The application of non- and minimally-invasive biological and environmental samples in wildlife surveillance and genetic research

Inaugural Dissertation
to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

Juan Li

from Zigong of Sichuan province, China

Berlin, 2023

This dissertation was done at the Leibniz Institute for Zoo and Wildlife Research in Berlin from October 2019 to October 2023 under the supervision of Prof. Alex D. Greenwood. It is submitted to the Department of Biology at the Freie Universität Berlin.

1st Reviewer: Prof. Alex D. Greenwood, Ph.D.
2nd Reviewer: Prof. Dr. Dino McMahon

Date of defense: February 01, 2024

This dissertation is based on the following manuscripts

1. **Juan Li**, Peter Seeber, Jan Axtner, Rachel Crouthers, Milou Groenenberg, Arnulf Koehncke, Alexandre Courtiol, Pin Chanratana, Alex D. Greenwood, **Monitoring terrestrial wildlife by combining hybridization capture and metabarcoding data from waterhole environmental DNA**, *Biological Conservation*, Volume 284, 2023, 110168, ISSN 0006-3207, https://doi.org/10.1016/j.biocon.2023.110168.

Chapter 2

---

2. **Juan Li**, Gabriela Ulloa Urizar, Pedro Mayor, Meddly L. Santolalla Robles, Alex D. Greenwood. **Nucleic acid degradation after long term dried blood spot storage**. Submitted manuscript. (Under review in *Molecular Ecology Resources*)

Chapter 3

---

3. **Juan Li**, Pedro Mayor, Meddly L. Santolalla Robles, Alex D. Greenwood. **The complete mitochondrial genome of the lowland paca (*Cuniculus paca*) and its phylogenetic relationship with other New World hystricognath rodents**. *Mitochondrial DNA Part B: Resources*. Accepted manuscript. https://doi.org/10.1080/23802359.2023.2275830. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).

Chapter 4

DECLARATION OF INDEPENDENCE

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me. This dissertation has not yet been presented to any other examination authority in the same or a similar form and has not yet been published.

Juan Li

Berlin Germany, October 26, 2023

CONTENTS

DECLARATION OF INDEPENDENCE III
CONTENTS IV
SUMMARY 1
ZUSAMMENFASSUNG 4
CHAPTER 1: General Introduction 7
1 Wildlife conservation 7
1.1 Wildlife Diversity 7
1.2 Evidence-based conservation 8
1.3 Monitoring data 9
2 Terrestrial wildlife monitoring 10
2.1 Observation-based approaches 10
2.2 Environmental DNA (eDNA)-based approaches 11
3 Wildlife sample collection and application in genetic research 13
4 Wildlife genomic research 14
5 Infectious disease surveillance 15
6 Study aims 16
References 18
CHAPTER 2: Manuscript 1 33
Monitoring terrestrial wildlife by combining hybridization capture and metabarcoding data from waterhole environmental DNA 33
CHAPTER 3: Manuscript 2 51
Nucleic acid degradation after long term dried blood spot storage 51
Benefit-Sharing Statement 69
CHAPTER 4: Manuscript 3 81
The complete mitochondrial genome of the lowland paca (*Cuniculus paca*) and its phylogenetic relationship with other New World hystricognath rodents 81
CHAPTER 5: General Discussion 91
Acknowledgements 102
List of publications and author contributions 103
Curriculum Vitae 104

SUMMARY

Over the last few decades, increased anthropogenic activities lead to dramatic biodiversity declines across the globe, particularly for wildlife animals, requiring effective conservation and management. Long-term monitoring is of paramount importance in this regard, allowing for characterizing how and why a given ecosystem and wildlife inhabiting it are changing in response to environmental changes and conservation interventions applied. Considering the complexity of ecosystems, a variety of information is required for monitoring wildlife, such as population dynamics over time and across space, genetic variation, and disease transmission. However, the implementation of in-field wildlife monitoring and biological sample collection for genetic research is often hindered due to the logistic barriers and ethical issues, especially for rare and threatened species.

In terrestrial wildlife monitoring, observation-based approaches are the most commonly used techniques that are often time-consuming and labor-intensive. The recent application of camera traps based on automatic recording mechanism has considerably scaled up the global monitoring, while the monitoring remains to be very limited for rare, cryptic and elusive species, such as small body size or low-density animals. Recently, environmental DNA (eDNA) methods have been applied to aquatic species monitoring, allowing for species detection, abundance estimation and genetic variation surveys, even for rare species. They have also recently been applied to free-ranging terrestrial animals. Terrestrial mammals are more difficult to work with in some cases than aquatic animals due to difficulty in collecting representative samples. Furthermore, eDNA approaches thus far are primarily based on PCR-based metabarcoding that frequently targets a short DNA region due to high eDNA degradation. Such approaches can fall short for closely related or molecularly uncharacterized species. Target capture sequencing using hybridization capture may complement by selectively enriching a wide range of target sequences. In **chapter 2**, to evaluate the applicability and efficiency of eDNA methods in terrestrial wildlife monitoring, environmental samples were collected from waterholes in a preserved areas in Cambodia during the dry season. Metabarcoding targeting a 93-bp 16S rRNA gene for tetrapod species and hybridization capture targeting complete mitochondrial genome of various representative mammalian species were applied simultaneously to water (N =46) and sediment (N =10) samples. Thirty-five samples were positive by metabarcoding and 28 samples by hybridization capture. In total, two methods detected seventeen species: one bird, two reptiles, three amphibians and 11 mammals, which is relatively consistent to the results from concurrent camera traps. All non-mammal species were exclusively detected by hybridization capture. Metabarcoding specifically detected four mammals, three of which were small rodents. Metabarcoding tend to be more sensitive to species shedding large amounts of DNA, such as Asian elephant and wild boar. In contrast,

hybridization capture retrieved more genetic information, in some cases, providing over 70% of mitochondrial genome coverage. Our results suggest that sampling from waterhole during dry season can provide representative samples for terrestrial wildlife monitoring using eDNA methods. Integration of multiple eDNA methods may effectively improve the sensitivity and inclusivity of monitoring in terrestrial habitats.

In wildlife disease surveillance, possible vulnerability of animals and the remoteness of sampling areas pose severe challenges to collecting biological samples. Dried blood spots (DBS) technology has been extensively used in medical diagnosis particularly in low-resource areas, due to the ease of use and transport, and high sample stability, although it often comes with low quality of nucleic acid. Recent technological innovation in sequencing techniques allow for applying degraded biological samples for a variety of genetic and genomic research. However, the value of DBS sampling remains underestimated in wildlife research. A comprehensive quality assessment of nucleic acid on wildlife DBS samples could demonstrate their utility and limitations. DBS samples have been effectively used for disease testing using various amplification-based approaches, demonstrating preservation of both viral DNA and RNA. By combining with advanced high-throughput sequencing, DBS samples may allow for viral detection and novel virus discovery from wildlife. In **chapter 3**, I evaluated the DNA isolated from DBS samples stored on two types of filter paper (FTA cards and Whatman filter paper no.3) for up to 15 years using three distinct isolation methods in comparison with blood samples stored in a cold chain. Meanwhile, target capture sequencing using a bait set targeting various DNA and RNA vertebrate viruses was conducted on DBS samples of 12 primates (4 species) for viral screening. DNA from DBS samples degraded significantly regardless of the storage time and paper types compared cold chain, preserved samples with DNA integrity highly dependent on extraction methods applied. However, relatively large DNA fragments (~20,000 bp) were preserved over 9 years of storage. FTA cards showed higher capacity to preserve large DNA fragment when storage period was extended, while nearly 20,000 bp DNA fragments were preserved on Whatman filter paper no.3 after nearly 15 years of storage. Overall, the two types of filter paper are comparable in DBS sample storage could be valuable resources for wildlife genetic research. However, no viruses were detected from the 12 primate samples. With the non-target sequencing data, I was able to *de novo* assemble seven mitochondrial genomes for three primate species, suggesting that DBS samples can be significantly important for wildlife genetic and genomic research.

Mitochondrial genomes are a major reference target in wildlife genetic research. Non- or minimal-invasively collected samples represent a majority of genetic resources in this regard, particularly for species that biological samples are hard to access, such as threatened species. In **chapter 4**, I *de novo* assembled a mitochondrial genome of lowland paca (*Cuniculus paca*) using DBS samples preserved

for nearly two years. The newly assembled mitochondrial genome expands our knowledge of the phylogenetic relationships of New World hystricognath rodents and demonstrated the value of DBS samples in wildlife genomic research.

My results demonstrate that non-invasive environmental samples can be effectively used for terrestrial wildlife, particularly when multiple eDNA methods are combined. Minimal-invasive DBS samples can provide invaluable biological resources for wildlife genetic and genomic research even after a prolonged storage time. Both sampling strategies may greatly facilitate the data collection to improve wildlife conservation and management.

ZUSAMMENFASSUNG

In den letzten Jahrzehnten haben die zunehmenden anthropogenen Aktivitäten zu einem dramatischen Rückgang der biologischen Vielfalt auf der ganzen Welt geführt, insbesondere bei wildlebenden Tieren, die einer wirksamen Erhaltung und Bewirtschaftung bedürfen. In diesem Zusammenhang ist die Langzeitüberwachung von größter Bedeutung. Dadurch kann charakterisiert werden, wie und warum sich ein bestimmtes Ökosystem sowie die darin lebende Tierwelt als Reaktion auf Umweltveränderungen und Naturschutzmaßnahmen verändern. In Anbetracht der Komplexität von Ökosystemen sind für Wildtierüberwachung eine Vielzahl von Informationen erforderlich, z. B. über die Populationsdynamik im zeitlichen und räumlichen Verlauf, die genetische Variation und die Übertragung von Krankheiten. Die Durchführung der Wildtierüberwachung vor Ort und die Entnahme biologischer Proben für die genetische Forschung wird jedoch häufig durch logistische Hindernisse und ethische Fragen erschwert, insbesondere bei seltenen und bedrohten Arten.

Bei der Überwachung von an Land lebenden Wildtieren werden in der Regel Observationsmethoden eingesetzt, die oft zeit- und arbeitsintensiv sind. Die jüngste Anwendung von Kamerafallen, die auf einem automatischen Aufzeichnungsmechanismus basieren, hat die globale Überwachung erheblich ausgeweitet, während die Überwachung von seltenen, kryptischen und schwer fassbaren Arten, wie z. B. Tieren mit geringer Körpergröße oder geringer Dichte, nach wie vor sehr begrenzt ist. In jüngster Zeit wurden Umwelt-DNA-Methoden (eDNA) bei der Überwachung aquatischer Arten angewandt, die die Erkennung von Arten, die Schätzung der Abundanz und Erhebungen der genetischen Variation, selbst bei seltenen Arten, ermöglichen. Seit kurzem werden sie auch bei freilebenden Landtieren angewandt. An Land lebende Säugetiere sind aufgrund der Schwierigkeit, repräsentative Proben zu sammeln, in einigen Fällen schwieriger zu bearbeiten als Wassertiere. Darüber hinaus basieren die bisherigen eDNA-Ansätze in erster Linie auf PCR-basiertem Metabarcoding, das aufgrund des hohen eDNA-Abbaus häufig auf eine kurze DNA-Region abzielt. Solche Ansätze können bei eng verwandten oder molekular nicht charakterisierten Arten zu kurz greifen. Die Zielsequenzierung (Target Capture) mittels Hybridisierung kann durch selektive Anreicherung eines breiten Spektrums von Zielsequenzen eine Ergänzung darstellen. In Kapitel 2 wurden zur Bewertung der Anwendbarkeit und Effizienz von eDNA-Methoden bei der Überwachung von landlebenden Wildtieren Umweltproben aus Wasserlöchern in einem Naturschutzgebiet in Kambodscha während der Trockenzeit entnommen. Metabarcoding, das auf ein 93-bp-16S-rRNA-Gen für Tetrapodenarten abzielt, und das Hybridisierungsverfahren, welches auf das komplette mitochondriale Genom verschiedener repräsentativer Säugetierarten ausgerichtet ist, wurden gleichzeitig auf Wasser- (N =46) und Sedimentproben (N =10) angewendet. Fünfunddreißig Proben waren per Metabarcoding positiv und

28 Proben per Hybridisierung. Insgesamt wurden mit beiden Methoden siebzehn Arten nachgewiesen: ein Vogel, zwei Reptilien, drei Amphibien und 11 Säugetiere, was relativ deckungsleich den Ergebnissen der zeitgleich eingesetzten Kamerafallen entspricht. Alle Nicht-Säugetierarten wurden ausschließlich durch Hybridisierung ermittelt. Durch Metabarcoding wurden vier Säugetiere nachgewiesen, von denen drei kleine Nagetiere waren. Das Metabarcoding ist tendenziell empfindlicher für Arten, die große Mengen an DNA absondern, was beispielsweise beim asiatischen Elefanten und dem Wildschwein der Fall ist. Im Gegensatz dazu erbrachte die Hybridisierung mehr genetische Informationen, die in einigen Fällen über 70 % des mitochondrialen Genoms abdeckten. Unsere Ergebnisse deuten darauf hin, dass die Entnahme von Proben aus Wasserlöchern während der Trockenzeit repräsentative Proben für die Überwachung landlebender Wildtiere durch eDNA-Methoden liefern kann. Die Integration mehrerer eDNA-Methoden kann die Sensitivität und Inklusivität der Überwachung in terrestrischen Lebensräumen effektiv verbessern.

Bei der Überwachung von Wildtierkrankheiten stellen die mögliche Anfälligkeit der Tiere und die Abgelegenheit der Probenahmegebiete eine große Herausforderung für die Sammlung biologischer Proben dar. Die Technologie der getrockneten Blutflecken (DBS) wurde in der medizinischen Diagnostik, insbesondere in ressourcenarmen Gebieten, aufgrund der einfachen Anwendung und des einfachen Transports sowie der hohen Probenstabilität ausgiebig genutzt, obwohl die Qualität der Nukleinsäure oft mangelhaft ist. Jüngste technologische Innovationen im Bereich der Sequenzierungstechniken ermöglichen es, degradierte biologische Proben für eine Vielzahl von genetischen und genomischen Untersuchungen zu verwenden. Allerdings wird der Wert von DBS-Proben in der Wildtierforschung nach wie vor unterschätzt. Eine umfassende Qualitätsbewertung der Nukleinsäure in DBS-Proben von Wildtieren könnte deren Nutzen und Grenzen aufzeigen. DBS-Proben wurden mit verschiedenen Amplifikationsverfahren effektiv für Krankheitstests eingesetzt, wobei sowohl virale DNA als auch RNA erhalten blieben. Durch die Kombination mit fortschrittlicher Hochdurchsatz-Sequenzierung könnten DBS-Proben den Nachweis von Viren und die Entdeckung neuer Viren bei Wildtieren ermöglichen. In Kapitel 3 bewertete ich die auf zwei Typen von Filterpapier (FTA-Karten und Whatman-Filterpapier Nr. 3) bis zu 15 Jahre lang aufbewahrte DNA aus DBS-Proben unter Anwendung von drei verschiedenen Isolierungsmethoden im Vergleich zu in einer Kühlkette aufbewahrten Blutproben. In der Zwischenzeit wurde die Zielsequenzierung mittels eines Ködersatzes, der auf verschiedene DNA- und RNA-Viren von Wirbeltieren abzielt, an DBS-Proben von 12 Primaten (4 Arten) zum Virenscreening durchgeführt. Die DNA aus DBS-Proben wurde unabhängig von der Lagerungsdauer und den Papiertypen im Vergleich zur Kühlkette erheblich abgebaut, wobei die DNA-Integrität stark von den angewandten Extraktionsmethoden abhing. Allerdings blieben relativ große DNA-Fragmente (~20.000 bp) über 9 Jahre Lagerung erhalten. FTA-Karten zeigten eine höhere Kapazität zur Erhaltung großer DNA-Fragmente, wenn die Lagerungszeit

verlängert wurde, während fast 20.000 bp DNA-Fragmente auf Whatman-Filterpapier Nr. 3 nach fast 15 Jahren Lagerung erhalten blieben. Insgesamt sind die beiden Arten von Filterpapier bei der Lagerung von DBS-Proben vergleichbar und könnten wertvolle Ressourcen für die wildtiergenetische Forschung darstellen. In den 12 Primatenproben wurden jedoch keine Viren nachgewiesen. Mit den Daten der nicht-zielgerichteten Sequenzierung (Non-Target) war ich in der Lage, sieben mitochondriale Genome für drei Primatenarten *de novo* zu assemblieren, was darauf hindeutet, dass DBS-Proben für die wildtiergenetische und genomische Forschung von großer Bedeutung sein können.

Mitochondriale Genome sind ein wichtiges Referenzziel in der Wildtiergenetikforschung. Nicht- oder minimal-invasiv gesammelte Proben stellen in dieser Hinsicht einen Großteil der genetischen Ressourcen dar, insbesondere bei Arten, bei denen biologische Proben nur schwer zugänglich sind, wie etwa bei bedrohten Arten. In Kapitel 4 habe ich ein mitochondriales Genom des Tieflandpakas (Cuniculus paca) unter Verwendung von DBS-Proben, die fast zwei Jahre lang aufbewahrt wurden, *de novo* assembliert. Das neu assemblierte mitochondriale Genom erweitert unser Wissen über die phylogenetischen Beziehungen der hystricognathen Nagetiere der Neuen Welt und demonstriert den Wert von DBS-Proben für die Genomforschung an Wildtieren.

Meine Ergebnisse zeigen, dass nicht-invasive Umweltproben effektiv für landlebende Wildtiere verwendet werden können, insbesondere wenn mehrere eDNA-Methoden kombiniert werden. Minimal-invasive DBS-Proben können auch nach längerer Lagerungszeit unschätzbare biologische Ressourcen für die genetische und genomische Wildtierforschung liefern. Beide Probenahmestrategien können die Datenerhebung zur Verbesserung der Erhaltung und Bewirtschaftung von Wildtieren erheblich erleichtern.

CHAPTER 1: General Introduction

1 Wildlife conservation

1.1 Wildlife Diversity

Maintaining global biodiversity is of paramount importance to support the health and resilience of ecosystems, the stability of economies, and the well-being of humanity. Greater biodiversity increases the resilience of ecosystems against disturbances and climate extremes, which then guarantee the delivery of ecosystem services (e.g. food supply, seed dispersal and nutrient cycling) (Isbell et al. 2015, Oliver et al. 2015, Sakschewski et al. 2016). Wildlife biodiversity plays a multitude of roles in regulating the health and stability of a given ecosystem. Herbivorous populations modify plant community diversity and structure by their browsing and foraging behaviors, and meanwhile, serve as prey to balance predator-prey dynamics (Clasen et al. 2015, Apollonio et al. 2017). The presence of carnivores influences the population size, structure, distribution and evolutionary adaptability of herbivores, and on the other hand, control pathogen spillover within prey communities by reducing their population densities (Karanth et al. 2004, Gray 2012, Buenavista and Palomares 2018, Say-Sallaz et al. 2023).

Despite the importance of biodiversity, global biodiversity is currently experiencing dramatic declines at a speed that is far greater than historical records in human history (Pimm et al. 2014, Ceballos et al. 2015, Ceballos et al. 2017). Given the known species losses over the last few centuries, the rate and magnitude of biodiversity declines may be comparable with mass extinctions that have occurred five times in the past 540 million years of the Earth history (Hooper et al. 2012, Dirzo et al. 2014, Ceballos et al. 2017). It is estimated that one million species may go extinct in the next few decades (Bongaarts 2019). Human-driven impacts are reported to be the most significant effects for the current biodiversity losses; wildlife have been considered the most profound victims. WWF Living Planet Report 2022 (WWF, 2022; https://livingplanet.panda.org/) shows that globally monitored wildlife declined an average of 69% in the relative abundance from 1970 to 2018. The severest global decline was observed in freshwater species as reflected by 83% of population decline. Latin America experienced the worst regional decline of losing 94% wildlife abundance. The situation is likely to be underestimated considering the many more undiscovered species (Pimm et al. 2014), which may not be documented before they vanish. Therefore, well-established conservation strategies are urgently needed to curb biodiversity declines.

Progress of species loss or extinction is mitigatable when conservation efforts have been made and deployed effectively (Simon et al. 2012, Mueller et al. 2020, Ghazi et al. 2021, Jhala et al. 2021). As a result of long-term practical conservation measures, the population status of Asian crested ibis (*Nipponia nippon*) has recovered from Critically Endangered in 1981 (7 individual observations) to Endangered in 2000, with 330 recorded mature individuals in China as assessed by IUCN in 2018 (Lin 1981). In the 1970s, India initiated the Tiger Conservation Project by implementing various management measures, such as setting up wildlife reserve areas, and restoring forest and prey communities (Jhala et al. 2021). The number of the iconic tiger species (*Panthera tigris*) in India has doubled from 1,411 in 2006 with an estimate of 3,167 individuals by 2022, the largest global wild tiger population (Jhala et al. 2021, Lamba et al. 2023). The monitoring data obtained over the course of intervention and policy implementation provide important guidance for the effectiveness evaluation and improvement of conservation management.

1.2 Evidence-based conservation

The concept of conservation can be dated back to the 17th century in the work of John Evelyn, who highlighted the importance of conserving forest resources for sustainable utilization. Conservation movement took off after the first state forest management project launched in the mid-19th century given the increased deforestation, while early conservation interventions were mainly for the sake of sustainably maintaining forest resources (i.e. timber or wood). Wildlife became the forefront victims along with growing human-driven changes in ecosystems (Hooper et al. 2012, Dirzo et al. 2014, Ceballos et al. 2017). Numerous wildlife species have undergone serious population declines even to the brink of extinction or extirpation (Hooper et al. 2012, Dirzo et al. 2014, Ceballos et al. 2017). Cambodia, which was known as “Asia’s Serengeti”, lost significant amounts of wildlife over the past few decades (Loucks et al. 2009, Gray et al. 2020, Groenenberg 2020). Many local wild ungulates and carnivore populations continue to decline even after protected areas were designated (Duckworth 1998, Steinmetz et al. 2010, Groenenberg 2020, Heinrich et al. 2020, Ladd et al. 2022). A charismatic herbivorous species giant panda (*Ailuropoda melanoleuca*) used to inhabit most regions in Southeast Asia, such as China, northern Vietnam, and northern Myanmar (Schaller 1993). Due to large-scale habitat destruction, the distribution of wild giant pandas shrank rapidly in the 20th century (Schaller 1993, Liu et al. 2005, Kong et al. 2017). From 1936 to 1998, the wild populations can only be traced in a few regions of China, with 25 geologically fragmented groups each with less than 20 individuals (Schaller 1993, Zang et al. 2017). Since the early 20th century, global wildlife conservation and management have been launched particularly in the regions with high biodiversity (e.g. Southern Asia, Africa). However, those preliminary managements were primarily species-focused, in some cases, leading to biased protection that failed to restore the entire ecosystem (Kong et al. 2017, Gray et al. 2020, Jhala et al. 2021). More recently, evidence-based conservation has been brought to forefront

with the aim of entire ecosystem restoration, which requires scientific evaluation based on dynamic monitoring.

1.3 Monitoring data

In conservation management, decisions on interventions or policies to implement are based on information that reflect the intricate and intertwined nature of biodiversity and ecosystem dynamics. Additionally, limited conservation resources require a cost-effective deployment to maximize the corresponding impacts, which rely on information from in-field surveillance. Long-term monitoring thus is essential to improve our understanding of ecological, genetic, social, and economic factors that shape ecosystems, and characterize how and why they are changing in response to environmental changes or applied conservation measures (Hughes et al. 2017). Given the intricacy of ecosystems, a variety of information needs to be gathered over the course of monitoring, such as species presence and absence, population size and distribution, genetic diversity, kinship structure, and cross-species disease transmission.

Monitoring population dynamics and trends provides fundamental knowledge to evaluating the health and conservation states of wildlife communities in a given ecosystem. For instance, a biodiversity baseline can be established using population data to gauge the management impacts. Numerous wildlife-vehicle collisions occur annually all around the world due to geographical fragmentation (Santos et al. 2018, Schwartz et al. 2020, Praill et al. 2023), requiring informed interventions based on collective data relevant to wildlife distribution, behaviors and activities. Invasive species control for ecosystem restoration relies on information relevant to species presence and abundance to develop proactive management strategies (Robertson et al. 2017, Ward et al. 2020, Duenas et al. 2021).

Decades of theoretical, empirical and data-simulative studies suggest that genetic variation strongly correlates with species adaptability (Fitzpatrick et al. 2020, Hohenlohe et al. 2021, Kardos et al. 2021). Genetic diversity and the associated adaptive capacity to evolve contribute to improved population-level fitness in response to environmental disturbance and selection (Spielman et al. 2004). Decreased genetic diversity results in the reduction in population growth and viability (Bowman 1960, Pacioni et al. 2015, Bull et al. 2016, Bozzuto et al. 2019), in some cases, increases the vulnerability of species to extinction or local extirpation (Saccheri et al. 1998, Wright et al. 2007). When genetic diversity is high in populations, the health of even small and isolated populations of threatened species can be improved (Fitzpatrick et al. 2020, Mueller et al. 2020, Figueiredo et al. 2021). The advances in sequencing technologies and bioinformatic techniques allow analyses that were impossible with traditional methods for more comprehensive genetic monitoring. Recently, the genetic monitoring data has been practically used to assess wildlife vulnerability and extinction risk and to provide guidance

for conservation and management (Forester et al. 2022). For instance, multi-year genetic investigations confirmed that inbreeding and loss of genetic diversity were responsible for the low success rate in lynx reintroductions (Linnell et al. 2009, Bull et al. 2016). The regional genetic assessment thereafter was used to guide the intervention of captive breeding to reinforce their genetic diversity (Simon et al. 2012). Thus, gathering and evaluating genetic information over time and across space is of importance to develop optimized conservation strategies for long-term wildlife population management.

Wildlife population declines are often initiated by diverse extrinsic non-genetic factors. Infectious diseases are increasingly being identified as a major risk factor, particularly for endangered species (Daszak et al. 2000, Smith et al. 2009). In some cases, disease can lead to significant amounts of species loss in a short period of time, such as facial tumor disease in Tasmanian devils (*Sarcophilus harrisi*) (Hawkins et al. 2006, McCallum et al. 2007). After its discovery in 1996 (Mozos et al. 1996), the disease rapidly swept the wild devil populations (Hawkins et al. 2006). McCallum et al. 2007 estimated that no wild individuals will be disease-free in 5 years without intervention and the entire population may die out within 15 years. A viral disease in rabbits, haemorrhagic disease, was first documented in 1984 in China (Liu et al. 1984, Xu and Chen 1989). In a few years, its swift spread has seriously threatened wild and domestic rabbits across the globe (O'Hara 2006, Schwensow et al. 2014, Hu et al. 2021, Tanikawa et al. 2023), and even indirectly affected European lynx reintroductions by decreasing prey populations (Figueiredo et al. 2021). On the other hand, pathogen spillover can pose severe threats to public health, husbandry and farming industry, and the global economy. Thus, disease surveillance and pathogen diagnosis in wildlife are necessary for both animals and human welfare.

2 Terrestrial wildlife monitoring

2.1 Observation-based approaches

Wildlife censuses have been conducted for centuries across different continents. For terrestrial animal monitoring, observation-based approaches have been commonly used, such as direct observations, vocalization monitoring, track and trail analysis, records from locals and line transects (Voss and Emmons 1996, Voss et al. 2001, Trolle 2003, Mendes Pontes 2004, Haugaasen and Peres 2005). The inventorying and monitoring data provides broad biological and ecological knowledge of terrestrial wildlife in distinct landscapes. However, those aforementioned approaches are labor-intensive and time-consuming, and often fail to gather information for elusive, cryptic and rare species (Voss and Emmons 1996), particularly in hard-to-access areas such as remote tropical jungles. Continued anthropogenic changes escalate biodiversity declines. Along with the rapid deforestation, urbanization, and hunting and poaching, terrestrial wildlife lost a significant amount of their diversity in the recent

few decades, emphasizing the need for continuous monitoring. In the 1980s, camera traps were developed to automatically photograph terrestrial animals in the field (Champion 1992, Griffiths and van Schaik 1993, Nichols et al. 1998, Cutler and Swann 1999, Tobler et al. 2013). The real-time recording technique has been widely applied to terrestrial animal monitoring, providing important information relevant to species presence and distribution, population size and dispersal, and animal behaviors (Karanth and Nichols 1998, Silveira et al. 2003, Maffei et al. 2005, Karanth et al. 2006). More recently, various new technologies have been integrated into camera trap protocols prior to or post data collection (Schneider et al. 2019, Corcoran et al. 2021), further advancing terrestrial wildlife monitoring.

Despite a growing understanding of wildlife dynamics due to the improvement of observation-based approaches, terrestrial animal monitoring and the associated conservation research are still limited to a small group of wildlife taxa (Johnson et al. 2017). In general, large and medium body size terrestrial animals are preferentially targeted, attributed to their suitability of being recorded by camera traps and other observational methods. In contrast, amphibians and reptiles so far are considerably under-investigated due to their cryptic and elusive nature, as are many other small mammals (MacKenzie et al. 2002, Raemy and Ursenbacher 2018, Matthias et al. 2021). Of note, more amphibian species (41%) are listed as threatened than all other vertebrates (mammals, birds, reptiles and fish) in the IUCN Red list as of the time of writing (2023); 21% reptiles are considered threatened, higher than birds (13%). The information deficit caused by technical limitations may severely underestimate the species loss and affect the effectiveness of conservation management, requiring complementary monitoring to fill in the data gaps.

2.2 Environmental DNA (eDNA)-based approaches

DNA-based approaches have been developed for terrestrial wildlife monitoring after several studies reported successfully retrieving DNA from non-invasive biological samples (e.g. feces, hair) (Hoss et al. 1992, Morin et al. 1992, Taberlet and Bouvet 1992). Despite being effective in population estimation, those non-invasive methods are generally limited to single-species detection (Gray et al. 2014). Given the rate and magnitude of biodiversity decline, current large-scale monitoring required in conservation management can benefit from developing genetic methods for multi-taxa detection across landscapes. In recent years, application of eDNA technology has increasingly gained attention in terrestrial wildlife monitoring due to the potential to assess an assemblage of taxa in a habitat without direct observation and contact (Beng and Corlett 2020, Andres et al. 2023).

eDNA represents a mixture of DNA isolated from environmental samples (Pawlowski et al. 2020), (Rodriguez-Ezpeleta et al. 2021). The initial application of eDNA was developed to determine

bacterial communities in sediment independently from culturing specimens (Ogram et al. 1987). In the early 2000s, eDNA approaches started to be used for detecting macro-organisms in freshwater systems (Martellini et al. 2005, Bhadury et al. 2006). With the advent of high-throughput sequencing, eDNA technology has been expanded to almost all types of aquatic ecosystems (Ficetola et al. 2008, Beng and Corlett 2020). The developed methodological workflows have been applied to biodiversity characterization and conservation surveillance (Alexander et al. 2019, Sales et al. 2020, Sales et al. 2021, Andres et al. 2023). More recently, non-invasive eDNA strategies have been extended to terrestrial habitats (Harper et al. 2019, Seeber et al. 2019, Lyet et al. 2021, Mena et al. 2021).

Genetic material released from terrestrial animals can enter and persist in their surrounding environments, which creates a bio-archive with a plethora of genetic information of taxa in the terrestrial habitat. One of the very first eDNA applications to detect the presence of terrestrial animals was published by Andersen et al. in 2012. Their results demonstrate that short mtDNA fragments of local large mammals can be amplified from soil DNA (Andersen et al. 2012). A growing number of studies suggest that eDNA methods can provide comparable results to camera traps (Ushio et al. 2017, Lyet et al. 2021, Mena et al. 2021, Farrell et al. 2022). In some cases, eDNA methods can perform better than traditional observation-based surveys, such as for cryptic, elusive and small animals (Deiner et al. 2016, Valentini et al. 2016, Harper et al. 2019). In this context, various sample types have been tested, such as environmental substrates (e.g. water, soil and air), and invertebrates-derived biological resources (e.g. haematophagous invertebrates) (Axtner et al. 2019). Recent progress demonstrates that eDNA approaches hold the great promise of comprehensive and indiscriminate monitoring of free-ranging terrestrial animals.

In eDNA-based surveys, species presence detection is often the main target of biomonitoring, primarily relying on metabarcoding methods, as a consequence of the low proportion of organismal DNA in eDNA (Seeber and Epp 2022). For taxonomic identification, eDNA metabarcoding can provide relatively high resolution, in some cases, even providing species relative abundance estimates (Pilliod et al. 2013, Axtner et al. 2019, Andres et al. 2023), despite generally being confined to target a short region of DNA sequences due to degradation typical of eDNA. In recent years, the amount of genetic information that could be obtained from eDNA approaches, has increased (Jensen et al. 2021, Andres et al. 2023). However, species classification based on short DNA regions is difficult for closely related or molecularly uncharacterized species. Besides, environmental contaminants in eDNA tend to increase the risk of PCR inhibition (Seeber and Epp 2022), leading to false negatives, particularly for low-density threatened species that shed less DNA. In contrast, target capture sequencing presents an alternative tool to enrich sequences of interests without requiring high cycles of amplification using hybridization capture based on specifically designed DNA or RNA baits

(Sigsgaard et al. 2020). The technique can flexibly choose target sequences depending on study purposes (Teer and Mullikin 2010, Ali et al. 2016) and tolerate fragmented genetic materials in small amounts (Hodges et al. 2007, Gnirke et al. 2009). The post-enrichment sequencing strategy has been widely tested and applied in ancient DNA research (Enk et al. 2014). Recent work demonstrates that hybridization capture applied to target marker genes, can identify more species than PCR-based eDNA approaches (Giebner et al. 2020) and reliably reflect relative species abundance (Wilcox et al., 2018) to survey aquatic ecosystems. Jensen et al 2021 used target capture to retrieve complete mitochondrial genome and various nuclear loci of target marine species from aquatic DNA samples. Seeber et al (2019) demonstrated that multi-taxa terrestrial free-ranging animals can be detected from waterhole environmental samples by applying a bait panel designed from multiple representative mammalian species. However, there is thus far a scarcity of studies to compare the different eDNA approaches in terrestrial animal monitoring and how much genetic information can be obtained for population genetic study purposes.

3 Wildlife sample collection and application in genetic research

Multiple noninvasive invasive biological samples, such as feces, hairs, droppings, saliva and urine can be sampled for DNA analysis (Solberg et al. 2006, De Barba et al. 2010, Andrews 2018, Carvajal-Agudelo et al. 2021). However, small amounts of starting materials are generally insufficient to provide complete and high-quality genomic results, which often occurs when using non-invasively collected biological samples. Well-preserved tissue or blood samples currently remain the standard resources to generate reliable baseline genetic data and genomic references that are fundamental to wildlife ecology and evolution research. In this context, collection, storage and transportation of biological samples from the field are required but challenging.

In genetic studies, frozen storage is desired (Camacho-Sanchez et al. 2013, Carvajal-Agudelo et al. 2021), which can effectively block enzymatic and environmental effects that occur shortly after sampling to avoid DNA/RNA strand breakage (Paabo 1989, Hofreiter 2001, Paabo et al. 2004). However, during fieldwork, it is often impossible to access cold-chain storage and transportation due to logistical barriers and the remoteness of sampling areas. Various ambient-preservation approaches thus have been developed to simplify the sample preservation in the field until samples are transported to a laboratory for further analysis, such as buffer (e.g. RNAlater, ethanol) (Riesgo et al. 2012, Camacho-Sanchez et al. 2013, Precioso et al. 2022) and filter paper (e.g. cellulose filter paper) storage (Jones et al. 2012, Molteni et al. 2013). Although degradation continues under the protection of preservation treatment, whole genome sequencing can still succeed after long term storage (Agrawal et al. 2021, Precioso et al. 2022).

Dried blood spots (DBS) stored on filter paper provide a powerful preservation tool due to the ease of collection, sample stability and minimal invasiveness (McClendon-Weary et al. 2020). DBS sampling acquires a small volume of blood applied onto filter paper. After drying at room temperature, the filter paper serves as a matrix to preserve blood in a stable and compact format, easing the field transportation. The preservation capacity of DBS technology was first discovered for phenylketonuria diagnosis in infants by metabolite inspection (Guthrie and Susi 1963, Pitt 2010, van Vliet et al. 2020). A wide range of biological components was also reported to be detectable from DBS samples, such as proteins, lipids, and nucleic acids (Lehmann et al. 2013, McClendon-Weary et al. 2020). Currently, DBS sampling has been well-established and integrated into the fields of clinics and medicine for screening rare genetic disease in newborns and detecting infectious disease in resource-limited areas, biobanking, and various medical and clinical studies (Demirev 2013, Lim 2018, McClendon-Weary et al. 2020, Russcher et al. 2020, Tuaillon et al. 2020). The resistance to harsh environmental conditions and long-term stability make DBS technologies highly suitable for wildlife research settings. However, the applicability of DBS technology thus far is under-studied, particularly its application in the field of wildlife genetic and genomic research (Samsonova et al. 2022a).

Flinders Technology Associates (FTA) cards are cellulose-based paper that is chemically treated to optimize the preservation of dried biological samples for DNA-oriented research, such as DBS samples (Hsiao et al. 1999, Smith and Burgoyne 2004). FTA cards were reported to preserve DNA samples for up to 16 years for the application of target amplicon amplification (Rahikainen et al. 2016). Blood samples from deceased patients were found to be efficiently preserved on FTA cards and suitable for DNA profiling analysis (Green et al. 2019). Despite the capacity to preserve valuable genetic materials, FTA cards were less frequently used in early wildlife research due to the high material expense than basic cellulose paper without chemical treatment. The samples preserved on this less-optimal filter paper often represent a majority of archived wildlife biological samples that may still hold significant research value. A comprehensive assessment of how well DNA is stored on distinct storage cards could provide important guidance for genomic DNA analyses based upon, especially for the samples preserved for substantial time periods.

4 Wildlife genomic research

Although genomic reference sequences of wildlife species have increased in number dramatically over the past few years, many wildlife species are missing in some cases due to lack of availability of biological samples. Such data deficiencies may hinder the development of many research areas and the associated monitoring applications. For instance, the wide application of eDNA approaches recently has helped to simplify wildlife monitoring in various habitats, while those approaches strongly rely on genomic reference databases to guarantee a reliable species identification and to allow population

genetic estimation. Recent work suggests that DBS samples can be used for genomic applications to provide comparable results to those obtained from well-preserved whole blood (Bassaganyas et al. 2018, Agrawal et al. 2021). In some cases, DBS is even applicable to RNA sequencing for transcriptome profiling (Bybjerg-Grauholm et al. 2017, Wu et al. 2019). However, to date, little research has been carried out to estimate the applicability of wildlife DBS samples in genomic sequencing.

Whole genome analysis has grown extensively for both model and non-model organisms, while mitochondrial genomes are still recognized as a backbone in many research fields, such as evolutionary genetics, phylogenetics and conservation genetics (Schmidt et al. 2017, Cardeñosa et al. 2020, Flores-Manzanero et al. 2021). This is particularly true in wildlife genetic research and conservation that often involve low-quality biological samples (Orkin et al. 2016, Arandjelovic and Vigilant 2018). Compared to nuclear DNA, multiple-copies mitochondrial DNA (mtDNA) tends to be more resistant during the period of long-term storage even in harsh environmental conditions (Bogenhagen and Clayton 1978, Schwarz et al. 2009, Allentoft et al. 2012), making mtDNA the most accessible and economic source of genetic data in wildlife research. DBS samples from wildlife can also be used to expand the taxonomic list of mitochondrial genome reference databases, providing baseline data in wildlife conservation management. For example, mitochondrial markers are particularly useful for species identification to allow individual-scale monitoring, and estimating population structure genetic diversity and inbreeding events (Serrao et al. 2021, Skovrind et al. 2021, Phukuntsi et al. 2022).

5 Infectious disease surveillance

Infectious diseases pose severe threats to both wildlife and humans. Most human viral pathogens are originally wildlife pathogens (Morse et al. 2012, Rosenberg et al. 2013, Rosenberg 2015). The discovery of zoonotic pathogens can predate their emergence (Rosenberg 2015). For example, Ebola (WHO, 2014) and Zika (Duffy et al. 2009) viruses both caused epidemics decades after their first identification. Foot and mouth disease was first identified in 1879 from livestock (Loeffler F 1897) but was recorded to infect humans until nearly 60 years later (Vetterlein 1954). Disease surveillance data can identify potential risks posed by viruses and direct epidemiological analysis to the areas where spillovers are suspected to occur (Bernstein et al. 2022). In the past, viral pathogen surveillance in wild animals has often relied on passive screening given the technical challenges to carry out large-scale surveillance (Schilling et al. 2022). This surveillance strategy can be highly biased and limited to a small set of etiological agents (Anthony et al. 2017, Schilling et al. 2022), leading to compromised disease prevention and control for both animal and human populations. In recent years, the introduction of high-throughput sequencing technology speeds up the discovery of novel viruses at a

rate that was impossible with culture-dependent approaches (Zhang et al. 2019, Bernstein et al. 2022). However, the aim to scale up pathogen surveillance in wild animal populations is still restricted by accessing and collecting biological samples. Minimal-invasive DBS samples can provide invaluable resources in this regard.

The application of DBS technology in clinics facilitates infectious disease screening in rural and remote areas, primarily in combination with pathogen-targeted techniques, such as PCR and serological assays (Cardona-Ospina et al. 2019, Samsonova et al. 2022b). DNA and RNA viruses remain detectable despite DBS samples being stored for an extended storage time in substandard conditions (Perozo et al. 2006, Narayanan et al. 2010, Abdelwhab et al. 2011, Awad et al. 2014, Jóźwiak et al. 2016). The availability of more powerful sequencing methods allows the use of low-quality biological samples for novel viral detections and viral diversity analysis (de Carvalho Ferreira et al. 2014, Bandoo et al. 2021, Hardmeier et al. 2021, Schilling et al. 2022). Anthony et al. (2015) identified 184 viruses in 14 viral families from fecal samples of wild macaques. A recent metagenomic study identified diverse viruses from 39 viral families in bat stool and nearly assembled a complete coronavirus with high similarity to MERS-CoV (Hardmeier et al. 2021). However, thus far, there are very few or no studies to assess the application of DBS samples on viral diversity characterization of wildlife animals.

Viral genetic materials in DBS constitute a small amount of total isolated nucleic acid unless a high titre viral infection occurs in animals when sampling is conducted. Over the course of storage, RNA, if any, presumably should be in a highly degraded state, requiring target enrichment to perform screening. A recent study demonstrates that a variety of known and unknown mammalian viruses can be detected from environmental and invertebrate samples by target capture sequencing (Alfano et al. 2021), showing 100-fold enrichment of targeted sequences in some cases. DBS samples resemble the degradation state of eDNA, and may thus benefit from the selective hybridization capture for viral targeted screening.

6 Study aims

The aim of this thesis is to investigate how non- and minimal-invasively collected biological and environmental samples can facilitate terrestrial wildlife monitoring and genetic research. In wildlife research, a particular challenge of using non-invasive sampling is to deal with low quality and low quantity samples, requiring methodological improvement for efficient and accurate genetic information retrieval. Thus, this study aims for providing guidelines of wildlife biological, evolutionary and monitoring research using low-quality noninvasive or minimally invasive biological samples.

To address the above questions, I conducted three studies:

1. In a methodological study (Chapter 2), I compared the efficacy of applying two eDNA methods in terrestrial wildlife detection: hybridization capture using a bait panel targeting complete mitochondrial genomes of a range of mammal species, and metabarcoding using a pair of vertebrate-specific primers targeting 93-bp 16S rRNA gene.
2. In a study (Chapter 3) on the stability assessment of nucleic acid in DBS samples, I evaluated the quality of nucleic acid isolated from wildlife DBS samples stored in two filter papers using three isolation methods. A set of RNA oligonucleotide baits was used for viral screening. Multiple bioinformatic analysis pipelines have been implemented for optimizing data interpretation from the perspectives of genetic and evolution, and pathogen surveillance.
3. A phylogenetic study (Chapter 4) was conducted to construct the evolutionary history of *Cuniculus paca* based on the mitochondrial genome directly assembled from next-generation sequencing data using low-quality DBS samples.

References

Abdelwhab, E. M., D. Lüschow, T. C. Harder, and H. M. Hafez. 2011. The use of FTA® filter papers for diagnosis of avian influenza virus. J Virol Methods **174**:120-122.

Agrawal, P., S. Katragadda, A. K. Hariharan, V. G. Raghavendrachar, A. Agarwal, R. Dayalu, D. Awasthy, S. C. Sharma, Y. K. Sivasamy, P. Lakshmana, A. Shanmugam, V. Veeramachaneni, V. Gupta, B. P. Vani, L. Subaiya, T. S. Syamala, R. Hariharan, V. Chandru, and D. E. Bloom. 2021. Validation of whole genome sequencing from dried blood spots. Bmc Medical Genomics **14**.

Alexander, J. B., M. Bunce, N. White, S. P. Wilkinson, A. A. S. Adam, T. Berry, M. Stat, L. Thomas, S. J. Newman, L. Dugal, and Z. T. Richards. 2019. Development of a multi-assay approach for monitoring coral diversity using eDNA metabarcoding. Coral Reefs **39**:159-171.

Alfano, N., A. Dayaram, J. Axtner, K. Tsangaras, M.-L. Kampmann, A. Mohamed, S. T. Wong, M. T. P. Gilbert, A. Wilting, and A. D. Greenwood. 2021. Non-invasive surveys of mammalian viruses using environmental DNA. **n/a**.

Ali, O. A., S. M. O'Rourke, S. J. Amish, M. H. Meek, G. Luikart, C. Jeffres, and M. R. Miller. 2016. RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping. Genetics **202**:389-400.

Allentoft, M. E., M. Collins, D. Harker, J. Haile, C. L. Oskam, M. L. Hale, P. F. Campos, J. A. Samaniego, M. T. Gilbert, E. Willerslev, G. Zhang, R. P. Scofield, R. N. Holdaway, and M. Bunce. 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. Proc Biol Sci **279**:4724-4733.

Andersen, K., K. L. Bird, M. Rasmussen, J. Haile, H. Breuning-Madsen, K. H. Kjaer, L. Orlando, M. T. Gilbert, and E. Willerslev. 2012. Meta-barcoding of 'dirt' DNA from soil reflects vertebrate biodiversity. Mol Ecol **21**:1966-1979.

Andres, K. J., D. M. Lodge, S. A. Sethi, and J. Andres. 2023. Detecting and analysing intraspecific genetic variation with eDNA: From population genetics to species abundance. Mol Ecol **32**:4118-4132.

Andrews, K. R., deBarba, M., Russello, M. A., & Waits, L. P. 2018. Advances in using non-invasive, archival, and environmental samples for population genomic studies. In O. P. Rajora (Ed.), Population Genomics: Wildlife. Springer Nature Switzerland AG.

Anthony, S. J., C. K. Johnson, D. J. Greig, S. Kramer, X. Y. Che, H. Wells, A. L. Hicks, D. O. Joly, N. D. Wolfe, P. Daszak, W. Karesh, W. I. Lipkin, S. S. Morse, J. A. K. Mazet, T. Goldstein, and P. Consortium. 2017. Global patterns in coronavirus diversity. Virus Evolution **3**.

Apollonio, M., V. V. Belkin, J. Borkowski, O. I. Borodin, T. Borowik, F. Cagnacci, A. A. Danilkin, P. I. Danilov, A. Faybich, F. Ferretti, J. M. Gaillard, M. Hayward, P. Heshtaut, M. Heurich, A. Hurynovich, A. Kashtalyan, G. I. H. Kerley, P. Kjellander, R. Kowalczyk, A. Kozorez, S.

Matveytchuk, J. M. Milner, A. Mysterud, J. Ozolins, D. V. Panchenko, W. Peters, T. Podgorski, B. Pokorny, C. M. Rolandsen, V. Ruusila, K. Schmidt, T. P. Sipko, R. Veeroja, P. Velihurau, and G. Yanuta. 2017. Challenges and science-based implications for modern management and conservation of European ungulate populations. Mammal Research **62**:209-217.

Arandjelovic, M., and L. Vigilant. 2018. Non-invasive genetic censusing and monitoring of primate populations. Am J Primatol **80**:e22743.

Awad, F., R. Chhabra, M. Baylis, and K. Ganapathy. 2014. An overview of infectious bronchitis virus in chickens. World's Poultry Science Journal **70**:375-384.

Axtner, J., A. Crampton-Platt, L. A. Horig, A. Mohamed, C. C. Y. Xu, D. W. Yu, and A. Wilting. 2019. An efficient and robust laboratory workflow and tetrapod database for larger scale environmental DNA studies. GigaScience **8**.

Bandoo, R. A., J. Bautista, M. Lund, E. Newkirk, J. Squires, A. Varsani, and S. Kraberger. 2021. Identification of novel circovirus and anelloviruses from wolverines using a non-invasive faecal sampling approach. Infect Genet Evol **93**:104914.

Bassaganyas, L., G. Freedman, D. Vaka, E. Wan, R. Lao, F. Chen, M. Kvale, R. J. Currier, J. M. Puck, and P. Y. Kwok. 2018. Whole exome and whole genome sequencing with dried blood spot DNA without whole genome amplification. Human Mutation **39**:167-171.

Beng, K. C., and R. T. Corlett. 2020. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. Biodiversity and Conservation **29**:2089-2121.

Bernstein, A. S., A. W. Ando, T. Loch-Temzelides, M. M. Vale, B. V. Li, H. Li, J. Busch, C. A. Chapman, M. Kinnaird, K. Nowak, M. C. Castro, C. Zambrana-Torrelio, J. A. Ahumada, L. Xiao, P. Roehrdanz, L. Kaufman, L. Hannah, P. Daszak, S. L. Pimm, and A. P. Dobson. 2022. The costs and benefits of primary prevention of zoonotic pandemics. Sci Adv **8**:eabl4183.

Bhadury, P., M. C. Austen, D. T. Bilton, P. J. D. Lambshead, A. D. Rogers, and G. R. Smerdon. 2006. Molecular detection of marine nematodes from environmental samples: overcoming eukaryotic interference. Aquatic Microbial Ecology **44**:97-103.

Bogenhagen, D., and D. A. Clayton. 1978. Mechanism of mitochondrial DNA replication in mouse L-cells: introduction of superhelical turns into newly replicated molecules. J Mol Biol **119**:69-81.

Bongaarts, J. 2019. IPBES, 2019. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Population and Development Review **45**:680-681.

Bowman, W. D. 1960. Area-Resource Development*. Transactions of the New York Academy of Sciences **22**:410-414.

Bozzuto, C., I. Biebach, S. Muff, A. R. Ives, and L. F. Keller. 2019. Inbreeding reduces long-term growth of Alpine ibex populations. Nat Ecol Evol **3**:1359-1364.

Buenavista, S., and F. Palomares. 2018. The role of exotic mammals in the diet of native carnivores from South America. Mammal Review **48**:37-47.

Bull, J. K., M. Heurich, A. P. Saveljev, K. Schmidt, J. Fickel, and D. W. Förster. 2016. The effect of reintroductions on the genetic variability in Eurasian lynx populations: the cases of Bohemian–Bavarian and Vosges–Palatinian populations. Conservation Genetics **17**:1229-1234.

Bybjerg-Grauholm, J., C. M. Hagen, S. K. Khoo, M. L. Johannesen, C. S. Hansen, M. Baekvad-Hansen, M. Christiansen, D. M. Hougaard, and M. V. Hollegaard. 2017. RNA sequencing of archived neonatal dried blood spots. Molecular Genetics and Metabolism Reports **10**:33-37.

Camacho-Sanchez, M., P. Burraco, I. Gomez-Mestre, and J. A. Leonard. 2013. Preservation of RNA and DNA from mammal samples under field conditions. Mol Ecol Resour **13**:663-673.

Cardeñosa, D., A. T. Fields, E. Babcock, S. K. H. Shea, K. A. Feldheim, D. W. Kraft, M. Hutchinson, M. A. Herrera, S. Caballero, and D. D. Chapman. 2020. Indo-Pacific origins of silky shark fins in major shark fin markets highlights supply chains and management bodies key for conservation. Conservation Letters **14**.

Cardona-Ospina, J. A., M. F. Villalba-Miranda, L. A. Palechor-Ocampo, L. I. Mancilla, and J. C. Sepulveda-Arias. 2019. A systematic review of FTA cards (R) as a tool for viral RNA preservation in fieldwork: Are they safe and effective? Preventive Veterinary Medicine **172**.

Carvajal-Agudelo, J. D., M. P. Trujillo-Betancur, D. Velásquez-Guarín, H. E. Ramírez-Chaves, J. E. Pérez-Cárdenas, and F. A. Rivera-Páez. 2021. Field blood preservation and DNA extraction from wild mammals: methods and key factors for biodiversity studies. Revista U.D.C.A Actualidad & Divulgación Científica **24**.

Ceballos, G., P. R. Ehrlich, A. D. Barnosky, A. Garcia, R. M. Pringle, and T. M. Palmer. 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. Sci Adv **1**:e1400253.

Ceballos, G., P. R. Ehrlich, and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proc Natl Acad Sci U S A **114**:E6089-E6096.

Champion, F. W. 1992. With a Camera in Tiger-Land. Catto and Windus,London, England, United Kingdom.

Clasen, C., M. Heurich, L. Glaesener, E. Kennel, and T. Knoke. 2015. What factors affect the survival of tree saplings under browsing, and how can a loss of admixed tree species be forecast? Ecological Modelling **305**:1-9.

Corcoran, E., M. Winsen, A. Sudholz, and G. Hamilton. 2021. Automated detection of wildlife using drones: Synthesis, opportunities and constraints. Methods in Ecology and Evolution **12**:1103-1114.

Cutler, T. L., and D. E. Swann. 1999. Using Remote Photography in Wildlife Ecology: A Review. Wildlife Society Bulletin (1973-2006) **27**:571-581.

Daszak, P., A. A. Cunningham, and A. D. Hyatt. 2000. Emerging infectious diseases of wildlife--threats to biodiversity and human health. Science **287**:443-449.

De Barba, M., L. P. Waits, E. O. Garton, P. Genovesi, E. Randi, A. Mustoni, and C. Groff. 2010. The power of genetic monitoring for studying demography, ecology and genetics of a reintroduced brown bear population. Mol Ecol **19**:3938-3951.

de Carvalho Ferreira, H. C., E. Weesendorp, S. Quak, J. A. Stegeman, and W. L. Loeffen. 2014. Suitability of faeces and tissue samples as a basis for non-invasive sampling for African swine fever in wild boar. Vet Microbiol **172**:449-454.

Deiner, K., E. A. Fronhofer, E. Machler, J.-C. Walser, and F. Altermatt. 2016. Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nature Communications **7**:1-9.

Demirev, P. A. 2013. Dried blood spots: analysis and applications. Anal Chem **85**:779-789.

Dirzo, R., H. S. Young, M. Galetti, G. Ceballos, N. J. Isaac, and B. Collen. 2014. Defaunation in the Anthropocene. Science **345**:401-406.

Duckworth, J. W. H., S. . 1998. Tracking tigers : a review of the status of tiger, Asian elephant, gaur, and banteng in Vietnam, Lao, Cambodia, and Yunnan (China), with recommendations for future conservation action. WWF Indochina, Hanoi, Vietnam **282pp**.

Duenas, M. A., D. J. Hemming, A. Roberts, and H. Diaz-Soltero. 2021. The threat of invasive species to IUCN-listed critically endangered species: A systematic review. Global Ecology and Conservation **26**.

Duffy, M. R., T. H. Chen, W. T. Hancock, A. M. Powers, J. L. Kool, R. S. Lanciotti, M. Pretrick, M. Marfel, S. Holzbauer, C. Dubray, L. Guillaumot, A. Griggs, M. Bel, A. J. Lambert, J. Laven, O. Kosoy, A. Panella, B. J. Biggerstaff, M. Fischer, and E. B. Hayes. 2009. Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med **360**:2536-2543.

Enk, J. M., A. M. Devault, M. Kuch, Y. E. Murgha, J. M. Rouillard, and H. N. Poinar. 2014. Ancient whole genome enrichment using baits built from modern DNA. Mol Biol Evol **31**:1292-1294.

Farrell, M. J., D. Govender, M. Hajibabaei, M. van der Bank, and T. J. Davies. 2022. Environmental DNA as a management tool for tracking artificial waterhole use in savanna ecosystems. Biological Conservation **274**.

Ficetola, G. F., C. Miaud, F. Pompanon, and P. Taberlet. 2008. Species detection using environmental DNA from water samples. Biology Letters **4**:423-425.

Figueiredo, A. M., L. M. de Carvalho, M. J. P. Gonzalez, R. T. Torres, S. Pla, J. C. Nunez-Arjona, C. Rueda, N. Vallverdu-Coll, F. Silvestre, J. Pena, D. Carmena, M. A. Habela, R. Calero-Bernal, C. Fonseca, and F. Najera. 2021. Parasites of the Reintroduced Iberian Lynx (Lynx pardinus) and Sympatric Mesocarnivores in Extremadura, Spain. Pathogens **10**.

Fitzpatrick, S. W., G. S. Bradburd, C. T. Kremer, P. E. Salerno, L. M. Angeloni, and W. C. Funk. 2020. Genomic and Fitness Consequences of Genetic Rescue in Wild Populations. Curr Biol **30**:517-522 e515.

Flores-Manzanero, A., D. Valenzuela-Galván, A. D. Cuarón, and E. Vázquez-Domínguez. 2021. Conservation genetics of two critically endangered island dwarf carnivores. Conservation Genetics **23**:35-49.

Forester, B. R., E. A. Beever, C. Darst, J. Szymanski, and W. C. Funk. 2022. Linking evolutionary potential to extinction risk: applications and future directions. Frontiers in Ecology and the Environment **20**:507-515.

Ghazi, M. G., S. P. Sharma, C. Tuboi, S. Angom, T. Gurumayum, P. Nigam, and S. A. Hussain. 2021. Population genetics and evolutionary history of the endangered Eld's deer (Rucervus eldii) with implications for planning species recovery. Sci Rep **11**:2564.

Giebner, H., K. Langen, S. J. Bourlat, S. Kukowka, C. Mayer, J. J. Astrin, B. Misof, and V. G. Fonseca. 2020. Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. Mol Ecol Resour **20**:1333-1345.

Gnirke, A., A. Melnikov, J. Maguire, P. Rogov, E. M. LeProust, W. Brockman, T. Fennell, G. Giannoukos, S. Fisher, C. Russ, S. Gabriel, D. B. Jaffe, E. S. Lander, and C. Nusbaum. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol **27**:182-189.

Gray, T. N. E. 2012. Studying Large Mammals With Imperfect Detection: Status and Habitat Preferences of Wild Cattle and Large Carnivores in Eastern Cambodia. Biotropica **44**:531-536.

Gray, T. N. E., M. J. Grainger, and R. Grosu. 2020. Conservation decision-making under uncertainty: Identifying when to reintroduce tiger Panthera tigris to Cambodia. Conservation Science and Practice **2**.

Gray, T. N. E., T. N. C. Vidya, S. Potdar, D. K. Bharti, and P. Sovanna. 2014. Population size estimation of an Asian elephant population in eastern Cambodia through non-invasive mark-recapture sampling. Conservation Genetics **15**:803-810.

Green, H., A. Tillmar, G. Pettersson, and K. Montelius. 2019. The use of FTA cards to acquire DNA profiles from postmortem cases. Int J Legal Med **133**:1651-1657.

Griffiths, M., and C. P. van Schaik. 1993. The Impact of Human Traffic on the Abundance and Activity Periods of Sumatran Rain Forest Wildlife. Conservation Biology **7**:623-626.

Groenenberg, M. C., R. Yoganand, K. 2020. Population status of ungulates in the Eastern Plains landscape. Srepok Wildlife Sanctuary and Phnom Prich Wildlife Sanctuary, Cambodia. Phnom Penh, Cambodia: WWF-Cambodia.

Guthrie, R., and A. Susi. 1963. A Simple Phenylalanine Method for Detecting Phenylketonuria in Large Populations of Newborn Infants. Pediatrics **32**:338-343.

Hardmeier, I., N. Aeberhard, W. Qi, K. Schoenbaechler, H. Kraettli, J. M. Hatt, C. Fraefel, and J. Kubacki. 2021. Metagenomic analysis of fecal and tissue samples from 18 endemic bat species in Switzerland revealed a diverse virus composition including potentially zoonotic viruses. PLoS One **16**:e0252534.

Harper, L. R., L. L. Handley, A. I. Carpenter, M. Ghazali, C. Di Muri, C. J. Macgregor, T. W. Logan, A. Law, T. Breithaupt, D. S. Read, A. D. McDevitt, and B. Hanfling. 2019. Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. Biological Conservation **238**:108225.

Haugaasen, T., and C. A. Peres. 2005. Primate assemblage structure in Amazonian flooded and unflooded forests. Am J Primatol **67**:243-258.

Hawkins, C. E., C. Baars, H. Hesterman, G. J. Hocking, M. E. Jones, B. Lazenby, D. Mann, N. Mooney, D. Pemberton, S. Pyecroft, M. Restani, and J. Wiersma. 2006. Emerging disease and population decline of an island endemic, the Tasmanian devil Sarcophilus harrisii. Biological Conservation **131**:307-324.

Heinrich, S., J. V. Ross, T. N. Gray, S. Delean, N. Marx, and P. J. B. C. Cassey. 2020. Plight of the commons: 17 years of wildlife trafficking in Cambodia. **241**:108379.

Hodges, E., Z. Xuan, V. Balija, M. Kramer, M. N. Molla, S. W. Smith, C. M. Middle, M. J. Rodesch, T. J. Albert, G. J. Hannon, and W. R. McCombie. 2007. Genome-wide in situ exon capture for selective resequencing. Nat Genet **39**:1522-1527.

Hofreiter, M. 2001. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Research **29**:4793-4799.

Hohenlohe, P. A., W. C. Funk, and O. P. Rajora. 2021. Population genomics for wildlife conservation and management. Molecular Ecology **30**:62-82.

Hooper, D. U., E. C. Adair, B. J. Cardinale, J. E. Byrnes, B. A. Hungate, K. L. Matulich, A. Gonzalez, J. E. Duffy, L. Gamfeldt, and M. I. O'Connor. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. Nature **486**:105-108.

Hoss, M., M. Kohn, S. Paabo, F. Knauer, and W. Schroder. 1992. Excrement analysis by PCR. Nature **359**:199.

Hsiao, K.-M., H.-M. Lin, H. Pan, T.-C. Li, S.-S. Chen, S.-B. Jou, Y.-L. Chiu, M.-F. Wu, C.-C. Lin, and S.-Y. Li. 1999. Application of FTA sample collection and DNA purification system on the

determination of CTG trinucleotide repeat size by PCR-based southern blotting. Journal of Clinical Laboratory Analysis **13**:188-193.

Hu, B., H. J. Wei, Z. Y. Fan, Y. H. Song, M. M. Chen, R. L. Qiu, W. F. Zhu, W. Z. Xu, J. B. Xue, and F. Wang. 2021. Emergence of rabbit haemorrhagic disease virus 2 in China in 2020. Veterinary Medicine and Science **7**:236-239.

Hughes, B. B., R. Beas-Luna, A. K. Barner, K. Brewitt, D. R. Brumbaugh, E. B. Cerny-Chipman, S. L. Close, K. E. Coblentz, K. L. de Nesnera, S. T. Drobnitch, J. D. Figurski, B. Focht, M. Friedman, J. Freiwald, K. K. Heady, W. N. Heady, A. Hettinger, A. Johnson, K. A. Karr, B. Mahoney, M. M. Moritsch, A.-M. K. Osterback, J. Reimer, J. Robinson, T. Rohrer, J. M. Rose, M. Sabal, L. M. Segui, C. Shen, J. Sullivan, R. Zuercher, P. T. Raimondi, B. A. Menge, K. Grorud-Colvert, M. Novak, and M. H. Carr. 2017. Long-Term Studies Contribute Disproportionately to Ecology and Policy. BioScience **67**:271-281.

Isbell, F., D. Craven, J. Connolly, M. Loreau, B. Schmid, C. Beierkuhnlein, T. M. Bezemer, C. Bonin, H. Bruelheide, E. de Luca, A. Ebeling, J. N. Griffin, Q. Guo, Y. Hautier, A. Hector, A. Jentsch, J. Kreyling, V. Lanta, P. Manning, S. T. Meyer, A. S. Mori, S. Naeem, P. A. Niklaus, H. W. Polley, P. B. Reich, C. Roscher, E. W. Seabloom, M. D. Smith, M. P. Thakur, D. Tilman, B. F. Tracy, W. H. van der Putten, J. van Ruijven, A. Weigelt, W. W. Weisser, B. Wilsey, and N. Eisenhauer. 2015. Biodiversity increases the resistance of ecosystem productivity to climate extremes. Nature **526**:574-577.

Jensen, M. R., E. E. Sigsgaard, S. Liu, A. Manica, S. S. Bach, M. M. Hansen, P. R. Moller, and P. F. Thomsen. 2021. Genome-scale target capture of mitochondrial and nuclear environmental DNA from water samples. Mol Ecol Resour **21**:690-702.

Jhala, Y., R. Gopal, V. Mathur, P. Ghosh, H. S. Negi, S. Narain, S. P. Yadav, A. Malik, R. Garawad, and Q. Qureshi. 2021. Recovery of tigers in India: Critical introspection and potential lessons. People and Nature **3**:281-293.

Johnson, C. N., A. Balmford, B. W. Brook, J. C. Buettel, M. Galetti, L. Guangchun, and J. M. Wilmshurst. 2017. Biodiversity losses and conservation responses in the Anthropocene. Science **356**:270-275.

Jones, S., C. J. Sutherland, C. Hermsen, T. Arens, K. Teelen, R. Hallett, P. Corran, M. van der Vegte-Bolmer, R. Sauerwein, C. J. Drakeley, and T. Bousema. 2012. Filter paper collection of Plasmodium falciparum mRNA for detecting low-density gametocytes. Malaria Journal **11**.

Jóźwiak, M., K. Wyrostek, K. Domańska-Blicharz, M. Olszewska-Tomczyk, K. Śmietanka, and Z. Minta. 2016. Application of FTA® Cards for detection and storage of avian influenza virus. Journal of Veterinary Research **60**:1-6.

Karanth, K. U., and J. D. Nichols. 1998. Estimation of tiger densities in India using photographic captures and recaptures. Ecology **79**:2852-2862.

Karanth, K. U., J. D. Nichols, N. S. Kumar, and J. E. Hines. 2006. Assessing Tiger Population Dynamics Using Photographic Capture–Recapture Sampling. Ecology **87**:2925-2937.

Karanth, K. U., J. D. Nichols, N. S. Kumar, W. A. Link, and J. E. J. P. o. t. N. A. o. S. Hines. 2004. Tigers and their prey: predicting carnivore densities from prey abundance. **101**:4854-4858.

Kardos, M., E. E. Armstrong, S. W. Fitzpatrick, S. Hauser, P. W. Hedrick, J. M. Miller, D. A. Tallmon, and W. C. Funk. 2021. The crucial role of genome-wide genetic variation in conservation. Proc Natl Acad Sci U S A **118**.

Kong, L. Q., W. H. Xu, L. Zhang, M. H. Gong, Y. Xiao, and Z. Y. Ouyang. 2017. Habitat conservation redlines for the giant pandas in China. Biological Conservation **210**:83-88.

Ladd, R., R. Crouthers, S. Brook, and J. C. Eames. 2022. Reviewing the status and demise of the Endangered Eld's deer and identifying priority sites and conservation actions in Cambodia. Mammalia **86**:407-421.

Lamba, A., H. C. Teo, R. Sreekar, Y. Zeng, L. R. Carrasco, and L. P. Koh. 2023. Climate co-benefits of tiger conservation. Nat Ecol Evol **7**:1104-1113.

Lehmann, S., C. Delaby, J. Vialaret, J. Ducos, and C. Hirtz. 2013. Current and future use of "dried blood spot" analyses in clinical chemistry. Clinical Chemistry and Laboratory Medicine **51**:1897-1909.

Lim, M. D. 2018. Dried Blood Spots for Global Health Diagnostics and Surveillance: Opportunities and Challenges. Am J Trop Med Hyg **99**:256-265.

Lin, Y. Z. 1981. RECOVERY OF JAPANESE CRESTED IBIS IN QIN-LING RANGE. Acta Zoologica Sinica **27**:273-273.

Linnell, J. D. C., U. Breitenmoser, C. Breitenmoser-Würsten, J. Odden, and M. von Arx. 2009. Recovery of Eurasian Lynx in Europe: What Part has Reintroduction Played? Pages 72-91 Reintroduction of Top-Order Predators.

Liu, S., H. Xue, B. Pu, and N. Qian. 1984. A new viral disease in rabbits. Animal Husbandry and Veterinary Medicine (Xumu yu Shouyi) **16**:253-255.

Liu, X. H., A. G. Toxopeus, A. K. Skidmore, X. M. Shao, G. D. Dang, T. J. Wang, and H. H. T. Prins. 2005. Giant panda habitat selection in Foping Nature Reserve, China. Journal of Wildlife Management **69**:1623-1632.

Loeffler F, F. P. 1897. Berichte der Kommission zur Erforschung der Maul-und Klauenseuche bei dem Institut für Infektionskrankheiten in Berlin. Zent Bakt Parasitkde Abt I **23**:371–391.

Loucks, C., M. B. Mascia, A. Maxwell, K. Huy, K. Duong, N. Chea, B. Long, N. Cox, and T. Seng. 2009. Wildlife decline in Cambodia, 1953-2005: exploring the legacy of armed conflict. Conservation Letters **2**:82-92.

Lyet, A., L. Pellissier, A. Valentini, T. Dejean, A. Hehmeyer, and R. Naidoo. 2021. eDNA sampled from stream networks correlates with camera trap detection rates of terrestrial mammals. Sci Rep **11**:11362.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm. 2002. Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One. Ecology **83**:2248-2255.

Maffei, L., A. J. Noss, E. Cuéllar, and D. I. Rumiz. 2005. Ocelot (Felis pardalis) population densities, activity, and ranging behaviour in the dry forests of eastern Bolivia: data from camera trapping. Journal of Tropical Ecology **21**:349-353.

Martellini, A., P. Payment, and R. Villemur. 2005. Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water Res **39**:541-548.

Matthias, L., M. J. Allison, C. Y. Maslovat, J. Hobbs, and C. C. Helbing. 2021. Improving ecological surveys for the detection of cryptic, fossorial snakes using eDNA on and under artificial cover objects. Ecological Indicators **131**.

McCallum, H., D. M. Tompkins, M. Jones, S. Lachish, S. Marvanek, B. Lazenby, G. Hocking, J. Wiersma, and C. E. Hawkins. 2007. Distribution and Impacts of Tasmanian Devil Facial Tumor Disease. Ecohealth **4**:318-325.

McClendon-Weary, B., D. L. Putnick, S. Robinson, and E. Yeung. 2020. Little to Give, Much to Gain-What Can You Do With a Dried Blood Spot? Curr Environ Health Rep **7**:211-221.

Mena, J. L., H. Yagui, V. Tejeda, E. Bonifaz, E. Bellemain, A. Valentini, M. W. Tobler, P. Sanchez-Vendizu, and A. Lyet. 2021. Environmental DNA metabarcoding as a useful tool for evaluating terrestrial mammal diversity in tropical forests. Ecol Appl **31**:e02335.

Mendes Pontes, A. R. 2004. Ecology of a community of mammals in a seasonailly dry forest in Roraima, Brazilian Amazon. Mammalian Biology **69**:319-336.

Molteni, C. G., L. Terranova, A. Zampiero, C. Galeone, N. Principi, and S. Esposito. 2013. Comparison of manual methods of extracting genomic DNA from dried blood spots collected on different cards: implications for clinical practice. International Journal of Immunopathology and Pharmacology **26**:779-783.

Morin, P. A., J. J. Moore, and D. S. Woodruff. 1992. Identification of chimpanzee subspecies with DNA from hair and allele-specific probes. Proc Biol Sci **249**:293-297.

Morse, S. S., J. A. K. Mazet, M. Woolhouse, C. R. Parrish, D. Carroll, W. B. Karesh, C. Zambrana-Torrelio, W. I. Lipkin, and P. Daszak. 2012. Zoonoses 3 Prediction and prevention of the next pandemic zoonosis. Lancet **380**:1956-1965.

Mozos, E., A. Mendez, J. C. Gomez-Villamandos, J. Martin De Las Mulas, and J. Perez. 1996. Immunohistochemical characterization of canine transmissible venereal tumor. Vet Pathol **33**:257-263.

Mueller, S. A., T. E. Reiners, T. L. Middelhoff, O. Anders, A. Kasperkiewicz, and C. Nowak. 2020. The rise of a large carnivore population in Central Europe: genetic evaluation of lynx reintroduction in the Harz Mountains. Conservation Genetics **21**:577-587.

Narayanan, M., M. Parthiban, P. Sathiya, K. Kathaperumal, and K. Narayanan. 2010. Molecular detection of Newcastle disease virus using Flinders Tehnology Associates-PCR. Veterinarski Arhiv **80**.

Nichols, J., B. Zevnik, K. Anastassiadis, H. Niwa, D. Klewe-Nebenius, I. Chambers, H. Scholer, and A. Smith. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell **95**:379-391.

O'Hara, P. 2006. The illegal introduction of rabbit haemorrhagic disease virus in New Zealand. Revue Scientifique Et Technique-Office International Des Epizooties **25**:119-123.

Ogram, A., G. S. Sayler, and T. Barkay. 1987. The extraction and purification of microbial DNA from sediments. Journal of Microbiological Methods **7**:57-66.

Oliver, T. H., M. S. Heard, N. J. B. Isaac, D. B. Roy, D. Procter, F. Eigenbrod, R. Freckleton, A. Hector, C. D. L. Orme, O. L. Petchey, V. Proenca, D. Raffaelli, K. B. Suttle, G. M. Mace, B. Martin-Lopez, B. A. Woodcock, and J. M. Bullock. 2015. Biodiversity and Resilience of Ecosystem Functions. Trends Ecol Evol **30**:673-684.

Orkin, J. D., Y. Yang, C. Yang, D. W. Yu, and X. Jiang. 2016. Cost-effective scat-detection dogs: unleashing a powerful new tool for international mammalian conservation biology. Sci Rep **6**:34758.

Paabo, S. 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A **86**:1939-1943.

Paabo, S., H. Poinar, D. Serre, V. Jaenicke-Despres, J. Hebler, N. Rohland, M. Kuch, J. Krause, L. Vigilant, and M. Hofreiter. 2004. Genetic analyses from ancient DNA. Annu Rev Genet **38**:645-679.

Pacioni, C., H. Hunt, M. E. Allentoft, T. G. Vaughan, A. F. Wayne, A. Baynes, D. Haouchar, J. Dortch, and M. Bunce. 2015. Genetic diversity loss in a biodiversity hotspot: ancient DNA quantifies genetic decline and former connectivity in a critically endangered marsupial. Mol Ecol **24**:5813-5828.

Pawlowski, J., L. Apotheloz-Perret-Gentil, and F. Altermatt. 2020. Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. Mol Ecol **29**:4258-4264.

Perozo, F., P. Villegas, C. Estevez, I. Alvarado, and L. B. Purvis. 2006. Use of FTA filter paper for the molecular detection of Newcastle disease virus. Avian Pathol **35**:93-98.

Phukuntsi, M. A., D. L. Dalton, M. Mwale, J. Selier, T. Cebekhulu, and M. T. Sethusa. 2022. Genetic patterns in three South African specialist antelope species: Threats, conservation management and their implications. Conservation Science and Practice **4**.

Pilliod, D. S., C. S. Goldberg, R. S. Arkle, and L. P. Waits. 2013. Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. Canadian Journal of Fisheries and Aquatic Sciences **70**:1123-1130.

Pimm, S. L., C. N. Jenkins, R. Abell, T. M. Brooks, J. L. Gittleman, L. N. Joppa, P. H. Raven, C. M. Roberts, and J. O. Sexton. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. Science **344**:1246752.

Pitt, J. J. 2010. Newborn screening. Clin Biochem Rev **2**:57-68.

Praill, L. C., T. M. Eppley, S. Shanee, P. M. K. Cunneyworth, F. D. Abra, N. Allgas, H. Al-Razi, M. Campera, S. M. Cheyne, W. Collinson, G. Donati, B. Linden, S. Manson, M. Maria, T. Q. Morcatty, K. A. I. Nekaris, L. I. Oklander, V. Nijman, and M. S. Svensson. 2023. Road Infrastructure and Primate Conservation: Introducing the Global Primate Roadkill Database. Animals **13**.

Precioso, M., M. Molina-Morales, D. A. Dawson, T. A. Burke, and J. G. Martinez. 2022. Effects of long-term ethanol storage of blood samples on the estimation of telomere length. Evolutionary Ecology **36**:915-931.

Raemy, M., and S. Ursenbacher. 2018. Detection of the European pond turtle (Emys orbicularis) by environmental DNA: is eDNA adequate for reptiles? Amphibia-Reptilia **39**:135-143.

Rahikainen, A. L., J. U. Palo, W. de Leeuw, B. Budowle, and A. Sajantila. 2016. DNA quality and quantity from up to 16 years old post-mortem blood stored on FTA cards. Forensic Sci Int **261**:148-153.

Riesgo, A., A. R. Perez-Porro, S. Carmona, S. P. Leys, and G. Giribet. 2012. Optimization of preservation and storage time of sponge tissues to obtain quality mRNA for next-generation sequencing. Mol Ecol Resour **12**:312-322.

Robertson, P. A., T. Adriaens, X. Lambin, A. Mill, S. Roy, C. M. Shuttleworth, and M. Sutton-Croft. 2017. The large-scale removal of mammalian invasive alien species in Northern Europe. Pest Management Science **73**:273-279.

Rodriguez-Ezpeleta, N., L. Zinger, A. Kinziger, H. M. Bik, A. Bonin, E. Coissac, B. C. Emerson, C. M. Lopes, T. A. Pelletier, P. Taberlet, and S. Narum. 2021. Biodiversity monitoring using environmental DNA. Mol Ecol Resour **21**:1405-1409.

Rosenberg, R. 2015. Detecting the emergence of novel, zoonotic viruses pathogenic to humans. Cell Mol Life Sci **72**:1115-1125.

Rosenberg, R., M. A. Johansson, A. M. Powers, and B. R. Miller. 2013. Search strategy has influenced the discovery rate of human viruses. Proc Natl Acad Sci U S A **110**:13961-13964.

Russcher, A., A. Enders, C. S. de Brouwer, D. Oepkes, R. Hahn, M. Enders, A. C. M. Kroes, and A. Vossen. 2020. Diagnosis of intrauterine parvovirus B19 infection at birth - Value of DNA detection in neonatal blood and dried blood spots. Journal of Clinical Virology **129**.

Saccheri, I., M. Kuussaari, M. Kankare, P. Vikman, W. Fortelius, and I. Hanski. 1998. Inbreeding and extinction in a butterfly metapopulation. Nature **392**:491-494.

Sakschewski, B., W. von Bloh, A. Boit, L. Poorter, M. Peña-Claros, J. Heinke, J. Joshi, and K. Thonicke. 2016. Resilience of Amazon forests emerges from plant trait diversity. Nature Climate Change **6**:1032-1036.

Sales, N. G., M. B. McKenzie, J. Drake, L. R. Harper, S. S. Browett, I. Coscia, O. S. Wangensteen, C. Baillie, E. Bryce, D. A. Dawson, E. Ochu, B. Haenfling, L. L. Handley, S. Mariani, X. Lambin, C. Sutherland, and A. D. McDevitt. 2020. Fishing for mammals: Landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from riverine systems. Journal of Applied Ecology **57**:707-716.

Sales, N. G., O. S. Wangensteen, D. C. Carvalho, K. Deiner, K. Praebel, I. Coscia, A. D. McDevitt, and S. Mariani. 2021. Space-time dynamics in monitoring neotropical fish communities using eDNA metabarcoding. Sci Total Environ **754**:142096.

Samsonova, J. V., N. Y. Saushkin, and A. P. Osipov. 2022a. Dried Blood Spots technology for veterinary applications and biological investigations: technical aspects, retrospective analysis, ongoing status and future perspectives. Vet Res Commun **46**:655-698.

Samsonova, J. V., N. Y. Saushkin, and A. P. Osipov. 2022b. Dried Blood Spots technology for veterinary applications and biological investigations: technical aspects, retrospective analysis, ongoing status and future perspectives. Veterinary Research Communications **46**:655-698.

Santos, R. A. L., M. Mota-Ferreira, L. M. S. Aguiar, and F. Ascensao. 2018. Predicting wildlife road-crossing probability from roadkill data using occupancy-detection models. Science of The Total Environment **642**:629-637.

Say-Sallaz, E., S. Chamaille-Jammes, S. Periquet, A. J. Loveridge, D. W. Macdonald, A. Antonio, H. Fritz, and M. Valeix. 2023. Large carnivore dangerousness affects the reactive spatial response of prey. Animal Behaviour **202**:149-162.

Schaller, G. B. 1993. The Last Panda. University of Chicago Press:291.

Schilling, A. K., M. V. Mazzamuto, and C. Romeo. 2022. A Review of Non-Invasive Sampling in Wildlife Disease and Health Research: What's New? Animals (Basel) **12**.

Schmidt, D. J., T. Espinoza, M. Connell, and J. M. Hughes. 2017. Conservation genetics of the Mary River turtle (Elusor macrurus) in natural and captive populations. Aquatic Conservation: Marine and Freshwater Ecosystems **28**:115-123.

Schneider, S., G. W. Taylor, S. Linquist, S. C. Kremer, and R. B. O'Hara. 2019. Past, present and future approaches using computer vision for animal re-identification from camera trap data. Methods in Ecology and Evolution **10**:461-470.

Schwartz, A. L. W., F. M. Shilling, and S. E. Perkins. 2020. The value of monitoring wildlife roadkill. European Journal of Wildlife Research **66**.

Schwarz, C., R. Debruyne, M. Kuch, E. McNally, H. Schwarcz, A. D. Aubrey, J. Bada, and H. Poinar. 2009. New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. Nucleic Acids Res **37**:3215-3229.

Schwensow, N. I., B. Cooke, J. Kovaliski, R. Sinclair, D. Peacock, J. Fickel, and S. Sommer. 2014. Rabbit haemorrhagic disease: virus persistence and adaptation in Australia. Evolutionary Applications **7**:1056-1067.

Seeber, P. A., and L. S. Epp. 2022. Environmental DNA and metagenomics of terrestrial mammals as keystone taxa of recent and past ecosystems. Mammal Review **52**:538-553.

Seeber, P. A., G. K. McEwen, U. Lober, D. W. Forster, M. L. East, J. Melzheimer, and A. D. Greenwood. 2019. Terrestrial mammal surveillance using hybridization capture of environmental DNA from African waterholes. Mol Ecol Resour **19**:1486-1496.

Serrao, N. R., J. K. Weckworth, K. S. McKelvey, J. C. Dysthe, and M. K. Schwartz. 2021. Molecular genetic analysis of air, water, and soil to detect big brown bats in North America. Biological Conservation **261**.

Sigsgaard, E. E., M. R. Jensen, I. E. Winkelmann, P. R. Moller, M. M. Hansen, and P. F. Thomsen. 2020. Population-level inferences from environmental DNA-Current status and future perspectives. Evol Appl **13**:245-262.

Silveira, L., A. T. A. Jácomo, and J. A. F. Diniz-Filho. 2003. Camera trap, line transect census and track surveys: a comparative evaluation. Biological Conservation **114**:351-355.

Simon, M. A., J. M. Gil-Sanchez, G. Ruiz, G. Garrote, E. B. McCain, L. Fernandez, M. Lopez-Parra, E. Rojas, R. Arenas-Rojas, T. D. Rey, M. Garcia-Tardio, and G. Lopez. 2012. Reverse of the decline of the endangered Iberian lynx. Conserv Biol **26**:731-736.

Skovrind, M., M. Louis, M. V. Westbury, C. Garilao, K. Kaschner, J. A. S. Castruita, S. Gopalakrishnan, S. W. Knudsen, J. S. Haile, L. Dalen, I. G. Meshchersky, O. V. Shpak, D. M. Glazov, V. V. Rozhnov, D. I. Litovka, V. V. Krasnova, A. D. Chernetsky, V. M. Bel'kovich, C. Lydersen, K. M. Kovacs, M. P. Heide-Jorgensen, L. Postma, S. H. Ferguson, and E. D. Lorenzen. 2021. Circumpolar phylogeography and demographic history of beluga whales reflect past climatic fluctuations. Mol Ecol **30**:2543-2559.

Smith, K. F., K. Acevedo-Whitehouse, and A. B. Pedersen. 2009. The role of infectious diseases in biological conservation. Animal Conservation **12**:1-12.

Smith, L. M., and L. A. Burgoyne. 2004. Collecting, archiving and processing DNA from wildlife samples using FTA databasing paper. BMC Ecol **4**:4.

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

Spielman, D., B. W. Brook, and R. Frankham. 2004. Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci U S A **101**:15261-15264.

Steinmetz, R., W. Chutipong, N. Seuaturien, E. Chirngsaard, and M. J. B. C. Khaengkhetkarn. 2010. Population recovery patterns of Southeast Asian ungulates after poaching. **143**:42-51.

Taberlet, P., and J. Bouvet. 1992. Bear conservation genetics. Nature **358**:197.

Tanikawa, T., S. Watanabe, O. Mikami, and A. Miyazaki. 2023. Genetics of the rabbit haemorrhagic disease virus strains responsible for rabbit haemorrhagic disease outbreaks in Japan between 2000 and 2020. Journal of General Virology **104**.

Teer, J. K., and J. C. Mullikin. 2010. Exome sequencing: the sweet spot before whole genomes. Hum Mol Genet **19**:R145-151.

Tobler, M. W., S. E. Carrillo-Percastegui, A. Zúñiga Hartley, and G. V. N. Powell. 2013. High jaguar densities and large population sizes in the core habitat of the southwestern Amazon. Biological Conservation **159**:375-381.

Trolle, M. 2003. Mammal survey in the southeastern Pantanal, Brazil. Biodiversity and Conservation **12**:823-836.

Tuaillon, E., D. Kania, A. Pisoni, K. Bollore, F. Taieb, E. N. O. Ngoyi, R. Schaub, J. C. Plantier, A. Makinson, and P. Van de Perre. 2020. Dried Blood Spot Tests for the Diagnosis and Therapeutic Monitoring of HIV and Viral Hepatitis B and C. Frontiers in Microbiology **11**.

Ushio, M., H. Fukuda, T. Inoue, K. Makoto, O. Kishida, K. Sato, K. Murata, M. Nikaido, T. Sado, Y. Sato, M. Takeshita, W. Iwasaki, H. Yamanaka, M. Kondoh, and M. Miya. 2017. Environmental DNA enables detection of terrestrial mammals from forest pond water. Mol Ecol Resour **17**:e63-e75.

Valentini, A., P. Taberlet, C. Miaud, R. Civade, J. Herder, P. F. Thomsen, E. Bellemain, A. Besnard, E. Coissac, F. Boyer, C. Gaboriaud, P. Jean, N. Poulet, N. Roset, G. H. Copp, P. Geniez, D. Pont, C. Argillier, J.-M. Baudoin, T. Peroux, A. J. Crivelli, A. Olivier, M. Acqueberge, M. Le Brun, P. R. Moller, E. Willerslev, and T. Dejean. 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology **25**:929-942.

van Vliet, K., W. G. van Ginkel, E. van Dam, P. de Blaauw, M. Koehorst, H. A. Kingma, F. J. van Spronsen, and M. R. Heiner-Fokkema. 2020. Dried blood spot versus venous blood sampling for phenylalanine and tyrosine. Orphanet J Rare Dis **15**:82.

Vetterlein, W. 1954. Das klinische bild der maul-und-klauenseuche beim menschen, aufgestelt aus den bisher experimentell gesicherten erkrankungen. Arch Exp Veterinarmed:541–564.

Voss, R. S., and L. H. Emmons. 1996. Mammalian diversity in Neotropical lowland rainforest: A preliminary assessment. Bull. Am. Mus. Nat. His. **230**:13-35.

Voss, R. S., D. P. Lunde, and N. B. Simmons. 2001. Mammals of Paracou, French Guiana: A Neotropical lowland rainforest fauna. Part 2: nonvolant species. Bull. Am. Mus. Nat. His. **263**:1-236.

Ward, A. I., S. Richardson, R. Macarthur, and A. C. Mill. 2020. Using and communicating uncertainty for the effective control of invasive non-native species. Mammal Review **50**:211-220.

Wright, L. I., T. Tregenza, and D. J. Hosken. 2007. Inbreeding, inbreeding depression and extinction. Conservation Genetics **9**:833-843.

Wu, S. G., W. J. Pan, H. N. Liu, M. L. Byrne-Steele, B. Brown, M. Depinet, X. H. Hou, J. Han, and S. Li. 2019. High throughput sequencing of T-cell receptor repertoire using dry blood spots. Journal of Translational Medicine **17**.

Xu, Z. J., and W. X. Chen. 1989. Viral haemorrhagic disease in rabbits: a review. Vet Res Commun **13**:205-212.

Zang, Z., G. Shen, G. Ren, C. Wang, C. Feng, W. Xu, Z. Xie, Q. Chen, X. Yang, and J. Li. 2017. Thermal habitat of giant panda has shrunk by climate warming over the past half century. Biological Conservation **211**:125-133.

Zhang, Y. Z., Y. M. Chen, W. Wang, X. C. Qin, and E. C. Holmes. 2019. Expanding the RNA Virosphere by Unbiased Metagenomics. Pages 119-139 *in* L. Enquist, D. DiMaio, and T. Demody, editors. Annual Review of Virology, Vol 6, 2019.

CHAPTER 2: Manuscript 1

Monitoring terrestrial wildlife by combining hybridization capture and metabarcoding data from waterhole environmental DNA

Juan Li[1, 8], Peter Seeber[1], Jan Axtner[2], Rachel Crouthers[3], Milou Groenenberg[3], Arnulf Koehncke[4], Alexandre Courtiol[5], Pin Chanratana[6], Alex D. Greenwood[1, 7*]

[1] Department of Wildlife Diseases, Leibniz Institute for Zoo and Wildlife Research, Berlin 10315, Germany

[2] Department of Ecological Dynamics, Leibniz Institute for Zoo and Wildlife Research, Berlin 10315, Germany

[3] World Wild Fund for Nature, Phnom Penh 120102, Cambodia

[4] WWF Germany, Berlin 10117, Germany

[5] Department of Evolutionary Genetics, Leibniz Institute for Zoo and Wildlife Research 10315, Berlin, Germany

[6] Ministry of Environment, Phnom Penh 120101, Cambodia

[7] Department of Veterinary Medicine, Freie Universität Berlin, Berlin 14163, Germany

[8] Department of Biologie, Chemie, and Pharmazie, Freie Universität Berlin, Berlin 14195, Germany

You have to purchase this part online:

Biological Conservation, Volume 284, 2023, 110168, ISSN 0006-3207

https://doi.org/10.1016/j.biocon.2023.110168

CHAPTER 3: Manuscript 2

Nucleic acid degradation after long term dried blood spot storage

Juan Li[1], Gabriela M. Ulloa[2,3], Pedro Mayor[2,4,5], Meddly L. Santolalla Robles[6], Alex D. Greenwood[1,7]*

[1] Department of Wildlife Diseases, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany

[2] Programa de Pós-Graduação em Saúde e Produção Animal na Amazônia, Universidade Federal Rural da Amazônia (UFRA), Av. Presidente Tancredo Neves 2501, Terra Firme, 66077-830, Belém-Pará, Brazil

[3] Grupo de Enfermedades Infecciosas Re-emergentes, Universidad Científica del Sur (UCSUR), Lima, Peru

[4] ComFauna, Comunidad de Manejo de Fauna Silvestre en la Amazonía y en Latinoamérica, Iquitos, Peru

[5] Departament de Sanitat i Anatomia Animals, Universitat Autònoma de Barcelona, Bellaterra, Spain

[6] Emerge, Emerging Diseases and Climate Change Research Unit, School of Public Health and Administration, Universidad Peruana Cayetano Heredia, Lima, Peru

[7] School of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

Abstract

Collecting and preserving biological samples in the field, particularly in remote areas in tropical forests, prior to laboratory analysis is challenging. Blood samples in many cases are used for nucleic acid based species determination, genomics or pathogen research. In most cases maintaining a cold chain is impossible and samples remain at ambient temperature for extended periods of time before controlled storage conditions available. Dried blood spot (DBS) storage, blood stored on cellulose-based paper, has been widely applied to facilitate sample collection and preservation in the field for decades. However, it is unclear how long term storage on this substrate affects nucleic acid concentration and integrity. We analyzed nucleic acid quality from DBS stored on Whatman filter paper no.3 and FTA cards for up to 15 years in comparison to cold chain stored samples using three nucleic acid extraction methods. We examined the ability to identify viral sequences from samples of twelve free-ranging primates in the Amazon forest, using targeted hybridization capture and determined if mitochondrial genomes could be retrieved. The results suggest that even after extended periods of storage, DBS will be suitable for some genomic applications but may be of limited use for viral pathogen research, particularly RNA viruses.

Keywords: DBS, nucleic acid, hybridization capture, virus, mitochondrial genome

Introduction

In-field preservation of biological samples is of paramount importance to their subsequent utility in laboratory settings. This is particularly true if the nucleic acids are the intended downstream analysis target. Various preservation methods have been employed to store samples for nucleic acid based research in the field including freezing where electricity or liquid nitrogen storage is possible, preservation in storage buffers such as RNALater and ethanol (Nsubuga et al. 2004, Soto-Calderon et al. 2009, Camacho-Sanchez et al. 2013), and several types of filter paper (Hsiao et al. 1999, Smith and Burgoyne 2004, Hearps et al. 2010, Samsonova et al. 2022). Snap freezing maintains the integrity of nucleic acids ideal for downstream genomic applications (Riesgo et al. 2012, Tang et al. 2012, Mareninov et al. 2013, da Cunha Santos 2018) as it effectively halts enzymatic activity, slows nucleic acid degradation and minimizes microbial growth (Naber 1996, Shabihkhani et al. 2014). However, its successful implementation requires overcoming the logistical challenges associated with maintaining low temperatures under various environmental conditions. In many cases, such as remote forest locations requiring weeks of travel, these challenges can be unsurmountable.

Storage buffers represent a good bridge method to preserve nucleic acids at ambient temperatures in the field for several days or weeks before transfer to a laboratory setting where they can be frozen (Camacho-Sanchez et al. 2013). However, preservation buffers preserve nucleic acids less effectively over time than freezing particularly under extreme environments (Riesgo et al. 2012, Camacho-Sanchez et al. 2013), making them more suited for short-term storage. Furthermore, storage buffer is relatively expensive and can be cost prohibitive for large field based studies.

Filter paper storage is easy to implement and transport even in remote field sites. The most widely applied sample type is dried blood spots (DBS) (McClendon-Weary et al. 2020, Samsonova et al. 2022). DBS samples were first applied in clinical settings to detect metabolites indicative of phenylketonuria in newborns (Guthrie and Susi 1963, Pitt 2010, van Vliet et al. 2020). Use of DBS have subsequently been extended to store samples in various medical and animal research settings (Lehmann et al. 2013, McClendon-Weary et al. 2020). Such studies have targeted proteins (Martin and Cooper 2014), lipids (Koulman et al. 2014), and nucleic acids (Tuaillon et al. 2020). DBS have been used extensively in medical diagnostics and bio-banking (Demirev 2013, Lim 2018, Capiau et al. 2019, McClendon-Weary et al. 2020, Russcher et al. 2020), while its potential significance in the fields of animal biology, particularly for wildlife may be underestimated (Samsonova et al. 2022).

DBS are obtained by spotting a small amount of blood onto filter paper, a stable and compact format for transportation and storage, and allowing it to dry at room temperature (McClendon-Weary et al. 2020). Multiple types of specialized filter paper have been designed to optimize the collection and stability of

biological materials of interest for a variety of research purposes, such as Whatman 903 protein saver, Flinders Technology Associates (FTA) Cards and Nucleic-Card (TM) (Uttayamakul et al. 2005, Molteni et al. 2013, Marchand et al. 2021). FTA cards are the best-established option in nucleic acid targeted research (Rahikainen et al. 2016, Tam et al. 2016, da Cunha Santos 2018, Cardona-Ospina et al. 2019, Hashimoto et al. 2019, Longhi et al. 2023). FTA cards rely on a porous matrix of cellulose fiber to trap nucleic acid molecules, while they are additionally impregnated with proprietary chemicals that can denature proteins and reduce microbial growth, preventing enzymatic degradation of nucleic acids. Untreated filter paper was reportedly comparable in preserving viral nucleic acids compared to more expensive chemically-treated FTA cards (Jones et al. 2012), even for prolonged storage (up to 9 months) under tropical environmental condition (Michaud et al. 2007). However, there is a scarcity of research on nucleic acid preservation on treated and non-treated filter paper over extended periods, particularly as it applies to wildlife and conservation research, fields which are particularly dependent on long term field storage of collected samples.

DBS on filter paper have been used in a number of animal genetic and pathogen studies. For example, rabies virus in dried brain tissue was stably detected on filter paper for up to 2 years at ambient condition (Wacharapluesadee et al. 2003, Rasolonjatovo et al. 2020). DNA from avian DBS on FTA cards were preserved for up to 44 months (Smith and Burgoyne 2004). Fowler et al. (2012) demonstrated that DNA on FTA cards can be used to amplify multiple single nucleotide polymorphism loci even after three years of DBS storage. The aforementioned studies were based on target-amplification approaches that can work with minimal starting DNA amounts and tolerate substantial DNA degradation. However, many samples that are retrieved and stored could benefit from recent advances in genomics methods or may represent the only resources for genomic information on rare species (Andrews 2018, Carroll et al. 2018, Schilling et al. 2022). In such cases, quality assessment in terms of fragment size and concentration of DNA or RNA are required for analysis. For pathogen testing, there is the additional variable of pathogen concentration in a given sample, which may be low in which case degradation may exacerbate general detection associated problems.

In the current study, we measured DNA integrity and concentration from wildlife DBSs collected on two types of filter paper (Whatman filter paper no.3 (W3) and FTA cards) and stored from 1-15 years, compared to immediately cold chain stored samples. We determined (1) how the quality of DNA and RNA on DBSs changed over time, compared to frozen human whole blood samples; (2) the storage performance of FTA and W3 cards; (3) the impact of nucleic acid extraction method; (4) virus sequence retrieval by target capture and sequencing; and (5) retrieval of non-target nuclear and mitochondrial DNA sequences.

Materials and methods

1. Sampling and storage

Dried blood spots (DBSs) were collected from 201 non-human primates (135 FTA and 66 W3) by subsistence indigenous hunters as part of a wildlife conservation program, taking advantage of the discarded material from legal subsistence hunting, in the Yavarí-Mirin River basin (04°19′53″ S, 71°57′33″ W) in the Peruvian Amazon. Approximately 0.8 mL blood was spotted on storage paper, followed by 2-3 hours ambient drying and storage in desiccants. DBS samples were kept in individual envelopes sealed in separate zipper closure bags at room temperature until sample transportation between 3 and 9 months (frequency being restricted by remoteness of the study areas and the poor transportation infrastructure). Once in the laboratory, DBSs were immediately archived at room temperature. Additionally, whole blood samples (2–3 mL) from 44 local inhabitants were collected by venipuncture from voluntary participants, after providing written informed consent, and were transported in liquid nitrogen to the research lab to be stored at -80 °C.

2. Nucleic acid isolation

Nucleic acid was extracted from cold-chain stored human whole blood and wildlife DBS samples. Four nucleic acid extraction protocols were applied:

(1) J_DNA: storage paper was cut into small pieces with sterile scissors and diluted with 500 µL of PBS buffer. Diluted DBS discs were centrifuged at 6000 rpm for 2×20s. After centrifugation, 200 µL supernatant was used. After adding 20 µL proteinase K and 200 µL Buffer AL in DNeasy Blood & Tissue Kit (Qiagen, Germany), the mixture was incubated at 56 °C for 10 min. Downstream extraction was conducted following the manufacturer's protocol;

(2) J_ R+D: scissor-sliced paper discs were diluted using 2 volumes (1120 µL) of AVL buffer with RNA carrier (Qiagen, Germany) were centrifuged at 2000 rpm for 10 min. Nucleic acid were isolated from the supernatant with a QIAamp Viral RNA Mini Kit (Qiagen, Germany) according to the downstream manufacturer's protocol. Human blood samples with sustained cold-chain storage were processed following this manufacturer's protocol throughout the entire procedure;

(3) G_D+R: filter papers were hole punched using 3 mm hole punch cutter. Between 10-12 paper discs, roughly equivalent to 50 ul of whole blood, was used for nucleic acid isolation. DBS hole punches were diluted with 600 µL of RLT Plus Buffer and then were incubated at 55°C at 2000 rpm for 30 min in a ThermoMixer. Column-based DNA isolations were performed following Qiagen Quick-Start Protocol AllPrep DNA/RNA Mini Kit.

(4) O_DNA: DNA were isolated for DNA targeted research from 2008 to 2018, with the isolation details unknown.

A total of 71 samples were processed by method one (J_DNA), 153 by method two (J_ R+D), and 15 by method three (G_D+R) and 36 by method four (O_DNA).

Nucleic acid extracts were quantified and qualified on an Agilent Tapestation 2200 system. Fragmentation was assessed using the DNA Integrity Number (DIN) algorithm, in which the fluorescent signals of fragment size distribution determine integrity scores (a range of 1 to 10 values). The higher the value, the higher the nucleic acid integrity. Fragment size peaks were obtained from the nucleic acid migration images generated.

3. Library preparation, hybridization capture and sequencing

Twelve nucleic acid extracts from four primate species with DIN ranging from 1.7 to 3.2 and fragment peaks from 433 to 7,778 bp were used for downstream viral screening using hybridization capture (Table S1). All twelve samples were kept on FTA cards and were extracted using the scissor sliced nucleic acid isolation method (described above as isolation method 2). To fully explore the potential existence of both DNA and RNA viruses, nucleic acid extracts were first reverse transcribed into double stranded cDNA with a two-step cDNA synthesis using Invitrogen SuperScript® Double-Stranded cDNA Synthesis Kit with SuperScript™ III Reverse Transcriptase and (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Neither DNAase nor RNAase was applied in order to detect both RNA and DNA viruses. Illumina sequencing libraries were prepared from the double stranded cDNA following the Meyer and Kircher 2010 protocol with slight modifications (Meyer and Kircher 2010; Seeber, McEwen et al. 2019). Briefly, reverse transcribed DNA was fragmented to 300 bp using an ultrasonicator in 50 µL volume (Covaris M220; Covaris, Woburn, MA, USA). Fragment size and quantity were assessed on an Agilent 2220 TapeStation with D1000 chips. End repair and adapter fill-in were performed on 42.5 µL sheared cDNA with the NEBNext kit (New England Biolabs, New England Biolabs, Ipswitch, MA, USA). Each sample was barcoded by adding a unique combination of 6-nucleotide tags with five amplification cycles and three technical replicates.

For hybridization capture reactions, a bait set (70-mer nucleotides) from the fifth-generation Virochip (Yozwiak et al. 2012) was used to target a wide range of vertebrate viruses after excluding viral families with only invertebrates host, endogenous retroviruses, plant viruses and bacteriophage. Three equimola libraries were pooled for each reaction run. The reaction mixtures were incubated for 48 h at 60 °C following the manufacturer's instructions. After bead washing, the capture products were enriched using 15 cycles of on-bead PCR with P5/P7 bridge primers with KAPA HiFi HotStart ReadyMix (Roche, Basel, Switzerland). The enriched capture products were then cleaned using a MinElute PCR Purification Kit (Qiagen, Hilden, Germany), quantified on both Agilent 4150 TapeStation and Qubit system, and pooled at equimolarity. The final pool of enriched capture products was diluted to 8 pM with a Phix DNA control spike-in of 1% and sequenced on an Illumina Miseq platform (Miseq v2, 300 cycles, 2 × 150 bp paired-end reads; Illumina, San Diego, CA, USA).

4. Bioinformatic pipelines

Demultiplexed raw reads were trimmed to remove adaptors with CUTADAPT v1.15 (Martin 2011) and were filtered using a quality score above 20 and read length above 50 bp with a single window of 10 bp using TRIMMOMATIC v0.38 (Bolger et al. 2014). The retained paired-end reads were subsequently merged using FLASH v1.2.11 (Magoc and Salzberg 2011) with a cutoff of 20 bp overlap. For comparison purposes, whole genome sequencing data of four primate species were downloaded from European Nucleotide Archive (ENA) under Project PRJEB59576: *Cacajao calvus* (ERR10941572), *Cebus albifrons* (ERR10942010), *Lagothrix lagotricha* (ERR10941540) and *Sapajus macrocephalus* (ERR10941507). The four ENA datasets were subsampled to incrementally generate a set of data with 50,000, 100,000, 150,000, 200,000, 300,000, 450,000, 600,000, 750,000, 900,000, 1050,000, 1,200,000, 1,350,000 and 1,500,000 reads by using seqtk (https://github.com/lh3/seqtk). The downloaded ENA data and corresponding subsampled data all went through the same quality control step as described above.
Direct mapping against mitochondrial genome references was performed using the pre-processed pair-end reads of the twelve samples to identify mitogenome sequences with BWA (Li and Durbin 2009). The bam files generated were processed to calculate the number of mapped reads, the coverage percentage and average coverage depth of mitogenome references using Samtools (Li et al. 2009) and bedtools (Quinlan 2014). All reads that failed to map to the mitogenome references were retained using Samtools, which were them mapped against nuclear genomes of the corresponding species or closely related primate species with Bowtie2 (Langdon 2015).

De novo mitochondrial genomes assemble were performed using the pair-end reads with Novoplasty (Dierckxsens et al. 2016), by taking complete CO1 gene of corresponding species as a seed sequence. The mitochondrial genome of respective species or the most closely related species were downloaded from NCBI nucleotide database as assembly references. Assembled mitochondrial genome sequences were curated by mapping paired reads against them using BWA. Mapping results were sorted using Samtools and manually checked in Tablet (Milne et al. 2013) to generate a final consensus mitochondrial genome sequence. Total coverage of each mitochondrial genome assembly was calculated with bedtools. The seven complete mitochondrial genomes were annotated with MitoFinder (Allio et al. 2020).

For viral sequence detection, the twelve primate sequencing datasets were processed using two viral classification pipelines: (1) direct classification and (2) post-assembly classification. Direct classification was performed by using merged reads to search against the BLAST viral genome reference database with Blastn (-evalue 1e-5) to identify candidate viral sequences. The reads with viral hits were retained after removing all bacteriophage hits. The remaining candidates were aligned against the complete NCBI nucleotide database (-evalue 1e-10) to confirm that matched reads were more similar to viral sequences than non-viral ones. The candidate viral sequences were then filtered with the threshold of overall query coverage above 90%, identity above 85%, and bit-score greater than 190. The retained

candidates were further curated by aligning candidate sequences to the corresponding virus genome. The reads identified as viral hits but with low complexity were removed due to high ambiguity. Post-assembly classification was performed by first removing host mitochondrial and nuclear genome sequences using Bowtie2 and Samtools and rRNA using SortMeRNA (Kopylova et al. 2012). Retrained reads were *de novo* assembled using the Spades assembler (Bankevich et al. 2012). Generated contigs were taxonomically classified following two rounds of blast search with a minimum e-value of 1e-5 for searching against the viral reference database and 1e-10 for searching against the complete nucleotide database.

5. Statistical analysis

To compare storage capacity of two types of paper, plus cold chain storage and nucleic acid extraction efficiency of three methods, a generalized linear model was fitted using the glm() function in R. The DIN (DNA integrity number) obtained from Tapestation system was used as the response variable. Logged interval days of storage from collection to isolation, logged DNA fragment peak, storage methods and extraction methods were taken as fixed effects, and individual sample ID as a random effect. We conducted anova tests using the R function anova() to evaluate the statistical significance of each fixed effect and using the rand() function to assess the random effect. Post hoc tests were performed to test the statistical differences of every two storage and extraction methods by using the R package emmeans (Searle et al. 2012) with tukey method.

Results

1. Storage and extraction methods effects on nucleic acid integrity

DNA integrity and maximal fragment size significantly deteriorated as storage time increased (Table 1). Storage and extraction methods both significantly affected the DNA quality retrieved (Table 1). In contrast, sample-specific effects (i.e. random effect from sample ID) did not contribute to the overall variation of the fitted model (Table 1).

The comparison of storage methods showed that cold chain storage significantly improved the integrity of DNA retrieved (Figure 1 c and Table 1). When compared to DBS samples of the same age or younger than cold-chain stored samples, the latter preserved longer DNA fragments (genomic DNA with fragment peaks above 15,000 bp) (Figure 1 b). In contrast, no significant differences were detected between the two types of filter paper used (Table 1 and Figure 1 c). Both FTA and W3 card stored samples yielded DNA fragments with maximal size of 20,000 bp (Figure 1 b). DNA integrity (i.e. DIN) slightly increased for W3 card stored samples that were extracted within 6-years after collection (Figure 1 a). For DBS samples stored longer than 9 years, DNA extracts had a DIN of ca 5.0 and maximal DNA fragment lengths below 10,000 bp (Figure 2 a and b).

Three DNA-targeted extractions (method 1, 3 and 4) yielded significantly more intact DNA compared with that isolated using RNA specific kit (Table 1 and Figure 1). Method 2 using RNA specific isolation kit yielded DNA with the lowest DNA integrity on average even for samples stored 1-2 years, and regardless of which storage paper was used (Table 1, and Figure 1 and 2 a). No significant differences were detected between the three DNA-specific extraction methods (Table 1 and Figure 1 c).

Table 1. Results from generalized linear model that predicts the effects of sample properties and processing on DNA integrity.

Linear mixed model to assess DNA integrity						
Fixed effects	Estimates(s.e)	Std. Error	Df	T value	Chi2_LR	P value
(Intercept)	8.736	1.356	234.846	6.443		
Storage interval days (log)	-0.424	0.128	211.574	-3.309	11.06	**0.0009** ***
DNA fragment peak (log)	0.309	0.062	241.357	4.973	24.274	**<.0001****
Storage[FTA]	-3.932	0.270	209.830	-14.580	146.413	**<.0001****
Storage[W3]	-3.796	0.332	223.081	-11.418		
Extraction[J_DNA]	-0.646	0.405	240.802	-1.596	79.928	**<.0001****
Extraction[J_R+D]	-2.570	0.431	237.947	-5.963		
Extraction[J_DNA]	-0.379	0.503	230.782	-0.752		
Random Effects					LRT	
Sample_ID (Intercept)	0.083	0.289			0.170	0.679
Residual	1.446	0.203				
ICC	0.05					
Marginal R2 / Conditional R2	0.779/ 0.791					

Post hoc Test						
Storage method	Contrast	Estimate	Std. Error	Df	T.ratio	P value
	CC-W3	3.932	0.271	204	14.532	**<.0001****
	CC-FTA	3.796	0.333	219	11.386	**<.0001****
	W3-FTA	-0.136	0.264	239	-0.517	0.8633
Extraction method	contrast	Estimate	Std. Error	Df	T.ratio	P value
	G_D+R - J_DNA	0.646	0.407	240	1.588	0.387
	G_D+R - J_R+D	2.570	0.433	237	5.940	**<.0001****
	G_D+R - O_DNA	0.379	0.506	228	0.749	0.877
	J_DNA - J_R+D	1.924	0.247	131	7.784	**<.0001****
	J_DNA - O_DNA	-0.267	0.364	229	-0.735	0.883
	J_D+R - O_DNA	-2.191	0.334	243	-6.556	**<.0001****

Types of filter paper used for sample storage: CC (cold chain), FTA (FTA card), W3 (Whatman filter paper no.3).

Nucleic acid extraction methods: J_DNA (extraction method 1: DNA specific), J_R+D (extraction method 2: RNA specific), G_D+R (extraction method 3: DNA and RNA combined), O_DNA (extraction method 4: DNA specific).

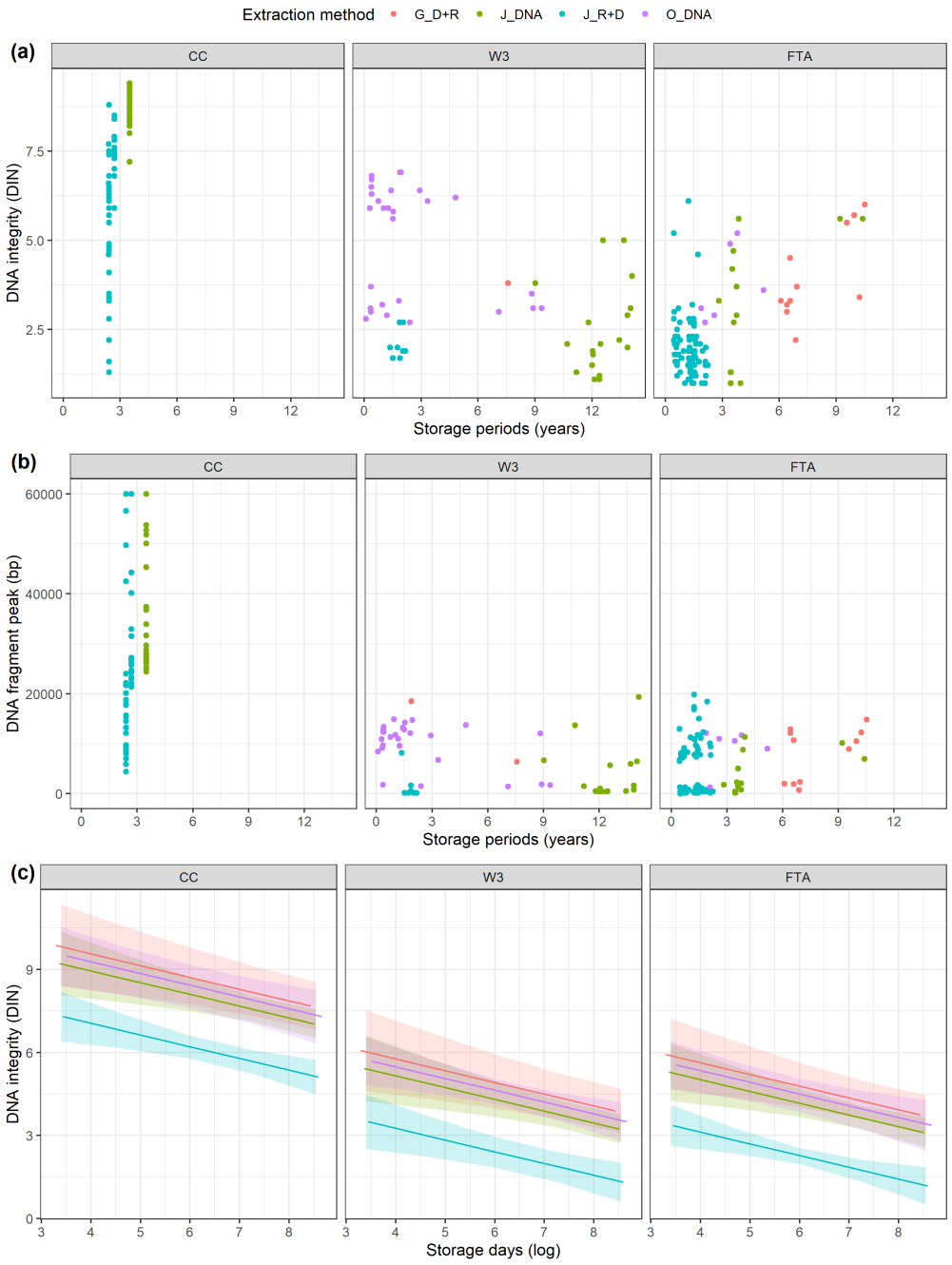

Figure 1. Plots from the isolated DNA and generalized linear models (GLM). Panel (a) shows the DNA integrity (i.e. DIN value from Tapestation system) of each DNA isolates. Panel (b) shows the maximal DNA fragment retained in each DNA samples. Panel (c) shows the GLM based predictions of the DNA integrity (DIN) versus logged storage days, storage methods and extraction methods. Storage days: the interval days from sample collection to isolation. Types of filter paper used for sample storage: CC (cold chain storage), FTA (FTA card storage), W3 (Whatman filter paper no.3). Nucleic acid extraction methods: J_DNA (extraction method 1: DNA specific), J_R+D (extraction method 2: RNA specific), G_D+R (extraction method 3: DNA and RNA combined), O_DNA (extraction method 4: DNA specific).

2. Nuclear genome and mitochondrial genome retrieval

Although virus target capture was performed, there was substantial non-targeted sequence generated per sample. The sequencing reaction generated 11.87 million reads for twelve primate samples with an average yield of 0.99 ± 0.35 million reads per sample, in which 11.24 million reads passed the quality control and could be merged into a single read (Table S2). The reads mapping to the host genome ranged from 0.64% to 51.21% of total reads from each sample, comprising 5,479-684,464 reads with a median of 121,748 reads (Table S2). None of the twelve samples obtained more than 0.01% nuclear genome coverage. In contrast, 7.83%-61.88% of total reads were mapped to mitochondrial genome references. Six samples achieved 100% mitochondrial genome coverage, and two were below 50% coverage.

Seven complete and three partial mitochondrial genomes out of twelve collected samples for four primate species were *de novo* assembled: complete mitochondrial genomes for *Cebus albifrons*, *Lagothrix lagothrica poeppigii* and *Cacajao calvus*; partial mitochondrial genomes for *Sapajus macrocephalus* (Table 2). A range of 120,805 to 659,177 reads were aligned to references in the *de novo* assembly pipeline, representing 17.05%-52.36% of the total sequencing data (Table 2). Mitochondrial genome assemblies ranged from 24× to 244,495 × base coverage (Table 2). Non-uniform base coverage was observed for all seven mitochondrial genome assemblies. The mapping against the corresponding mitochondrial genome assemblies showed a slight increase in the number of aligned reads and average coverage depth compared to the results obtained from directly mapping against downloaded mitogenome references. The number of aligned reads in *de novo* assembly was correlated to the number of reads mapped to NCBI mitochondrial genome references (Table 2 and Figure S1 a, R^2=1, P value<0.0001). The coverage percentage of the respective references was positively correlated with the *de novo* assembly success (Table 2). In contrast, the indiscriminate whole genome sequencing obtained significantly lower percentage of reads that were aligned to references, ranging from 0.43% to 1.31% depending upon species.

Table 2. Results from *de novo* assembled mitochondrial genomes and mapping against mitochondrial genome references.

Species	Sample ID/ENA dataset ID	Mitogenome *de novo* assemble								Mitogenome mapping to NCBI references (BWA)			
		De novo assemble with Novoplasty					Mapping reads to assemblies (BWA)						
		Assembled	Length (bp)	Aligned reads	Percentage of assembled reads	Average organelle coverage	Aligned reads	Percentage of mapped reads	Average coverage depth	Aligned reads	Percentage of mapped reads	Coverage percentage	NCBI Accession of references
Sapajus macrocephalus (Large-headed capuchin)	151320002	Partial	16541	370679	29.37%	6768	457509	35.77%	6651.2	460123	34.42%	100.00%	
	151320004	No								194314	22.72%	42.40%	NC_064167.1
	151320010	No								88734	7.83%	35.90%	
	ERR10941507	Complete	16544	424724	0.44%	3879							
Cebus albifrons (White-fronted capuchin)	151320038	No								425781	60.53%	87.00%	NC_002763.1
	151320039	Complete	16552	203081	41.79%	3705	287005	58.15%	3904.63	277246	53.28%	99.90%	
	ERR10942010	Complete	16552	438362	0.43%	3999							
Lagothrix poeppigii (Silvery woolly monkey)	151320054	Complete	16594	342405	37.85%	6232	439258	46.38%	6014.44	435214	44.09%	100.00%	NC_021951.1
	151320055	Complete	16594	659177	47.85%	11997	805386	57.04%	11240.68	794557	54.35%	100.00%	
	151320056	Complete	16595	430654	51.82%	7837	534059	63.42%	8019.77	534631	61.88%	100.00%	
	151320065	No								371125	28.13%	95.10%	
	151320068	Complete	16595	120805	31.81%	2198	170511	44.91%	2429.74	169277	43.26%	100.00%	
	151320086	Complete	16595	305024	41.27%	5551	355957	46.82%	5470.5	356346	45.12%	100.00%	
Lagothrix lagotricha	ERR10941540	Complete	16595	1115096	1.31%	10146							
Cacajao calvus (uakari)	151320109	Complete	16712	250435	19.69%	49311	351124	26.92%	4977.01	349481	25.83%	99.96%	NC_021967.1
	ERR10941572	Complete	16700	702212	1.17%	6349							

Compared with the unselective whole genome sequencing data, the mitochondrial genome assembled from hybridization capture data obtained highly fluctuating coverage across the entire assemblies (Figure 2 a, b and d). Coverage depth peaked at different gene coding regions across all three species (Figure 2 a and b). Ribosome RNA regions had the highest coverage, especially for the end regions of the 16S rRNA gene (Figure 2 a). By using subsampled whole genome sequencing data, we observed that the sequencing effort required for successful *de novo* mitogenome assemble was highly related to the percentage of mitogenome sequences (Figure 2 c). The two species with above 1.1% mitogenome sequence required at least 100,000 reads to successfully *de novo* assemble complete mitochondrial genome with above 20× base coverage (Figure 2 c and Table 2). The other two species with around 0.4% mitochondrial sequence required 300,000 reads to reach similar per base coverage (Figure 2 c and Table 2).

3. Virus classification

Between 2.38%-24.34% of sequencing data with a median of 118,448 reads per sample aligned to viral references. The number of reads aligned to viral references was positively correlated to the number of reads aligned to mitochondrial genomes despite no statistical significance (Figure S1 c, R^2=0.21, P-value=0.1305). Of those, 54.59%-93.47% of candidate viral reads were aligned to bacteriophage, ranging from 17,437 to 269,815 reads per sample (Table S3). After a second round of blast searches against the complete NCBI nucleotide database and stringent filtering, 24 reads from four samples were retained as potential candidate viral reads after removing bacteriophage reads. However, all 24 reads were low-complexity sequences and therefore cannot be confirmed as viral derived.

Post sequence assembly classification results showed that a total of 1,087 to 82,854 contigs were assembled from nuclear genome and rRNA removed data. Ten out of twelve samples yielded 51 candidate viral contigs (Table S3). Four were retained after filtering out bacteriophage-like contigs, but were also excluded as viral hits due to high ambiguity, low query coverage and low sequence identity.

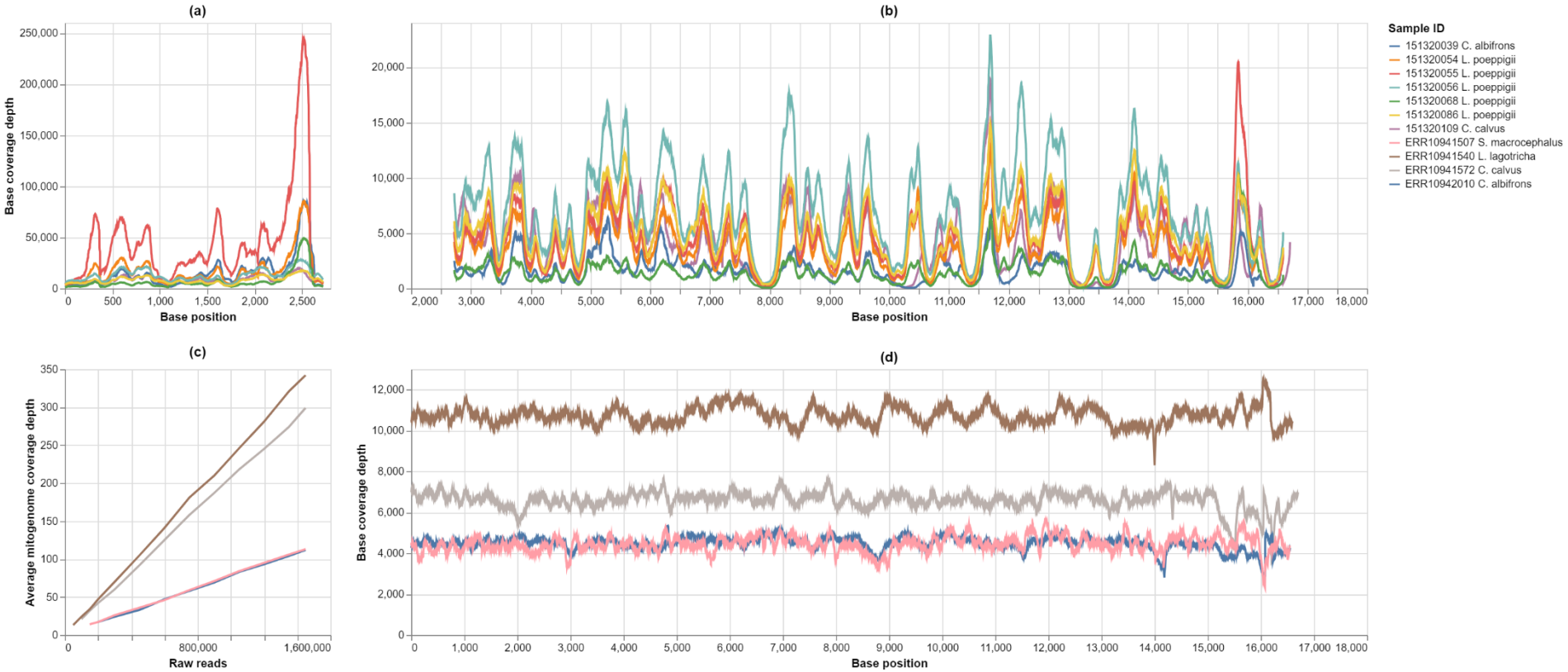

Figure 2. Results from the mapping against mitochondrial genome assemblies. Plot (a) shows the base coverage of seven mitochondrial genome assemblies that *de novo* assembled from target capture sequencing data from position zero to 2,700. Plots (b) shows the base coverage of seven mitochondrial genome assemblies that *de novo* assembled from target capture sequencing data from position 2,700 to the end position of each assemblies. Plot (c) shows the average coverage depth of four mitochondrial genome assemblies that *de novo* assembled from downloaded ENA data versus the number of downsamples sequencing reads. Plot (d) shows the base coverage of four mitochondrial genome assemblies that *de novo* assembled from whole genome sequencing data downloaded from European Nucleotide Archive (ENA).

Discussion

1. Nucleic acid storage and stabilization

In the present study, DNA integrity was strongly dependent on extraction methods applied. High-temperature incubation protocols prior to nucleic acid isolation significantly improved the quality of DNA isolated from DBS preserved on both FTA cards and W3 paper. Kumar et al. (2019) reported that no differences were observed in DNA yields between two incubation-based isolation protocols consistent with our results (Kumar et al. 2019). In the process of DBS sampling, the DNA is absorbed, and captured within the porous and uniform matrix of the fibers once biological fluid is applied to the cellulose-base paper. Under high-temperature incubation prior to isolation (generally applied in DNA specific isolation), the physical nucleic acid bond to the fiber matrix is dissolved, which likely improves DNA dilution efficacy and DNA retrieval. In contrast, RNA-specific isolation kits can isolate DNA and RNA simultaneously but tend to result in DNA fragmentation, including for cold chain stored samples.

DBS samples have been extensively applied for amplification-dependent analysis for a variety of research and diagnosis purposes (Hollegaard et al. 2009, van Schalkwyk et al. 2019, Tuaillon et al. 2020). More recently, whole genome sequencing without selective amplification has been applied to DBS samples (Bassaganyas et al. 2018, Wu et al. 2019, Agrawal et al. 2021), with the results highly dependent on DNA fragmentation. A previous study demonstrates that relatively large DNA fragment (e.g. 1000 bp) can be preserved and amplified even for up to 11 years storage at ambient temperature in tropical regions, but amplification efficacy decreased when samples were stored longer than 9 years (Chaisomchit S et al. 2005). PCR-based target detection performed differently on DNA isolated different preservation paper (Bezerra et al. 2021), in some cases with improved sensitivity when FTA cards were used (Tam et al. 2016, Hashimoto et al. 2019, Longhi et al. 2023). In the current study, DNA integrity of DBS deteriorated significantly compared to that of cold chain stored samples, but retained at maximum 20,000 bp fragments regardless of the storage period and filter paper used. Shorter storage periods within 3 years of collection preserved DNA fragments with an average length of 10,000 bp, even in some cases when less optimal RNA-specific isolation method were applied. DBS samples stored longer than 9 years in a few cases preserved large DNA fragments, particularly for the samples preserved on FTA cards but the results were highly sample dependent. This may be attributed to the chemical protection of the cards, which likely prevented large DNA fragment from enzymatic and environmental effects when storage was prolonged. However, one case of large-size DNA preservation (e.g. 20,000 bp DNA) was observed in DBS stored on W3 after nearly 15 years storage. Therefore, minimal invasive DBS technology regardless of filter paper used may hold great promise for genomic research on wildlife animals.

Collectively, our results suggest that DNA preservation of DBS on cellulose-based W3 was comparable to that of chemically-treated FTA cards under long-term storage conditions, with DNA retrieval efficiency strongly depending on the nucleic acid isolation methods applied. DNA fragmentation tend to deteriorate more seriously for samples stored on W3 over time, which may limit the application of the DBS samples in genomic-targeted research. Considering material expense and ease of implementation, DBS samples preserved on basic untreated filter paper could be an option for DNA based projects.

When comparing the samples stored and extracted with the same methods, much of the observed degradation was stochastic and sample specific, especially for FTA stored samples. Local hunters collected the samples and the protocol was basic and uncomplicated. However, they remained in the field for days to weeks introducing substantial among sample preservation variation e.g. samples collected later in a given season would be better preserved than the first samples collected due to semiannual or annual sampling transporting to laboratory settings. Such factors may have contributed to the variable degradation profiles even though samples were subsequently stored under the same conditions at each sample site. Besides, minor collection volume differences, microclimatic differences or species specific blood integrity may be responsible for the stochasticity but remain to be explored.

Overall, the sampling strategies with minimal invasiveness provide invaluable biological resource for wildlife research. Our sample-collection method was integrated within a community-based participatory program aimed at improving the conservation and sustainability of natural resources and the livelihoods of local Amazon communities, who rely on subsistence hunting for food security (Jacob et al. 2023). The collection of blood samples from wild animals hunted for self-consumption by the local population allows obtaining valuable biological material from wildlife in natural habitats without taking additional samples to those normally collected (Mayor et al. 2017). In our study area, the local population discards such biological samples. Therefore, the collection of biological samples is compatible with local culture; however, training is required to properly obtain, identify, and preserve the samples. This simple and low-cost method allows for the efficient collection of a high diversity of wild species hunted for self-consumption, allowing to improve the understanding direct health risks of local communities.

2. Mining of capture sequencing data generated from degraded DNA

While metagenomic approaches enable rapid and comprehensive discovery and characterization of viruses, the approaches tend to be less suitable and less cost-efficient for highly degraded wildlife genetic material, particularly when the genome size of the study subject is large (McCartney-Melstad et al. 2016). Target capture sequencing can improve results from such samples as it enriches the yields of

target sequences (Mamanova et al. 2010) and reduces background when applied to historical DNA (Willerslev and Cooper 2005). Extensive studies have been conducted for evaluating the practicability of both DNA and RNA virus screening on DBS samples by using amplification-dependent methods (Smit et al. 2014, Lange et al. 2017, Cardona-Ospina et al. 2019). We targeted DNA and RNA viruses using a panel of RNA baits targeting a wide spectrum of vertebrate viruses. Our results showed that a median of ~15.94% sequencing data per sample were viral candidate sequences based on blasting against the NCBI viral genome reference database, while ~83.39% of the candidates were bacteriophage hits. Neither mammalian RNA nor DNA viruses, after excluding bacteriophage reads, were identified using two distinct taxonomical workflows with stringent filtering steps. A previous large-cohort study in humans without any infectious disease indicated only a tiny fraction of ~5% reads in 42% of the study participants were viral reads (Moustafa et al. 2017). Among them, the number of RNA virus reads was particularly low (Moustafa et al. 2017). Viral titre in uncharacterized wildlife is unknown but can presumably be low in presumably healthy sampled wildlife. In addition, although RNA viruses can be amplified and detected after months of storage on FTA cards (Abdelwhab et al. 2011, Cardona-Ospina et al. 2019), RNA was more likely to degrade more severely than DNA. The highly degraded nature of the DBS blood spots likely precluded the detection of RNA viruses. The absence of DNA viruses such as herpes or poxviruses (though positive for several by targeted PCR screening, data not shown) could suggest that none of the animals were infected, the titre was below detection limits or the DNA was too degraded to detect them.

In target capture sequencing experiments, non-target sequences are commonly generated, especially when using broad capture panels or when capturing sequences with high similarity to unrelated genomic regions (Picardi and Pesole 2012, Diroma et al. 2014). Previous studies targeting exome or nuclear ultra-conserved elements showed that mitochondrial sequences tend to be highly-represented in the non-target sequences (Picardi and Pesole 2012, Diroma et al. 2014, Raposo do Amaral et al. 2015, Dashti et al. 2021), consisting of approximately 1–5% of reads in exome sequencing data (Samuels et al. 2013). A recent study demonstrates that DBS samples are robust for WGS analysis, allowing for the effective retrieval of whole mitochondrial genomes (Agrawal et al. 2021). In the current study, a median of ~43.67% total viral-target sequencing data per sample were mitochondrial sequences. Sequences were heavily biased to the ribosomeal RNA regions consistently in all seven mitochondrial assemblies, particularly in the16S rRNA gene, leading to an average of ~10× higher coverage depth than the other regions. Hybridization capture allows for a high proportion of mismatched sites, which can improve the performance of detecting species that are highly divergent from those represented by reference sequences. Mitochondrial DNA experiences a slower degradation rate compared to nuclear DNA (Schwarz et al. 2009, Allentoft et al. 2012) likely due to the relative abundance of mitochondrial DNA compared to nuclear DNA (Bogenhagen and Clayton 1978). We observed a positive correlation between

the number of reads mapped to virus reference databases and those mapping to mitochondrial genome references though this did not reach statistical significance. This may indicate that viral baits with limited homology to mitochondrial sequences enriched them. Overall, our results demonstrated that non-target sequences generated in the sequencing of capturing non-host sequences could also be biological valuable.

Conclusion

Whatman filter paper no.3 and FTA cards can be used for long-term DNA storage, with retrieval efficiency strongly depending on the nucleic acid isolation methods applied. DNA-specific isolation protocols with high temperature incubation steps likely increase the efficiency of DNA retrieval from DBS samples. Overall DNA quality of DBS samples was significantly lower than that of cold chain stored samples, while relatively large DNA fragment could still be preserved even after nearly 15 years storage on basic filter paper without chemical protection. However, the application of DBS to viral detection with capture based approaches, particularly for RNA viruses, was unsuccessful. Furthermore, off target mitochondrial genome sequences could be effectively assembled from the target capture sequencing data, indicating the potential utility of DBS samples in wildlife genetic and genomic research.

Ethical statements and considerations

The human study was performed with the approval of the Ethics Committee of the Universidad Peruana Cayetano Heredia (102142), the Ethics Committee of the Universidad Autónoma de Barcelona (CEEAH 4829), and the Ethics Committee of the Hospital Clínic de Barcelona (HCB/2019/1107). The study was submitted to the local authorities of Nueva Esperanza, as well as to the regional and local authorities (n° 267-2019-GRL-DRSL/30.09.01). Research protocols for wildlife sampling were approved by the Peruvian Forest and Wildlife Service (n° 258-2019-MINAGRI-SERFOR-DGGSPFFS) and the Institutional Ethical Committee for Animal Use of the Universidad Peruana Cayetano Heredia (ref. 102142). Dried wildlife blood samples on filter paper were exported from Peru to Germany with CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) approval through the Peruvian Forestry and Wildlife Service - SERFOR (N° 003258/SP, N° 003260/SP, N° 003568-SERFOR, N° 003579-SERFOR) in accordance with the Nagoya Protocol.

Acknowledgement

This work has received financial support from the ERANet-LAC [grant 17/HLH-0271] within the research project (Contract N° 136-2018-FONDECYT; AC18/00054 Instituto de Salud Carlos; CNPq - 400800/2019-5). JL was supported by the China Scholarship Council. GMU had two fellowships from

the program PPGSPAA by the CNPq in GD modality (140312/2020-0) and SWE modality (201546/2020-5).

Data accessibility

The sequencing data generated in the current study are deposit in NCBI SRA under the BioPorject number PRJNA1024660. *De novo* assembled mitochondrial genomes of three primate species were deposit in NCBI GenBank with accession number OR689233-OR689239.

Author contributions

ADG, PM, MLSR, JL and GMU developed the project. JL and GMU performed nucleic acid isolations and assessment. JL conduct target capture sequencing, bioinformatics analysis and statistical analysis. MLSR and GMU organized the sample collected and project implementation in the field. PM provided expertise in Peruvian wildlife. JL and ADG wrote the paper. All authors read, approved and agreed on the final version of the manuscript.

Declaration of competing interest

The authors declare that no conflicts of interest exist to influence the work reported in this paper.

Benefit-Sharing Statement

Benefits generated: Permission was obtained from the regional health authorities through a presentation and a detailed copy of the research project. We requested permission to enter the Apu, the main authority, of the Nueva Esperanza indigenous community to proceed with the fieldwork. In a general meeting, we explained to the community members the objectives of the project and the procedure for collecting biological samples from the participating population. The informed consent form was read to each participant according to the regulations of the Ethics Committee, and a copy of the document signed by both parties was given. Indigenous peoples practice subsistence hunting as a cultural heritage and for food security. Therefore, as part of a community participatory program aimed at improving the conservation and sustainability of natural resources and livelihoods of indigenous peoples, community hunters were trained to collect blood from the hunted animal on filter paper and to keep a hunting register with prey data. The contributions of all persons involved in the research, including the participating community members and hunters, are described in the METHODS AND ACKNOWLEDGEMENTS, and a report of the field work has been submitted to the relevant health authorities, as well as the communication of the results to the population of the community of Nueva Esperanza. An important collaboration has been developed between researchers from Peru, Spain, Brazil and Germany, the collaborators are included as co-authors, the research covers a priority concern, as the collection of blood from locally hunted wild animals allows to evaluate the health status of wild populations in natural habitats without incentivizing hunting and taking advantage of discarded material. Finally, as described above, all data have been made available to the public through appropriate biological databases.

References

Abdelwhab, E. M., D. Lüschow, T. C. Harder, and H. M. Hafez. 2011. The use of FTA® filter papers for diagnosis of avian influenza virus. J Virol Methods **174**:120-122.

Agrawal, P., S. Katragadda, A. K. Hariharan, V. G. Raghavendrachar, A. Agarwal, R. Dayalu, D. Awasthy, S. C. Sharma, Y. K. Sivasamy, P. Lakshmana, A. Shanmugam, V. Veeramachaneni, V. Gupta, B. P. Vani, L. Subaiya, T. S. Syamala, R. Hariharan, V. Chandru, and D. E. Bloom. 2021. Validation of whole genome sequencing from dried blood spots. Bmc Medical Genomics **14**.

Allentoft, M. E., M. Collins, D. Harker, J. Haile, C. L. Oskam, M. L. Hale, P. F. Campos, J. A. Samaniego, M. T. Gilbert, E. Willerslev, G. Zhang, R. P. Scofield, R. N. Holdaway, and M. Bunce. 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. Proc Biol Sci **279**:4724-4733.

Allio, R., A. Schomaker-Bastos, J. Romiguier, F. Prosdocimi, B. Nabholz, and F. Delsuc. 2020. MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. Molecular Ecology Resources **20**:892-905.

Andrews, K. R., deBarba, M., Russello, M. A., & Waits, L. P. 2018. Advances in using non-invasive, archival, and environmental samples for population genomic studies. In O. P. Rajora (Ed.), Population Genomics: Wildlife. Springer Nature Switzerland AG.

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, and P. A. Pevzner. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology **19**:455-477.

Bassaganyas, L., G. Freedman, D. Vaka, E. Wan, R. Lao, F. Chen, M. Kvale, R. J. Currier, J. M. Puck, and P. Y. Kwok. 2018. Whole exome and whole genome sequencing with dried blood spot DNA without whole genome amplification. Human Mutation **39**:167-171.

Bezerra, C. S., M. M. Portilho, C. C. Frota, and L. M. Villar. 2021. Comparison of four extraction methods for the detection of hepatitis B virus DNA in dried blood spot samples. Microbiologyopen **10**.

Bogenhagen, D., and D. A. Clayton. 1978. Mechanism of mitochondrial DNA replication in mouse L-cells: introduction of superhelical turns into newly replicated molecules. J Mol Biol **119**:69-81.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**:2114-2120.

Camacho-Sanchez, M., P. Burraco, I. Gomez-Mestre, and J. A. Leonard. 2013. Preservation of RNA and DNA from mammal samples under field conditions. Mol Ecol Resour **13**:663-673.

Capiau, S., H. Veenhof, R. A. Koster, Y. Bergqvist, M. Boettcher, O. Halmingh, B. G. Keevil, B. C. P. Koch, R. Linden, C. Pistos, L. M. Stolk, D. J. Touw, C. P. Stove, and J. W. C. Alffenaar. 2019. Official International Association for Therapeutic Drug Monitoring and Clinical Toxicology Guideline: Development and Validation of Dried Blood Spot-Based Methods for Therapeutic Drug Monitoring. Therapeutic Drug Monitoring **41**:409-430.

Cardona-Ospina, J. A., M. F. Villalba-Miranda, L. A. Palechor-Ocampo, L. I. Mancilla, and J. C. Sepulveda-Arias. 2019. A systematic review of FTA cards (R) as a tool for viral RNA preservation in fieldwork: Are they safe and effective? Preventive Veterinary Medicine **172**.

Carroll, E. L., M. W. Bruford, J. A. DeWoody, G. Leroy, A. Strand, L. Waits, and J. Wang. 2018. Genetic and genomic monitoring with minimally invasive sampling methods. Evol Appl **11**:1094-1119.

Chaisomchit S, Wichajarn R, and C. W. Janejai N. 2005. Stability of genomic DNA in dried blood spots stored on filter paper. Southeast Asian J Trop Med Public Health **36**:270-273.

da Cunha Santos, G. 2018. FTA Cards for Preservation of Nucleic Acids for Molecular Assays: A Review on the Use of Cytologic/Tissue Samples. Arch Pathol Lab Med **142**:308-312.

Dashti, M., H. Alsaleh, M. Eaaswarkhanth, S. E. John, R. Nizam, M. Melhem, P. Hebbar, P. Sharma, F. Al-Mulla, and T. A. Thanaraj. 2021. Delineation of Mitochondrial DNA Variants From Exome Sequencing Data and Association of Haplogroups With Obesity in Kuwait. Frontiers in Genetics **12**.

Demirev, P. A. 2013. Dried blood spots: analysis and applications. Anal Chem **85**:779-789.

Dierckxsens, N., P. Mardulyn, and G. Smits. 2016. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Research **45**:e18-e18.

Diroma, M. A., C. Calabrese, D. Simone, M. Santorsola, F. M. Calabrese, G. Gasparre, and M. Attimonelli. 2014. Extraction and annotation of human mitochondrial genomes from 1000 Genomes Whole Exome Sequencing data. Bmc Genomics **15**.

Fowler, K. E., C. P. Reitter, G. A. Walling, and D. K. Griffin. 2012. Novel approach for deriving genome wide SNP analysis data from archived blood spots. BMC Res Notes **5**:503.

Guthrie, R., and A. Susi. 1963. A Simple Phenylalanine Method for Detecting Phenylketonuria in Large Populations of Newborn Infants. Pediatrics **32**:338-343.

Hashimoto, M., M. Bando, J. Kido, K. Yokota, T. Mita, K. Kajimoto, and M. Kataoka. 2019. Nucleic acid purification from dried blood spot on FTA Elute Card provides template for polymerase chain reaction for highly sensitive Plasmodium detection. Parasitology International **73**.

Hearps, A. C., C. E. Ryan, L. M. Morris, M. M. Plate, V. Greengrass, and S. M. Crowe. 2010. Stability of Dried Blood Spots for HIV-1 Drug Resistance Analysis. Current Hiv Research **8**:134-140.

Hollegaard, M. V., P. Thorsen, B. Norgaard-Pedersen, and D. M. Hougaard. 2009. Genotyping whole-genome-amplified DNA from 3- to 25-year-old neonatal dried blood spot samples with reference to fresh genomic DNA. Electrophoresis **30**:2532-2535.

Hsiao, K.-M., H.-M. Lin, H. Pan, T.-C. Li, S.-S. Chen, S.-B. Jou, Y.-L. Chiu, M.-F. Wu, C.-C. Lin, and S.-Y. Li. 1999. Application of FTA sample collection and DNA purification system on the determination of CTG trinucleotide repeat size by PCR-based southern blotting. Journal of Clinical Laboratory Analysis **13**:188-193.

Jacob, M. C. M., A. Medeiros Souza, A. Martins de Carvalho, C. F. Alves de Vasconcelos Neto, D. Tregidgo, D. Hunter, F. de Oliveira Pereira, G. Ros Brull, H. V. Kunhlein, L. Juliane Guedes da Silva, L. Mont'Alverne Jucá Seabra, M. d. P. Drewinski, N. Menolli Jr, P. Carignano Torres, P. Mayor, P. F. M. Lopes, R. R. Vasconcelos da Silva, S. Marcelino Gomes, and J. K. da Silva-Maia. 2023. Food Biodiversity as an Opportunity to Address the Challenge of Improving Human Diets and Food Security. Ethnobiology and Conservation.

Jones, S., C. J. Sutherland, C. Hermsen, T. Arens, K. Teelen, R. Hallett, P. Corran, M. van der Vegte-Bolmer, R. Sauerwein, C. J. Drakeley, and T. Bousema. 2012. Filter paper collection of Plasmodium falciparum mRNA for detecting low-density gametocytes. Malaria Journal **11**.

Kopylova, E., L. Noe, and H. Touzet. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics **28**:3211-3217.

Koulman, A., P. Prentice, M. C. Y. Wong, L. Matthews, N. J. Bond, M. Eiden, J. L. Griffin, and D. B. Dunger. 2014. The development and validation of a fast and robust dried blood spot based lipid profiling method to study infant metabolism. Metabolomics **10**:1018-1025.

Kumar, A., S. Mhatre, S. Godbole, P. Jha, and R. Dikshit. 2019. Optimization of extraction of genomic DNA from archived dried blood spot (DBS): potential application in epidemiological research & bio banking. Gates Open Research **2**.

Langdon, W. B. 2015. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. BioData Min **8**:1.

Lange, B., T. Roberts, J. Cohn, J. Greenman, J. Camp, A. Ishizaki, L. Messac, E. Tuaillon, P. van de Perre, C. Pichler, C. M. Denkinger, and P. Easterbrook. 2017. Diagnostic accuracy of detection and quantification of HBV-DNA and HCV-RNA using dried blood spot (DBS) samples - a systematic review and meta-analysis. Bmc Infectious Diseases **17**:71-85.

Lehmann, S., C. Delaby, J. Vialaret, J. Ducos, and C. Hirtz. 2013. Current and future use of "dried blood spot" analyses in clinical chemistry. Clinical Chemistry and Laboratory Medicine **51**:1897-1909.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**:1754-1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and S. Genome Project Data Processing. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078-2079.

Lim, M. D. 2018. Dried Blood Spots for Global Health Diagnostics and Surveillance: Opportunities and Challenges. Am J Trop Med Hyg **99**:256-265.

Longhi, S. A., L. J. G. Casares, A. A. Munoz-Calderon, J. Alonso-Padilla, and A. G. Schijman. 2023. Combination of ultra-rapid DNA purification (PURE) and loop-mediated isothermal amplification (LAMP) for rapid detection of Trypanosoma cruzi DNA in dried blood spots. Plos Neglected Tropical Diseases **17**.

Magoc, T., and S. L. Salzberg. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics **27**:2957-2963.

Mamanova, L., A. J. Coffey, C. E. Scott, I. Kozarewa, E. H. Turner, A. Kumar, E. Howard, J. Shendure, and D. J. Turner. 2010. Target-enrichment strategies for next-generation sequencing. Nat Methods **7**:111-118.

Marchand, A., I. Roulland, F. Semence, and M. Ericsson. 2021. EPO transgene detection in dried blood spots for antidoping application. Drug Testing and Analysis **13**:1888-1896.

Mareninov, S., J. De Jesus, D. E. Sanchez, A. B. Kay, R. W. Wilson, I. Babic, W. Chen, D. Telesca, J. J. Lou, L. Mirsadraei, T. P. Gardner, N. Khanlou, H. V. Vinters, B. B. Shafa, A. Lai, L. M. Liau, P. S. Mischel, T. F. Cloughesy, and W. H. Yong. 2013. Lyophilized brain tumor specimens can be used for histologic, nucleic acid, and protein analyses after 1 year of room temperature storage. J Neurooncol **113**:365-373.

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

Martin, N. J., and H. J. Cooper. 2014. Challenges and opportunities in mass spectrometric analysis of proteins from dried blood spots. Expert Review of Proteomics **11**:685-695.

Mayor, P., H. El Bizri, R. E. Bodmer, and M. Bowler. 2017. Assessment of mammal reproduction for hunting sustainability through community-based sampling of species in the wild. Conserv Biol **31**:912-923.

McCartney-Melstad, E., G. G. Mount, and H. B. Shaffer. 2016. Exon capture optimization in amphibians with large genomes. Mol Ecol Resour **16**:1084-1094.

McClendon-Weary, B., D. L. Putnick, S. Robinson, and E. Yeung. 2020. Little to Give, Much to Gain-What Can You Do With a Dried Blood Spot? Curr Environ Health Rep **7**:211-221.

Michaud, V., P. Gil, O. Kwiatek, S. Prome, L. Dixon, L. Romero, M. F. Le Potier, M. Arias, E. Couacy-Hymann, F. Roger, G. Libeau, and E. Albina. 2007. Long-term storage at tropical temperature of dried-blood filter papers for detection and genotyping of RNA and DNA viruses by direct PCR. J Virol Methods **146**:257-265.

Milne, I., G. Stephen, M. Bayer, P. J. Cock, L. Pritchard, L. Cardle, P. D. Shaw, and D. Marshall. 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform **14**:193-202.

Molteni, C. G., L. Terranova, A. Zampiero, C. Galeone, N. Principi, and S. Esposito. 2013. Comparison of manual methods of extracting genomic DNA from dried blood spots collected on different cards: implications for clinical practice. International Journal of Immunopathology and Pharmacology **26**:779-783.

Moustafa, A., C. Xie, E. Kirkness, W. Biggs, E. Wong, Y. Turpaz, K. Bloom, E. Delwart, K. E. Nelson, J. C. Venter, and A. Telenti. 2017. The blood DNA virome in 8,000 humans. Plos Pathogens **13**.

Naber, S. P. 1996. Continuing Role of a Frozen-tissue Bank in Molecular Pathology. Diagnostic Molecular Pathology **5**:253-259.

Nsubuga, A. M., M. M. Robbins, A. D. Roeder, P. A. Morin, C. Boesch, and L. Vigilant. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Molecular Ecology **13**:2089-2094.

Picardi, E., and G. Pesole. 2012. Mitochondrial genomes gleaned from human whole-exome sequencing. Nat Methods **9**:523-524.

Pitt, J. J. 2010. Newborn screening. Clin Biochem Rev **2**:57-68.

Quinlan, A. R. 2014. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinformatics **47**:11 12 11-34.

Rahikainen, A. L., J. U. Palo, W. de Leeuw, B. Budowle, and A. Sajantila. 2016. DNA quality and quantity from up to 16 years old post-mortem blood stored on FTA cards. Forensic Sci Int **261**:148-153.

Raposo do Amaral, F., L. G. Neves, M. F. Resende, Jr., F. Mobili, C. Y. Miyaki, K. C. Pellegrino, and C. Biondo. 2015. Ultraconserved Elements Sequencing as a Low-Cost Source of Complete Mitochondrial Genomes and Microsatellite Markers in Non-Model Amniotes. PLoS One **10**:e0138446.

Rasolonjatovo, F. S., H. Guis, M. Rajeev, L. Dacheux, L. Arivony Nomenjanahary, G. Razafitrimo, J. T. Rafisandrantantsoa, C. Cêtre-Sossah, J. M. Heraud, and S. F. Andriamandimby. 2020. Enabling animal rabies diagnostic in low-access areas: Sensitivity and specificity of a molecular diagnostic test from cerebral tissue dried on filter paper. PLoS Negl Trop Dis **14**:e0008116.

Riesgo, A., A. R. Perez-Porro, S. Carmona, S. P. Leys, and G. Giribet. 2012. Optimization of preservation and storage time of sponge tissues to obtain quality mRNA for next-generation sequencing. Mol Ecol Resour **12**:312-322.

Russcher, A., A. Enders, C. S. de Brouwer, D. Oepkes, R. Hahn, M. Enders, A. C. M. Kroes, and A. Vossen. 2020. Diagnosis of intrauterine parvovirus B19 infection at birth - Value of DNA detection in neonatal blood and dried blood spots. Journal of Clinical Virology **129**.

Samsonova, J. V., N. Y. Saushkin, and A. P. Osipov. 2022. Dried Blood Spots technology for veterinary applications and biological investigations: technical aspects, retrospective analysis, ongoing status and future perspectives. Vet Res Commun **46**:655-698.

Samuels, D. C., L. Han, J. Li, Q. H. Sheng, T. A. Clark, Y. Shyr, and Y. Guo. 2013. Finding the lost treasures in exome sequencing data. Trends in Genetics **29**:593-599.

Schilling, A. K., M. V. Mazzamuto, and C. Romeo. 2022. A Review of Non-Invasive Sampling in Wildlife Disease and Health Research: What's New? Animals (Basel) **12**.

Schwarz, C., R. Debruyne, M. Kuch, E. McNally, H. Schwarcz, A. D. Aubrey, J. Bada, and H. Poinar. 2009. New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. Nucleic Acids Res **37**:3215-3229.

Searle, S. R., F. M. Speed, and G. A. Milliken. 2012. Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. The American Statistician **34**:216-221.

Shabihkhani, M., G. M. Lucey, B. Wei, S. Mareninov, J. J. Lou, H. V. Vinters, E. J. Singer, T. F. Cloughesy, and W. H. Yong. 2014. The procurement, storage, and quality assurance of frozen blood and tissue biospecimens in pathology, biorepository, and biobank settings. Clin Biochem **47**:258-266.

Smit, P. W., I. Elliott, R. W. Peeling, D. Mabey, and P. N. Newton. 2014. Review Article: An Overview of the Clinical Use of Filter Paper in the Diagnosis of Tropical Diseases. American Journal of Tropical Medicine and Hygiene **90**:195-210.

Smith, L. M., and L. A. Burgoyne. 2004. Collecting, archiving and processing DNA from wildlife samples using FTA databasing paper. BMC Ecol **4**:4.

Soto-Calderon, I. D., S. Ntie, P. Mickala, F. Maisels, E. J. Wickings, and N. M. Anthony. 2009. Effects of storage type and time on DNA amplification success in tropical ungulate faeces. Molecular Ecology Resources **9**:471-479.

Tam, P. Y. I., N. Hernandez-Alvarado, M. R. Schleiss, F. Hassan-Hanga, C. Onuchukwu, D. Umoru, and S. K. Obaro. 2016. Molecular Detection of Streptococcus pneumoniae on Dried Blood Spots from Febrile Nigerian Children Compared to Culture. PLoS One **11**.

Tang, W., Z. Hu, H. Muallem, and M. L. Gulley. 2012. Quality assurance of RNA expression profiling in clinical laboratories. J Mol Diagn **14**:1-11.

Tuaillon, E., D. Kania, A. Pisoni, K. Bollore, F. Taieb, E. N. O. Ngoyi, R. Schaub, J. C. Plantier, A. Makinson, and P. Van de Perre. 2020. Dried Blood Spot Tests for the Diagnosis and Therapeutic Monitoring of HIV and Viral Hepatitis B and C. Frontiers in Microbiology **11**.

Uttayamakul, S., S. Likanonsakul, R. Sunthornkachit, K. Kuntiranont, S. Louisirirotchanakul, A. Chaovavanich, V. Thiamchai, S. Tanprasertsuk, and R. Sutthent. 2005. Usage of dried blood spots for molecular diagnosis and monitoring HIV-1 infection. Journal of Virological Methods **128**:128-134.

van Schalkwyk, C., J. Maritz, G. U. van Zyl, W. Preiser, and A. Welte. 2019. Pooled PCR testing of dried blood spots for infant HIV diagnosis is cost efficient and accurate. Bmc Infectious Diseases **19**.

van Vliet, K., W. G. van Ginkel, E. van Dam, P. de Blaauw, M. Koehorst, H. A. Kingma, F. J. van Spronsen, and M. R. Heiner-Fokkema. 2020. Dried blood spot versus venous blood sampling for phenylalanine and tyrosine. Orphanet J Rare Dis **15**:82.

Wacharapluesadee, S., P. Phumesin, B. Lumlertdaecha, and T. Hemachudha. 2003. Diagnosis of rabies by use of brain tissue dried on filter paper. Clin Infect Dis **36**:674-675.

Willerslev, E., and A. Cooper. 2005. Ancient DNA. Proc Biol Sci **272**:3-16.

Wu, S. G., W. J. Pan, H. N. Liu, M. L. Byrne-Steele, B. Brown, M. Depinet, X. H. Hou, J. Han, and S. Li. 2019. High throughput sequencing of T-cell receptor repertoire using dry blood spots. Journal of Translational Medicine **17**.

Yozwiak, N. L., P. Skewes-Cox, M. D. Stenglein, A. Balmaseda, E. Harris, and J. L. DeRisi. 2012. Virus identification in unknown tropical febrile illness cases using deep sequencing. PLoS Negl Trop Dis **6**:e1485.

Supplementary materials

Table S1. Information of samples used for target capture sequencing

Species	Sample ID	Collected date(mm/dd/yyyy)	DNA isolation date(mm/dd/yyyy)	Storage days	DIN	Concentration (ng/μl)	DNA fragment peak/bp
Sapajus macrocephalus (Large-headed capuchin)	151320002	09/15/2019	04/22/2021	585	2.1	11.4	475
	151320004	10/16/2019	04/22/2021	554	2.6	86.9	1131
	151320010	01/8/2020	04/29/2021	477	2.7	127	1323
Cebus albifrons (White-fronted capuchin)	151320038	10/27/2019	04/22/2021	543	3.8	47	1422
	151320039	11/18/2019	04/22/2021	521	3.2	25.1	1695
Lagothrix poeppigii (Silvery woolly monkey)	151320054	09/05/2019	04/22/2021	595	1.9	18.7	433
	151320055	09/25/2019	04/22/2021	575	1.7	11.9	451
	151320056	09/28/2019	05/12/2021	592	2.2	43.5	898
	151320065	01/31/2020	05/12/2021	467	2.3	21.5	444
	151320068	02/01/2020	05/12/2021	466	1.7	35	482
	151320086	2019	05/17/2021	Not collected	2.4	16	7778
Cacajao calvus (Bald uakari)	151320109	10/03/2019	04/22/2021	567	2.0	33.8	7747

Table S2. Samples used for target capture sequencing and the associated sequencing data.

Species	Sample ID	Raw data	Percentage of capturable sequences per sample	Quality controled reads	Nuclear genome mapping			
					Alignment reads Bowtie2	Percentage	Reference species	NCBI Accession
Sapajus macrocephalus (Large-headed capuchin)	151320002	1355268	44.98%	1336610	684464	52.21%	*Sapajus apella*	GCF_009761245.1
	151320004	868893	27.97%	855366	5479	0.64%		
	151320010	1151788	41.37%	1132646	10168	0.90%		
Cebus albifrons (White-fronted capuchin)	151320038	709778	25.49%	703448	96799	13.76%	*Cebus albifrons*	GCA_023783575.1
	151320039	526776	18.92%	520339	161187	30.98%		
Lagothrix l. poeppigii (Silvery woolly monkey)	151320054	995900	31.67%	987201	104014	10.54%	*Pithecia pithecia*	GCA_004026645.1
	151320055	1475833	49.83%	1461862	157486	10.77%		
	151320056	871668	28.06%	863916	139483	16.15%		
	151320065	1347626	42.86%	1319420	328930	24.93%		
	151320068	395759	14.21%	391272	91547	23.40%		
	151320086	800604	25.46%	789853	65671	8.31%		
Cacajao calvus (Bald uakari)	151320109	1365747	43.97%	1352834	292611	21.63%		

Table S3. Results from two viral classification pipelines.

Species	Sample ID	Meged reads	Direct blasting					Contig_blasting	
			Total hits	Viral hits over total reads	Phage hits	Phage hits over total viral hits	Viral candidate reads against nt database	Assemble contigs	Viral candidate contigs against nt database
Sapajus macrocephalus	151320002	1278954	56726	4.79%	30967	54.59%		2384	1
	151320004	811355	139547	19.02%	130384	93.43%	15	37349	1
	151320010	1071452	22704	2.38%	17437	76.80%	2	82854	2
Cebus albifrons	151320038	683402	152449	24.34%	142609	93.55%		6278	1
	151320039	493602	97349	21.79%	88707	91.12%		1087	
Lagothrix poeppigii	151320054	946998	166999	18.71%	148564	88.96%		8183	9
	151320055	1411956	288654	22.07%	269815	93.47%	2	12974	1
	151320056	842047	57511	7.72%	41495	72.15%	2	9171	1
	151320065	1253901	142743	12.16%	116317	81.49%		15509	6
	151320068	379661	49231	14.86%	40878	83.03%	1	7881	
	151320086	760330	63755	9.36%	50204	78.75%	1	17005	19
Cacajao calvus	151320109	1304223	200551	17.03%	167970	83.75%	1	6973	10

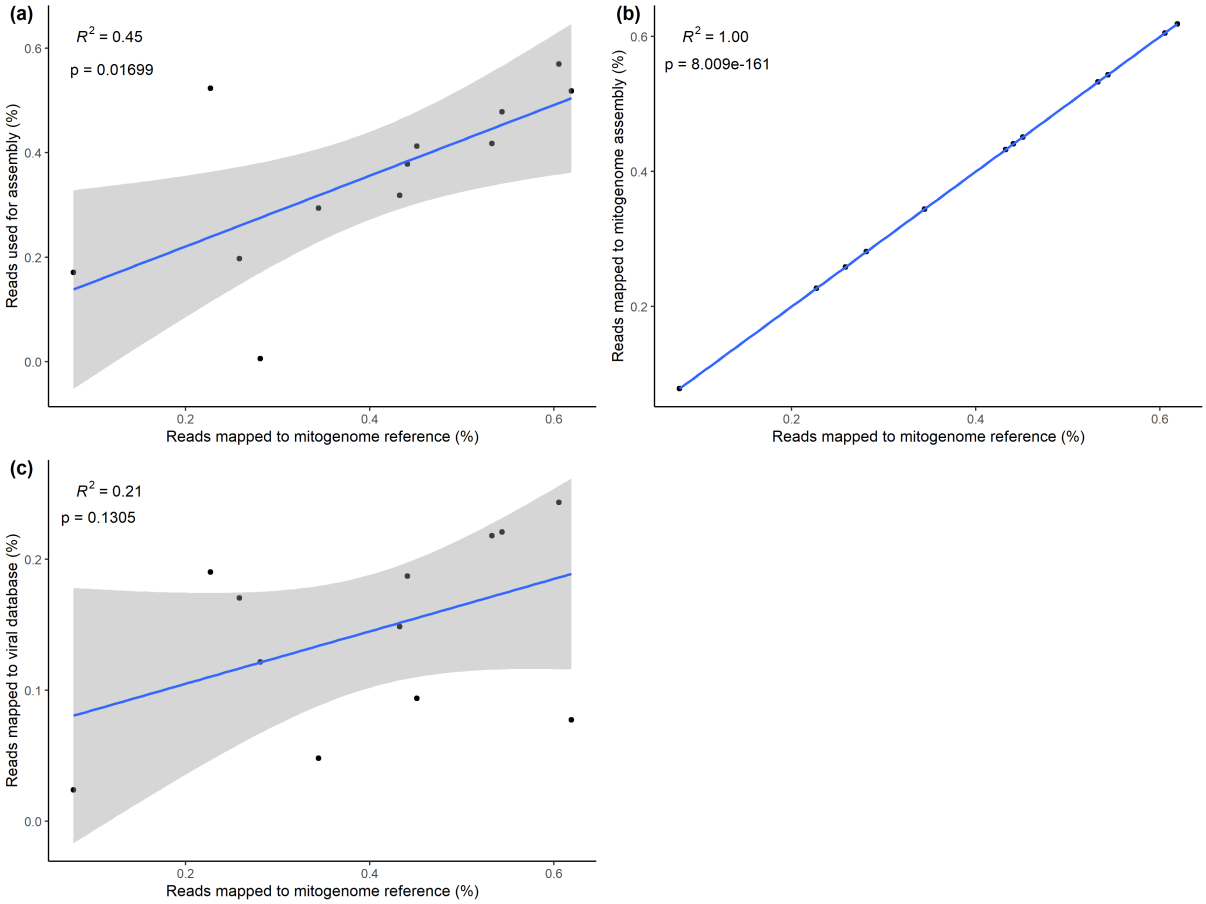

Figure S1. Plots from the linear models (LM) are shown. Plot (a) shows the percentage of reads mapped to mitochondrial genome references versus the percentage of reads used for *de novo* assemblies by Novoplasty. Plot (b) shows the percentage of reads mapped to mitochondrial genome references versus the percentage of reads mapped to *de novo* assembled mitochondrial genomes. Plot (c) the percentage of reads mapped to mitochondrial genome references versus the percentage of reads mapped to viral reference database.

CHAPTER 4: Manuscript 3

The complete mitochondrial genome of the lowland paca (*Cuniculus paca*) and its phylogenetic relationship with other New World hystricognath rodents

Juan Li [1, 5], Pedro Mayor[2, 3], Meddly L. Santolalla Robles[4], Alex D. Greenwood [1, 6*]

[1] Department of Wildlife Diseases, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany

[2] Departament de Sanitat i Anatomia Animals, Universitat Autònoma de Barcelona, Bellaterra, Spain

[3] ComFauna, Comunidad de Manejo de Fauna Silvestre en la Amazonía y en Latinoamérica, Iquitos, Peru

[4] Emerge, Emerging Diseases and Climate Change Research Unit, School of Public Health and Administration, Universidad Peruana Cayetano Heredia, Lima, Peru

[5] Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

[6] Department of Biologie, Chemie, and Pharmazie, Freie Universität Berlin, Berlin, Germany

Abstract

The lowland paca (*Cuniculus paca*) is a nocturnal, widespread and solitary large-sized rodent in the family Cuniculidae, and one of the most frequently hunted mammals in the Neotropical forests of Latin America. We assembled the first complete mitochondrial genome of lowland paca using three closely related hystricognath species as reference sequences. The mitochondrial genome is 16,770 basepairs (bp) in length, with similar characteristics of vertebrate mitochondrial genomes. We performed phylogenetic analyses using 26 mitochondrial genome of hystricognath species based on thirteen protein-coding genes. The result confirms the taxonomical placement among the New World hystricognath rodents with high support. The placement is consistent with previous phylogenetic studies based on individual mitochondrial and nuclear genes. The current study improves the phylogenic resolution of hystricognath rodents.

Keywords: mitogenome, hystricognath, *Cuniculus*, phylogenetics

Introduction

The lowland paca (*Cuniculus paca*, Linnaeus, 1766) is the living species of *Cuniculus*, the only genus in the family Cuniculidae. The pacas were previously placed among the *Agouti* species in the family Dasyprotidae, subfamily Agoutinae, but has been defined as an independent family due to morphological differences (Grzimek 2003). Lowland pacas (Figure 1) are nocturnal, herbivorous and solitary large-sized (6–12 kg) rodents with a relatively low fecundity for rodents (Mayor et al. 2013, Mayor et al. 2017). The species inhabit a variety of forest types in tropical areas within Central and South America from eastern and southern Mexico to northeast Argentina (Patton 2015). Introductions into Cuba and the Lesser Antilles have occurred recently (Patton 2015). Although this species currently is the most hunted Neotropical rodent species (El Bizri et al. 2020a, El Bizri et al. 2020b), wild populations are stable and are listed as Least Concern by The IUCN Red List of Threatened Species in 2016 (Emmons 2016).

Figure 1. Images of male (A) and female (B) *Cuniculus paca* in captivity. The pictures were offered from the coauthor Pedro Mayor in this study.

Materials and methods

A volume of ~100 µl blood sample was collected by local hunters in the Yavarí-Mirin River basin (04°19′53″ S, 71°57′33″ W) in the Peruvian Amazon and was spotted on Whatman FTA card for storage. The specimen and DNA isolate were deposited in the frozen sample archive of Leibniz institute for zoo and wildlife research (IZW, https://www.izw-berlin.de/en/home.html; contact: Alex D. Greenwood, greenwood@izw-berlin.de) under the voucher number 151310178. We sliced the FTA card, which was diluted using AVL buffer in QIAamp Viral RNA Mini Kit (Qiagen, Germany). After 10 min of vortex at 2000 rpm, genomic DNA was isolated using DNeasy Blood & Tissue Kit (Qiagen, Germany) following the manufacturer's protocol. Illumina DNA sequencing library following Meyer

and Kircher (2010) was built with an average insert size of 300 bp, and was subsequently sequenced on the illumina MiSeq platform using MiSeq Reagent Kit v2 (PE, 2×150 cycles).

We removed adapter sequences with Cutadapt v.1.15 (Martin 2011) and trimmed low quality reads with Trimmomatic v.0.38 (Bolger et al. 2014). Complete circular mitochondrial genome was assembled with Mitofinder (Allio et al. 2020) by using three complete mitogenomes of hystricognath species as reference sequences: *Cavia aperea* (NC_046949), *Cavia porcellus* (NC_000884) and *Coendou insidiosus* (NC_021387). All reads were then mapped backed to the *de novo* assembled mitogenome sequence using BWA v. 0.7.17 (Li and Durbin 2009). We called consensus sequence with default parameters using bcftools v.1.7 (Danecek and McCarthy 2017). The resultant mitochondrial genome was annotated using Mitofinder (Allio et al. 2020) after reverse complementing.

A phylogenetic analysis was performed using 13 PCGs of the lowland paca and 26 hystricognath rodents. Each gene from different species was aligned individually using MAFFT (-linsi) v7.310 (Katoh and Standley 2013). The multiple sequence alignments of the 13 genes were concatenated using catfasta2phyml (https://github.com/nylander/catfasta2phyml). A maximum likelihood tree was estimated using IQ-TREE v.2.0.3 (Minh et al. 2020) with 1000 bootstrap replicates. We applied TVM+I+G as the substitution model for ND6 gene and GTR+I+G for other 12 genes. The substitution models were determined by PartitionFinder v.2.1.1 (Lanfear et al. 2016).

Results

In total, 45,552 merged reads out of 1,658,046 read pairs were mapped back to the assembled mitogenome sequence. The mitochondrial genome sequence has 100% total coverage with an average of 370× base coverage, ranging from 28× to 1890× (Figure S1). The complete mitochondrial genome of the lowland paca is 16,770 bp length with GC content of 40.81%, constituting a control region (position:15,585-16,770) and a set of 37 genes: 13 protein-coding genes (PCGs), 2 ribosomal RNA (rRNA) genes and 22 transfer RNA (tRNA) genes (Figure 2). The nucleotides of the mitogenome are composed of 32.98% adenine (A), 26.21% thymine (T), 13.33% guanine (G) and 27.47% cytosine (C). Majority of the PCGs initiate with common vertebrate start codon ATG, except ND2 and ND3 with ATT, and ND5 and ND6 with ATC. Ten PCGs terminate with stop codons either TAG (ND1, ND2, COX1, ATP8 and ND3) or TAA (COX2, ATP6, ND4L, ND5 and ND6), while CYTB gene terminates with arginine coding codon AGA, COX3 with a single T and ND4 with TA. Phylogenetic analyses shows the lowland paca (*Cuniculus paca*) mitogenome clusters with the genus *Cavia* and is placed in the infraorder Hystricognathi within the hystricognath rodents (Figure 3).

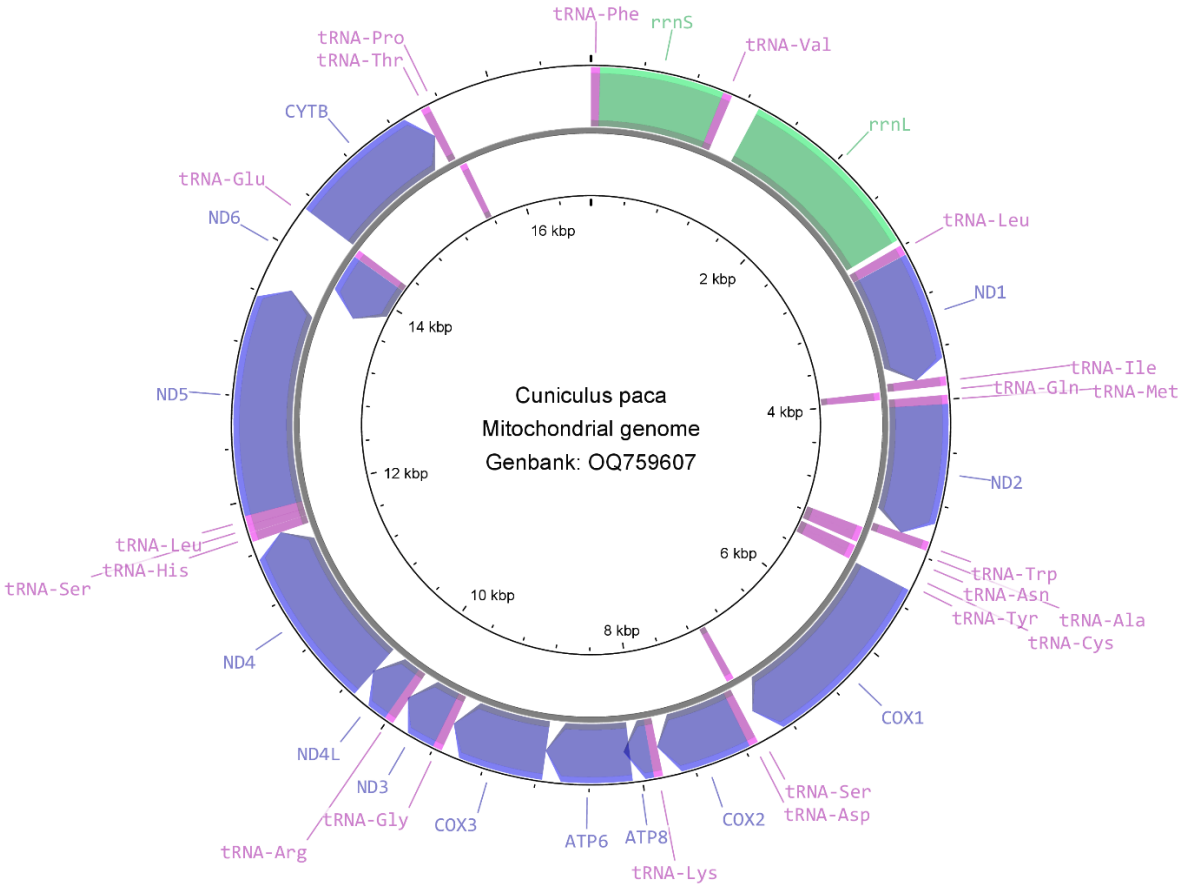

Figure 2. Clockwise view of the mitochondrial genome of the *Cuniculus paca* generated from CGView (Stothard and Wishart 2005).

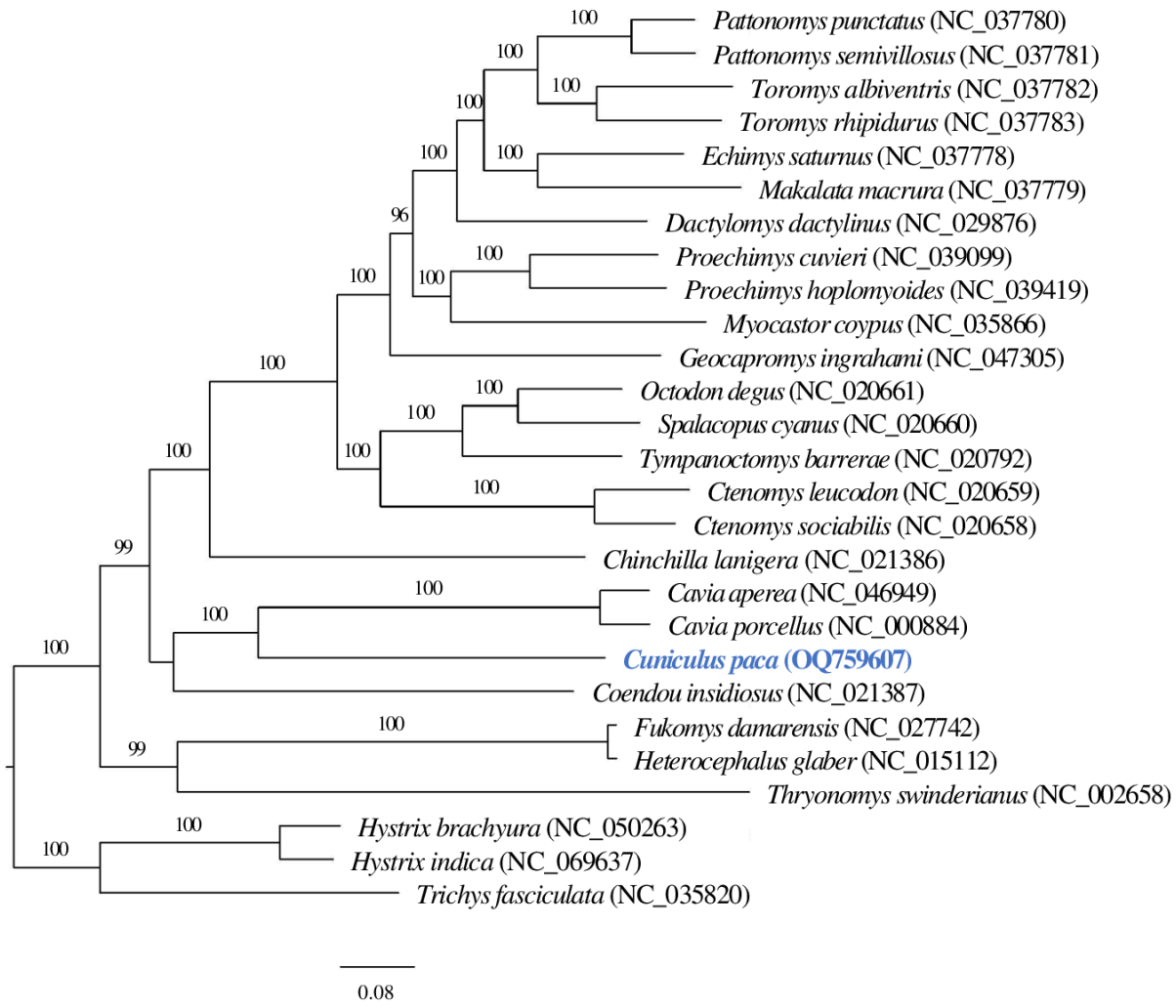

Figure 3. The phylogenetic relationship of lowland paca (*Cuniculus paca)* with 26 hystricognath rodent species are inferred from maximum likelihood estimation based on 13 protein-coding genes with 1000 bootstrap replicates. Support values were given above each branch. The NCBI accession of mitogenome references used were displayed following species scientific names in paracenteses. The tip marked in blue depicts the mitochondrial genome of *Cuniculus paca* generated in this study.

Discussion and conclusion

Phylogenetic analyses confirms the taxonomical position of lowland paca (*Cuniculus paca*) mitogenome in the infraorder Hystricognathi within the hystricognath rodents, showing that *Cuniculus paca* is closely related to the genus *Cavia* (Figure 3). The results are consistent with previous analyses based on multiple mitochondrial and nuclear genes (Voloch et al. 2013). The current study *de novo* assembles a mitochondrial genome of lowland paca (*Cuniculus paca*) and improves the phylogenic resolution of hystricognath rodents.

Data access

The mitochondrial genome sequence supporting this study are publicly released at NCBI GenBank with the accession no. OQ759607. The associated sequencing data is deposited at NCBI Sequencing Reads Archive (SRA) under the BioProject PRJNA950048 (accession number SRR24006478 and BioSample number SAMN33969573).

Ethics statement

Research protocols for NHPs sampling were approved by the Peruvian Forestry and Wildlife Agency (N° 258-2019-MINAGRI-SERFOR-DGGSPFFS) and the Institutional Animal Use Ethics Committee of the Universidad Peruana Cayetano Heredia (ref. SIDISI 102142).

Acknowledgement

This work has received financial support from the ERANet-LAC [grant 17/HLH-0271] within the research project (Contract N° 136-2018-FONDECYT). The authors thank Dorina Meneghini for providing bioinformatic support. JL acknowledges the support of the China Scholarship Council. Finally, the authors would like to acknowledge local people of Nueva Esperanza in the Yavari-Mirin River for their help in collecting biological samples.

Authors' contribution

ADG, PM, MLSR and JL developed the project. JL developed and conducted the laboratory and bioinformatics workflow. PM and MLSR helped the project management and sample collection in Peru, and provided the expertise about the species. JL, ADG and PM wrote the article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Reference

Allio, R., A. Schomaker-Bastos, J. Romiguier, F. Prosdocimi, B. Nabholz, and F. Delsuc. 2020. MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. Molecular Ecology Resources **20**:892-905.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**:2114-2120.

Danecek, P., and S. A. McCarthy. 2017. BCFtools/csq: haplotype-aware variant consequences. Bioinformatics **33**:2037-2039.

El Bizri, H. R., T. Q. Morcatty, J. C. Ferreira, P. Mayor, C. F. A. Vasconcelos Neto, J. Valsecchi, V. Nijman, and J. E. Fa. 2020a. Social and Biological Correlates of Wild Meat Consumption and Trade by Rural Communities in the Jutaí River Basin, Central Amazonia. Journal of Ethnobiology **40**.

El Bizri, H. R., T. Q. Morcatty, J. Valsecchi, P. Mayor, J. E. S. Ribeiro, C. F. A. Vasconcelos Neto, J. S. Oliveira, K. M. Furtado, U. C. Ferreira, C. F. S. Miranda, C. H. Silva, V. L. Lopes, G. P. Lopes, C. C. F. Florindo, R. C. Chagas, V. Nijman, and J. E. Fa. 2020b. Urban wild meat consumption and trade in central Amazonia. Conserv Biol **34**:438-448.

Emmons, L. 2016. *Cuniculus paca*. The IUCN Red List of Threatened Species 2016: e.T699A22197347.

Grzimek, B. 2003. Grzimek's Animal Life Encyclopedia, Volume 16: Mammals V. Farmington Hills, MI: Gale Group.

Katoh, K., and D. M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution **30**:772-780.

Lanfear, R., P. B. Frandsen, A. M. Wright, T. Senfeld, and B. Calcott. 2016. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Molecular Biology and Evolution **34**:772-773.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**:1754-1760.

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

Mayor, P., H. El Bizri, R. E. Bodmer, and M. Bowler. 2017. Assessment of mammal reproduction for hunting sustainability through community-based sampling of species in the wild. Conserv Biol **31**:912-923.

Mayor, P., D. A. Guimaraes, and C. Lopez. 2013. Functional morphology of the genital organs in the wild paca (Cuniculus paca) female. Anim Reprod Sci **140**:206-215.

Meyer, M., and M. Kircher. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc **2010**:pdb prot5448.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution **37**:1530-1534.

Patton, J. L. P., U.F.J. D'Elía, G. 2015. Mammals of South America, Volume 2 Rodents. University of Chicago Press.

Voloch, C. M., J. F. Vilela, L. Loss-Oliveira, and C. G. Schrago. 2013. Phylogeny and chronology of the major lineages of New World hystricognath rodents: insights on the biogeography of the Eocene/Oligocene arrival of mammals in South America. BMC Res Notes **6**:160.

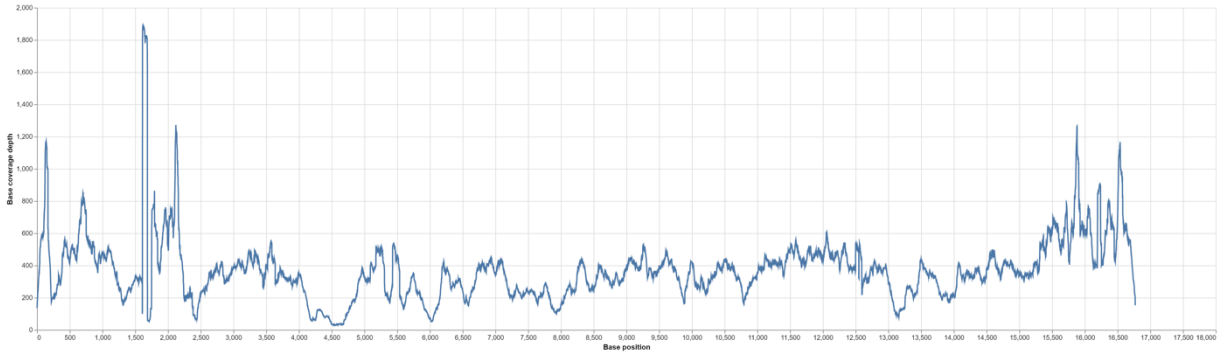

Figure S1. The base coverage depth of the assembled mitochondrial genome sequence of *Cuniculus paca*.

CHAPTER 5: General Discussion

In this thesis, I aimed to evaluate the use of non-invasive environmental and minimally-invasive biological samples in wildlife monitoring and biological research, and to optimize the associated lab workflow and downstream data exploration. For wildlife monitoring, I sampled water and sediment samples over the dry season from waterholes in Cambodia. Two independent eDNA methods have been applied to retrieve genetic information of terrestrial wildlife: hybridization capture and metabarcoding. Hybridization capture was conducted by using a set of RNA oligonucleotide baits referred from the complete mitochondrial genomes of 38 representative terrestrial mammalian species, aiming for capturing as many mitochondrial sequences as possible. Metabarcoding was carried out by using a pair of 16S rRNA primers that specifically target tetrapod species. Hybridization capture provided more genetic information, in some cases, capturing over 70% of mitochondrial genomes. Metabarcoding can be more sensitive to abundant and large species. The results demonstrated that environmental samples can collectively be used for detecting the presence of terrestrial wildlife, which can be more sensitive to small species that are overlooked in observation-based monitoring. In order to evaluate minimal-invasive DBS sampling and the stability of preserved nucleic acid, I used the DNA quality data (DNA integrity) generated from DBS samples stored in two types of filter paper (FTA card and Whatman no.3) for 1-15 years using three different isolation methods. The preservation capacity was assessed in comparison with samples stored in a cold chain. A subset of isolated samples were used for viral screening by applying target enrichment prior to sequencing with hybridization capture. In the bioinformatic analyses, I explored both target sequences for viral detection and non-target sequences for mitochondrial genome retrieval. The results demonstrated that DBS sampling can preserve DNA samples regardless of which filter paper was used albeit in a highly degraded state. Complete mitochondrial genomes were assembled using non-target sequencing data. However, the preservation technology is relatively limited in terms of viral screening research likely due to the low RNA preservation capacity. For assessing the use of low quality DBS samples in genomic research, I performed next-generation sequencing on DBS samples of *Cuniculus paca* for *de novo* mitochondrial genome assembly. One complete mitochondrial genome of the lowland paca was successfully *de novo* assembled.

All living creatures shed DNA into their surrounding environment (e.g. feces, urine, skin, or hairs), which forms a natural pool of the genetic materials of all taxa in the habitat. The non-invasive sampling strategy has been widely used for aquatic species monitoring primarily by metabarcoding methods (Seeber and Epp 2022), such as invasive species detection (Rees et al. 2014, Bowers et al. 2021) and species diversity estimate (Adams et al. 2019, Alexander et al. 2019). More recently,

eDNA-based approaches have been proposed for terrestrial wildlife monitoring (Ushio et al. 2017, Williams et al. 2018, Harper et al. 2019, Seeber et al. 2019, Leempoel et al. 2020, Mena et al. 2021), while non-representative sampling can challenge the in-field implementation. In seasonal ecosystems, the dry due to water scarcity, limited water bodies tend to concentrate shed DNA after animals congregate at water sources. In **chapter 2**, I collected water and sediment samples from waterholes in Srepok wildlife sanctuary during the dry season of Cambodia. Two independent eDNA methods were used, which detected mammals, reptiles, amphibians and birds with a total of 17 species. Both eDNA methods performed better for targeted mammals than other taxa, as reflected by dominant detections of mammal species. The rate and frequency of species detection were consistent with the contemporary camera trap records (Pin et al. 2020). Ungulates were the most frequently detected species. The detection frequency of individual ungulate species was positively correlated to the baseline monitoring results (Groenenberg 2020, Pin et al. 2020), indicating eDNA methods can reliably reflect the relative abundance of terrestrial animals, even for low-density species (e.g Eld's deer). No significant differences were observed between the two eDNA methods in terms of species detection, while metabarcoding showed a higher sensitivity to detect abundant and large animals (e.g wild boar, Asian elephant). In comparison with the contemporary camera trap results (Pin et al. 2020), the eDNA methods detected ten small body size species, eight of which Were exclusive detections. Although the camera trap method provides dynamic recording in the field, small animals often are overlooked, particularly reptiles and amphibians. eDNA methods with more tailored lab workflows may help fill the monitoring gaps, constructing a more complete surveillance network for forest wildlife. Furthermore, in wildlife monitoring, other than species presence detection, genetic monitoring is of great importance to estimate the diversity and conservation status of wildlife populations (Andres et al. 2023). In our study, the hybridization capture pipeline enabled us to retrieve more genomic information: two detected species with over 70% of mitochondrial genome coverage and four with over 10% coverage. With enough sequencing effort, the retrieved genomic information may allow for further genetic exploration, such as individual identification, or population genetic analysis in the future (Sigsgaard et al. 2017, Jensen et al. 2021). Our results demonstrated that the seasonal sampling strategy in tropical forest can provide representative samples for eDNA-based terrestrial wildlife monitoring.

When using eDNA-based approaches, post-enrichment sequencing data are generally more difficult to parse than metabarcoding data, requiring customized bioinformatic pipeline tailored to the specific project. The limitation becomes more obvious when the monitoring or surveillance are scaled up. In **chapter 2**, I developed a bioinformatic pipeline to optimize the data exploration for more accurate species classification and effective genetic information retrievals by combining mapping against the customized database and re-aligning via blast searches against the complete reference database.

Compared to metabarcoding and camera trap results, the pipeline can effectively detect the species presence but is restricted to the availability of reference sequences. Our results suggest that the expansion of the reference database over time may further empower eDNA-based approaches in terrestrial wildlife monitoring.

Genetic analyses requiring preserved nucleic acid samples are routinely conducted in many fields of wildlife research, providing information on biology, ecology and pathology of wild animals relevant to conservation and management (Fitzpatrick et al. 2020, Hohenlohe et al. 2021, Kardos et al. 2021, Benham and Bowie 2023). A particular challenge in wildlife genetic research is sample collection and preservation in the fieldwork that does not allow for lab-standard sample storage. Moreover, invasive biological sampling can be detrimental to wild animal populations, particularly for threatened species. Noninvasive or minimal-invasive samplings thus have been commonly used due to the ease of access and use (Fowler et al. 2012, Andrews et al. 2016, Staunstrup et al. 2016, Agrawal et al. 2021), while the sample types often come with the drawback of being highly degraded. DBS technology provides a solution considering its low sample volume requirement, and high sample stability and mobility although sample degradation persists (McClendon-Weary et al. 2020). Currently, DBS preservation is commonly used in clinical diagnosis and medical research (Demirev 2013, Lim 2018, McClendon-Weary et al. 2020, Russcher et al. 2020, Tuaillon et al. 2020), while the significant value in the area of wildlife genetic and genomic research is underexplored. Multiple materials can be used for DBS sample preservation. FTA card, chemically treated cellulose-based paper, is widely used in nucleic acid target research (da Cunha Santos 2018, Hashimoto et al. 2019, Longhi et al. 2023), but with higher cost than basic untreated Whatman cellulose filter paper. DNA preservation of DBS on the two most common preservation papers is under characterized. In **chapter 3**, I isolated DNA from DBS samples stored on FTA card and Whatman no.3 filter paper for up to 15 years in comparison with blood samples stored with a cold chain using three different isolation workflows. Altough nucleic acids were highly degraded for all DBS samples, the elongated storage period did not significantly affect the overall quality of DNA preserved on either of the filter papers. The economic Whatman no.3 paper showed a comparable DNA storage performance to FTA card, which was statistically correlated with the isolation methods applied. The high-temperature incubation based approach greatly improved the integrity of DNA isolated from both preservation papers, while the DNA from FTA cards tended to retain larger DNA fragments. The results demonstrated that DBS storage can be used for biobanking of wildlife species. Considering the significant amounts of species loss in the recent few centuries, the banked DBS samples could become the last biological material available for some species.

Continuous pathogen surveillance in wildlife is needed considering the high frequency and rate of zoonotic epidemics in the recent few decades (Morse et al. 2012, Rosenberg et al. 2013, Rosenberg 2015). In **chapter 3**, twelve nucleic acid isolates from DNA-RNA combined isolation protocol were used for viral screening by applying hybridization capture with a bait panel targeting a wide range of vertebrate viruses. When broadly screening the presence of viral-candidate reads, a median of ~15.94% sequencing data per sample were aligned to the viral reference database, but majority of the hits were categorized as bacteriophage-like sequences. None of the reads aligned to viral references were confirmed as viral sequences after stringent data filtering. A previous cohort viral screening project in the healthy human population demonstrated that human blood likely carried a very tiny amount of viral genetic materials when no infectious symptoms were observed (Moustafa et al. 2017). Over half of study individuals were detected to be viral-free (Moustafa et al. 2017). For wildlife individuals, the majority are presumably healthy, unless animals presenting specific symptoms are targeted. This suggests most DBS samples will have low titre of any infectious agents. RNA viruses are more likely to degrade during prolonged storage. This may explain the negative viral detection. Overall, our results suggest that DBS samples may not be an ideal sample type for virus screening with high-throughput sequencing, although it shows relatively good preservation performance for PCR-based screening for some pathogens (Kraus et al. 2011, Hall-Mendelin et al. 2017, Rasolonjatovo et al. 2020).

Target capture sequencing provides a cost-effective technique to enrich sequences of interest (Teer and Mullikin 2010, Ali et al. 2016), and also avoids contamination when historical samples are used (Enk et al. 2014). However, off-target sequencing data are commonly generated, which have been reported to additionally provide precious genomic information (Picardi and Pesole 2012, Diroma et al. 2014). In **chapter 3**, I used a bait panel for screening the presence of viruses in twelve wild primates, while none of two our viral taxonomical pipelines identified any virological sequences from twelve capture-sequenced samples. I then further explored the data for other genomic applications. Seven complete mitochondrial genomes of three primate species were successfully *de novo* assembled from the non-target data. In comparison with shotgun sequencing data (~1%), the mitogenome sequences increased in percentage (i.e. mitogenome reads/total reads) to a maximum of ~50%, with a non-uniform coverage distribution across the mitochondrial genome. The coverage depth consistently peaked at similar loci of rRNA and protein-coding genes across seven mitogenome assemblies. Our bait panel was designed by using a subset of viral genomes as references, primarily targeting coding regions. Hybridization capture may be able to tolerate the sequence mismatch when no viral candidate sequences exist to bind with, leading to highly discrepant coverage depth. The results demonstrate that off-target data generated from pathogen-targeted capture sequencing can be used to assemble

mitochondrial genomes, providing a valuable source for genomic and population genetic investigations.

In **chapter 4**, I used illumina sequencing data to *de novo* assemble a mitochondrial genome of lowland paca (*Cuniculus paca*) using degraded DNA isolates from DBS. The phylogenetic analysis based upon confirmed the taxonomic placement of lowland paca in the infraorder Hystricognathi within the hystricognath rodents (Voloch et al. 2013), indicating DBS samples can be considered valuable resources for genomic scale studies. The improvement in DNA sequencing over time may lead to explicit progress that allows for recovering genomic information from the degraded biological samples.

Noninvasive and minimal-invasive environmental and biological samples remain invaluable resources in genetic monitoring and biological applications, such as eDNA and DBS samples. The use of those sample types has been improved tremendously along with recent technological innovation in sequencing techniques (Bybjerg-Grauholm et al. 2017, Bassaganyas et al. 2018, Wu et al. 2019, Agrawal et al. 2021), in some cases, allowing for recovering complete or partial genomes (Racimo et al. 2020). For terrestrial wildlife monitoring, more studies have recently been carried out to test the efficiency of eDNA-based methods compared to that of camera traps both empirically and in the field (Ushio et al. 2017, Williams et al. 2018, Harper et al. 2019, Seeber et al. 2019, Leempoel et al. 2020, Mena et al. 2021). Given that different molecular methods may provide different results and complement one another, our comparison study of two distinct eDNA methods suggest that method integration can significantly improve the efficiency and accuracy of eDNA-based terrestrial wildlife monitoring. To improve monitoring performance and application value of eDNA approaches, sample collections need to be customized by taking characteristics of the local ecosystem and species of interest into consideration.

In animal and wildlife genetic research, minimal-invasive DBS technology so far remains under represented in genomic applications (Samsonova et al. 2022) despite being widely used for stabilizing nucleic acid in medical applications. Our study demonstrated that DNA preserved in DBS from wildlife can stay stable for an extended period of storage time during the fieldwork under sub-standard conditions and in the lab. Although nucleic acid degradation persists, the DNA isolated from DBS can provide valuable genomic-scale information, while the application in viral screening via high-throughput sequencing may be limited.

This thesis shows the significant application potential and limitation of noninvasive and minimal-invasive environmental and biological samples in wildlife biological research and conservation, in

combination with the advanced molecular and sequencing technologies. The results provide important guidance for the application of the abovementioned sampling strategies in wildlife research and conservation. However, in order to put the developed strategies into practical use, a more standardized and simplified workflow and pipeline is necessary. Besides, combining information from all aspects is key for informed management and policy making, such as population dynamics from camera trap records, eDNA detection and the associated genetic diversity estimate, and genetic markers archived for local species individuals. This may help construct a super intertwined bio-archive and dynamic monitoring, empowering evidence-based conservation interventions.

References

Adams, C. I. M., M. Knapp, N. J. Gemmell, G. J. Jeunen, M. Bunce, M. D. Lamare, and H. R. Taylor. 2019. Beyond Biodiversity: Can Environmental DNA (eDNA) Cut It as a Population Genetics Tool? Genes **10**:20.

Agrawal, P., S. Katragadda, A. K. Hariharan, V. G. Raghavendrachar, A. Agarwal, R. Dayalu, D. Awasthy, S. C. Sharma, Y. K. Sivasamy, P. Lakshmana, A. Shanmugam, V. Veeramachaneni, V. Gupta, B. P. Vani, L. Subaiya, T. S. Syamala, R. Hariharan, V. Chandru, and D. E. Bloom. 2021. Validation of whole genome sequencing from dried blood spots. Bmc Medical Genomics **14**.

Alexander, J. B., M. Bunce, N. White, S. P. Wilkinson, A. A. S. Adam, T. Berry, M. Stat, L. Thomas, S. J. Newman, L. Dugal, and Z. T. Richards. 2019. Development of a multi-assay approach for monitoring coral diversity using eDNA metabarcoding. Coral Reefs **39**:159-171.

Ali, O. A., S. M. O'Rourke, S. J. Amish, M. H. Meek, G. Luikart, C. Jeffres, and M. R. Miller. 2016. RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping. Genetics **202**:389-400.

Andres, K. J., D. M. Lodge, S. A. Sethi, and J. Andres. 2023. Detecting and analysing intraspecific genetic variation with eDNA: From population genetics to species abundance. Mol Ecol **32**:4118-4132.

Andrews, K. R., J. M. Good, M. R. Miller, G. Luikart, and P. A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet **17**:81-92.

Bassaganyas, L., G. Freedman, D. Vaka, E. Wan, R. Lao, F. Chen, M. Kvale, R. J. Currier, J. M. Puck, and P. Y. Kwok. 2018. Whole exome and whole genome sequencing with dried blood spot DNA without whole genome amplification. Human Mutation **39**:167-171.

Benham, P. M., and R. C. K. Bowie. 2023. Natural history collections as a resource for conservation genomics: Understanding the past to preserve the future. Journal of Heredity **114**:367-384.

Bowers, H. A., X. Pochon, U. von Ammon, N. Gemmell, J. A. L. Stanton, G. J. Jeunen, C. D. H. Sherman, and A. Zaiko. 2021. Towards the Optimization of eDNA/eRNA Sampling Technologies for Marine Biosecurity Surveillance. Water **13**.

Bybjerg-Grauholm, J., C. M. Hagen, S. K. Khoo, M. L. Johannesen, C. S. Hansen, M. Baekvad-Hansen, M. Christiansen, D. M. Hougaard, and M. V. Hollegaard. 2017. RNA sequencing of archived neonatal dried blood spots. Molecular Genetics and Metabolism Reports **10**:33-37.

da Cunha Santos, G. 2018. FTA Cards for Preservation of Nucleic Acids for Molecular Assays: A Review on the Use of Cytologic/Tissue Samples. Arch Pathol Lab Med **142**:308-312.

Demirev, P. A. 2013. Dried blood spots: analysis and applications. Anal Chem **85**:779-789.

Diroma, M. A., C. Calabrese, D. Simone, M. Santorsola, F. M. Calabrese, G. Gasparre, and M. Attimonelli. 2014. Extraction and annotation of human mitochondrial genomes from 1000 Genomes Whole Exome Sequencing data. Bmc Genomics **15**.

Enk, J. M., A. M. Devault, M. Kuch, Y. E. Murgha, J. M. Rouillard, and H. N. Poinar. 2014. Ancient whole genome enrichment using baits built from modern DNA. Mol Biol Evol **31**:1292-1294.

Fitzpatrick, S. W., G. S. Bradburd, C. T. Kremer, P. E. Salerno, L. M. Angeloni, and W. C. Funk. 2020. Genomic and Fitness Consequences of Genetic Rescue in Wild Populations. Curr Biol **30**:517-522 e515.

Fowler, K. E., C. P. Reitter, G. A. Walling, and D. K. Griffin. 2012. Novel approach for deriving genome wide SNP analysis data from archived blood spots. BMC Res Notes **5**:503.

Groenenberg, M. C., R. Yoganand, K. 2020. Population status of ungulates in the Eastern Plains landscape. Srepok Wildlife Sanctuary and Phnom Prich Wildlife Sanctuary, Cambodia. Phnom Penh, Cambodia: WWF-Cambodia.

Hall-Mendelin, S., G. R. Hewitson, D. Genge, P. J. Burtonclay, A. J. De Jong, A. T. Pyke, and A. F. van den Hurk. 2017. FTA Cards Facilitate Storage, Shipment, and Detection of Arboviruses in Infected Aedes aegypti Collected in Adult Mosquito Traps. Am J Trop Med Hyg **96**:1241-1243.

Harper, L. R., L. L. Handley, A. I. Carpenter, M. Ghazali, C. Di Muri, C. J. Macgregor, T. W. Logan, A. Law, T. Breithaupt, D. S. Read, A. D. McDevitt, and B. Hanfling. 2019. Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. Biological Conservation **238**:108225.

Hashimoto, M., M. Bando, J. Kido, K. Yokota, T. Mita, K. Kajimoto, and M. Kataoka. 2019. Nucleic acid purification from dried blood spot on FTA Elute Card provides template for polymerase chain reaction for highly sensitive Plasmodium detection. Parasitology International **73**.

Hohenlohe, P. A., W. C. Funk, and O. P. Rajora. 2021. Population genomics for wildlife conservation and management. Molecular Ecology **30**:62-82.

Jensen, M. R., E. E. Sigsgaard, S. Liu, A. Manica, S. S. Bach, M. M. Hansen, P. R. Moller, and P. F. Thomsen. 2021. Genome-scale target capture of mitochondrial and nuclear environmental DNA from water samples. Mol Ecol Resour **21**:690-702.

Kardos, M., E. E. Armstrong, S. W. Fitzpatrick, S. Hauser, P. W. Hedrick, J. M. Miller, D. A. Tallmon, and W. C. Funk. 2021. The crucial role of genome-wide genetic variation in conservation. Proc Natl Acad Sci U S A **118**.

Kraus, R. H., P. van Hooft, J. Waldenström, N. Latorre-Margalef, R. C. Ydenberg, and H. H. Prins. 2011. Avian influenza surveillance with FTA cards: field methods, biosafety, and transportation issues solved. J Vis Exp.

Leempoel, K., T. Hebert, and E. A. Hadly. 2020. A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. Proc Biol Sci **287**:20192353.

Lim, M. D. 2018. Dried Blood Spots for Global Health Diagnostics and Surveillance: Opportunities and Challenges. Am J Trop Med Hyg **99**:256-265.

Longhi, S. A., L. J. G. Casares, A. A. Munoz-Calderon, J. Alonso-Padilla, and A. G. Schijman. 2023. Combination of ultra-rapid DNA purification (PURE) and loop-mediated isothermal amplification (LAMP) for rapid detection of Trypanosoma cruzi DNA in dried blood spots. Plos Neglected Tropical Diseases **17**.

McClendon-Weary, B., D. L. Putnick, S. Robinson, and E. Yeung. 2020. Little to Give, Much to Gain-What Can You Do With a Dried Blood Spot? Curr Environ Health Rep **7**:211-221.

Mena, J. L., H. Yagui, V. Tejeda, E. Bonifaz, E. Bellemain, A. Valentini, M. W. Tobler, P. Sanchez-Vendizu, and A. Lyet. 2021. Environmental DNA metabarcoding as a useful tool for evaluating terrestrial mammal diversity in tropical forests. Ecol Appl **31**:e02335.

Morse, S. S., J. A. K. Mazet, M. Woolhouse, C. R. Parrish, D. Carroll, W. B. Karesh, C. Zambrana-Torrelio, W. I. Lipkin, and P. Daszak. 2012. Zoonoses 3 Prediction and prevention of the next pandemic zoonosis. Lancet **380**:1956-1965.

Moustafa, A., C. Xie, E. Kirkness, W. Biggs, E. Wong, Y. Turpaz, K. Bloom, E. Delwart, K. E. Nelson, J. C. Venter, and A. Telenti. 2017. The blood DNA virome in 8,000 humans. Plos Pathogens **13**.

Picardi, E., and G. Pesole. 2012. Mitochondrial genomes gleaned from human whole-exome sequencing. Nat Methods **9**:523-524.

Pin, C., D. Ngoprasert, T. N. E. Gray, T. Savini, R. Crouthers, and G. A. Gale. 2020. Utilization of waterholes by globally threatened species in deciduous dipterocarp forest of the Eastern Plains Landscape of Cambodia. Oryx **54**:572-582.

Racimo, F., M. Sikora, M. Vander Linden, H. Schroeder, and C. Lalueza-Fox. 2020. Beyond broad strokes: sociocultural insights from the study of ancient genomes. Nat Rev Genet **21**:355-366.

Rasolonjatovo, F. S., H. Guis, M. Rajeev, L. Dacheux, L. Arivony Nomenjanahary, G. Razafitrimo, J. T. Rafisandrantantsoa, C. Cêtre-Sossah, J. M. Heraud, and S. F. Andriamandimby. 2020. Enabling animal rabies diagnostic in low-access areas: Sensitivity and specificity of a molecular diagnostic test from cerebral tissue dried on filter paper. PLoS Negl Trop Dis **14**:e0008116.

Rees, H. C., B. C. Maddison, D. J. Middleditch, J. R. M. Patmore, and K. C. Gough. 2014. REVIEW The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. Journal of Applied Ecology **51**:1450-1459.

Rosenberg, R. 2015. Detecting the emergence of novel, zoonotic viruses pathogenic to humans. Cell Mol Life Sci **72**:1115-1125.

Rosenberg, R., M. A. Johansson, A. M. Powers, and B. R. Miller. 2013. Search strategy has influenced the discovery rate of human viruses. Proc Natl Acad Sci U S A **110**:13961-13964.

Russcher, A., A. Enders, C. S. de Brouwer, D. Oepkes, R. Hahn, M. Enders, A. C. M. Kroes, and A. Vossen. 2020. Diagnosis of intrauterine parvovirus B19 infection at birth - Value of DNA detection in neonatal blood and dried blood spots. Journal of Clinical Virology **129**.

Samsonova, J. V., N. Y. Saushkin, and A. P. Osipov. 2022. Dried Blood Spots technology for veterinary applications and biological investigations: technical aspects, retrospective analysis, ongoing status and future perspectives. Vet Res Commun **46**:655-698.

Seeber, P. A., and L. S. Epp. 2022. Environmental DNA and metagenomics of terrestrial mammals as keystone taxa of recent and past ecosystems. Mammal Review **52**:538-553.

Seeber, P. A., G. K. McEwen, U. Lober, D. W. Forster, M. L. East, J. Melzheimer, and A. D. Greenwood. 2019. Terrestrial mammal surveillance using hybridization capture of environmental DNA from African waterholes. Mol Ecol Resour **19**:1486-1496.

Sigsgaard, E. E., I. B. Nielsen, S. S. Bach, E. D. Lorenzen, D. P. Robinson, S. W. Knudsen, M. W. Pedersen, M. Jaidah, L. Orlando, E. Willerslev, P. R. Moller, and P. F. Thomsen. 2017. Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. Nat Ecol Evol **1**.

Staunstrup, N. H., A. Starnawska, M. Nyegaard, L. Christiansen, A. L. Nielsen, A. Borglum, and O. Mors. 2016. Genome-wide DNA methylation profiling with MeDIP-seq using archived dried blood spots. Clinical Epigenetics **8**.

Teer, J. K., and J. C. Mullikin. 2010. Exome sequencing: the sweet spot before whole genomes. Hum Mol Genet **19**:R145-151.

Tuaillon, E., D. Kania, A. Pisoni, K. Bollore, F. Taieb, E. N. O. Ngoyi, R. Schaub, J. C. Plantier, A. Makinson, and P. Van de Perre. 2020. Dried Blood Spot Tests for the Diagnosis and Therapeutic Monitoring of HIV and Viral Hepatitis B and C. Frontiers in Microbiology **11**.

Ushio, M., H. Fukuda, T. Inoue, K. Makoto, O. Kishida, K. Sato, K. Murata, M. Nikaido, T. Sado, Y. Sato, M. Takeshita, W. Iwasaki, H. Yamanaka, M. Kondoh, and M. Miya. 2017. Environmental DNA enables detection of terrestrial mammals from forest pond water. Mol Ecol Resour **17**:e63-e75.

Voloch, C. M., J. F. Vilela, L. Loss-Oliveira, and C. G. Schrago. 2013. Phylogeny and chronology of the major lineages of New World hystricognath rodents: insights on the biogeography of the Eocene/Oligocene arrival of mammals in South America. BMC Res Notes **6**:160.

Williams, K. E., K. P. Huyvaert, K. C. Vercauteren, A. J. Davis, and A. J. Piaggio. 2018. Detection and persistence of environmental DNA from an invasive, terrestrial mammal. Ecol Evol **8**:688-695.

Wu, S. G., W. J. Pan, H. N. Liu, M. L. Byrne-Steele, B. Brown, M. Depinet, X. H. Hou, J. Han, and S. Li. 2019. High throughput sequencing of T-cell receptor repertoire using dry blood spots. Journal of Translational Medicine **17**.

Acknowledgements

Herein, I would like to express my great appreciation to everyone who helped and supported me to accomplish the journey of my doctoral study in Berlin in the last four years.

First and foremost, I would like to thank my supervisor who gave me a graduate position in the department of Wildlife Diseases at the IZW, where I embark on my doctoral journey in Berlin four years ago and found my passion towards wildlife research. Highly appreciate for patiently giving guidance and encouragement all the way along, as well as quick feedbacks when I needed.

I would like to show my special appreciation to Dorina Meneghini who kindly offered to be my mentor in bioinformatics learning process and meanwhile my dear friend in Berlin. Thank her for giving all the helps whenever I confronted technical challenges and always trying her best to solve my issues. A big thank to Alexandre Courtiol, who helped and supported me to improve my ability to perform good scientific statistical analysis. My sincere gratitude to Gábor Czirják, for his mental support in my challenged times and for his great work as a PhD counsellor.

I am very grateful to the entire team of the Department of Wildlife Diseases for creating a warm and friendly working atmosphere. Many thanks to all the great colleagues whom I met throughout my study time at the IZW and who added an extra credit to this experience. Thank Karin Hönig for maintaining and organizing the lab settings and offering all the technical supports in this regard. In this respective, a big thank to Katja Pohle, Anke Schmidt, and Jan Axtner for offering lab supports and willing to sharing lab equipment. A big thank to Peter Seeber for helping me from the very beginning to the end in Cambodian eDNA project.

My sincere gratitude to all dear friends I met at the IZW, who have been my precious companions and given me great encouragement to walk through this tough journey: my godparents Saba Mottaghinia and Miguel Veiga, John, Lilla Jordan, Seth Wong, Rohit Chakravarty, Morgane Gicquel.

Lastly, many thanks to my family and my boyfriend, right now my husband for bearing my bad temper when I felt lonely and got depressed, and for giving me mental support.

Finally, I am very grateful to the CSC: Chinese Scholarship Council for awarding me with a competitive scholarship for full Doctoral studies in Germany.

List of publications and author contributions

First author:

Li, J., Seeber, P., Axtner, J., Crouthers, R., Groenenberg, M., Koehncke, A., Courtiol, A., Chanratana, P., and Greenwood, A.D. (2023). Monitoring terrestrial wildlife by combining hybridization capture and metabarcoding data from waterhole environmental DNA. Biological Conservation 284. https://doi.org/10.1016/j.biocon.2023.110168

Li, J., Jiang, H., Li, L., Zhang, X., and Chen, J. (2019). The Effect of Disease and Season to Hepatopancreas and Intestinal Mycobiota of Litopenaeus vannamei. Front Microbiol 10, 889. https://doi.org/10.3389/fmicb.2019.00889

Li, J., Li, L., Jiang, H., Yuan, L., Zhang, L., Ma, J.E., Zhang, X., Cheng, M., and Chen, J. (2018). Fecal Bacteriome and Mycobiome in Bats with Diverse Diets in South China. Curr Microbiol 75, 1352-1361. https://doi.org/10.1007/s00284-018-1530-0

Li, J., Mayor, P., Robles, M. L.S., Greenwood, A.D. The complete mitochondrial genome of the lowland paca (*Cuniculus paca*) and its phylogenetic relationship with other New World hystricognath rodents. (2023). Mitochondrial DNA Part B: Resources. Accepted manuscript.

Li, J., Urizar, G.U., Mayor, P., Robles, M. L.S., Greenwood, A.D. Nucleic acid degradation after long term dried blood spot storage. Submitted manuscript.

Coauthor:

Ma, J.E., Li, L.M., Jiang, H.Y., Zhang, X.J., **Li, J**., Li, G.Y., and Chen, J.P. (2018). Acidic mammalian chitinase gene is highly expressed in the special oxyntic glands of Manis javanica. FEBS Open Bio 8, 1247-1255. https://doi.org/10.1002/2211-5463.12461

Zhang, X., Li, L., Jiang, H., Ma, J.E., **Li, J**., and Chen, J. (2018). Identification and differential expression of microRNAs in testis and ovary of Amur sturgeon (Acipenser schrenckii). Gene 658, 36-46. https://doi.org/10.1016/j.gene.2018.03.014

Jiang, H., **Li, J**., Li, L., Zhang, X., Yuan, L., and Chen, J. (2017). Selective evolution of Toll-like receptors 3, 7, 8, and 9 in bats. Immunogenetics 69, 271-285. https://doi.org/10.1007/s00251-016-0966-2.

Jiang, H.Y., Ma, J.E., **Li, J**., Zhang, X.J., Li, L.M., He, N., Liu, H.Y., Luo, S.Y., Wu, Z.J., Han, R.C., and Chen, J.P. (2017). Diets Alter the Gut Microbiome of Crocodile Lizards. Front Microbiol 8, 2073. https://doi.org/10.3389/fmicb.2017.02073

Ma, J.E., Li, L.M., Jiang, H.Y., Zhang, X.J., **Li, J**., Li, G.Y., Yuan, L.H., Wu, J., and Chen, J.P. (2017). Transcriptomic analysis identifies genes and pathways related to myrmecophagy in the Malayan pangolin (Manis javanica). PeerJ 5, https://doi.org/10.7717/peerj.4140

Curriculum Vitae

For reasons of data protection, the curriculum vitae is not published in the electronic version.

For reasons of data protection, the curriculum vitae is not published in the electronic version.