murcielago Blanco and Asociacion Ecuatoriana de Mastozoologia, Quito and Mexico, DF.
- 354 Mariette J, Escudié F, Allias N *et al.* (2012) NG6: Integrated next generation sequencing storage and processing
355 environment. *BMC genomics*, **13**, 462.
- 356 Moreno M, Saavedra MP, Bickersmith SA *et al.* (2017) Intensive trapping of blood-fed *Anopheles darlingi* in
357 Amazonian Peru reveals unexpectedly high proportions of avian blood-meals. *PLOS Neglected Tropical*
358 *Diseases*, **11**, e0005337.
- 359 Ostfeld RS, Keesing F (2012) Effects of Host Diversity on Infectious Disease. *Annual Review of Ecology,*
360 *Evolution, and Systematics*, **43**, 157–182.
- 361 Pereira HM, Ferrier S, Walters M *et al.* (2013) Essential Biodiversity Variables. *Science*, **339**, 277–278.
- 362 Riaz T, Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers from whole
363 genome sequence analysis. *Nucleic Acids Research*, **39**, 1–11.
- 364 Rotureau B (2006) Ecology of the *Leishmania* species in the Guianan ecoregion complex. *The American journal*
365 *of tropical medicine and hygiene*, **74**, 81–96.
- 366 Rovero F, Martin E, Rosa M, Ahumada JA, Spitale D (2014) Estimating species richness and modelling habitat
367 preferences of tropical forest mammals from camera trap data. *PLOS ONE*, **9**, e103300.
- 368 Sales KG da S, Costa PL, de Morais RCS *et al.* (2015) Identification of phlebotomine sand fly blood meals by real-
369 time PCR. *Parasites & Vectors*, **8**.
- 370 Schnell IB, Bohmann K, Gilbert MTP (2015) Tag jumps illuminated - reducing sequence-to-sample
371 misidentifications in metabarcoding studies. *Molecular Ecology Resources*, **15**, 1289–1303.
- 372 Schnell IB, Thomsen PF, Wilkinson N *et al.* (2012) Screening mammal biodiversity using DNA from leeches.
373 *Current Biology*, **22**, R262–R263.
- 374 Simpson JE, Hurtado PJ, Medlock J *et al.* (2012) Vector host-feeding preferences drive transmission of multi-
375 host pathogens: West Nile virus as a model system. *Proceedings of the Royal Society of London B:*
376 *Biological Sciences*, **279**, 925–933.
- 377 Smallwood KS, Fitzhugh EL (1995) A track count for estimating mountain lion *Felis concolor californica*
378 population trend. *Biological Conservation*, **71**, 251–259.
- 379 Stamatakis A (2014) RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.
380 *Bioinformatics*, **30**, 1312–1313.

- 381 Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity
382 assessment using DNA metabarcoding. *Molecular ecology*, **21**, 2045–2050.
- 383 Talaga, S., Leroy, C., Guidez, A., Dusfour, I., Girod, R., Dejean, A., & Murienne, J. (2017a). DNA reference
384 libraries of French Guianese mosquitoes for barcoding and metabarcoding. *PLOS ONE*, **12**, e0176993.
- 385 Talaga S, Dézerald O, Carteron A *et al.* (2017b) Urbanization impacts the taxonomic and functional structure of
386 aquatic macroinvertebrate communities in a small Neotropical city. *Urban Ecosystems*, 1–9.
- 387 Tempelis CH (1975) Review article: host-feeding patterns of mosquitoes, with a review of advances in analysis
388 of blood meals by serology. *Journal of medical entomology*, **11**, 635–653.
- 389 Tesh RB, Chaniotis BN, Aronson MD, Johnson KM, others (1971) Natural host preferences of Panamanian
390 phlebotomine sandflies as determined by precipitin test. *American Journal of Tropical Medicine and*
391 *Hygiene*, **20**, 150–156.
- 392 Tesh RB, Chaniotis BN, Carrera BR, Johnson KM (1972) Further studies on the natural host preferences of
393 Panamanian phlebotomine sandflies. *American Journal of Epidemiology*, **95**, 88–93.
- 394 de Thoisy B, Brosse S, Dubois MA (2008) Assessment of large-vertebrate species richness and relative
395 abundance in Neotropical forest using line-transect censuses: what is the minimal effort required?
396 *Biodiversity and Conservation*, **17**, 2627–2644.
- 397 de Thoisy B, Fayad I, Clément L *et al.* (2016) Predators, Prey and Habitat Structure: Can Key Conservation Areas
398 and Early Signs of Population Collapse Be Detected in Neotropical Forests? *PLOS ONE*, **11**, e0165362.
- 399 de Thoisy B, Renoux F, Julliot C (2005) Hunting in northern French Guiana and its impact on primate
400 communities. *Oryx*, **39**, 149–157.
- 401 de Thoisy B, Richard-Hansen C, Goguillon B *et al.* (2010) Rapid evaluation of threats to biodiversity: human
402 footprint score and large vertebrate species responses in French Guiana. *Biodiversity and Conservation*,
403 **19**, 1567–1584.
- 404 Turner IM (1996) Species Loss in Fragments of Tropical Rain Forest: A Review of the Evidence. *Journal of*
405 *Applied Ecology*, **33**, 200–209.
- 406 Valentini A, Taberlet P, Miaud C *et al.* (2016) Next-generation monitoring of aquatic biodiversity using
407 environmental DNA metabarcoding. *Molecular Ecology*, **25**, 929–942.
- 408 Vieira EM, Monteiro-Filho ELA (2003) Vertical stratification of small mammals in the Atlantic rain forest of
409 south-eastern Brazil. *Journal of Tropical Ecology*, **19**, 501–507.
- 410 Von Ihering R (1968) *Dicionário dos animais do Brasil*. Edit. Univ. de Brasília, Brasília.
- 411 Voss RS, Emmons LH (1996) *Mammalian diversity in Neotropical lowland rainforests: a preliminary assessment*.
412 American Museum of Natural History, New York.
- 413 Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to
414 phylogenetic placements. *Bioinformatics*, **29**, 2869–2876.

415 **Data availability**

416 All data and scripts are available in the Supporting Information.

417 **Author Contributions**

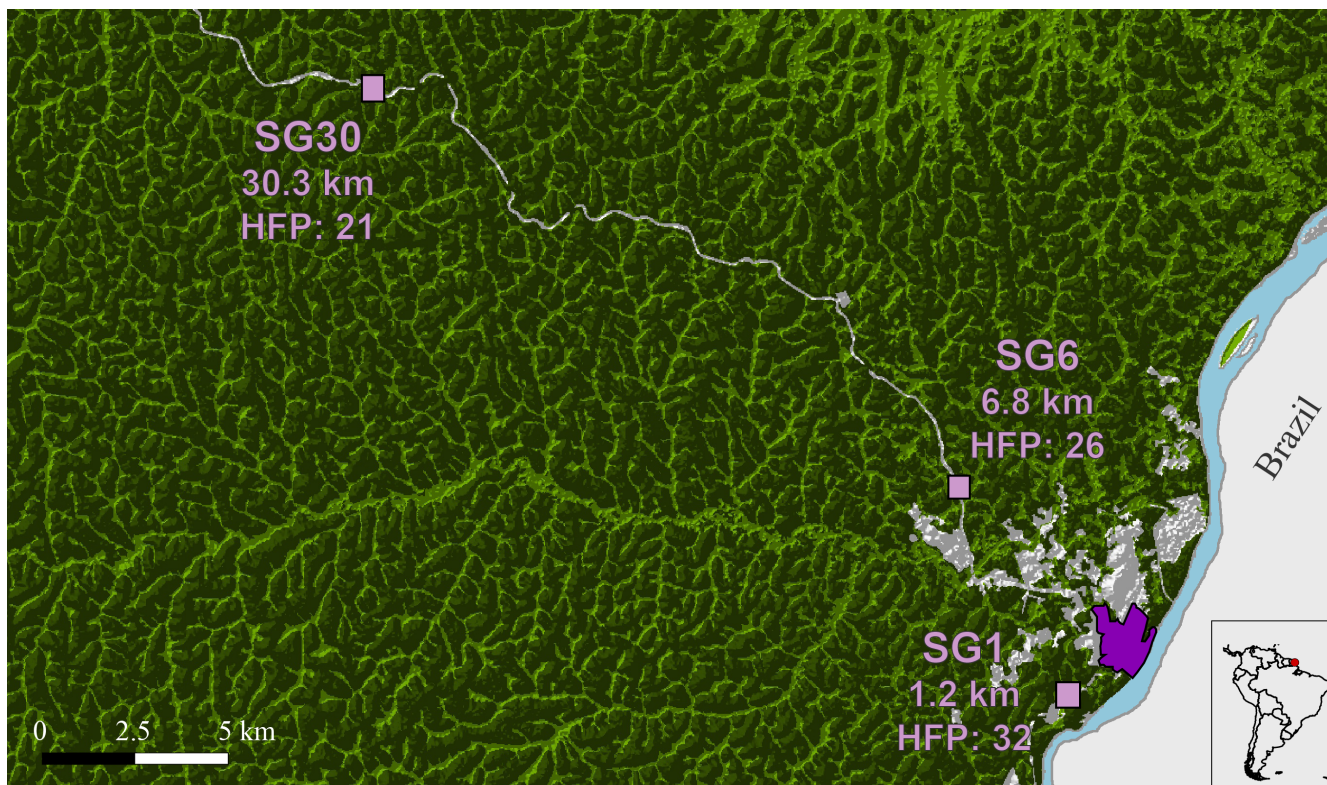
418 AK, A-L B, BT and JM designed research; AK, JM, FC and BT performed field work; AK and SV performed
419 laboratory work; AK analysed the data; AK wrote the paper and all authors contributed to its improvement.

420 **Tables**

421 Table 1: Identifications of blood meals in sand flies and mosquitoes collected the area of Saint-Georges de
 422 l'Oyapock in French Guiana. Only the 152 specimens for which both dipteran species and blood meals were
 423 successfully identified are included.

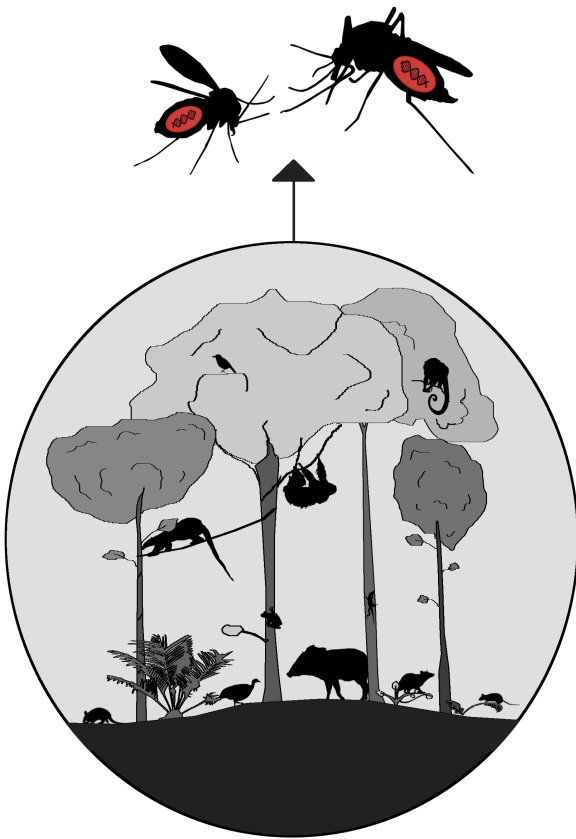
Group	Species	Nb	Blood meals
Phlebotominae	<i>Bichromomyia flaviscutellata</i>	4	<i>Echimys chrysurus</i> (1); <i>Proechimys cuvieri</i> (2); <i>Pr. MOTU 1</i> (1)
	<i>Evandromyia brachyphalla</i>	2	<i>Pr. cuvieri</i> (1); <i>Pr. guyannensis</i> (1)
	<i>Ev. infraspinoso</i>	1	<i>Dasyprocta leporina</i> (1)
	<i>Ev. sericea</i>	1	<i>Cuniculus paca</i> (1)
	<i>Ev. walkeri</i>	1	<i>Isothrix sinnamariensis</i> (1)
	<i>Nyssomyia</i> MOTU 1	3	<i>Choloepus didactylus</i> (2); <i>Tamandua tetradactyla</i> (1)
	<i>Ny. umbratilis</i>	12	<i>Alouatta macconelli</i> (1); <i>Ch. didactylus</i> (7); <i>Coendou melanurus</i> (2); <i>Pecari tajacu</i> (1); <i>Ta. tetradactyla</i> (1)
	<i>Pintomyia damascenoi</i>	1	<i>Ta. tetradactyla</i> (1)
	<i>Psychodopygus amazonensis</i>	3	<i>Dasytus kappleri</i> (1); <i>Du. novemcinctus</i> (2)
	<i>Ps. ayrozai</i>	17	<i>Du. kappleri</i> (2); <i>Du. novemcinctus</i> (14); <i>Pe. tajacu</i> (1)
	<i>Ps. claustrai</i>	1	<i>Dr. leporina</i> (1)
	<i>Ps. hirsuta</i>	5	<i>Du. novemcinctus</i> (5)
	<i>Ps. s. maripaensis</i>	8	<i>Du. novemcinctus</i> (7); <i>Pe. tajacu</i> (1)
	<i>Ps. MOTU 1</i>	5	<i>Du. novemcinctus</i> (5)
	<i>Sciopemyia sordellii</i>	1	<i>Pr. cuvieri</i> (1)
Culicidae	<i>Aedes serratus</i>	1	<i>Thamnophilus</i> MOTU 1 (1)
	<i>Culex mollis</i>	1	<i>Tinamus major</i> (1)
	<i>Culex</i> MOTU 1	78	<i>Ch. didactylus</i> (3); <i>Co. melanurus</i> (1); <i>Cu. paca</i> (6); <i>D. leporina</i> (2); <i>Du. novemcinctus</i> (1); <i>M. nudicaudatus</i> (3); <i>Pe. tajacu</i> (1); <i>Pr. cuvieri</i> (37); <i>Pr. guyannensis</i> (1); <i>Pr. MOTU 1</i> (4); <i>Th. nigrocinereus</i> (16); <i>Th. MOTU 1</i> (3)
	<i>Culex</i> MOTU 2	1	<i>Th. nigrocinereus</i> (1)
	<i>Culex</i> MOTU 3	2	<i>Th. nigrocinereus</i> (2)
	<i>Culex</i> MOTU 4	3	<i>Kentropyx calcarata</i> (1); <i>Osteocephalus</i> MOTU (1); <i>Polychrus marmoratus</i> (1)
	<i>Wyeomyia ypsipola</i>	1	<i>Ti. major</i> (1)

424



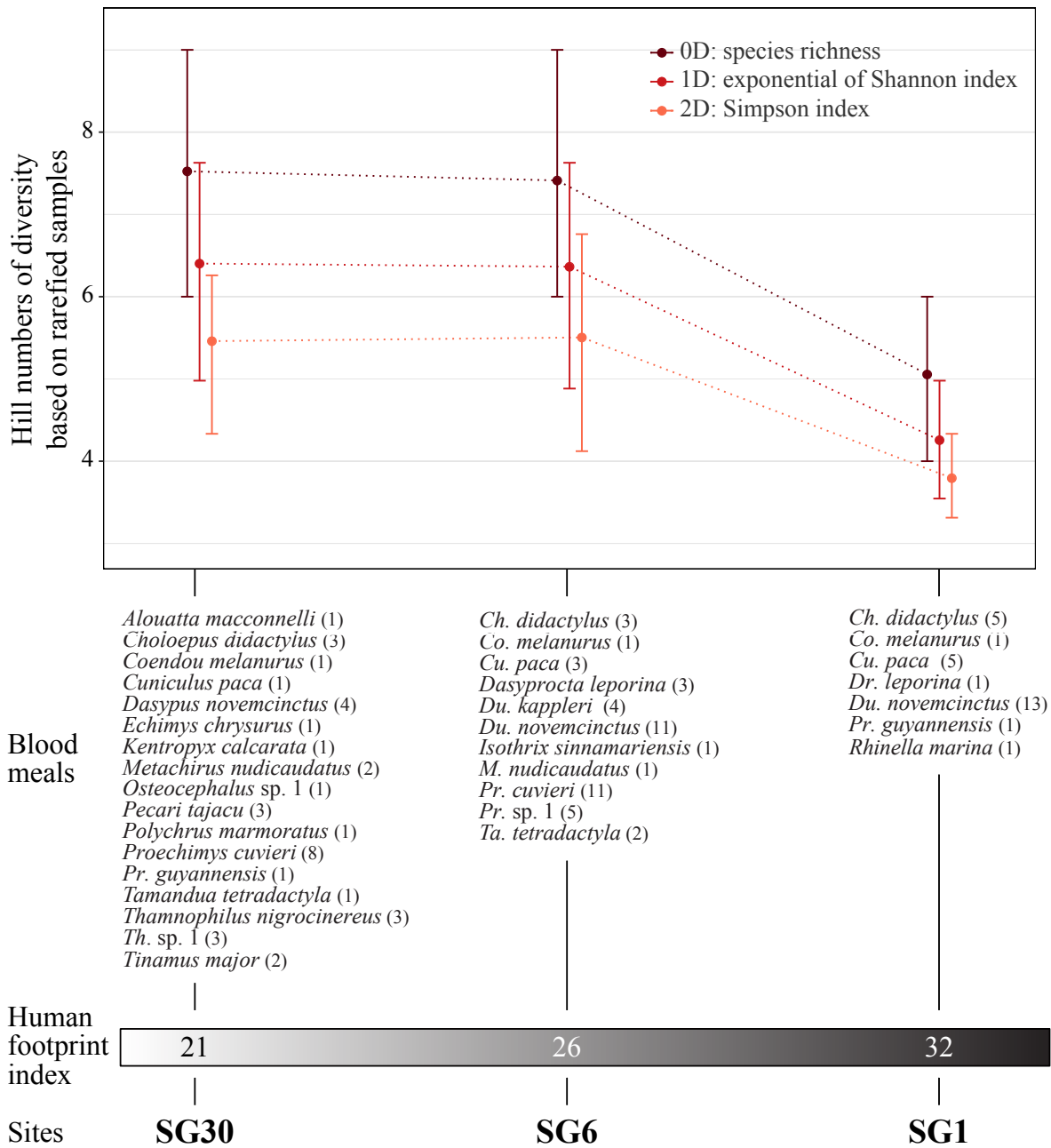
426

427 Figure 1: Geographical localisation of sampling sites in the area of Saint-Georges de l'Oyapock, French Guiana.
428 Human footprint index values (HFP, de Thoisy *et al.* 2010) and distance from town are indicated. Forested areas
429 are represented in green. Saint-Georges town is displayed in dark purple.



430

431 Figure 2: Schematization of the methodology employed in this study. Dipteran blood meal analyses allowed
432 identifying arboreal and terrestrial mammals of various body size as well as birds, lizards and amphibians. This
433 would typically require the implementation of multiple observational methods, wide taxonomic expertise and
434 considerable fieldwork



435

436

437

438

439

Figure 3: Vertebrates identified in dipteran blood meals (numbers in brackets indicate occurrences per trap-night) and estimations of mean Hill numbers of diversity (^qD) on rarefied blood meal pools among sites, along an anthropogenic pressure gradient defined by the human footprint index (de Thoisy *et al.* 2010). Error bars indicate the middle 95% of rarefied Hill number distributions.