*Bulletin of Entomological Research*

**cambridge.org/ber**

# Monitoring mosquito richness in an understudied area: can environmental DNA metabarcoding be a complementary approach to adult trapping?

Rafael Gutiérrez-López[1,2], Bastian Egeter[1], Christophe Paupy[3], Nil Rahola[3], Boris Makanga[4], Davy Jiolle[3], Vincent Bourret[1,5], Martim Melo[1,6,7] and Claire Loiseau[1,8]

[1]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBio, Laboratório Associado, University of Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal; [2]Animal Health Research Center, National Food and Agriculture Research and Technology Institute (INIA-CISA-CSIC), Valdeolmos, Spain; [3]MIVEGEC, Univ. Montpellier, CNRS, IRD, Montpellier 34394, France; [4]Institut de Recherche en Écologie Tropicale/CENAREST, BP 13354 Libreville, Gabon; [5]INRAE – Université de Toulouse UR 0035 CEFS, 31326 Castanet Tolosan, France; [6]MHNC-UP – Natural History and Science Museum of the University of Porto, Porto, Portugal; [7]FitzPatrick Institute of African Ornithology, University of Cape Town, Cape Town, South Africa and [8]CEFE, Université de Montpellier, CNRS, Montpellier, France

**Research Paper**

**Cite this article:** Gutiérrez-López R, Egeter B, Paupy C, Rahola N, Makanga B, Jiolle D, Bourret V, Melo M, Loiseau C (2023). Monitoring mosquito richness in an understudied area: can environmental DNA metabarcoding be a complementary approach to adult trapping? *Bulletin of Entomological Research* **113**, 456–468. https://doi.org/10.1017/S0007485323000147

Received: 6 September 2022
Revised: 28 December 2022
Accepted: 19 March 2023
First published online: 15 May 2023

**Keywords:**
Invasive species; metabarcoding; oil palm plantation; rainforest; vectors

**Corresponding author:**
Rafael Gutierrez Lopez,
Email: rafael.gutierrez@inia.csic.es

## Abstract

Mosquito surveillance programmes are essential to assess the risks of local vector-borne disease outbreaks as well as for early detection of mosquito invasion events. Surveys are usually performed with traditional sampling tools (i.e., ovitraps and dipping method for immature stages or light or decoy traps for adults). Over the past decade, numerous studies have highlighted that environmental DNA (eDNA) sampling can enhance invertebrate species detection and provide community composition metrics. However, the usefulness of eDNA for detection of mosquito species has, to date, been largely neglected. Here, we sampled water from potential larval breeding sites along a gradient of anthropogenic perturbations, from the core of an oil palm plantation to the rainforest on São Tomé Island (Gulf of Guinea, Africa). We showed that (i) species of mosquitoes could be detected via metabarcoding mostly when larvae were visible, (ii) larvae species richness was greater using eDNA than visual identification and (iii) new mosquito species were also detected by the eDNA approach. We provide a critical discussion of the pros and cons of eDNA metabarcoding for monitoring mosquito species diversity and recommendations for future research directions that could facilitate the adoption of eDNA as a tool for assessing insect vector communities.

## Introduction

Factors associated with global change, such as temperature increase, land-use change and the increasing spread of invasive species, are leading to a considerable loss and reorganization of biodiversity (Harold, 2000; Segan *et al.*, 2016; Eriksson and Hillebrand, 2019), with important consequences for the emergence of infectious diseases that affect wildlife, livestock and human populations (Caminade *et al.*, 2019; Smith *et al.*, 2019). Among global emerging infectious disease events, vector-borne diseases are disproportionately over-represented (Swei *et al.*, 2020) and constant efforts for monitoring insect vector populations should be carried out in locations at risk (Pedersen *et al.*, 2009). Among insect vectors, mosquitoes (Culicidae), with more than 3500 described species widely distributed around the world (Harbach, 2013), are considered among the main insect vectors involved in the transmission of pathogens including viruses, protozoans and nematodes. Three main genera, *Anopheles*, *Aedes* and *Culex* are considered of medical importance for humans and transmit pathogens causing diseases to more than 700 million people annually, resulting in over 1 million deaths (World Health Organization, 2020). In the last few decades, the rapid worldwide spread of the invasive yellow fever mosquito *Aedes aegypti* and the Asian tiger mosquito *Aedes albopictus* is producing novel epidemiological scenarios (Bonizzoni *et al.*, 2013; Iwamura *et al.*, 2020). Early detection of mosquito invasion events, as well as continued surveillance of such invasions, is becoming essential to assess the risks of local mosquito-borne disease outbreaks. In addition, it seems essential to understand the ecological interactions between mosquito species at breeding sites to evaluate the competitiveness of indigenous species (Juliano, 2009).

To date, surveys of mosquito species have been performed with traditional sampling using ovitraps and dipping method for immature stages or with light/decoy traps and human

landing catches for adults (Focks, 2004). Skilled entomologists are able to identify specimens using morphological traits (Besansky *et al.*, 2003; Hajibabaei *et al.*, 2007), however some species are indistinguishable morphologically (e.g., cryptic species of *Anopheles*) (Coetzee and Koekemoer, 2013). In addition, the identification of different mosquito stages (i.e., eggs, larvae and adult mosquito specimens) needs solid knowledge from experts in entomology. Identification can be time consuming, especially when there are many species such as in the tropics, but also characterization at the species level becomes very difficult, if not impossible, when specimens are too damaged (Foley *et al.*, 2007). Developments in molecular techniques over the past decade, coupled with reduced sequencing costs, have made use of environmental DNA (eDNA) as an approach with a huge potential to survey micro-biodiversity in the field. eDNA is DNA that is shed by organisms (e.g., through faecal waste, dead skin, gastrointestinal tract cells, gametes or via post-mortem degradation), and it has formed the basis of numerous studies focussed on vertebrate detection (Ficetola *et al.*, 2008; Goldberg *et al.*, 2011; Jerde *et al.*, 2011; Minamoto *et al.*, 2012; Thomsen *et al.*, 2012; Spear *et al.*, 2015; Egeter *et al.*, 2018), and more recently for invasive invertebrates (Clusa *et al.*, 2017; Klymus *et al.*, 2017; Mychek-Londer *et al.*, 2020). In natural habitats, eDNA is affected by a variety of factors, such as temperature, microbial activity, pH (Seymour *et al.*, 2018), conductivity (Collins *et al.*, 2018), water chemistry or ultraviolet radiation. It is degraded over time, but can remain at detectable levels weeks after an organism's removal (Dejean *et al.*, 2011; Barnes *et al.*, 2014; Pilliod *et al.*, 2014). Hence, most eDNA detection is expected to indicate a current or recent colonization of the habitat (Piaggio *et al.*, 2014), making it a potentially suitable method for contemporary surveillance of aquatic populations, such as mosquito aquatic stages.

Although studies have shown the usefulness of eDNA metabarcoding for the monitoring of numerous invertebrate species, to the best of our knowledge, only few studies have demonstrated the usefulness of this technique for detection of mosquito species in particular. Schneider *et al.* (2016) analysed the potential of eDNA for the detection of invasive *Aedes* mosquitoes in Europe. They collected water samples in the field and used both quantitative real-time polymerase chain reaction (qPCR) and eDNA metabarcoding of a short fragment of the 16S rRNA gene of the Culicidae family. Both molecular methods gave comparable results and performed better than the traditional survey methods, however, the detection capacity decreased by half 10 days after the removal of the larvae. Those authors recommended for the eDNA approach to be used as a complement to traditional captures. Two other studies compared the effectiveness of eDNA approaches with traditional sampling techniques to detect mosquito larvae diversity in the field (Boerlijst *et al.*, 2019; Krol *et al.*, 2019). These studies both used eDNA primers targeting the cytochrome oxidase I (COI) gene. Boerlijst *et al.* (2019) found that 98% of the Culicidae species were correctly identified using eDNA, suggesting that eDNA-based approaches are reliable and can be even more reliable than traditional dipping methods for certain species. However, both studies yielded only a subset of the adult community known in their field sites. Species that were detected with eDNA were generally the most abundant species in the traps indicating that the eDNA metabarcoding method was more likely to pick up more abundant species than rare mosquito species (Krol *et al.*, 2019). Although eDNA metabarcoding can increase the accuracy of identification, while reducing the cost and time, compared to classical barcoding, it must be integrated with classical taxonomy and molecular methods for comprehensive ecological studies (Ruppert *et al.*, 2019). The use of eDNA is a booming technique, but also has many limitations, including the degradation of eDNA in the environment, especially in tropical regions, as well as the sample preservation methods. In addition, one of the important considerations in eDNA metabarcoding studies is the primer design (Ruppert *et al.*, 2019). Primers for different genes vary in coverage, resolution and inter-taxon bias. COI gene is the standard gene for barcoding, but other regions such as 12S or 16S ribosomal RNA may be more appropriate for different taxa (Epp *et al.*, 2012; Taberlet *et al.*, 2012; Deiner *et al.*, 2017; Hering *et al.*, 2018). Primers for eDNA metabarcoding must be short enough to amplify degraded samples, identical for the same species, but variable between species, allowing amplification of a variety of species (Epp *et al.*, 2012).

In our study on São Tomé Island, Gulf of Guinea (Africa), we wanted to evaluate the richness of mosquito species along a gradient of anthropogenic disturbances in order to compare the assemblage of species between human habitation areas (i.e., village with domestic animals), intensive agricultural areas (i.e., oil palm plantations) and natural neighbouring forested areas. To assess the mosquito richness at these three habitat types, we collected (i) water from larval breeding sites to perform eDNA metabarcoding using COI and (ii) adult specimens using CDC light traps set up in trees. The aims of this study were (i) to compare our metabarcoding results with the visual identification of larvae and the light traps captures, taking into account the samples' characteristics (i.e., water turbidity, containers), (ii) to identify the assemblage of mosquito species along a gradient of anthropogenic disturbance, (iii) to detect the presence of the invasive tiger mosquito *Ae. albopictus* which recently colonized the island (Reis *et al.*, 2017) and finally (iv) to perform a short review of the pros and cons of the eDNA metabarcoding as a complementary methodological approach to traditional ones.

## Materials and methods

### Study sites and sampling

Water sampling took place in three different types of habitats in October 2019 on São Tomé Island (Gulf of Guinea, Africa): (i) a small village located in the middle of the oil palm plantation (0°6′57.308″ N; 6°35′33.414″ E), (ii) the oil palm plantation that surrounds the village and (iii) the secondary rainforest adjacent to the plantation at 1 km from the village (fig. 1).

We collected 37 water samples (30 ml each, with 10 ml of Longmire solution added for preservation) (Williams *et al.*, 2016), from a variety of containers, either natural or artificial, that presented variation in water turbidity (defined as either clean or dirty; fig. 2, table 1). Eighteen (48.65%) of the water samples were taken in larval development sites where larvae were present, while 19 samples (51.35%) came from sites with no larvae detected. When larvae were visually detected, they were identified at least at the genus level (table 1), except for three samples for which a correct *de visu* identification was not possible.

A total of 47 CDC light traps were set up to collect adult mosquitoes three consecutive nights in each habitat in parallel of the water sampling (fig. 2). Eleven traps were in the village, 18 in the oil palm plantation and 18 in the forested areas. Every morning, traps were gathered and placed in a freezer for 15 min. Then all arthropods were sorted and dipterans of interest were identified

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

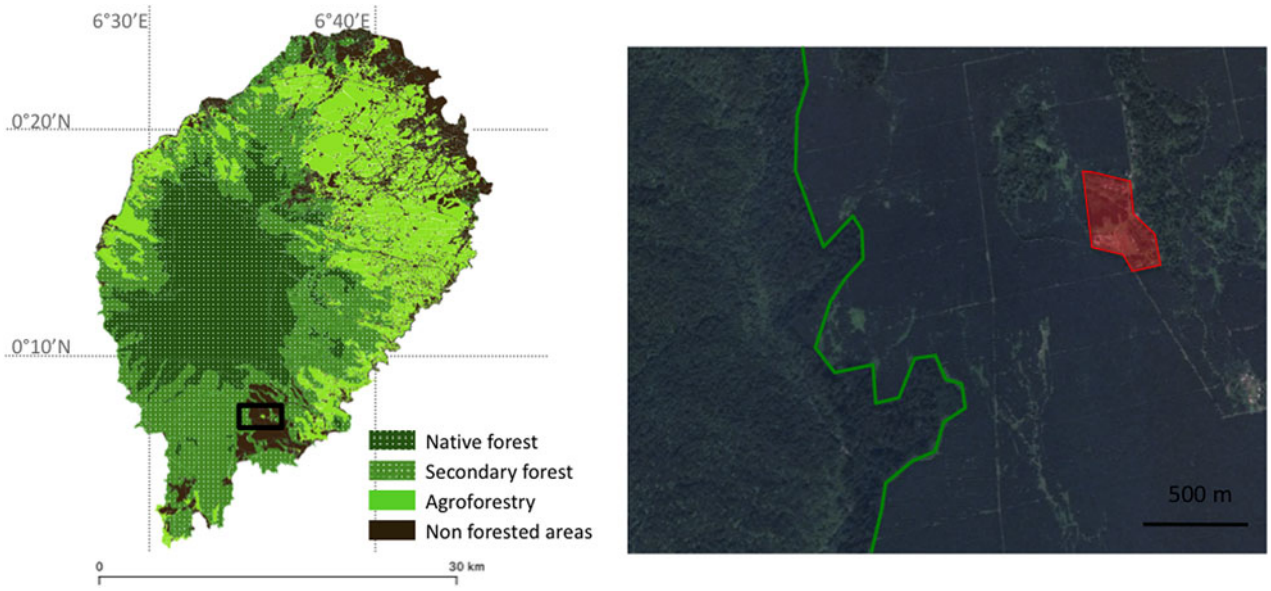


**Figure 1.** Left: Map of São Tomé Island (Gulf of Guinea, Africa), with the black frame representing the sampling area in the southeast of the island. Right: Satellite picture of that area, with the village (circled in red), surrounding by the oil palm plantation; the green line being the border between the oil palm plantation and secondary forest.

morphologically using a Leica S9E stereomicroscope (Leica Microsystems GmbH, Germany). Adults and larvae mosquito were identified to species or species group using different morphological keys and detailed descriptions are provided in Edwards (1941), Hopkins (1952), Gillies and Coetzee (1987), Service (1990) and Ribeiro *et al.* (1998). Our sources of data on species naming were based on that recorded in the Walter Reed Biosystematics Unit Mosquito Catalogue (http://www.mosquitocatalog.org).

### *Molecular methods*

DNA extractions were performed in a laboratory 'clean-room' (at CIBIO, Portugal) equipped with UV radiation where strict protocols are followed for the prevention of contaminations (disposable laboratory clothing, UV sterilization of all equipment before entering the laboratory and laboratory cleaning with a 60% dilution of bleach between extraction batches). Prior to filtration, the water samples were manually shaken for 5 min (Civade *et al.*,


**Figure 2.** Photograph representing the sampling methods used in our study: (A) sampling water in an artificial container, (B) sampling in a natural rock hole and (C) a CDC light trap in the oil palm plantation.

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

**Table 1.** Characteristics of water samples ($n = 37$) collected on São Tomé Island (village $n = 17$; oil palm plantation $n = 8$; forest $n = 12$; with A or N for artificial or natural containers respectively) and the species identification, either visually or by metabarcoding (COI marker).

| | Field | | | Visual | | eDNA metabarcoding | |
|---|---|---|---|---|---|---|---|
| ID | Location | Container | Water | Larvae (quantity) | Species | Seq | Species assignment |
| V1 | Village | N: puddle | Clean | Yes (1) | Unknown | Yes | *Ae. Albopictus* |
| V2 | Village | A: tyre | Dirty | Yes (>20) | *Ae. albopictus*, *Eretmapodites* sp. | Yes | *Ae. albopictus*, *An. coluzzii*, *Er. intermedius*, *Cx. cambournaci* |
| V3 | Village | A: pot | Clean | Yes (>20) | *Ae. albopictus* | Yes | *Ae. albopictus*, *An. coluzzii*, *Er. intermedius*, *Cx. cambournaci* |
| V4 | Village | A: tyre | Clean | Yes (>20) | *Ae. albopictus*, *Culex* sp., *Eretmapodites* sp. | Yes | *An. coluzzii*, *Cx. decens* |
| V5 | Village | A: tyre | Dirty | Yes (>20) | *Eretmapodites* sp. | Yes | *Ae. albopictus*, *An. coluzzii*, *Er. intermedius* |
| V6 | Village | A: tyre | Clean | Yes (>20) | *Ae. albopictus*, *Culex* sp. | Yes | *An. coluzzii*, *Cx. decens* |
| V7 | Village | A: tyre | Clean | Yes (>20) | *Ae. albopictus*, *Eretmapodites* sp. | Yes | *Ae. albopictus*, *Cx. decens* |
| V8 | Village | N: bamboo | Clean | Yes (<10) | *Ae. albopictus* | Yes | *Ae. albopictus* |
| V9 | Village | N: fruit shell | Dirty | Yes (<10) | *Ae. albopictus*, *Eretmapodites* sp. | Yes | *Ae. albopictus*, *Cx. cambournaci*, *Er. intermedius* |
| V10 | Village | A: bottle | Clean | Yes (1) | Unknown | No | |
| V11 | Village | N: crab hole | Dirty | No | | No | |
| V12 | Village | A: can | Clean | No | | No | |
| V13 | Village | N: puddle | Clean | Yes (<10) | *Anopheles* sp. | Yes | *An. coluzzii* |
| V14 | Village | N: puddle | Clean | Yes (<10) | *Anopheles* sp. | Yes | *An. coluzzii* |
| V15 | Village | A: tyre | Clean | Yes (>20) | *Ae. albopictus* | Yes | *Ae. albopictus*, *Cx. cambournaci* |
| V16 | Village | N: puddle | Clean | Yes (<10) | *Anopheles* sp. | Yes | Other family |
| V17 | Village | A: bottle | Clean | Yes (1) | Unknown | Yes | Other family |
| P1 | Plantation | N: puddle | Clean | Yes (1) | *Anopheles* sp. | No | |
| P2 | Plantation | N: puddle | Clean | No | | No | |
| P3 | Plantation | N: rocks | Dirty | No | | No | |
| P4 | Plantation | N: puddle | Clean | No | | Yes | Other family |
| P5 | Plantation | N: puddle | Clean | No | | No | |
| P6 | Plantation | N: leaf | Clean | No | | No | |
| P7 | Plantation | N: rocks | Dirty | No | | No | |
| P8 | Plantation | N: rocks | Clean | No | | No | |
| F1 | Forest | N: leaf | Clean | No | | Yes | Other family |
| F2 | Forest | N: rocks | Clean | No | | No | |
| F3 | Forest | N: tree hole | Dirty | No | | No | |
| F4 | Forest | N: river | Clean | No | | No | |
| F5 | Forest | N: tree hole | Dirty | No | | No | |
| F6 | Forest | N: rocks | Clean | No | | No | |
| F7 | Forest | N: rocks | Clean | No | | No | |
| F8 | Forest | N: rocks | Clean | No | | No | |
| F9 | Forest | N: tree hole | Clean | Yes (1) | *Culex* sp. | Yes | Other family |
| F10 | Forest | N: tree hole | Dirty | No | | No | |
| F11 | Forest | N: tree hole | Dirty | No | | No | |
| F12 | Forest | N: bamboo | Dirty | Yes (<10) | *Ae. albopictus* + *Culex* sp. | Yes | *Ae. aegypti*; *Cx. sasai* |

In five sequenced samples, we did not detect Culicidae species but other arthropod families (see fig. 3).

2016; Lopes *et al.*, 2017) to homogenize the water column within the 50 ml Falcons tube. To concentrate material to a suitable volume for subsequent extraction, we filtered each sample (40 ml; water + Longmire) by pouring it into a sterile container (100-ml filtering cup; Nalgene Polysulphone Filter Holder with funnel, Thermo Scientific, USA) through sterile 47 mm nitrocellulose disc filters, 0.45 µm pore size (Whatman, UK), using a vacuum pump. The disc filters were cut into small pieces and placed in a 50 ml Falcon tube with 1.5 ml 3 M sodium acetate and 33 ml absolute ethanol for the water samples. These samples were placed in a carousel rotating shaker for 2 h at room temperature to homogenize the samples. Subsequently, the water samples were stored for 24 h at −20°C. Filter manipulation was performed with sterilized forceps between samples. Subsequently, the samples were centrifuged at 3184*g* for 45 min, at 10°C to recover the precipitated DNA and/or cell debris (Peixoto *et al.*, 2021). The supernatant was discarded (Valiere and Taberlet, 2000) and we performed DNA extraction on the pellet using the Dneasy Blood and Tissue Kit following the manufacturer's instructions (Qiagen, Hilden, Germany) (Gutiérrez-López *et al.*, 2015). The pellet was exposed to enzymatic lysis using proteinase K in a carousel rotating shaker for 1 h at 56°C and the supernatant was spun through the column purification of DNA. We included a negative control (distilled water) in each set of extractions to monitor potential contaminations. The DNA was eluted in 80 µl of ultrapure sterilized MilliQ water. After extraction, DNA was quantified using the Qubit high-sensitivity dsDNA assay (Thermo Fisher Scientific). DNA metabarcoding libraries were prepared by amplifying a 200 bp fragment of the COI genomic region using the following primers: eCul-F (5′ GGRKCHGGDACWGGDTGAAC 3′) and eCul-R (5′ GATCAWACAAATAAAGGTAWTCGATC 3′) (Krol *et al.*, 2019). Illumina sequencing primer sequences were attached to the 5′ ends of PCR primers with i7 and i5 as indexes (*Index 1 (i7) Adapter: P7-P5'CAAGCAGAAGACGGCATACGAGAT*[*i7*] *GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC; Index 2 (i5) Adapter: P5-P7'AATGATACGGCGACCACCGAGATCTACAC*[*i5*] *ACACTCTTTCCCTACACGACGCTCTTCCGATCT).* PCRs were carried out in a final volume of 25 µl, containing 2.5 µl of template DNA, 0.5 µM of each primer, 12.5 µl of Supreme NZYTaq 2× Green Master Mix (NZYTech) and ultrapure water up to 25 µl. The thermocycler programme for DNA amplification started with an initial denaturation step at 95°C for 5 min, followed by 40 cycles of 95°C for 30 s, 58°C for 45 s, 72°C for 30 s and a final extension step at 72°C for 10 min.

The oligonucleotide indices, which are required for multiplexing different libraries in the same sequencing pool, were attached in a second PCR round with identical conditions but for only ten cycles and 60°C as the annealing temperature. We used in-house designed indexes, which are a combinatorial set of 24 i5 and 24 i7 indexes, which we have pre-mixed and randomized. They are 8-bp long and the Levenshtein distance between any two indexes is at least 3. A negative control containing no DNA (MilliQ water) was included in every PCR round to check for contamination during library preparation. The libraries were run on 2% agarose gels stained with GreenSafe (NZYTech), and imaged under UV light to verify the library size. Libraries were purified using the Mag-Bind RXNPure Plus magnetic beads (Omega Biotek). We pooled the samples and purified the resulting pool following the same method (1× of magnetic beads). The purified pool was run through a size-select eGel to precisely select the band of interest. Libraries were quantified using the Qubit high-sensitivity dsDNA assay (Thermo Fisher Scientific).

Very low library quantification was detected in 18 samples that were removed for sequencing. These samples corresponded to water samples in which only one larva ($n = 2$) or none were detected visually ($n = 16$; table 1). Therefore, 19 samples were selected for sequencing and were pooled in equimolar amounts and re-purified for double size selection in an e-gel system (Life Technologies) for primer dimer elimination. The pool was sequenced in a fraction (1/16) of a MiSeq PE300 run (Illumina). Library preparation and sequencing were carried out by AllGenetics & Biology SL (www.allgenetics.eu).

### *Bioinformatics analyses and taxonomic assignment*

Illumina paired-end raw files consist of forward (R1) and reverse (R2) reads sorted by library and their quality scores. The indices and sequencing primers were trimmed from the samples using software CUTADAPT (Martin, 2011) and the quality of the FASTQ files was checked using the software FastQC (Andrews, 2010). Plots summarizing the quality across bases of R1 and R2 reads were generated by using MultiQC (Ewels *et al.*, 2016) (see Supplementary file). The merging of the R1 and R2 reads was performed with FLASH2 (Magoč and Salzberg, 2011). The mismatch resolution in the overlapping region (minimum overlap of 30 base pairs) was accomplished by keeping the base with the higher quality score. We used CUTADAPT 1.3 software (Martin, 2011) to remove sequences that did not contain the PCR primers (allowing up to two mismatches) and sequences that ended up being shorter than 145 nucleotides and larger than 210 nucleotides. The sequences were quality-filtered (minimum Phred quality score of 20), then were dereplicated (-derep fulllength) and clustered at a similarity threshold of 97% (-cluster fast, -centroids option) and sorted (-sortbysize) using VSEARCH (Rognes *et al.*, 2016). *De novo* chimaera detection was carried out using the UCHIME algorithm (Edgar *et al.*, 2011) implemented in VSEARCH.

We conducted the taxonomic assignment of each operational taxonomic unit (OTU) (I) using a customized taxonomic COI reference database. The database included (i) newly generated mosquito sequences of four species sampled during the fieldwork using light traps (*Aedes nigricephalus*, *Culex cambournaci*, *Uranotaenia micromelas* and *Uranotaenia connali*; Genbank accession numbers ON504276–ON504279), and (ii) sequences downloaded from the National Center for Biotechnology Information (NCBI) and the Barcode of Life Data System (BOLD) databases (Ratnasingham and Hebert, 2007) (accessed on March 2022). These mosquito sequences (from mosquito species known to be present on São Tomé; table S1) were added to the database build using RESCRIPt (Robeson *et al.*, 2021) (last version on July 2020) based on the BOLD reference database (Ratnasingham and Hebert, 2007).

We employed the script *feature-classifier classify-consensus-vsearch* implemented in Qiime2 (Bokulich *et al.*, 2018) and the VSEARCH algorithm (Rognes *et al.*, 2016) with a sequence similarity threshold of 95%. In addition, we used the *top-hits-only* option in the VSEARCH command to recover only the hit with the highest percentage of identity. In spite of the multiple top hits used in the consensus taxonomic assignment carried out by VSEARCH, this option allows the assignment of the query to the closest reference sequence. The table resulting from this step lists the number of sequences from each OTU found in each sample and their corresponding taxonomic information (table S2 – before OTU filtering). Subsequently, based on the results of this table, we applied several different filters. We removed singletons

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

(i.e., OTUs containing only one-member sequence in the whole data set). In DNA metabarcoding studies, it has been observed that a low percentage of the reads of a library can be assigned to another library. This phenomenon, referred to as mistagging, tag jumping, index hopping, index jumping, etc. is the result of the misassignment of the indices during library preparation, sequencing and/or demultiplexing steps (Esling *et al.*, 2015; Bartram *et al.*, 2016; Guardiola *et al.*, 2016; Illumina, 2017). In order to correct for this phenomenon, OTUs occurring at a frequency below 0.01% in each sample were removed. Finally, only the OTUs that matched any reference sequence in the database at a minimum similarity threshold of 85% were kept in the OTU table. Therefore, the unidentified OTUs ('Unassigned') were removed from the OTU table for downstream analysis (table S3 – after OTUs filtering). Six samples (V16, V17, P4, F1, F7, F9) had no OTUs assigned to the family Culicidae.

The alpha rarefaction plots show the number of OTUs obtained with a rarefied number of sequences in each sample. These plots were generated using the OTU table before (table S2) and after (table S3) the OTU filtering (fig. S1). The vertical axis displays the number of OTUs observed at different subsampling depths. When the rarefaction curves tend towards saturation, the sequencing depth is considered to be sufficient to retrieve most of the taxa diversity. We have to note that curve from sample V8 did not reach the plateau in the number of OTUs observed (see fig. S1.B, rarefaction plot after the OTU filtering).

In order to easily visualize the breakdown of taxonomic classification, stacked bar plots showing the relative abundance of each OTU in each sample were generated at the order, family and species levels (fig. 3). In DNA metabarcoding studies, OTU relative abundance is defined as the number of reads assigned to that OTU divided by the total number of reads. Note that the PCR may cause biases due to differences in primer specificity. These biases can cause taxa with low representation in the original DNA sample to become more abundant in the final results. As a result, this bias prevents from correctly inferring the abundance of species in the original DNA sample. For example, if SPECIES A is represented by the 35% of the sequences in SAMPLE 1, and SPECIES B is represented by the 50% of the sequences in the same sample, we cannot reliably conclude that there was more SPECIES B DNA in the original sample. That being said, it is expected that, within the same study, the PCR bias always go in the same direction. Therefore, it is possible to compare how the abundance of a given taxon varies across different samples with a similar composition. For example, if SPECIES A is represented by the 35% of the sequences in SAMPLE 1 and by the 10% in SAMPLE 2, we can conclude that there was less SPECIES A DNA in SAMPLE 2 (Geisen *et al.*, 2015; Thomas *et al.*, 2016; Matesanz *et al.*, 2019).

Finally, we extracted the representative sequences for each of the picked OTUs before and after the OTU filtering process. For the particular case of the taxonomic assignment of OTUs to *Eretmapodites intermedius*, we performed a blast in NCBI and the results are shown in fig. S2.

DNA metabarcoding analyses were carried out by AllGenetics & Biology SL (www.allgenetics.eu).

## Results

### Visual and genetic detection

Of the 19 water samples collected from sites where no larvae were detected visually, one was positive for *Ae. albopictus* (5%; table 1), two others were found with chironomids (Diptera) or coleopterans (10%) and 16 could not be sequenced because of the low library DNA quantities (84%). Of the 18 water samples in which larvae were seen, eDNA metabarcoding detected Culicidae in 13 (72%), three of which had detections of other dipterans and branchiopodans (16%), and two could not be sequenced because of the low library DNA quantities (11%; table 1). When larvae were present at the collection site, one or two Culicidae genera were identified visually in each sample, whereas eDNA metabarcoding detected up to four genera per sample (table 1).

We recovered DNA sequences in 14 water samples out of the 26 considered as clean (53.8%), and in four out of seven considered as dirty (57%). Although our sample sizes remain small, we found that the turbidity of the water did not appear to be a limitation for eDNA metabarcoding (chi-squared test, $\chi^2 = 0.33$).

Overall, the taxonomic assignments revealed four orders of arthropods that comprised of 13 families. Within Culicidae, taxonomic assignments at the species level for the genus *Anopheles* returned *An. coluzzii*, the main human malaria vector on the island (Chen *et al.*, 2019). For the genus *Aedes*, the taxonomic assignments at the species level returned the invasive tiger mosquitoes *Ae. albopictus* and *Ae. aegypti*. All OTUs that matched the genus *Eretmapodites*, an endemic genus of the Afrotropical region and vector of various viruses (Bamou *et al.*, 2021), were assigned to *Er. intermedius* (Supplementary fig. S2). As for the *Culex* genus, OTUs were assigned to *Cx. cambournaci*, *Cx. decens* and *Cx. sasai*.

In summary, 12 species of Culicidae were detected, seven with eDNA metabarcoding and nine with CDC light traps. Four species were common to both approaches: *Ae. albopictus*, *An. coluzzii/gambiae*, *Cx. Cambournaci* and *Cx. decens*, all collected in the village (table 2; fig. 4).

### Habitat effects on species detection

In the village, five orders and eight families of arthropods were found. The Culicidae was the dominant family found in the village, with 78% of the total reads from the village attributed to the genera *Aedes*, *Anopheles*, *Culex* and *Eretmapodites*. The invasive mosquito *Ae. albopictus* and the malaria vectors *An. coluzzii* were present respectively in 57% ($N = 8$) and 50% ($N = 7$) of the samples collected in the village that led to amplification. *Ae. albopictus* was found in both artificial and natural breeding sites, while *Ae. aegypti* was totally absent from the village, a pattern that had already been noted in previous surveys (Reis *et al.*, 2017). *Culex* spp. were present in half of the village samples that could be sequenced (7 out of 14; fig. 3).

In the plantation, in the eight potential breeding sites that were sampled, we did not detect any larvae visually. The only sample whose amplification worked gave two OTUs affiliated to the Chironomidae family (order Diptera; see Supplementary tables S2 and S3).

In the forest, four orders and four families of arthropods were found, with the Chironomidae being the dominant family with 73% of the reads (fig. 3). In the forest, *Cx. sasai* and *Ae. aegypti* were detected in the same sample (fig. 3).

## Discussion

Our study showed that eDNA metabarcoding could be a complementary method to the light or decoy traps to recover mosquito

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

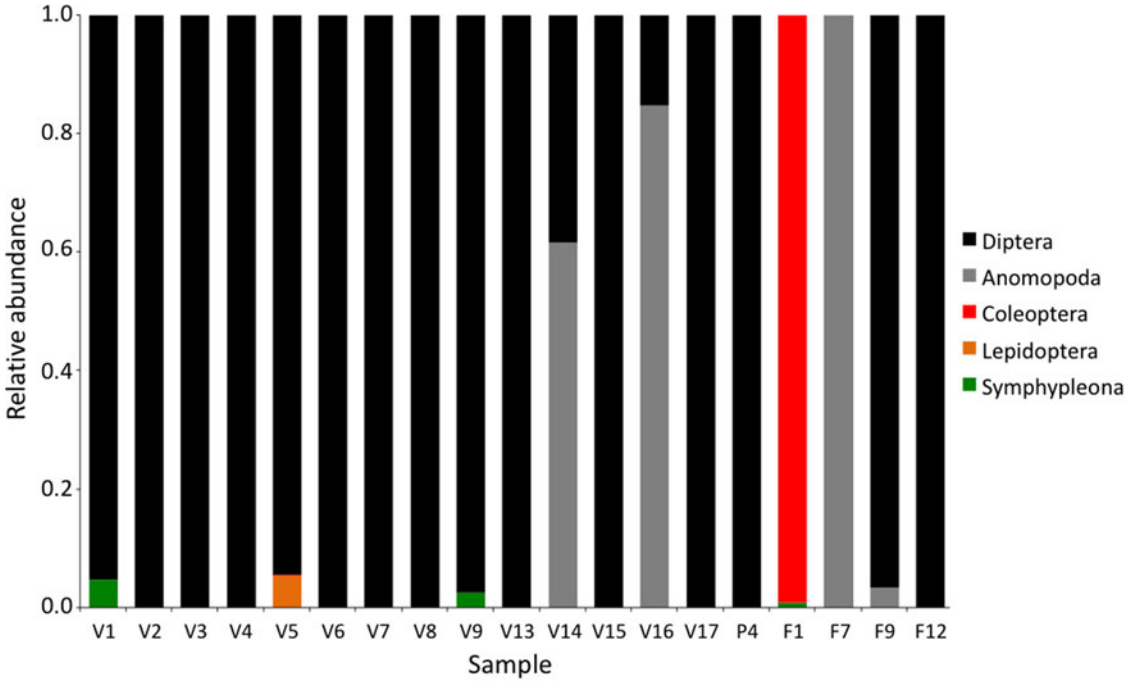


**Figure 3.** Stacked bar plots of the various arthropods detected along the anthropogenic gradient using eDNA metabarcoding (COI marker): (a) order level, (b) family level for the Diptera order and (c) species level for the family Culicidae (V, village; P, plantation; F, forest).

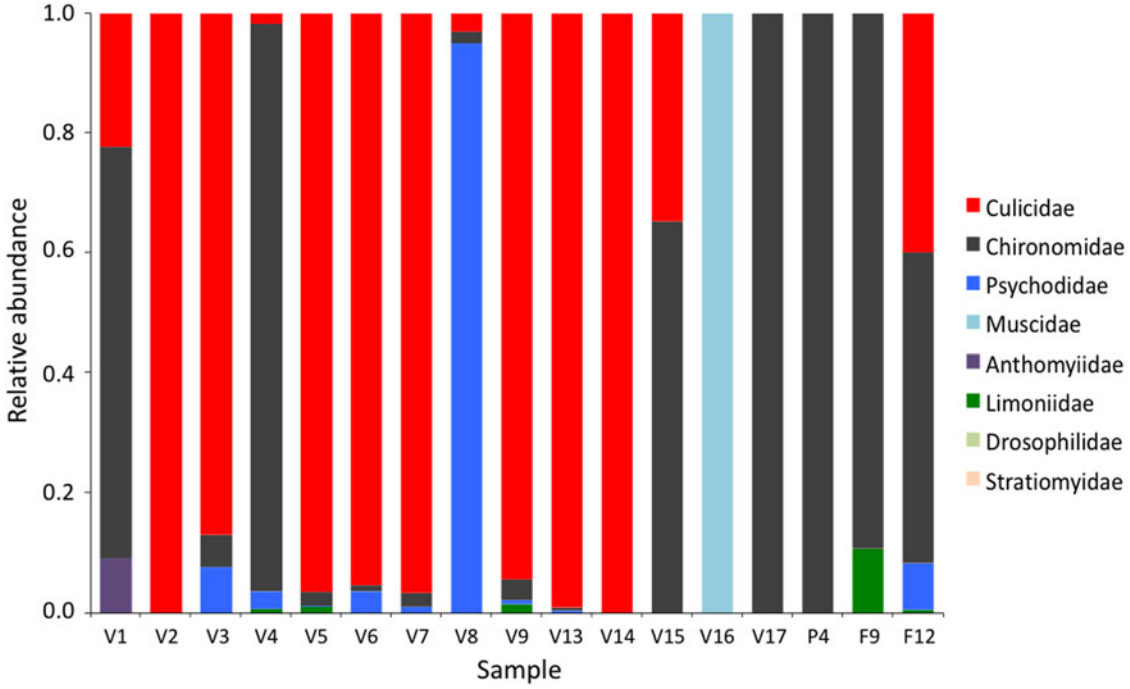


**Figure 3.** Continued.

diversity, and help to evaluate the assemblage of species using the same breeding sites. In particular, eDNA metabarcoding was able to detect species that were not captured with light traps and picked up different assemblage of mosquito species associated with the degree of anthropogenic disturbance.

In the oil palm plantation, we found larvae of mosquitoes by *de visu* at one sampling location. eDNA metabarcoding detected only one family of DIPTERA (Chironomidae) with very few reads, but no mosquito species. This result is not surprising and is consistent with the view that the core of oil palm plantations is overall poor in terms of arthropod diversity (Koh and Wilcove, 2008; Turner and Foster, 2009; Fayle *et al.*, 2010; Ghazali *et al.*, 2016). Recently, Young *et al.* (2021) also found that mosquito abundance in oil palm plantations in Borneo was lower than that in the forest. On the contrary, in the village, the arthropod diversity was much higher than that in the surrounding

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

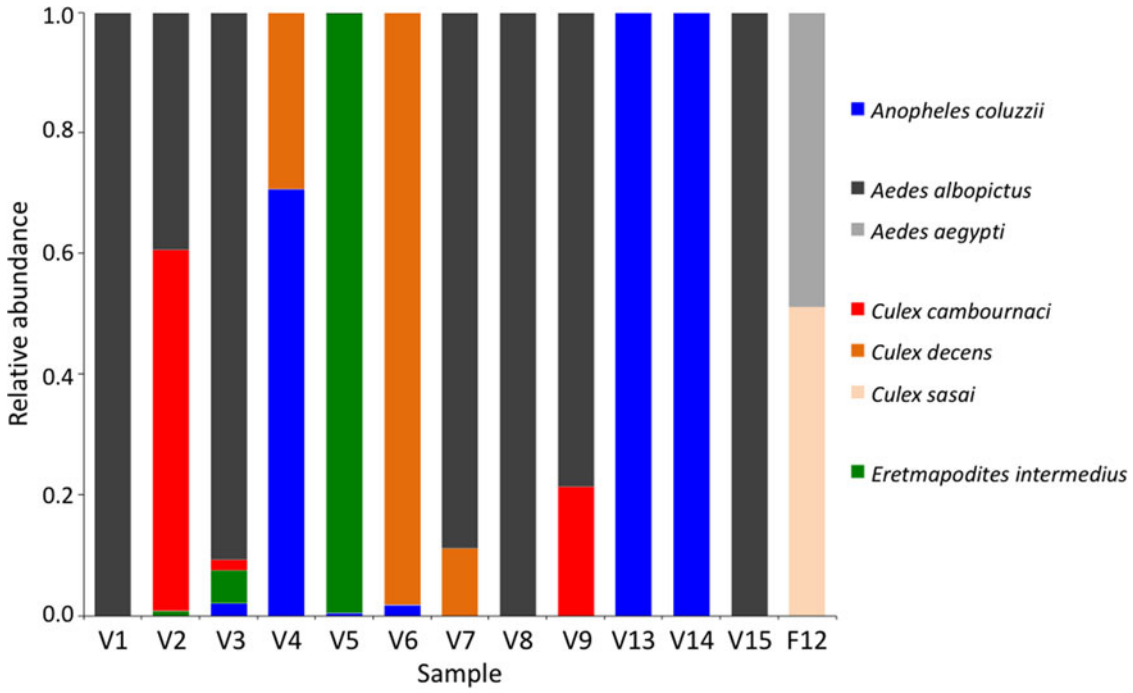


**Figure 3.** Continued.

**Table 2.** List of mosquito species detected in the water samples, visually and with eDNA metabarcoding, and with CDC traps along the gradient of anthropogenic disturbance in São Tomé Island

| | Water sample | | CDC light trap |
|---|---|---|---|
| | Visual identification (larvae) | eDNA | Visual identification (adults) |
| Village | *Ae. albopictus*<br>*Anopheles* sp.<br>*Culex* sp.<br>*Eretmapodites* sp. | *Ae. albopictus*<br>*An. coluzzii*<br>*Cx. cambournaci*<br>*Cx. decens*<br>*Er. intermedius* | *Ae. albopictus*<br>*An. coluzzii/gambiae*<br>*Cx. cambournaci*<br>*Cx. decens*<br>*Culex micolo*<br>*Ur. connali*<br>*Ur. micromelas* |
| Oil palm plantation | *Anopheles* sp. | None | *Anopheles coustani*<br>*Cx. cambournaci*<br>*Lutzia tigripes*<br>*Ur. micromelas* |
| Forest | *Ae. albopictus*<br>*Culex* sp. | *Ae. aegypti*<br>*Cx. sasai*[a] | *Cx. cambournaci*<br>*Ur. micromelas* |

See also fig. 4 for illustration.
[a]Incorrect taxonomic assignment likely due to incomplete molecular reference database.

plantations with eight families of Diptera recorded. Culicidae was the predominant family: *Ae. albopictus* accounted for 36% of the reads, followed by *Culex* species (33.5%), while *Anopheles* genus was the least abundant, with 3.3% of the reads. Although more surveys are needed, *Ae. albopictus*, which recently colonized the island (Reis *et al.*, 2017), shared breeding sites with *Culex*, *Eretmapodites* and *Anopheles* species. Co-occurrence with the latter was less expected since these species do not usually use the same niche. Finally, in the forest, among the four families of Diptera detected, Chironomids were the predominant one, with 73% of the reads, while mosquito species were found in lower abundance (17%). Interestingly, the yellow fever mosquito *Ae. aegypti* was detected in only one sample, inside a bamboo stalk. It used to be very common and widespread on the island, and found equally in both natural and artificial breeding sites (Ribeiro *et al.*, 1998). However, recent on-going mosquito projects and, surveys on the island revealed that *Ae. aegypti* became quite rare and seems to have been replaced by *Ae. albopictus* in lowland and disturbed habitats (Reis *et al.*, 2017; Loiseau *et al.*, 2022). This replacement pattern has been largely documented in Florida, USA (Yang *et al.*, 2021) but is less evident in mainland Central Africa (Simard *et al.*, 2005; Paupy *et al.*, 2010; Kamgang *et al.*, 2013; Tedjou *et al.*, 2019). Nonetheless, our eDNA metabarcoding approach corroborates the actual known distribution of these two *Aedes* species on the island (Loiseau *et al.*, 2022). Finally, the other Culicidae species found in the forest was *Cx. sasai*. It

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

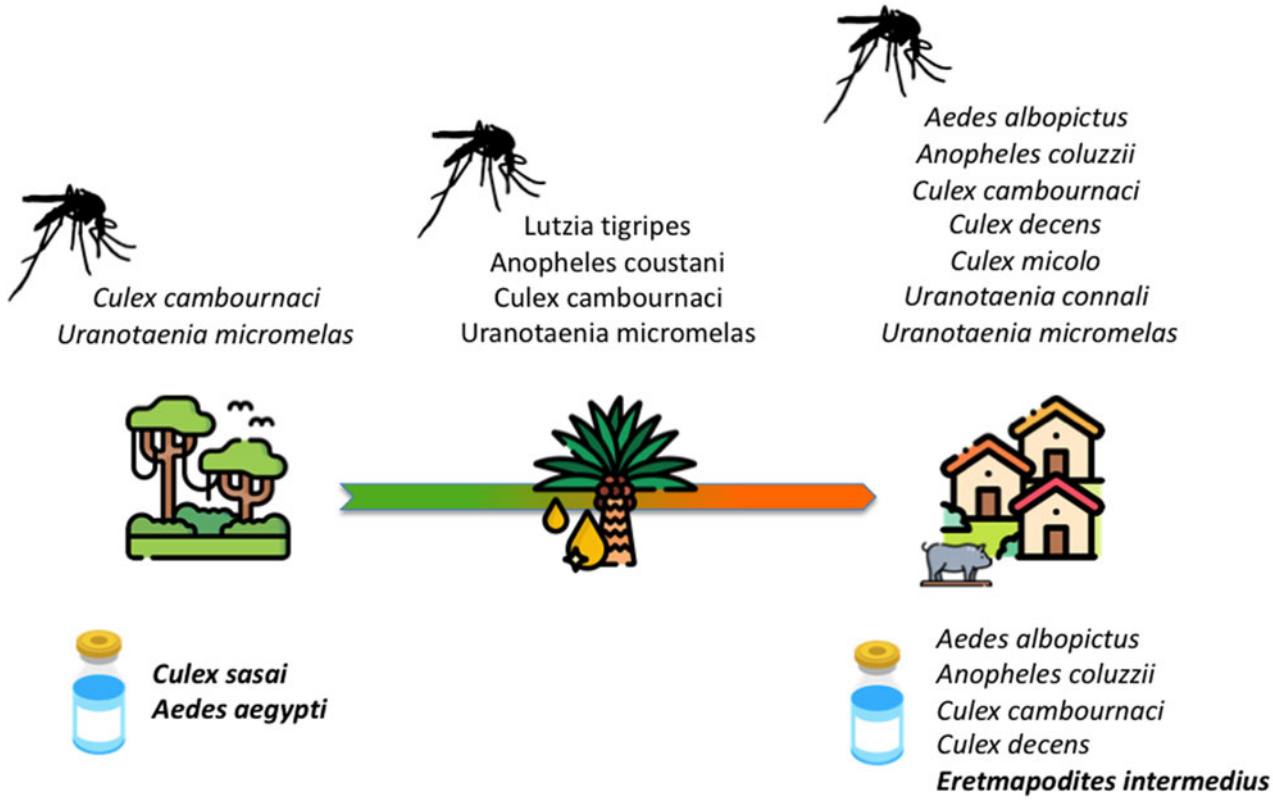


**Figure 4.** List of mosquito species by habitat recovered using both methods: CDC traps (depicted by the adult mosquito) and eDNA metabarcoding (depicted by the water bottle). Species detected uniquely with eDNA metabarcoding are in bold (see also table 2). Icons from Freepik.

is highly unlikely that this mosquito is present on the island, since to date it has been detected only in Asia (Phanitchakun *et al.*, 2017), and is not known to be present on São Tomé Island (Loiseau *et al.*, 2022). Because *Cx. sasai* belongs to the *Culiciomyia* subgenus, we probably detected here a mosquito species belonging to this same subgenus. There are actually four species of this subgenus on São Tomé Island: *Cx. cambournaci*, *Culex nebulosus*, *Culex cinerellus* and *Culex macfiei* (Loiseau *et al.*, 2022), with only two having barcoding sequences on online databases (*Cx. cambournaci* and *Cx. nebulosus*). One could speculate that the species found in this forest sample could be either *Cx. cinerellus* or *Cx. macfiei* and not *Cx. sasai*. This error highlights one of the limitations of the eDNA metabarcoding approach which is discussed below, i.e., incomplete reference databases.

### *Challenges of eDNA metabarcoding: sample quality and taxonomic assignment issues*

As with any new methods, some weaknesses and concerns need to be addressed. Some critical factors for the application of eDNA methods to detect aquatic species have already been reviewed (Goldberg *et al.*, 2016), including contamination in the field and in the laboratory, choosing appropriate sample analysis methods, validating assays or testing for sample inhibition. Here, we highlight concerns that are specific for insect vector monitoring using eDNA approaches.

First, mosquito larvae are mostly found in small and turbid breeding sites or in stagnant water bodies. While water from some larval breeding sites (e.g., rock pools, puddles, artificial containers) is easy to sample, it can be difficult to collect from other sites (e.g., tree holes, plant axils). Traps and sampling procedures, such as aspiration of resting mosquitoes, collection on human or animal bait, allow collecting a greater diversity of species. For inventory purposes, eDNA techniques may need a great water sampling effort in order to be comparable to other techniques (Krol *et al.*, 2019). In addition, sampling small volumes of water can lead to false-negative detection when the density of targeted organisms is low (Ulibarri *et al.*, 2017). Another potential sampling issue is the large amount of soil and humic substances found in breeding sites that may act as PCR inhibitors, increasing the chance to obtain false-negative results (Buxton *et al.*, 2017). In our case, we managed to amplify COI even from dirty samples, although these samples contained many larvae. One study experimentally tested the success of PCR detection of eDNA samples from containers with two different water volumes (50 ml and 1 litre) (Odero *et al.*, 2018). They found that the volume of water required in relation to the density of larvae has an effect on the mosquito detection by eDNA analysis. The detection was better when the samples had many larvae at low densities than few larvae at higher densities (Odero *et al.*, 2018). In addition, the effect of different substrates in the eDNA analysis as well as the preservation methods are parameters that should not be overlooked since metabarcoding analyses require good DNA quality (Ball and Armstrong, 2014).

Second, it seems appealing to evaluate and compare mosquito diversity from different types of samples (water vs. bulk samples) using the metabarcoding approach because traditional dipping methods to survey larvae in breeding sites may not always reflect the adult diversity that can be found with CDC traps (and inversely). In fact, in our survey, only four species were shared

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

between the two techniques (eDNA vs. CDC traps). It is worth noting that some species may be very difficult to detect with traditional trapping because not all insect vector species are equally attracted to dry ice or light (Reisen and Lothrop, 1999). It is especially true for daytime biting mosquitoes. On the other hand, it might be difficult to sample water in breeding sites, such as plant axils or tree holes, which can be high up. More investigations under controlled conditions are needed to compare the efficacy of metabarcoding water samples with trapped adults to characterize insect-vector communities.

Third, in the BOLD, of about 3500 species of Culicidae known globally, barcodes are only available for 1329 species (38%; accessed on 2021-05-25) and, among the 41 known mosquito genera, three genera alone (*Aedes*, *Anopheles* and *Culex*) account for 78% of the occurrences. Similar patterns are found when gathering data on different genes in NCBI (COI, 18S rRNA and 28S rRNA). While *Aedes*, *Culex* and *Anopheles* species account for only 60% of the total mosquito species, 90% of the sequences on average correspond to these three genera (see fig. S3 for illustration of these data). Sequences belonging to unknown taxa are still a common problem in eDNA barcoding and therefore, when starting a new monitoring programme to assess the mosquito diversity in a region or locality, creating a good quality reference database is an indispensable first step. This means that a considerable amount of essential taxonomic work is required to setup eDNA-based monitoring protocols. In this study, we managed to get DNA sequences of four mosquito species that were not deposited in online databases yet. Eleven species out of the 34 known on the island (Loiseau *et al*., 2022) still have to be captured and sequenced to have a full reference database for future research work. Taking all this into account, and considering that certain limitations can be surpassed, then eDNA metabarcoding can have significant advantages for mosquito surveys.

### *Advantages of eDNA metabarcoding: easy sampling and less entomological expertise required*

Sampling for eDNA can be as simple as collecting freshwater samples in tubes and adding preservation buffers (Williams *et al*., 2016), which drastically reduces the cost and time allocated to fieldwork, as well as equipment and resources required for sampling. This is particularly relevant for research projects carried out in remote regions. The effort required for the traditional trapping methods is substantial. Logistically it requires the transport of traps and batteries (which are voluminous and heavy), the availability of freezers (to kill mosquitoes before identification) and of high-quality stereomicroscopes. Once this material is in the field, traps must be set up for several hours, with light that attracts mosquitoes together with a wide range of flying insects, or with traps containing odour products to attract more specifically females (BG-Sentinel or Gravid Mosquito traps). Since light traps are not selective, a great amount of time is spent on sorting all the flying insects from the mosquitoes, separating engorged individuals and labelling individual tubes. Once back in the laboratory, experts may spend a great amount of time at the microscope identifying and dissecting individuals. Identification of mosquito eggs and larvae implies mounting, which is time consuming, and require a specific training. Although an alternative solution could be rearing larvae into adults for unambiguous identification, this is logistically challenging when doing fieldwork in remote places. In addition, for the identification of many adult insect vectors, dissecting male genitalia is required, which is the case for example for most of the species of the African genus *Eretmapodites* (Service, 1990). Molecular identification of eDNA is able to circumvent time-consuming morphological investigation and to detect the presence of species without requiring a strong entomology expertise. The efficacy of eDNA-based surveys will increase as reference databases become more complete. Interestingly, in our study, we detected the species *Er. intermedius* for the first time on the island, as until now *Eretmapodites chrysogater* was the only known representative of this genus on the island (Ribeiro *et al*., 1998). This detection would have been almost impossible using traditional light traps since *Eretmapodites* species are day-biting mosquitoes and males are generally less attracted to them. Finally, the ease of water sampling procedures for eDNA protocols will allow developing large-scale citizen science monitoring programmes and integrating non-specialists in research projects (Biggs *et al*., 2015).

## Concluding remarks

To date, numerous studies have demonstrated that eDNA sampling generally provides greater detection probabilities than traditional techniques (Thomsen *et al*., 2012; McKelvey *et al*., 2016; Valentini *et al*., 2016), but it still remains to be formally demonstrated for mosquito communities. In fact, eDNA methods could surely help in applied medical and veterinary entomology and significantly improve (i) the detection of invasive species and (ii) the evaluation of the composition of mosquito communities in understudied regions. In our study, we showed that CDC light traps and adult identification methods recovered more species than the eDNA metabarcoding per habitat. However, eDNA metabarcoding was able to detect (i) more species at a mosquito breeding site than *de visu* larval identification, and (ii) different species than traditional methods. Therefore, our results highlight the fact that it is best to use in conjunction traditional survey methods and eDNA metabarcoding to enhance detection rates and increase confidence in the monitoring results.

Like any ecological survey tool, eDNA metabarcoding will always suffer biases and uncertainties which have to be taken into account at each step of the study (i.e., fieldwork, labwork, bioinformatics analyses). The building up of the BOLD is required to expand the potential of eDNA metabarcoding, a task where taxonomic expertise will be essential. However, the relative simplicity of field sampling protocols can create opportunities to collect samples using volunteers and even to develop citizen science programmes such as (i) for monitoring and surveillance of invasive species such as *Ae. albopictus*, and (ii) for improving our understanding of ecological systems (competition and predation at breeding sites) that could definitely help in vector control management (Dambach, 2020).

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S0007485323000147.

**Acknowledgements.** We are grateful to the field assistants on São Tomé: Ricardo 'Mito' Fonseca, Martim Veiga and Sidney 'Dulay' Samba, and we thank Arlindo Carvalho, former Director of the Department of the Environment of São Tomé and Príncipe for granting us the permits to conduct the research. We thank Antón Vizcaíno, Ania Pino-Querido and Neus Marí-Mena for their work in the lab and with the bioinformatics analyses. We also thank the two anonymous reviewers for their constructive comments on the manuscript.

**Financial support.** This work is funded by through FCT – Foundation for Science and Technology (Portugal) under the PTDC/BIA-EVL/29390/

2017 DEEP Research Project (C.L.) and via structural funding for CIBIO-InBIO (UIDB/50027/2021). R.G.L. was supported by the project PTDC/BIA-EVL/29390/2017 DEEP Research Project (FCT – Foundation for Science and Technology (Portugal)) and by the grant Juan de la Cierva 2019 Formación from the Ministry of Science and Innovation (REF: FJC2019-041291-I). B.E. was supported via the European Union's Horizon 2020 research and innovation programme under grant agreement No. 668981. C.P., N.R. and D.J. were supported by the French National Research Agency (ANR PRC TIGERBRIDGE, grant No. 16-CE35-0010-01). M.M. was supported via the European Union's Horizon 2020 research and innovation programme under grant agreement No. 854248.

## References

**Andrews S** (2010) FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

**Ball SL and Armstrong KF** (2014) Rapid, one-step DNA extraction for insect pest identification by using DNA barcodes. *Journal of Economic Entomology* **101**, 523–532.

**Bamou R, Mayi MPA, Djiappi-Tchamen B, Nana-Ndjangwo SM, Nchoutpouen E, Cornel AJ, Awono-Ambene FParola P, Tchuinkam T and Antonio-Nkondjio C** (2021) An update on the mosquito fauna and mosquito-borne diseases distribution in Cameroon. *Parasites and Vectors* **14**(1), 1–15.

**Barnes MA, Turner CR, Jerde CL, Renshaw MA, Chadderton WL and Lodge DM** (2014) Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science Technology* **48**, 1819–1827.

**Bartram J, Mountjoy E, Brooks T, Hancock J, Williamson H, Wright G, Moppett J, Goulden N and Hubank M** (2016) Accurate sample assignment in a multiplexed, ultrasensitive, high-throughput sequencing assay for minimal residual disease. *The Journal of Molecular Diagnostics* **18**, 494–506.

**Besansky NJ, Severson DW and Ferdig MT** (2003) DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you. *Trends in Parasitology* **19**, 545–546.

**Biggs J, Ewald N, Valentini A, Gaboriaud C, Dejean T, Griffiths RA, Foster J, Wilkinson JW, Arnell A, Brotherton P, Williams P and Dunn F** (2015) Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). *Biological Conservation* **183**, 19–28.

**Boerlijst SP, Trimbos KB, Van der Beek JG, Dijkstra KDB, Van der Hoorn BB and Schrama M** (2019) Field evaluation of DNA based biodiversity monitoring of Caribbean mosquitoes. *Frontier in Ecology and Evolution* **7**, 240.

**Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA and Gregory Caporaso J** (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**(1), 1–17.

**Bonizzoni M, Gasperi G, Chen X and James AA** (2013) The invasive mosquito species *Aedes albopictus*: current knowledge and future perspectives. *Trends in Parasitology* **29**, 460–468.

**Buxton AS, Groombridge JJ and Griffiths RA** (2017) Is the detection of aquatic environmental DNA influenced by substrate type? *PLoS ONE* **12**, e0183371.

**Caminade C, McIntyre KM and Jones AE** (2019) Impact of recent and future climate change on vector-borne diseases. *Annals of the New York Academy of Science* **1436**, 157–173.

**Chen YA, Lien JC, Tseng LF, Cheng CF, Lin WY, Wang HY and Tsai KH** (2019) Effects of indoor residual spraying and outdoor larval control on *Anopheles coluzzii* from São Tomé and Príncipe, two islands with pre-eliminated malaria. *Malaria Journal* **18**, 405.

**Civade R, Dejean T, Valentini A, Roset N, Raymond JC, Bonin A, *et al.*** (2016) Spatial representativeness of environmental DNA metabarcoding signal for fish biodiversity assessment in a natural freshwater system. *PLoS ONE* **11**, e0157366.

**Clusa L, Miralles L, Basanta A, Escot C and García-Vázquez E** (2017) eDNA for detection of five highly invasive molluscs. A case study in urban rivers from the Iberian Peninsula. *PLoS ONE* **12**, e0188126.

**Coetzee M and Koekemoer LL** (2013) Molecular systematics and insecticide resistance in the major African malaria vector *Anopheles funestus*. *Annual Review in Entomology* **58**, 393–412.

**Collins RA, Wangensteen OS and O'Gorman EJ** (2018) Persistence of environmental DNA in marine systems. *Communications Biology* **1**, 185.

**Dambach P** (2020) The use of aquatic predators for larval control of mosquito disease vectors: opportunities and limitations. *Biological Control* **150**, 104357.

**Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, *et al.*** (2017) Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular Ecology* **26**, 5872–5895.

**Dejean T, Valentini A, Duparc A, Pellier-Cuit S, Pompanon F, Taberlet P and Miaud C** (2011) Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE* **6**, e23398.

**Edgar RC, Haas BJ, Clemente JC, Quince C and Knight R** (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics (Oxford, England)* **27**, 2194–2200.

**Edwards FW** (1941) *Mosquitoes of the Ethiopian Region. III.-Culicine Adults and Pupae*. London, UK: British Museum (Natural History).

**Egeter B, Peixoto S, Brito JC, Jarman S, Puppo P and Velo-Antón G** (2018) Challenges for assessing vertebrate diversity in turbid Saharan water-bodies using environmental DNA. *Genome* **61**, 807–814.

**Epp LS, Boessenkool S, Bellemain EP, Haile J, Esposito A, Riaz T, *et al.*** (2012) New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology* **21**, 1821–1833.

**Eriksson BK and Hillebrand H** (2019) Rapid reorganization of global biodiversity. *Science (New York, N.Y.)* **366**, 308–309.

**Esling P, Lejzerowicz F and Pawlowski J** (2015) Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research* **43**, 2513–2524.

**Ewels P, Magnusson M, Lundin S and Käller M** (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics (Oxford, England)* **32**, 3047–3048.

**Fayle TM, Turner EC, Snaddon JL, Chey VK, Chung AYC, Eggleton P and Foster WA** (2010) Oil palm expansion into rain forest greatly reduces ant biodiversity in canopy, epiphytes and leaf-litter. *Basic and Applied Ecology* **11**, 337–345.

**Ficetola GF, Miaud C, Pompanon F and Taberlet P** (2008) Species detection using environmental DNA from water samples. *Biology Letters* **4**, 423–425.

**Focks DA** (2004) *A Review of Entomological Sampling Methods and Indicators for Dengue Vectors*. Special Programme for Research and training in Tropical Diseases (TDR). 2003. TDR/IDE/DEN/03.1. Geneva: World Health Organization

**Foley DH, Rueda LM and Wilkerson RC** (2007) Insight into global mosquito biogeography from country species records. *Journal of Medical Entomology* **44**, 554–567.

**Geisen S, Laros I, Vizcaíno A, Bonkowski M and De Groot G** (2015) Not all are free-living: high-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa. *Molecular Ecology* **24**, 4556–4569.

**Ghazali A, Asmah S, Syafiq M, Yahya MS, Aziz N, Tan LP, Norhisham AR, Puan CL, Turner EC and Azhar B** (2016) Effects of monoculture and polyculture farming in oil palm smallholdings on terrestrial arthropod diversity. *Journal of Asia-Pacific Entomology* **19**, 415–421.

**Gillies MY and Coetzee M** (1987) *A Supplement to the Anophelinae of Africa South of the Sahara*. Johannesburg, South Africa: The South African Institute for Medical Research.

**Goldberg CS, Pilliod DS, Arkle RS and Waits LP** (2011) Molecular detection of vertebrates in stream water: a demonstration using Rocky Mountain tailed frogs and Idaho giant salamanders. *PLoS ONE*. **6**, e22746.

**Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, *et al.*** (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution* **7**, 1299–1307.

**Guardiola M, Wangensteen OS, Taberlet P, Coissac E, Uriz MJ and Turon X** (2016) Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. *PeerJ* **4**, e2807.

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

**Gutiérrez-López R, Martínez-de la Puente J, Gangoso L, Soriguer RC and Figuerola J** (2015) Comparison of manual and semi-automatic DNA extraction protocols for the barcoding characterization of hematophagous louse flies (Diptera: Hippoboscidae). *Journal of Vector Ecology* **40**, 11–15.

**Hajibabaei M, Singer GAC, Hebert PDN and Hickey DA** (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics, and population genetics. *Trends in Genetics* **23**, 167–172.

**Harbach RE** (2013) Mosquito taxonomic inventory. Available at http://mosquito-taxonomic-inventory info/.

**Hering D, Borja A, Jones JI, Pont D, Boets P, Bouchez A, *et al.*** (2018) Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research* **138**, 192–205.

**Harold AM and Richard JH** (2000) *Invasive species in a changing world*. Washington DC. Covelo, California: Island press.

**Hopkins GHE** (1952) *Mosquitoes of the Ethiopian Region I. – Larval Bionomics of Mosquitoes and Taxonomy of Culicine Larvae*, 2nd Edn. London, UK: British Museum (Natural History).

**Illumina I** (2017) Effects of index misassignment on multiplexing and downstream analysis. Available at https://www.illumina.com.

**Iwamura T, Guzman-Holst A and Murray KA** (2020) Accelerating invasion potential of disease vector *Aedes aegypti* under climate change. *Nature Communication* **11**, 2130.

**Jerde CL, Mahon AR, Chadderton WL and Lodge DM** (2011) Sight-unseen detection of rare aquatic species using environmental DNA. *Conservation Letters* **4**, 150–157.

**Juliano SA** (2009) Species interactions among larval mosquitoes: context dependence across habitat gradients. *Annual Review of Entomology* **54**, 37–56.

**Kamgang B, Ngoagouni C, Manirakiza A, Nakouné E, Paupy C and Kazanji M** (2013) Temporal patterns of abundance of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) and mitochondrial DNA analysis of *Ae. albopictus* in the Central African Republic. *PLoS Neglected Tropical Diseases* **7**, e2590.

**Klymus KE, Marshall NT and Stepien CA** (2017) Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLoS ONE* **12**, e0177643.

**Koh LP and Wilcove DS** (2008) Is oil palm agriculture really destroying tropical biodiversity? *Conservation Letters* **1**, 60–64.

**Krol L, Van der Hoorn B, Gorsich EE, Trimbos K, Bodegom PM and Schrama M** (2019) How does eDNA compare to traditional trapping? Detecting mosquito communities in South-African freshwater ponds. *Frontiers in Ecology and Evolution* **7**, 260.

**Loiseau, C, Gutiérrez-López, R, Mathieu, B, Makanga, BK, Paupy, C, Rahola, N and Cornel, AJ** (2022) Diversity and Distribution of the Arthropod Vectors of the Gulf of Guinea Oceanic Islands. In Ceríaco LMP, de Lima RF, Melo M and Bell RC (eds), *Biodiversity of the Gulf of Guinea Oceanic Islands*. Springer, Cham. pp. 383–405.

**Lopes CM, Sasso T, Valentini A, Dejean T, Martins M, Zamudio KR and Haddad CF** (2017) eDNA metabarcoding: a promising method for anuran surveys in highly diverse tropical forests. *Molecular Ecology Resources* **17**, 904–914.

**Magoč T and Salzberg SL** (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics (Oxford, England)* **27**, 2957–2963.

**Martin M** (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* **17**, 10–12.

**Matesanz S, Pescador DS, Pías B, Sánchez AM, Chacón-Labella J, Illuminati A, de la Cruz M, López-Angulo J, Marí-Mena N, Vizcaíno A, *et al.*** (2019) Estimating belowground plant abundance with DNA metabarcoding. *Molecular Ecology Resources* **19**, 1265–1277.

**McKelvey KS, Young MK, Knotek WL, Carim KJ, Wilcox TM, Padgett-Stewart TM and Schwartz MK** (2016) Sampling large geographic areas for rare species using environmental DNA: a study of bull trout *Salvelinus confluentus* occupancy in western Montana. *Journal of Fish Biology* **88**, 1215–1222.

**Minamoto T, Yamanaka H, Takahara T, Honjo MN and Kawabata ZI** (2012) Surveillance of fish species composition using environmental DNA. *Limnology* **13**, 193–197.

**Mychek-Londer JG, Balasingham KD and Heath DD** (2020) Using environmental DNA metabarcoding to map invasive and native invertebrates in two Great Lakes tributaries. *Environmental DNA* **2**(3), 283–297.

**Odero J, Gomes B, Fillinger U and Weetman D** (2018) Detection and quantification of *Anopheles gambiae* sensu lato mosquito larvae in experimental aquatic habitats using environmental DNA (eDNA). *Wellcome Open Research* **3**, 26.

**Paupy C, Ollomo B, Kamgang B, Moutailler S, Rousset D, Demanou M, Hervé JP, Leroy E and Simard F** (2010) Comparative role of *Aedes albopictus* and *Aedes aegypti* in the emergence of dengue and Chikungunya in Central Africa. *Vector Borne and Zoonotic Diseases* **10**, 259–266.

**Pedersen EM, Stolk W, Laney S and Michael E** (2009) The role of monitoring mosquito infection in the global programme to eliminate lymphatic filariasis. *Trends in Parasitology* **25**, 319–327.

**Peixoto S, Chaves C, Velo-Antón G, Beja P and Egeter B** (2021) Species detection from aquatic eDNA: assessing the importance of capture methods. *Environmental DNA* **3**, 435–448.

**Phanitchakun T, Wilai P, Saingamsook J, Namgay R, Drukpa T, Tsuda Y, Walton C, Harbach RE and Somboon P** (2017) *Culex* (*Culiciomyia*) *sasai* (Diptera: Culicidae), senior synonym of *Cx. spiculothorax* and a new country record for Bhutan. *Acta Tropica* **171**, 194–198.

**Piaggio AJ, Engeman RM, Hopken MW, Humphrey JS, Keacher KL, Bruce WE and Avery ML** (2014) Detecting an elusive invasive species: a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular Ecology Research* **14**, 374–380.

**Pilliod DS, Goldberg CS, Arkle RS and Waits LP** (2014) Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Research* **14**, 109–116.

**Ratnasingham S and Hebert PD** (2007) BOLD: the barcode of life data system (http://www.barcodinglife.org). *Molecular Ecology Notes* **7**, 355–364.

**Reis S, Cornel AJ, Melo M, Pereira H and Loiseau C** (2017) First record of *Aedes albopictus* (Skuse 1894) on São Tomé Island. *Acta Tropica* **171**, 86–89.

**Reisen WK and Lothrop HD** (1999) Effects of sampling design on the estimation of adult mosquito abundance. *Journal of the American Mosquito Control Association* **15**, 105–114.

**Ribeiro H, Da Cunha Ramos E, Capela R and Alves Pires C** (1998) Os mosquitos (Diptera: Culicidae) da Ilha de São Tomé. *Garcia de Orta Serie de Zoologia* **22**, 1–20.

**Robeson MS II, Rourke DR, Kaehler BD, Ziemski M, Dillon MR, Foster JT and Bokulich NA** (2021) RESCRIPt: reproducible sequence taxonomy reference database management. *PLoS Computational Biology* **17**, e1009581.

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

**Ruppert KM, Kline RJ and Rahman MS** (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation* **17**, e00547.

**Schneider J, Valentini A, Dejean T, Montarsi F, Taberlet P, Glaizot O and Fumagalli L** (2016) Detection of invasive mosquito vectors using environmental DNA (eDNA) from water samples. *PLoS ONE* **11**, e0162493.

**Segan DB, Murray KA and Watson JEM** (2016) A global assessment of current and future biodiversity vulnerability to habitat loss–climate change interactions. *Global Ecology and Conservation* **5**, 12–21.

**Service MW** (1990) *Handbook to the Afrotropical Toxorhynchitine and Culicine Mosquitoes, Excepting Aedes and Culex*. London: British Museum (Natural History), pp. 1–207.

**Seymour M, Durance I, Cosby BJ, Ransom-Jones E, Deiner K, Ormerod SJ, *et al.*** (2018) Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. *Communication Biology* **1**, 4.

**Simard F, Nchoutpouen E, Toto JC and Fontenille D** (2005) Geographic distribution and breeding site preference of *Aedes albopictus* and *Aedes aegypti* (Diptera: Culicidae) in Cameroon, Central Africa. *Journal of Medical Entomology* **42**, 726–731.

**Smith KM, Machalaba CC, Seifman R, Feferholtz Y and Karesh WB** (2019) Infectious disease and economics: the case for considering multi-sectoral impacts. *One Health (Amsterdam, Netherlands)* **7**, 100080.

**Spear SF, Groves JD, Williams LA and Waits LP** (2015) Using environmental DNA methods to improve detectability in a hellbender (*Cryptobranchus alleganiensis*) monitoring program. *Biological Conservation* **183**, 38–45.

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press

**Swei A, Couper LI, Coffey l, Kapan D and Bennett S** (2020) Patterns, drivers, and challenges of vector-borne disease emergence. *Vector Borne and Zoonotic Diseases* **20**, 159–170.

**Taberlet P, Prud'Homme SM, Campione E, Roy J, Miquel C, Shehzad W, *et al.*** (2012) Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology* **21**, 1816–1820.

**Tedjou AN, Kamgang B, Yougang AP, Njiokou F and Wondji CS** (2019) Update on the geographical distribution and prevalence of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae), two major arbovirus vectors in Cameroon. *PLoS Neglected Tropical Diseases* **13**, e0007137.

**Thomas AC, Deagle BE, Eveson JP, Harsch CH and Trites AW** (2016) Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. *Molecular Ecology Resources* **16**, 714–726.

**Thomsen PF, Kielgast J, Iversen LL, Wiuf C, Rasmussen M, Gilbert MTP, Orlando L and Willerslev E** (2012) Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology* **21**, 2565–2573.

**Turner E and Foster W** (2009) The impact of forest conversion to oil palm on arthropod abundance and biomass in Sabah, Malaysia. *Journal of Tropical Ecology* **25**, 23–30.

**Ulibarri RM, Bonar SA, Rees C, Amberg J, Ladell B and Jackson C** (2017) Comparing efficiency of American Fisheries Society standard snorkeling techniques to environmental DNA sampling techniques. *North American Journal of Fisheries Management* **37**, 644–651.

**Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, *et al.*** (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology* **25**, 929–942.

**Valiere N and Taberlet P** (2000) Urine collected in the field as a source of DNA for species and individual identification. *Molecular Ecology* **9**, 2150–2152.

**Williams KE, Huyvaert KP and Piaggio AJ** (2016) No filters, no fridges: a method for preservation of water samples for eDNA analysis. *BMC Research Notes* **9**, 298.

**World Health Organization** (2020) World malaria report 2020: 20 years of global progress and challenges.

**Yang B, Borgert BA, Alto BW, Boohene CK, Brew J, Deutsch K, *et al.*** (2021) Modelling distributions of *Aedes aegypti* and *Aedes albopictus* using climate, host density and interspecies competition. *PLoS Neglected Tropical Diseases* **15**, e0009063.

**Young KI, Buenemann M, Vasilakis N, Perera D and Hanley KA** (2021) Shifts in mosquito diversity and abundance along a gradient from oil palm plantations to conterminous forests in Borneo. *Ecosphere (Washington, D.C.)* **12**, e03463.

https://doi.org/10.1017/S0007485323000147 Published online by Cambridge University Press