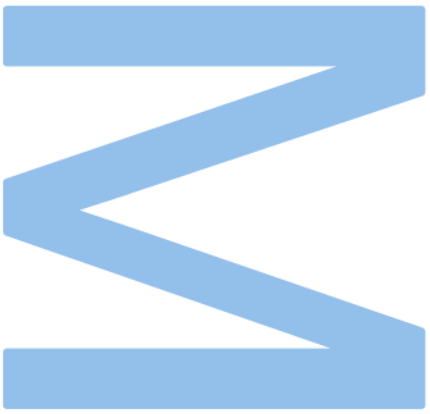
Rapid bioassessment of terrestrial vertebrate diversity using Tsetse flies' diet metabarcoding: a case study in Corubal river gallery forest in Guinea Bissau

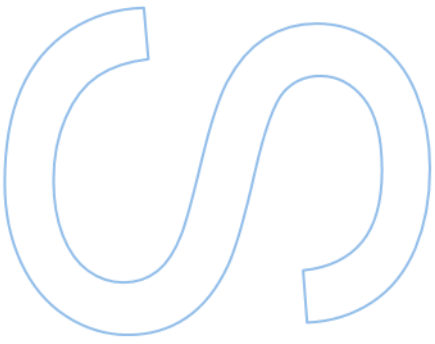
Vasco Rafael da Silva Fernandes
Mestrado em Biodiversidade, Genética e Evolução
Departamento de Biologia
2022

Supervisor
Dr. Manuel Lopes Lima, Researcher, CIBIO-InBIO/BIOPOLIS

Co-supervisor
Dr. Vanessa Mata, Researcher, CIBIO-InBIO/BIOPOLIS
Dr. Sónia Ferreira, Researcher, CIBIO-InBIO/BIOPOLIS


U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO





Sworn Statement

I, Vasco Rafael da Silva Fernandes, enrolled in the Master Degree Biodiversidade Genética e Evolução at the Faculty of Sciences of the University of Porto hereby declare, in accordance with the provisions of paragraph a) of Article 14 of the Code of Ethical Conduct of the University of Porto, that the content of this dissertation reflects perspectives, research work and my own interpretations at the time of its submission.

By submitting this dissertation, I also declare that it contains the results of my own research work and contributions that have not been previously submitted to this or any other institution.

I further declare that all references to other authors fully comply with the rules of attribution and are referenced in the text by citation and identified in the bibliographic references section. This dissertation does not include any content whose reproduction is protected by copyright laws.

I am aware that the practice of plagiarism and self-plagiarism constitute a form of academic offense.

Vasco Fernandes

31/10/2022

Acknowledgements

Gostava de expressar a minha gratidão para com:

O meu orientador Dr. Manuel Lopes Lima e ás minhas co-orientadoras Dra. Vanessa Mata e Dra. Sónia Ferreira pela grande paciência, dedicação e aprendizagem ao longo destes últimos meses.

A toda a minha família, amigos e colegas por estes anos de faculdade.

A toda a equipa que realizou o trabalho de campo, por toda a experiência partilhada.

A todos os membros do CTM do CiBio, por toda a ajuda no trabalho de laboratório.

Ao TROPIBIO por financiar este projeto. "This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 854248".

Resumo

África abriga um quarto da biodiversidade global, incluindo os maiores conjuntos de grandes mamíferos da Terra. No entanto, o continente está a sofrer grandes níveis de crescimento populacional, de urbanização e desenvolvimento agrícola, que estão a causar impactos crescentes sobre a vida selvagem e os respetivos habitats naturais. As bioavaliações são uma ferramenta importante para avaliar o estado de conservação de espécies e ecossistemas. No entanto, é um desafio realizar essas avaliações em áreas remotas e de difícil acesso. Por esta razão, a Guiné-Bissau e o rio Corubal são áreas pouco estudadas com supostos relatos de alta biodiversidade. Portanto, é muito importante recolher dados dessas regiões para orientar futuros esforços de conservação.

Ferramentas moleculares têm sido cada vez mais utilizadas para auxiliar bioavaliações rápidas, principalmente usando fontes ambientais de DNA, como água, solo ou mesmo ar. O DNA metabarcoding tem sido utilizado com sucesso noutras regiões para estudar dietas de moscas tsé-tsé (*Glossina*) para deteção de espécies.

Neste estudo, moscas tsé-tsé foram apanhadas por armadilhas em 8 locais ao longo da bacia do Corubal e o seu conteúdo intestinal foi analisado através de DNA metabarcoding para testar o seu uso numa bioavaliação rápida de vertebrados terrestres. 116 moscas (*Glossina palpalis* e *Glossina morsitans*) foram apanhadas e 90 amostras de 4 dos 8 locais foram analisadas. Os taxa mais detetados foram *Homo sapiens* e *Varanus niloticus*. 25 outros taxa foram também detetados, mas com menor frequência.

Os resultados mostram que o DNA metabarcoding de moscas tsé-tsé pode ser uma ferramenta útil na deteção de alguns grupos de espécies na região de Corubal na Guiné-Bissau, mas o protocolo utilizado ainda necessita de otimização para ser utilizado com maior rigor em estudos futuros de bioavaliação.

Palavras-chave: *Glossina*, Corubal, Guinea Bissau, Rapid bioassessment, DNA metabarcoding, Biodiversity, iDNA, *Varanus niloticus*

Abstract

Africa hosts a quarter of the global biodiversity including the earth's largest intact assemblages of large mammals. However, the continent is facing exceptional rates of population growth, urbanisation, and agricultural development, which are bringing increasing impacts on wildlife and their natural habitats.

Bioassessments are an important tool to evaluate the conservation status of species and ecosystems. However, it is challenging to perform these assessments in remote areas that are difficult to access. For this reason, Guinea Bissau and the Corubal river are poorly studied areas with anecdotal reports of high biodiversity. It is therefore very important to collect data from these regions to guide future conservation efforts.

Molecular tools have been increasingly used to assist in rapid bioassessments, especially using environmental DNA sources such as water, soil, or even air. The use of DNA metabarcoding has been successfully used in other regions to study tsetse flies (*Glossina*) diets for species detections.

In this study, tsetse flies were collected from traps in 8 sites along the Corubal basin and their gut content was analysed through DNA metabarcoding to test its use in a rapid bioassessment of terrestrial vertebrates. 116 flies (*Glossina palpalis* and *Glossina morsitans*) were collected and 90 samples from 4 of the 8 sites were analysed. The most detected taxa were *Homo sapiens* and *Varanus niloticus*. 25 other taxa were also detected but less frequently.

The results show that DNA metabarcoding of tsetse flies can be a useful tool in the detection of some species groups in the Corubal region in Guinea-Bissau, but the protocol used still needs optimization to be used more successfully in future bioassessment studies.

Keywords: *Glossina*, Corubal, Guinea Bissau, Rapid bioassessment, DNA metabarcoding, Biodiversity, iDNA, *Varanus niloticus*

Table of Contents

List of Tables

List of Figures

List of Abbreviations

FCUP	FACULTY OF SCIENCES OF THE UNIVERSITY OF PORTO
UP	UNIVERSITY OF PORTO
DNA	DEOXYRIBONUCLEIC ACID
IUCN	INTERNATIONAL UNION FOR CONSERVATION OF NATURE
NCBI	NATIONAL CENTER OF BIOTECHNOLOGY INFORMATION
NGS	NEXT GENERATION SEQUENCING
OTU	OPERATIONAL TAXONOMIC UNIT
PCR	POLYMERASE CHAIN REACTION
MTDNA	MITOCHONDRIAL DNA
RDNA	RIBOSOMAL DNA
IDNA	INVERTEBRATE-DERIVED DNA
HAT	HUMAN AFRICAN TRYPANOSOMIASIS
AAT	AFRICAN ANIMAL TRYPANOSOMIAISIS
SSA	SUB-SAHARAN AFRICA
DBT COMPLEX	DULOMBI-BOÉ-XECHÉ COMPLEX

Introduction

Biodiversity

The world's natural forests are facing a fast decline, which is concerning because these important ecosystems support wildlife and human populations. With the exponential growth of human populations, the biodiversity around the globe is strongly declining at an alarming rate and the threats to wildlife conservation are getting more serious and broader (Brooks et al. 2006). Destruction and degradation of natural habitats are widespread and profound, leading the remaining natural areas into progressively smaller and more isolated fragments (Saunders et al. 1991). Humans´ intensive demand for land and resources is predicted to increase in the future, consequently decreasing the future availability of habitats and other resources for other animals (Lee, 2006).

Africa is known to host high biodiversity with many endemic and endangered charismatic species, especially the mammal megafauna, such as lions, elephants, hippos, and antelopes. However, its richness and diversity are decreasing rapidly, mainly due to direct anthropogenic threats, such as transformation and fragmentation of wild habitats, overexploitation of important resources, as well as other indirect impacts such as climate change (Bellard et al. 2012).

The knowledge about the biodiversity of an area relies primarily on baseline surveys or bioassessments (Torres et al. 2010). Bioassessments are important tools not only to evaluate the conservation status of species and ecosystems, but also for ecological studies, for instance, giving insights about the dynamics of community structure and provide an ecological framework to search for functional links, environmental variation, and ecosystem diversity (Willerslev, et al. 2014).

In remote areas, assessing and surveying biodiversity is difficult and expensive due to the isolation and hard-working conditions (Strange et al. 2007). For this reason, in many regions of Africa there is a lack of basic information about species diversity and population abundances (Álvares & Brito, 2006), highlighting the urgent need for comprehensive bioassessments in these areas. However, due to limitations in time and human resources, as well as funds availability, it is practically unfeasible to survey all species in every study area. Therefore, it is important to apply surveying methods that are fast and inexpensive and can be standardised across a wide range of species and habitats. Due to this, multiple rapid bioassessment protocols have been developed over

the last decade, targeting various biological groups and habitats (e.g. Larsen, 2016). However, these protocols are not intended to provide comprehensive data on rare and difficult to find species, but to indicate a general idea of the communities (Hammond, 1994, Gaston, 1995).

Rapid bioassessment protocols have been developed based on multiple surveying methodologies, e.g. acoustic profilers and nets for bats and birds, camera trapping for mammals, electrofishing and angling for fish, see examples in Larsen (2016). Over the last two decades molecular tools have also been developed to assist bioassessments. The collection of non-invasive genetic samples, such as hairs (Massey et al. 2019), faeces (Deagle et al. 2010) and feathers (Clare et al. 2011) or DNA from the environment, also known as eDNA, have been increasingly used not only for species detections, but also to obtain inventories of species using DNA metabarcoding methodologies.

DNA Metabarcoding

DNA metabarcoding consists in the amplification, sequencing, and identification of a target gene allowing for the simultaneous identification of multiple taxa in a single sample. It is a cost-effective approach used for species identification studies (Taberlet et al. 2012), for distinct aims, for example, to obtain present and past species inventories from eDNA of distinct sources, including air, soil and water samples (Haile et al. 2009), and in diet analyses by examining gut or faecal content (Deagle et al. 2010). The main purpose of the DNA metabarcoding is to characterise communities, although relative abundance of species can sometimes be measured with sequence frequency (Aizpurua et al., 2018)(Aizpurua et al. 2018; Deagle et al. 2018b).

Over the last decades, DNA metabarcoding tools have been replacing traditional barcoding methods because it is able to identify multiple species from a mixed sample of DNA based on high-throughput sequencing (HTS) of a specific DNA marker, instead of the conventional DNA barcoding that identifies individual specimens only. Metabarcoding can be very handy in large ecological studies, relieving its costs and improving the identification accuracy in species with similar morphologies (Chariton et al. 2010). The recent boost in the number of HTS machines and decrease in sequencing costs, led to the exponential growth of HTS projects, including metabarcoding studies (Binladen et al. 2007; Valentini et al. 2009).

To obtain successful results with this technique, it is important to follow protocols carefully and understand its limitations. A high DNA quality is crucial to get reliable

results. For that, to ensure DNA integrity and preservation and avoid external contamination, it is very important to carefully choose the sample collection, handling, and transport protocols from the field to the laboratory. Biases in DNA metabarcoding in species detection can also occur during DNA extraction and PCR amplification or even during the bioinformatics phase. When handling the necessary equipment, either in the field or in the laboratory it is important to avoid contamination. To prevent contamination, immersing the equipment in 10% bleach for at least 10 min should be enough for DNA decontamination (Kemp & Smith, 2005). The transportation and storage of DNA metabarcoding sources for long periods may be a risk for the preservation of the DNA and should be avoided. The protocols to prevent DNA degradation, including e.g. preservation in ethanol, RNAlater or silica-gel, must also be adapted depending on the type of DNA source. During the bioinformatic steps, the main limitations are potential primer biases and incomplete barcoding reference libraries. Primer biases may occur due to the preferential affinity of the selected primers to DNA templates of specific species or groups, leading to a biased amplification of these taxa. The lack of complete barcoding reference libraries may bring difficulties in the identification of taxa to the species level, with identifications being many times made only at family or genus level instead (Clausen et al. 1998). These limitations highlight the need to use several distinct primers in metabarcoding studies and the urgent need of published and available genomic resources across a wide range of taxa. The selection of the DNA marker also plays an important role in metabarcoding studies (Deagle et al. 2014). First, it is important to search for the availability of target taxa in publicly available DNA databases before choosing the right markers for the study. Then, the ideal marker should have some features, the ability to successfully amplify the target groups, but not the non-target ones. For that, it should have sufficiently conserved flanking primer-binding sites only to target species. Marker size is also important, it should be long enough for better taxonomic resolution, but not too extensive because it may decrease the success of amplification, since it must often bind to fragmented DNA (Mingxin Liu et al. 2020).

In studies working with closely related vertebrates, it appears that the mitochondrial DNA (mtDNA) is a better choice, since it contains a high proportion of evolutionary-linked nucleotide substitutions, improving its level of discrimination (Kocher et al. 1989).

The 16S and 12S rRNA genes are relatively conserved mitochondrial genes, generally evolving more slowly than the whole mitogenome (Di Finizio et al. 2007). Both genes contain highly conserved and variable regions, providing a good framework for primer development. However, 16S and 12S rRNA sequences are only available

comprehensively in charismatic groups such as large mammals, but not in neglected and poorly studied taxa, such as insects, bats and molluscs, not allowing for great accuracy in species identifications using these primers (Linacre & Tobe, 2011).

iDNA

Free DNA molecules are always present in the environment, they can occur in faeces, urine, blood, skin, or body fluids of the organisms and can be collected from other sources such as water and soil samples (Taberlet, et al. 2012). The use of invertebrates in fauna monitoring and detection has a high potential, since many of them are widespread and abundant, and can occur in many habitats around the world (Siddig, et al. 2016). Also, numerous invertebrate species feed on vertebrates, allowing storage of their DNA inside their bodies, so it is possible in DNA metabarcoding studies to apply invertebrate-derived DNA (iDNA) to monitor and identify vertebrate species (Calvignac-Spencer et al. 2013). The analysis of the invertebrates gut to extract the host's DNA is a promising tool in DNA metabarcoding regarding biodiversity monitoring and can be extracted from many invertebrate taxa, such as leeches (Drinkwater et al. 2021), mosquitoes (Kocher, et al. 2017), ticks (Gariepy et al. 2012), marine copepods (Meekan et al. 2017), sand flies (Kocher, et al. 2017), blow and flesh flies (Calvignac-Spencer et al. 2013; Hoffmann et al. 2018), and also from tsetse flies (Muturi et al. 2011). The analysis of the intestinal tract of invertebrate taxa to detect vertebrate DNA has been growing in wildlife detection as a very useful tool (Calvignac-Spencer et al. 2013). Invertebrate-derived DNA (iDNA) is an offshoot of eDNA (Schnell et al. 2012).

Tsetse flies are terrestrial invertebrates and in most of the studies it is used >95% ethanol as a preservative, although there are other alternatives like food-grade propylene glycol (Ferro & Park, 2013; Patrick et al. 2016). Gut contents analysis of tsetse flies has been applied in metabarcoding and monitoring studies (Muturi et al. 2011).

Tsetse flies

Tsetse flies are bloodsucking flies from the genus *Glossina*, occurring in sub-Saharan tropical Africa and are the vectors that transmit unicellular protozoa of genus *Trypanosoma* that cause both human and animal African Trypanosomoses, better known as sleeping sickness or Human African Trypanosomiasis (HAT) for humans, and as Nagana or African Animal Trypanosomiasis (AAT) for livestock (Simarro et al. 2008).

These animals are responsible for the transmission of the fatal human disease ("Sleeping sickness"), responsible for 50,000 to 70,000 deaths annually (Chappuis et al. 2010). However, due to recent improvements in case detection, treatment, and vector management, the number of cases decreased around 90% (Franco, 2014).

Tsetse flies and their ability to spread diseases through infected hosts have further impact in pastoral farming, especially in humans and domestic animals that are not trypanotorelant. The areas that are highly infested by tsetse flies are usually deprived of milk and meat because of the high number of cattle casualties caused by Trypanosomoses (Jordan, 1985), therefore, limiting pasturage regions to uninfected areas (Malele et al. 2011).

Glossina species are distributed exclusively in Africa (Atouguia, 1999). Their habitat preference is tropical humid forest associated with water sources, such as lakes, lagoons, rivers and swamps. They are considered hygrophilous species, not because of their direct need for water, but for the conditions these moist habitats provide, including humidity and shady places from the dense vegetation that are favourable for larvae deposition. The highest population density of *Glossina* is in West and Central Africa (Despommier et al. 2005). *Glossina* may be divided in three main species groups, i.e. the palpalis, morsitans and fusca groups (Despommier et al. 2005). The glossinas from the palpalis group, the focus of this study, are distributed in high humidity areas of West and Central Africa (Odete & Afonso, 2012), inhabiting transition zones from forest to herbaceous savanna associated with different water bodies, ranging from rivers, lakes, lagoons to swamps and mangroves. They are also found in peri-urban settlements by coffee, cocoa and cashew plantations (Buxton, 1955; Odete & Afonso, 2012). Their diet is mostly based on Primates (38%), Bovides (22%) and Reptiles (28%), and more specifically on Men, Bushbuck and monitor lizards, species that co-occur with tsetse flies around riverine habitats (Weitz, 1963).

The life cycle of these flies is quite unique (Figure 1), since they reproduce by adrenotropic viviparity (Aksoy & Rio, 2005). The female tsetse, instead of laying eggs, produces a larva which is developed inside the uterus for a period of roughly 10 days and posteriorly deposited on moist soil under shaded conditions. The larva then turns into a pupa and 22-60 days later, a fly emerges (Rozendaal, 1997).

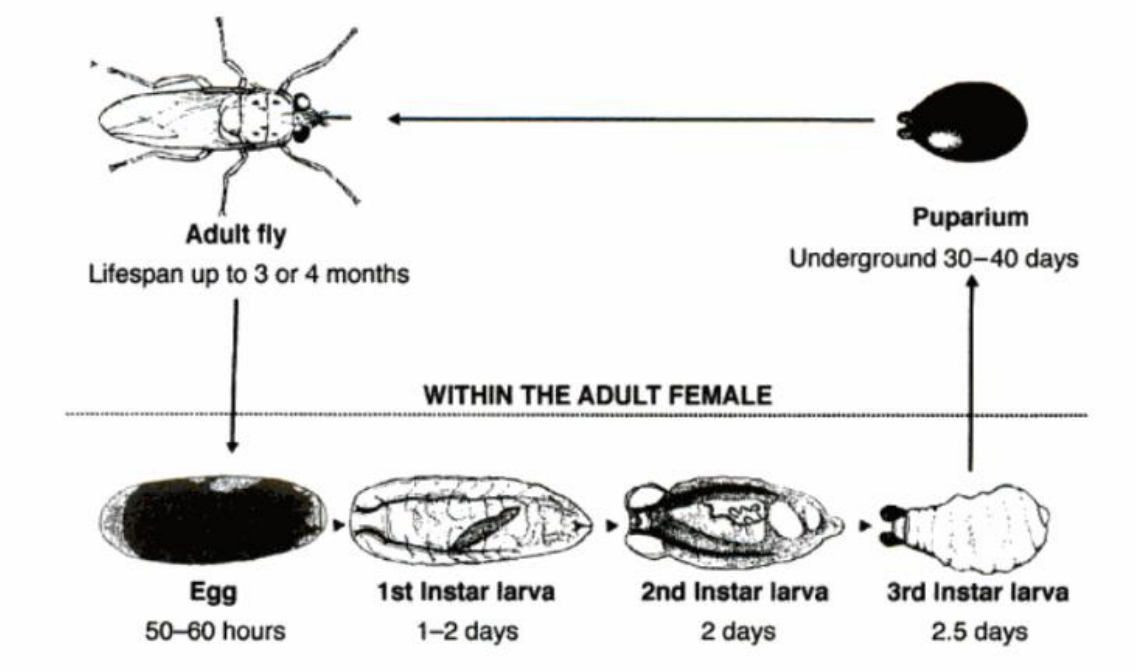


Figure 1 - Life cycle of a Tsetse Fly Leak, S.G.A. (1998)

Despite species differences in food source preferences, all tsetse flies rely on blood meals, usually from animals rather than humans. Their bloodsucking ability is achieved through their anatomy. Their head is divided in three parts; the labium, which is the stinging organ equipped with bilobed strong denticles in the distal end, that allows the bite, the labrum is the organ responsible for suction and transportation of the blood to the pharynx; and the hypopharynx, which is a tubular structure that takes out the saliva during the bite (Hoare, 1972).

Study area

Guinea-Bissau is one of the smallest countries in western Africa, located between 10º 59' - 12º 20' N and 13º 40' - 16º 43' W, with a total area of 36.125 km² at the coast of the Eastern Atlantic Ocean. This country shares borders with Senegal and the Republic of Guinea, including an archipelago, the Bijagós islands. After the war for independence against Portugal, ending in 1973, it has been developing slowly, mainly due to political and military instability.

The climate in Guinea-Bissau has two marked seasons with the rainy season lasting from June to October and the dry season from November to May. The southwest and the northeast of the country are very distinct in terms of annual rainfall, varying between 2,400-2,600 mm to 1,200-1,400 mm respectively. On the contrary, the annual temperature is more stable across the country, having a constant range between 25.9ºC

and 27.1ºC, with a mean temperature of 26.5ºC (Catarino et al. 2001). There are nearly seven months with little or no rainfall in which water resources for wildlife are heavily restricted (Mendes, 2018). The remaining water sources are presumably critical to the survival and persistence of many water-dependent species, among which the forest elephant.

Guinea-Bissau has 46% of the total area (including the protected coastal waters) designated as protected areas (IUCN/WCPA, 2007). Nevertheless, this protection has not fully prevented the degradation of habitats, caused by the increasing human encroachment, and the consequent expansion of cashew orchards, and forest degradation, that modified the savanna and forest into agricultural areas (White, 1993). The consequences associated with unsustainable land clearing and also expansion of cashew plantations (Anacardium occidentale L.) are accumulating, since these practices are increasing and becoming the main export of Guinea-Bissau, consequently occupying more land and leading to deforestations. This increase in crop production is related to the fact that this practice is key for Sub-Saharan Africa (SSA). Human–wildlife conflicts are rising across Africa (Hill, 1997) and the incomes obtained through these practices are fundamental to fight poverty and to establish economic growth, as well as for food security to SSA, so it is challenging to reduce economic and poverty problems, while maintaining these agricultural activities and finding sustainable solutions at the same time (Robbins & van der Straeten, 1996).

Additionally, overexploitation by unsustainable hunting is provoking a pantropical conservation crisis and especially in Guinea-Bissau (Milner-Gullanda et al. 2002; Macdonald et al. 2012). Bushmeat is considered one of the most valuable tropical forest products after timber (Wilkie & Carpenter, 1999). Therefore, its practice has contributed to primates and other large/medium mammals' populations to decline severely (Alliance, 2006).

Guinea Bissau lies at the transitional zone between the rainforests of the Guinean Forest Block and the Sudanian savannas. Dry, open and closed forests are the dominant land cover, small areas of sub-humid primary tropical forest also persist (White, 1993) in the southwest (Tombali and Quínara regions). Further to the north with rainfall below 600 mm, the vegetation cover diminishes to bush savanna or scrub steppe, which is a transitional between the Sudanian savanna and the Sahara Desert, therefore this zone often has floristical features of the Sudanian savanna, interspersed with the fire-resistant trees typical of the Guinean savanna. The Boé-Dulombi region, where the study was performed, including the Boé and Dulombi National Parks and related wildlife corridors

of Xeché, Salifo, and Cuntabani (DBT Complex), is located at the northwestern distributions limit for several species of conservation concern in West Africa. This region holds a rich mammal community, including large-bodied species such as the West African chimpanzee (*Pan troglodytes verus*) and small communities of forest elephants (*Loxodonta cyclotis*) and lions, which are among the most threatened large mammal species in West Africa (Sarmiento & Butynski 1996). Other large mammals are also reported from this region including, leopards (*Panthera pardus*), buffalos (*Syncerus caffer*), Roan antelopes (*Hippotragus equinus*), warthogs (*Phacochoerus africanus*), bush pigs (*Potamochoerus larvatus*) and duikers (tribe Cepholophinae, most of which are also endangered throughout West Africa and poorly known in terms of distribution and ecological partition. The fauna of Guinea-Bissau is still poorly studied, despite being extremely rich. There are over 1,000 species of vertebrates surveyed (Dodman et al. 2004), and much remain to be discovered. Because they are indigenous species, it is very important to ensure their protection. The Corubal River lies between 11◦ N and 12◦30′ N and between 12◦ W and 14◦30′ W and is a West African transboundary Basin, shared between Guinea and Guinea-Bissau, with a coverage area of 20 876.4 km2. Its origin begins in the Fouta Djallon highlands in Middle Guinea, entering in Guinea-Bissau through its eastern border and flowing West until emptying into the Geba estuary. Along the way, it passes the towns of Xeché, Xitole upstream from Bissau, with two main tributaries in Guinea Bissau, the Fefine and the Seli rivers. In the Corubal River, temporary lagoons surrounded by small vegetation occur in the Eastern part of the country. These water bodies are also known as Wendus and are connected to the tributaries of the Corubal, depending on the annual season, especially during reduced pluviosity periods. The vegetation around the Wendus is composed of dense forests, dry forest covered annually by bush fires, wooded savanna and woodland, crops, fallow fields, etc. The importance of the Corubal for biodiversity connectivity is provided by the ecological corridors formed, connecting two of the largest Guinea-Bissau National Parks (Fernandes, 2019).

The Corubal also plays an important economic and social role for the human communities (Alves, 2007), providing the necessary sources for agriculture, fishing, cattle raising and for the villagers themselves. The intensive growth of the population and its needs has demanded a lot of the catchment's natural and water assets capacities, that could be exceeded by climate change and other events, such as the construction of dams (Fernandes, 2019).

Objectives

The main goal of this study is to develop a protocol for testing the use of DNA metabarcoding from tsetse flies' diet in rapid bioassessments in the Corubal region of Guinea Bissau. The study also aimed to detect limitations and flaws on the designed protocol from, flies capture, sample storage and transport, to DNA extraction and amplification, primers choice and analyses, in order to help improving the success of future studies using this methodology.

Methods and materials

Fieldwork

In a field expedition from February 13th to February 22nd by the Corubal river and associated freshwater habitats, two tsetse traps were placed in eight locations in the Dulombi-Boé-Xeché (DBT) complex (Figure 2) to collect tsetse flies for a metabarcoding analysis of the terrestrial vertebrate biodiversity in this area. Five stations were set along the main channel of the Corubal River, i.e. Saltinho (A), Cussilinta (B), Xitole (C), Suto Manca (D), and Xeché (H), two in tributaries, i.e. the rivers Quinhiqué (E) and Fefiné (F), and one in the Wendu Tcham (G).

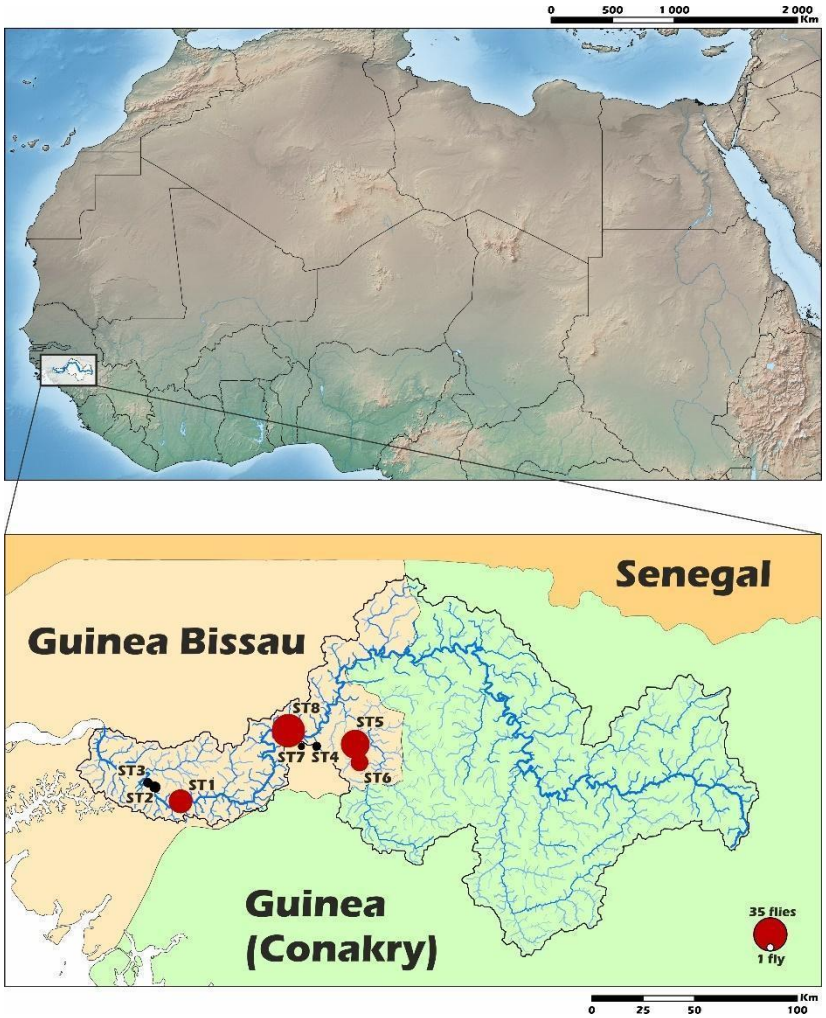


Figure 2 - – Map of the sampled sites in the Dulombi-Boé-Xeché (DBT) complex. ST1- Saltinho, ST2- Cussilinca, ST3- Xitole, ST4- Suto Manca, ST5- Quinhiqué, ST6-Fefiné, ST7- Wendu Tcham, ST8- Xeché. The two traps used for captures were handmade and based on the Vavoua model (Figure 3). Traps were set during a period of 5-7 hours in each location, usually from 11:00-14:00 to 16:00-19:00, at a minimum distance of 20 meters. The placement spots were open and with high sunlight cover.

The dimensions of the two traps were identical, with 80 cm in diameter and a height of 120 cm, the traps were settled with the bottom of the trap about 30 cm above the ground.

It was a conic trap consisting of a cone of mosquito netting attached to three black and blue screens made of cotton fabric joined together at angles of 120°. The traps were fixed to the ground with a wooden stake equipped with an empty plastic bottle with three holes on the sides and one hole on the base. This bottle was placed on top of the stake inside the trap, leaving a passage between the inside of the trap and the outside where the other half of the bottle was placed with a bag net around it to collect the flies. The entrances to the trap were triangular and opened under the mosquito net cone. After the capture, each fly was collected and preserved in individual tubes with 96% ethanol with the respective label added externally and brought to Portugal for the laboratory and bioinformatic analysis. Each fly was preserved individually because the success rate of vertebrate detection increases when in individual collection and extraction instead of pooling the invertebrates (Schnell et al. 2012; Rodgers et al. 2017).

All the samples followed the same labelling system, each one was assigned with a letter representing the site they were captured, followed by a number representing the trap from where they were collected and after a dot, a number identifying the specific fly. For example, a sample labelled with "A1.001" is the first fly caught on the trap number one in Saltinho.

Figure 3 – Vavoua trap in the field (Laveissiere & Grebaut 1990)

Laboratory analysis

All the flies were taxonomically identified in terms of species and their sex was also determined using the Olympus SZX2-ILLT model Led Transmitted Light Illuminator Stand following Pollock (1982). Flies were then prepared for DNA extraction in the laboratory. To extract and isolate the DNA from the samples, the abdomen of each fly was removed. Only abdomens were used instead of other body parts, such as legs and

wings, to ensure we were working with DNA from hosts present in the diet of the flies and not with external contaminations. Each abdomen was posteriorly cut into smaller pieces, placed in a 1.5 ml centrifuge tube, and its DNA extracted using the EZ-10 Spin Column Genomic DNA Miniprep Kit. DNA extraction followed the manufacturer's protocol, except that lysis incubation was done at 55°C overnight, until the tissue was completely lysed. DNA extracts were stored in the fridge at 4°C overnight. The extractions were conducted in a regular laboratory inside a hotte with previous decontamination under UV light, constant decontamination of the laboratory bench and material was done using bleach to reduce contamination. DNA extractions were conducted in batches of up to 23 samples plus a negative control in which no sample was added.

For DNA amplification, primers were selected considering their capacity of amplifying vertebrate prey without amplifying the invertebrate host. RNA genes like mitochondrial 12S (Riaz et al. 2011) or 16S (Taylor, 1996) rDNA, which are located on non-protein coding sequences were used since these often possess variable regions flanked by highly conserved regions that are appropriate for targeted metabarcodes. These markers have been shown to have great resolution power for genus and species identification across numerous vertebrate taxa (Litvaitis, 2000) and with detection rates between 21%-100% (Calvignac-Spencer et al. 2013). The selected forward and reverse primers for 16S were: Chord_16S_F_mod 5'-CGAGAAGACCCTDTGRAGCT-3' and Chord_16S_R_mod 5'-RGATTGCGCTGTTATCCCT-3', both modified from Deagle et al. (2009). The forward and reverse primers for 12S were: Vert01_F 5'-TTAGATACCCCACTATGC-3' and Vert01_R 5'-TAGAACAGGCTCCTCTAG-3' (Riaz et al, 2011). All primers were further modified to contain Illumina adaptors at the 5' end of the sequence.

To prepare the library, a two-step PCR protocol was employed using specific amplification of vertebrate 12S and 16S mitochondrial rRNA genes, followed by an Illumina-based indexing PCR and clean-up steps.

The first PCR reactions (both 12S and 16S) were performed using 5 µL of the Multiplex Master Mix from Qiagen, 0.3 µL of each forward and reverse primers, 2.4 µL of distilled water and 2 µL of DNA for a total volume of 10 µL. The 16S PCR reactions were run at 95 °C for 15 minutes followed by 35 cycles of 30 seconds at 95 °C, 30 seconds at 45 °C and 30 seconds at 72 °C and a final extension of 10 minutes at 72 °C. The 12S PCR reactions were run at 95 °C for 15 min followed by 5 cycles of touch down PCR decreasing 0.5 °C each annealing step from 51 °C to 49 °C, each cycle started at 95 °C

for 30 seconds, 30 seconds at 51 °C (decreasing 0.5 °C until 49 °C) and 15 seconds at 72°C. Then 35 cycles were run at 95 °C for 30 seconds, 49 °C for 30 seconds and 72 °C for 15 seconds and an extension of 10 minutes at 60 °C ending the final cycle at 10 °C. PCR products were visualized on 2% agarose gels (E-gel 96; Invitrogen), stained with GelRedTM.

The second PCR was done to add Illumina P5 and P7 adaptors, along with unique combinations of 7bp long barcodes for sample tagging. PCR reactions consisted in 14 µL reactions containing 7 µL KAPA HiFi HotStart ReadyMix (Rocher, KAPA Biosystems, Basel, Switzerland), 2.8 µL of ultra-pure water, 0.7 µL of each P5 and P7 index and 2.8 µL of 1:4 diluted PCR product. The temperature and cycling profile included incubation at 95 °C for 3 min, followed by 10 cycles at 95 °C for 30 sec, 55 °C for 30 seconds and 72 °C for 30 seconds with a final extension at 72 °C for 5 min and finally holding at 4 °C. A PCR clean-up was done using 1.4 µL AMPure XP beads (Beckman Coulter, Brea, California, USA) per 1.0 µL of sample. Clean PCR products were quantified using Nanodrop. Then all PCR products were pooled equimolarly per marker and each pool was run in TapeStation (Agilent, California, USA) for library quality control. Both pools were further quantified using qPCR, diluted to 4nM and sequenced in a Illumina Miseq using a v3 kit (600 cycles) , along with samples from other projects).

Bioinformatics analysis

Bioinformatic analysis was done using OBITOOLS (Boyer et al. 2016), VSEARCH (Rognes et al. 2016), and LULU (Frøslev et al. 2017). The first step was to merge the paired-end reads in order to align and assemble them (ILLUMINAPAIREDEND). The next step was to add the sample information and by allowing a maximum of four mismatches on primers (NGSFILTER).

Afterwards it was applied OBIUNIQ + OBIANNOTATE to collapse the reads into haplotypes and then the singletons were discarded so sequences likely to have been PCR or sequencing errors could be removed, and all the identical sequences were merged.

In Vsearch, all haplotypes from different samples were collapsed into unique haplotypes. The next step was to do the denoising of the sequences and the removal of the chimeric sequences. Now for the 16S data, the clustering of the OTUs was done at 99% and then was mapping the original reads to retained haplotypes and an OTU table was produced. Finally, a pair-wise identify file was created for LULU. In the 12S run the process was similar with only a difference in the clustering step, where in the 12S case was not done.

The LULU script was run in order to remove potential errors as well as nuclear copies of the mitochondria that may have passed the filtering. To do so, for the 16S data an 84% divergence (minimum_match) was allowed for sequences that cooccurred in 95% of the samples (minimum_relative_cooccurrence). For the 12S data a 95% divergence (minimum_match) was allowed for sequences that cooccurred in 95% of the samples (minimum_relative_cooccurrence).

The generated OTUs were assigned to taxa by comparison with sequences deposited in NCBI, where BLAST identification engine was achieved searching for the closest match based on the Max score, percentage of Query coverage and percent identit, being used 98% similarity to define species level, 95% similarity for genus level and 91% for family level.

Then in an excel sheet, the OTUs and respective matches were classified as diet or /non-diet and as errors (for example if the final species identified was a species not present in the studied area). Finally, all reads observed in the extraction and PCR blanks were removed from the respective samples, along with OTUs not classified as diet, as well as all OTUs with reads representing less than 1% of the dietary reads.

Data analysis

Dietary analysis was based on the presence/absence of taxa per dropping analysed, considering 3 different taxonomic levels: OTU (all taxonomic units identified to the most possibly resolved taxonomic level, even if the unit was classified only up to family or order level), family and order. The OTUs were considered as the most resolved taxonomy instead of species.

Results

Tsetse flies

From all the locations, 116 tsetse flies were captured and preserved. From these flies, 5 were *Glossina morsitans*, 109 were *Glossina palpalis* and 2 were not possible to identify. From *Glossina morsitans*, 1 male and 3 females were identified, while in 1 individual it was not possible to be determine the sex. In the case of *Glossina palpalis*, 9 males and 99 females were identified, while in 1 individual it was also not possible to determine the sex. In the case of the undetermined species, 1 was male and the other was female (Table 1).

In this study, 90 *Glossina palpalis* (87 females and 3 males) from 4 sites (Saltinho, Quinhiqué, Fefiné and Xeché) were selected for gut content analysis, in order to allow minimally robust comparisons across sites.

Table 1- Traps' locations and respective coordinates. Number of species and sex of the collected flies. *G. palpalis/morsitans* stands for *Glossina palpalis/morsitans*

Site	Location	Lat/Long (1)	Lat/Long (2)	Nº Flies	Species	Sex
Site 1	Saltinho	11.6206; -14.6860	11.6210;-14.6855	29	27 *G. palpalis* / 2 *G. morsitans*	24 ♀/ 5 ♂
Site 2	Cussilinta	11.6825; -14.7979	-	5	4 G. palpalis/ 1 G. morsitans	1 ♀ / 2 ♂/ 2 undetermined
Site 3	Suto Manca	11.8619;-14.0901	-	2	1 G. palpalis/ 1 G. morsitans	2♀
Site 4	Xitole	11.7024; -14.8294	-	3	3 *G. palpalis*	3 ♀
Site 5	Quinhiqué	11.8707; -13.9213	-	29	28 *G. palpalis*/ *1 undetermined*	28 ♀/ 1 ♂
Site 6	Fefiné	11.7909;-13.9025	-	12	11 *G. palpalis*/ *1 undetermined*	12 ♀
Site 7	Wendu Tcham	11.8610, -14.1564	-	1	1 *G. palpalis*	1 ♀
Site 8	Xeché	11.9303; -14.2138	11.9303;-14.2134	35	34 *G. palpalis*/ 1 *G. morsitans*	32 ♀/ 3 ♂
TOTAL	-	-	-	116	109 G. palpalis/ 5 G. morsitans/ 2 undetermined	103 ♀/ 11 ♂/ 2 undetermined

Unfiltered taxa

Before the filtration steps in the bioinformatics, there was taxa present in the samples that were considered as diet and as field or laboratory contamination (Table 2). In the table are present 74 different taxa in which 34 were considered as part of the diet of the tsetse flies, 16 were considered laboratory contamination, 12 were considered field contamination and 12 were considered non-target amplifications for both primers. In

total, 70 haplotypes and 1 458 561 reads were generated in the 16S marker and 189 haplotypes and 2 932 593 reads were generated in the 12S run.

Table 2- Filtered taxa in 16S run and 12S run.

			16S		12S	
Order	**Family**	**OTU**	**No. OTUs**	**No. reads**	**No. OTUs**	**No. reads**
Anura	Arthroleptidae	*Leptopelis sp.*	-	-	1	1803
	Hyperoliidae	*Hyperolius sp.*	-	-	1	819
	Phrynobatrachidae	*Phrynobatrachus francisci*	-	-	1	2331
Artiodactyla	Bovidae	*Bos taurus*	1	210	1	6144
		Bovidae	1	301	1	186
		Ovis aries	1	904	1	4459
		Tragelaphus scriptus	1	119598	5	226491
	Suidae	*Phacochoerus africanus*	1	80931	1	123154
		Potamochoerus porcus	1	27384	1	69345
		Sus scrofa	2	123160	1	210185
Carnivora	Canidae	*Canis familiaris*	1	3616	4	21202
Crocodylia	Crocodylidae	*Crocodylus suchus*	1	1188	-	-
		Crocodylus sp.	-	-	1	5547
Galliformes	Phasianidae	*Gallus gallus*	1	13	2	64698
Passeriformes	Ploceidae	*Ploceus sp.*	1	3081	-	-
		Passeridae	-	-	1	11781
Pelecaniformes	Ardeidae	*Ardea cinerea*	-	-	1	12278
		Ardea purpurea	1	22	1	35121
		Ardea sp.	-	-	1	1590
Perissodactyla	Equidae	*Equus asinus*	-	-	1	678
Primates	Hominidae	*Homo sapiens*	6	106224	25	361407
Rodentia	Hystricidae	*Hystrix sp.*	1	4192	-	-
	Thryonomyidae	*Thryonomys swinderianus*	1	548	1	4673
Squamata	Elapidae	*Naja sp.*	2	142596	-	-
	Pythonidae	*Pythonidae*	1	1784	1	46861
		Python sp.	-	-	1	1418
	Varanidae	*Varanus niloticus*	1	684380	2	928724
	Gekkonidae	*Gekkonidae*	-	-	1	88
Testudines	Pelomedusidae	*Pelusios castaneus*	1	4380	1	26013
TOTAL			**26**	**1304512**	**58**	**2166996**

For the taxa considered as diet, 17 taxa were detected with both markers. For the taxa considered contaminations, they were separated into field and laboratory contaminations based on the distribution of the respective taxa. If the taxa occurred in the study area, it was considered field contamination and the taxa inhabiting outside the study area was considered laboratory contaminations.

For the 16S, 38 different taxa were detected, while in the 12S, 61 different taxa were detected. Between the two runs, 25 different taxa were shared successful detection in both primers.

From the 75 different taxa, 38 were able to be detected at species level. In the 16S marker, 12 taxa had more than 1 haplotype, while in the 12S was 25.

Filtered taxa

After the filtering process, 1,304,512 million reads, 26 OTUs and 26 haplotypes were retained for the 16S marker and 2,166,996 million reads, 58 OTUs and 58 haplotypes for the 12S marker (Figure 4).

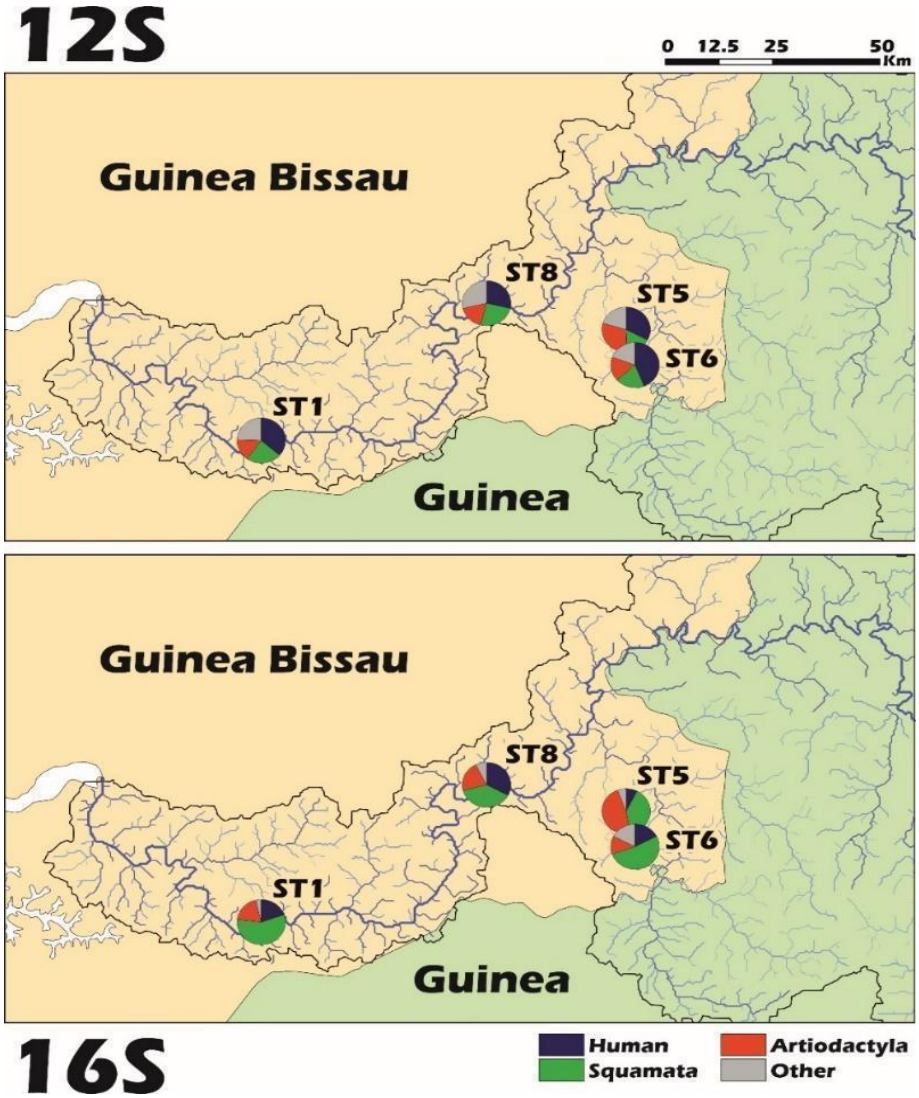


Figure 4 – Map of the study area with indication of the four sampled sites (ST1-Saltinho, ST5-Quinhiqué, ST6-Fefiné and ST8-Xeché) and respective piecharts, representing the three main orders detected: Primates (Human), Squamata, Artiodactyla and the other orders.

Diet analysis

Regarding host diversity per sample, the number of hosts in the 16S marker varied between 1 to 5 different taxa per sample, while in the 12S marker, the number of different hosts per sample ranged from 1 to 9 (Figure 5).

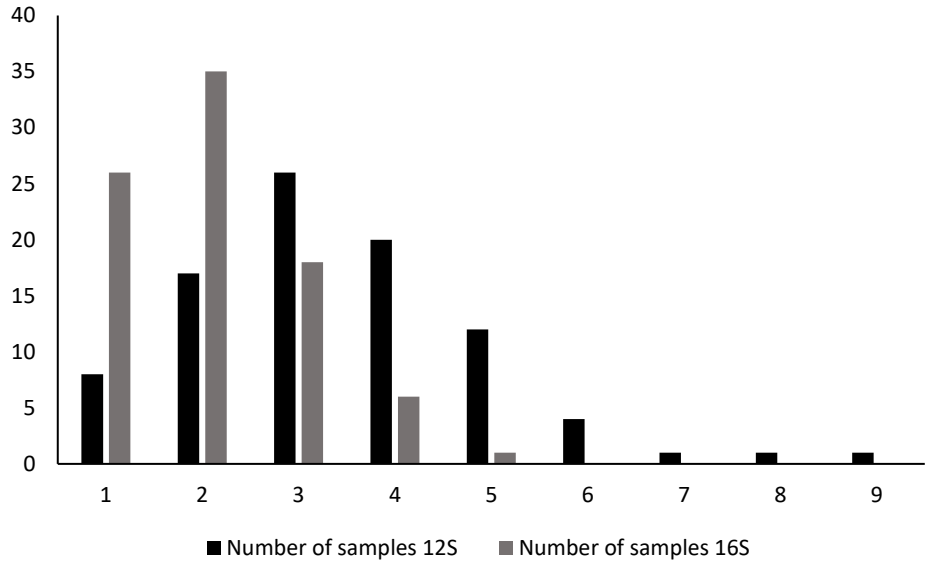


Figure 5 – Number of different hosts present in each sample for both primers (the 12S and the 16S)

In terms of detected taxa, in a general view of the 16S, it was possible to detect the presence of Squamata in 84 samples, the presence of Primates (*Homo sapiens*) in 33 samples, Artiodactyla in 49 samples in all locations combined, the class Aves was detected in 3 samples and other taxa in 10 samples (Figure 6). The locations Fefiné (8) and Saltinho (7) had the lowest number of detetcted taxa when compared to Quinhiqué (11) and Xeché (11). The order Squamata (mainly *Varanus niloticus*) was caught the most in Fefiné, Xeché, Saltinho and in general, while Artiodactyla was detected the most in Quinhiqué. In Fefiné the order Squamata was detected in 12 samples, the order Artiodactyla was detected in 3 samples, the order Primates was detected in 4 samples and the other orders were detected in 3 samples. In Quinhiqué the order Squamata was detected in 21 samples, the order Artiodactyla was detected in 25 samples, the order Primates was detected in 4 samples and the other orders were detected in 3 samples. In Saltinho, the order Squamata was detected in 30 samples, the order Artiodactyla was detected in 9 samples, the order Primates was detected in 10 samples and the other orders were detected in 2 samples. In Xeché the order Squamata was detected in 21 samples, the order Artiodactyla was detected in 12 samples, the order Primates was detected in 15 samples and the other orders were detected in 3 samples (Figure 7, Attachment 1 Figure A1 and Figure A2).

In terms of detected taxa, in a general view of the 12S, it was possible to detect the presence of Squamata in 75 samples, the presence of Primates (Homo sapiens) in 63 samples, Artiodactyla in 76 samples, the class Aves was detected in 56 samples and other taxa in 38 samples. The locations Fefiné (10) had the lowest number of detetcted taxa when compared to and Saltinho (13), Quinhiqué (14) and Xeché (22) had the highest number of detected taxa. In Fefiné the order Squamata was detected in 10 samples, the order Artiodactyla was detected in 10 samples, the order Primates was detected in 7 samples and the other orders were detected in 11 samples. In Quinhiqué the order Squamata was detected in 17 samples, the order Artiodactyla was detected in 32 samples, the order Primates was detected in 17 samples and the other orders were detected in 11 samples. In Saltinho the order Squamata was detected in 25 samples, the order Artiodactyla was detected in 17 samples, the order Primates was detected in 19 samples and the other orders were detected in 32 samples. In Xeché the order Squamata was detected in 29 samples, the order Artiodactyla was detected in 17 samples, the order Primates was detected in 20 samples and the other orders were detected in 32 samples (Attachment 1 Figure A1 and Figure A2).

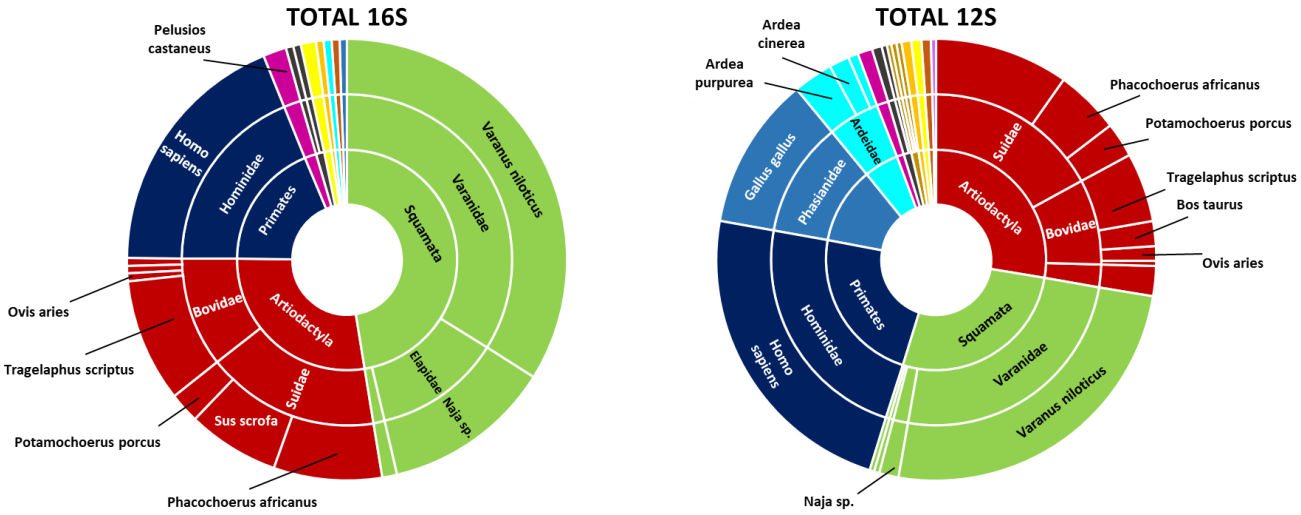


Figure 6 – Detected taxa in total for both primers (16S and 12S)

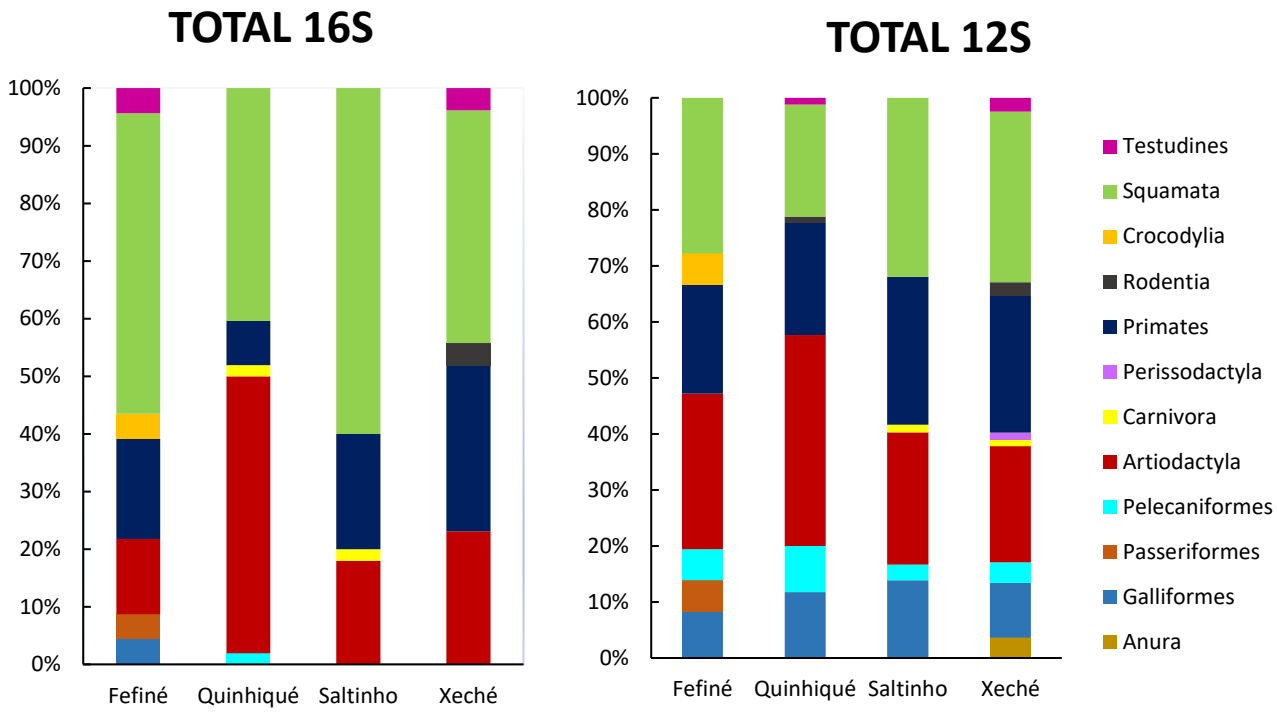


Figure 7 – Percentage of the detected orders in all locations for both primers.

The number of different species caught in the 12S is superior to the number of species caught in the 16S, as well as the number of samples with detected taxa.

In total, it was possible to reach 12 different orders, 14 families and 14 species for the 16S. For the 12S, it was possible to reach 19 different orders, 19 families and 17 species (Attachment 1 Table A1).

In the 16S primer, 2 sites had 5 orders detected (Xeché and Quinhiqué), Saltinho had the lowest number of orders (4) and Fefiné was the location with the most orders detected (7). In terms of taxa to reach species level, in Fefiné 6 different species were detected, in Quinhiqué were 9 species, in Saltinho were 6 species and in Xeché were 8 species. In terms of the dominant species detected, in all locations was *Varanus niloticus*.

In the 12S primer, 2 sites had 7 orders detected (Fefiné and Quinhiqué), in Saltinho 6 orders were detected and Xeché had the most orders detected (10). When looking at the detection in species level, in Fefiné 7 different species were detected, in Quinhiqué were 11 species, in Saltinho were 11 species as well and in Xeché were 16species. In terms of the dominant species, *Homo sapiens* was the dominant species in Quinhiqué, while in the other locations was *Varanus niloticus*.

Discussion

The results from this study give an insight about the occurrence of some mammalian, birds and reptile groups in the study area. The constructed traps were able to capture tsetse flies with a total of 116 flies. A bigger number of traps covering more area would probably lead to a greater number of trapped flies for the study. Additionally, a longer period of exposure should also increase the number of catches. The number of flies captured per trap was highly variable, ranging from 0 to 32. Only 4 out of the 8 sampled locations had a substantial number of flies caught in the traps. Both traps have worked with the same efficiency, and success of captures seem to be linked to the microhabitat features around each trap and time of the capture. Since we did not use enough replicates to do statistical analyses and because no microhabitat features were characterised, it is not possible for us to identify which feature is important to explain capture efficiency. To optimize this protocol, future studies should include a minimum of 5 traps per station and a detailed microhabitat characterization should be done with at least 10 meters radius around each trap site. The experimental set should be repeated during different periods of the day using a standardized period across experiments. In the study of Makhulu EE (2021) performed in Tanzania between June and July of 2016, the number of catches was higher (1167) when compared with the flies caught in this study. The experiment was done in 10 locations using Ngu traps baited with acetone and cow urine and the traps were set in the morning with a period of 24 hours before the trapped flies were collected again. The number of sites was superior, the period of exposure was higher, the trap model was different, the sampling season was in other period of the year and attractive odours were added to the traps. Perhaps these factors may have contributed to a higher number of collected flies.

The trapping was performed only during the dry season. Since flies' abundance could vary between the rainy and dry season, this also applying to the distribution of the hosts, the bloodmeals composition may have varied if the study was performed in other time of the year. Therefore, it is recommended to repeat this study at least once per season. The number of flies caught during the dry season (263) in the study of Mulugeta et al. 2013 in Ethiopia, was inferior to the late rainy (582), early rainy (322) and wet (379) seasons, showing a different approach in the potential abundance of tsetse flies during different seasons.

The gradient of habitat humanisation caused by urbanisation, grazing, and agriculture, may also influence tsetse flies' abundance and composition of potential hosts (both wild

and domestic animals). Other principal habitat gradients may also influence both the number of tsetse flies and the hosts, as well as their interactions. Future studies should be carried at different gradients of humanisation and in distinct habitats to increase our knowledge to this biotic interaction. According to Imna (2011), the human population growth and land expansion (for agriculture and grazing) are destroying the breeding and resting sites of tsetse flies, leading to the vanishment of tsetse flies' population from their initial favoured habitats.

Considering the species of tsetse flies collected, the number of *Glossina morsitans* caught in the traps was much lower than the number of *Glossina palpalis*. This can be explained given that *G. morsitans* are considered xerophilous, preferring drier and more pristine habitats, being more abundant in East Africa. Conversely, *G. palpalis* is more tolerant of human presence, preferring more humid habitats and are more prevalent in West Africa (Buxton, 1955; Odete & Afonso, 2012).

There were more females collected comparing to the males. Female tsetse flies tend to emerge first than the male's tsetse flies (Itard & Bauer, 1984). Sex may also influence the number and variety of hosts. However, due to the low number of males collected, no statistical analysis was possible to test this. In future studies, we propose using a higher number of males and females in the same sites to test if sex has any influence on host choice. Similar results were detected in Mohamed & Dairri (1987) and Leak (1999), where female population composed between 70 to 80% of the mean population of the collected flies.

The amplification of gene fragments with both primers was successful. This means that the quality of the vertebrate DNA inside the flies still hold a good quality, after sample preservation, manipulation and a month of storage.

A wide variety of hosts (n=29) was detected from the DNA metabarcoding results, including humans, domestic (e.g. cows, pigs, sheep, or dogs) and wild animals (e.g. cape bushbuck, monitor lizards, or mud turtles).

For both primers, 3 orders were clearly more represented than the rest (Primates, Squamata and Artiodactyla) in all locations. These results are concordant with other published results, also identified these taxa as the base diet of tsetse flies (Weitz, 1963).

In every location, the most detected taxa in 16S belong to the order Squamata (mainly Varanus niloticus), while in 12S, the order Primates (only Homo sapiens) was more

represented only in Quinhiqué. The number of different species detected with the 12S fragment is superior to the number of species detected with the 16S fragment.

As for site specific detection, the highest abundance of Artiodactyla was noticed in Quinhiqué, especially of wild species. This can be explained by the higher distance to human settlements compared with the other 3 locations isolations. The lower frequency of human activities in this area, also explains the lower detection of human DNA in tsetse flies' diet from this site. Conversely, in the other locations closer to human populations, the number of domestic animals and human DNA detected is higher.

According to Bouyer et al. (2007), the first meal of the tsetse fly has a great influence in the next meal, so if a particular taxon is very common in a given area, the probability of the next meal to be the same taxa should be higher.

As for the number of detected hosts in each fly, it ranged from 1 to 9 hosts, and in most cases more than 1 host was detected in a single fly. *Glossina palpalis* are riparian flies (Odete & Afonso, 2012) that easily adapt to host abundance and are not conditioned to a strict feeding preference, so they can have a more varied diet depending on host availability and the conditions that the habitat provides (Clausen et al. 1998; de la Rocque et al. 2001). Given that G. palpalis only feed every 2 days, this means we were able to detect possible meals up to 18 days of age (Mellanby,1936).

A variety of hosts (n=75 different taxa) was also detected from the DNA metabarcoding results, but with a small number of reads and other species that are unlikely to be hosts of tsetse flies. For the former case, taxa with less than 1% of reads were discarded. For the latter, potential contamination was the most probable reason for these species detections. The most probable contaminations might have been originated in the field or the laboratory. During the field work, the transportation of the traps between sites was done together with other materials. For example, traps were transported inside buckets that contained or had contained other animals collected for other studies (explaining the African fish detected), also the traps were not cleaned or decontaminated after being used. Given DNA extractions were performed from a lysate of the whole abdomen of each fly without washing its exterior, this would allow an easy DNA transfer from potentially contaminated nets. This could lead to DNA traces of animals that were not hosts of the collected tsetse flies. In future studies, more careful transport and maintenance of traps, and washing of the flies before DNA extraction is recommended to avoid possible field contaminations.

DNA extractions were performed in a standard molecular laboratory, which can be potentially contaminated with DNA sources from other studies. This can explain the residual detection of the Iberian lynx and the European fire salamander. The sterilization of the laboratory bench and equipment may not been 100% effective, and the appearance of species from different locations around the world can be explained through external contamination in the laboratory.

There are many vertebrate species known to be present in the study area, but these species were not detected in the metabarcoding analyses. For example, although there are many species of primated from the family Cercopithecoidea inhabiting the study area, none of them was detected using the metabarcoding technique. Undetected species from these family do not necessarily mean their absence in the study area, because the tsetse flies may not have fed on these primates (Alberdi et al. 2019). Another explanation could be the possibility of some incompleteness in the DNA reference database.

According to Rodgers et al. (2017), the 16S primer should have detected more often the presence of Primates when compared to the 12S, but in this study the percentage of Primates was higher only in Quinhiqué, while the Squamata had the highest percentage in both markers. This shows the complementary relation between the two primers in species detection in the same study area. In future studies the addition of more primers could lead to a wider detection of different vertebrate taxa. The 16S had fewer species comparing to the 12S, although there was some overlapping of species between both primers. The differential detection of taxa between primers is not very high and maybe explained by the different affinities of each primer to DNA templates of each group, also known as primer biases (Elbrecht et al. 2015).

It was possible to detect some bird at the family and even at species level, whereas in other tsetse diet' metabarcoding studies this was not so common. Since birds are not recognised as part of tsetse flies' diets, further studies are needed to ascertain their role as hosts of tsetse flies.

Although not tested in this study, the metabarcoding of tsetse flies' diet may also be used to detect trypanosomes, that cause sleep disease, as well as other vertebrate parasites. Future studies in the prevalence of diseases caused by these parasites should also be tested using the protocol here presented.

Conclusion

In the present study, we developed a metabarcoding protocol to detect a varied diversity of hosts from tsetse flies' diet. We also were able to identify flaws and bottlenecks that limit its efficiency and utility, giving suggestions on how to improve each stage. A special attention was given to field and laboratory contamination, and suggestions steps were given to improve the protocol to overcome this issue.

From the case study in the Corubal river basin of Guinea Bissau, we conclude that tsetse flies' diet metabarcoding analysis may be used in rapid bioassessments of specific vertebrate groups, such as the Artiodactyla and some groups of large reptiles.

Both markers used, successfully amplified 25 different species in 4 different locations. The three most detected orders were Squamata, Artiodactyla and Primates. The most detected species were Homo sapiens and Varanus niloticus. Most of the hosts groups previously known to be present in tsetse flies' diet were detected, including other species not so frequently reported (e.g. some reptile species). Both primers generally obtained the same results, although with small differences, showing that the combination of different primers increase the range of taxa detection. Some known species in the study area were not detected, therefore further studies with a more comprehensive sampling coupled with other assessments should clarify the efficiency of this methodology.

Given that tsetse flies' metabarcoding mainly target specific groups, for a wider biodiversity of terrestrial vertebrates, it should be used as a compliment for other survey methods such, as camera trapping or soil/water eDNA metabarcoding.

The study presented in this thesis is an important first step for the study of tsetse flies' diet in Guinea Bissau and provide a baseline protocol to be used in further bioassessments of terrestrial vertebrates and potentially in public health issues caused by vertebrates' parasites.

References

Aizpurua, O., Budinski, I., Georgiakakis, P., Gopalakrishnan, S., Ibañez, C., Mata, V., Rebelo, H., Russo, D., Szodoray-Parádi, F., Zhelyazkova, V., Zrncic, V., Gilbert, M. T. P., & Alberdi, A. (2018). Agriculture shapes the trophic niche of a bat preying on multiple pest arthropods across Europe: Evidence from DNA metabarcoding. *Molecular Ecology*, *27*(3), 815–825. https://doi.org/10.1111/MEC.14474

Aksoy, S., & Rio, R. V. M. (2005). Interactions among multiple genomes: Tsetse, its symbionts and trypanosomes. *Insect Biochemistry and Molecular Biology*, *35*(7), 691–698. https://doi.org/10.1016/J.IBMB.2005.02.012

Alberdi, A., Aizpurua, O., Bohmann, K., Gopalakrishnan, S., Lynggaard, C., Nielsen, M., & Gilbert, M. T. P. (2019). Promises and pitfalls of using high-throughput sequencing for diet analysis. *Molecular Ecology Resources*, *19*(2), 327–348. https://doi.org/10.1111/1755-0998.12960

Atouguia, J., & Costa, J. (1999). Therapy of Human African Trypanosomiasis: Current Situation. *Memórias Do Instituto Oswaldo Cruz*, *94*(2), 221–224. https://doi.org/10.1590/S0074-02761999000200018

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, *15*(4), 365–377. https://doi.org/10.1111/J.1461-0248.2011.01736.X

Binladen, J., Gilbert, M. T. P., Bollback, J. P., Panitz, F., Bendixen, C., Nielsen, R., & Willerslev, E. (2007). The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing. *PLOS ONE*, *2*(2), e197. https://doi.org/10.1371/JOURNAL.PONE.0000197

Boyer, F., Mercier, C., Bonin, A., le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: a unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182. https://doi.org/10.1111/1755-0998.12428

Brooks, T. M., Mittermeier, R. A., da Fonseca, G. A. B., Gerlach, J., Hoffmann, M., Lamoreux, J. F., Mittermeier, C. G., Pilgrim, J. D., & Rodrigues, A. S. L. (2006). Global biodiversity conservation priorities. *Science*, *313*(5783), 58–61. https://doi.org/10.1126/SCIENCE.1127609/SUPPL_FILE/BROOKS.SOM.PDF

Buxton, P. A. (1955). The Natural History of Tsetse Flies. An Account of the Biology of the Genus Glossina (Diptera). *The Natural History of Tsetse Flies. An Account of the Biology of the Genus Glossina (Diptera).*

Calvignac-Spencer, S., Leendertz, F. H., Gilbert, M. T. P., & Schubert, G. (2013). An invertebrate stomach's view on vertebrate ecology. *BioEssays*, *35*(11), 1004–1013. https://doi.org/10.1002/BIES.201300060

Catarino, L., Martins, E. S., & Moreira, I. (2001). Influence of environmental features in the phytogeographic framework of Guinea-Bissau. *Systematics and Geography of Plants*, *71*(2), 1079–1086. https://doi.org/10.2307/3668740

Chappuis, F., Lima, M. A., Flevaud, L., & Ritmeijer, K. (2010). Human African Trypanosomiasis in Areas without Surveillance. *Emerging Infectious Diseases*, *16*(2), 354. https://doi.org/10.3201/EID1602.090967

Chariton, A. A., Court, L. N., Hartley, D. M., Colloff, M. J., & Hardy, C. M. (2010). Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. *Frontiers in Ecology and the Environment*, *8*(5), 233–238. https://doi.org/10.1890/090115

Clare, E. L., Barber, B. R., Sweeney, B. W., Hebert, P. D. N., & Fenton, M. B. (2011). Eating local: influences of habitat on the diet of little brown bats (Myotis lucifugus). *Molecular Ecology*, *20*(8), 1772–1780. https://doi.org/10.1111/J.1365-294X.2011.05040.X

Clausen, P. H., Adeyemi, I., Bauer, B., Breloeer, M., Salchow, F., & Staak, C. (1998). Host preferences of tsetse (Diptera: Glossinidae) based on bloodmeal identifications. *Medical and Veterinary Entomology*, *12*(2), 169–180. https://doi.org/10.1046/J.1365-2915.1998.00097.X

Deagle, B. E., Chiaradia, A., McInnes, J., & Jarman, S. N. (2010). Pyrosequencing faecal DNA to determine diet of little penguins: Is what goes in what comes out? *Conservation Genetics*, *11*(5), 2039–2048. https://doi.org/10.1007/S10592-010-0096-6/METRICS

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, *10*(9). https://doi.org/10.1098/RSBL.2014.0562

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., Kartzinel, T. R., & Eveson, J. P. (2019). Counting with DNA in metabarcoding studies:

How should we convert sequence reads to dietary data? *Molecular Ecology*, *28*(2), 391–406. https://doi.org/10.1111/MEC.14734

DESPOMMIER, D. (2005). Emerging Zoonoses and Pathogens of Public Health Concern. World Organization for Animal Health. Volume 23 (2), 2004. pp. 310. ISBN 92 9044 621 8. € 50 (airmail postage is included for all countries). *Parasitology*, *131*(1), 70–70. https://doi.org/10.1017/S0031182005007869

Desta, M., Menkir, S., & Kebede, A. (2013). *Journal of Veterinary Medicine and Animal Health The study on tsetse fly (Glossina species) and their role in the trypanosome infection rate in Birbir valley, Baro Akobo River system, western Ethiopia The study was conducted in Birbir valley of Oromia Regional State, Western Ethiopia from November*. *5*(7), 186–194. https://doi.org/10.5897/JVMAH2013.0223

di Finizio, A., Guerriero, G., Russo, G. L., & Ciarcia, G. (2007). Identification of gadoid species (Pisces, Gadidae) by sequencing and PCR-RFLP analysis of mitochondrial 12S and 16S rRNA gene fragments. *European Food Research and Technology*, *225*(3–4), 337–344. https://doi.org/10.1007/S00217-006-0420-Z/METRICS

Dodman, T., & Sá, J. (n.d.). *Monitorização de aves aquáticas no Arquipélago dos Bijagós, Guiné-Bissau*. Retrieved December 21, 2022, from www.nhbs.co.uk

Drinkwater, R., Jucker, T., Potter, J. H. T., Swinfield, T., Coomes, D. A., Slade, E. M., Gilbert, M. T. P., Lewis, O. T., Bernard, H., Struebig, M. J., Clare, E. L., & Rossiter, S. J. (2021). Leech blood-meal invertebrate-derived DNA reveals differences in Bornean mammal diversity across habitats. *Molecular Ecology*, *30*(13), 3299–3312. https://doi.org/10.1111/MEC.15724

Elbrecht, V., & Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. *PLOS ONE*, *10*(7), e0130324. https://doi.org/10.1371/JOURNAL.PONE.0130324

Ferro, M. L., & Park, J.-S. (2013). Effect of Propylene Glycol Concentration on Mid-Term DNA Preservation of Coleoptera. *Https://Doi.Org/10.1649/0010-065X-67.4.581*, *67*(4), 581–586. https://doi.org/10.1649/0010-065X-67.4.581

Franco, J. R., Simarro, P. P., Diarra, A., & Jannin, J. G. (2014). Epidemiology of human African trypanosomiasis. *Clinical Epidemiology*, *6*(1), 257. https://doi.org/10.2147/CLEP.S39728

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications 2017 8:1*, *8*(1), 1–11. https://doi.org/10.1038/s41467-017-01312-x

Gariepy, T. D., Lindsay, R., Ogden, N., & Gregory, T. R. (2012). Identifying the last supper: utility of the DNA barcode library for bloodmeal identification in ticks. *Molecular Ecology Resources*, *12*(4), 646–652. https://doi.org/10.1111/J.1755-0998.2012.03140.X

GASTON, K. J., SCOBLE, M. J., & CROOK, A. (1995). Patterns in species description: a case study using the Geometridae (Lepidoptera). *Biological Journal of the Linnean Society*, *55*(3), 225–237. https://doi.org/10.1111/J.1095-8312.1995.TB01061.X

Gorillas of Bwindi-Impenetrable Forest and the Virunga Volcanoes: Taxonomic implications of morphological and ecological differences - Sarmiento - 1996 - American Journal of Primatology - Wiley Online Library. (n.d.). Retrieved December 21, 2022, from https://onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1098-2345(1996)40:1%3C1::AID-AJP1%3E3.0.CO;2-1

Haile, J., Froese, D. G., MacPhee, R. D. E., Roberts, R. G., Arnold, L. J., Reyes, A. v., Rasmussen, M., Nielsen, R., Brook, B. W., Robinson, S., Demuro, M., Gilbert, M. T. P., Munch, K., Austin, J. J., Cooper, A., Barnes, I., Möller, P., & Willerslev, E. (2009). Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22352–22357. https://doi.org/10.1073/PNAS.0912510106/SUPPL_FILE/0912510106SI.PDF

Hammond, P. M. (1994). Practical approaches to the estimation of the extent of biodiversity in speciose groups. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 119–136. https://doi.org/10.1098/RSTB.1994.0092

Hill, C. M. (2010). Crop-raiding by wild vertebrates: The farmer's perspective in an agricultural community in western Uganda. *Http://Dx.Doi.Org/10.1080/096708797229022*, *43*(1), 77–84. https://doi.org/10.1080/096708797229022

Hoare, C. A. (1972). The trypanosomes of mammals. A zoological monograph. *The Trypanosomes of Mammals. A Zoological Monograph.*

Hoffmann, C., Merkel, K., Sachse, A., Rodríguez, P., Leendertz, F. H., & Calvignac-Spencer, S. (2018). Blow flies as urban wildlife sensors. *Molecular Ecology Resources*, *18*(3), 502–510. https://doi.org/10.1111/1755-0998.12754

Itard, J., & Bauer, B. (1984). Elevages de glossines. Synthèse. *Revue d'Elevage et de Médecine Vétérinaire Des Pays Tropicaux*.

Jordan, A. M. (1985a). THE VECTORS OF AFRICAN TRYPANOSOMIASIS: RESEARCH TOWARDS NON-INSECTICIDAL METHODS OF CONTROL. *British Medical Bulletin*, *41*(2), 181–186. https://doi.org/10.1093/OXFORDJOURNALS.BMB.A072047

Jordan, A. M. (1985b). THE VECTORS OF AFRICAN TRYPANOSOMIASIS: RESEARCH TOWARDS NON-INSECTICIDAL METHODS OF CONTROL. *British Medical Bulletin*, *41*(2), 181–186. https://doi.org/10.1093/OXFORDJOURNALS.BMB.A072047

Kemp, B. M., & Smith, D. G. (2005). Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. *Forensic Science International*, *154*(1), 53–61. https://doi.org/10.1016/J.FORSCIINT.2004.11.017

Kocher, A., de Thoisy, B., Catzeflis, F., Valière, S., Bañuls, A. L., & Murienne, J. (2017). iDNA screening: Disease vectors as vertebrate samplers. *Molecular Ecology*, *26*(22), 6478–6486. https://doi.org/10.1111/MEC.14362

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. v., Paabo, S., Villablanca, F. X., & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences*, *86*(16), 6196–6200. https://doi.org/10.1073/PNAS.86.16.6196

Larsen, T. H., & Travis Thyberg, by. (n.d.). *Core Standardized MethodS for rapid BiologiCal field aSSeSSMent*. Retrieved December 26, 2022, from www.conservation.org

Laveissiere, C., & Grebaut, P. (1990). [The trapping of tsetse flies (Diptera: Glossinidae). Improvement of a model: the Vavoua trap]. *Tropical Medicine and Parasitology : Official Organ of Deutsche Tropenmedizinische Gesellschaft and of Deutsche Gesellschaft Fur Technische Zusammenarbeit (GTZ)*, *41*(2), 185–192. https://europepmc.org/article/med/2166330

Lee, P. C., & Graham, M. D. (2006). African elephants Loxodonta africana and human-elephant interactions: implications for conservation. *International Zoo Yearbook*, *40*(1), 9–19. https://doi.org/10.1111/J.1748-1090.2006.00009.X

Linacre, A., & Tobe, S. S. (2011). An overview to the investigative approach to species testing in wildlife forensic science. *Investigative Genetics*, *2*(1), 1–9. https://doi.org/10.1186/2041-2223-2-2/FIGURES/2

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, *45*(3), 373–385. https://doi.org/10.1111/EEN.12831

Macdonald, D. W., Johnson, P. J., Albrechtsen, L., Seymour, S., Dupain, J., Hall, A., & Fa, J. E. (2012). Bushmeat trade in the Cross–Sanaga rivers region: Evidence for the importance of protected areas. *Biological Conservation*, *147*(1), 107–114. https://doi.org/10.1016/J.BIOCON.2011.12.018

Makhulu, E. E., Villinger, J., Adunga, V. O., Jeneby, M. M., Kimathi, E. M., Mararo, E., Oundo, J. W., Musa, A. A., & Wambua, L. (2021). Tsetse blood-meal sources, endosymbionts and trypanosome-associations in the Maasai Mara National Reserve, a wildlife-human-livestock interface. *PLOS Neglected Tropical Diseases*, *15*(1), e0008267. https://doi.org/10.1371/JOURNAL.PNTD.0008267

Malele, I., Nyingilili, H., & Msangi, A. (2011a). Factors defining the distribution limit of tsetse infestation and the implication for livestock sector in Tanzania. *African Journal of Agricultural Research*, *6*(10), 2341–2347. https://doi.org/10.5897/AJAR10.1050

Massey, A., Roffler, G., Vermeul, T., Allen, J., & Levi, T. (2019). Comparison of mechanical sorting and DNA metabarcoding for diet analysis with degraded wolf scats. *BioRxiv*, 2019.12.13.875898. https://doi.org/10.1101/2019.12.13.875898

Meekan, M., Austin, C. M., Tan, M. H., Wei, N. W. v., Miller, A., Pierce, S. J., Rowat, D., Stevens, G., Davies, T. K., Ponzo, A., & Gan, H. M. (2017). iDNA at sea: Recovery of whale shark (Rhincodon typus) mitochondrial DNA sequences from the whale shark copepod (Pandarus rhincodonicus) confirms global population structure. *Frontiers in Marine Science*, *4*(DEC), 420. https://doi.org/10.3389/FMARS.2017.00420/BIBTEX

Mellanby, K. (1936). Experimental Work with the Tsetse-fly, Glossina palpalis, in Uganda. *Bulletin of Entomological Research*, *27*(4), 611–632. https://doi.org/10.1017/S0007485300058703

Mendes, O. (2018). *Indicadores das mudanças climáticas no leste da Guiné-Bissau e adaptação camponesa*. https://recil.ensinolusofona.pt/handle/10437/12023

Milner-Gulland, E. J., Bennett, E. L., Abernethy, K., Bakarr, M., Bennett, E., Bodmer, R., Brashares, J., Cowlishaw, G., Elkan, P., Eves, H., Fa, J., Milner-Gulland, E. J., Peres, C., Roberts, C., Robinson, J., Rowcliffe, M., & Wilkie, D. (2003). Wild meat: the bigger picture. *Trends in Ecology & Evolution*, *18*(7), 351–357. https://doi.org/10.1016/S0169-5347(03)00123-X

Mohamed Ahmed, M. M., & Dairri, M. F. (1987). Trypanosome infection rate of Glossina pallidipes during wet and dry seasons in Somalia. *Tropical Animal Health and Production*, *19*(1), 11–20. https://doi.org/10.1007/BF02250840/METRICS

Muturi, C. N., Ouma, J. O., Malele, I. I., Ngure, R. M., Rutto, J. J., Mithöfer, K. M., Enyaru, J., & Masiga, D. K. (2011). Tracking the Feeding Patterns of Tsetse Flies (Glossina Genus) by Analysis of Bloodmeals Using Mitochondrial Cytochromes Genes. *PLOS ONE*, *6*(2), e17284. https://doi.org/10.1371/JOURNAL.PONE.0017284

Odete, M., & Afonso, A. (. (2012). BIOECOLOGIA E IMPORTÂNCIA MÉDICA DAS GLOSSINAS (Diptera, Glossinidae) E DOS FLEBÓTOMOS (Diptera, Psychodidae) COMO VETORES DE TRIPANOSSOMATÍDEOS. *Anais Do Instituto de Higiene e Medicina Tropical*, *11*, 85–89. https://doi.org/10.25761/ANAISIHMT.219

Patrick, H. J. H., Chomič, A., & Armstrong, K. F. (2016). Cooled Propylene Glycol as a Pragmatic Choice for Preservation of DNA From Remote Field-Collected Diptera for Next-Generation Sequence Analysis. *Journal of Economic Entomology*, *109*(3), 1469–1473. https://doi.org/10.1093/JEE/TOW047

Pollock, J. N. (1992). *Training Manual for TSETSE CONTROL PERSONNEL Volume 1*.

Poulton, S., Birks, J. D. S., & Messenger, J. E. (n.d.). *A QUALITY-SCORING SYSTEM FOR USING SIGHTINGS DATA TO ASSESS PINE MARTEN DISTRIBUTION AT LOW DENSITIES The Vincent Wildlife Trust, 3 & 4 Bronsil Courtyard*.

Research Techniques in Animal Ecology: Controversies and Consequences - Luigi Boitani, T. K. Fuller - Google Livros. (n.d.). Retrieved December 21, 2022, from https://books.google.pt/books?hl=pt-PT&lr=&id=5c1xxoBc0zAC&oi=fnd&pg=PA165&dq=Litvaitis,+J.+A.+(2000).+Investigating+food+habits+of+terrestrial+vertebrates.+Research+techniques+in+animal+ecology:+controversies+and+consequences.+Columbia+University+Press,+New+York,+165-190.+&ots=PzMAdwLNNe&sig=enuMR2QkYhkx2iQvGwx6iEyrL0c&redir_esc=y#v=onepage&q&f=false

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, *39*(21), e145–e145. https://doi.org/10.1093/NAR/GKR732

Robbins, C. B., & van der Straeten, E. (1996). Small mammals of Togo and Benin. II. Rodentia. *Mammalia*, *60*(2), 231–242. https://doi.org/10.1515/MAMM.1996.60.2.231/MACHINEREADABLECITATION/RIS

Rocque, S. de la, Augusseau, X., Guillobez, S., Michel, V., Wispelaere, G. de, Bauer, B., & Cuisance, D. (2001). The changing distribution of two riverine tsetse flies over 15 years in an increasingly cultivated area of Burkina Faso. *Bulletin of Entomological Research*, *91*(3), 157–166. https://doi.org/10.1079/BER200190

Rodgers, T. W., Xu, C. C. Y., Giacalone, J., Kapheim, K. M., Saltonstall, K., Vargas, M., Yu, D. W., Somervuo, P., McMillan, W. O., & Jansen, P. A. (2017). Carrion fly-derived DNA metabarcoding is an effective tool for mammal surveys: Evidence from a known tropical mammal community. *Molecular Ecology Resources*, *17*(6), e133–e145. https://doi.org/10.1111/1755-0998.12701

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *2016*(10), e2584. https://doi.org/10.7717/PEERJ.2584/FIG-7

SAUNDERS, D. A., HOBBS, R. J., & MARGULES, C. R. (1991). Biological Consequences of Ecosystem Fragmentation: A Review. *Conservation Biology*, *5*(1), 18–32. https://doi.org/10.1111/J.1523-1739.1991.TB00384.X

Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R. D., Willerslev, E., Bertelsen, M. F., & Gilbert, M. T. P. (2012). Screening mammal biodiversity using DNA from leeches. *Current Biology*, *22*(8), R262–R263. https://doi.org/10.1016/J.CUB.2012.02.058

Siddig, A. A. H., Ellison, A. M., Ochs, A., Villar-Leeman, C., & Lau, M. K. (2016). How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in Ecological Indicators. *Ecological Indicators*, *60*, 223–230. https://doi.org/10.1016/J.ECOLIND.2015.06.036

Simarro, P. P., Jannin, J., & Cattand, P. (2008). Eliminating Human African Trypanosomiasis: Where Do We Stand and What Comes Next? *PLOS Medicine*, *5*(2), e55. https://doi.org/10.1371/JOURNAL.PMED.0050055

Strange, N., Theilade, I., Thea, S., Sloth, A., & Helles, F. (2007). Integration of species persistence, costs and conflicts: An evaluation of tree conservation strategies in Cambodia. *Biological Conservation*, *137*(2), 223–236. https://doi.org/10.1016/J.BIOCON.2007.02.006

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793. https://doi.org/10.1111/J.1365-294X.2012.05542.X

Taylor, P. G. (1996). Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Article in Molecular Biology and Evolution*. https://doi.org/10.1093/oxfordjournals.molbev.a025566

Torres, J., Brito, J. C., Vasconcelos, M. J., Catarino, L., Gonçalves, J., & Honrado, J. (2010). Ensemble models of habitat suitability relate chimpanzee (Pan troglodytes) conservation to forest and landscape dynamics in Western Africa. *Biological Conservation*, *143*(2), 416–425. https://doi.org/10.1016/J.BIOCON.2009.11.007

Tsetse Biology and Ecology: Their Role in the Epidemiology and Control of ... - Stephen G. A. Leak - Google Livros. (n.d.). Retrieved December 21, 2022, from https://books.google.pt/books?hl=pt-PT&lr=&id=J1latecW1BQC&oi=fnd&pg=PR16&dq=Leak,+S.G.A+(1999).+Tsetse+Biology+and+Ecology:+Their+Role+in+the+Epidemiology+and+control+of+Trypanosomosis.+CABI+publishing+inassociation+with+the+ILRI.+pp.152-210.&ots=-j5P8UEQfP&sig=JTnZZUOPufU3DqEs0OhPyQjFdz8&redir_esc=y#v=onepage&q&f=false

UNIVERSIDADE DA INTEGRAÇÃO INTERNACIONAL DA LUSOFONIA AFRO-BRASILEIRA INSTITUTO DE HUMANIDADES E LETRAS DOS MALÊS BACHARELADO EM HUMANIDADES HELCIM DANAMON FERNANDO FERNANDES A POLÍTICA AMBIENTAL E DESENVOLVIMENTO SUSTENTÁVELNA GUINÉ-BISSAU: UM ESTUDO SOBRE O INSTITUTO DA BIODIVERSIDADE DAS ÁREAS PROTEGIDAS (IBAP). (n.d.).

Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E., Coissac, E., Pompanon, F., Gielly, L., Cruaud, C., Nascetti, G., Wincker, P., Swenson, J. E., & Taberlet, P. (2009). New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Molecular Ecology Resources*, *9*(1), 51–60. https://doi.org/10.1111/J.1755-0998.2008.02352.X

Vector Control: Methods for Use by Individuals and Communities - Jan A. Rozendaal, World Health Organization Staff, World Health Organization, Organisation mondiale de la santé - Google Livros. (n.d.). Retrieved December 21, 2022, from https://books.google.pt/books?hl=pt-

PT&lr=&id=z6H5gwdddREC&oi=fnd&pg=PR9&dq=Rozendaal,+J.+A.+(1997).+Vector+control:+methods+for+use+by+individuals+and+communities.+World+Health+Organization.&ots=-zbCBkfkYu&sig=VQql5k_gn_06NkqKoQlnXprUD_Q&redir_esc=y#v=onepage&q=Rozendaal%2C%20J.%20A.%20(1997).%20Vector%20control%3A%20methods%20for%20use%20by%20individuals%20and%20communities.%20World%20Health%20Organization.&f=false

White, F. (1993). Refuge theory, ice-age aridity and the history of tropical biotas: an essay in plant geography. *Fragmenta Floristica et Geobotanica. Supplementum*, *02*(2).

Wilkie, D. S., & Carpenter, J. F. (1999). Bushmeat hunting in the Congo Basin: An assessment of impacts and options for mitigation. *Biodiversity and Conservation*, *8*(7), 927–955. https://doi.org/10.1023/A:1008877309871/METRICS

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergård, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Bråthen, K. A., Yoccoz, N., … Taberlet, P. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature 2014 506:7486*, *506*(7486), 47–51. https://doi.org/10.1038/nature12921

Wld, B., Org, H., & Weitz, B. (1963). The feeding habits of Glossina. *Bulletin of the World Health Organization*, *28*(5–6), 711. /pmc/articles/PMC2554947/?report=abstract

Attachments

Attachment 1

Table A1- Unfiltered taxa in 16S run and 12S run. The taxa with an asterisk (*) were considered field contamination, the taxa with two asterisks (**) were considered laboratory contamination, the taxa with three asterisks (***) were considered non-target amplifications.

Order	Family	OTU	16S		12S	
			No. OTUs	No. reads	No. OTUs.	No. reads.
		Chordata***	2	99	1	60
		Mammalia***	-	-	2	489
Artiodactyla	Bovidae	*Bos taurus*	1	226	1	6431
		Bovidae	1	301	1	193
		Ovis aries	1	923	2	4764
		*Bison sp.***	-	-	1	11
		Tragelaphus scriptus	2	120053	18	230558
	Suidae	*Phacochoerus africanus*	1	81229	3	124156
		Potamochoerus porcus	1	27595	1	69563
		Sus scrofa	3	123689	11	212747
		Artiodactyla***	-	-	2	5199
Carnivora	Canidae	*Canis familiaris*	-	-	5	29887
		*Vulpes vulpes***	-	-	3	6655
		Canis lupus	1	3793	-	-
	Felidae	*Felis catus****	-	-	5	6588
		*Lynx pardinus***	-	-	1	1546
	Hyaenidae	*Crocuta crocuta*	-	-	1	15
	Viverridae	*Civettictis civetta*	-	-	1	116
Perissodactyla	Equidae	*Equus asinus*	-	-	1	679
Primates	Hominidae	*Homo sapiens*	13	194193	56	802817
Rodentia	Hystricidae	*Hystrix sp.*	1	4213	-	-
		*Hystrix brachyura***	-	-	1	736
	Muridae	*Mus musculus*	-	-	1	30
	Thryonomyidae	*Thryonomys swinderianus*	1	550	1	4687
Lagomorpha	Leporidae	*Oryctolagus cuniculus***	1	380	5	17989
		Leporidae***	-	-	1	95
Soricomorpha	Talpidae	*Galemys pyrenaicus***	-	-	1	5631
		Aves***	2	42160	-	-
Galliformes	Phasianidae	*Gallus gallus*	1	56	2	65559
		Pavo cristatus	-	-	1	738
		*Alectoris rufa***	-	-	1	480
		Meleagris gallopavo	1	53	-	-
		Passeridae	-	-	2	11921
Pelecaniformes	Ardeidae	*Ardea cinerea*	1	24	1	12316
		Ardea purpurea	1	23	1	35410

		Ardea sp.	-	-	2	1635
		*Gorsachius sp.**	-	-	2	3204
Anseriformes	Anatidae	*Anas platyrhynchos**	-	-	1	581
Anura	Arthroleptidae	*Leptopelis sp.*	-	-	1	1812
	Hyperoliidae	*Hyperolius sp.*	-	-	1	880
	Phrynobatrachidae	*Phrynobatrachus francisci*	1	50	1	2340
Caudata	Salamandridae	*Salamandra salamandra***	-	-	1	4961
Testudines	Pelomedusidae	*Pelusios castaneus*	1	4410	2	59689
Crocodylia	Crocodylidae	*Crocodylus suchus*	1	1191	1	5553
		Crocodylus sp.				
Squamata	Elapidae	*Naja sp.*	2	143743	2	48313
	Pythonidae	Pythonidae	1	1793	-	-
		Python sp.	-	-	1	1433
	Varanidae	*Varanus niloticus*	1	701434	7	1002303
	Colubridae	*Coronella austriaca***	1	35	-	-
	Gekkonidae	Gekkonidae	-	-	2	104
		Actinopterygii**	2	35	3	1952
Carcharhiniformes	Scyliorhinidae	*Scyliorhinus canicula**	-	-	1	529
Characiformes	Hepsetidae	Hepsetidae*	-	-	1	821
		Charadriiformes***	-	-	1	93704
Cyprinodontiformes	Nothobranchiidae	*Epiplatys sp.***	3	3984	1	713
		Nothobranchiidae*	-	-	2	13490
		*Scriptaphyosemion sp.***	1	20	-	-
Perciformes	Anabantidae	*Ctenopoma acutirostre**	-	-	1	2406
	Cichlidae	*Oreochromis sp.***	1	242	2	961
		*Coptodon rheophila**	2	210	-	-
		Cichlidae*	-	-	1	1082
	Latidae	*Lates niloticus**	-	-	1	890
		*Lates sp.**	1	398	-	-
	Moronidae	*Dicentrarchus labrax***	-	-	1	1389
Siluriformes	Clariidae	*Clarias sp.***	-	-	1	570
	Claroteidae	*Chrysichthys nigrodigitatus**	-	-	1	145
		Claroteidae*	-	-	1	1035
	Schilbeidae	*Pareutropius sp.***	1	27	-	-
		Arthropoda***	-	-	1	25
		Insecta***	5	162	7	672
Diptera	Glossinidae	*Glossina fuscipes****	5	568	1	20166
		*Glossina morsitans****	1	54	-	-
		*Glossina****	5	645	4	1169
TOTAL			**70**	**1458561**	**189**	**2932593**

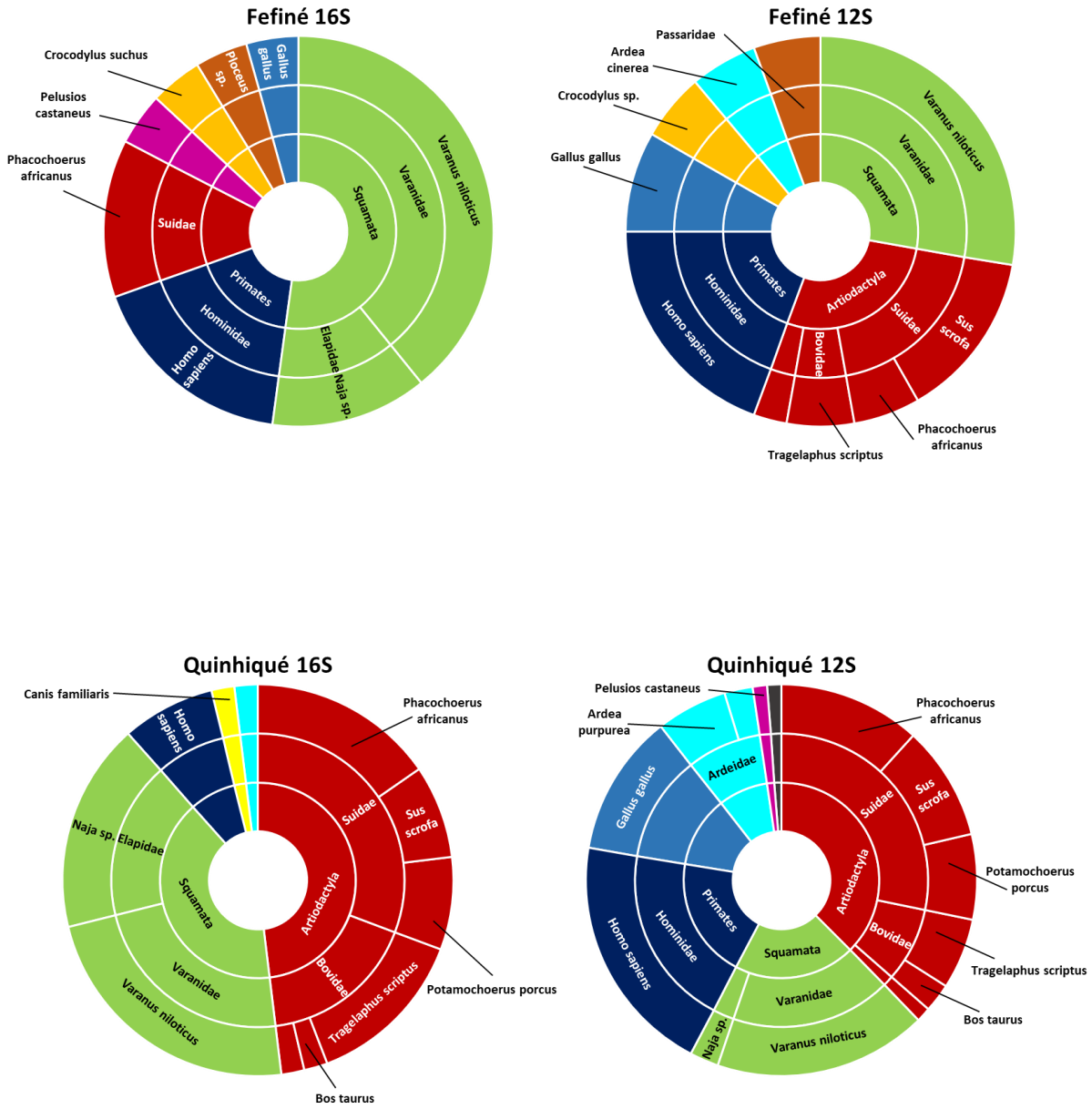


Figure A1 - Detected taxa in the small rivers (Quinhiqué and Saltinho) for both primers (16S and 12S)

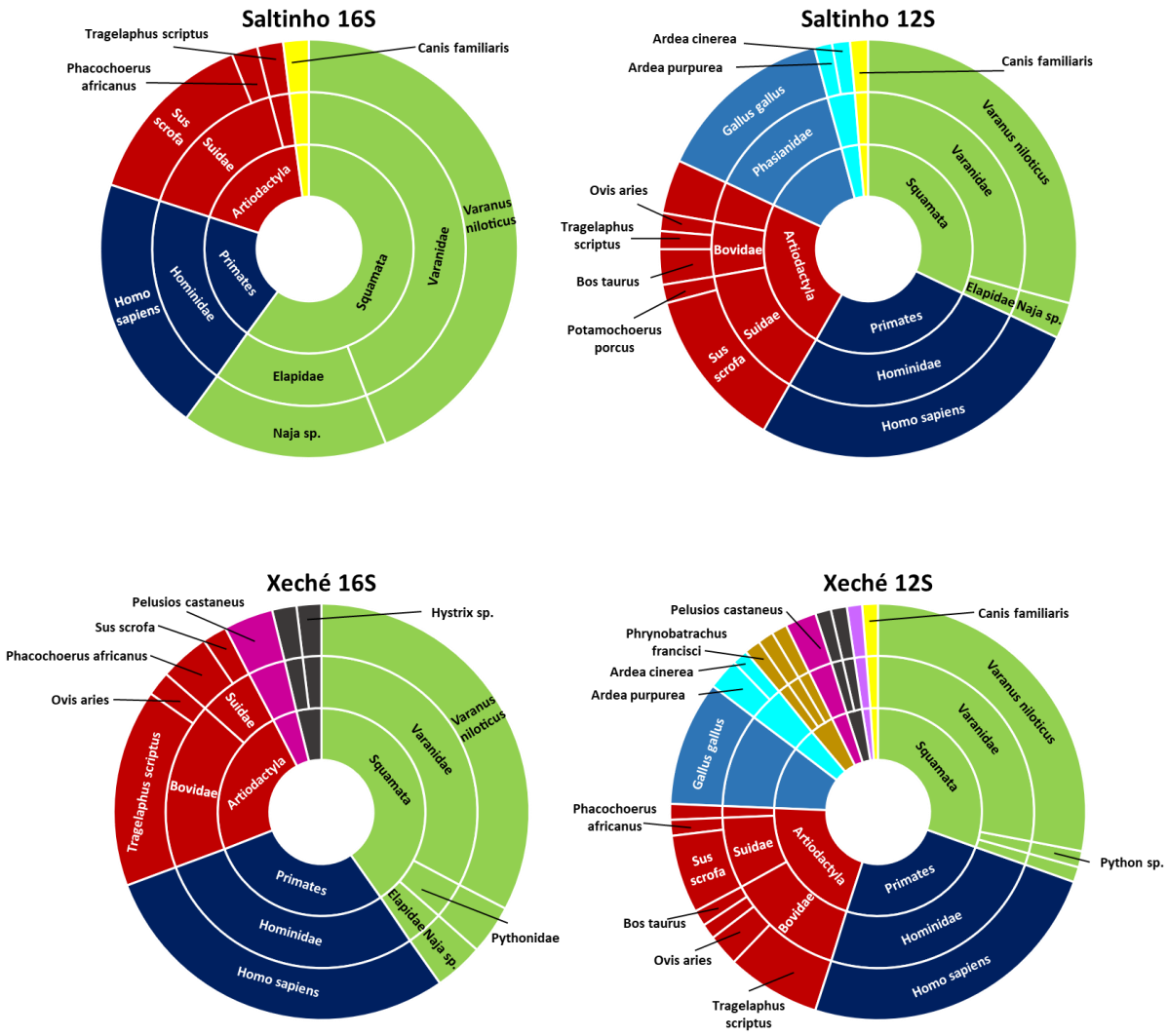


Figure A2 - Detected taxa in the big rivers (Saltinho and Xéché) for both primers (16S and 12S)