MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2014) doi: 10.1111/1755-0998.12306

# Targeted detection of mammalian species using carrion fly-derived DNA

GRIT SCHUBERT,*[1] MELANIE STOCKHAUSEN,*†[1] CONSTANZE HOFFMANN,* KEVIN MERKEL,* LINDA VIGILANT,‡ FABIAN H. LEENDERTZ* and SÉBASTIEN CALVIGNAC-SPENCER*
*Epidemiology of highly pathogenic microorganisms, Robert Koch Institute, Nordufer 20, 13353 Berlin, Germany, †Faculty of Biology, University of Freiburg, Fahnenbergplatz, 79085 Freiburg, Germany, ‡Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany*

## Abstract

**DNA analysis from carrion flies (iDNA analysis) has recently been promoted as a powerful tool for cost- and time-efficient monitoring of wildlife. While originally applied to identify any mammalian species present in an area, it should also allow for targeted detection of species and individuals. Using carrion flies captured in the Taï National Park, Côte d'Ivoire, we assessed this possibility by (i) screening carrion fly DNA extracts with nonspecific and species-specific PCR systems, respectively, targeting mitochondrial DNA (mtDNA) fragments of any mammal or of Jentink's duiker (*Cephalophus jentinki*), three colobine monkeys (subfamily Colobinae) and sooty mangabey (*Cercocebus atys*); and (ii) genotyping carrion fly extracts containing sooty mangabey mtDNA. In comparison with the nonspecific PCR assay, the use of specific PCRs increased the frequency of detection of target species up to threefold. Detection rates partially reflected relative abundances of target species in the area. Amplification of seven microsatellite loci from carrion flies positive for sooty mangabey mtDNA yielded an average PCR success of 46%, showing that the identification of individuals is, to some extent, possible. Regression analysis of microsatellite PCR success and mtDNA concentration revealed that, among all carrion flies analysed for this study, 1% contained amounts of mammal mtDNA sufficient to attempt genotyping with potentially high success. We conclude that carrion fly-derived DNA analysis represents a promising tool for targeted monitoring of mammals in their natural habitat.**

*Keywords*: biodiversity, carrion fly, iDNA, microsatellites, quantitative PCR, targeted PCR

*Received 5 May 2014; revision received 11 July 2014; accepted 16 July 2014*

## Introduction

Global biodiversity is declining at a dramatic pace, and numerous mammal species are listed as endangered or critically endangered by the International Union for Conservation of Nature (IUCN, http://www.iucnredlist.org). Action plans aiming at conserving local faunas often target such threatened species, which first requires assessing a species' distribution and population density in the area of interest to later guide conservation actions (Nicholson *et al.* 2012).

Traditional monitoring methods, including transect studies, depend on the direct observation of animals or some of the traces they leave behind, for example, footprints or nests (Kouakou *et al.* 2009). Novel acoustic or visual device-based approaches circumvent the need for 'direct' data collection, while demanding relatively low logistical effort in the field and providing information on several species at the same time (Ahumada *et al.* 2011). Molecular methods using noninvasively collected samples of the desired species take wildlife monitoring a step further by also delivering in-depth information on many aspects of species population ecology (Schwartz *et al.* 2007). For instance, individual-based methods relying on typing of microsatellite markers allow estimation of population sizes through genetic capture–recapture studies, as well as inference of population genetic structure or population viability in primates (Arandjelovic *et al.* 2009; Schubert *et al.* 2011), carnivores (Solberg *et al.* 2006; Bellemain *et al.* 2007) and numerous other animal taxa. However, finding and sampling the ephemeral noninvasive samples appropriate for DNA analyses, such as faeces, is a task that demands much time and effort, especially for elusive animals, and may even be impossible for small mammals.

Correspondence: Grit Schubert and Sébastien Calvignac-Spencer, Fax: +49(0)30187542181; E-mails: schubertg@rki.de; calvignacs@rki.de

[1]These authors contributed equally to this manuscript.

Novel sources of DNA allowing for a broad screening of local mammalian biodiversity with reduced costs and effort have recently been identified. Specifically, great potential is seen in the use of DNA gathered from the environment (eDNA; Taberlet *et al.* 2012) or contained by invertebrates feeding on vertebrates (iDNA; Bohmann *et al.* 2013; Calvignac-Spencer *et al.* 2013a,b; Schnell *et al.* 2012). In one of the latter proof of principle studies, Calvignac-Spencer and coworkers focused on blow and flesh flies (families Calliphoridae and Sarcophagidae; later collectively referred to as 'carrion flies'), which are globally distributed and are known to feed on various mammalian-borne materials (carcasses, open wounds and faeces; Norris 1965). Screening carrion flies collected from two tropical ecosystems with pan-mammal and pan-vertebrate PCR assays, they identified a large number of mammal species that use different forest strata (e.g. terrestrial porcupine, *Hystrix cristata*, and arboreal Diana monkey, *Cercopithecus diana*) and exhibit different body masses (e.g. pygmy hippopotamus, *Hexaprotodon liberiensis*, and swamp rat, *Malacomys* sp.). While this illustrates the broad range of taxa exploited by carrion flies, carrion fly iDNA approaches could also be applied to targeted species monitoring. Here, the use of species-specific PCR assays might allow for more accurate investigation of target species local abundance or biomass as compared to generic approaches, which might be biased by, for example, the preferential amplification of one species over another in iDNA extracts containing a mixture of mammalian DNA resulting from multiple meals of the fly (Calvignac-Spencer *et al.* 2013b).

Using iDNA extracted from carrion flies captured in the evergreen rain forest of Taï National Park, Côte d'Ivoire, we specifically addressed the questions of (i) the relative effectiveness of a mammal-specific and several species-specific PCR systems to detect five mammal species with different population densities; and (ii) the feasibility of individual mammal genotyping. Our results suggest that species-specific PCRs increase target species detection rates from carrion fly iDNA and that these rates reflect, to some extent, local species abundance. In addition, we also show that individual mammal genotyping might be possible from a very small proportion of carrion flies.

## Materials and methods

### *Field site and sample collection*

A total of 533 carrion flies were captured at the Taï National Park (TNP) in Côte d'Ivoire. Of these, 491 carrion flies were caught at 58 random locations ('forest' flies), using a mosquito net trap and a bowl containing either a piece of meat or a commercial fly bait (Fig. S1, supporting information), or using customized plastic bottles holding the same bait types (see Calvignac-Spencer *et al.* 2013b). For microsatellite analyses (see below), we also used 42 carrion flies caught under a mosquito net during the dissection of a sooty mangabey (*Cercocebus atys*, 'net' flies; Calvignac-Spencer *et al.* 2013b). Carrion flies were either directly placed in 2-mL cryotubes and snap-frozen in liquid nitrogen on the day of collection or first kept in 15-mL tubes containing ethanol before being transferred into fresh 15-mL tubes containing silica and kept at ambient temperature. At their arrival in the laboratory, samples were then stored at −80°C or 4°C, respectively. It should be noted that frozen flies allowed for more frequent detection of mammalian DNA (52% success rate) than dried flies (22% success rate; two-tailed two-proportion *Z*-test, *P*-value = 0).

### *DNA extraction*

Carrion flies were cut into small pieces using sterilized scissors and homogenized in FastPrep® lysing matrix tubes (MP Biomedicals, Illkirch, France) filled with 250 μL phosphate buffer saline, using a FastPrep® system (with three times agitation for five-seconds). Nucleic acids were extracted from 100 μL carrion fly mixture using the GeneMATRIX Stool DNA Purification Kit (Roboklon, Berlin, Germany) following manufacturer's instructions. The efficiency of this protocol in removing PCR inhibitors had already been assessed on 40 fly DNA extracts in a previous study (Calvignac-Spencer *et al.* 2013b). To confirm these results, 40 DNA extracts generated for this study were also submitted to the same quantitative inhibition test, which did not reveal measurable inhibition.

DNA was also extracted from tissue of the dissected sooty mangabey which served as a positive control for microsatellite analyses, using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) according to manufacturer's instructions. Total DNA concentrations of all extracts were measured with a Nanodrop device (Thermo Scientific, Waltham, MA, USA). It should be noted that total DNA concentrations were only determined so as to ensure PCR screening would be performed on approximately comparable amounts of fly material (see below); these concentrations do not correlate with the (most often negligible) amount of mammal genetic material carried by flies.

### *Comparison of nonspecific and species-specific PCR assays*

We used DNA extracted from a subset of 'forest' flies ($N = 116$) to compare the performance of a nonspecific

mammal PCR assay (NSA) with that of PCR systems targeting individual species or small sets of species (species-specific assays; SSA). Target species were chosen because they exhibit different population densities in TNP, have different average adult body mass and exploit different strata of the forest. The following five species were therefore targeted: (i) the Jentink's duiker (*Cephalophus jentinki*; low density, 70 kg, terrestrial; East 1998); (ii) three species of colobine monkeys (Western black-and-white colobus, *Colobus polykomos*; Western red colobus, *Piliocolobus badius*; olive colobus, *Procolobus verus*; moderate to high densities, 4–8 kg, arboreal; Refisch & Koné 2005; Smith *et al.* 2003); and (iii) the sooty mangabey (*Cercocebus atys*; moderate density, 6 kg, terrestrial; Refisch & Koné 2005).

The NSA targeted a approximately 130-bp fragment of mitochondrial 16S DNA (16S; Taylor 1996). This assay included blocking primers aimed at preventing the amplification of human (primer 16Smam_blkhum3; Boessenkool *et al.* 2012) and pig sequences (primer 16Smam_blkpig1; Calvignac-Spencer *et al.* 2013b; Vestheim & Jarman 2008), which are common contaminants whose amplification can mask the presence of DNA fragments from other species (Boessenkool *et al.* 2012). Three SSA were initially designed using alignments of publicly available sequences from the target(s) and related taxa. The Jentink's duiker assay was designed to amplify a 143-bp fragment of the mitochondrial cytochrome b region (CytB); the colobine assay was designed to amplify a 122-bp fragment of the mitochondrial 12S ribosomal RNA gene (12S) from any colobine species (in our case from any of the three above-mentioned species); and the sooty mangabey assay was designed to amplify a 79-bp fragment of the mitochondrial cytochrome oxidase subunit 2 gene (COII). A sooty mangabey assay amplifying a longer fragment (162 bp) was also designed at a later stage, so as to allow for a better comparison to the results obtained with the Jentink's duiker and colobine assays (fragment length can be expected to influence the likelihood of recovery). This assay was only performed on flies found positive with the first sooty mangabey assay. All four assays were first tested using a series of target and nontarget DNA extracts and shown to exhibit the expected level of specificity (only species occurring in TNP were included in this assessment). All PCRs were performed in a final volume of 25 $\mu$L, containing 0.2 $\mu$M of each primer, 200 $\mu$M dNTP (using dUTP instead of dTTP to reduce possible carry-over contamination with PCR products), 0.3 U AmpErase® uracil N-glycosylase (Invitrogen, Carlsbad, CA, USA), 4 mM $MgCl_2$, 1× PCR buffer and 1.25 U Platinum® Taq polymerase (Invitrogen), adding 200 ng of DNA extract or, for DNA extracts containing <40 ng/$\mu$L, 5 $\mu$L of the extract. For the NSA, 1 $\mu$M of each blocking primer was added. Primer sequences and cycling conditions are given in Table S1 (supporting information).

PCR products of the appropriate length upon visualization on an agarose gel were treated with ExoSAP-IT® (Affymetrix, Santa Clara, CA, USA) and Sanger sequenced in both directions. All chromatograms were evaluated with the software GENEIOUS PRO v5.4.3 (Kearse *et al.* 2012). For the first sooty mangabey assay, PCR products were too short for direct sequencing. Hence, a subset ($N$ = 17) were cloned using the Topo TA Cloning ® Kit (Invitrogen), colony PCR performed using primers M13_F and M13_ R (Kilger *et al.* 1997) and PCR products Sanger sequenced on both strands. As species identity was confirmed for 16/17 of PCR products, positivity of the remaining PCRs was evaluated from gel photographs only. Five PCR products obtained with the NSA exhibited mixed chromatograms, possibly indicating that DNA remnants of multiple meals were co-amplified. To be able to incorporate the corresponding flies into our comparison of NSA and SSA, several NSA PCR products were generated for each of these flies in end-point dilution conditions, that is, DNA extract dilutions were used to seed 10–20 NSA replicates per fly. When less than half of the replicates were found positive, they were all Sanger sequenced.

Taxa were assigned to sequences by comparing them to the NCBI nonredundant database and a local database of sequences from TNP mammals using BLAST (Altschul *et al.* 1990). The criteria used for assignment are detailed in Calvignac-Spencer *et al.* 2013b. In brief, most species identifications were based on identity levels of ≥98%. In instances where lower identity levels were observed, we identified species or families on a case-per-case basis combining BLAST information (e.g. second hit exhibiting a clear drop in e-value) and biological information (e.g. phylogenetic relationships or reported presence in TNP). A complete list of BLAST results and the corresponding assignments is available as Table S2 (supporting information). We calculated detection rates of the target species for all assays and for all species by dividing the number of carrion fly extracts producing DNA sequences targeted by the assay by the total number of carrion fly extracts analysed with this assay. Potential differences in the detection rate between NSA and SSA were examined using McNemar's test with exact $P$-values. All statistical analyses were performed in R (R Development Core Team 2012).

### *Feasibility of microsatellite analysis*

*Microsatellite analysis.* We attempted to genotype sooty mangabeys using 42 'net' flies that had all been shown to contain sooty mangabey mtDNA (this study

and Calvignac-Spencer *et al.* 2013b). As a positive control, we used a DNA extract generated from a tissue sample of the mangabey that had been dissected under the very same net. We genotyped DNA extracts at seven autosomal microsatellite loci, among which six had already been shown to be sufficiently polymorphic to discriminate individuals (D2s433, D2s1326, D2s1333, D4s243, D9s905, D18s536; Santiago *et al.* 2005), while the seventh locus (D4s1627) had been used to genotype other primate species (Arandjelovic *et al.* 2009). Average allele lengths were 196, 267, 284, 174, 287, 159 and 203 bp, respectively. The corresponding primer sequences are listed in Table S3 (supporting information). We used a two-step amplification method as detailed in Arandjelovic *et al.* 2009. All primer pairs were combined with template DNA in a first multiplex PCR reaction. Amplification was carried out in 20 μL reaction volumes consisting of 1× SuperTaq buffer (HT Biotechnology Ltd, Cambridge, UK), 1.75 mM MgCl2, 0.15 mM of each forward (unlabelled) and reverse primer, 0.125 mM of each dNTP, 0.016 mg bovine serum albumin (BSA), 0.1 μL SuperTaq (HT Biotechnology) premixed 2:1 with TaqStart Antibody (BD Biosciences, Franklin Lakes, NJ, USA), and 5 μL template DNA. For PCR thermocycling, initial denaturation was for 9 min at 94°C, 30 cycles of 20 s at 94°C, 30 s at 55°C and 30 s at 72°C, with a final extension of 4 min at 72°C. In subsequent singleplex PCRs, we then diluted the initial PCR products 1:100 for amplification of each individual locus using fluorescently labelled forward primers. Unlike in Arandjelovic *et al.* 2009, the reverse primers were not nested. The total reaction volume of 10 μL consisted of 1× SuperTaq buffer, 0.875 mM MgCl2, 0.25 mM of each forward and reverse primer, 0.125 mM of each dNTP, 0.016 mg BSA, 0.04 μL SuperTaq premixed 2:1 with TaqStart Antibody and 2.5 μL diluted first round product. PCR thermocycling conditions were as above, using annealing temperatures between 58 and 61°C for the individual primers (also reported in Table S3).

Six independent amplifications of each extract were attempted in a 96-well plate along with a total of ten negative controls (controls always revealed negative). Up to three different PCR products from the singleplex amplification step were combined and electrophoresed on an ABI PRISM 3100 Genetic Analyser, and alleles were sized relative to an internal size standard (ROX labelled HD400) using GENEMAPPER Software version 3.7 (Applied Biosystems Inc., Foster City, CA, USA). As we did not determine the amount of nuclear DNA in the extracts, at least five replicates were required to confirm homozygous genotypes with >99% confidence (Arandjelovic *et al.* 2009). We considered genotypes to be heterozygous after we observed each allele in at least two independent PCRs. Genotypes are given in Table S4 (supporting information). While D4s1627 did not work in the positive control, it produced one identical allele in all but two tested 'net' flies and was therefore still included in subsequent analyses.

For extracts that only produced alleles that matched the positive control (see below), we calculated PCR success for each extract by dividing the sum of successful PCRs across all loci by the total number of attempted PCRs for all loci. Overall, PCR success was calculated by averaging all extracts' PCR success. There was no correlation between success rate and fragment length.

*Quantitative PCR and correlation with PCR success from microsatellite analysis.* For all 'forest' and 'net' fly extracts ($N$ = 491 and 42, respectively), mtDNA copy numbers were determined using quantitative PCR (qPCR). For this quantitative assay, we employed the primers used for the NSA along with the human blocking primer (nonspecific mammal qPCR). PCRs were performed in 12.5 μL reaction volumes, consisting of 6.25 μL of GoTaq® qPCR Master Mix (Promega, Fitchburg, WI, USA), 0.2 μM of each amplification primer, 1 μM of blocking primer, 1 μL extracted DNA or 1 μL of a diluted African palm civet (*Nandinia binotata*) PCR product ($10^1$–$10^5$ molecules per μL) for standard reactions. All 'forest' flies found qPCR positive were also analysed with NSA and the corresponding PCR products Sanger sequenced, so as to unveil potential multiple meals (none but these already reported for the sample set used for NSA/SSA comparison was detected). For extracts which had been genotyped, we also performed a sooty mangabey-specific qPCR, using the same primers as for nonquantitative sooty mangabey DNA amplification. PCR conditions were as for the NSA, except that 7.5 μL of GoTaq® qPCR Master Mix were used and no blocking primer was added. Runs were performed in a Stratagene qPCR machine (Stratagene, La Jolla, CA, USA). Cycling conditions are given in Table S1. Each sample was tested in duplicate, and fluorescence signals were evaluated with the software MXPRO (http://www.genomics.agilent.com/).

We aimed at using nonspecific mammal qPCR content as a predictor of genotyping success for our entire sample of 'forest' flies. As genotyping success was only available from sooty mangabey-positive carrion flies, we first verified that the nonspecific mammal quantification assay appropriately reflected sooty mangabey mtDNA content (and was not massively influenced, for instance, by remnant molecules from a previous meal), using a Spearman correlation. We next performed logistic regression analysis on extracts' PCR success (including again only extracts that produced alleles identical to the positive control) and mtDNA copy number from nonspecific mammal qPCR. We set the PCR success threshold at 60% for potentially good candidate extracts for genotyping

with the rationale that, when some additional effort for filling gaps in the genotypes is taken, genotypes would be sufficiently complete to allow for subsequent analyses (i.e. identity analysis). Finally, we applied this threshold to our entire sample of 'forest' flies.

### *Fly species identification*

Most of the 'forest' flies used for the comparison of NSA and SSA were identified using DNA barcoding ($n$ = 104). For this, PCRs were prepared as above-mentioned with a primer pair widely used to target a ca. 700 bp COI fragment in invertebrates (Folmer *et al.* 1994). All reactions were seeded with 50 ng DNA. Primer sequences and cycling conditions are given in Table S1. PCR products were visualized and Sanger sequenced. To assign sequences to fly species, we first performed phylogenetic analyses using BEAST v1.8.0 under a relaxed clock model and modelling tree shape with a coalescent model (constant population size; Drummond *et al.* 2012). The model of nucleotide substitution used in these analyses had first been identified through comparisons of many model maximum-likelihood estimates, using JMODELTEST v2.1.3 (Darriba *et al.* 2012). Two Markov chain Monte Carlo analyses were run with BEAST, convergence of the chains and appropriate sampling of the posterior were checked for in TRACER v1.5 (available at http://tree.bio.ed.ac.uk/software/tracer/), and the corresponding tree files were merged with LOGCOMBINER v1.8.0 (excluding their respective burn-in; software distributed within the BEAST package) before being summarized onto the maximum clade credibility tree (MCC tree) using TREEANNOTATOR v1.8.0 (also distributed with BEAST). The MCC tree was used as input for sequence-based species delineation using the general mixed Yule coalescent method (GMYC; Fujisawa & Barraclough 2013; Pons *et al.* 2006). The identified species were anchored to the lower taxon possible using a BLAST search and applying criteria similar to those mentioned above. Fly species assignment is reported in Table S2.

## Results

### *Assessing the performance of NSA and SSA in amplifying target mammal species' DNA*

A first set of 116 'forest' flies was used to compare the performance of the assays. All carrion fly extracts were tested using the NSA, while 106, 99 and 108 were also analysed using the Jentink's duiker assay, the colobine assay and the sooty mangabey assay, respectively (fly material was not available to test all 119 'forest' flies of this set with any SSA). The NSA allowed for the identification of mammal DNA from 66 carrion fly extracts (including five extracts that originally yielded mixed chromatograms and were further examined by end-point dilution PCR). Sequences were assigned to 16 mammal taxa (among which 11 species were identified) and one species of frog (excluding sequences considered as contaminant, e.g. human sequences; Table S2). Twenty-seven of those sequences matched one of our five target species (Jentink's duiker: 1, Western black-and-white colobus: 2, Western red colobus: 13, olive colobus: 3, sooty mangabey: 8; Table 1).

Using the specific Jentink's duiker assay, this species' DNA was only detected from the unique carrion fly that had been found positive using the NSA. In addition, one carrion fly contained DNA of another closely related duiker species of the same genus, the black duiker (*Cephalophus niger*), which we could not include in our initial tests of specificity. In the NSA, mammal identity for this fly was only determined on the subfamily level (Cephalophinae). With the colobine specific assay, the overall detection rate (i.e. the cumulative detection rate of all three species) increased significantly by a factor of 1.8 (McNemar's test, $P$-value = 0.018), resulting in 32% of all carrion fly extracts being positive for at least one of the species (Table 1). This increase was mainly driven by the increased detection of Western red colobus (factor 1.8; McNemar's test, $P$-value = 0.045), while for Western black-and-white colobus and olive colobus, no significant detection rate increase was observed, although the number of carrion flies positive for the first species was tripled. With the sooty mangabey-specific assay, detection rate increased significantly when considering both the short and long PCR assays (McNemar's tests, $P$-value < 0.0001 and $P$-value = 0.045, respectively), with the short PCR assay resulting in a higher detection rate increase (factor 3.3) than the long one (factor 1.9). These PCR assays, respectively, identified 25 and 13.9% of all tested DNA extracts as positive for this species.

More fly traps were found to contain carrion flies that produced target species' DNA sequences upon SSA, that is, the number of 'positive' traps increased in comparison with NSA results (Table 1). The absence of any amplification and the preferential amplification of nontarget species' DNA were approximately equal contributors to NSA failures in detecting target species' DNA (where SSA succeeded; Table 2).

### *Delineating carrion fly species and assessing their relative sampling performance*

General mixed yule coalescent identified five carrion fly species (confidence interval: 5–12). Of all flies, 92% could

**Table 1** Species detection rates with nonspecific mammal (NSA) and species-specific (SSA) PCR assays. [l]long sooty mangabey assay; [s]short sooty mangabey assay; *with the Jentink's duiker assay, one additional, nontarget species of the same subfamily (Cephalophinae) was identified; **assignment only on subfamily level. Other unambiguously assigned mammal species detected with the NSA from this set of flies are reported at the end of the table

| Taxon/species | # Carrion flies analysed | # Carrion flies with target species DNA (# fly traps containing those flies) | | Increase in detection rate |
|---|---|---|---|---|
| | | NSA | SSA | |
| Jentink's duiker | 106 | 1 (1) | 1 (1) | 0 |
| Colobinae | 99 | 18 (12) | 32 (21) | 1.8 |
| Black-and-white colobus | | 2 (2) | 6 (5) | 3 |
| Red colobus | | 13 (10) | 24 (12) | 1.8 |
| Olive colobus | | 3 (3) | 2 (2) | 0.7 |
| Sooty mangabey | 108 | 8 (6) | $15^{l}$ (7); $27^{s}$ (11) | $1.9^{l}$ ($3.4^{s}$) |
| Total | | 27 | $48^{l}$; $60^{s}$ | – |
| Black duiker * | | 1** (1) | 1 (1) | – |
| Campbell's monkey | | 4 (4) | – | – |
| Diana monkey | | 7 (6) | – | – |
| Greater spot-nosed monkey | | 1 (1) | – | – |
| Hammer-headed fruit bat | | 2 (2) | – | – |
| Lesser spot-nosed monkey | | 1 (1) | – | – |
| Little collared fruit bat | | 1 (1) | – | – |
| Maxwell's duiker | | 1 (1) | – | – |
| Pygmy hippopotamus | | 1 (1) | – | – |

**Table 2** Results of sequencing of nonspecific PCR products in cases where target species were identified in the species-specific assay but not the nonspecific mammal assay

| Taxon/species | No amplicon | Nontarget species |
|---|---|---|
| Colobinae | 10 | 12 |
| Sooty mangabey | 5 | 10 |
| Total | 15 | 22 |

**Table 3** Carrion fly species-specific detection rates with nonspecific mammal and species-specific assays. The last column reports the results of a two-tailed two-proportion *Z*-test

| Assay | *Chrysomya putoria* | *Chrysomya inclinata* | *Z*-score/ *P*-value |
|---|---|---|---|
| Nonspecific mammal | 24/44 (54%) | 22/37 (59%) | −0.44/0.66 |
| Colobine specific | 11/36 (31%) | 9/29 (31%) | −0.04/0.97 |
| Sooty mangabey specific | 5/38 (13%) | 11/35 (31%) | −1.88/0.06 |

be assigned to three species belonging to the family Calliphoridae and the rest to two species belonging to the family Sarcophagidae. Two Calliphoridae species clearly dominated this assemblage: 52% of all individuals were identified as *Chrysomya putoria* (52 individuals found in 16 traps) and 36% as *Chrysomya inclinata* (37 individuals in 11 traps). For the three other species, less than five individuals were sampled in a maximum of three traps.

We could examine the relative performance of the two most represented species with respect to NSA and SSA. NSA success rates were comparable (*C.putoria* 54%, *C. inclinata* 59%) as well as colobine SSA (31% for both species; Table 3). In contrast, sooty mangabey SSA detection rate was higher in *C. inclinata* (31%) than *C. putoria* (13%). All two-proportion z-tests failed to reject the null hypothesis of equal proportions of positives in both species at a 0.05 significance level, although for the sooty mangabey SSA, the *P*-value was close to this threshold (*P*-value = 0.059; Table 3).

### *Predicting mammal genotyping success from carrion fly DNA extracts*

We attempted to genotype 42 'net' flies, all of which had been shown to contain sooty mangabey mtDNA, at seven microsatellite loci (for genotypes, see Table S4). For 30 fly extracts, no mismatches were observed among the genotypes of 'net' flies and the mangabey being dissected

under the same mosquito net. The remaining 12 fly extracts produced genotypes with different alleles at some loci but potentially in the allelic range of sooty mangabeys at Taï National Park (i.e. not more than 40 bp apart from alleles of the positive control, Table S4) and therefore might originate from additional sooty mangabey individuals. We had, however, no means to verify this as only seven individuals from TNP have been genotyped up to now, resulting in a limited allelic range (Santiago *et al.* 2005). To be conservative, PCR success was thus only calculated from the 30 'matching' flies, averaging 45.6% (range 2.4–100%).

We then quantified mtDNA with a nonspecific mammal qPCR assay in DNA extracts from 'net' flies and 'forest' flies ($N$ = 479, including 104 flies used for the comparison of NSA and SSA). For 'net' flies, we first determined that mtDNA copy numbers estimated through the nonspecific mammal qPCR were correlated with those obtained with a sooty mangabey-specific qPCR ($P$-value < $1e^{-6}$, Spearman's rho = 0.81; Fig. 1). For the same DNA extracts, we then compared PCR success (for microsatellites) and mtDNA copy numbers determined from nonspecific mammal qPCR using a logistic regression analysis (estimate: 0.14, SD: 0.02, $z$-value = 7.14, $P$-value < $1e^{-12}$; Fig. 2). Using this regression, the threshold of a minimum PCR success of 60% was reached at ca. 700 copies/$\mu$L DNA extract. Five of 479 'forest' flies (1%) from which nonspecific mammal qPCR results were available were found to contain DNA quantities above this threshold, indicating their potential suitability for microsatellite analysis.

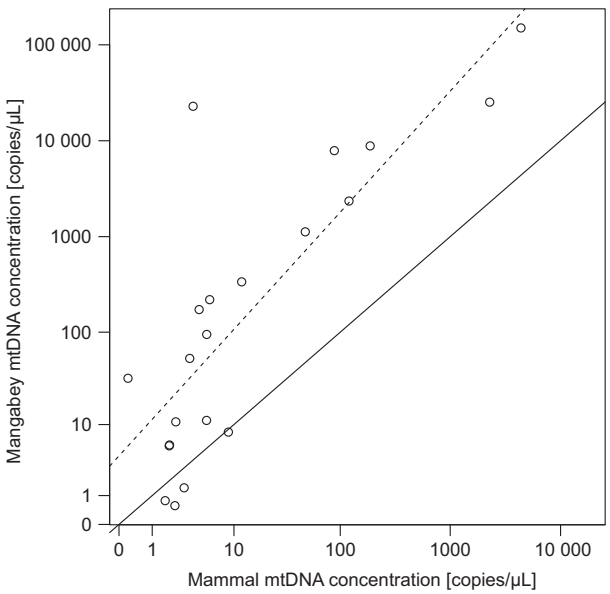


**Fig. 1** Spearman correlation between mammal mitochondrial DNA (mtDNA) concentration and mangabey mtDNA concentration. $P$-value < 0.001, Spearman's rho = 0.81, the dashed line depicts the result of a standard linear regression with both variables being log-transformed. The solid line represents represent a (theoretical) perfect agreement of the estimates of mtDNA counts with both methods. The fact that most data points lie above that line points at a relatively lower sensitivity of the nonspecific mammal assay.

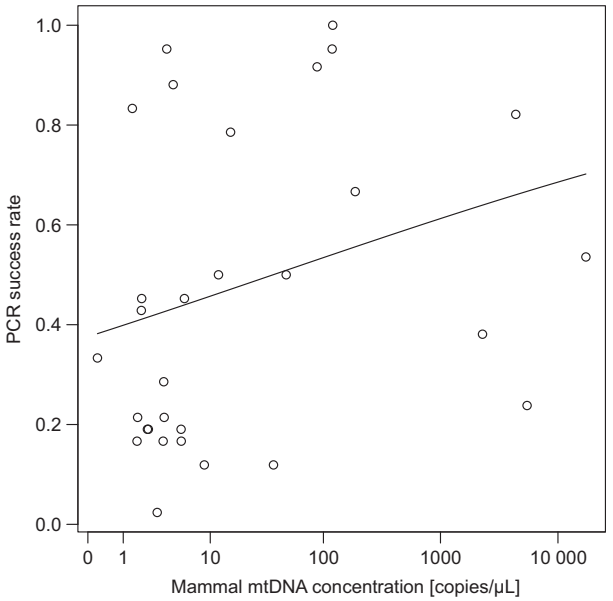


**Fig. 2** Logistic regression between mammal mitochondrial DNA (mtDNA) concentration and PCR success. LP (linear predictor) = −0.5071 + 0.1397*log($x$). $Z$-value = 7.14. $P$-value < 0.0001.

## Discussion

In a comparison of the performance of nonspecific mammal (NSA) and species-specific (SSA) PCR assays in detecting three target mammal taxa from carrion fly iDNA, we found SSA to be more successful. On our sample of carrion flies collected in the rain forest of the Taï National Park, Côte d'Ivoire, the detection rates of iDNA from two species of colobine monkeys and sooty mangabey measurably increased, reaching statistical significance for both Western red colobus and sooty mangabey. Of note, up to 25% of all tested flies were positive for Western red colobus (Table 1). Numerous studies have made use of haematophagous insects to examine transmission dynamics of parasites relevant to human and veterinary medicine (Kent 2009). While most of this work has adopted generic approaches so as to determine the spectrum of potential hosts of invertebrates (and thus of the parasites itself), some also applied specific assays comparable to ours. Thus, 24.3% of stable flies (*Stomoxys calcitrans*) collected in equine facilities had fed on horses (Pitzer *et al.* 2011) and between 29.5% and 97.1% of

engorged biting midges (Culicoides) collected near livestock farms contained DNA assignable to the host species of the farm (Garros *et al.* 2011). It is notable that, despite targeting undomesticated wildlife living at lower densities, our study yielded a comparable success rate for the most abundant target species (i.e. Western red colobus).

It is also interesting that distinct fly species exhibited statistically indistinguishable NSA and SSA detection rates (Table 3) and seem equally useful as iDNA samplers. This is well in line with the notion that many fly species are opportunistic generalists. However, the ecology of carrion flies remains largely unknown, and especially so in pristine tropical areas. For example, while carrion flies living in temperate areas were shown to fly up to several kilometres per day (Norris 1965), estimates of daily dispersal by their tropical forest counterparts are not available. More detailed information on carrion flies' ecology (e.g. daily dispersal distances, life expectancy) would clearly be beneficial to evaluate their potential for iDNA analysis.

When examining the sources of differences in detection rates between NSA and SSA, besides a relative and expected increase in sensitivity with specific systems (as indicated by nonspecific mammal negative samples turning positive with species-specific PCR), new positives often stemmed from samples which were previously positive for another species with the NSA (Table 2). This is in line with previous findings showing that a mixture of vertebrate DNA sequences may sometimes be detectable from individual carrion flies (Calvignac-Spencer *et al.* 2013b). This has important consequences for carrion fly-based analyses as the presence of multiple templates might result in systematic detection biases when using NSA, which might favour the amplification of some templates over the others (as seen with 16S PCR applied to bacterial communities; Polz & Cavanaugh 1998). While such biases might end up with some taxa being simply undetectable, at a more general level, this would introduce much uncertainty to iDNA detection rate-based estimates of species abundance. For the latter kind of application, a preliminary examination of biases associated with the generic system of choice (possibly by tests performed on DNA mixtures) and corroboration of parts of the results with alternative generic systems or specific systems would likely be desirable.

Our data also provide some preliminary insight into the potential of carrion fly-based analyses with respect to species abundance estimation. The detection of extremely low-density mammals was possible. Jentink's duiker and black duiker locally reach maximal population densities of 1 and 2 individuals/km$^2$, respectively (East 1998), which translates into biomasses of only 70 and 40 kg/km$^2$ (based on average adult mass; Smith *et al.* 2003). Detection rates of these duiker species were correspondingly low in our analyses (one individual of each species in the SSA). Similarly, we had previously observed a low detection rate when using a chimpanzee (*Pan troglodytes verus*)-specific PCR system (one fly positive among 150 tested; Calvignac-Spencer *et al.* 2013b), this species also occurring at densities close to 1 individual/km$^2$ in this area, with an estimated biomass of 60 kg/km$^2$ (Smith *et al.* 2003). When considering results obtained with the SSA, the frequency of detection of colobine species (in terms of individual carrion flies and 'positive' traps) followed their respective biomasses in TNP: 1170 kg/km$^2$ for Western red colobus, 390 kg/km$^2$ for Western black-and-white colobus and < 215 kg/km$^2$ for olive colobus (derived from Korstjens 2001; Refisch & Koné 2005). This indicates that detection rates from carrion fly iDNA may correlate with local species biomass. However, while representing the second most detected monkey species in our study (when considering the long fragment assay results), the sooty mangabey exhibits only the third highest biomass at TNP, at best (155 kg/km$^2$), which is less than half the Western black-and-white colobus biomass. We note that, contrary to colobines, sooty mangabeys are terrestrial. Relative abundance may therefore only be meaningful for species known to exploit the same forest strata. With this limitation in mind, carrion flies which, unlike many blood-sucking insects, do not appear to have strong host preferences (Calvignac-Spencer *et al.* 2013b) could therefore be used not only to reveal the presence of a given species in a given habitat, but also to estimate relative biomasses (translatable into species' abundance) of multiple target species, or of different subpopulations of the same species. This possibility would, however, warrant further investigation, notably in truly comparative settings where contemporaneous estimates of species abundance are independently derived.

Our finding that microsatellite amplification of iDNA extracts is possible supports the notion that yet more in-depth, individual-based analyses of species' population ecology are imaginable from carrion fly iDNA. However, average genotyping PCR success for extracts derived from 'net' flies was only 45.6%. This is substantially lower than success rates from other, 'fresh', noninvasive DNA sources such as faeces, which might range between 79% and 94% in primates (Arandjelovic *et al.* 2009). Accordingly, direct genotyping from specimens positive for a given taxon (as shown, e.g. by sequencing a short mtDNA fragment) is unlikely to stand as a reasonable strategy. This prompted us to determine a threshold of mammal mtDNA concentration above which genotyping might be attempted (i.e. average PCR success is > 60%). Of the 479 tested

extracts, only 1% were found to be suitable for genotyping. This low rate does indicate that, while genotyping from carrion flies is possible, carrion flies will only rarely allow the amplification of microsatellites from iDNA. Preselection of extracts through qPCR will be an absolute prerequisite. Maximizing genotyping success from carrion flies is likely to be possible, for instance, by focusing on microsatellites that prove particularly amplifiable for a given species. Other sources of iDNA might stand as promising alternatives for iDNA genotyping. In semi-controlled settings, some studies obtained genotyping success rates from preselected, engorged haematophagous insects of up to 100% (e.g. mosquitoes of the family Culicidae feeding on bird hosts; Ligon *et al.* 2009). Leeches, which can take up impressive blood volumes and were recently used for a broad assessment of biodiversity in a Vietnamese forest (Schnell *et al.* 2012), may also represent good candidates for further iDNA-based genotyping attempts.

All in all, our results suggest that carrion fly-derived iDNA is a promising substrate for targeted detection of mammalian species and possibly a new tool for estimating their abundance. We also show that carrion fly DNA extracts can sometimes be used for genotyping individual mammals, although the very small proportion of suitable flies will most likely prevent population-scale applications. This source of mammalian DNA therefore allows assessing mammalian biodiversity at multiple levels.

## Acknowledgements

We thank the Ivorian authorities for long-term support and granting sampling permission, especially the Ministry of the Environment and Forests, the Ministry of Research, the directorship of the Taï National Park and the Swiss Research Center in Abidjan. We also thank the members of the Taï chimpanzee project for field support and Kathrin Nowak and Sonja Metzger for sampling of carrion fly specimens. We are grateful for practical help from the team of the sequencing laboratory of the Robert Koch Institute and for advice on laboratory analyses at the Max Planck Institute for Evolutionary Anthropology by Veronika Städele, Tillmann Fünfstück, Jörg Hans and Mimi Arandjelovic. We also warmly thank Rahel Sollmann for her critical reading of this manuscript and SivAina Leendertz and Roger Mundry for statistical advice.

## References

Ahumada JA, Silva CE, Gajapersad K *et al.* (2011) Community structure and diversity of tropical forest mammals: data from a global camera trap network. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **366**, 2703–2711.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Arandjelovic M, Guschanski K, Schubert G *et al.* (2009) Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. *Molecular Ecology Resources*, **9**, 28–36.

Bellemain E, Nawaz MA, Valentini A, Swenson JE, Taberlet P (2007) Genetic tracking of the brown bear in northern Pakistan and implications for conservation. *Biological Conservation*, **134**, 537–547.

Boessenkool S, Epp LS, Haile J *et al.* (2012) Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. *Molecular Ecology*, **21**, 1806–1815.

Bohmann K, Schnell IB, Gilbert M (2013) When bugs reveal biodiversity. *Molecular Ecology*, **22**, 909–911.

Calvignac-Spencer S, Leendertz FH, Gilbert MT, Schubert G (2013a) An invertebrate stomach's view on vertebrate ecology: certain invertebrates could be used as "vertebrate samplers" and deliver DNA-based information on many aspects of vertebrate ecology. *BioEssays*, **35**, 1004–1013.

Calvignac-Spencer S, Merkel K, Kutzner N *et al.* (2013b) Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Molecular Ecology*, **22**, 915–924.

Darriba D, Taboada GL, Doallo R, Posada D (2012) JMODELTEST 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

East R (1998) *African Antelope Database 1998* (IUCN Species Survival Commission Paper), IUCN, Gland, Switzerland, Cambridge, UK.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology Biotechnology*, **3**, 294–299.

Fujisawa T, Barraclough TG (2013) Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology*, **62**, 707–724.

Garros C, Gardès L, Allène X *et al.* (2011) Adaptation of a species-specific multiplex PCR assay for the identification of blood meal source in Culicoides (Ceratopogonidae: Diptera): applications on Palaearctic biting midge species, vectors of Orbiviruses. *Infection, Genetics and Evolution*, **11**, 1103–1110.

Kearse M, Moir R, Wilson A *et al.* (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.

Kent RJ (2009) Molecular methods for arthropod bloodmeal identification and applications to ecological and vector-borne disease studies. *Molecular Ecology Resources*, **9**, 4–18.

Kilger C, Krings M, Poinar H, Paabo S (1997) "Colony sequencing": direct sequencing of plasmid DNA from bacterial colonies. *BioTechniques*, **22**, 412–414, 416, 418.

Korstjens A (2001) *The mob, the Secret Sorority, and the Phantoms. An Analysis of the Socio-Ecological Strategies of the Three Colobines of Taï*. Utrecht University, Utrecht, The Netherlands.

Kouakou CY, Boesch C, Kuehl H (2009) Estimating chimpanzee population size with nest counts: validating methods in Taï National Park. *American Journal of Primatology*, **71**, 447–457.

Ligon RA, Burkett-Cadena ND, Liu M *et al.* (2009) Assessing mosquito feeding patterns on nestling and brooding adult birds using microsatellite markers. *The American Journal of Tropical Medicine and Hygiene*, **81**, 534–537.

Nicholson E, Collen B, Barausse A *et al.* (2012) Making robust policy decisions using global biodiversity indicators. *PLoS ONE*, **7**, e41128.

Norris K (1965) The bionomics of blow flies. *Annual Review of Entomology*, **10**, 47–68.

Pitzer JB, Kaufman PE, Tenbroeck SH, Maruniak JE (2011) Host blood meal identification by multiplex polymerase chain reaction for dispersal evidence of stable flies (Diptera: Muscidae) between livestock facilities. *Journal of Medical Entomology*, **48**, 53–60.

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, **64**, 3724–3730.

Pons J, Barraclough TG, Gomez-Zurita J *et al.* (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595–609.

R Development Core Team (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Refisch J, Koné I (2005) Impact of commercial hunting on monkey populations in the Taï region, Côte d'Ivoire1. *Biotropica*, **37**, 136–144.

Santiago ML, Range F, Keele BF *et al.* (2005) Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Taï Forest, Côte d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *Journal of Virology*, **79**, 12515–12527.

Schnell IB, Thomsen PF, Wilkinson N *et al.* (2012) Screening mammal biodiversity using DNA from leeches. *Current Biology*, **22**, R262–R263.

Schubert G, Stoneking CJ, Arandjelovic M *et al.* (2011) Male-mediated gene flow in patrilocal primates. *PLoS ONE*, **6**, e21514.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution*, **22**, 25–33.

Smith FA, Lyons SK, Ernest SKM *et al.* (2003) Body mass of late quaternary animals. *Ecology*, **84**, 3403.

Solberg KH, Bellemain E, Drageset O-M, Taberlet P, Swenson JE (2006) An evaluation of field and non-invasive genetic methods to estimate brown bear (*Ursus arctos*) population size. *Biological Conservation*, **128**, 158–168.

Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012) Environmental DNA. *Molecular Ecology*, **21**, 1789–1793.

Taylor PG (1996) Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Molecular Biology and Evolution*, **13**, 283–285.

Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. *Frontier of Zoology*, **5**, 12.

G.S., L.V., F.L. and S.C.S. supervised the study; G.S., M.S. and S.C.S. designed experiments; G.S., M.S., C.H. and K.M. performed laboratory experiments; G.S., M.S. and S.C.S. analysed the data; G.S, M.S. and S.C.S. drafted the article; all authors revised the manuscript.

## Data accessibility

Sanger sequences are provided as Appendix S1 (supporting information).

Microsatellite genotypes and quantitative PCR results Figs S1 and S2 are provided as Table S4.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** FASTA file comprising all mammalian sequences generated for this study.

**Table S1** Primer sequences and PCR thermocycling conditions used to amplify mammalian mitochondrial DNA fragments.

**Table S2** Summary of all PCR results.

**Table S3** Primers used for microsatellite analysis of carrion fly DNA extracts containing sooty mangabey mtDNA.

**Table S4** Microsatellite typing results.