Characterization of the herbivore and carnivore faecal microbiota with special emphasis on diet and host-specificity

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt dem Fachbereich Biowissenschaften

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Franziska Zölzer

aus Dorsten

Frankfurt 2024

(D30)

Vom Fachbereich der Biowissenschaften der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Sven Klimpel

Gutachter: Prof. Dr. Paul W. Dierkes

Zweitgutachterin: Prof. Dr. Lisa Schulte

Datum der Disputation:

Table of Contents

I. Zusammenfassung

In den letzten zwei Dekaden hat die Erforschung des menschlichen und tierischen Mikrobioms erheblich an Bedeutung gewonnen. Mittlerweile ist es unumstritten, dass Darmbakterien unverzichtbar für die Gesundheit sind, da sie an einer Vielzahl von physiologischen Prozessen im Organismus beteiligt sind welche sich nicht nur auf das Verdauungssystem beschränken. So übernehmen gastrointestinale Mikroorganismen neben der Verdauung auch Funktionen im Immunsystem oder beeinflussen Aspekte des menschlichen und tierischen Verhaltens.

Der Begriff Mikrobiom beschreibt die Gesamtheit aller Mikroorganismen sowie deren Gene und Stoffwechselprodukte in einer abgeschlossenen Lebensgemeinschaft, während das Mikrobiota sich auf die Zusammensetzung der Mikroorganismen an sich bezieht (Lane-Petter, 1962; Lederberg and McCray, 2001). Im letzten Jahrzehnt hat sich die Forschung auf viele dieser Lebensgemeinschaften konzentriert, beispielsweise im Boden (Fierer, 2017; Islam et al., 2020), im Wasser (Sunagawa et al., 2015; Hull et al., 2019), in der Luft (Leung et al., 2014; Drautz-Moses et al., 2022) oder auf und im menschlichen und tierischen Organismus. Letztere bezieht sich zum Großteil auf Mikroorganismen im Verdauungssystem, zum einen da diese eine Vielzahl an physiologischen Aufgaben übernehmen, zum anderen da diese über fäkale Proben nicht invasiv detektiert werden können.

Innerhalb der Säugetiere sind Mikroorganismen insbesondere für den Abbau von pflanzlichen Zellbestandteilen im Darm verantwortlich. Diese bestehen hauptsächlich aus Zellulose, welche in der Nahrung von herbivoren Tierarten bis zu 28% der Futter-Trockenmasse ausmachen kann und welche vom Tier aufgrund der fehlenden Zellulase nicht selbstständig abgebaut werden kann (Choct, 2015). Für den Abbau der Zellulose sind zellolytische Mikroorganismen, beispielsweise Bakterien aus den Familien der Prevotellaceae, Fibrobacteraceae oder Spirochaetaceae nötig (Yatsunenko et al., 2012; Kartzinel et al., 2019; Quercia et al., 2019). Diese spalten die β-1,4-glykosidischen Bindungen zwischen den D-Glucose Molekülen und nutzen die entstandene Glucose für den mikrobiellen Stoffwechsel. Als Nebenprodukt werden kurzkettige Fettsäuren wie Acetat oder Butyrat produziert, welche wiederum dem Host-Organismus als Energiequelle zur Verfügung stehen (Patterson et al., 2014; Froidurot and Julliand, 2022).

Neben ihrer Funktion als Energiequelle für den Host-Organismus können kurzkettige Fettsäuren auch mit dem Immunsystem interagieren. Beispielsweise kann eine verringerte Butyrat- oder Acetat-Konzentration die Zytokinproduktion von T-Helferzellen verändern (Bird et al., 1998) oder auch die Integrität des Darmepithels verringern (Peng et al., 2007; Maslowski et al., 2009). Dies zeigt, wie eng das gastrointestinale Mikrobiota auch mit der Regulierung des

Immunsystems verknüpft ist. Zusätzlich kann das Mikrobiom auch verschiedene Verhaltensweisen direkt oder indirekt beeinflussen. Die chemische Kommunikation beruht bei verschiedenen Tierarten auf Metaboliten welche direkt vom Mikrobiom produziert werden (Gorman, 1976). Diese Metabolite enthalten unterschiedliche Informationen über den Host-Organismus wie beispielsweise Alter, Geschlecht oder Fertilität und sind somit verantwortlich für das Sozialverhalten der jeweiligen Tierarten (Sin et al., 2012; Theis et al., 2013; Leclaire et al., 2014). Neben der direkten Rolle in der olfaktorischen Kommunikation zwischen Individuen beeinflusst das Mikrobiom auch indirekt das Verhalten einzelner Individuen, beispielsweise in Stresssituationen. Veränderungen im Mikrobiom wie eine verringerte Diversität oder eine Dysbiose im Verhältnis zwischen Firmicutes und Bacteroidetes können die Produktion der kurzkettigen Fettsäuren verändern. Der Host Organismus reagiert auf diese Veränderungen häufig mit Stress- und Angstzuständen und beeinflusst so das Wohlbefinden des Individuums (Jiang et al., 2018; Yuan et al., 2021; Malan-Muller et al., 2022).

Die genannten Aspekte, welche über das Darmmikrobiom beeinflusst werden können, sind auch wichtige Indikatoren im Tier- und Artenschutz. Dieser ist heutzutage aktueller als jemals zuvor. Während die International Union for Conservation of Nature (IUCN) mehr als 42.100 Tierarten als vom Aussterben bedroht einstuft (IUCN, 2022), hat die Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) in einem zusammenfassenden Report aus über 15.000 Studien deutliche Hinweise für ein fortschreitendes weltweites Massenaussterben vorgelegt (Díaz et al., 2019). Diese Entwicklung zeigt auch die Dringlichkeit von Erhaltungszuchtprogrammen um den Bestand gefährdeter Arten weitestgehend sicher zu stellen. Die Ausführung dieser liegt bei zoologischen Einrichtungen, deren Hauptaufgabe unter anderem der Schutz der Artenvielfalt und eine stetige Kontrolle des Tierwohls sind. Die Mikrobiomforschung bietet hier eine zugängliche Methode um den Tier- und Artenschutz nicht invasiv zu unterstützen. Es ist bekannt, dass ausgehend von einer fäkalen Probe mikrobielle Marker detektiert werden können, welche zum Beispiel Aufschlüsse über den Gesundheitszustand und Reproduktionszyklus der Tiere geben (Weingrill et al., 2004; Antwis et al., 2019).

Hier liegt auch der Ansatzpunkt dieser Thesis. Um diese eventuell artspezifischen Marker zu identifizieren und zu charakterisieren, bieten zoologische Einrichtungen eine ideale Forschungsumgebung. Zum einen ist die Artenvielfalt hier sehr hoch, sodass auch sehr seltene und stark gefährdete Arten in die Studien mit einbezogen werden können. Zum anderen können Arten und Individuen aus verschiedenen Einrichtungen miteinander, oder zu verschiedenen Zeitpunkten untereinander verglichen werden. Außerdem sind alle relevanten Metadaten wie

Alter, Geschlecht, Verwandtschaft oder medizinische Behandlungen dokumentiert und können zur Auswertung herangezogen werden. Die Ziele dieser Thesis sind demnach zunächst die Charakterisierung des Mikrobioms verschiedenster Säugetierarten um eine Datenbank als Grundlage für die weitere Forschung zu erstellen. Zusätzlich sollen hier auch tägliche mikrobielle Schwankungen mit einbezogen werden, um eine genauere Beschreibung des individuellen Mikrobioms zu gewährleisten und Rückschlüsse auf das Sammelschema der fäkalen Proben zu ziehen. Da die Mikrobiomforschung in Zukunft auch im Freiland durchgeführt werden soll, um auch hier einen Beitrag zum Artenschutz zu liefern, soll abschließend eine Methode entwickelt werden um Host-spezifische Metadaten aus unbekannten fäkalen Proben zu generieren.

Die vorliegende Arbeit in drei Teilstudien gegliedert, die dazu beitragen, das fäkale Mikrobiom verschiedener Arten zu verstehen und Host-spezifische Informationen aus diesem zu erlangen. In der ersten Studie mit dem Titel „Unravelling differences in faecal microbiota stability in mammals: from highly variable carnivores and consistently stable herbivores“ geht es um die Charakterisierung des fäkalen Mikrobioms von 31 Säugetierarten. Das Hauptaugenmerk liegt hier auf einem aussagekräftigen Datensatz von mindestens fünf Proben pro Tierart aus mindestens drei verschiedenen Zoos. In erster Linie sind hier die Unterschiede im fäkalen Mikrobiom zwischen karnivoren (Canoidea und Feloidea) und herbivoren (Ruminantia und Perissodactyla) Arten anhand von 621 Proben aufgezeigt worden. Diese erste Studie zeigt signifikante Unterschiede im Mikrobiom von karnivoren und herbivoren Tierarten - sowohl in der mikrobiellen Zusammensetzung als auch in der Diversität. Das jeweilige Mikrobiota ist in seiner Zusammensetzung stark an die Ernährung der jeweiligen Art angepasst. So ist das karnivore Mikrobiota geprägt von Protein-abbauenden Bakterienfamilien wie etwa Fusobacteriaceae und Clostridiaceae. Gleichzeitig ist die mikrobielle Diversität hier sehr gering, da das einfache Verdauungssystem dieser Tiere keine spezifischen Anpassungen an die bakterielle Fermentation bietet. Im Gegensatz hierzu finden sich im herbivoren Mikrobiom zellolytische Bakterienfamilien wie z.B. Spirochaetaceae, Oscillospiraceae und Lachnospiraceae. Im Vergleich zu den karnivoren Vertretern zeigen herbivore Arten eine signifikant höhere mikrobielle Diversität, welche die Anpassung an die morphologisch komplexeren Verdauungssysteme der Vorder- und Enddarmfermentierer verdeutlicht. Ein weiteres Ergebnis dieser Publikation ist, dass das Mikrobiota herbivorer Arten deutlich homogener ist als das der karnivoren Arten. Sowohl die Proben der Ruminantia als auch die der Perissodactyla bilden deutlich abgegrenzte Cluster. Im Gegensatz hierzu, zeigen die Proben der

Carnivora eine deutlich höhere Streuung und es ist keine klare Abgrenzung zwischen den Feloidea und Canoidea zu erkennen. Diese hohe Variabilität innerhalb der Carnivora wird neben der Hauptkoordinatenanalyse auch in den hohen Variationskoeffizienten einzelner Mikroorganismen deutlich. Hierdurch ergeben sich Implikationen für das Sammelschema weiterer Studien. Um die gesamte Variabilität im Mikrobiom einer Tierart abzudecken, sind einzelne Proben wenig aussagekräftig, insbesondere bei der Analyse von karnivoren Tierarten. Aufbauend auf diesen Erkenntnissen beschäftigt sich die zweite Teilstudie „Time series cluster analysis reveals individual assignment of microbiota in captive tiger (*Panthera tigris*) and wildebeest (*Connochaetes taurinus*)" mit den täglichen Oszillationen im fäkalen Mikrobiota. Das Ziel dieser Studie ist es ein Clusterverfahren zu entwickeln mit welchem zwischen verschiedenen Individuen einer Art differenziert werden kann. Für diesen Ansatz sind über acht Tage hinweg Proben von verschiedenen *Panthera tigris* und *Connochaetes taurinus* Individuen gesammelt worden. Um die Fragestellung zu beantworten, sind zwei verschiedene methodische Ansätze entwickelt worden. Zum einen sind im einfachen Clustering Prozess alle Proben einer individuellen Zeitreihe separat in die beiden Cluster-Algorithmen - Ward's Linkage (Ward, 1963) und Community Detection (Newman, 2004) – eingefügt worden. Zum anderen sind Proben einer Zeitreihe zunächst mittels dynamischer Zeitnormierung kombiniert und erst im Anschluss über die Cluster-Algorithmen analysiert worden. Die Ergebnisse zeigen, dass sowohl die Art- als auch die Individuen Erkennung deutlich zuverlässiger funktioniert, wenn zuvor eine Zeitnormierung angewendet wird. Für die Arterkennung ist Ward's Linkage besser geeignet da hier beide Arten ein eigenes Cluster bilden. Um zwischen einzelnen Individuen einer Art zu differenzieren, eignet sich hingegen der Community Detection Algorithmus besser. Durch diesen können Individuen von *Panthera tigris* etwas besser identifiziert werden als *Connochaetes taurinus* Individuen. Zusätzlich können in dieser Studie die Bakterienfamilien bestimmt werden, die für die einzelnen Cluster verantwortlich sind. Für die Arterkennung sind dies hauptsächlich Taxa, welche für die Verdauung von pflanzlicher (Spirochaetaceae, Methanobacteriaceae) oder fleischlicher (Fusobacteriaceae, Clostridiaceae) Nahrung verantwortlich sind. Im Falle der individuellen Zuordnung handelt es sich um eine Kombination aus täglich auftretenden sowie individuell spezifischen mikrobiellen Taxa. Beispielsweise können *Panthera tigris* Individuen anhand der hoch abundanten Prevotellaceae und Clostridiaceae identifiziert werden. Im Gegensatz hierzu ist eine individuelle Differenzierung bei *Connochaetes taurinus* nur anhand von sehr gering abundanten Taxa wie beispielsweise Acidaminococcaceae und Muribaculaceae erfolgreich.

In der dritten Teilstudie ist, basierend auf den vorherigen Erkenntnissen, eine Methodik entwickelt worden, um den Artenschutz in Zukunft auch im Freiland durch Mikrobiomforschung unterstützen zu können. Insbesondere die Identifikation des Hosts über dessen fäkales Mikrobiota steht bei dieser Arbeit im Vordergrund. Zunächst ist über eine Korrelationsanalyse überprüft worden, welche Taxa mit der Nahrung (herbivor, karnivor, omnivor), dem Verdauungssystem (simpel, Wiederkäuer, Enddarmfermentierer) und/ oder der Host-Familie korrelieren. In Bezug auf den Nahrungstyp wird das Ergebnis aus den vorangegangenen Teilstudien bestätigt, welches besagt, dass das karnivore Mikrobiota positiv mit Protein-abbauenden und das herbivore Mikrobiota mit zellolytischen Taxa korreliert. In Anbetracht der verschiedenen Verdauungssysteme korrelieren beispielsweise Ruminococcaceae (R=0,77) mit den Wiederkäuern, Spirochaetaceae (R=0,86) mit den Enddarmfermentierern und Clostridiaceae (R=0,56) mit dem einfachen Verdauungssystem. Auf Grundlage dieser Ergebnisse ist ein logistisches Regressionsmodel entwickelt worden, welches in der Lage ist, verschiedene Host-spezifischen Eigenschaften anhand der mikrobiellen Zusammensetzung einer fäkalen Probe vorherzusagen. Im ersten Schritt unterscheidet das Modell mit einer Genauigkeit von bis zu 99% zwischen karnivorem, herbivorem oder omnivorem Host. Handelt es sich um einen karnivoren Host, kann im Folgenden mit einer Genauigkeit von bis zu 96% die Host Familie (Canidae, Felidae, Ursidae, Herpestidae) vorhergesagt. Bei herbivoren Hosts erreicht das Modell für das jeweilige Verdauungssystem (Wiederkäuer, Enddarmfermentierer, einfach) sogar bis zu 100% Genauigkeit. Solche Modellierungsansätze sind auf dem Gebiet der Mikrobiomforschung sehr vielversprechend, da sie neue Möglichkeiten bieten, die Analyse mikrobieller Daten mit Host-Metadaten zu kombinieren. Insbesondere ist dies für die Freilandforschung interessant, da die Gewinnung von Host-spezifischen Informationen aus fäkalen Proben sehr aufwändig ist. Gegenwärtig erfolgt dies in der Regel durch Mikrosatellitenanalysen, die sehr kostenintensiv sind. Eine weitere Einschränkung ist die hierfür notwendige, qualitativ hochwertige Host-DNA, welche in fäkalen Proben zum Großteil fragmentiert vorliegt. Der in dieser Teilstudie entwickelte Modellierungsansatz zur Gewinnung möglichst vieler Host-spezifischer Informationen aus der Zusammensetzung des fäkalen Mikrobioms ist kosten- und arbeitseffizienter als herkömmliche Ansätze und eröffnet somit ein neues und zugänglicheres Feld für die Mikrobiomforschung im Freiland.

II. Summary

Research on the human and animal microbiome has become increasingly important in recent years. It is now widely accepted the gut microbiome is of crucial importance to health, as it is involved in a large number of physiological processes. The term ‘microbiome’ refers to the all living microorganisms including their genes and metabolites in a defined environment, while the specific composition of microorganisms consisting of bacteria, archaea and protozoa is referred to as the ‘microbiota’ (Lane-Petter, 1962; Lederberg and McCray, 2001).

In recent years, research has focused on various of these communities in the soil (Fierer, 2017), water (Sunagawa et al., 2015), air (Leung et al., 2014) and especially in the human gut. However, this topic is also becoming increasingly relevant for the conservation of endangered species. In the face of global mass extinctions and the listing of over 42,000 animal species as ‘critically endangered’, conservation breeding programmes are more important than ever (Díaz et al., 2019; IUCN, 2022). The responsibility for these tasks lies with zoological institutions, which are dedicated to animal conservation and the continuous monitoring of animal welfare. Microbiome research offers a non-invasive method to support species conservation. By analysing faecal samples, microbial markers can be identified that provide important information about the health status and reproductive cycle of animals (Weingrill et al., 2004; Antwis et al., 2019). Zoological facilities also provide an ideal research environment for comparing individuals from different habitats. In addition, all necessary metadata such as age, sex, kinship or medical treatment are documented and can be used for the analysis.

This is the starting point for this thesis. In order to identify such microbial markers, it is necessary to understand the microbiome of a variety of animal species. The first aim is therefore to characterise the faecal microbiota of 31 mammalian species, focusing on herbivores and carnivores. It could be shown that they differ significantly in terms of both microbial diversity and microbiota composition. Herbivorous species express a very diverse microbial composition, consisting mainly of cellulose-degrading taxa of the families Fibrobacteraceae or Spirochaetaceae. In contrast, the microbiota of carnivorous species is less diverse and is dominated by protein-degrading Fusobacteriaceae and Clostridiaceae. In addition, this thesis proves that the microbiota of herbivorous species is highly consistent, whereas the microbiota of carnivorous species is highly variable. The results of this study provide important insights for the sampling scheme of future projects. Especially when analysing carnivorous species, single samples are not sufficient to capture the full variability of the microbiome.

These results lead to the question of whether this variability can be explained by daily fluctuations in the individual microbiome and whether this can be used to distinguish between species or individuals. Using individual longitudinal data and a combined approach of clustering algorithms and dynamic time warping, it is shown that such a distinction is possible at the species and individual level. This was confirmed for both a carnivorous (*Panthera tigris*) and a herbivorous (*Connochaetes taurinus*) species. These results confirm the influence of the host individual on the faecal microbiota, in addition to the often described influence of diet (Ley et al., 2008a; Kartzinel et al., 2019).

Based on the knowledge gained from these studies, a methodology has been developed that will enable the conservation of species in the field to be supported by microbiome research in the future. The focus here lays on the identification of host-specific metadata based on the faecal microbiota. The developed regression model is able to distinguish between carnivorous, herbivorous and omnivorous hosts with up to 99% accuracy. In addition, a more accurate phylogenetic classification of the family (Canidae, Felidae, Ursidae, Herpestidae) can be made for carnivorous hosts. For herbivorous hosts, the model can predict the respective digestive system with up to 100% accuracy, distinguishing between ruminants, hindgut fermenters and a simple digestive system. The acquisition of host-specific metadata from an unknown faecal sample is an important step towards establishing microbiome research in species conservation. Field studies in particular will benefit from such new methods. Usually, costly microsatellite analysis and high-quality host DNA are required to obtain host-specific information from faecal samples. The newly developed method offers a less costly and labour-intensive alternative to conventional techniques and opens up a more accessible field for microbiome research in the field.

III. Research Project

This section of the thesis provides a comprehensive overview of the research project. Firstly, the specific objectives and aims of each study are presented, together with the underlying theoretical framework. This is followed by a detailed description of the methodologies employed in the research. Finally, the results of the study are outlined and critically examined.

1. Theoretical background

The term ‘microbiome’ was popularised by the Nobel Laureate Dr Joshua Lederberg in 2001, who defined the microbiome as a community of microorganisms in a particular habitat including their genomes and all environmental conditions (Lederberg and McCray, 2001). Although the term was used in a similar context in 1988 (Whipps et al.), Lederberg’s definition is the one that has underpinned the development of microbiome research since 2001. Simultaneously the term ‘microbiota’, common in microbiology since the 1960s (Lane-Petter, 1962), has come to the fore again. This term refers to the organisms living within a microbiome, including bacteria, archaea, viruses, fungi and protists.

It was the development of next-generation sequencing (NGS) methods during this period that led to the breakthrough in microbiome research in the 2000s. Until then, Sanger’s chain termination method (Sanger and Coulson, 1975) was the only technique for DNA sequencing. Although automated methods were available to replace the gel electrophoresis step, such as the ABI Prism (Applied Biosystems®, Massachusetts, USA), the costs of sequencing a genome was extremely high. When the first human draft genome was sequenced in 2001 using this method, the total cost of the research project was around $3 billion (Venter et al., 2001). With the development of second-generation sequencing methods in 2005 and 2006, the costs of sequencing fell dramatically, outpacing Moore's law in early 2008, reducing the costs of a human genome to around $50,000 in 2010 (Wetterstrand, 2023). One of the sequencing companies speeding up DNA sequencing was Illumina® Incorporation. The main advantages over first-generation sequencing is the ability to sequence millions of DNA fragments (reads) simultaneously. This method is based on bridge amplification, followed by sequencing by synthesis, where fluorescence-labelled nucleotides are directly detected as they are incorporated into the DNA (Figure 3). Especially for short DNA fragments, e.g. for 16S amplicons, this method has become the most widely used and has been improved since then (Fadrosh et al., 2014; Sinclair et al., 2015; Pichler et al., 2018). Today, the cost of sequencing a genome is approaching $1,000, which has been made possible largely by the third-generation sequencing

methods (Wetterstrand, 2023). Since 2011, Pacific Biosciences Inc. and, since 2015, Oxford Nanopore Technologies plc have developed new sequencing methods that enable single-molecule real-time sequencing (SMRT), which means that long DNA fragments up to complete genomes can be sequenced in one run without a PCR step. With the MinIon, Nanopore has developed a sequencer that requires very little work in the laboratory and, thanks to its portability, can also be used for field applications.

1.1. Relevance of microbiome research

In the context of the Sustainable Development Goals (SDG) set by the United Nations General Assembly in 2015, microbiome research can contribute to many aspects of human and ecosystem health (D'Hondt et al., 2021). Firstly, the link between the gut microbiome and the human or animal health is well described. In particular, a balanced diet rich in complex carbohydrates and unsaturated fatty acids, commonly referred to as the Mediterranean diet, has a positive impact on the gut microbiome. Consumption of whole plant materials promotes the growth of fibre-digesting bacteria (Grundy et al., 2016). The most important products of this microbial fermentation are short-chain fatty acids (SCFAs), such as acetate, propionate and butyrate, which serve as an energy source for colonocytes (Cummings, 1981; Rérat et al., 1987; Sunvold et al., 1995). In contrast, the production of SCFAs is significantly reduced following a high-fat westernised diet (David et al., 2014; Ojo et al., 2016). This may also negatively affect the permeability of the intestinal barrier, as SCFAs are involved in the production of mucin and various tight junction proteins necessary for barrier function (Lam et al., 2012; Maffeis et al., 2016). A damaged gut barrier allows microbial products such as endotoxins and neuroactive substances (serotonin, kynurenine) to enter the bloodstream (Griffiths et al., 2004; Lyte, 2011; Barrett et al., 2012). Furthermore, these substances can damage the blood-brain barrier, leading to hippocampal dysfunctions (Kanoski et al., 2010; Freeman and Granholm, 2012). Therefore, not only gastrointestinal diseases such as inflammatory bowel disease are linked to the gut microbiota (Lucke et al., 2006; Wexler, 2007), but also neurodegenerative diseases such as Alzheimer's disease, anxiety or depression can be affected by gut microbes (Widner et al., 2000; Gulaj et al., 2010; Lach et al., 2018; Peirce and Alviña, 2019; Arora et al., 2020).

Secondly, microbiome research can help improve food supply chains by promoting animal health and a sustainable agriculture. In particular, pre- and probiotics are often used in animal production either to support animal health or to improve animal production (Lallès et al., 2007; Heo et al., 2013; Stanley et al., 2016; Roselli et al., 2017; Maki et al., 2019; Du et al., 2020;

Marmion et al., 2021). Another emerging aspect is the methane production in dairy cows, which is also influenced by the individual's microbiome (Roehe et al., 2016; Difford et al., 2018). In addition to the animal microbiome, the soil and root microbiome are also important factors in agricultural research. Due to the intensification of agriculture with synthetic fertilisers, soil erosion and biodiversity loss, the need for biofertilisation is increasing (Matson et al., 1997; Vitousek et al., 1997; Stoate et al., 2001). Many studies have shown that biofertilisers are able to increase soil bacterial diversity and thereby creating an unfavourable environment for plant pathogens (Garbeva et al., 2004; Dey et al., 2012; Qiu et al., 2012). In the rhizosphere, the soil microbiome directly interacts with the plant root microbiome, which is necessary for stress tolerance and pathogen protection (Berendsen et al., 2012; Bulgarelli et al., 2015). Therefore, increasing soil diversity positively affects the root microbiome, which is important for plant growth and crop yield.

Thirdly, microbiome research can contribute to the goal of a clean environment. Wastewater treatment is an important contributor and especially during the COVID-19 pandemic, microbiome surveillance of wastewater was used to detect viral mutations within cities or communities (Gallardo-Escárate et al., 2021; Brumfield et al., 2022). But it's not just the pandemic that has brought the wastewater microbiome into focus. More generally, wastewater treatment plants are seen as a source of antibiotic-resistant bacteria and antibiotic resistance genes, which can be identified using certain microbiome research approaches (Buelow et al., 2018; Leroy-Freitas et al., 2022).

Lastly, besides human health, food supply and a clean environment, the preservation of biodiversity is another major goal of the SDG's. Over 42,100 animal species are in danger of extinction, which makes captive breeding programs essential for species conservation (IUCN, 2022). As previously mentioned, the gut microbiome is closely related to the host health status – also in animals. To improve the reproductive success and animal welfare in captivity, microbiome analysis is a very useful tool. Although much research has been conducted in the field of animal microbiomes, this was mainly done on farm animals due to their importance in agriculture (Snelling et al., 2019; Wang et al., 2019; O'Hara et al., 2020). Rather less is known on the gut microbiota of threatened species, even if some faecal microbiomes are characterized in single studies on e.g. *Sarcophilus harrisii* (Cheng et al., 2015), *Apteryx mantelli* (San Juan et al., 2021) or *Ailurus fulgens* (Kong et al., 2014). This requires a large database of species-

specific microbiome reference data to rapidly detect deviations that may cause disease or change in behaviour.

1.2. Role of host related factors

Not only is the microbiota closely linked to the health status of the host, but also do other host-specific factors influence the gut microbiome, such as age, phylogenetic relationship and the host's diet. The microbiome of all mammals is shaped twice by age, once in the first year of life and again in the senior stage of life. New-borns rapidly come into contact with a wide variety of microorganisms, leading to rapid colonisation of their gastrointestinal system. This is characterised by a steady increase in alpha diversity with a change in microbial composition, when the juvenile is no longer weaned and starts to eat solid food (Jami et al., 2013; Guzman et al., 2016; Guevarra et al., 2019; Wang et al., 2019). During this period, the microorganisms switch from simple to complex carbohydrate digestion, enriched production of SCFAs and amino acid biosynthesis (Derrien et al., 2019). The elderly life stage in mammals is associated with a reduction in gut microbial diversity and SCFA production (Gavini et al., 2001; Dougal et al., 2014; Pellanda et al., 2021). This shift is mainly observed in the Firmicutes to Bacteroidetes ratio in favour of Firmicutes (Woodmansey et al., 2004; Mariat et al., 2009).
Another factor that shapes the microbiome is the host phylogeny. This relationship between the host phylogeny and the associated microbiome is known as phylosymbiosis (Lim and Bordenstein, 2020). In recent years, phylosymbiosis has been described for a variety of animal groups spanning the entire tree of life, from sponges (Easson and Thacker, 2014) to birds (Trevelline et al., 2020) and various mammals, including primates (Ochman et al., 2010). These findings show that the closer the host species are related, such as species within the same genera, the more similar their microbiota are (Li et al., 2018; Knowles et al., 2019). This co-evolution of gut microbes and host species could be driven by different factors such as different dietary (Ley et al., 2008a), geographical (Moeller et al., 2017) or ecological preferences (Grieneisen et al., 2019).

In addition to phylogeny, diet stands out as another important factor influencing both the gut and faecal microbiome. Of particular note is the role of dietary fibre, which is derived from carbohydrate intake and metabolised into SCFAs by gut microbes (O'Keefe, 2019). As explained in the previous section, a diet characterized low in fat but high in carbohydrate has a positive impact on the human gut microbiota. The influence of diet is further supported by the remarkable adaptability of the human microbiota, which can rapidly adapt to both plant-based

and meat-based diets within 24 hours. This adaptability triggers changes in the composition of the microbiota, resulting in changes in the products of carbohydrate and amino acid fermentation (Wu et al., 2011; David et al., 2014). This adaptability to a different diet has also been described in animals. The composition and diversity of the microbiota is strongly influenced by a herbivorous, carnivorous or omnivorous diet, favouring microbial taxa that can best degrade the respective nutrient inputs (Ley et al., 2008a; Muegge et al., 2011; Youngblut et al., 2019).

1.3. Role of microbes in carnivore and herbivore digestion

The evolution of the different digestive systems dates back to the first appearance of mammals in the Late Triassic about 200 million years ago (Lillegraven et al., 1979). Unable to compete with the dominant reptiles, their rise was characterised by endothermy, which allowed them to feed nocturnally (Crompton, 1980). There is increasing evidence that these first mammals were carnivores and that carnivory may have been the most ancestral state in animals for more than 800 million years (Vermeij and Lindberg, 2000; Román-Palacios et al., 2019). The digestive system of the 284 known Carnivora species is short and not very complex. As shown in Figure 1A, it consists of a simple stomach followed by a short intestine, a reduced colon and a small cecum (Stevens and Hume, 1995). Despite being hindgut fermenters, meaning that the highest microbial density and fermentation occurs in the large intestine, carnivores show only modest adaptations to microbial fermentation (Mackie, 2002; Suchodolski et al., 2008). This adaptation reflects their specialised dietary needs, such as low glucose requirements and the efficient processing of protein-rich food (Buffington, 2008; Verbrugghe et al., 2012). Accordingly, the carnivore microbiota consists mainly of bacterial taxa associated with this diet. For example, *Fusobacterium* and *Clostridium* are able to utilise ingested amino acids to produce the SFCAs acetate and butyrate (Olsen, 2014; Dahlstrand Rudin et al., 2021) and are therefore common inhabitants of the carnivore gut microbiota (Smith and Macfarlane, 1998; Vital et al., 2015; Milani et al., 2020).

Herbivory first evolved about 55 million years ago, with the first herbivores being small hindgut fermenting rodents with a limited intestinal capacity (Romer, 1966). Hindgut fermentation then evolved in the perissodactyls with modern specifications, e.g. in horses, tapirs or rhinos (Janis, 1976). Like carnivores, hindgut fermenters have a simple stomach (Figure 1B). In order to extract sufficient energy from the plant biomass, food retention time is extended by an enlarged intestine and cecum.

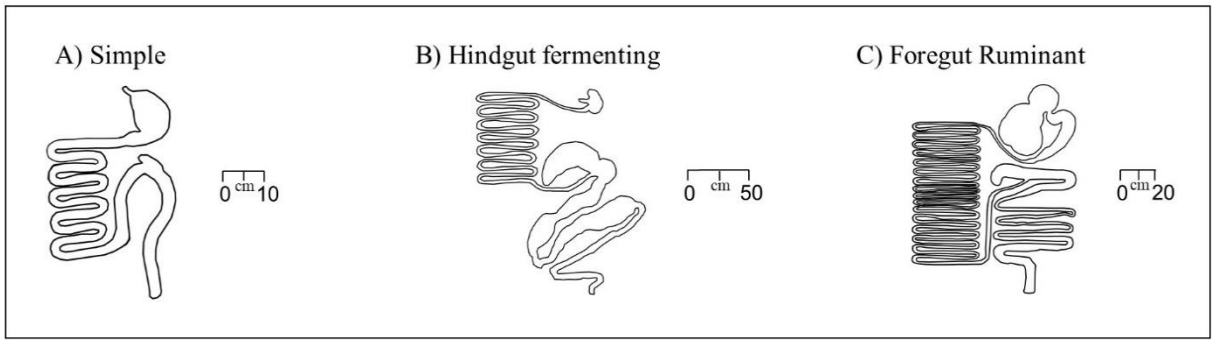

Figure 1: Comparison of the simple, hindgut fermenting and foregut ruminant digestive system. A) shows the monogastric digestive system of carnivores, characterized by a short and simple intestine and colon. B) shows the hindgut fermentative digestive system, which is common in e.g. Equidae species. Here, the microbial fermentation takes place in the enlarged cecum and hindgut, the monogastric stomach is similar to that of carnivores. C) shows the gastrointestinal system of ruminants. Microbial fermentation takes place in the highly compartmented stomach, particularly in the rumen. The figure is modified according to (Stevens and Hume, 1995).

As mammals are unable to break down cellulose, herbivores depend on a symbiosis with microorganisms that digest the plant material. In the case of hindgut fermenters, the main microbial fermentation takes place in the proximal cecum (Stevens and Hume, 1995). Some of the microbial fibre fermenters are Prevotellaceae or Spirochaetaceae species (Yatsunenko et al., 2012; Kartzinel et al., 2019; Quercia et al., 2019). Furthermore, *Fibrobacter succinogenes* plays an important role in the hindgut fermentation by degrading monosaccharides, glycosides and plant cell wall components.

The other prominent form of herbivory is the ruminant digestive system (Figure 1C). The expansion and compartmentalisation of the stomach first evolved in the Oligocene, about 38 million years ago, as the climate became drier. In order to adapt to this new environment, retention time is further increased and the ruminant process makes the nutrients more accessible (Gordon and Illius, 1994; Stevens and Hume, 1995). The first ruminants were the Camelidae. Other ruminants, such as the Bovidae, first evolved 38 million years ago (Janis, 1976). Unlike monogastrics, ruminants have a compartmentalised stomach consisting of the rumen, reticulum, omasum and abomasum. As shown in Figure 1B and Figure 1C, the large intestine is comparable to the hindgut fermenters while the cecum is much reduced in size (Mitchell, 1905; Douglas, 2018). The rumen is the main site of microbial fermentation. Here, cellolytic and amylolytic bacteria such as *Bacillus succinogenes* and *Ruminococcus spp* produce SCFAs as formate, acetate and succinate. Methanobacteriaceae are also important in the digestive process of ruminants as they use CO_2 and H_2 to produce methane. This mechanism is essential to prevent an increase in the partial pressure of H_2 in the rumen and to maintain constant environmental conditions for microbial digestive enzymes (Balch et al., 1979; Morgavi et al., 2010; Delzenne and Cani, 2011; Patra et al., 2017).

The evolution of the different digestive systems is accompanied by the adaptation of different gut and faecal microbial symbionts. These differ in their composition, diversity and functional profile based on the respective host gut morphology and dietary specializations (Ley et al., 2008a; Muegge et al., 2011; Nishida and Ochman, 2018).

1.4. Microbiome analyses as a tool for animal welfare in zoos

In animal conservation, microbiome research has emerged as a new aspect of understanding and improving the well-being of animals. Over the past decade, research has shed light on the relationship between the gut microbiome and various facets of animal health and behaviour in the context of zoo environments. For example, stress which may occur in captivity (Morgan and Tromborg, 2007; Mason, 2010) can alter the gut microbiome (Madden et al., 2022). Furthermore, the microbiome can also affect the reproductive success, as some microbial genera are highly correlated with specific hormones, such as progestogen, which is important for pregnancy (Antwis et al., 2019). Reproductive success may also be influenced by diet, as some phytoestrogens in commonly fed alfalfa may lead to low reproductive rates because the microbiota may not be able to digest these metabolites (Tubbs et al., 2016; Tubbs et al., 2020). Another important aspect of animal conservation is the reintroduction of certain animal species. It is unclear to what extent the microbiomes of captive and wild animals differ from one another. Some studies have reported a decrease in diversity in *Equus caballus* (Metcalf et al., 2017) or *Vombatus ursinus* (Eisenhofer et al., 2021), while others have found no differences or even a higher microbial diversity in zoo-housed *Mirounga leonine* and *Hydrurga leptonyx* (Nelson et al., 2013), *Diceros bicornis* (Gibson et al., 2019), *Papio kindae* and *Papio ursinus* (Tsukayama et al., 2018) and different species within the Giraffidae and Bovidae (McKenzie et al., 2017). Nevertheless, it is important to consider the microbiome of wild populations before releasing captive individuals. Due to habitat fragmentation, the microbiome of wild species may have adapted to the new environment, for example, red colobus monkeys have lost the ability to detoxify plant xenobiotics (Barelli et al., 2015). Therefore, it is necessary to ensure that captive individuals are able to cope with the environment prior to release.

Overall, microbiome research has great potential in animal conservation, mainly because it is a non-invasive tool to assess the health status of the host in terms of stress, disease and even fertility. Faecal samples are easy to collect and, with ongoing advances in sequencing methods, this approach is also cost effective. The fundamental requirement for microbiome research as an animal welfare tool, is a complex and well-curated database of information on the

composition and diversity of the microbiome in a variety of species. This should include as many individuals as possible of a given species and, in the best case, individual longitudinal data to capture normal variation in the microbiota. Zoos provide an ideal environment for this research for two reasons. First, they house a variety of species, including rare and endangered species where faecal sampling is easily integrated into the keepers' daily cleaning routine. Second, zoos provide a variety of sample metadata information such as age, sex, pedigree, medical treatments and feeding protocols. All of these factors can be used to establish a species-specific microbiota profile from which individual deviations can be easily detected.

1.5. Research questions and scope of the studies

The aim of this thesis is to lay the foundations for a microbiome database for a wide range of animal species. In order to achieve this, the first step is to characterise the faecal microbiota of a number of mammalian species kept in zoos in terms of their composition and diversity. The focus lays on comparing herbivores and carnivores, as diet is one of the most important factors influencing the faecal microbiome (Ley et al., 2008a). The second step involves analysing the microbiota of different individuals, including the daily oscillations that make longitudinal data necessary for future studies. Finally, to make this research accessible in the field and transferable to wildlife, a method is developed to identify host-specific metadata from the microbial composition of an unknown faecal sample.

1.5.1. Publication A: Unravelling differences in faecal microbiota stability in mammals: from highly variable carnivores and consistently stable herbivores

Microbiome research has gained considerable popularity in the last decade, but studies on animals have largely been carried out on livestock to increase food production. Studies on exotic or endangered species are rare, focusing mainly on popular species such as *Phascolarctos cinereus* (Alfano et al., 2015; Brice et al., 2019) or *Diceros bicornis* (Gibson et al., 2019; Burnham et al., 2023). In addition, two different sampling strategies have typically been used, either analysing many samples of one species or a few samples of many species (Youngblut et al., 2019; Milani et al., 2020). To create a reliable database of microbiome data, this first study analyses faecal samples from 31 mammalian species. The first aim is to characterise the differences in the microbial composition of Canoidea, Feloidea, Ruminantia and Perissodactyla and to apply an indicator family analysis. Second, the microbial diversity within this groups is

calculated. Third, the variation of the main bacterial families is analysed by calculating the coefficient of variation. This variation is important to have in mind when the number of samples per species is under consideration. Overall, this first study aims to provide data from different species that can be used to build a faecal microbiome database, focusing on dietary differences between carnivores and herbivores.

1.5.2. Publication B: Time series cluster analysis reveals individual assignment of microbiota in captive tiger (*Panthera tigris*) and wildebeest (*Connochaetes taurinus*)

Beyond the simple characterisation of the faecal microbiota, there are few studies describing the natural variation of an animal's microbiota. Most of the longitudinal studies focus on either dietary changes (Lyu et al., 2018; Butowski et al., 2019), juvenile development (Guevarra et al., 2019; Wang et al., 2019; Amin and Seifert, 2021) or disease progression and treatment (Mamun et al., 2020; Ayoub et al., 2022), mainly in farm animals. Therefore, this second study concentrates on the natural oscillations in the microbiota of a herbivorous (*Connochaetes taurinus*) and a carnivorous (*Panthera tigris*) host species. The main research questions to be answered here are whether there are species- or individual-specific oscillations in the faecal microbiota and which microbial taxa are mainly responsible for these. Additionally, the results of **Publication A** are further investigated to determine whether, for example, the high variability in the microbiome of carnivores is species- or individual-specific. Therefore, this study applies two approaches. First, each sample of a time series is analysed individually using two clustering algorithms. Second, the time series data per individual are merged by dynamic time warping prior to clustering. Finally, the respective cluster-specific taxa are identified to infer a species- and individual-specific microbiota.

1.5.3. Manuscript C: Development and evaluation of an ensemble model to identify host-related metadata from fecal microbiota of zoo-housed mammals

The complex interplay of diet, digestive system (simple, hindgut, ruminant) and host species, with a focus on phylosymbiosis, are considered to be key factors influencing the composition of the faecal microbiota (Ley et al., 2008a; Muegge et al., 2011; Nishida and Ochman, 2018). **Manuscript C** aims to first investigate the correlation between microbial taxa and these three factors using faecal samples from 14 species. In a next step, these correlations are used to compute a regression model that predicts host-specific metadata from a given microbiota

composition at the three levels of diet, digestive system and host family. The overall goal of this manuscript is to open up new opportunities in microbiome research, especially for field studies. As it is difficult to obtain host-specific metadata in the field, it would be a great advantage if these could be obtained from faecal samples. The combination of microbiome analysis and host metadata offers a cost-effective alternative to conventional, resource-intensive microsatellite analysis and represents a promising approach for future research and conservation efforts.

2. Methods

This section of the thesis provides an overview of the sampling and sequencing procedure and the applied research methods. Within this context, the calculation of diversity measurements, cluster analysis and regression model development are described by the use of different statistical methods.

2.1 Faecal Sample collection

A total of 669 faecal samples were collected in 20 German zoos. This includes 31 mammalian host species from the three orders Perissodactyla, Artiodactyla and Carnivora. Within the Artiodactyla, the focus of the studies lies on the suborder of Ruminantia and, within the Carnivora, on the two suborders of Canoidea and Feloidea (Figure 2).

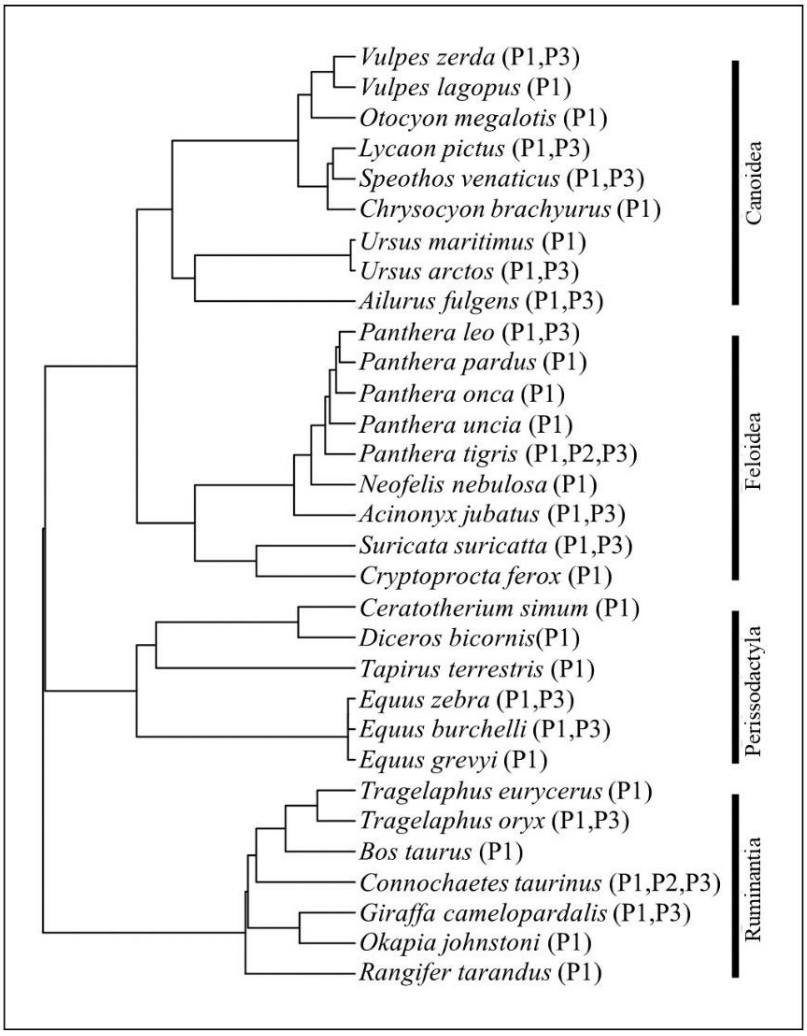

Figure 2: Phylogenetic tree of all analysed host species. The (sub-)order of each species that is analysed in this thesis is indicated by the line on the right dividing the dataset into four clades. The letters in brackets highlight the publication in which a species is analysed. The tree is generated using TimeTree (Kumar et al., 2022) and the implemented hierarchical average linkage method to estimate species divergence times.

From this total dataset, partially overlapping data subsets were created for the three studies. Regarding the different research questions, those subsets contain either a large variety of host species (Publication A, Manuscript C) or focus on a few host species but on different individuals within these (Publication B) (Table 1).

Table 1: Subsets of the total dataset used for each publication. The number of host species, zoos, and faecal samples for the three publications/manuscript is given in the columns.

Publication/ Manuscript	Number of host species	Number of zoos	Number of samples
A	31	20	621
B	2	5	95
C	14	17	525

The sampling procedure was carried out by the zoo keepers non-invasively during the daily enclosure cleaning. Where possible, sampling of individual animals was preferred, otherwise as many samples as there were individuals in the enclosure were taken. Faecal samples were collected in sterile 2mL cryotubes and were stored in liquid nitrogen as soon as possible. Subsequent sample processing was performed by StarSeq GmbH in Mainz, Germany. Here, all samples were homogenised in a Precellys® Evolution Homogenizer (Bertin Instruments, Rockville, USA) before DNA extraction was performed applying the QIAmp® PowerFaecal DNA kit (Qiagen, Hilden, Germany) and the DNA concentration was measured by NanoDrop spectrophotometer (Thermofisher, Massachusetts, USA).

2.2 16S rRNA gene

The 16S rRNA (*rrs*) gene is located on the *rrn* locus along with the 23S (*rrl*) and 5S (*rrf*) genes (Pace, 1973; Bram et al., 1980). This gene is commonly used for microbial taxa identification because it meets all the requirements for a reliable marker gene. First, it is ubiquitous in prokaryotic organisms and occurs in high copy numbers because it forms the 30S subunit of prokaryotic ribosomes together with other proteins and therefore is essential in protein metabolism (Fogel et al., 1999; Klappenbach et al., 2001; Rosselló-Mora, 2001). Second, its evolutionary rate is low and it is little affected by horizontal gene transfer and external environmental factors (Jain et al., 1999; Daubin et al., 2003; Pontes et al., 2007). Third, the 16S rRNA gene contains both, highly conserved and highly variable regions. With a total length of about 1,500bp, it consists of nine highly variable regions (V1-V9) suitable for taxa identification. Those variable regions are interspaced by conserved regions which are used as primer binding sites (Brosius et al., 1978; van de Peer et al., 1996; Janda and Abbott, 2007; Kim et al., 2011). For this thesis, a combination of the V3 and V4 region was used, because this is known to exhibit the highest nucleotide diversity. Within the nine variable regions, the V4 region is more conserved because it interacts directly with the tRNA in the translational process than the faster evolving V3 region (Schluenzen et al., 2000; Morosyuk et al., 2001). This enables both, the detection of higher-level taxa as well as the identification on family or genus level (Bukin et al., 2019).

2.3 Illumina® sequencing of the 16S rRNA gene

Illumina® sequencing is a next-generation sequencing method that allows high-throughput sequencing of many parallel samples. Sequencing is performed on a flow cell to which different oligopeptides (forward and reverse primers) are attached (Figure 3).

In a first step, the DNA is sheared into small fragments. Next, different sequencing adapters, indices and sequencing binding sites are added to both ends. These DNA amplicons then bind to the complementary primer on the flow cell. The DNA polymerase then binds to the target sequence and synthesises the complementary strand. The resulting double strand is denatured and the template is discarded. The second step is bridge amplification. The resulting single strand overlaps and hybridises with a complementary primer on the flow cell. Again, the polymerase synthesises the complementary strand, resulting in two single strands bound to the flow cell.

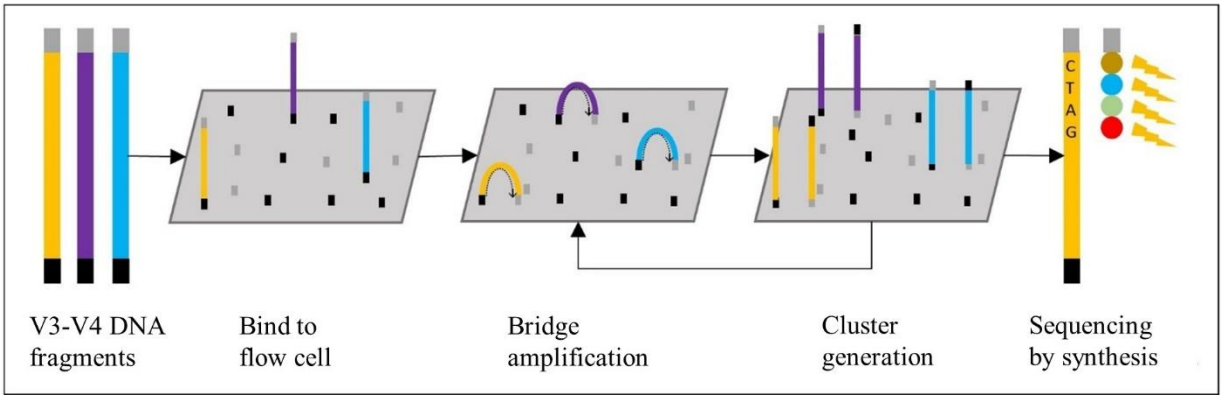

Figure 3: Illumina® sequencing workflow. Amplified DNA fragments from the V3-V4 region of 16S rRNA are attached to the flow cell by complementary primer pairs. The strands overlap, forming a bridge to free primers on the flow cell. The DNA polymerase then binds and constructs the complementary DNA strand. This process is repeated until millions of DNA clusters are formed on the flow cell. Fluorescent nucleotides are then added, which emit specific light signals when they bind to the template sequence. These signals are captured by a camera and the resulting DNA sequence is stored. Figure created according to Lu et al. 2016.

This process of bridge amplification is repeated until millions of sequence clusters are generated across the flow cell. The third step is the sequencing by synthesis of the forward and reverse strands. The reverse strands are washed away, the 3' ends of the forward reads are blocked and the sequencing primer binds to the adapter region. The key components of this sequencing method are fluorescent nucleotides that emit a nucleotide-specific light signal when they bind to the template sequence. The wavelength and intensity of the emitted light is captured by a camera for each cluster and the corresponding nucleotide is determined. Once the forward read has been sequenced, it is discarded and the blocked 3' end is released. The bridge amplification is then repeated to produce the reverse reads (Illumina Inc, 2010; Lu et al., 2016).

For this research, DNA extraction, library preparation and Illumina® sequencing were outsourced to StarSeq GmbH in Mainz, Germany. Here, a dual-index strategy for sequencing the V3-V4 region of the 16S rRNA gene was used, following the protocol described in (Caporaso et al., 2012) with slight modifications. Amplicon amplification was performed in a single-step PCR with 33 cycles, using the primer combination 341f and 806b (Takahashi et al., 2014; Apprill et al., 2015). The final library was then sequenced on the Illumina® MiSeq platform in paired-end mode at 300bp each, a sequencing depth of 100,000 reads and a 25% PhiX control library for quality control purposes.

2.4 Data analyses

Demultiplexed, paired-end reads are generated by StarSeq GmbH. The following pre-processing is applied to each of the three datasets prior to the in-depth analysis. All samples are processed following the Qiime2 pipeline (Bolyen et al., 2019). In a first step, quality filtering and amplicon sequence variant (ASV) calling is applied using the integrated DADA2 plugin (Callahan et al., 2016). During the creation of FastQ files, phred scores (Q) are calculated for each nucleotide in a read with the following formula (Ewing et al., 1998):

$$Q = -10 * \log_{10}(p)$$

p is the probability that a nucleotide was sequenced incorrectly. A phred score of Q=30 therefore has a 1:1000 chance that this nucleotide was misidentified. The read quality usually drops towards the end of the reads due to emission interferences during the sequencing process. Those occur when some molecules per cluster fail to bind a nucleotide during a cycle. Due to the resulting asynchrony within the cluster, different light signals are released at the same time. Towards the end of the read, these interferences accumulate, reducing the quality of the read (Schirmer et al., 2015).

2.4.1 Read pre-processing and taxonomic classification

The first step in pre-processing is the read quality trimming implemented in the DADA2 plugin. Reads with a base quality below 20 are trimmed off from their ends. In a second step, DADA2 performs a dereplication of the read-set using a modified divisive portioning algorithm (Rosen et al., 2012). This involves grouping reads with the same nucleotide sequence into a centre sequence, including abundance information. All reads are then compared to the centre sequences and nucleotide deviations are calculated. This information, together with the type of mismatch, is used to compute and train a dataset-specific error model. In a third step, the reads

are denoised using the divisive amplicon denoising algorithm, which is based on the previously trained error model (Callahan et al., 2016). This algorithm calculates the probability that an ASV is too frequent to be explained by sequencing errors. Singletons are also discarded. In a fourth step, paired-end reads are merged with a Needleman-Wunsch sequence alignment (Needleman and Wunsch, 1970) and a minimum overlap of 12 nucleotides. Finally, ASV's that occur less than 10 times and chimeric sequences are removed.

The remaining high-quality reads are then used for taxonomic classification. Here, a pre-trained naïve Bayes classifier (Bokulich et al., 2018; Robeson et al., 2020) is applied to compare the dataset to the SILVA 138 full-length database (Quast et al., 2013). Subsequently, chloroplast and mitochondrial sequences are removed from the dataset.

2.4.2 Alpha and beta diversity analysis

To calculate alpha and beta diversity measures, the filtered reads are aligned to create a phylogenetic tree for subsequent diversity measurements. This is done with the tool mafft (Katoh et al., 2002), which performs a de novo multiple sequence alignment on the dataset. The aligned sequences are then the input data for FastTree (Price et al., 2010) to create a phylogenetic tree of all taxa being present in the samples. This tool works with a heuristic neighbour joining approach, an agglomerative clustering method, to create a rough tree and improves this by likelihood rearrangements. The distances between sequences are estimated by the Jukes Cantor nucleotide model (Jukes and Cantor, 2013).

Alpha diversity is a measurement to describe the microbial diversity within a sample. In order to compensate for fluctuations in the sequencing depth and not to discard any samples, a multiple rarefaction is performed to the read count of the smallest sample per dataset. Subsequently the Shannon index and effective number of species (ENS) are calculated. The Shannon index is a quantitative diversity measurement that relies on the species abundances. It is calculated with the following formula:

$$H = -\sum_{i=1}^{s} (p_i * \ln(p_i))$$

In this formula, s is the number of microbial taxa and p_i the relative abundance of taxon i. High values indicate greater diversity within a sample and vice versa. H=0 is the minimum value and indicates that a sample only consists of one taxon. The Shannon index reaches its maximum when all taxa are represented equally.

Additionally, the ENS is calculated for each sample. This measurement aims to identify the number of equally-common taxa within a sample (Jost, 2006). To convert the Shannon index to the ENS, the following formula is applied:

$$\mathrm{ENS} = \exp(\mathrm{H})$$

The resulting data is statistically tested in R version 3.4.1 (R core Team, 2020) applying the packages vegan (Jari Oksanen et al., 2012) and FSA (Ogle et al., 2020). To test for normal distribution, the Shapiro-Wilk test (Shapiro and Wilk, 1965) is used. Furthermore, differences between metadata categories are analysed with the Kruskal-Wallis test (Kruskal and Wallis, 1952) followed by a Dunn's test with Bonferroni correction (Dunn, 1964).

While the previously described alpha diversity only refers to one sample, beta diversity describes the similarity between all samples. Therefore, this metric relies on a distance matrix. In this case, this has been calculated on the qualitative unweighted (u) and quantitative weighted (W) unique fraction metrics (UniFrac) (Lozupone and Knight, 2005; Lozupone et al., 2007). Both metrics calculate a distance matrix based on a given phylogenetic tree of all sequences and focus on branch length that lead to taxa which are unique for a given environment or sample with the formulas:

$$u = \frac{\sum_{i=1}^{N} l_i |A_i - B_i|}{\sum_{i=1}^{N} l_i \max(A_i, B_i)} \qquad W = \frac{\sum_{i=1}^{N} l_i \frac{A_i}{A_T} - \frac{B_i}{B_T}}{\sum_{j=i}^{S} L_j}$$

Here, N refers to the numbers of nodes in the tree and l_i to the branch length between a node i and its parent. A and B are descendants from the respective node. In the unweighted UniFrac formula, A and B are defined as present or absent. Regarding the weighted UniFrac, S indicates the total number of sequences in the tree and L_J is the overall branch length. In this case, A_i and B_i are sequence numbers descending from a node and A_T and B_T are the total amount of sequences from environment A and B.

Overall, the unweighted UniFrac searches for presence and absence of taxa to highlight structural differences in the phylogenetic tree, the weighted UniFrac includes abundance information to weight the branch lengths.

2.4.3 Time series analyses

For a deeper understanding of microbial fluctuations within an individual, longitudinal data is necessary. In **Publication B**, time series data over a period of eight days are analysed in a clustering approach with and without dynamic time warping (Figure 4). All analyses are performed in Matlab v9.11 (The MathWorks Inc., 2020a) using the software CASE (Schneider et al., 2022). Dynamic time warping synchronises time series with different characteristics, e.g. different length due to discrepancies in sampling points. To find an optimal alignment of the time series, they are divided into equal points. Subsequently, the distances between the first point of time series A and all points from time series B are calculated and the smallest distance is stored (Sakoe and Chiba, 1978; Paliwal et al., 1982). This process is repeated until all smallest distances between both time series are stored in a similarity matrix, which is the input for the clustering algorithms. Two different hierarchical cluster algorithms were applied, namely Ward's linkage (Ward, 1963) and a community detection algorithm (Newman, 2004). Ward's linkage evaluates the proximity between two clusters using a linkage function. This detects the increase in the error sum of squares that occurs when two clusters are merged into a single entity. The core objective of Ward's method is to minimise the spread of errors at each clustering step. This is achieved by judiciously selecting the most appropriate clustering steps.

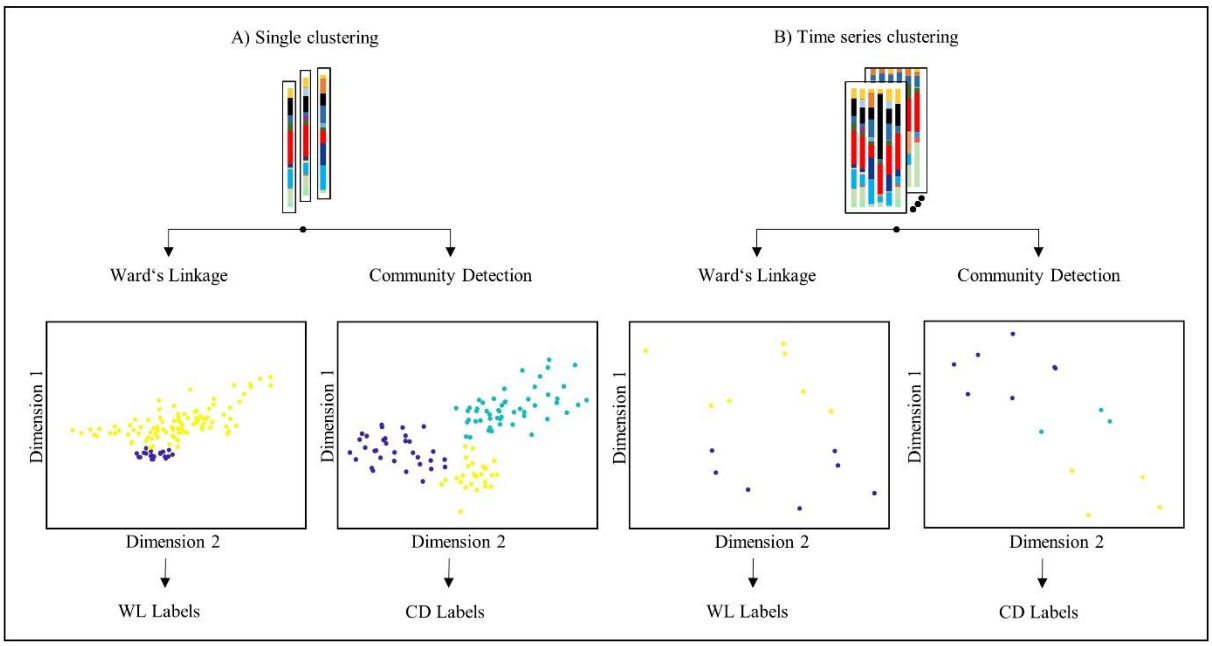

Figure 4: Clustering pipeline of longitudinal microbiome datasets. In A), the single clustering pipeline is shown, for which all eight samples of an individual are treated as single samples. B) shows the time series clustering approach, in which the eight samples per individual undergo dynamic time warping prior to clustering. Both approaches are treated similarly during the clustering process. Ward's Linkage and Community detection algorithms are applied to each dataset and the clustering results are visualised. Additionally, the labels for each cluster in each approach are extracted including the normalised mutual information (NMI).

Community detection is frequently used in network analyses. In microbiome data, the network is defined as follows: the microbial composition of the samples are the nodes and the distance between samples the edges. The aim of community detection is to group vertices into clusters. First, each vertex is considered as a separate community, which are then merged based on the number of edges. The results are clusters with many internal edges defining the community and few edges connecting vertices from different clusters. Finally, the normalized mutual information (NMI) is calculated as a measure of cluster reliability. Furthermore, the LASSO algorithm (Tibshirani, 1996) is used to identify taxa that cause the cluster-specific differences as described in more detail in **Publication B**.

2.4.4 Regression model development for host identification

In **Manuscript C**, a logistic ensemble model is developed to identify the host diet, digestive system and family. For this approach, species with less than 20 samples are discarded from the dataset. To test for the influence of the three categories prior to model development, a Pearson correlation is calculated. A general linear model (glm) is then used to assess the statistical significance of moderately and strongly correlated bacterial families as factors explaining whether or not a given microbiome composition belongs to one of the specified categories. The resulting dataset is then divided into trainings (64%), development (16%) and test (20%) subsets. The preparation of the data confirms the assumption that the data set is independent and identically distributed.

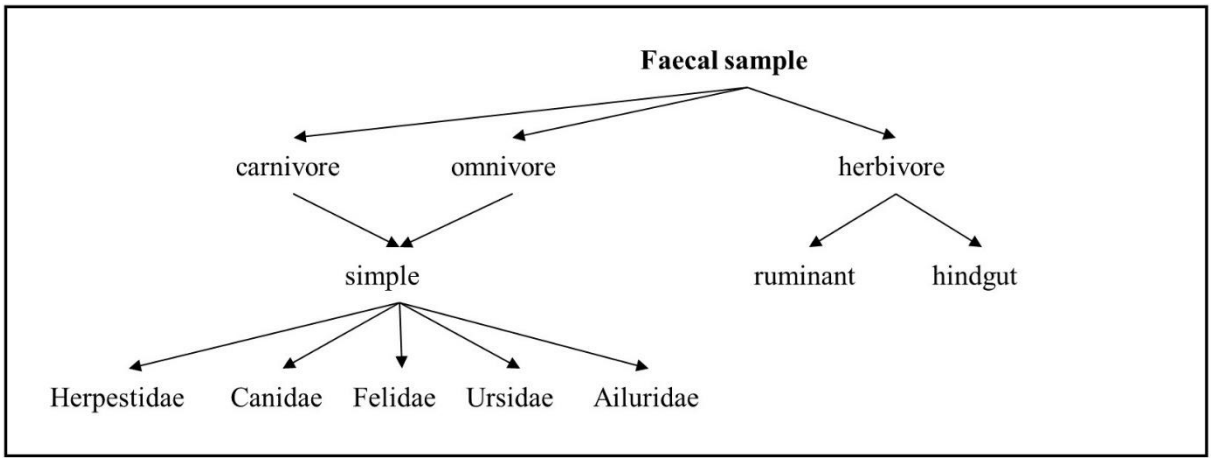

Figure 5: Development of the ensemble logistic regression model. Based on the microbiota composition of a faecal sample, logistic model performs different steps to characterize host-specific metadata. First, the model identifies the diet type being carnivore, omnivore or herbivore. Second, for herbivorous hosts, the model decides between a ruminant or a hindgut fermenting digestive system. In contrast, for carnivorous and omnivorous hosts, the model distinguishes between the host family.

The model development includes different steps. First, a logistic regression model ist trained to distinguish between a carnivorous, omnivorous and herbivorous host. Due to their digestive systems, herbivorous hosts are treated separately from omnivorous/carnivorous hosts in the following steps (Figure 5). Regarding the herbivores, a second logistic regression model is built to distinguish between the digestive systems (simple, hindgut, ruminant). For carnivores which all have the same digestive morphology, a third logistic regression model is built to distinguish between the host families. All of these models are validated using the development dataset. To further improve the reliability, the probabilities of the models are aggregated to an ensemble model by taking the average of each single regression model. Finally, the performance of this combined model is validated by applying the test dataset. The performance is quantified by using the F1 scores which combines the accuracy and precision of the models.

3. Results

Given these methodological considerations, this chapter aims to summarise the results of the three studies briefly. As all studies used different datasets, methodologies and research questions, the findings are described individually.

3.1 Publication A

This publication, which includes the largest dataset of 621 samples, focuses on describing the differences in faecal microbiota between herbivores and carnivores. More specifically, the suborders Canoidea and Feloidea are compared with the orders Perissodactyla and Artiodactyla. The main results of this publication are that carnivores and herbivores differ significantly in their microbiota composition (ANOSIM statistics: $R=0.50$, $p<0.001$, permutations:999, distance= Bray Curtis), diversity (Kruskal-Wallis: $p<0.001$, $df=3$, Dunn's test with Bonferroni correction $p>0.001$) and consistency. The main bacterial families contributing to the herbivore microbiota are Spirochaetaceae, Lachnospiraceae, Rikenellaceae and Oscillospiraceae. In contrast, carnivores are dominated by Fusobacteriaceae and Clostridiaceae. In terms of microbial diversity, herbivores show a significantly higher alpha diversity in all measurements compared to carnivores. Beta diversity analyses further reveal a separation of carnivores from herbivores. Additionally, the herbivores form two distinct clusters representing Perissodactyla and Ruminantia. In contrast, the carnivore samples are more diverse and heterogeneous, indicating no visible separation of Feloidea and Canoidea. This high variation within Carnivora is further supported by the high coefficient of variation of bacterial taxa in this order.

3.2 Publication B

To gain a deeper understanding of the variation within the faecal microbiota of carnivores and herbivores, this study analyses individual time series data from a herbivore (*Connochaetes taurinus*) and a carnivore (*Panthera tigris*) species. Therefore, two clustering algorithms, namely Ward's linkage and community detection, are applied with and without prior dynamic time warping. Species-specific clustering reaches the best results when dynamic time warping is applied before clustering the data with Ward's linkage. In this case, perfect clustering (NMI=1) is achieved, splitting both species into two separate clusters. In contrast, single clustering without dynamic time warping fails in species recognition (NMI_{Com}=0.69, NMI_{Ward}=0.25). In a second step, individual detection is tested for both species. Again, individual identification fails using the single-sample approach. In contrast, the application of dynamic time warping results in reliable individual clustering. For both species, the best results are obtained by applying the community detection algorithm. This results in a NMI_{Com}=0.97 and seven clusters for the *Panthera tigris* individuals and NMI_{Com}=0.82 and six clusters for *Connochaetes taurinus*. Limitations arise when considering two time series of an individual that are further apart in time. These are not assigned to a single individual, as the algorithm groups them into separate clusters.

In addition, cluster-specific bacterial taxa are identified using the LASSO algorithm as implemented in the CASE software. The species-specific taxa are Clostridiaceae and Fusobacteriaceae for *Panthera tigris* and Spirochaetes and Methanobacteriaceae for *Connochaetes taurinus*. Individual-specific bacterial taxa are characterised in both species as a combination of daily occurring core bacteria, e.g. Clostridiaceae and Enterobacteriaceae in the *Panthera tigris* or Spirochaetes in *Connochaetes taurinus*, and individual low abundance bacterial families that show greater daily variability.

3.3 Manuscript C

As the former publications suggest diet, digestive system and host species to have a strong effect on the faecal microbiota, this manuscript aims to develop a model to predict those host-specific factors. In a first step a Pearson correlation is performed on the dataset to analyse the influence of diet and digestive system on the bacterial taxa. Herbivores show many positive correlations with the strongest correlational values being Rikenellaceae (R=0.73) and Ruminococcaceae (R=0.71) while carnivores show less positive correlations (e.g. Fusobacteriaceae: R=0.55). Omnivore species in contrast, do not have any strong correlations.

In regard of the digestive system, ruminants are strongly correlated with Ruminococcaceae (R=0.77) and Methanobacteriaceae (R=0.64) while hindgut fermenters show strong correlational values with Spirochaetaceae (R=0.86). Carnivores, having a simple digestive system, correlate positively with Clostridiaceae (R=0.56) and Fusobacteriaceae (R=0.52).

In the next step, a logistic regression model is built to identify the diet, digestive system and host family from a microbiota composition of an undisclosed faecal sample. First, the model distinguishes between a carnivorous, herbivorous or omnivorous host with an overall accuracy of 88%. The F1 scores vary between 0.73 for omnivores, 0.87 for carnivores and 0.93 for herbivores. Within the herbivores, the model then distinguishes between a simple, hindgut fermenting or foregut ruminant digestive system. The overall accuracy is 98%, with perfect results for the hindgut fermenter (F1=1.00), followed by ruminants (F1=0.98) and the simple digestive system (F1=0.92). Due to the uniformity of simple digestive systems across all carnivores, it is not possible to discriminate the digestive system. Instead, the model decides between the carnivorous host families with an overall accuracy of 79%. Canidae are identified with a high F1 score of 0.93, while Felidae and Ursidae are identified with a slightly lower reliability of 0.83 and 0.79 respectively. Only the Herpestidae could not be distinguished from the other families, with an F1 score of 0.00.

4. Discussion

4.1 Challenges and difficulties

Faecal samples are established as a proxy for microbiome analysis because the microbial density increases along the gastrointestinal system, with more than 98% of the human microbiota living in the colon (Haller and Hörmannsperger, 2015; Douglas, 2018). As they have a similar digestive system, the same is true for carnivores and hindgut fermenters. In contrast, ruminants show the highest abundance of microorganisms in the rumen, but due to lacking comparability, many studies also focus on faecal samples of ruminants (Tanca et al., 2017; Clemmons et al., 2019; O'Hara et al., 2020). Another reason for using faecal samples for microbiome research, especially in animals, is the non-invasive sampling method, which allows for easily repeated longitudinal sampling. For research on zoo animals, the EAZA (European Association of Zoos and Aquaria) research strategy focuses on non-invasive research approaches to assess the health status of different endangered animal species (European Association of Zoos and Aquaria, 2022). The use of faecal samples requires some attention to handling and storage procedures. On the one hand, there are many different recommendations

on how a given storage temperature affects the microbiota composition of the sample. It is unclear how long faecal samples can be stored at room temperature or at 4°C without affecting the microbial composition, often varying between one and five days (Roesch et al., 2009; Carroll et al., 2012; Choo et al., 2015; Tedjo et al., 2015; Ezzy et al., 2019). For long-term storage, a temperature of -80°C is most suitable to maintain a stable microbiota composition for at least six months (Carroll et al., 2012; Fouhy et al., 2015). Temperature-induced changes are mostly observed at lower taxonomic levels such as species or genus (Cardona et al., 2012). Therefore, all taxonomic analyses in this thesis have been carried out at the microbial family level. On the other hand, repeated cycles of freezing and thawing must be avoided as this can lead to a degradation of the microbial DNA and to changes in the cellular structure of gram-positive bacteria (Bahl et al., 2012; Fouhy et al., 2015).

To minimise these storage-related effects on the microbiota, the samples for all studies are processed according to the same collection and storage scheme. Keepers collected the samples in the morning so that the time between defecation and sampling was a maximum of 12 to 24 hours. To further reduce the influence of oxygen on the surface of the sample, a subsample is taken immediately from the core of the faeces. This was done to avoid bias in the ratio of aerobic to anaerobic bacteria. Even if all these things are taken into account, there are some factors that could not be considered in the studies. For example, the rate of change of the microbiota also depends on the nutrient composition of the individual samples or the host species. Similarly, contamination with soil bacteria or weather effects cannot be completely prevented (Roesch et al., 2009; Menke et al., 2015).

Due to the rapid development of a large number of sequencing methods and bioinformatic analysis tools, the comparability of microbiome studies is often limited. For this reason, for all three studies in this thesis the same laboratory equipment (DNA extraction kit, library preparation, sequencing method) and a standardised analysis approach of DNA data are used.

4.2 Comparison of the carnivore and herbivore faecal microbiota

Based on their completely different diets and digestive morphology, carnivores and herbivores have significantly different microbiomes (Ley et al., 2008a; Muegge et al., 2011; Youngblut et al., 2019; Guo et al., 2020; Milani et al., 2020). This is being confirmed throughout this thesis, both in terms of microbiota composition and diversity (Figures 6,7). Herbivores lack plant degrading enzymes such as glycoside hydrolases and polysaccharide lyases and are therefore dependent on microbial fermentation (Stevens and Hume, 1995; Russel and Rychlik, 2001).

Bacterial families that are responsible for cellulose degradation are Prevotellaceae, Rikenellaceae, Oscillospiraceae and Spirochaetaceae. In this thesis, they were found in high abundance in all herbivorous species, which is consistent with previous studies (Kartzinel et al., 2019; Milani et al., 2020). In the human gut microbiome, Prevotellaceae represent a distinct enterotype based on a low-fat, high fibre diet (Arumugam et al., 2011). Members of this family are involved in the glucose metabolism, which is a necessary step in polysaccharide digestion (Wang et al., 2016). Furthermore, Rikenellaceae, Spirochaetaceae and Lachnospiraceae have been found in the faeces of many herbivorous species as they also play an important role in carbohydrate fermentation (Cwyk and Canale-Parola, 1979; Yatsunenko et al., 2012; DeLong et al., 2014; Obregon-Tito et al., 2015; La Reau and Suen, 2018; Lau et al., 2018; Vacca et al., 2020). In addition to the plant diet, herbivorous animal species differ in the site of microbial fermentation. Fermentation occurs either in the hindgut after enzymatic digestion or in the foregut prior to enzymatic digestion. Some studies have shown that the microbiome is adapted to the type of digestive system, but the data set was very limited and included only single samples per species (Ley et al., 2008a; Muegge et al., 2011). In this thesis, these differences in the microbiota are confirmed using a variety of species with multiple samples. On the one hand, these differences occur in the composition of the microbiota, e.g. in the presence of methanogenic archaea in ruminants, especially Bovidae (Figure 6). On the other hand, diet and digestive system can be perfectly predicted using a logistic ensemble model. These results further indicate the influence of diet and digestive morphology on the microbiome.

In contrast to herbivores, this thesis shows that the microbiota of carnivores is less complex and diverse (Figure 6,7). It is mainly adapted to digest a high fat and protein-rich diet with a simple digestive system. The microbial taxa that perform amino acid fermentation to produce SCFAs are mainly members of Clostridiaceae, Fusobacteriaceae and Peptostreptococcaceae (Wiegel et al., 2006; DeLong et al., 2014; Slobodkin, 2014). This adaptation to protein metabolism is consistent with previous studies showing that carnivores express fewer carbohydrate and more choline and trimethylamine degradation pathways (Milani et al., 2020).

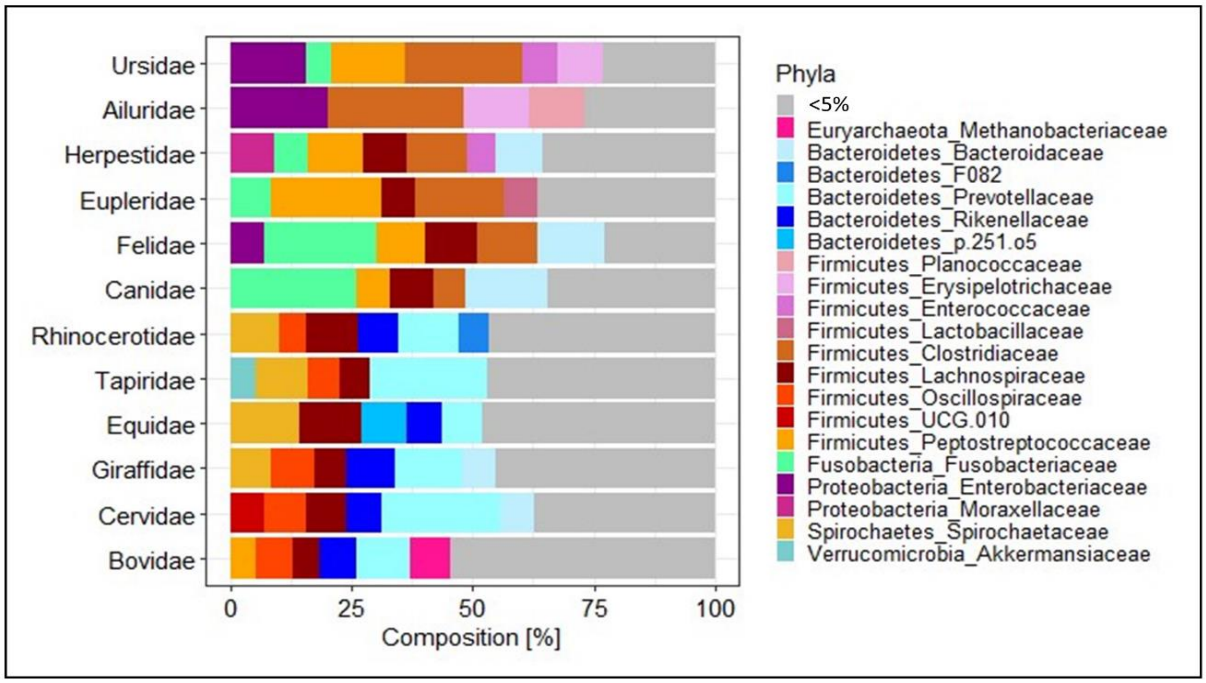

Figure 6: Taxonomic composition of the carnivore and herbivore faecal microbiota based on the SILVA database. The microbial composition is averaged for each mammalian family. Taxa that occur in less than 5% of the overall composition are summed up. The legend indicates the respective phylum and family of each microbial taxon. Figure created according to Zoelzer et al. 2021.

In addition to the composition, the microbial alpha diversity differs significantly between carnivores and herbivores (Figure 7). All of the herbivorous host families assessed in this thesis show a significantly higher diversity in their faecal microbiota compared to carnivores, regardless of the measurement. Figure 7 shows that both, the number of different taxa and the phylogenetic range of these taxa, are significantly increased. This is mainly due to the need for different plant-degrading microorganisms and a complex structure of the digestive system that is adapted to microbial fermentation (Vital et al., 2015; Guo et al., 2020).

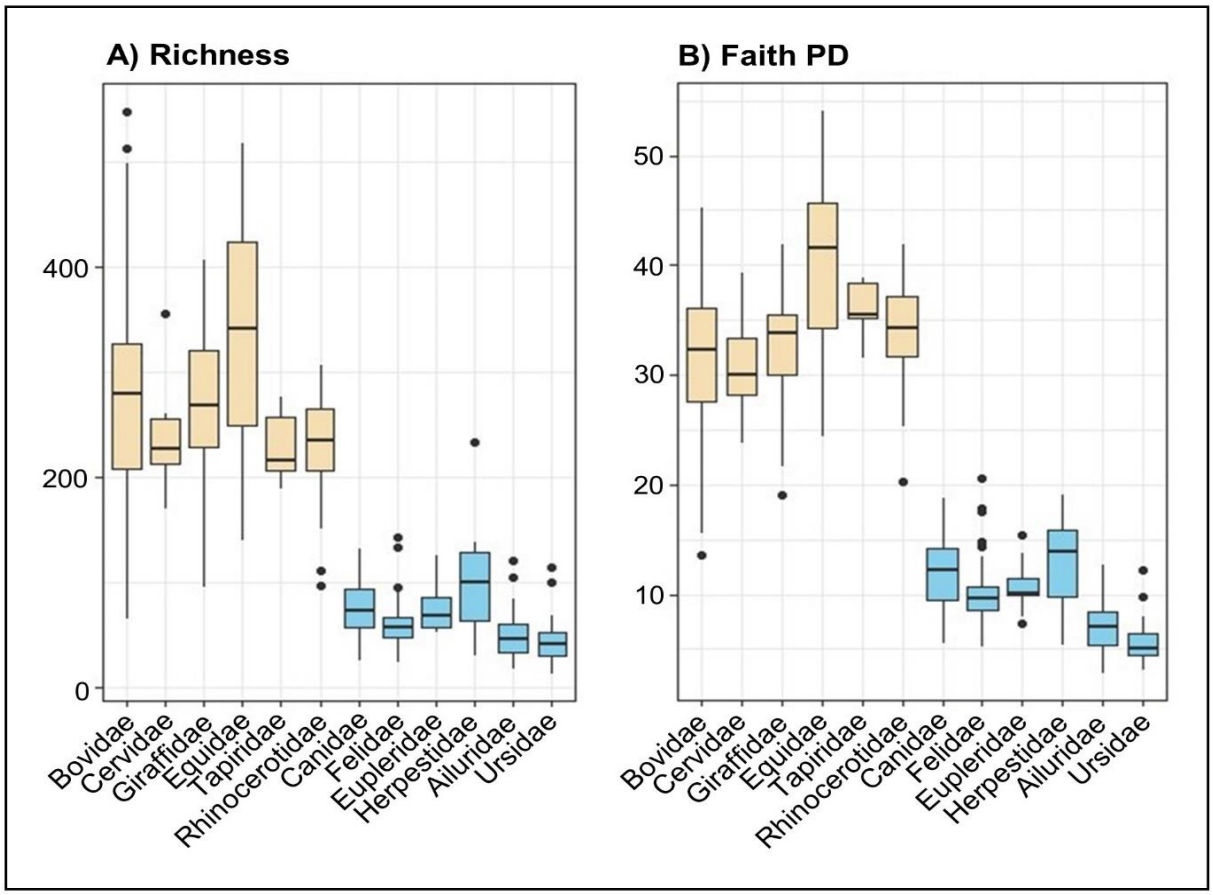

Figure 7: Alpha diversity of carnivores and herbivores. The left plot shows the ASV richness, indicating the number of different ASV's. On the right, Faith's phylogenetic diversity is shown, which includes the amount of the microbial phylogenetic tree being covered by the ASV's. Figure created according to Zoelzer et al. 2021.

Furthermore, this thesis shows a greater similarity within the herbivore microbiota in contrast to a highly diverse carnivore microbiota (Figure 8). As already discussed, previous studies have shown a clustering of herbivore and carnivore host species, but the extent and distribution of these clusters have not been the focus of research to date. The large dispersion of data points within the carnivores is further reinforced by higher coefficients of variation for all microbial taxa analysed in this group. In contrast, the coefficient of variation within herbivores is rather low. The close similarity of herbivore species has been demonstrated for other species such as *Giraffa camelopardalis* (AlZahal et al., 2016) and different antelopes (Kartzinel et al., 2019; Guo et al., 2020).

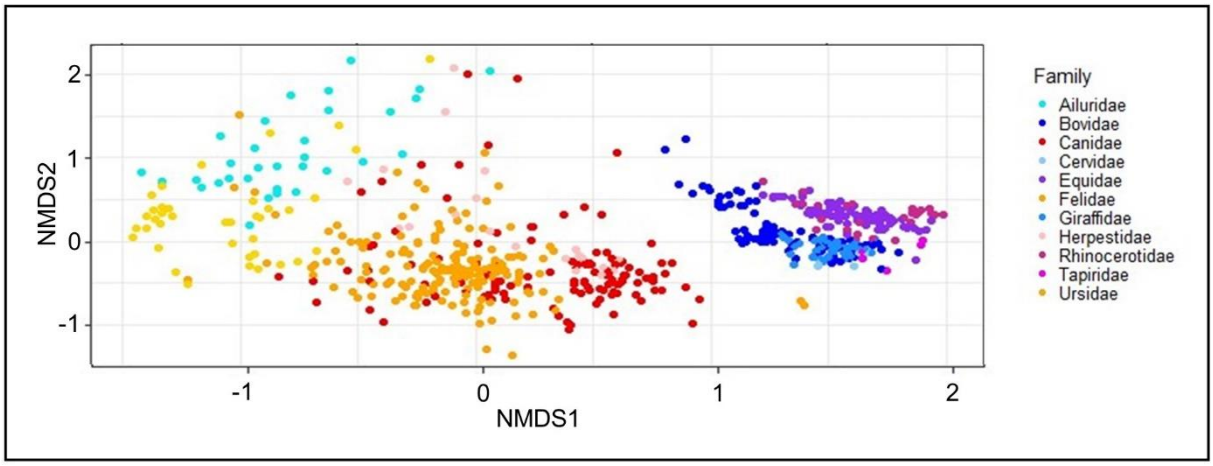

Figure 8: Beta diversity of carnivores and herbivores, expressed as non-metric dimensional scaling (NMDS) based on a Bray-Curtis distance matrix. Each point represents one sample and the distance between points indicates the similarity in the microbiota composition. The colouring refers to the host phylogeny. Figure created according to Zoelzer et al. 2021.

In addition, this thesis highlights these differences by analysing time series data from one herbivorous (*Connochaetes taurinus*) and one carnivorous (*Panthera tigris*) host species. While the different *Connochaetes taurinus* individuals show a very homogenous pattern of microbiota composition, the microbiota of *Panthera tigris* shows daily fluctuations. These findings may be related to differences in diets. Herbivores are fed on hay ad libitum, with seasonal variations in alfalfa and grass across all zoos. In contrast, the diet of carnivores consists of a variety of fresh and kibbled meats, insects and vegetable or fruit supplements. Even species classified as strict carnivores such as *Speothos venaticus* or *Lycaon pictus* experience daily variations in the origin or preparation of their meat, including options such as whole body or sheared meat. In well-studied companion animals such as *Canis lupus* and *Felis silvestris*, changes in carbohydrate or protein intake affect the faecal microbiota and the production of SCFAs. In particular, *Clostridium* and *Fusobacterium*, both members of highly abundant microbial families are known to be affected by dietary changes (Bermingham et al., 2017; Lyu et al., 2018; Bragg et al., 2020). The abundance of these taxa increase when the diet consists mainly of fresh meat. In contrast, *Prevotella* has previously been discussed as a plant-degrading bacterium that increases in a starch-containing meat-based diet (Butowski et al., 2019).

Those differences in the microbiomes of carnivores and herbivores lead to methodological aspects for future studies. Especially for carnivores, there seems to be a strong need to include multiple samples per species to compensate for the variation in the microbiota. Studies analysing the microbiota of *Acinonyx jubatus* differ in their results, e.g. for Fusobacteria, when including either two or up to 50 samples of this species (Becker et al., 2014; Wasimuddin et al., 2017).

4.3 Influence of the host on the faecal microbiota

In addition to the influence of diet and the digestive system, the host species also affects the microbiome. This interaction between the microbiome and the host organism is defined as phylosymbiosis (Lim and Bordenstein, 2020). On a broader scale, the beta diversity results provide a first insight into a possible phylosymbiosis across the analysed mammal species (Figure 8). In particular, the Ailuridae should be considered due to their unique combination of diet and digestive system. This family is characterised by a simple carnivorous digestive system, but is fed a herbivorous diet. In this thesis, the Ailuridae show a closer similarity to the other carnivores, especially to the Ursidae, to which they have a closer phylogenetic kinship. This shows that in some groups, the influence of phylogeny and of the digestive system's morphology outweighs the influence of the diet. This was previously suggested in a study with a smaller sample size and is being confirmed in this thesis (Ley et al., 2008a). To further test the host-specific influence on the faecal microbiota, a predictive logistic regression model was developed. With this model it is possible to extract host-specific information from the microbiota composition of a faecal sample. In particular, the diet type (carnivore, herbivore, omnivore) and the digestive system (ruminant, hindgut, simple) can be predicted with a high accuracy of up to 100%. This further strengthens the influence of the host's diet and digestive system on the faecal microbiota. Furthermore, it is possible to predict carnivorous host families based on the composition of the microbiota. While a distinct clustering of diet groups and digestive systems has already been described in a variety of studies (Ley et al., 2008a; Ley et al., 2008b; O' Donnell et al., 2017), a differentiation between carnivorous host species is largely unexplored mainly due to the highly variable microbiota within these animal species (Ley et al., 2008a; Guo et al., 2020). With the developed regression model, it is possible to distinguish between Canidae (F1=0.93), Felidae (F1=0.82) and Ursidae host families. This model can easily be improved for future studies, as the differentiation of herbivore host families currently fails due to an insufficient number of samples. By increasing the number of samples to 50 per host species, it will be possible to determine not only the host family but also the species. Phylosymbiosis has previously been described in various rodents (Knowles et al., 2019), Cervinae (Li et al., 2018) and *Bos mutus* (Fu et al., 2021), and throughout this thesis it is also proven for Canidae, Felidae and Ursidae host families. In addition, first evidences occur that phylosymbiosis is not restricted to the family or even species level. Analyses of the longitudinal data from single individuals of *Connochaetes taurinus* and *Panthera tigris* in this thesis show that it is possible to identify species and even individuals based on a time series of microbiota data in both species.

The combination of logistic regression models and clustering approaches of individual time series data is very promising for future studies to extract host-specific metadata from a faecal sample. Some studies were successful in predicting other host-specific factors such as age (Biagi et al., 2012; Yatsunenko et al., 2012; Björk et al., 2019) or health-status (Greenblum et al., 2012; Gupta et al., 2020) from faecal samples, mainly based on humans or primates. On the one hand, further development of this method can lead to non-invasive monitoring of the health status of a wide range of animals, thereby improving animal welfare in zoos. On the other hand, this approach opens up new possibilities of microbiome research in the field. The combination of portable sequencing technologies such as the MinION™ and the ability to extract host-specific metadata from faecal samples in a cost- and labour-efficient approach facilitates non-invasive wildlife monitoring.

5. Conclusion and future implementations

This thesis provides an important contribution to the field of animal microbiome research, including species conservation efforts. A wide range of knowledge is gained on a variety of zoo-housed animal species, with a focus on mammalian carnivores and herbivores. It was possible to collect sufficient data containing information on the composition and diversity of the microbiota of 31 mammal species, based on over 600 faecal samples. With this information it is possible to fill the knowledge gaps, as previous studies either focused on a singular species (mainly farm animals) with many samples (Clemmons et al., 2019; Wang et al., 2019; O'Hara et al., 2020), or on many species with only a few samples per species (Ley et al., 2008a; Kartzinel et al., 2019; Youngblut et al., 2019; Milani et al., 2020).

The main differences between carnivores and herbivores lie in a significantly higher microbial diversity but more homogenous microbial composition in herbivores, as opposed to low diversity and high variability in carnivores. This thesis finds strong evidences for diet, digestive morphology and host phylogeny being the most important factors influencing the faecal microbiota. Overall, highly-abundant bacterial taxa are responsible for plant or meat digestion, regardless of whether herbivores or carnivores are considered. In contrast, low abundance taxa appear to be individual-related and are therefore important to consider for individual discrimination. This thesis also provides new insights into best practice sampling methods for future microbiome research, depending on the research objective. It is shown that longitudinal data is necessary for a reliable representation of an individual's faecal microbiota, especially in carnivores. A minimum of ten samples is recommended for this approach, but ideally this

sampling should be carried out several times a year to compensate for possible seasonal differences. In contrast, to establish a reliable database of the microbiota of different host species a minimum of 50 samples per species is recommended. This is necessary to develop predictive, species-specific regression models or to extract further host-related information in future studies.

In addition, new methods are proposed that are important for future research. Combining clustering algorithms with logistic regression models provides the possibility to extend the microbiome research to wild-living animals. In particular, the identification of species or individuals benefits from these approaches. It is easy to collect faecal samples in the field, e.g. at waterholes or in grazing areas, but obtaining host-specific information is often challenging, especially in a non-invasive way. It is common practice to apply specific microsatellites to the host DNA, which needs to be isolated from the faeces. Unfortunately, this is very costly and time-consuming, especially if the microsatellites still have to be established (Kurose et al., 2005; Miller et al., 2016; Walker et al., 2019). Therefore, the combination of clustering and modelling approaches to extract as much host-specific information as possible from the microbiota is very promising. Possible applications in the field include wildlife monitoring, analysis of mixed species group composition or disease detection from microbiota samples. For the latter, the established database needs to be expanded to include samples from wild species. A common criticism of microbiome analyses of zoo animals is the influence of captivity, which can lead to a reduced diversity (Chi et al., 2019; Jiang et al., 2020). However, this has not been conclusively proven and only seems to be true for some animal families (McKenzie et al., 2017). Nevertheless, this is an important point to consider for future research. With the dataset established in this study, it is possible to compare samples from wild living animals with zoo-housed animals. The knowledge of differences or similarities between the microbiota of zoo-housed species and their wild counterparts is particularly important in the field of reintroduction programmes. Prior to release, the microbiome of these individuals should be able to cope with the new environment and especially the nutritional requirements. This is more important than ever in times of climate change and ongoing habitat fragmentation.

Future studies will benefit not only from the latest research, but also from technical advances in new sequencing methods. As this thesis is based on second-generation sequencing, third-generation sequencing methods are now well established and offer reliable and error-tolerant advancement. In particular, the MinION™ sequencer, established by Oxford Nanopore Technology. This sequencing method, which is based on ionic voltage differences within a

nanopore, runs on a device with the size of a USB stick. The company's goal is to make sequencing available to anyone, anywhere. Therefore, the entire library preparation requires very basic laboratory equipment and the sequencing can be done offline, making it a perfect tool for sequencing DNA directly in the field. This technology, combined with the knowledge of sampling procedures and the bioinformatical tools, is perfectly suited to apply the results of this thesis to wildlife faecal samples.

6. References

Alfano, N., Courtiol, A., Vielgrader, H., Timms, P., Roca, A. L., and Greenwood, A. D. (2015). Variation in koala microbiomes within and between individuals: effect of body region and captivity status. *Sci Rep* 5, 10189. doi: 10.1038/srep10189

AlZahal, O., Valdes, E. V., and McBride, B. W. (2016). Analysis of the distal gut bacterial community by 454-pyrosequencing in captive giraffes (Giraffa camelopardalis). *Zoo Biol* 35, 42–50. doi: 10.1002/zoo.21252

Amin, N., and Seifert, J. (2021). Dynamic progression of the calf's microbiome and its influence on host health. *Comput Struct Biotechnol J* 19, 989–1001. doi: 10.1016/j.csbj.2021.01.035

Antwis, R. E., Edwards, K. L., Unwin, B., Walker, S. L., and Shultz, S. (2019). Rare gut microbiota associated with breeding success, hormone metabolites and ovarian cycle phase in the critically endangered eastern black rhino. *Microbiome* 7, 27. doi: 10.1186/s40168-019-0639-0

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* 75, 129–137. doi: 10.3354/ame01753

Arora, K., Green, M., and Prakash, S. (2020). The Microbiome and Alzheimer's Disease: Potential and Limitations of Prebiotic, Synbiotic, and Probiotic Formulations. *Front Bioeng Biotechnol* 8, 537847. doi: 10.3389/fbioe.2020.537847

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944

Ayoub, C., Arroyo, L. G., MacNicol, J. L., Renaud, D., Weese, J. S., and Gomez, D. E. (2022). Fecal microbiota of horses with colitis and its association with laminitis and survival during hospitalization. *J Vet Intern Med* 36, 2213–2223. doi: 10.1111/jvim.16562

Bahl, M. I., Bergström, A., and Licht, T. R. (2012). Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. *FEMS Microbiol Lett* 329, 193–197. doi: 10.1111/j.1574-6968.2012.02523.x

Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R., and Wolfe, R. S. (1979). Methanogens: reevaluation of a unique biological group. *Microbiol Rev* 43, 260–296. doi: 10.1128/mr.43.2.260-296.1979

Barelli, C., Albanese, D., Donati, C., Pindo, M., Dallago, C., Rovero, F., et al. (2015). Habitat fragmentation is associated to gut microbiota diversity of an endangered primate: implications for conservation. *Sci Rep* 5, 14862. doi: 10.1038/srep14862

Barrett, E., Ross, R. P., O'Toole, P. W., Fitzgerald, G. F., and Stanton, C. (2012). γ-Aminobutyric acid production by culturable bacteria from the human intestine. *J Appl Microbiol* 113, 411–417. doi: 10.1111/j.1365-2672.2012.05344.x

Becker, A. A. M. J., Hesta, M., Hollants, J., Janssens, G. P. J., and Huys, G. (2014). Phylogenetic analysis of faecal microbiota from captive cheetahs reveals underrepresentation of Bacteroidetes and Bifidobacteriaceae. *BMC Microbiol* 14, 43. doi: 10.1186/1471-2180-14-43

Berendsen, R. L., Pieterse, C. M. J., and Bakker, P. A. H. M. (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci* 17, 478–486. doi: 10.1016/j.tplants.2012.04.001

Bermingham, E. N., Maclean, P., Thomas, D. G., Cave, N. J., and Young, W. (2017). Key bacterial families (Clostridiaceae, Erysipelotrichaceae and Bacteroidaceae) are related to the digestion of protein and energy in dogs. *PeerJ* 5, e3019. doi: 10.7717/peerj.3019

Biagi, E., Candela, M., Fairweather-Tait, S., Franceschi, C., and Brigidi, P. (2012). Aging of the human metaorganism: the microbial counterpart. *Age (Dordr)* 34, 247–267. doi: 10.1007/s11357-011-9217-5

Bird, J. J., Brown, D. R., Mullen, A. C., Moskowitz, N. H., Mahowald, M. A., Sider, J. R., et al. (1998). Helper T cell differentiation is controlled by the cell cycle. *Immunity* 9, 229–237. doi: 10.1016/s1074-7613(00)80605-6

Björk, J. R., Dasari, M., Grieneisen, L., and Archie, E. A. (2019). Primate microbiomes over time: Longitudinal answers to standing questions in microbiome research. *Am J Primatol* 81, e22970. doi: 10.1002/ajp.22970

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6, 90. doi: 10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Bragg, M., Freeman, E. W., Lim, H. C., Songsasen, N., and Muletz-Wolz, C. R. (2020). Gut Microbiomes Differ Among Dietary Types and Stool Consistency in the Captive Red Wolf (Canis rufus). *Front. Microbiol.* 11, 590212. doi: 10.3389/fmicb.2020.590212

Bram, R. J., Young, R. A., and Steitz, J. A. (1980). The ribonuclease III site flanking 23S sequences in the 30S ribosomal precursor RNA of E. coli. *Cell* 19, 393–401. doi: 10.1016/0092-8674(80)90513-9

Brice, K. L., Trivedi, P., Jeffries, T. C., Blyton, M. D. J., Mitchell, C., Singh, B. K., et al. (2019). The Koala (Phascolarctos cinereus) faecal microbiome differs with diet in a wild population. *PeerJ* 7, e6534. doi: 10.7717/peerj.6534

Brosius, J., Palmer, M. L., Kennedy, P. J., and Noller, H. F. (1978). Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. *Proc Natl Acad Sci U S A* 75, 4801–4805. doi: 10.1073/pnas.75.10.4801

Brumfield, K. D., Leddy, M., Usmani, M., Cotruvo, J. A., Tien, C.-T., Dorsey, S., et al. (2022). Microbiome Analysis for Wastewater Surveillance during COVID-19. *mBio* 13, e0059122. doi: 10.1128/mbio.00591-22

Buelow, E., Bayjanov, J. R., Majoor, E., Willems, R. J., Bonten, M. J., Schmitt, H., et al. (2018). Limited influence of hospital wastewater on the microbiome and resistome of wastewater in a community sewerage system. *FEMS Microbiol Ecol* 94. doi: 10.1093/femsec/fiy087

Buffington, C. T. (2008). Dry foods and risk of disease in cats. *Can Vet J* 49, 561–563.

Bukin, Y. S., Galachyants, Y. P., Morozov, I. V., Bukin, S. V., Zakharenko, A. S., and Zemskaya, T. I. (2019). The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Sci Data* 6, 190007. doi: 10.1038/sdata.2019.7

Bulgarelli, D., Garrido-Oter, R., Münch, P. C., Weiman, A., Dröge, J., Pan, Y., et al. (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17, 392–403. doi: 10.1016/j.chom.2015.01.011

Burnham, C. M., Ange-van Heugten, K., McKenney, E. A., Minter, L. J., and Trivedi, S. (2023). Conservation Innovations and Future Directions for the Study of Rhinoceros Gut Microbiome. *JZBG* 4, 396–412. doi: 10.3390/jzbg4020030

Butowski, C. F., Thomas, D. G., Young, W., Cave, N. J., McKenzie, C. M., Rosendale, D. I., et al. (2019). Addition of plant dietary fibre to a raw red meat high protein, high fat diet, alters the faecal bacteriome and organic acid profiles of the domestic cat (Felis catus). *PLoS One* 14, e0216072. doi: 10.1371/journal.pone.0216072

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6, 1621–1624. doi: 10.1038/ismej.2012.8

Cardona, S., Eck, A., Cassellas, M., Gallart, M., Alastrue, C., Dore, J., et al. (2012). Storage conditions of intestinal microbiota matter in metagenomic analysis. *BMC Microbiol* 12, 158. doi: 10.1186/1471-2180-12-158

Carroll, I. M., Ringel-Kulka, T., Siddle, J. P., Klaenhammer, T. R., and Ringel, Y. (2012). Characterization of the fecal microbiota using high-throughput sequencing reveals a stable microbial community during storage. *PLoS One* 7, e46953. doi: 10.1371/journal.pone.0046953

Cheng, Y., Fox, S., Pemberton, D., Hogg, C., Papenfuss, A. T., and Belov, K. (2015). The Tasmanian devil microbiome-implications for conservation and management. *Microbiome* 3, 76. doi: 10.1186/s40168-015-0143-0

Chi, X., Gao, H., Wu, G., Qin, W., Song, P., Wang, L., et al. (2019). Comparison of gut microbiota diversity between wild and captive bharals (Pseudois nayaur). *BMC Vet Res* 15, 243. doi: 10.1186/s12917-019-1993-7

Choct, M. (2015). Feed non-starch polysaccharides for monogastric animals: classification and function. *Animal Production Science* 55, 1360. doi: 10.1071/AN15276

Choo, J. M., Leong, L. E. X., and Rogers, G. B. (2015). Sample storage conditions significantly influence faecal microbiome profiles. *Sci Rep* 5, 16350. doi: 10.1038/srep16350

Clemmons, B. A., Martino, C., Schneider, L. G., Lefler, J., Embree, M. M., and Myer, P. R. (2019). Temporal Stability of the Ruminal Bacterial Communities in Beef Steers. *Sci Rep* 9, 9522. doi: 10.1038/s41598-019-45995-2

Crompton, A. W. (1980). "Biology of the earliest mammals," in *Comparative physiology: primitive mammals*, eds. K. Schmidt-Nielsen, L. Bolis, and C. R. Taylor (Cambridge: Cambridge Univ. Press), 1–12.

Cummings, J. H. (1981). Short chain fatty acids in the human colon. *Gut* 22, 763–779.

Cwyk, W. M., and Canale-Parola, E. (1979). Treponema succinifaciens sp. nov., an anaerobic spirochete from the swine intestine. *Arch Microbiol* 122, 231–239. doi: 10.1007/BF00411285

Dahlstrand Rudin, A., Khamzeh, A., Venkatakrishnan, V., Basic, A., Christenson, K., and Bylund, J. (2021). Short chain fatty acids released by Fusobacterium nucleatum are neutrophil chemoattractants acting via free fatty acid receptor 2 (FFAR2). *Cell Microbiol* 23, e13348. doi: 10.1111/cmi.13348

Daubin, V., Moran, N. A., and Ochman, H. (2003). Phylogenetics and the cohesion of bacterial genomes. *Science* 301, 829–832. doi: 10.1126/science.1086568

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. doi: 10.1038/nature12820

DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F., and Rosenberg, E., eds (2014). *The Prokaryotes: Firmicutes and Tenericutes*. Berlin, Heidelberg: Springer.

Delzenne, N. M., and Cani, P. D. (2011). Interaction between obesity and the gut microbiota: relevance in nutrition. *Annu Rev Nutr* 31, 15–31. doi: 10.1146/annurev-nutr-072610-145146

Derrien, M., Alvarez, A.-S., and Vos, W. M. de (2019). The Gut Microbiota in the First Decade of Life. *Trends Microbiol* 27, 997–1010. doi: 10.1016/j.tim.2019.08.001

Dey, R., Pal, K. K., and Tilak, K. V. B. R. (2012). Influence of Soil and Plant Types on Diversity of Rhizobacteria. *Proc. Natl. Acad. Sci., India, Sect. B Biol. Sci.* 82, 341–352. doi: 10.1007/s40011-012-0030-4

D'Hondt, K., Kostic, T., McDowell, R., Eudes, F., Singh, B. K., Sarkar, S., et al. (2021). Microbiome innovations for a sustainable future. *Nat Microbiol* 6, 138–142. doi: 10.1038/s41564-020-00857-w

Díaz, S., Settele, J., Brondízio, E.S., Ngo, H. T., Guèze, M., Agard, J., et al. (2019). *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.*

Difford, G. F., Plichta, D. R., Løvendahl, P., Lassen, J., Noel, S. J., Højberg, O., et al. (2018). Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS Genet* 14, e1007580. doi: 10.1371/journal.pgen.1007580

Dougal, K., La Fuente, G. de, Harris, P. A., Girdwood, S. E., Pinloche, E., Geor, R. J., et al. (2014). Characterisation of the faecal bacterial community in adult and elderly horses fed a high fibre, high oil or high starch diet using 454 pyrosequencing. *PLoS One* 9, e87424. doi: 10.1371/journal.pone.0087424

Douglas, A. E. (2018). *Fundamentals of Microbiome Science: How Microbes Shape Animal Biology*. Princeton University Press.

Drautz-Moses, D. I., Luhung, I., Gusareva, E. S., Kee, C., Gaultier, N. E., Premkrishnan, B. N. V., et al. (2022). Vertical stratification of the air microbiome in the lower troposphere. *Proc Natl Acad Sci U S A* 119. doi: 10.1073/pnas.2117293119

Du, W., Deng, J., Yang, Z., Zeng, L., and Yang, X. (2020). Metagenomic analysis reveals linkages between cecal microbiota and feed efficiency in Xiayan chickens. *Poult Sci* 99, 7066–7075. doi: 10.1016/j.psj.2020.09.076

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics* 6, 241–252. doi: 10.1080/00401706.1964.10490181

Easson, C. G., and Thacker, R. W. (2014). Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front Microbiol* 5, 532. doi: 10.3389/fmicb.2014.00532

Eisenhofer, R., Helgen, K. M., and Taggart, D. (2021). Signatures of landscape and captivity in the gut microbiota of Southern Hairy-nosed Wombats (Lasiorhinus latifrons). *Anim Microbiome* 3, 4. doi: 10.1186/s42523-020-00068-y

European Association of Zoos and Aquaria (2022). *EAZA Research Standards*. Accessed December 06, 2023, https://www.eaza.net/assets/Uploads/EAZA-Documents-2022/2022-04-EAZA-Research-Standards.pdf

Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8, 175–185. doi: 10.1101/gr.8.3.175

Ezzy, A. C., Hagstrom, A. D., George, C., Hamlin, A. S., Pereg, L., Murphy, A. J., et al. (2019). Storage and handling of human faecal samples affect the gut microbiome composition: A feasibility study. *J Microbiol Methods* 164, 105668. doi: 10.1016/j.mimet.2019.105668

Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., et al. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* 2, 6. doi: 10.1186/2049-2618-2-6

Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15, 579–590. doi: 10.1038/nrmicro.2017.87

Fogel, G. B., Collins, C. R., Li, J., and Brunk, C. F. (1999). Prokaryotic Genome Size and SSU rDNA Copy Number: Estimation of Microbial Relative Abundance from a Mixed Population. *Microb Ecol* 38, 93–113. doi: 10.1007/s002489900162

Fouhy, F., Deane, J., Rea, M. C., O'Sullivan, Ó., Ross, R. P., O'Callaghan, G., et al. (2015). The effects of freezing on faecal microbiota as determined using MiSeq sequencing and culture-based investigations. *PLoS One* 10, e0119355. doi: 10.1371/journal.pone.0119355

Freeman, L. R., and Granholm, A.-C. E. (2012). Vascular changes in rat hippocampus following a high saturated fat and cholesterol diet. *J Cereb Blood Flow Metab* 32, 643–653. doi: 10.1038/jcbfm.2011.168

Froidurot, A., and Julliand, V. (2022). Cellulolytic bacteria in the large intestine of mammals. *Gut Microbes* 14, 2031694. doi: 10.1080/19490976.2022.2031694

Fu, H., Zhang, L., Fan, C., Liu, C., Li, W., Cheng, Q., et al. (2021). Environment and host species identity shape gut microbiota diversity in sympatric herbivorous mammals. *Microb Biotechnol* 14, 1300–1315. doi: 10.1111/1751-7915.13687

Gallardo-Escárate, C., Valenzuela-Muñoz, V., Núñez-Acuña, G., Valenzuela-Miranda, D., Benaventel, B. P., Sáez-Vera, C., et al. (2021). The wastewater microbiome: A novel insight for COVID-19 surveillance. *Sci Total Environ* 764, 142867. doi: 10.1016/j.scitotenv.2020.142867

Garbeva, P., van Veen, J. A., and van Elsas, J. D. (2004). Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. *Annu Rev Phytopathol* 42, 243–270. doi: 10.1146/annurev.phyto.42.012604.135455

Gavini, F., Cayuela, C., Antoine, J.-M., Lecoq, C., Lefebvre, B., Membré, J.-M., et al. (2001). Differences in the Distribution of Bifidobacterial and Enterobacterial Species in Human Faecal Microflora of Three Different (Children, Adults, Elderly) Age Groups. *Microbial Ecology in Health and Disease* 13, 40–45. doi: 10.1080/089106001750071690

Gibson, K. M., Nguyen, B. N., Neumann, L. M., Miller, M., Buss, P., Daniels, S., et al. (2019). Gut microbiome differences between wild and captive black rhinoceros - implications for rhino health. *Sci Rep* 9, 7570. doi: 10.1038/s41598-019-43875-3

Gordon, I. J., and Illius, A. W. (1994). The functional significance of the browser-grazer dichotomy in African ruminants. *Oecologia* 98, 167–175. doi: 10.1007/BF00341469

Gorman, M. L. (1976). A mechanism for individual recognition by odour in Herpestes auropunctatus (Carnivora: Viverridae). *Animal Behaviour* 24, 141–145. doi: 10.1016/S0003-3472(76)80107-8

Greenblum, S., Turnbaugh, P. J., and Borenstein, E. (2012). Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc Natl Acad Sci U S A* 109, 594–599. doi: 10.1073/pnas.1116053109

Grieneisen, L. E., Charpentier, M. J. E., Alberts, S. C., Blekhman, R., Bradburd, G., Tung, J., et al. (2019). Genes, geology and germs: gut microbiota across a primate hybrid zone are explained by site soil properties, not host species. *Proc Biol Sci* 286, 20190431. doi: 10.1098/rspb.2019.0431

Griffiths, E. A., Duffy, L. C., Schanbacher, F. L., Qiao, H., Dryja, D., Leavens, A., et al. (2004). In vivo effects of bifidobacteria and lactoferrin on gut endotoxin concentration and mucosal

immunity in Balb/c mice. *Dig Dis Sci* 49, 579–589. doi: 10.1023/b:ddas.0000026302.92898.ae

Grundy, M. M.-L., Edwards, C. H., Mackie, A. R., Gidley, M. J., Butterworth, P. J., and Ellis, P. R. (2016). Re-evaluation of the mechanisms of dietary fibre and implications for macronutrient bioaccessibility, digestion and postprandial metabolism. *Br J Nutr* 116, 816–833. doi: 10.1017/S0007114516002610

Guevarra, R. B., Lee, J. H., Lee, S. H., Seok, M.-J., Kim, D. W., Kang, B. N., et al. (2019). Piglet gut microbial shifts early in life: causes and effects. *J Anim Sci Biotechnol* 10. doi: 10.1186/s40104-018-0308-3

Gulaj, E., Pawlak, K., Bien, B., and Pawlak, D. (2010). Kynurenine and its metabolites in Alzheimer's disease patients. *Adv Med Sci* 55, 204–211. doi: 10.2478/v10039-010-0023-6

Guo, X., Lei, H., Zhang, K., Ke, F., and Song, C. (2020). Diversification of animal gut microbes and NRPS gene clusters in some carnivores, herbivores and omnivores. *Biotechnology & Biotechnological Equipment* 34, 1280–1287. doi: 10.1080/13102818.2020.1835536

Gupta, V. K., Kim, M., Bakshi, U., Cunningham, K. Y., Davis, J. M., Lazaridis, K. N., et al. (2020). A predictive index for health status using species-level gut microbiome profiling. *Nat Commun* 11, 4635. doi: 10.1038/s41467-020-18476-8

Guzman, C. E., Bereza-Malcolm, L. T., Groef, B. de, and Franks, A. E. (2016). Uptake of milk with and without solid feed during the monogastric phase: Effect on fibrolytic and methanogenic microorganisms in the gastrointestinal tract of calves. *Anim Sci J* 87, 378–388. doi: 10.1111/asj.12429

Haller, D., and Hörmannsperger, G. (2015). *Darmgesundheit und Mikrobiota*. Wiesbaden: Springer Fachmedien Wiesbaden.

Heo, J. M., Opapeju, F. O., Pluske, J. R., Kim, J. C., Hampson, D. J., and Nyachoti, C. M. (2013). Gastrointestinal health and function in weaned pigs: a review of feeding strategies to control post-weaning diarrhoea without using in-feed antimicrobial compounds. *J Anim Physiol Anim Nutr (Berl)* 97, 207–237. doi: 10.1111/j.1439-0396.2012.01284.x

Hull, N. M., Ling, F., Pinto, A. J., Albertsen, M., Jang, H. G., Hong, P.-Y., et al. (2019). Drinking Water Microbiome Project: Is it Time? *Trends Microbiol* 27, 670–677. doi: 10.1016/j.tim.2019.03.011

Illumina Inc (2010). *Illumina Sequencing Technology: Highest data accuracy, simple workflow, and a broad range of application.*, https://www.illumina.com/documents/products/techspotlights/techspotlight_sequencing.pdf

Islam, W., Noman, A., Naveed, H., Huang, Z., and Chen, H. Y. H. (2020). Role of environmental factors in shaping the soil microbiome. *Environ Sci Pollut Res Int* 27, 41225–41247. doi: 10.1007/s11356-020-10471-2

IUCN (2022). *The IUCN Red List of Threatened Species*. Accessed November 27, 2023, https://www.iucnredlist.org

Jain, R., Rivera, M. C., and Lake, J. A. (1999). Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci U S A* 96, 3801–3806. doi: 10.1073/pnas.96.7.3801

Jami, E., Israel, A., Kotser, A., and Mizrahi, I. (2013). Exploring the bovine rumen bacterial community from birth to adulthood. *ISME J* 7, 1069–1079. doi: 10.1038/ismej.2013.2

Janda, J. M., and Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 45, 2761–2764. doi: 10.1128/JCM.01228-07

Janis, C. (1976). The Evolutionary Strategy of the Equidae and the Origins of Rumen and Cecal Digestion. *Evolution* 30, 757. doi: 10.2307/2407816

Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, et al. (2012). *vegan: Community Ecology Package*.

Jiang, D., He, X., Valitutto, M., Chen, L., Xu, Q., Yao, Y., et al. (2020). Gut microbiota composition and metabolomic profiles of wild and captive Chinese monals (Lophophorus lhuysii). *Front Zool* 17, 36. doi: 10.1186/s12983-020-00381-x

Jiang, H.-Y., Zhang, X., Yu, Z.-H., Zhang, Z., Deng, M., Zhao, J.-H., et al. (2018). Altered gut microbiota profile in patients with generalized anxiety disorder. *J Psychiatr Res* 104, 130–136. doi: 10.1016/j.jpsychires.2018.07.007

Jost, L. (2006). Entropy and diversity. *Oikos* 113, 363–375. doi: 10.1111/j.2006.0030-1299.14714.x

Jukes, T. H., and Cantor, C. R. (2013). "Evolution of Protein Molecules," in *Mammalian Protein Metabolism: Volume III*, ed. H. N. Munro (Burlington: Elsevier Science), 21–132.

Kanoski, S. E., Zhang, Y., Zheng, W., and Davidson, T. L. (2010). The effects of a high-energy diet on hippocampal function and blood-brain barrier integrity in the rat. *J Alzheimers Dis* 21, 207–219. doi: 10.3233/JAD-2010-091414

Kartzinel, T. R., Hsing, J. C., Musili, P. M., Brown, B. R. P., and Pringle, R. M. (2019). Covariation of diet and gut microbiome in African megafauna. *Proc Natl Acad Sci U S A* 116, 23588–23593. doi: 10.1073/pnas.1905666116

Katoh, K., Misawa, K., Kuma, K.-i., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30, 3059–3066. doi: 10.1093/nar/gkf436

Kim, M., Morrison, M., and Yu, Z. (2011). Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods* 84, 81–87. doi: 10.1016/j.mimet.2010.10.020

Klappenbach, J. A., Saxman, P. R., Cole, J. R., and Schmidt, T. M. (2001). rrndb: the Ribosomal RNA Operon Copy Number Database. *Nucleic Acids Res* 29, 181–184. doi: 10.1093/nar/29.1.181

Knowles, S. C. L., Eccles, R. M., and Baltrūnaitė, L. (2019). Species identity dominates over environment in shaping the microbiota of small mammals. *Ecol Lett* 22, 826–837. doi: 10.1111/ele.13240

Kong, F., Zhao, J., Han, S., Zeng, B., Yang, J., Si, X., et al. (2014). Characterization of the gut microbiota in the red panda (Ailurus fulgens). *PLoS One* 9, e87885. doi: 10.1371/journal.pone.0087885

Kruskal, W. H., and Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association* 47, 583–621. doi: 10.1080/01621459.1952.10483441

Kumar, S., Suleski, M., Craig, J. M., Kasprowicz, A. E., Sanderford, M., Li, M., et al. (2022). TimeTree 5: An Expanded Resource for Species Divergence Times. *Mol Biol Evol* 39. doi: 10.1093/molbev/msac174

Kurose, N., Masuda, R., and Tatara, M. (2005). Fecal DNA analysis for identifying species and sex of sympatric carnivores: a noninvasive method for conservation on the Tsushima Islands, Japan. *J Hered* 96, 688–697. doi: 10.1093/jhered/esi124

La Reau, A. J., and Suen, G. (2018). The Ruminococci: key symbionts of the gut ecosystem. *J Microbiol* 56, 199–208. doi: 10.1007/s12275-018-8024-4

Lach, G., Schellekens, H., Dinan, T. G., and Cryan, J. F. (2018). Anxiety, Depression, and the Microbiome: A Role for Gut Peptides. *Neurotherapeutics* 15, 36–59. doi: 10.1007/s13311-017-0585-0

Lallès, J.-P., Bosi, P., Smidt, H., and Stokes, C. R. (2007). Nutritional management of gut health in pigs around weaning. *Proc Nutr Soc* 66, 260–268. doi: 10.1017/S0029665107005484

Lam, Y. Y., Ha, C. W. Y., Campbell, C. R., Mitchell, A. J., Dinudom, A., Oscarsson, J., et al. (2012). Increased gut permeability and microbiota change associate with mesenteric fat

inflammation and metabolic dysfunction in diet-induced obese mice. *PLoS One* 7, e34233. doi: 10.1371/journal.pone.0034233

Lane-Petter, W. (1962). The Provision and Use of Pathogen-free Laboratory Animals. *Proc R Soc Med* 55, 253–263.

Lau, S.K.P., Teng, J.L.L., Ho, T., Chiu, Elaine, C., Alan K.L., T., et al. (2018). Differential Microbial Communities of Omnivorous and Herbivorous Cattle in Southern China. *Comput Struct Biotechnol J* 16, 54. doi: 10.1016/j.csbj.2018.02.004

Leclaire, S., Nielsen, J. F., and Drea, C. M. (2014). Bacterial communities in meerkat anal scent secretions vary with host sex, age, and group membership. *Behavioral Ecology* 25, 996–1004. doi: 10.1093/beheco/aru074

Lederberg, J., and McCray, A. T. (2001). Ome sweet omics: a genealogical treasury of words. *Scientist*, 8.

Leroy-Freitas, D., Machado, E. C., Torres-Franco, A. F., Dias, M. F., Leal, C. D., and Araújo, J. C. (2022). Exploring the microbiome, antibiotic resistance genes, mobile genetic element, and potential resistant pathogens in municipal wastewater treatment plants in Brazil. *Sci Total Environ* 842, 156773. doi: 10.1016/j.scitotenv.2022.156773

Leung, M. H. Y., Wilkins, D., Li, E. K. T., Kong, F. K. F., and Lee, P. K. H. (2014). Indoor-air microbiome in an urban subway network: diversity and dynamics. *Appl Environ Microbiol* 80, 6760–6770. doi: 10.1128/AEM.02244-14

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., et al. (2008a). Evolution of mammals and their gut microbes. *Science* 320, 1647–1651. doi: 10.1126/science.1155725

Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R., and Gordon, J. I. (2008b). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* 6, 776–788. doi: 10.1038/nrmicro1978

Li, J., Zhan, S., Liu, X., Lin, Q., Jiang, J., and Li, X. (2018). Divergence of Fecal Microbiota and Their Associations With Host Phylogeny in Cervinae. *Front Microbiol* 9, 1823. doi: 10.3389/fmicb.2018.01823

Lillegraven, J. A., Kraus, M. J., and Brown, T. M. (1979). "Paleogeography of the world of the Mesozoik," in *Mesozoic mammals: The first two-thirds of mammalian history*, ed. J. A. Lillegraven (Berkeley: Univ. of California Press), 277–308.

Lim, S. J., and Bordenstein, S. R. (2020). An introduction to phylosymbiosis. *Proc Biol Sci* 287, 20192900. doi: 10.1098/rspb.2019.2900

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71, 8228–8235. doi: 10.1128/AEM.71.12.8228-8235.2005

Lozupone, C. A., Hamady, M., Kelley, S. T., and Knight, R. (2007). Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Appl Environ Microbiol* 73, 1576–1585. doi: 10.1128/AEM.01996-06

Lu, Y., Shen, Y., Warren, W., and Walter, R. (2016). "Next Generation Sequencing in Aquatic Models," in *Next Generation Sequencing - Advances, Applications and Challenges*, ed. J. K. Kulski (Erscheinungsort nicht ermittelbar: InTechOpen).

Lucke, K., Miehlke, S., Jacobs, E., and Schuppler, M. (2006). Prevalence of Bacteroides and Prevotella spp. in ulcerative colitis. *J Med Microbiol* 55, 617–624. doi: 10.1099/jmm.0.46198-0

Lyte, M. (2011). Probiotics function mechanistically as delivery vehicles for neuroactive compounds: Microbial endocrinology in the design and use of probiotics. *Bioessays* 33, 574–581. doi: 10.1002/bies.201100024

Lyu, T., Liu, G., Zhang, H., Wang, L., Zhou, S., Dou, H., et al. (2018). Changes in feeding habits promoted the differentiation of the composition and function of gut microbiotas between domestic dogs (Canis lupus familiaris) and gray wolves (Canis lupus). *AMB Express* 8, 123. doi: 10.1186/s13568-018-0652-x

Mackie, R. I. (2002). Mutualistic fermentative digestion in the gastrointestinal tract: diversity and evolution. *Integr Comp Biol* 42, 319–326. doi: 10.1093/icb/42.2.319

Madden, A. A., Oliverio, A. M., Kearns, P. J., Henley, J. B., Fierer, N., Starks, P. T. B., et al. (2022). Chronic stress and captivity alter the cloacal microbiome of a wild songbird. *J Exp Biol* 225. doi: 10.1242/jeb.243176

Maffeis, C., Martina, A., Corradi, M., Quarella, S., Nori, N., Torriani, S., et al. (2016). Association between intestinal permeability and faecal microbiota composition in Italian children with beta cell autoimmunity at risk for type 1 diabetes. *Diabetes Metab Res Rev* 32, 700–709. doi: 10.1002/dmrr.2790

Maki, J. J., Klima, C. L., Sylte, M. J., and Looft, T. (2019). The Microbial Pecking Order: Utilization of Intestinal Microbiota for Poultry Health. *Microorganisms* 7. doi: 10.3390/microorganisms7100376

Malan-Muller, S., Valles-Colomer, M., Foxx, C. L., Vieira-Silva, S., van den Heuvel, L. L., Raes, J., et al. (2022). Exploring the relationship between the gut microbiome and mental

health outcomes in a posttraumatic stress disorder cohort relative to trauma-exposed controls. *Eur Neuropsychopharmacol* 56, 24–38. doi: 10.1016/j.euroneuro.2021.11.009

Mamun, M. A. A., Sandeman, M., Rayment, P., Brook-Carter, P., Scholes, E., Kasinadhuni, N., et al. (2020). Variation in gut bacterial composition is associated with Haemonchus contortus parasite infection of sheep. *Anim Microbiome* 2, 3. doi: 10.1186/s42523-020-0021-3

Mariat, D., Firmesse, O., Levenez, F., Guimarães, V., Sokol, H., Doré, J., et al. (2009). The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiol* 9, 123. doi: 10.1186/1471-2180-9-123

Marmion, M., Ferone, M. T., Whyte, P., and Scannell, A. G. M. (2021). The changing microbiome of poultry meat; from farm to fridge. *Food Microbiol* 99, 103823. doi: 10.1016/j.fm.2021.103823

Maslowski, K. M., Vieira, A. T., Ng, A., Kranich, J., Sierro, F., Di Yu, et al. (2009). Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* 461, 1282–1286. doi: 10.1038/nature08530

Mason, G. J. (2010). Species differences in responses to captivity: stress, welfare and the comparative method. *Trends Ecol Evol* 25, 713–721. doi: 10.1016/j.tree.2010.08.011

Matson, P. A., Parton, W. J., Power, A. G., and Swift, M. J. (1997). Agricultural intensification and ecosystem properties. *Science* 277, 504–509. doi: 10.1126/science.277.5325.504

McKenzie, V. J., Song, S. J., Delsuc, F., Prest, T. L., Oliverio, A. M., Korpita, T. M., et al. (2017). The Effects of Captivity on the Mammalian Gut Microbiome. *Integr Comp Biol* 57, 690–704. doi: 10.1093/icb/icx090

Menke, S., Meier, M., and Sommer, S. (2015). Shifts in the gut microbiome observed in wildlife faecal samples exposed to natural weather conditions: lessons from time-series analyses using next-generation sequencing for application in field studies. *Methods Ecol Evol* 6, 1080–1087. doi: 10.1111/2041-210X.12394

Metcalf, J. L., Song, S. J., Morton, J. T., Weiss, S., Seguin-Orlando, A., Joly, F., et al. (2017). Evaluating the impact of domestication and captivity on the horse gut microbiome. *Sci Rep* 7, 15497. doi: 10.1038/s41598-017-15375-9

Milani, C., Alessandri, G., Mancabelli, L., Mangifesta, M., Lugli, G. A., Viappiani, A., et al. (2020). Multi-omics Approaches To Decipher the Impact of Diet and Host Physiology on the Mammalian Gut Microbiome. *Appl Environ Microbiol* 86. doi: 10.1128/AEM.01864-20

Miller, S. M., Clarke, A. B., Bloomer, P., Guthrie, A. J., and Harper, C. K. (2016). Evaluation of microsatellites for common ungulates in the South African wildlife industry. *Conservation Genet Resour* 8, 329–341. doi: 10.1007/s12686-016-0554-7

Mitchell, P. C. (1905). On the Intestinal Tract of Mammals. *The Transactions of the Zoological Society of London* 17, 437–536. doi: 10.1111/j.1096-3642.1905.tb00036.x

Moeller, A. H., Suzuki, T. A., Lin, D., Lacey, E. A., Wasser, S. K., and Nachman, M. W. (2017). Dispersal limitation promotes the diversification of the mammalian gut microbiota. *Proc Natl Acad Sci U S A* 114, 13768–13773. doi: 10.1073/pnas.1700122114

Morgan, K. N., and Tromborg, C. T. (2007). Sources of stress in captivity. *Applied Animal Behaviour Science* 102, 262–302. doi: 10.1016/j.applanim.2006.05.032

Morgavi, D. P., Forano, E., Martin, C., and Newbold, C. J. (2010). Microbial ecosystem and methanogenesis in ruminants. *Animal* 4, 1024–1036. doi: 10.1017/S1751731110000546

Morosyuk, S. V., Cunningham, P. R., and SantaLucia, J. (2001). Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA. II. NMR solution structure. *J Mol Biol* 307, 197–211. doi: 10.1006/jmbi.2000.4431

Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., González, A., Fontana, L., et al. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 332, 970–974. doi: 10.1126/science.1198719

Needleman, S. B., and Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 48, 443–453. doi: 10.1016/0022-2836(70)90057-4

Nelson, T. M., Rogers, T. L., Carlini, A. R., and Brown, M. V. (2013). Diet and phylogeny shape the gut microbiota of Antarctic seals: a comparison of wild and captive animals. *Environ Microbiol* 15, 1132–1145. doi: 10.1111/1462-2920.12022

Newman, M. E. J. (2004). Detecting community structure in networks. *The European Physical Journal B - Condensed Matter* 38, 321–330. doi: 10.1140/epjb/e2004-00124-y

Nishida, A. H., and Ochman, H. (2018). Rates of gut microbiome divergence in mammals. *Mol Ecol* 27, 1884–1897. doi: 10.1111/mec.14473

O' Donnell, M. M., Harris, H. M. B., Ross, R. P., and O'Toole, P. W. (2017). Core fecal microbiota of domesticated herbivorous ruminant, hindgut fermenters, and monogastric animals. *Microbiologyopen* 6. doi: 10.1002/mbo3.509

Obregon-Tito, A. J., Tito, R. Y., Metcalf, J., Sankaranarayanan, K., Clemente, J. C., Ursell, L. K., et al. (2015). Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat Commun* 6, 6505. doi: 10.1038/ncomms7505

Ochman, H., Worobey, M., Kuo, C.-H., Ndjango, J.-B. N., Peeters, M., Hahn, B. H., et al. (2010). Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol* 8, e1000546. doi: 10.1371/journal.pbio.1000546

Ogle, D. H., Wheeler, P., and Dinno, A. (2020). *FSA: Fisheries Stock Analysis R package*.

O'Hara, E., Neves, A. L. A., Song, Y., and Le Guan, L. (2020). The Role of the Gut Microbiome in Cattle Production and Health: Driver or Passenger? *Annu Rev Anim Biosci* 8, 199–220. doi: 10.1146/annurev-animal-021419-083952

Ojo, B., El-Rassi, G. D., Payton, M. E., Perkins-Veazie, P., Clarke, S., Smith, B. J., et al. (2016). Mango Supplementation Modulates Gut Microbial Dysbiosis and Short-Chain Fatty Acid Production Independent of Body Weight Reduction in C57BL/6 Mice Fed a High-Fat Diet. *J Nutr* 146, 1483–1491. doi: 10.3945/jn.115.226688

O'Keefe, S. J. D. (2019). Plant-based foods and the microbiome in the preservation of health and prevention of disease. *Am J Clin Nutr* 110, 265–266. doi: 10.1093/ajcn/nqz127

Olsen, I. (2014). "The Family Fusobacteriaceae," in *The Prokaryotes: Firmicutes and Tenericutes*, eds. E. F. DeLong, S. Lory, E. Stackebrandt, F. Thompson, and E. Rosenberg (Berlin, Heidelberg: Springer), 109–132.

Pace, N. R. (1973). Structure and synthesis of the ribosomal ribonucleic acid of prokaryotes. *Bacteriol Rev* 37, 562–603.

Paliwal, K. K., Agarwal, A., and Sinha, S. S. (1982). A modification over Sakoe and Chiba's dynamic time warping algorithm for isolated word recognition. *Signal Processing* 4, 329–333. doi: 10.1016/0165-1684(82)90009-3

Patra, A., Park, T., Kim, M., and Yu, Z. (2017). Rumen methanogens and mitigation of methane emission by anti-methanogenic compounds and substances. *J Anim Sci Biotechnol* 8, 13. doi: 10.1186/s40104-017-0145-9

Patterson, E., O' Doherty, R. M., Murphy, E. F., Wall, R., O' Sullivan, O., Nilaweera, K., et al. (2014). Impact of dietary fatty acids on metabolic activity and host intestinal microbiota composition in C57BL/6J mice. *Br J Nutr* 111, 1905–1917. doi: 10.1017/S0007114514000117

Peirce, J. M., and Alviña, K. (2019). The role of inflammation and the gut microbiome in depression and anxiety. *J Neurosci Res* 97, 1223–1241. doi: 10.1002/jnr.24476

Pellanda, P., Ghosh, T. S., and O'Toole, P. W. (2021). Understanding the impact of age-related changes in the gut microbiome on chronic diseases and the prospect of elderly-specific dietary interventions. *Curr Opin Biotechnol* 70, 48–55. doi: 10.1016/j.copbio.2020.11.001

Peng, G., Wang, H. Y., Peng, W., Kiniwa, Y., Seo, K. H., and Wang, R.-F. (2007). Tumor-infiltrating gammadelta T cells suppress T and dendritic cell function via mechanisms controlled by a unique toll-like receptor signaling pathway. *Immunity* 27, 334–348. doi: 10.1016/j.immuni.2007.05.020

Pichler, M., Coskun, Ö. K., Ortega-Arbulú, A.-S., Conci, N., Wörheide, G., Vargas, S., et al. (2018). A 16S rRNA gene sequencing and analysis protocol for the Illumina MiniSeq platform. *Microbiologyopen* 7, e00611. doi: 10.1002/mbo3.611

Pontes, D. S., Lima-Bittencourt, C. I., Chartone-Souza, E., and Amaral Nascimento, A. M. (2007). Molecular approaches: advantages and artifacts in assessing bacterial diversity. *J Ind Microbiol Biotechnol* 34, 463–473. doi: 10.1007/s10295-007-0219-3

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490. doi: 10.1371/journal.pone.0009490

Qiu, M., Zhang, R., Xue, C., Zhang, S., Li, S., Zhang, N., et al. (2012). Application of bio-organic fertilizer can control Fusarium wilt of cucumber plants by regulating microbial community of rhizosphere soil. *Biol Fertil Soils* 48, 807–816. doi: 10.1007/s00374-012-0675-4

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41, D590-6. doi: 10.1093/nar/gks1219

Quercia, S., Freccero, F., Castagnetti, C., Soverini, M., Turroni, S., Biagi, E., et al. (2019). Early colonisation and temporal dynamics of the gut microbial ecosystem in Standardbred foals. *Equine Vet J* 51, 231–237. doi: 10.1111/evj.12983

R core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rérat, A., Fiszlewicz, M., Giusi, A., and Vaugelade, P. (1987). Influence of meal frequency on postprandial variations in the production and absorption of volatile fatty acids in the digestive tract of conscious pigs. *J Anim Sci* 64, 448–456. doi: 10.2527/jas1987.642448x

Robeson, M. S., O'Rourke, D. R., Kaehler, B. D., Ziemski, M., Dillon, M. R., Foster, J. T., et al. (2020). *RESCRIPt: Reproducible sequence taxonomy reference database management for the masses*.

Roehe, R., Dewhurst, R. J., Duthie, C.-A., Rooke, J. A., McKain, N., Ross, D. W., et al. (2016). Bovine Host Genetic Variation Influences Rumen Microbial Methane Production with Best Selection Criterion for Low Methane Emitting and Efficiently Feed Converting Hosts Based on Metagenomic Gene Abundance. *PLoS Genet* 12, e1005846. doi: 10.1371/journal.pgen.1005846

Roesch, L. F. W., Casella, G., Simell, O., Krischer, J., Wasserfall, C. H., Schatz, D., et al. (2009). Influence of fecal sample storage on bacterial community diversity. *Open Microbiol J* 3, 40–46. doi: 10.2174/1874285800903010040

Román-Palacios, C., Scholl, J. P., and Wiens, J. J. (2019). Evolution of diet across the animal tree of life. *Evol Lett* 3, 339–347. doi: 10.1002/evl3.127

Romer, A. S. (1966). *Vertebrate Paleontology*.

Roselli, M., Pieper, R., Rogel-Gaillard, C., Vries, H. de, Bailey, M., Smidt, H., et al. (2017). Immunomodulating effects of probiotics for microbiota modulation, gut health and disease resistance in pigs. *Animal Feed Science and Technology* 233, 104–119. doi: 10.1016/j.anifeedsci.2017.07.011

Rosen, M. J., Callahan, B. J., Fisher, D. S., and Holmes, S. P. (2012). Denoising PCR-amplified metagenome data. *BMC Bioinformatics* 13, 283. doi: 10.1186/1471-2105-13-283

Rosselló-Mora, R. (2001). The species concept for prokaryotes. *FEMS Microbiology Reviews* 25, 39–67. doi: 10.1016/S0168-6445(00)00040-1

Russel, J. B., and Rychlik, J. L. (2001). Factors That Alter Rumen Microbial Ecology. *Science* 292, 1119–1122. doi: 10.1126/science.1058830

Sakoe, H., and Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. Acoust., Speech, Signal Process.* 26, 43–49. doi: 10.1109/TASSP.1978.1163055

San Juan, P. A., Castro, I., and Dhami, M. K. (2021). Captivity reduces diversity and shifts composition of the Brown Kiwi microbiome. *Anim Microbiome* 3, 48. doi: 10.1186/s42523-021-00109-0

Sanger, F., and Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 94, 441–448. doi: 10.1016/0022-2836(75)90213-2

Schirmer, M., Ijaz, U. Z., D'Amore, R., Hall, N., Sloan, W. T., and Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res* 43, e37. doi: 10.1093/nar/gku1341

Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., et al. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* 102, 615–623. doi: 10.1016/s0092-8674(00)00084-2

Schneider, S., Hammerschmidt, K., and Dierkes, P. W. (2022). Introducing the Software CASE (Cluster and Analyze Sound Events) by Comparing Different Clustering Methods and Audio Transformation Techniques Using Animal Vocalizations. *Animals (Basel)* 12. doi: 10.3390/ani12162020

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591–611. doi: 10.1093/biomet/52.3-4.591

Sin, Y. W., Buesching, C. D., Burke, T., and Macdonald, D. W. (2012). Molecular characterization of the microbial communities in the subcaudal gland secretion of the European badger (Meles meles). *FEMS Microbiol Ecol* 81, 648–659. doi: 10.1111/j.1574-6941.2012.01396.x

Sinclair, L., Osman, O. A., Bertilsson, S., and Eiler, A. (2015). Microbial community composition and diversity via 16S rRNA gene amplicons: evaluating the illumina platform. *PLoS One* 10, e0116955. doi: 10.1371/journal.pone.0116955

Slobodkin, A. (2014). "The Family Peptostreptococcaceae," in *The Prokaryotes: Firmicutes and Tenericutes*, eds. E. F. DeLong, S. Lory, E. Stackebrandt, F. Thompson, and E. Rosenberg (Berlin, Heidelberg: Springer), 291–302.

Smith, E.A., and Macfarlane, G.T. (1998). Enumeration of amino acid fermenting bacteria in the human large intestine: effects of pH and starch on peptide metabolism and dissimilation of amino acids. *FEMS Microbiol Ecol* 25, 355–368. doi: 10.1111/j.1574-6941.1998.tb00487.x

Snelling, T. J., Auffret, M. D., Duthie, C.-A., Stewart, R. D., Watson, M., Dewhurst, R. J., et al. (2019). Temporal stability of the rumen microbiota in beef cattle, and response to diet and supplements. *anim microbiome* 1, 1–14. doi: 10.1186/s42523-019-0018-y

Stanley, D., Hughes, R. J., Geier, M. S., and Moore, R. J. (2016). Bacteria within the Gastrointestinal Tract Microbiota Correlated with Improved Growth and Feed Conversion: Challenges Presented for the Identification of Performance Enhancing Probiotic Bacteria. *Front Microbiol* 7, 187. doi: 10.3389/fmicb.2016.00187

Stevens, C. E., and Hume, I. D. (1995). *Comparative physiology of the vertebrate digestive system*. Cambridge: Cambridge University Press.

Stoate, C., Boatman, N. D., Borralho, R. J., Carvalho, C. R., Snoo, G. R. de, and Eden, P. (2001). Ecological impacts of arable intensification in Europe. *J Environ Manage* 63, 337–365. doi: 10.1006/jema.2001.0473

Suchodolski, J. S., Camacho, J., and Steiner, J. M. (2008). Analysis of bacterial diversity in the canine duodenum, jejunum, ileum, and colon by comparative 16S rRNA gene analysis. *FEMS Microbiol Ecol* 66, 567–578. doi: 10.1111/j.1574-6941.2008.00521.x

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., et al. (2015). Ocean plankton. Structure and function of the global ocean microbiome. *Science* 348, 1261359. doi: 10.1126/science.1261359

Sunvold, G. D., Hussein, H. S., Fahey, G. C., Merchen, N. R., and Reinhart, G. A. (1995). In vitro fermentation of cellulose, beet pulp, citrus pulp, and citrus pectin using fecal inoculum

from cats, dogs, horses, humans, and pigs and ruminal fluid from cattle. *J Anim Sci* 73. doi: 10.2527/1995.73123639x

Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., and Nishijima, M. (2014). Development of a prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-generation sequencing. *PLoS One* 9, e105592. doi: 10.1371/journal-pone.0105592

Tanca, A., Fraumene, C., Manghina, V., Palomba, A., Abbondio, M., Deligios, M., et al. (2017). Diversity and functions of the sheep faecal microbiota: a multi-omic characterization. *Microb Biotechnol* 10, 541–554. doi: 10.1111/1751-7915.12462

Tedjo, D. I., Jonkers, D. M. A. E., Savelkoul, P. H., Masclee, A. A., van Best, N., Pierik, M. J., et al. (2015). The effect of sampling and storage on the fecal microbiota composition in healthy and diseased subjects. *PLoS One* 10, e0126685. doi: 10.1371/journal.pone.0126685

The MathWorks Inc. (2020a). *MathWorks MATLAB*.

Theis, K. R., Venkataraman, A., Dycus, J. A., Koonter, K. D., Schmitt-Matzen, E. N., Wagner, A. P., et al. (2013). Symbiotic bacteria appear to mediate hyena social odors. *Proc Natl Acad Sci U S A* 110, 19832–19837. doi: 10.1073/pnas.1306477110

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x

Trevelline, B. K., Sosa, J., Hartup, B. K., and Kohl, K. D. (2020). A bird's-eye view of phylosymbiosis: weak signatures of phylosymbiosis among all 15 species of cranes. *Proc Biol Sci* 287, 20192988. doi: 10.1098/rspb.2019.2988

Tsukayama, P., Boolchandani, M., Patel, S., Pehrsson, E. C., Gibson, M. K., Chiou, K. L., et al. (2018). Characterization of Wild and Captive Baboon Gut Microbiota and Their Antibiotic Resistomes. *mSystems* 3. doi: 10.1128/mSystems.00016-18

Tubbs, C., Durrant, B., and Milnes, M. (2020). Reconsidering the use of soy and alfalfa in southern white rhinoceros diets. *Pachyderm* 58, 135–139.

Tubbs, C. W., Moley, L. A., Ivy, J. A., Metrione, L. C., LaClaire, S., Felton, R. G., et al. (2016). Estrogenicity of captive southern white rhinoceros diets and their association with fertility. *Gen Comp Endocrinol* 238, 32–38. doi: 10.1016/j.ygcen.2016.05.004

Vacca, M., Celano, G., Calabrese, F. M., Portincasa, P., Gobbetti, M., and Angelis, M. de (2020). The Controversial Role of Human Gut Lachnospiraceae. *Microorganisms* 8. doi: 10.3390/microorganisms8040573

van de Peer, Y., Chapelle, S., and Wachter, R. de (1996). A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Res* 24, 3381–3391. doi: 10.1093/nar/24.17.3381

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., et al. (2001). The sequence of the human genome. *Science* 291, 1304–1351. doi: 10.1126/science.1058040

Verbrugghe, A., Hesta, M., Daminet, S., and Janssens, G. P. J. (2012). Nutritional modulation of insulin resistance in the true carnivorous cat: a review. *Crit Rev Food Sci Nutr* 52, 172–182. doi: 10.1080/10408398.2010.499763

Vermeij, G. J., and Lindberg, D. R. (2000). Delayed herbivory and the assembly of marine benthic ecosystems. *Paleobiology* 26, 419–430. doi: 10.1666/0094-8373(2000)026<0419:DHATAO>2.0.CO;2

Vital, M., Gao, J., Rizzo, M., Harrison, T., and Tiedje, J. M. (2015). Diet is a major factor governing the fecal butyrate-producing community structure across Mammalia, Aves and Reptilia. *ISME J* 9, 832–843. doi: 10.1038/ismej.2014.179

Vitousek, P. M., Mooney, H. A., Lubchenco, J., and Melillo, J. M. (1997). Human Domination of Earth's Ecosystems. *Science* 277, 494–499. doi: 10.1126/science.277.5325.494

Walker, F. M., Tobin, A., Simmons, N. B., Sobek, C. J., Sanchez, D. E., Chambers, C. L., et al. (2019). A fecal sequel: Testing the limits of a genetic assay for bat species identification. *PLoS One* 14, e0224969. doi: 10.1371/journal.pone.0224969

Wang, X., Tsai, T., Deng, F., Wei, X., Chai, J., Knapp, J., et al. (2019). Longitudinal investigation of the swine gut microbiome from birth to market reveals stage and growth performance associated bacteria. *Microbiome* 7, 109. doi: 10.1186/s40168-019-0721-7

Wang, Y., Ames, N. P., Tun, H. M., Tosh, S. M., Jones, P. J., and Khafipour, E. (2016). High Molecular Weight Barley β-Glucan Alters Gut Microbiota Toward Reduced Cardiovascular Disease Risk. *Front. Microbiol.* 7, 171563. doi: 10.3389/fmicb.2016.00129

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* 58, 236. doi: 10.2307/2282967

Wasimuddin, Menke, S., Melzheimer, J., Thalwitzer, S., Heinrich, S., Wachter, B., et al. (2017). Gut microbiomes of free-ranging and captive Namibian cheetahs: Diversity, putative functions and occurrence of potential pathogens. *Mol Ecol* 26, 5515–5527. doi: 10.1111/mec.14278

Weingrill, T., Gray, D. A., Barrett, L., and Henzi, S. P. (2004). Fecal cortisol levels in free-ranging female chacma baboons: relationship to dominance, reproductive state and environmental factors. *Horm Behav* 45, 259–269. doi: 10.1016/j.yhbeh.2003.12.004

Wetterstrand, K. A. (2023). *DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP)*. Accessed November 23, 2023, www.genome.gov/sequencingcostsdata

Wexler, H. M. (2007). Bacteroides: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev* 20, 593–621. doi: 10.1128/CMR.00008-07

Whipps, J. M., Lewis, K., and Cooke, R. C. "Mycoparasitism and plant disease control," in *Burge, NM: Fungi in biological control systems*, 161–187.

Widner, B., Leblhuber, F., Walli, J., Tilz, G. P., Demel, U., and Fuchs, D. (2000). Tryptophan degradation and immune activation in Alzheimer's disease. *J Neural Transm (Vienna)* 107, 343–353. doi: 10.1007/s007020050029

Wiegel, J., Tanner, R., and Rainey, F. A. (2006). "An Introduction to the Family Clostridiaceae," in *The Prokaryotes: Volume 4: Bacteria: Firmicutes, Cyanobacteria*, eds. M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (New York, NY: Springer US), 654–678.

Woodmansey, E. J., McMurdo, M. E. T., Macfarlane, G. T., and Macfarlane, S. (2004). Comparison of Compositions and Metabolic Activities of Fecal Microbiotas in Young Adults and in Antibiotic-Treated and Non-Antibiotic-Treated Elderly Subjects. *Appl Environ Microbiol* 70, 6113–6122. doi: 10.1128/AEM.70.10.6113-6122.2004

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S. A., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334, 105–108. doi: 10.1126/science.1208344

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi: 10.1038/nature11053

Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., et al. (2019). Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nat Commun* 10, 2200. doi: 10.1038/s41467-019-10191-3

Yuan, X., Chen, B., Duan, Z., Xia, Z., Ding, Y., Chen, T., et al. (2021). Depression and anxiety in patients with active ulcerative colitis: crosstalk of gut microbiota, metabolomics and proteomics. *Gut Microbes* 13, 1987779. doi: 10.1080/19490976.2021.1987779

7. List of figures

8. List of tables

9. Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich in den letzten Jahren während meiner Promotion unterstützt haben.

Mein besonderer Dank geht zunächst an meinen betreuenden Professor Paul Dierkes. Danke Paul, dass du mir die Möglichkeit gegeben hast das Forschungsgebiet innerhalb deiner Arbeitsgruppe nach der Bachelorarbeit zu wechseln. Auch wenn es eine Erfahrung war das Nachtverhalten von 15 Giraffen parallel zu beobachten hat mich die Mikrobiomforschung dann doch mehr fasziniert, auch wenn der Einstieg nicht immer einfach war. Danke dass du mich immer unterstützt hast und immer ein offenes Ohr für Fragen und Probleme hattest. Ich freue schon auf die weitere spannende Zusammenarbeit.

Vielen Dank auch an Professor Lisa Schulte für die Übernahme des Zweitgutachtens dieser Arbeit und für die tolle Zusammenarbeit während des Mastermoduls in den letzten Jahren.

Besonders dankbar bin ich auch für die tolle Arbeitsatmosphäre in diesem Arbeitskreis und dem Zoo Büro, welche neben der ständigen Versorgung durch Kekse und Kaffee auch zu konstruktiven Gesprächen und fachübergreifender Expertise geführt hat. Insbesondere danke ich Jonas, Isi und Matthias dafür, dass sie mir in den letzten Jahren immer zur Seite gestanden haben.

Außerdem möchte ich mich bei dem Team des Opel-Zoos Kronberg und der „von Opel Hessischen Zoostiftung“ bedanken, ohne die diese Arbeit nicht möglich gewesen wäre. Insbesondere Danke an Dr. Martin Becker und Dr. Tanja Spengler für die Unterstützung beim Proben sammeln. Zusätzlich danke ich auch den Zoos in Neuwied, Landau und Heidelberg für die tolle Zusammenarbeit und die umfassenden und spannenden Einblicke in die Zoowelt.

Abschließend danke ich meiner Familie und Freunden für die emotionale Unterstützung in den letzten Jahren. Ihr wart immer für mich da und habt mir Kraft und Verständnis gegeben, wenn es mal nicht so gut lief.

IV. Veröffentlichungen

Publication A

Declaration of author contributions to the publication Unravelling differences in faecal microbiota stability in mammals: from highly variable carnivores and consistently stable herbivores:

Status: *accepted*
Starting with *accepted:* name of journal: Animal Microbiome

Contributing authors:
Franziska Zölzer (FZ, PhD candidate)
Anna Lena Burger (ALB)
Paul Wilhelm Dierkes (PWD)

What are the contributions of the doctoral candidate and his co-authors?

(1) Concept and design
Doctoral candidate FZ: 70
ALB: 10
PWD: 20

(2) Conducting tests and experiments
Doctoral candidate FZ: 100

(3) Compilation of data sets and figures
Doctoral candidate FZ: 90
PWD: 10

(4) Analysis and interpretation of data
Doctoral candidate FZ: 70
ALB: 10
PWD: 20

(5) Drafting of manuscript
Doctoral candidate FZ: 70
ALB: 10
PWD: 20

I hereby certify that the information above is correct.

Date and place	Signature doctoral candidate
Date and place	Signature supervisor

Zoelzer *et al. anim microbiome* (2021) 3:77
https://doi.org/10.1186/s42523-021-00141-0

Animal Microbiome

RESEARCH ARTICLE Open Access

Unraveling differences in fecal microbiota stability in mammals: from high variable carnivores and consistently stable herbivores

Franziska Zoelzer*, Anna Lena Burger and Paul Wilhelm Dierkes

Abstract

Background: Through the rapid development in DNA sequencing methods and tools, microbiome studies on a various number of species were performed during the last decade. This advance makes it possible to analyze hundreds of samples from different species at the same time in order to obtain a general overview of the microbiota. However, there is still uncertainty on the variability of the microbiota of different animal orders and on whether certain bacteria within a species are subject to greater fluctuations than others. This is largely due to the fact that the analysis in most extensive comparative studies is based on only a few samples per species or per study site. In our study, we aim to close this knowledge gap by analyzing multiple individual samples per species including two carnivore suborders Canoidea and Feloidea as well as the orders of herbivore Perissodactyla and Artiodactyla held in different zoos. To assess microbial diversity, 621 fecal samples from 31 species were characterized by sequencing the V3–V4 region of the 16S rRNA gene using Illumina MiSeq.

Results: We found significant differences in the consistency of microbiota composition and in fecal microbial diversity between carnivore and herbivore species. Whereas the microbiota of Carnivora is highly variable and inconsistent within and between species, Perissodactyla and Ruminantia show fewer differences across species boundaries. Furthermore, low-abundance bacterial families show higher fluctuations in the fecal microbiota than high-abundance ones.

Conclusions: Our data suggest that microbial diversity is significantly higher in herbivores than in carnivores, whereas the microbiota in carnivores, unlike in herbivores, varies widely even within species. This high variability has methodological implications and underlines the need to analyze a minimum amount of about 10 samples per species. In our study, we found considerable differences in the occurrence of different bacterial families when looking at just three and six samples. However, from a sample number of 10 onwards, these within-species fluctuations balanced out in most cases and led to constant and more reliable results.

Keywords: 16S rRNA gene, Microbiota, Herbivores, Carnivores, Variability

Background

Due to intensive research in the field of microbiome science and further development of DNA sequencing, the tasks and importance of gastrointestinal microorganisms, especially the production of short-chain fatty acids (SCFA) serving the host organism as energy supply, are now well described [1–3]. In recent years, a lot of research has been conducted to analyze the composition and factors influencing the microbiome for various species using two different approaches. The first often-used study design focuses on a single species or on a specific taxonomic classification. Here, multiple samples per

*Correspondence: zoelzer@bio.uni-frankfurt.de
Bioscience Education and Zoo Biology, Goethe University Frankfurt, Max-von-Laue-Str. 13, 60438 Frankfurt am Main, Germany

individual or species are analyzed, representing one or several time points. Especially farm animals e.g. cattle [4–6], pigs [7–9] or sheep [10, 11], have been largely analyzed due to their importance in agriculture. The microbiota of some wild species, especially highly endangered species such as black rhinos [12], koalas [13] and Tasmanian devils [14], has also been described in more detailed studies. The advantage of this study design is that the microbial composition and diversity of the species studied can be compiled in detail and comprehensively. Moreover, further factors influencing the microbial composition can also be determined in in-depth statistical analyses.

The second study design focuses on an overall comparison within or between groups of animals e.g. terrestrial [15–19] and marine mammals [20], amphibians [21] or birds [22]. In contrast to the former approach, studies involving numerous species are usually based on a smaller number of samples per species or collection site. A possible disadvantage of this approach could be non-representative results of these analyses due to the limited number of samples per species studied. Especially for studies on Carnivora, there are notable inconsistencies across different studies. For instance, two lion samples show a dominance of *Fusobacteria* and *Firmicutes* in one study [17], while three lion samples of another one lack of *Fusobacteria* and instead contain *Actinobacteria* [16]. A similar pattern occurs in studies on different tiger and fox subspecies. While about half of the samples in one study [17] consist of *Proteobacteria* and *Fusobacteria* respectively, another study found large differences for those microbial families [raref. The above-mentioned examples raise the question whether a minimum number of samples is needed to describe the microbiota of a carnivore species. In addition, the issue remains whether there are taxa that are more susceptible to microbial fluctuations, or whether this is due to specific bacterial species.

We aim to integrate the above-mentioned approaches by analyzing a comprehensive dataset of four major mammalian (sub-)orders (Canoidea, Feloidea, Perissodactyla and Ruminantia) to identify differences within or between those. As those animals each have a characteristic digestive system and rely on a different diet, they are well suited to test for variation in their microbial composition. The digestive tract of the Carnivora is short and—beside that of the Insectivora—also one of the least complex among mammals. It is characterized by a short intestine and colon, as well as a small cecum. Carnivora are among the hindgut fermenter, which have the highest microbial density in the appendix, colon and rectum [23, 24]. In general, individuals of this order show only slight adaptations to microbial fermentation, since they rely on easily digestible protein-rich nutrition and have lower glucose needs [25, 26]. Analyses of 16S rRNA gene have shown a low bacterial diversity in the stomach of carnivores, but that diversity increases steadily within the distal intestinal sections [27]. In contrast to carnivores, herbivores such as Perissodactyla and Ruminantia depend on microbial fermentation for cellulose and hemicellulose degradation. Perissodactyla, as hindgut fermenters, are characterized by a simple stomach similar to Carnivora, but in contrast have an enlarged large intestine to extend the retention time of food, as well as an enlarged cecum as the main place of microbial fermentation. Compared to monogastric animals, ruminants have a segmented stomach consisting of the rumen, reticulum, omasum and abomasum. In contrast to the Perissodactyla, ruminants are foregut fermenters, in which microbial fermentation mainly takes place in the rumen. While the small and large intestines are similar in size to the Perissodactyla, the cecum is reduced [28, 29].

In order to create such a widespread dataset, microbiome analyses of zoo-housed animals are suitable in different ways. First, it is necessary to know as many individual and environmental influencing factors as possible to create a representative dataset using multiple samples per species, individuals and collection sites. In this regard, zoos offer a nearly perfect environment because the general conditions such as nutrition, age and pedigree of the animals are well-known. Second, microbiome research is of great interest for the zoos to improve animal welfare. Finally, the microbiota influences a variety of physiological and behavioral processes and, accordingly, a healthy microbiota is correlated with an animal's fitness. Other aspects that are largely unclear so far include possible changes in the microbiome in specific situations such as animal transport, animal socialization or feed conversion. With a meaningful dataset, deviations from the species-specific references can be identified and potential treatments initiated.

Results

In total, we analyzed 621 fecal samples of 31 zoo-housed carnivore and herbivore species, performing Illumina MiSeq paired-end sequencing of the V3–V4 region of the 16S rRNA gene. After quality filtering and read merging, the dataset consists of 12,651,811 sequences (2315–134,440 sequences per sample) with an average of 20,308 sequences per sample. Following the DADA2 pipeline in QIIME 2, we identified 21,058 different amplicon sequence variants (ASV), across all samples (2315 to 134,414 ASV's per sample). The most common classified ASV represented 453,104 times in 329 samples and belongs to a *Clostridium perfringens* strain.

Composition of fecal microbiota of major mammalian (sub-)orders

We found significant differences between herbivores and carnivores in the microbial composition (ANOSIM statistic: R = 0.50, $p < 0.001$, number of permutations: 999, distance = "bray") as shown in Fig. 1B. As can be seen in this figure, the four major bacterial families across all herbivoresare *Spirochaetaceae* (Average ± standard deviation: 15.3 ± 9.0%), *Lachnospiraceae* (15.3 ± 5.8%), *Rikenellaceae* (14.5 ± 4.4%) and *Oscillospiraceae* (12.4 ± 4.3%) (Additional file 2). Within the herbivores, *Spirochaetaceae* are more than twice as common in Perissodactyla (23.2 ± 4.4%) than in ruminants (8.5 ± 5.8%). While this family is equally distributed across perissodactylan species, within the ruminants it only occurs in larger proportions in giraffes (14.3%) and okapis (15.0%). In contrast, we found on average 20.2 ± 3.9% of *Lachnospiraceae* in Perissodactyla and only 11.1 ± 3.4% in ruminants, where larger proportions were observed in reindeer (18.9%). *Rikenellaceae*, the third most-common family in herbivorous species, constitutes on average to 16.1 ± 4.1% of the fecal microbiota of ruminants and to 12.6 ± 4.0% that of Perissodactyla. With respect to the *Oscillospiraceae*, we found notable differences between Ruminantia and Perissodactyla. While this family is equally abundant across nearly all ruminants (14.8 ± 2.7%), it only appears

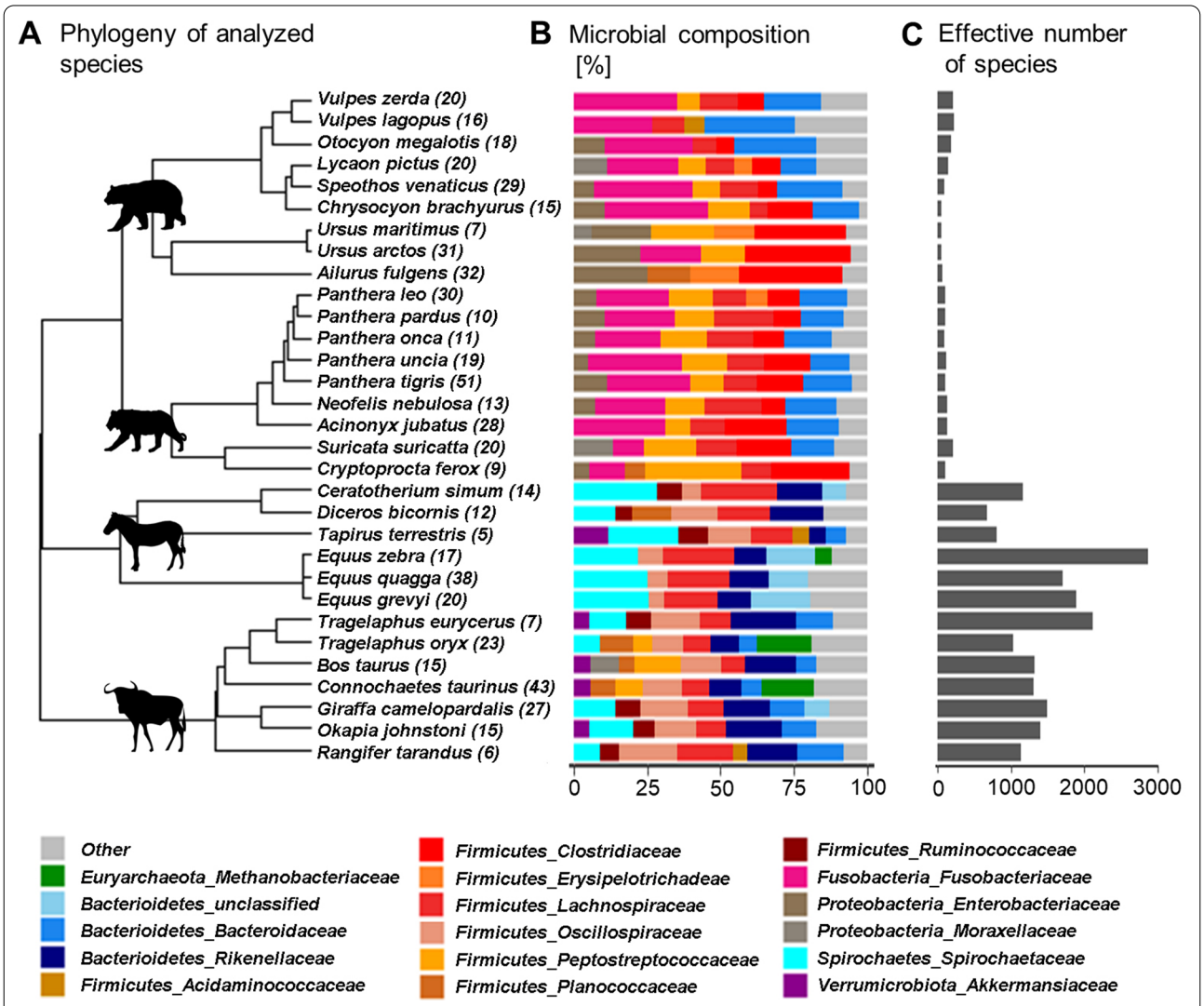

Fig. 1 Variation in the fecal microbiota of mammals. **A** Phylogeny of the analyzed 31 species based on TimeTree database [30]. The total number of samples per species is shown in brackets. **B** Average composition of the fecal microbiota per species. Microbes that occur in less than 5% are summarized under "Other". ANOSIM on the four groups: permutations = 999, distance = bray, R = 0.496, $p < 0.001$. **C** Average fecal diversity per species presented as number of effective species. Kruskal–Wallis on the four groups: $p < 0.01$, df = 3, Dunn Test with Bonferroni correction $p < 0.001$

in tapirs (14.3%) and black rhinoceros (15.8%) in greater proportions of all Perissodactyla (9.6 ± 4.0%). Besides those four major families, we identified *Bacteroidaceae* in many ruminants (10.2 ± 3.4%) and an uncultured bacterium *p-251-o5* of the Bacteroidales order in Perissodactyla (9.7 ± 7.7%), especially in the grevy's zebras (20.3%). Other bacterial families such as *Tanerellaceae*, *Erysipelotrichaceae*, *Clostridiaceae*, *Fusobacteriaceae* and *Enterobacteriaceae* constitute on average less than 5% of the microbiota across all herbivore species.

Furthermore, Fig. 1B illustrates that *Fusobacteriaceae* is the most dominant bacterial family in Carnivora species, occurring on average in 23.2 ± 7.1% of all Feloidea and in 22.38 ± 13.1% of all Canoidea. However, within the Canoidea, this family is low-abundant in red pandas and brown bears as it constitutes to less than 5% of both fecal microbiota. The distribution of *Clostridiaceae* (15.9 ± 10.1%), the second dominant family within the Carnivora, is on average similar for Feloidea (15.2 ± 5.8%) and Canoidea (16.6 ± 13.0%). *Clostridiaceae* form a large proportion of the microbiota, accounting for more than 30%, in both bears and red pandas. Those species also differ from other Canoidea with regard to *Bacteroidaceae*. Whereas this family is frequently found in most Carnivora (14.2 ± 8.9%), it is low-abundant (< 5%) in the red pandas, brown bears, polar bears and fossas. Additionally, we found on average 16.0 ± 6.5% *Peptostreptococcaceae* in Feloidea and only 8.5 ± 7.1% of this family in Canoidea, but the value calculated for Felidae is mostly influenced by its high abundance of 33.0% in fossas. Beside these major bacterial families, some others are largely represented in both bear species and red pandas. For example, we found that *Enterobacteriaceae* contribute on average 25.3% to the fecal microbial composition in red pandas, to 22.7% in polar bears and to 20.4% in brown bears. Furthermore, *Erysipelotrichaceae* are more dominant in brown bears (13.7%) and red pandas (16.5%) than in other Canoidea (4.0 ± 6.2%). With regard to the Felidae, *Lachnospiraceae* (14.0 ± 3.5%) are another dominant family, being equally distributed across all sampled felid species. Other bacterial families such as *Spirochaetaceae*, *Rikenellaceae* and *Oscillospiraceae*, which were dominant in herbivorous species, accounted for less than 5% of the carnivore microbiota.

Microbial diversity within and between herbivores and carnivores

The microbial diversity measured by effective number of species differs significantly between carnivores and herbivores as shown in Figs. 1C and Fig. 2C

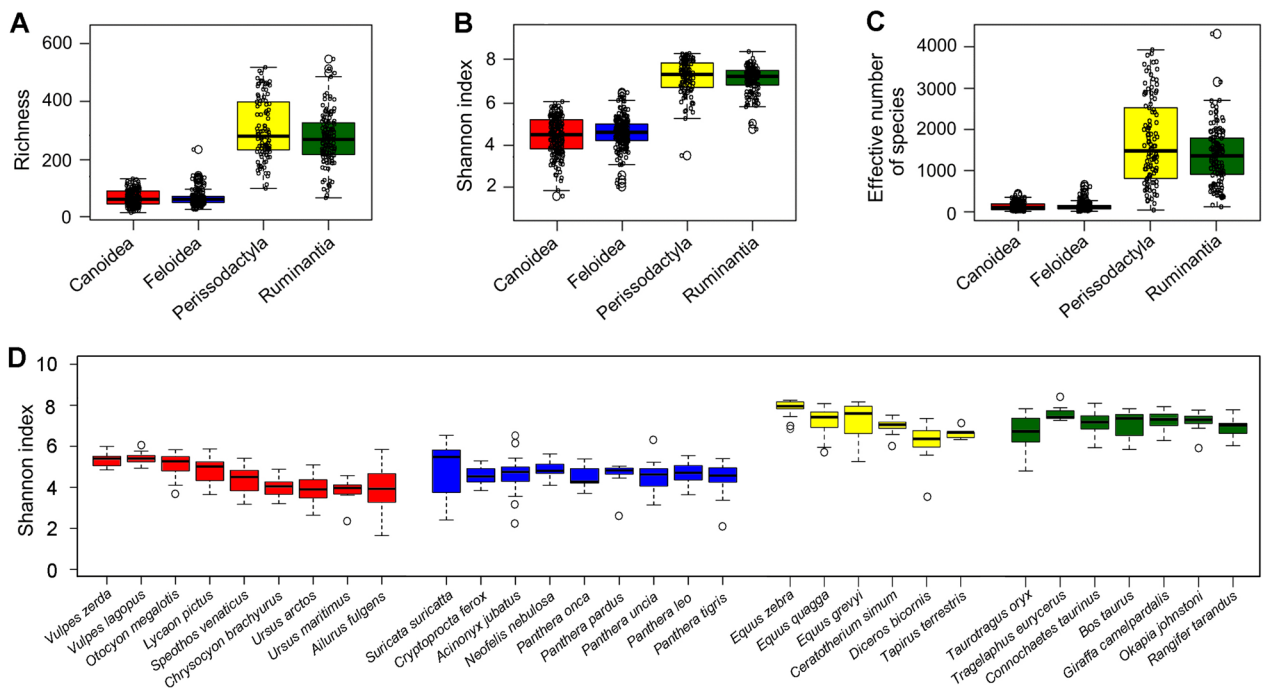

Fig. 2 Alpha diversity of carnivores (Canoidea and Feloidea) and herbivores (Ruminantia and Perissodactyla) measured as species richness (**A**), Shannon index (**B**) and effective number of species (**C**). Part (**D**) shows the Shannon index for all analyzed species. Kruskal–Wallis $p < 0.01$, df = 3, Dunn Test with Bonferroni correction $p < 0.001$ for all diversity indices Statistical results for pairwise comparisons are presented in an Additional file 4)

(Kruskal–Wallis: $p < 0.001$, df = 3, Dunn Test with Bonferroni correction $p < 0.001$), while there are no significant differences between Canoidea (90.0 ± 88.2) and Feloidea (101.1 ± 93.9) as well as between Perissodactyla (1475.9 ± 1030.5) and Ruminantia (1350.4 ± 673.3). Besides the ENS, those significant differences between carnivores and herbivores are further illustrated in the Shannon index and species richness (Fig. 2A, B). The median of the Shannon index is 4.5 ± 0.9 for Canoidea which is similar to Feloidea (4.6 ± 0.7) and significantly different ($p < 0.01$) to Perissodactyla (7.3 ± 0.8) and Ruminantia (7.2 ± 0.6). Furthermore, comparable results are obtained with the species richness ($p < 0.01$), which is more than four times higher in perissodactylan (279.0 ± 103.5) and ruminant species (268.5 ± 87.3) than in Canoidea (61.0 ± 27.0) and Feloidea (60.0 ± 25.5). Consequently, Carnivora species show a reduced microbial diversity over all measurements compared to Perissodactyla and Ruminantia species. Regarding the Shannon index across all species within a (sub-)order, further differences become visible (Fig. 2D). Within the Canoidea the greatest variation is found within the red pandas (3.9 ± 0.9). Additionally, the red panda samples show a significantly lower Shannon index compared to the Vulpini species represented by the fennec fox (5.4 ± 0.3, $p < 0.001$), arctic fox (5.4 ± 0.3, $p < 0.001$) and bat-eared fox (5.2 ± 0.6, $p = 0.004$). These three species generally show the highest alpha diversity within the Canoidea and differ significantly from the brown bear (3.9 ± 0.6, $p < 0.001$) and maned wolf samples (4.0 ± 0.4, $p < 0.001$). The Shannon index within the Feloidea species is very similar among species, and just the suricate samples show greater deviations (5.4 ± 1.2). Compared to some big cat species as the cheetah (4.7 ± 0.8), lion (4.7 ± 0.5), snow leopard (4.6 ± 0.7) or tiger (4.5 ± 0.5), the suricate samples show a significantly greater alpha diversity ($p < 0.05$). The zebras show the highest alpha diversity within the Perissodactyla, with the mountain zebra having a significant higher diversity (8.0 ± 0.4) compared to the plains zebra (7.4 ± 0.6, $p < 0.05$), tapir (6.7 ± 0.3, $p < 0.05$), black (6.5 ± 1.0, $p < 0.001$) and white rhino (7.1 ± 0.4, $p < 0.001$). Additionally, the highest variation was found within the grevy's zebra (7.6 ± 0.9). The Shannon index within the analyzed ruminants is similar across all species. Only the elands (6.7 ± 0.8) show a significantly lower Shannon index compared to bongos (7.4 ± 0.4, $p < 0.05$) and wildebeests (7.1 ± 0.5, $p < 0.05$).

Regarding the beta diversity, the principal coordinate analysis (PCoA) of the unweighted UniFrac distance matrix explains a total of 46.3% of data variability within the first three main axes (Fig. 3A, B), while the weighted UniFrac matrix explains a total of 63.3% of the data (Fig. 3C, D). Both plots show a clear separation between carnivores and herbivores, indicating a general difference in bacterial composition between these two groups. Homogeneity of dispersion is given within the four (sub-) orders (F = 0.670, $p = 0.570$, permutations = 999) and the ADONIS test shows significant differences in the fecal microbial composition between Canoidea, Feloidea, Ruminantia and Perissodactyla ($R^2 = 0.020$, $p < 0.001$, permutations = 999). This is also confirmed by the PCoA of the unweighted UniFrac measurement (Fig. 3A, B). Similar to the weighted UniFrac, the homogeneity of dispersion is given for the animal (sub-)order (F = 0.670, $p = 0.570$, permutations = 999) and also for this metric, we found significant differences in the fecal microbial composition between the four(sub-)orders ($R^2 = 0.020$, $p < 0.001$, permutations = 999).

Regarding the Perissodactyla and Ruminantia, both form clearer clusters in the unweighted UniFrac than in the weighted UniFrac measurement. This suggests that both, Perissodactyla and Ruminantia, can be distinguished by their general bacterial composition. Furthermore, in combination with the pattern observed in the weighted UniFrac plot, some differences within Perissodactyla and Ruminantia become visible which can be explained by the different abundance of some bacterial taxa. Thus, both herbivore groups consist of a similar microbiota that differs in the abundance of certain bacterial taxa. In contrast, there is no clear separation between Canoidea and Feloidea in either plot, indicating a differing bacterial composition within the Carnivora. In the unweighted UniFrac plot of the Carnivora (Fig. 3B), a slight pattern becomes visible. At the order-specific level, the Carnivora are divided into three clusters (Fig. 3B). The first cluster, closest to the Perissodactyla, consists of the polar and brown bear as well as the red panda samples. A little distant from these lies the center of the second cluster, made of the big and small cats as well as the South American Cerdocyonina represented by the bush dog and maned wolf samples. Finally, the third cluster, which is most distant from the herbivorous species, is composed of the Vulpini (fennec fox, arctic fox, bat-eared fox) and the African wild dog samples. Since these clusters are based on the unweighted UniFrac method, they can be distinguished from each other by a generally different bacterial composition. Since these clusters are less clear in the weighted UniFrac plot (Fig. 3D), these differences might be explained by the occurrence of low-abundant bacterial taxa, which are not visible when bacterial abundances are taken into account. Noticeably, four samples fall between the herbivores and carnivores, which belong to two elands from the same zoo and two cheetahs respectively. Since these animals were apparently healthy and did not differ in any other way from other sampled herd members, these outliers can at best be explained

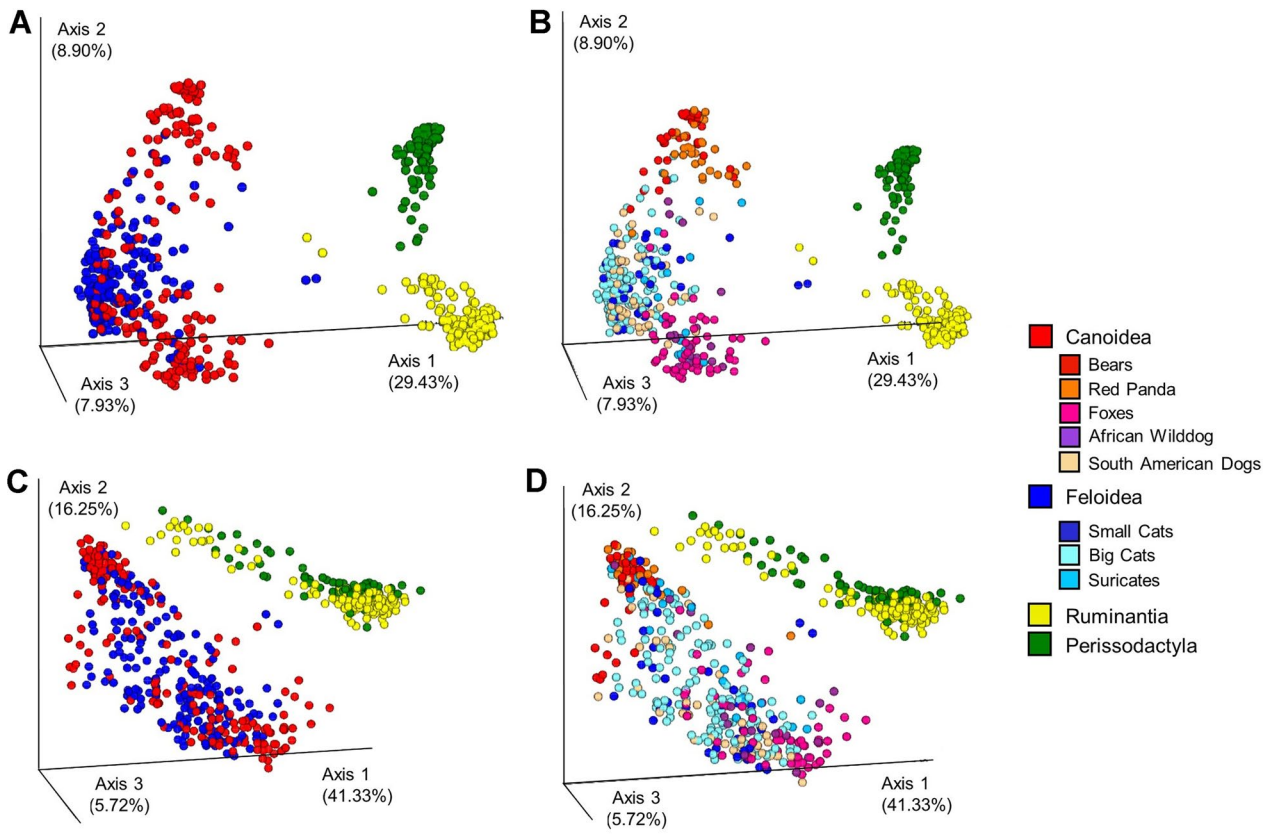

Fig. 3 Principal Coordinates Analysis on the differences between carnivores and herbivores based on an unweighted and weighted UniFrac distance matrix. Differences in samples are shown based on the four analyzed groups (**A**, **C**) and on a more detailed division of carnivore groups (**B**, **D**) as shown in the figure legend. The proportion of data explained by this measurement is shown in brackets for each axis

by a reduced read count (8204 and 7631 sequences for elands and 6,521 and 10,028 sequences). Regarding the two cheetahs, the general variability within the small cats is very high (Fig. 3B) and those samples might just underlie these deviations.

For a more detailed analysis of the variation within the Carnivora, we focused on fluctuations within the most common bacterial families, calculated as coefficient of variation (CV). The CV is defined as the ratio of standard deviation to the mean. Figure 4 shows the CV plotted against the number of samples and against the total percentage of occurrence of herbivores (4A) and carnivores (4B). These figures show three main results. First, the CV is in general lower for the illustrated bacterial families in carnivores compared to herbivores. Whereas the CV for the most dominant bacterial families within herbivores mostly not exceeds values of 1.0, the respective values within carnivores are about twice as high, e.g. for *Peptostreptococcaceae*, indicating higher variation within this bacterial family. Second, the relative variation (CV) of the low-abundant bacterial families (e.g., *Enterobacteriaceae*) is significantly greater on average per species than the variation of the high-abundant families (e.g., *Clostridiaceae* and *Fusobacteriaceae*), although the absolute variation of these bacterial families within the species studied is similar. Third, it is noticeable that the CV does not necessarily decrease with regard to a larger number of samples being analyzed, at least not when all herbivores or all carnivores are considered together. To examine whether this effect is possibly affected by species-specific differences, we created randomized subsets of bacterial abundance data for different sample numbers (n = 3, 6, 10, 15, 20, 25) with three replicates each, of three carnivore and herbivore species. For this purpose, we used bacterial families that occur in more than 7% of all herbivore or carnivore species, because low-abundant families seem to have a higher variability per se as shown before. Within all species, this results in a decreased coefficient of variation as the number of samples increases (Fig. 5). This clearly shows that when analyzing only a few samples per species (n = 3 or 6), there is generally greater variability in bacterial abundance

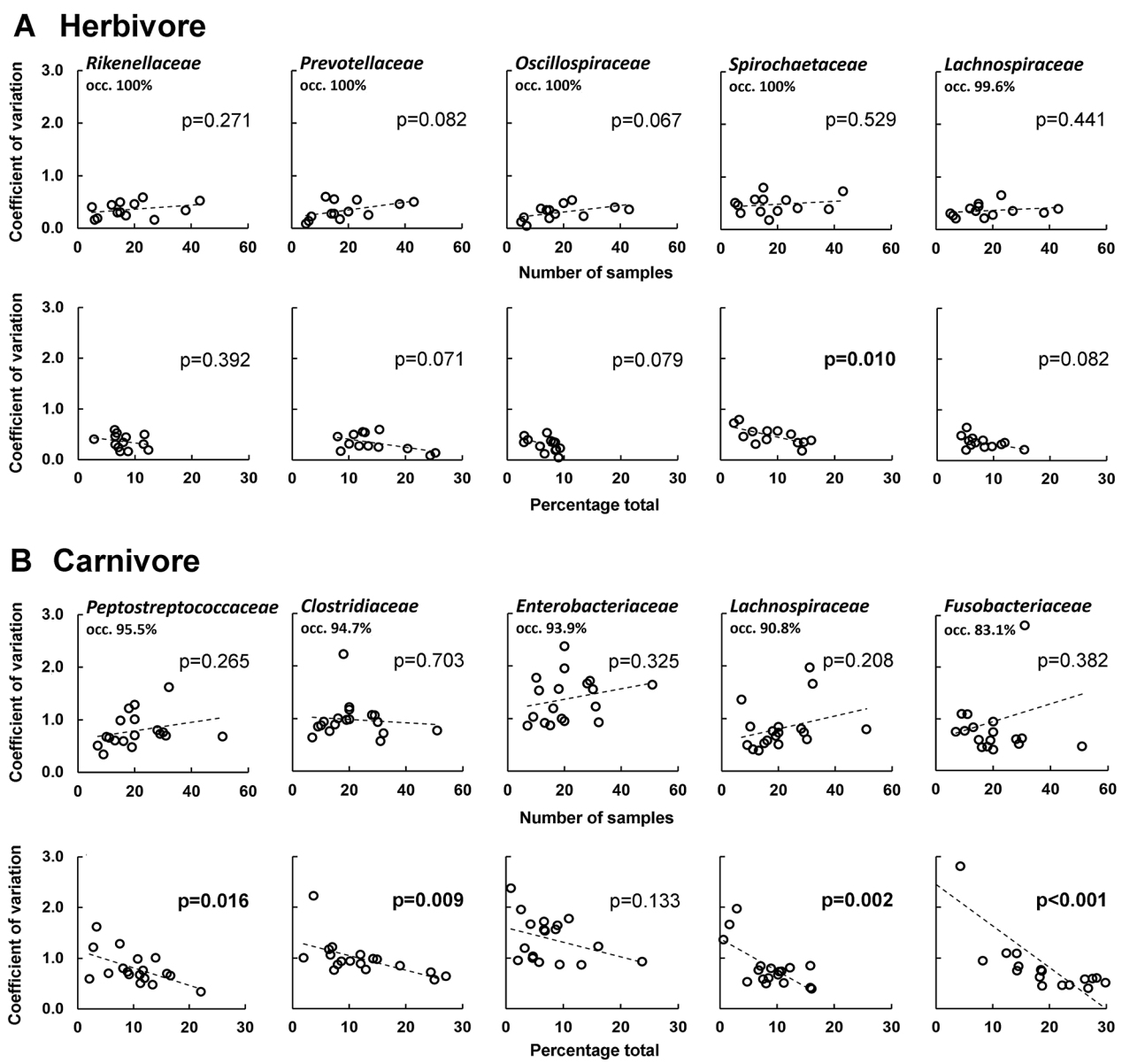

Fig. 4 Coefficient of variation of different bacterial families for all herbivores (**A**) and carnivores (**B**) plotted against the number of samples (top row in each case) and against the total percentage of occurrence averaged per species (bottom row in each case). The tendency is indicated by a linear regression line and significant *p*-values are indicated in bold. The occurrence in the total sample is given for each bacterial family

data between samples than when using larger numbers of samples (n = 20 or 25). In addition, species-specific differences become visible. For example, giraffes show a constantly low variability in both bacterial families, even when only a few samples are considered. In contrast, wildebeests and plains zebras are more variable when only a few samples are taken into account and first stabilize at a sample number of 15 in both analyzed bacterial families. Within carnivores, the tiger samples show a constant CV for all bacterial families from a sample number of n = 10. Even if the variability within the lion samples is higher compared to the tiger ones, they also become stable from a sample number of 10 onwards. Besides species-specific differences, we also found differences in the variability between bacterial families in the brown bear. While the pattern for *Peptostreptococcaceae* and *Clostridiaceae* is the same as in tigers and lions, the high CV values of the *Fusobacteriaceae* is not noticeably declining with an increased sample size. Detailed results are shown in the Additional file 3.

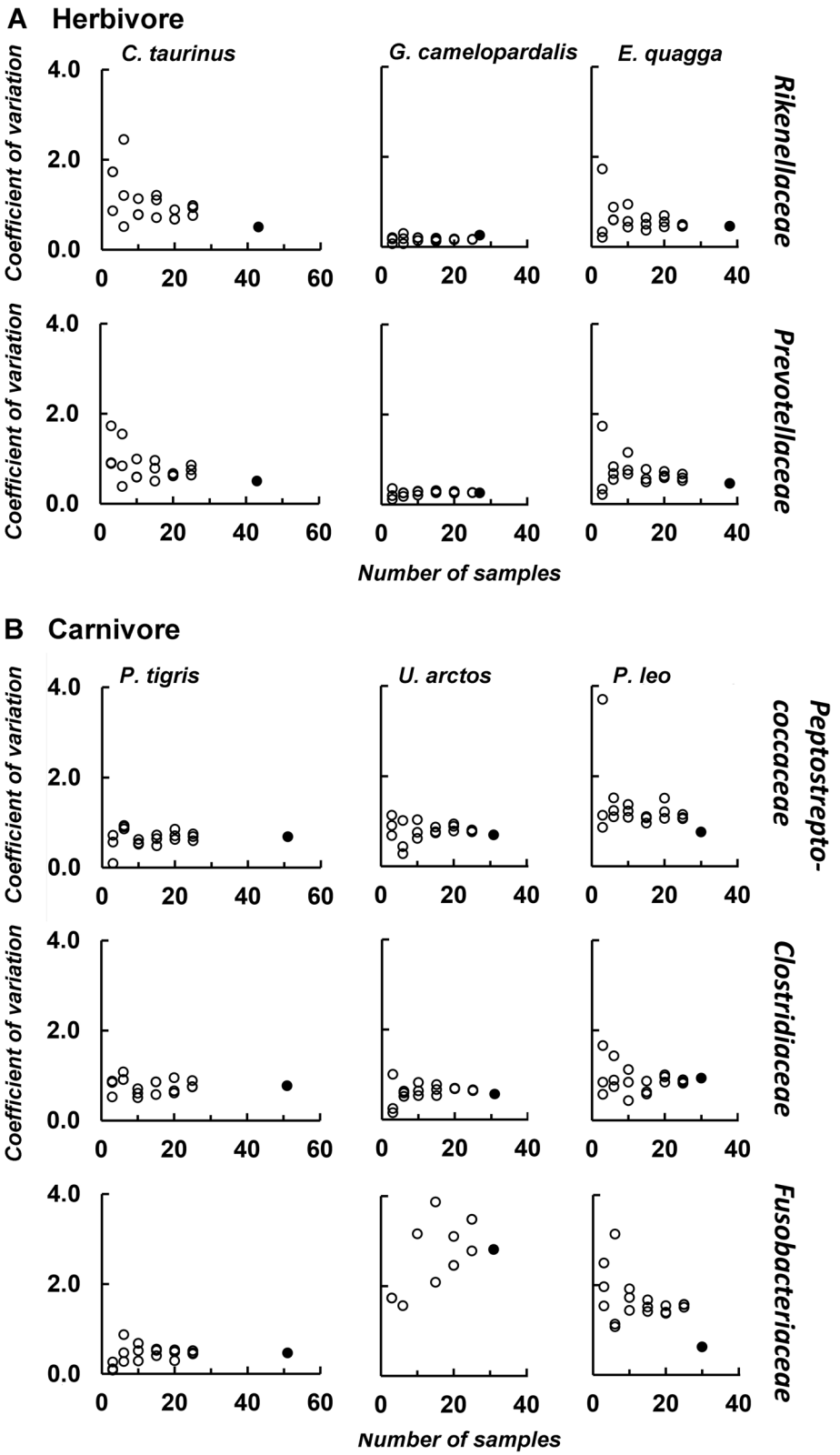

Fig. 5 Coefficient of variation of different bacterial families for selected herbivorous (**A**) and carnivorous (**B**) species. Shown are randomized subsets (unfilled circles) for a different number of samples, as well as the entire data set (filled circles)

To control for the zoo habitat as a possible influencing factor on the fecal microbiota, we performed a multinomial regression model on the microbial differential abundance data. The evaluation of the model setting 'zoo' against a null model obtained a Q^2 value of 0.13. Compared to that, the model containing only 'species' as explanatory variable obtained a Q^2 value of 0.33. A combined model ('zoo' + 'species') results in a slightly higher Q^2 value of 0.43. In order to distinguish the obtained zoo effect more precisely from the effect of the variable 'species', we compared the model including 'zoo' as a variable against a baseline model containing 'species' as a variable. This results in a negative Q^2 score, illustrating that the variable 'zoo' does not improve the model when 'species' is set as a baseline.

Microbial indicators for herbivore and carnivore animals

Indicator families were analyzed for each of the four (sub-)orders and each possible combination using the IndVal.g function. We identified a total of 276 indicator families, most of them for herbivores, especially for Perissodactyla (Table 1). With 18 indicator families, Canoidea and Feloidea share less indicators than Perissodactyla and Ruminantia and only minor proportions of indicator families were found in combinations of herbivore and carnivore species. The complete results are presented in the Additional file 5.

Almost all predicted indicator families show high A values, meaning that this indicator only occurs in the tested (sub-)order, but is not necessarily spread across all of its members. In contrast, the B values, showing the distribution of an indicator across all taxa, are much more variable. Indicator families restricted to Canoidea are *Gemellaceae* (A = 1.00, B = 0.03) and *Xiphinematobacteraceae* (A = 1.00, B = 0.02), but they do not occur in all the samples. Regarding the Feloidea, no exclusive indicators were found. However, *Coriobacteriaceae* (A = 0.88, B = 0.88) are strongly related to this suborder and distributed among nearly all members. In general, all indicator families associated to the Carnivora show low B values, which might be a further indication of greater diversity within the two suborders as seen in the PCoA analysis. However, this view changes when one considers the indicator families that occur in both the Feloidea and the Canoidea. In particular, *Enterobacteriaceae* (A = 0.98, B = 0.94), *Clostridiaceae* (A = 0.96, B = 0.95) and *Fusobacteriaceae* (A = 0.99, B = 0.83) occur in almost all Carnivora species and appear to be clear indicator families for those in general. Additionally, these families are also the most dominant ones in the Carnivora fecal microbiota composition (Fig. 1b).

In contrast, more indicator families were found in Perissodactyla and Ruminantia. *Fibrobacteraceae* (A = 0.81, B = 0.97), *Synergistaceae* (A = 1.00, B = 0.75), *Defluviitaleaceae* (A = 0.88, B = 0.80) and *Methanocorpusculaceae* (A = 0.79, B = 0.88) occur almost exclusively in Perissodactyla and are present in almost all species. For ruminant species, one of the most prominent indicators are *Barnesiellaceae* (A = 0.89, B = 0.72) and *Atopobiaceae* (A = 0.73, B = 0.46), which occur in many members of this suborder. Looking at the combined indicators of Perissodactyla and ruminants, many microbial families are found almost exclusively in those two (sub-)orders and are present in all taxa. Again, those indicator families are among the most dominant ones in the taxonomy plot (Fig. 1b) i.e. *Spirochaetaceae* (A = 0.99, B = 1.00), *Rikenellaceae* (A = 0.96, B = 0.99) and *Oscillospiraceae* (A = 0.87, B = 0.90).

Table 1 Microbial indicators for different animal (sub-)orders and their combination

(sub-)order	Number of indicator species
Canoidea	10
Feloidea	6
Perissodactyla	43
Ruminantia	16
Canoidea + Feloidea	18
Perissodactyla + Ruminantia	42
Canoidea + Perissodactyla	1
Canoidea + Ruminantia	3
Feloidea + Perissodactyla	2
Canoidea + Feloidea + Perissodactyla	3
Canoidea + Feloidea + Ruminantia	6
Canoidea + Perissodactyla + Ruminantia	4
Feloidea + Perissodactyla + Ruminantia	2

Indicators were assigned at microbial family level

Discussion

The aim of this work was to conduct a study on the variability of the microbiota of zoo-housed carnivore and herbivore species, with a focus on the four (sub-)orders Canoidea, Feloidea, Perissodactyla and Ruminantia. In contrast to previous studies using just a few samples per species, we analyzed multiple samples per species and compared the microbiota of species from different locations. Our study results in two main findings. Firstly, we found significant differences in the microbiota composition of carnivorous and herbivorous species, as well as a significant higher alpha diversity in herbivores. Secondly, we found closer similarities and less variability in the fecal microbiota of Perissodactyla and Ruminantia compared to higher deviations in Carnivora, which has some

important methodological implications as discussed below.

Differences in the microbiota composition between carnivores and herbivores

We found significant differences in the fecal microbial composition between herbivore and carnivore species. The most dominant bacterial families found in herbivore species are *Spirochaetaceae*, *Lachnospiraceae*, *Rikenellaceae* and *Oscillospiraceae*. The first two mentioned occur more frequently in Perissodactyla, whereas the latter two appear on average more often in ruminants. Those results are in line with the in-depth study on African herbivores [19], who also found *Oscillospiraceae* as the most dominant family in ruminants such as giraffes, cattle or hartebeests. Nevertheless, our study showed greater proportions of *Rikenellaceae* in ruminants. Both, *Oscillospiraceae* and *Rikenellaceae*, have recently been characterized as herbivore specific bacteria in a covariance network analysis [18], with *Oscillospiraceae* being a major player in cellulose degradation and therefore being related to a herbivore and fiber-rich diet [31]. Another link to the study on African herbivores [19] is the appearance of *Spirochaetaceae*, especially in zebras, as representatives of Perissodactyla. Similar to *Oscillospiraceae*, this family is responsible for fiber digestion and therefore essential for the herbivore digestive system [32, 33]. Besides *Spirochaetaceae*, we found *Lachnospiraceae* as another main family in Perissodactyla. This family has been detected in the human intestine as well as in the rumen and digestive system of different mammals [34, 35]. Bacteria belonging to this family such as *Roseburia* or *Lachnospira* are involved in the production of SCFAs by hydrolyzing sugars (e.g. starch) and were found to be associated with the consumption of plant protein and fiber [36, 37]. Additionally, the abundance of *Lachnospiraceae* can decrease with regard to a high-protein diet, indicating a minor role in protein metabolism [38]. Those major bacterial families found in herbivorous animals are mainly capable of carbohydrate digestion like starch or maltose, allowing the host to gain enough energy from the plant-based diet.

In contrast, the main bacterial families found in Carnivora are *Fusobacteriaceae*, *Clostridiaceae*, *Bacteroidaceae* and *Peptostreptococcaceae*. *Fusobacteriaceae* are often linked to a high-fat and protein-based diet and were observed in different carnivores, with *Fusobacterium* previously being classified as a carnivore specific bacterium [18, 39]. This bacterial family is able to produce SCFAs using carbohydrates or amino acids [40] and it has been shown that *Fusobacteriaceae* are more common in carnivorous Carnivora than in omnivorous or herbivorous Carnivora [41], which is consistent with our study. Both, *Clostridiaceae* and *Bacteroidaceae*, being dominant in carnivore families in our study have already been detected in the gastrointestinal microbiota of different predators [18, 39, 42, 43]. While *Clostridiaceae* appear to be important for protein metabolism, *Bacteroidaceae* occur in combination with a fiber-rich diet and are not affected by protein intake [44–46]. In summary, our results show the highest proportion of *Bacteroidaceae* in bat-eared foxes as well as the highest proportion of *Clostridiaceae* in polar bears, which partly matches this theory. However, we could not find major differences for these two bacterial families.

Beside significant differences in the microbial taxonomic assignment between carnivorous and herbivorous mammals, we also found a significantly higher microbial alpha diversity in Ruminantia and Perissodactyla compared to Carnivora. This might be due to the more complex digestive system of herbivorous species and their dependence on microbes to break down cellulose. This relationship has been shown previously for several species [35, 39, 41, 47, 48]. Furthermore, herbivorous mammals are known to rely on microbial metabolic pathways to a greater extent than carnivores [18].

In addition to confirming previous studies on the carnivore microbiota, we have also found some species that deviate from previous assumptions, namely both bear species, the red panda and the fossa. Contrary to the other Carnivora, *Fusobacteriaceae* only occur in minor proportions within red pandas and brown bears, but *Erysipelochtrichaceae* are enriched in these animals. Furthermore, both bear species and the red pandas consist of major proportions of *Enterobacteriaceae* but only of minor proportions of *Bacteroidaceae*—similar to the fossa. Within the PCoA plot of beta diversity (Fig. 3B), the fossa samples lie within those of other felids, whereas the two bear species as well as the red pandas form a separate cluster apart from the Feloidea and the Canidae. The most influencing factors for fecal microbiota composition are described to be diet and phylogeny [17, 32, 49]. Because the omnivorous diet of the analyzed bears was similar to that of the other Canoidea as e.g. the Vulpini species which form an own cluster, and even the red pandas were fed an omnivore diet in half of the analyzed zoos, it is unlikely, that this separation is mainly influenced by diet. Another factor influencing the microbial composition is the host phylogeny. Bears, red pandas and fossa all evolved separated from other members of the respective suborder. The fossa as a Malagasy carnivore evolved distinct from other Felidae as a sister clade to the Herpestidae about 18–24 Mya ago [50, 51]. Regarding the Caniformia, the Arctoidea clade split in a rapid radiation about 43 Mya in three superfamilies Ursoidea, Pinnipedia and Musteloidea. Within these, the Ursidae

evolved about 18 Mya ago, whereas the Ailuridae evolved about 33 MYA ago as a sister clade to Mephitidae, Procyonidae and Mustelidae [52–55]. In recent years, the theory of co-evolution between host and microbes arose and continues to be proven. It states that bacterial symbionts adapt to e.g. dietary changes of the host and the host in turn adapts to the changed microbiota or that allopatric speciation of the host might even lead to co-phylogenetic patterns between microbes and host [16, 17, 56–58]. Although this was not analyzed in this study, our results may suggest a co-evolution between gut microbes and host phylogeny in different mammalian (sub-)orders. Furthermore, the results indicate that there are clear differences between herbivore and carnivore species but that there are several deviations from previously published gut microbiota.

Close similarity in the fecal microbiota of herbivores and great diversity within the carnivores

Beside significant differences between herbivore and carnivore species, our results reveal a closer similarity in the fecal microbiota of Perissodactyla and Ruminantia compared to higher deviations in Carnivora. Although there are several studies that describe either a distinct clustering of herbivores and carnivores due to differences in diet or phylogeny or a clustering of herbivorous carnivores to other Carnivora [16, 39, 41, 47, 49], none of them has yet referred to the variability of the microbiota within these taxa. A first indication of greater uniformity within Perissodactyla and Ruminantia is the larger variety of indicator species than for Carnivora, which can be explained with an overall higher alpha diversity as well as a closer similarity of the fecal microbes in herbivores. Furthermore, the Carnivora indicators are not distributed equally across all species, indicating a greater intra- and interspecies variation within this order. These differences are further illustrated in Figs. 4 and 5 showing the coefficient of variation within herbivores and carnivores. Here, the coefficient of variation is much higher in low-abundant microbial families compared to high-abundant families. One explanatory approach for the higher deviation within the Carnivora is the diet. While the analyzed herbivores are mostly fed on hay, alfalfa or grass throughout the year, the diet and its composition is more variable in carnivores. Especially omnivorous Carnivora such as most Canoidea are fed on a variety of food sources as fresh and kibble meat, fruits, vegetables or insects. But even hypercarnivore species undergo daily changes in meat origin or preparation (e.g. whole-body or sheer meat). For canids and felids, it is shown that the fecal microbiota is greatly altered by diet and dietary changes. Especially changes in the proportion of carbohydrates and protein influence the necessary gut bacteria, i.e. *Prevotella* or *Fusobacteria* respectively [42, 43, 59–61].

These differences in the microbial variability of carnivorous fecal samples also have important methodological implications. It is therefore necessary to adapt the number of samples being analyzed to the species to be studied in order to obtain meaningful results. Herbivores are very similar in terms of their microbial composition. In ruminants, *Oscillospiraceae*, *Lachnospiraceae,* and *Rikenellaceae* appear to dominate as the major bacterial families [16, 19] and this is evident in studies using different sample sizes. For example, the core results of a study on five giraffe samples are consistent to a similar study on more than 50 giraffe samples and the same pattern can be seen in regard to studies on elands [19, 47, 62] or zebras representing Perissodactyla [16, 19]. These results are in line with the low CV that we found in herbivores. Nevertheless, we found species-specific differences for the major bacterial families within herbivores as well. Giraffes show very low variability in *Rikenellaceae* and *Prevotellaceae,* so these differences should be visible even in very few samples analyzed. In contrast, wildebeest samples are highly variable for those two families, resulting in the need to analyze at least 15 samples to control for these variations.

The Carnivora microbiota in general is much more variable, which is expressed in a higher CV compared to that of Perissodactyla and Ruminantia. Especially within this order, it is therefore important to analyze a reliable number of samples in order to characterize the microbiota. This is also illustrated by the fact that previous studies on carnivores yield significantly different results on the composition of the fecal microbiota. For example, studies using just two or three fecal fox, polar bear or bush dog samples [16, 47] found great differences in the proportion of *Prevotellaceae* and *Fusobacteriaceae*. The same pattern was observed for Feloidea, in studies on just a few cheetah and lion samples which could only detect minor proportions of *Fusobacteria*, whereas a study using more than 60 animals reported about 20% *Fusobacteria* in cheetahs [16, 63–66]. In this study, we found *Fusobacteriaceae* across all Carnivora species in highly different proportions. Within the brown bear samples, this family is present on average in 4.3%, which explains the high coefficient of variation even when using a high amount of samples. But also within the lion and tiger samples, in which the proportion of *Fusobacteriaceae* with an average of 18.3% and 23.5% is considerably higher, the CV for this family only becomes constant with 10 samples being analyzed (Fig. 5). This strengthens our finding that low-abundant bacterial families are more variable in the fecal

microbiota of mammals, and the necessity of analyzing multiple samples to reduce uncertainties that can occur with small sample numbers (n = 3 or 6).

Considering the highly variable microbiota of Canoidea and Feloidea and the more constant microbiota of Ruminantia and Perissodactyla, it is important to select an appropriate number of samples for further analysis. Depending on the methodological approach, it should be noted that low-abundant bacterial families are often subject to greater fluctuations than high-abundant ones, and that there seem to be species-specific differences in microbiota variability within these animal (sub-)orders.

External influencing factors on the microbiota of zoo-housed animals

An often mentioned criticism on the microbiome analysis of zoo animals is the fact that captivity might lead to a reduced microbial alpha diversity in some species [67, 68]. Reasons for this may include a different dietary composition, the use of additives and medicines, or the artificial enclosure design. To address this point of criticism, we have compared some of our data with the methodologically comparable study by McKenzie et al. [69]. They stated, that not all mammalian families are affected equally by a loss of microbial diversity as an effect of captivity. For example, the authors found a significant decrease in the Shannon index in canids. In our dataset, canids of the same species show a Shannon index which is higher than that of their captive samples and which is even more similar to the wild samples. Furthermore, the authors mentioned *Bovidae* and *Giraffidae* not to be impacted by captivity as they obtained comparable Shannon values in the wild and in captivity. Here too, our results are comparable with the diversity measurements of their wild samples. Another interesting finding of their study is that the alpha diversity of captive *Rhinocerotidae* is even increased, we calculated a Shannon index that is very similar to those enlarged value for captive rhinos. Even though the alpha diversity is only one component in the analysis of the fecal microbiota, and a comprehensive comparison would of course need to include the sample's taxonomic composition as well as beta diversity, these results provide first indications for a better understanding of the microbiota diversity of zoo animals.

Nonetheless, our primary goal is to generate a dataset that contains numerous mammalian species, with a defined number of samples per species from different locations (zoos) to get an overall view of species-specific deviations in the fecal microbiota. Even if some species are subject to the captivity effect of reduced microbial diversity, all the samples are equally affected by this and therefore the results themselves are not biased. Rather, the respective zoo could be an external influencing factor on the fecal microbiota and to control for this effect, we conducted a multinomial regression. Regarding the whole dataset of microbial abundance data, the species-specific effect outweighs the effect of the housing location (zoo). Nevertheless, the respective zoo has slight influence on the fecal microbiota which can be caused by for example different feeding regimes, co-habitation and interaction of different species or the enclosure equipment. Furthermore, this zoo-specific effect differs between species and ranges from zero effects (e.g. Cheetah, Red panda) to greater effects in wildebeests or suricates. However, as we only focus on zoo-housed animals and our main focus in this study is not to compare those samples to samples from free-ranging animals, the housing location as influencing factor should balance out across all zoos. Nevertheless, we are aware that the microbiota of wild animals may differ from our results, and our findings clearly relate to captive animals. For them, however, they provide a comprehensive database on which further research can be conducted.

Conclusions

To the best of our knowledge, this is the first study focusing on the microbiota variability of a wide range of carnivore and herbivore mammals by analyzing multiple samples per species in different locations. Our results support already existing theories such as a greater alpha diversity in herbivores or the general description of major bacterial families in Perissodactyla and Ruminantia species. Additionally, we found some species as the brown and polar bear, red panda or fossa that deviate from other members of their diet group. Phylogeny and host-microbe co-evolution may have a greater effect on fecal microbial composition here. In addition, we show that the microbiota of ruminants and Perissodactyla is more similar within the respective (sub-)order than within Carnivora. This results in a lower minimum number of samples that need to be analyzed to decipher the total fecal microbial diversity. For most of the bacterial families and animal species studied, our results show larger deviations when only a few samples (n = 3 or 6) are considered. In general, these deviations become smaller when 10 samples or more are considered and should thus be sufficient to provide a good insight into the fecal microbiota.

For further research, it will be interesting to investigate whether the greater variability of the Carnivora microbiota also applies in short-term time series analyses of a few days and which bacterial families remain constant or contribute to daily fluctuations in the fecal microbial composition.

Methods

Sample collection

Between April 2018 and August 2020, 621 samples were taken from 31 carnivore and herbivore species in a total of 20 German zoos (see Additional file 1). Non-invasive sampling was mostly performed during the daily cleaning routines of the enclosures in cooperation with the keepers. The samples were collected across four animal (sub-)orders, including Canoidea and Feloidea as representatives of the Carnivora, as well as Perissodactyla and Artiodactyla (only Ruminantia) as herbivores. For each species, a minimum of five samples across at least three different zoos was collected (except for *Vulpes lagopus*, *Equus zebra* and *Panthera onca*) When individual differentiation was not possible, fresh samples were collected from different locations in the enclosure to increase the likelihood that the samples are derived from different individuals. Only fresh fecal samples of different individuals were collected in previously disinfected 50 mL centrifuge tubes using sterile inoculation loops. In the next step, a subsample was taken from the center of the feces and transferred to a sterile 2 mL cryotube, which was then immediately stored in liquid nitrogen. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed by the zoos. For further processing, the samples were delivered to StarSEQ GmbH in Mainz, Germany. Here, the samples were preprocessed with the Precellys® Evolution Homogenizer (Bertin Instruments, Rockville, USA) and DNA extraction was performed using the QIAamp® PowerFecal DNA Kit (Qiagen, Hilden, Germany). The DNA concentration in all extracts was measured using a NanoDrop spectrophotometer (Thermofisher, Massachusetts, USA).

16S rRNA gene sequencing and data processing

PCR amplicons for the V3–V4 region of the 16S rRNA gene were generated with primer pair 341F and 806R. Pooled amplicons were sequenced with the Illumina MiSeq 2 × 250 v3 kit for 600 cycles at StarSEQ GmbH. To control for sequencing quality, a 25% PhiX control library was added to the run. Samples were processed following the QIIME 2 [70] pipeline. After demultiplexing, DADA2 [71] was used to call amplicon sequence variants (ASVs) which reflect the biological sequence without clustering similar sequences on a given threshold. A phylogenetic tree was inferred for all sequences based on a sequence alignment generated by MAFFT and low-abundant ASV's that occurred less than 10 times in the total data set as well as chloroplast and mitochondrial sequences were removed from the dataset. The taxonomic assignment of ASVs was performed using a pre-trained naive Bayes classifier [72] based on SILVA 138 full-length sequences [73]. The following statistics were performed in R version 3.6.3 [74] using the packages vegan [75] and FSA [76]. To test for differences in the taxonomic composition between the four mammalian (sub-)orders, ANOSIM test was performed on dissimilarity matrices with Bray–Curtis distances. Alpha diversity was determined by Shannon index, the effective number of species (ENS) [77, 78] and richness which were calculated using QIIME2 after rarefying the number of reads per sample to a total of 2,300 reads. Afterwards, differences between groups were tested using the Kruskal–Wallis test, followed by a post-hoc Dunn Test with Bonferroni correction in R. Beta diversity was also calculated in Qiime2 core-metrics on the rarefied ASV table using unweighted and weighted UniFrac distances. Subsequently, a test for homogeneity of dispersion and the Adonis test for differences between groups was performed on the four (sub-) orders as well as on diet type (herbivore, carnivore). To calculate differences in the occurrence of bacterial families within carnivores and herbivores, the coefficient of variation (CV) was calculated for the respective major bacterial families. The coefficient of variation is a measure of relative variability of sample data and is calculated as the ratio of the standard deviation to the mean. An advantage of this measurement is that it is unitless and independent of the data scaling, which makes it particularly well suited to describe the dispersion of a parameter (here the abundances of individual bacteria families). For further analyses, subsets of the taxonomic assignment of the wildebeest, giraffe, plains zebra, tiger, brown bear and lion were created. Samples of the respective species were randomly drawn until a total sample number of 3,6, 10, 15 and 25 was reached. In addition, three replicates were created for each of these subsets.

For the most-abundant bacterial families, the CV was calculated on those replicates. To control for zoo as a possible influencing factor on the fecal microbiota, we performed a multinomial regression model on differential abundances using Songbird [79]. On the one hand we applied the model on the whole dataset setting 'zoo' and 'species' as explanatory variables and evaluated this against a null model. On the other hand, the same regression was performed on a species-specific subset of microbial abundance data set as dependent variable and 'zoo' as explanatory variable. Furthermore, indicator species for each (sub-)order were identified using the indicspecies R package [80]. The IndVal value calculates the associations between species and sites, followed by a permutation significance test ($n = 999$, $\alpha = 0.05$). Indicators were assigned at microbial family level. To create an approximate host phylogeny, the TimeTree database was used on the involved species names.

Abbreviations
SCFA: Short chain fatty acids; ASV: Amplicon sequence variant; ENS: Effective number of species; CV: Coefficient of variation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42523-021-00141-0.

Additional file 1. Sample metadata. Metadata of analyzed fecal samples according to MIMARKS host-associated package (version 5.0).

Additional file 2. Taxonomic assignment. Taxonomic assignment according to SILVA database (version 1.38) provided for each species. Furthermore, the microbial composition for all samples of one representative species per analyzed (sub-)order is shown. Those representatives are the brown bear (Canoidea), lion (Feloidea), wildebeest (Ruminantia) and plains zebra (Perissodactyla).

Additional file 3. Randomized datasets on the most abundant bacterial families in selected herbivore and carnivore species. For each species (plains zebra, giraffe, wildebeest, brown bear, lion, tiger) random sample ID's were chosen for n = 3, 6, 10, 15, 20 and 25 samples to calculate the coefficient of variation for the bacterial families.

Additional file 4. Statistical tests on alpha diversity between herbivores and carnivores. Results of the significance tests for richness, Shannon index and effective number of species tested on Canoidea, Feloidea, Perissodactyla and Ruminantia. Non-parametric Kruskal-Wallis rank sum test and post-hoc Dunn test were used to analyze for differences between those groups.

Additional file 5. Indicator species results. The data describes the results of the indicator species analysis. A and B values as well as the respective p-value are presented for each microbial indicator. The analysis is based on microbial family level.

Acknowledgements
We gratefully acknowledge support from the directors, curators and especially keepers of Cologne Zoo, Duisburg Zoo, Erlebniszoo Hannover, Frankfurt Zoo, Munich Hellabrunn Zoo, Neunkircher Zoo, Opel-Zoo Kronberg, Schwerin Zoo, Tierpark Berlin, Tiergarten Nuremberg, Vivarium Darmstadt, Wilhelma Stuttgart, Zoo Dortmund, Zoo Heidelberg, Zoo Krefeld, Zoo Landau in der Pfalz, Zoo Leipzig, Zoologischer Garten Berlin, Zoom Erlebniswelt Gelsenkirchen, Zoo Neuwied and Zoo Wuppertal who enabled this project.

Authors' contributions
Conceptualization and Study design: FZ, ALB and PWD; Writing original draft: FZ; Data collection: FZ; Bioinformatics and statistical analysis: FZ; Data interpretation and Visualization: FZ, ALB and PWD; Project administration and Funding acquisition: PWD; Review and editing: FZ, ALB and PWD. All authors read and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. F. Zoelzer, A.L. Burger and P.W. Dierkes were supported by Opel-Zoo Foundation Professorship in Zoo Biology from the "von Opel Hessische Zoostiftung".

Availability of data materials
Raw amplicon sequencing data have been deposited on NCBI's SRA (sequence read archive; accession PRJNA716130). All other data generated or analyzed during the current study are included in the manuscript and its additional files. Reviewer Link: https://dataview.ncbi.nlm.nih.gov/object/PRJNA716130?reviewer=nddpiikt6t8ivk912j6kmi1nib

Declarations

Ethical approval and consent to participate
All procedures were performed in accordance with international guidelines and regulations for the use of animals in research.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 28 April 2021 Accepted: 18 October 2021
Published online: 04 November 2021

References

1. Koh A, Vadder DF, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. Cell. 2016;165:1332–45.
2. Ríos-Covián D, Ruas-Madiedo P, Margolles A, Gueimonde M, Reyes-Gavilán DLCG, Salazar N. Intestinal short chain fatty acids and their link with diet and human health. Front Microbiol. 2016;7:185.
3. Sanna S, Zuydam VNR, Mahajan A, Kurilshikov A, Vila AV, Võsa U, et al. Causal relationships between gut microbiome, short-chain fatty acids and metabolic diseases. Nat Genet. 2019;51:600–5.
4. O'Hara E, Neves ALA, Song Y, Guan LL. The role of the gut microbiome in cattle production and health: driver or passenger? Annu Rev Anim Biosci. 2020;8:1999–2220.
5. Clemmons BA, Martino C, Schneider LG, Lefler J, Embree MM, Myer PR. Temporal stability of the ruminal bacterial communities in beef steers. Sci Rep. 2019;9:9522.
6. Huang S, Ji S, Wang F, Huang J, Alugongo GW, Li S. Dynamic changes of the fecal bacterial community in dairy cows during early lactation. AMB Express. 2020;10:167.
7. Wang X, Tsai T, Deng F, Wei X, Chai J, Knapp J, et al. Longitudinal investigation of the swine gut microbiome from birth to market reveals stage and growth performance associated bacteria. Microbiome. 2019;7:109.
8. Knecht D, Cholewinska P, Jankowska-Makosa A, Czyz K. Development of swine's digestive tract microbiota and its relation to production indices—a review. Animals. 2020;10:527.
9. Guevarra RB, Lee JH, Lee SH, Seok MJ, Kim DW, Kang BN, et al. Piglet gut microbial shifts early in life: causes and effects. J Anim Sci Biotechnol. 2019;10:1.
10. Mamun MAA, Sandeman M, Rayment P, Brook-Carter P, Scholes E, Kasinadhuni N, et al. Variation in gut bacterial composition is associated with *Haemonchus contortus* parasite infection of sheep. Anim Microbiome. 2020;2:3.
11. Tanca A, Fraumene C, Manghina V, Palomba A, Abbondio M, Deligios M, et al. Diversity and functions of the sheep faecal microbiota: a multi-omic characterization. Microb Biotechnol. 2016;10:541–54.
12. Gibson KM, Nguyen BN, Neumann LM, Miller M, Buss P, Daniels S, et al. Gut microbiome differences between wild and captive black rhinoceros—implications for rhino health. Sci Rep. 2019;9:7570.
13. Brice KL, Trivedi P, Jeffries TC, Blyton MDJ, Mitchell C, Singh BK, et al. The koala (*Phascolarctos cinereus*) faecal microbiome differs with diet in a wild population. PeerJ. 2019. https://doi.org/10.7717/peerj.6534.
14. Cheng Y, Fox S, Pemberton D, Hogg C, Papenfuss AT, Belov K. The tasmanian devil microbiome—implications for conservation and management. Microbiome. 2015;3:76.
15. Matthews C, Crispie F, Lewis E, Reid M, O'Toole PW, Cotter PD. The rumen microbiome: a crucial consideration when optimising milk and meat production and nitrogen utilization efficiency. Gut Microbes. 2019;10:115–32.
16. Ley RE, Hamady M, Lozupone C, Turnbaugh P, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. Science. 2008;320:1647–51.
17. Youngblut ND, Reischer GH, Walters W, Schuster N, Walzer C, Stalder G, et al. Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. Nat Commun. 2019;10:2200.
18. Milani C, Alessandri G, Mancabelli L, Mangifesta M, Lugli GA, Viappiani A, et al. Multi-omics approaches to decipher the impact of diet and host physiology on the mammalian gut microbiome. Appl Environ Microbiol. 2020;86:23.
19. Kartzinel TA, Hsingb JC, Musilic PM, Browna BRP, Pringl RM. Covariation of diet and gut microbiome in african megafauna. PNAS. 2019;116:47.

20. Nelson TM, Apprill A, Mann J, Rogers TL, Brown MV. The marine mammal microbiome: current knowledge and future directions. Microbiol Aust. 2015;36:8–13.
21. Kueneman JG, Wegener Parfrey L, Woodhams DC, Archer HM, Knight R, McKenzie VJ. The amphibian skin-associated microbiome across species, space and life history stages. Mol Ecol. 2014;6:1238–50.
22. Song SJ, Sanders JG, Delsuc F, Metcalf J, Amato K, Taylor MW, et al. Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. MBio. 2020;11:1.
23. Stevens CE, Hume ID. Comparative physiology of the vertebrate digestive system. 2nd ed. Cambridge: Cambridge University Press; 1995.
24. Mackie RI. Mutualistic fermentative digestion in the gastrointestinal tract: diversity and evolution. Integr Comp Biol. 2002;42:319–26.
25. Buffington CA. Dry foods and risk of disease in cats. Can Vet J. 2008;49:561–3.
26. Verbrugghe A, Hesta M, Daminet S, Janssens GP. Nutritional modulation of insulin resistance in the true carnivorous cat: a review. Crit Rev Food Sci Nutr. 2012;52:172–82.
27. Suchodolski JS, Camacho J, Steiner JM. Analysis of bacterial diversity in the canine duodenum, jejunum, ileum, and colon by comparative 16S rRNA gene analysis. FEMS Microbiol Ecol. 2008;66:567–78.
28. Douglas AE. Fundamentals of microbiome science. How microbes shape animal biology. New Jersey: Princeton University Press; 2018.
29. Mitchell PC. On the intestinal tract of mammals. Trans Zool Soc Lond. 1905;17:437–536.
30. Kumar S, Stecher G, Suleski M, Hedges SB. Timetree: a resource for timelines, timetrees, and divergence times. Mol Biol Evol. 2017;34:1812–9.
31. Reau LAJ, Suen G. The *Ruminococci*: key symbionts of the gut ecosystem. Microbiol. 2018;56:199–208.
32. Yatsunenko T, Rey F, Manary M, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486:222–322.
33. Obregon-Tito A, Tito R, Metcalf J, Sankaranarayanan K, Clemente JC, Ursell LK, et al. Subsistence strategies in traditional societies distinguish gut microbiomes. Nat Commun. 2015;6:6505.
34. The SE, Lachnospiraceae F. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The prokaryotes. Berlin: Springer; 2014. p. 197–201.
35. Vacca M, Celano G, Calabrese FM, Portincasa P, Gobbetti M, Angelis DM. The controversial role of human gut *Lachnospiraceae*. Microorganisms. 2020;8:573.
36. Biddle A, Stewart L, Blanchard J, Leschine S. Untangling the genetic basis of fibrolytic specialization by *Lachnospiraceae* and *Oscillospiraceae* in diverse gut communities. Diversity. 2013;5:627–40.
37. Iorio DBR, Rocchetti MT, Angelis DM, Cosola C, Marzocco S, Micco DL, et al. Nutritional therapy modulates intestinal microbiota and reduces serum levels of total and free indoxyl sulfate and p-cresyl sulfate in chronic kidney disease (Medika Study). J Clin Med. 2019;8:1424.
38. Kim E, Kim DB, Park JY. Changes of mouse gut microbiota diversity and composition by modulating dietary protein and carbohydrate contents: a pilot study. Prev Nutr Food Sci. 2016;21:57.
39. Vital M, Gao J, Rizzo M, Harrison T, Tiedje JM. Diet is a major factor governing the fecal butyrate-producing community structure across Mammalia, Aves and Reptilia. ISME J. 2015;9:832–43.
40. The OI, Fusobacteriaceae F. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The prokaryotes. Berlin: Springer; 2014. p. 109–32.
41. Zhu L, Wu Q, Deng C, Zhang M, Zhang C, Chen H, et al. Adaptive evolution to a high purine and fat diet of carnivorans revealed by gut microbiomes and host genomes. Environ Microbiol. 2018;20:1711–22.
42. Bermingham EN, Maclean P, Thomas DG, Cave NJ, Young W. Key bacterial families (*Clostridiaceae*, *Erysipelotrichaceae* and *Bacteroidaceae*) are related to the digestion of protein and energy in dogs. PeerJ. 2017. https://doi.org/10.7717/peerj.3019.
43. Bragg M, Freeman EW, Lim HC, Songsasen N, Muletz-Wolz CR. Gut microbiomes differ among dietary types and stool consistency in the captive red wolf (*Canis rufus*). Front Microbiol. 2020;11:2777.
44. Kerr KR, Forster G, Dowd SE, Ryan EP, Swanson KS. Effects of dietary cooked navy bean on the fecal microbiome of healthy companion dogs. PLoS ONE. 2013;8:9.
45. Panasevich MR, Kerr KR, Dilger RN, Fahey GC, Guérin-Deremaux L, Lynch GL, et al. Modulation of the faecal microbiome of healthy adult dogs by inclusion of potato fiber in the diet. Br J Nutr. 2015;113:125–33.
46. Vazquez-Baeza Y, Hyde ER, Suchodolski JS, Knight R. Dog and human inflammatory bowel disease rely on overlapping yet distinct dysbiosis networks. Nat Microbiol. 2016;1:16177.
47. Guo X, Lei H, Zhang K, Ke F, Song C. Diversification of animal gut microbes and NRPS gene clusters in some carnivores, herbivores and omnivores. Biotechnol Equip. 2020;34:1280–7.
48. Nishida AH, Ochman H. Rates of gut microbiome divergence in mammals. Mol Ecol. 2018;27:1884–97.
49. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science. 2011;332:970–4.
50. Yoder A, Burns MM, Zehr S, Delefosse T, Veron G, Goodman SM, et al. Single origin of malagasy carnivora from an african ancestor. Nature. 2003;421:734–7.
51. Flynn JJ, Finarelli JA, Zehr S, Hsu J, Nedbal MA. Molecular phylogeny of the carnivora (mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst Biol. 2005;54:317–37.
52. Eizirik E, Murphy WJ. Carnivores (Carnivora). In: Hedges SB, Kumar S, editors. The timetree of life. Oxford: Oxford University Press; 2009. p. 504–7.
53. Delisle I, Strobeck C. A phylogeny of the Caniformia (order Carnivora) based on 12 complete protein-coding mitochondrial genes. Mol Phylogenet Evol. 2005;37:192–201.
54. Arnason U, Gullberg A, Janke A, Kullberg M. Mitogenomic analyses of caniform relationships. Mol Phylogenet Evol. 2007;45:863–74.
55. Nyakatura K, Bininda-Emonds OR. Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. BMC Biol. 2012;10:12.
56. Zhao S, Lieberman TD, Poyet M, Kauffman KM, Gibbons SM, Groussin M, et al. Adaptive evolution within gut microbiomes of healthy people. Cell Host Microb. 2019;25:656–67.
57. Groussin M, Mazel F, Alm EJ. Co-evolution and co-speciation of host-gut bacteria systems. Cell. 2020;28:12–22.
58. Groussin M, Mazel F, Alm EJ. Co-evolution and co-speciation of host-gut bacteria systems. Cell Host Microb. 2020;28:12–22.
59. Lyu T, Liu G, Zhang H, Wang L, Zhou S, Dou H, et al. Changes in feeding habits promoted the differentiation of the composition and function of gut microbiotas between domestic dogs (*Canis lupus familiaris*) and gray wolves (*Canis lupus*). AMB Express. 2018;8:1.
60. Butowski CF, Thomas DG, Young W, Cave NJ, McKenzie CM, Rosendale DI, et al. Addition of plant dietary fibre to a raw red meat high protein, high fat diet, alters the faecal bacteriome and organic acid profiles of the domestic cat (*Felis catus*). PLoS ONE. 2019;14:5.
61. Wernimont SM, Radosevich J, Jackson MI, Ephraim E, Badri DV, MacLeay JM, et al. The effects of nutrition on the gastrointestinal microbiome of cats and dogs: Impact on health and disease. Front Microbiol. 2020;11:1266.
62. Zahal AO, Valdes EV, McBride BW. Analysis of the distal gut bacterial community by 454-pyrosequencing in captive giraffes (*Giraffa camelopardalis*). Zoo Biol. 2016;35:42–50.
63. Becker AA, Hesta M, Hollants J, Janssens GPJ, Huys G. Phylogenetic analysis of faecal microbiota from captive cheetahs reveals underrepresentation of *Bacteroidetes* and *Bifidobacteriaceae*. BMC Microbiol. 2014;14:43.
64. Uddin W, Menke S, Melzheimer J, Thalwitzer S, Heinrich S, Wachter B, Sommer S. Gut microbiomes of free-ranging and captive Namibian cheetahs: diversity, putative functions and occurrence of potential pathogens. Mol Ecol. 2017;26:5515–27.
65. Mittal P, Saxena R, Gupta A, Mahajan S, Sharma VK. The gene catalog and comparative analysis of gut microbiome of big cats provide new insights on *Panthera* species. Front Microbiol. 2020;11:1012.
66. He F, Liu D, Zhang L, Zhai J, Ma Y, Yu Y, et al. Metagenomic analysis of captive amur tiger faecal microbiome. BMC Vet Res. 2018;14:379.
67. Jiang D, He X, Valitutto M, Chen L, Xu Q, Yao Y, Hou R, Wang H. Gut microbiota composition and metabolomic profiles of wild and captive Chinese monals (*Lophophorus lhuysii*). Front Zool. 2020;17:1–12.
68. Chi X, Gao H, Wu G, Qin W, Song P, Wang L, Chen J, Cai Z, Zhang T. Comparison of gut microbiota diversity between wild and captive bharals (*Pseudois nayaur*). BMC Vet Res. 2019;15:1–18.

69. McKenzie VJ, Song SJ, Delsuc F, Prest TL, Oliverio AM, Korpita TM, et al. The effects of captivity on the mammalian gut microbiome. Integr Comp Biol. 2017;57:690–704.
70. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:1091.
71. Callahan B, McMurdie P, Rosen M, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3.
72. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome. 2018;6:90.
73. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:590–6.
74. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria;2020.
75. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community Ecology Package. 2019. https://CRAN.R-project.org/package=vegan. Accessed 25 Mar 2021.
76. Ogle DH, Wheeler P, Dinno A. FSA: Fisheries Stock Analysis. R package.2020. https://github.com/droglenc/FSA. Accessed 28 Jul 2021.
77. Jost L. Entropy and diversity. Oikos. 2006;113:110–6.
78. Jost L. Partitioning diversity into independent alpha and beta components. Ecology. 2007;88:2427–39.
79. Morton JT, Marotz C, Washburne A, Silverman J, Zaramela LS, Edlund A, Zengler K, Knight R. Establishing microbial composition measurement standards with reference frames. Nat Commun. 2019;10:2719.
80. Cáceres MD, Legendre P. Associations between species and groups of sites: indices and statistical inference. Ecology. 2009;90:3566–74.

Publisher's Note

Publication B

Declaration of author contributions to the publication Time series cluster analysis reveals individual assignment of microbiota in captive tiger (*Panthera tigris*) and wildebeest (*Connochaetes taurinus*):

Status: *accepted*
Starting with *accepted:* name of journal: Ecology and Evolution

Contributing authors:

Franziska Zölzer (FZ, PhD candidate)
Sebastian Schneider (SS)
Paul Wilhelm Dierkes (PWD)

What are the contributions of the doctoral candidate and his co-authors?

(1) Concept and design
Doctoral candidate FZ: 90
PWD: 10

(2) Conducting tests and experiments
Doctoral candidate FZ: 100

(3) Compilation of data sets and figures
Doctoral candidate FZ: 100

(4) Analysis and interpretation of data
Doctoral candidate FZ: 70
SS: 20
PWD: 10

(5) Drafting of manuscript
Doctoral candidate FZ: 80
SS: 10
PWD: 10

I hereby certify that the information above is correct.

_____ _____

Date and place Signature doctoral candidate

_____ _____

Date and place Signature supervisor

Received: 20 December 2022 | Revised: 18 April 2023 | Accepted: 24 April 2023

DOI: 10.1002/ece3.10066

Ecology and Evolution Open Access WILEY

RESEARCH ARTICLE

Time series cluster analysis reveals individual assignment of microbiota in captive tiger (*Panthera tigris*) and wildebeest (*Connochaetes taurinus*)

Franziska Zoelzer | **Sebastian Schneider** | **Paul Wilhelm Dierkes**

Bioscience Education and Zoo Biology, Goethe University Frankfurt, Frankfurt am Main, Germany

Correspondence
Franziska Zoelzer, Bioscience Education and Zoo Biology, Goethe University Frankfurt, Max-von-Laue-Str.13, 60438 Frankfurt, Germany.
Email: zoelzer@bio.uni-frankfurt.de

Funding information
Vereinigung von Freunden und Förderern der Goethe-Universität e.V.; von Opel Hessische Zoostiftung

Abstract

Fecal microbiota variability and individuality are well studied in humans and also in farm animals (related to diet- or disease-specific influences), but very little is known for exotic zoo-housed animals. This includes a wide range of species that differ greatly in microbiota composition and variation. For example, herbivorous species show a very similar and constant fecal microbiota over time, whereas carnivorous species appear to be highly variable in fecal microbial diversity and composition. Our objective was to determine whether species-specific and individual-specific clustering patterns were observed in the fecal microbiota of wildebeest (*Connochaetes taurinus*) and tigers (*Panthera tigris*). We collected 95 fecal samples of 11 animal individuals that were each sampled over eight consecutive days and analyzed those with Illumina MiSeq sequencing of the V3–V4 region of the 16SrRNA gene. In order to identify species or individual clusters, we applied two different agglomerative hierarchical clustering algorithms – a community detection algorithm and Ward's linkage. Our results showed that both, species-specific and individual-specific clustering is possible, but more reliable results were achieved when applying dynamic time warping which finds the optimal alignment between different time series. Furthermore, the bacterial families that distinguish individuals from each other in both species included daily occurring core bacteria (e.g., Acidaminococcaceae in wildebeests or Clostridiaceae in tigers) as well as individual dependent and more fluctuating bacterial families. Our results suggest that while it is necessary to consider multiple consecutive samples per individual, it is then possible to characterize individual abundance patterns in fecal microbiota in both herbivorous and carnivorous species. This would allow establishing individual microbiota profiles of animals housed in zoos, which is a basic prerequisite to quickly detect deviations and use microbiome analysis as a non-invasive and cost-effective tool in animal welfare.

KEYWORDS
16S rRNA gene, abundance pattern, cluster analysis, microbiota

TAXONOMY CLASSIFICATION
Microbiomics

1 | INTRODUCTION

In recent years, much research has been conducted to analyze the composition and diversity of gastrointestinal microorganisms and their impact and interaction with the host organism for various animal species as well as for phylogenetic and dietary groups (Koh et al., 2016; Ley et al., 2008; Milani et al., 2020; Nelson et al., 2015; Sanna et al., 2019; Youngblut et al., 2019). In addition to simply characterizing the species-specific microbiome, other questions include whether microbiomes remain stable or are subject to fluctuations over time, the frequency at which these fluctuations occur, and which bacterial taxa are affected by them. Particularly in the early years of microbiome research, some studies proposed an individual long-term stable microbiome in humans, with some bacterial taxa being persistent over a yearlong sampling interval (Björk et al., 2019; Faith et al., 2013; Hildebrand et al., 2021; Martínez et al., 2013; Schloissnig et al., 2013). In contrast, other longitudinal studies suggest that the individual human microbiome is highly variable over time (Caporaso et al., 2011; Olsson et al., 2022). Especially intra-individual variation seems to outweigh inter-individual variation with regard to daily fluctuation, as the majority of bacterial taxa show great shifts in abundance. Furthermore, high-abundant taxa seem to express less variation than low-abundant taxa and the extent of variation is constant over time (Vandeputte et al., 2021; Zoelzer et al., 2021). In contrast, there are few time-series data on the natural variation in the fecal microbiome in various animal species. The studies conducted here (primarily on farm animals) mostly refer to the influence of dietary changes (Butowski et al., 2019; Lyu et al., 2018), impact and courses of diseases (Ayoub et al., 2022; Mamun et al., 2020) or the development of juvenile animals (Amin & Seifert, 2021; Guevarra et al., 2019; Wang et al., 2019). Only some studies refer to natural oscillations of different bacterial taxa (Björk et al., 2022; Rojas et al., 2023), finding evidence for diurnal rhythmicity in microbial diversity and composition, e.g. in dairy cows (Shaani et al., 2018) or meerkats (Risely et al., 2021).

For this reason, our study examined the daily course of the fecal microbiota in two animal species with fundamentally different digestive systems, namely tiger (*Panthera tigris*) and wildebeest (*Connochaetes taurinus*) in order to identify natural microbial abundance patterns. These two species are especially suited for time series analyses for several reasons. First, ruminants depend on bacterial fermentation to digest cellulose and, therefore, show a high microbial fecal diversity whereas carnivores with a less complex digestive system have a lower microbial diversity (Guo et al., 2020; Milani et al., 2020; Vital et al., 2015). Second, ruminants seem to have a high similarity in their microbiota, whereas the microbiota of felids is highly variable (Petri et al., 2013; Snelling et al., 2019; Zoelzer et al., 2021). Here, we survey the fecal microbiotas of tigers and wildebeest and determine whether they are species-and individual-specific, and whether specific bacterial families can help distinguish the different groupings.

Cluster algorithms are now widely used not only in social (Hoffman et al., 2018) or technological (Faloutsos et al., 1999) but also in biological and health-related (Bhar et al., 2022; Fell & Wagner, 2000) network analysis. The fecal microbiota can also be considered a network in which the microbial composition of a sample represents the nodes and the distance between the samples and the respective edges of the network. In this network, closely connected nodes form a community that shares only a few edges with neighboring communities. To identify microbial communities, we used two different agglomerative hierarchical clustering algorithms—a community detection algorithm and Ward's linkage. Community detection tries to find groups of nodes that are highly connected to each other forming a cluster while Ward's linkage is based on the distance between clusters, aiming to minimize the variance within a cluster (Newman, 2004; Ward, 1963). In order to enhance the clustering results, we utilized dynamic time warping to synchronize the time series datasets and adjust for any discrepancies in sampling points. As a few studies have used dynamic time warping followed by clustering algorithms to measure the similarity between individual time series or to identify the abundance pattern of bacterial taxa over time (Armoni & Borenstein, 2022; Muinck & Trosvik, 2018; Ponziani et al., 2022), we would like to extend this approach. After clustering the individual time series, we try to identify the correct species or individual based on the microbiota composition.

Applying this approach to individual time series data of two species, we developed two hypotheses. First, we expect a clear species-specific clustering due to the previously described significant differences in fecal microbiota composition and diversity between carnivore and herbivore species. Second, individual-specific microbiota clustering works more reliably in tigers than in wildebeests because the herbivore microbiota is too stable within individuals to identify characteristic individual variation. Nevertheless, if not only species-specific but also individual abundance patterns can be detected in different bacterial families and thus an individual can be identified over a time variable microbiota, this leads to several future application areas. In zoo animal husbandry, animal welfare plays a major role, e.g. in veterinary care. Using an individual fecal microbiota profile, deviations from natural fluctuations can be detected easily, inexpensively, and non-invasively. This would provide an additional and easily accessible monitoring tool for the health of zoo animals.

2 | MATERIALS AND METHODS

2.1 | Sample collection

In the period from May 2018 to November 2020, 95 fecal samples were collected from 11 individuals, six tiger, and five wildebeests, housed in five German zoos (Table S1). The collection plan included a time series of eight consecutive days in which one fecal sample per day and individual was collected if available (Table A1). At least two individuals per species were sampled from each zoo. Due to sample availability, we included a total of two time series from three individuals (one tiger and two wildebeests) in the analysis to capture

possible temporal variation in the microbiota. Sampling was performed non-invasively by animal caretakers during the daily enclosure cleaning routine. The samples were immediately transferred to sterile cryotubes and stored in liquid nitrogen until further processing. We followed the EAZA research standard guidelines for the care and use of animals.

Further preparation of the samples was carried out by StarSEQ GmbH in Mainz, Germany. First, the samples were homogenized (Precellys® Evolution Homogenizer, Bertin Instruments, Rockville, USA) and subsequently DNA extraction was performed using the QIAamp® PowerFecal DNA Kit (Qiagen, Hilden, Germany). A NanoDrop spectrophotometer (Thermofisher, Massachusetts, USA) was used to measure the DNA concentration.

2.2 | 16S rRNA gene sequencing and data processing

At StarSEQ GmbH, the V3–V4 region of the 16S rRNA gene was sequenced by a dual-index strategy based on the protocol of (Caporaso et al., 2012) with minor modifications. Amplicons were generated by a single-step of 33 cycles using the primer combination 341f and 806bR (Apprill et al., 2015; Takahashi et al., 2014). The final library was sequenced on the Illumina MiSeq platform in paired-end mode (300 nt) with a 25% PhiX control library. Samples were analyzed according to the QIIME 2 pipeline (Bolyen et al., 2019). As described in previous work (Zoelzer et al., 2021), DADA2 (Callahan et al., 2016) was applied to determine amplicon sequence variants (ASVs), and a phylogenetic tree was constructed for all sequences using MAFFT sequence alignments (Katoh et al., 2002). Low abundance features that are covered by less than 10 sequences, chloroplast, and mitochondrial sequences were removed from the dataset. Taxonomic assignment of ASVs was performed using a pre-trained Naive Bayes classifier (Bokulich et al., 2018) based on the SILVA 138 full-length database (Quast et al., 2013). The following statistical analyses were performed in R version 3.4.1 (R Core Team, 2020) using the packages vegan (Oksanen et al., 2019) and FSA (Ogle et al., 2022). To test for individual and interspecific differences in the microbial composition, ANOSIM test was performed on dissimilarity matrices with Bray-Curtis distances. Differences in microbial richness between and within species were tested using ANOVA, followed by a post hoc pairwise t-Test with Bonferroni correction.

Most methods for time series analysis require that the time intervals between samples are equidistant, which can be difficult if the animals (mainly carnivores) do not reliably defecate on a daily basis. In this study, we solved the problem using a combined approach of dynamic time warping followed by a clustering algorithm to ensure the comparability of individual time series data and contrast the results with a method using k-nearest-neighbor as the classifier (Figure 1). The whole clustering pipeline is implemented in Matlab version 9.11 (The MathWorks Inc., 2020) using the software CASE (Schneider et al., 2022). Accordingly, we applied two different approaches to cluster the data (Figure 2). On the one hand, all samples were clustered individually (Single clustering). This was done by using k-nearest-neighbor-search (Friedman et al., 1977) which determines for each object Ni (fecal sample) its k-nearest neighbors with the smallest Euclidean distance and creates a distance matrix which serves as an input for the subsequent creation of a Jaccard similarity matrix. On the other hand, samples of the same individual were combined as a time series (Time series clustering) and first compared using dynamic time warping (DTW). The DTW algorithm (Paliwal et al., 1982; Sakoe & Chiba, 1978) is designed to compare two time series by calculating the Euclidean distance between them. To achieve this, each element of the two time series (or columns for matrices) is repeated until the Euclidean distance is minimized. The output is a distance matrix, which again is used to create a Jaccard similarity matrix.

In both cases, Ward's linkage and a community detection algorithm are applied to the dataset. Ward's linkage is a type of hierarchical cluster analysis technique that involves evaluating the distance between two clusters through the linkage function. This function is computed by measuring the increase in error sum of squares (ESS) that occurs when two clusters are merged into one. Ward's method aims to minimize the increase in ESS during each clustering step by selecting the most appropriate clustering steps (Ward, 1963). Here, the number of clusters was determined automatically by estimating the most consistent cluster solution in the dendrogram. Cluster solutions with two or less clusters were ignored except for the species-specific clustering, as two clusters were to be expected. Furthermore, we used a community detection algorithm (Newman, 2004) as implemented in the software CASE (Schneider et al., 2022). This agglomerative hierarchical clustering algorithm groups vertices into clusters. It results in a hierarchical dendrogram, where at the beginning each vertex is considered as a separate community. As the algorithm progresses, it merges pairs of communities together based on the number of edges connecting their vertices, resulting in clusters with many internal edges and relatively few edges connecting vertices from different clusters (Fortunato, 2010).

To check the reliability of the cluster results, the normalized mutual information (NMI) was calculated. The NMI compares the determined labels of Ward's linkage and community detection with each other and outputs a value between 0 and 1, where 1 corresponds to an optimal match. Additionally, to evaluate the clustering in terms of individual discrimination, the NMI was also calculated against the true label. In this case, the true label refers to either the correct species (wildebeest or tiger) of a sample or the correct individual within a species. This label is used within the pipeline to compare the clustering results with the true and correct results. The true labels could be clearly determined because the associated individual was known when the fecal samples were collected. The features (bacterial families) that best describe the differences between the species and individuals in the dataset were calculated from the true labels using the LASSO algorithm (Tibshirani, 1996). LASSO is a type of shrinkage method that automatically reduces the influence of less relevant features by making them smaller and less significant.

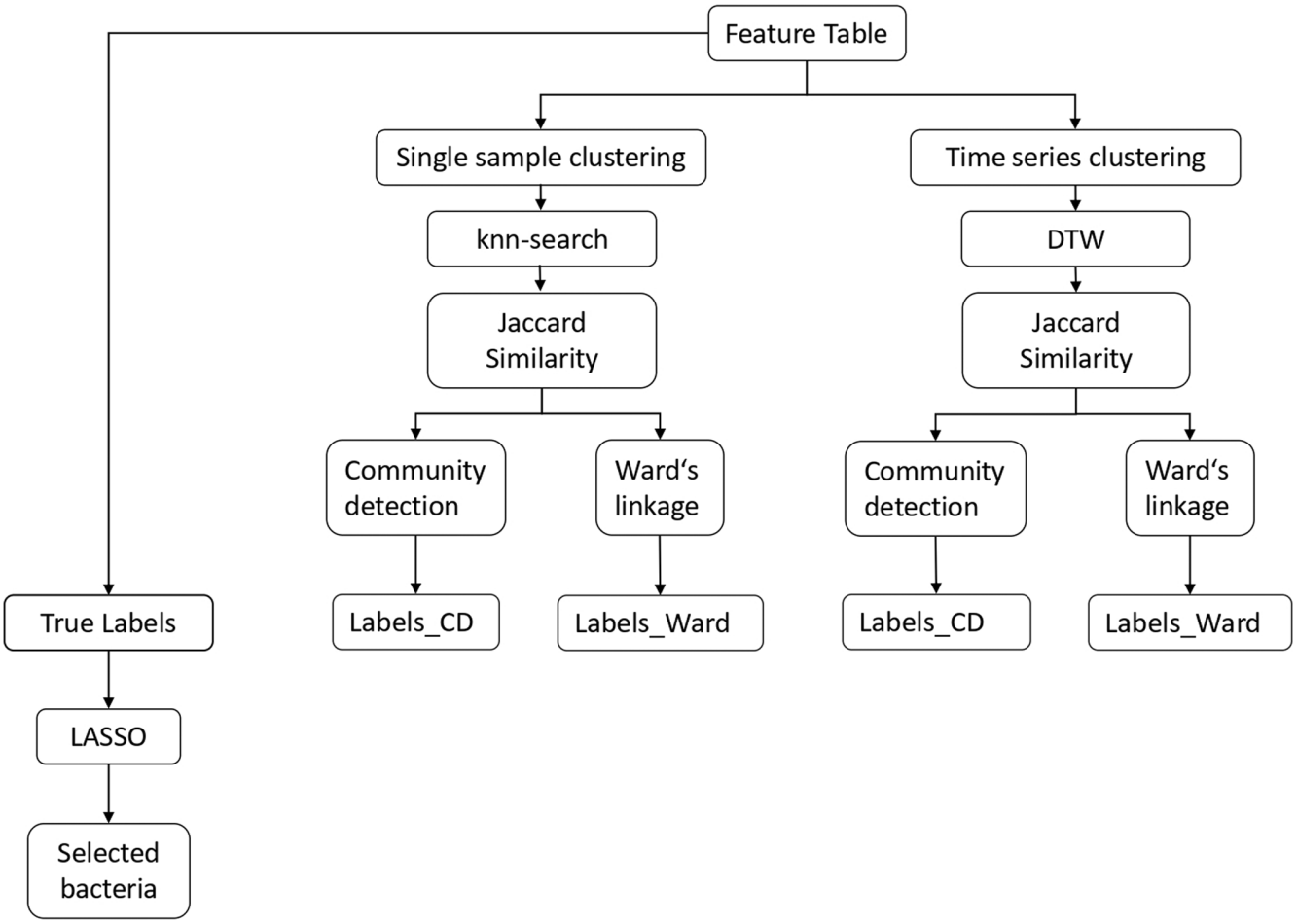

FIGURE 1 Workflow of the clustering pipeline including single and time series clustering. Starting from a feature table, both methods rely on a similarity matrix as input for the two clustering algorithms, Ward's linkage and community detection. As output, the samples within a cluster are characterized by specific labels. In addition, the true labels (known species or individuals) were used as input to the LASSO algorithm to identify individual- and species-specific bacterial families.

Additionally, it can perform variable selection by setting irrelevant variables to zero. The whole pipeline was performed on the entire dataset as well as on the wildebeest- and tiger-specific datasets to test for intraspecific variation.

3 | RESULTS

In total, we analyzed 95 fecal samples of five wildebeests and six tigers, performing Illumina MiSeq paired-end sequencing of the V3–V4 region of the 16S rRNA gene. After preprocessing, the dataset consisted of 5,662,914 sequences (5836–230,928 sequences per sample) with an average of 59,610 sequences per sample. We found a significantly higher species richness in wildebeests (ANOVA statistic: $F = 137.10$, $p < .001$) than in tigers and very different microbial composition between these two species (ANOSIM statistic: $R = .89$, $p < .001$, number of permutations: 999, distance = "bray"). Within the respective species, the individuals analyzed had a significantly different microbiota, with some overlap (Wildebeest ANOSIM statistic: $R = .18$, $p = .001$; Tigers ANOSIM statistic: $R = .29$, $p = .003$, number of permutations: 999, distance = "bray").

3.1 | Species-specific clustering of the microbiota

We found the highest support and strongest clustering of samples by species using dynamic time warping prior to clustering with Ward's linkage. However, when using just single samples without dynamic time warping, species-specific clustering was not observed and samples were given incorrect species assignments (Figure A1). In the best scenario, a species-specific assignment of all individual samples would result in two cluster solutions—wildebeests and tigers. This was not possible with either Ward's linkage ($NMI_{Ward/True} = 0.25$) or community detection algorithm ($NMI_{Com/True} = 0.69$) (Table 1). When applying dynamic time warping to the dataset which synchronizes the time series samples for individuals, the overall quality of clustering increased. Regarding species discrimination, both algorithms reached

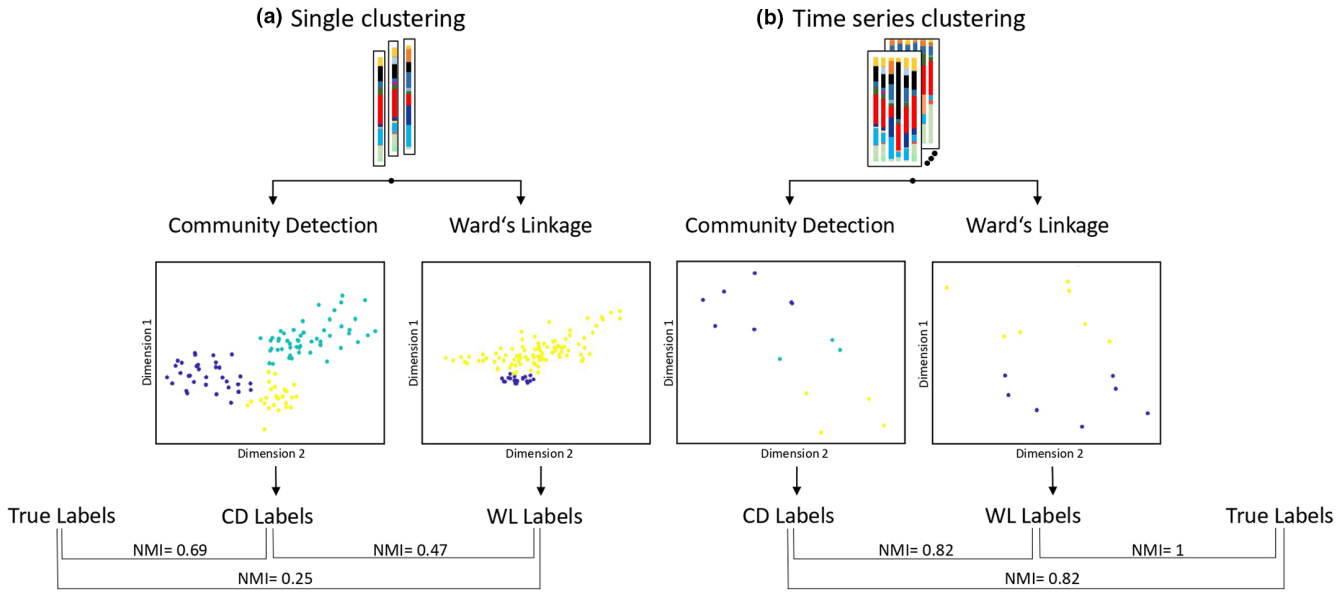

FIGURE 2 Comparison of the two clustering approaches. (a) For the single clustering approach, each sample is considered individually. The microbial composition of each sample as a single data set is used here as input for clustering. (b) For the time series clustering, a whole time series of each individual consisting of up to eight samples are used as input data for the clustering pipeline. In both approaches, the output cluster labels are compared with each other as well as with the true labels (correct species or individual per sample) by calculating the NMI value.

TABLE 1 Clustering results of single and time series clustering.

	NMI (com/Ward)	NMI (Ward/$True_{Ind}$)	NMI (com/$True_{Ind}$)	Number cluster ($Ward_{Ind}$)	Number cluster (Com_{Ind})	NMI (com/Ward)	NMI (Ward/$True_{Spe}$)	NMI (com/$True_{Spe}$)	Number cluster ($Ward_{Spe}$)	Number cluster (Com_{Spe})
Single clustering										
Total						0.47	0.25	0.69	2	3
Tiger	0.68	0.18	0.01	4	2					
Wilde-beest	0.81	0.39	0.54	3	4					
Time series clustering										
Total						0.82	1	0.82	2	3
Tiger	0.64	0.51	0.95	3	7					
Wilde-beest	1	0.82	0.82	6	6					

Note: The table shows the reliability of both algorithms (NMI) to each other and against the true species/individual label. Furthermore, the number of calculated clusters for species detection (Spe) as well as for individual discrimination (Ind) is represented. The upper part of the table represents the results for the single clustering approach and the lower part shows the results for the time series clustering. For individual discrimination, a minimum of three cluster solutions is required while for species detection results with two cluster solutions are allowed.

comparable results ($NMI_{Ward/Com}$ = 0.82). While community detection ($NMI_{Com/True}$ = 0.82) resulted in three clusters, in which all wildebeest individuals fall into one cluster and the tigers were split up into two clusters, Ward's linkage led to a correct assignment of species ($NMI_{Ward/True}$ = 1) (Figure A1).

3.2 | Individual-specific clustering of the microbiota

We then sought out to test whether our clustering approaches could detect individual-specific clustering of the microbiota in tigers and wildebeest. Due to significant differences in microbial diversity and composition between the two species, the whole dataset was separated into species-specific sets to analyze individual differences within both species. For individual differentiation within the species, the optimal result would consist of six cluster solutions for the tigers and five for the wildebeests, one per individual. Comparing the labels determined by single clustering with the true individual labels, a low level of consistency was found (Table 1). For wildebeest, we found that the best results were achieved using community detection ($NMI_{Com/True}$ = 0.54) which resulted in four different clusters (Figure A1). In contrast, a correct individual assignment within the tigers ($NMI_{Ward/True}$ = 0.18, $NMI_{Com/True}$ = 0.01) was not possible using

the single clustering approach. When dynamic time warping was applied prior to clustering, individual discrimination was improved. For wildebeests, both algorithms ($NMI_{Ward/Com}=1$) led to the same results. Except for two individuals, each time series was assigned to a separate cluster here. For tigers, individual-specific clustering led to more reliable results, especially when applying a community detection algorithm ($NMI_{Ind}=0.95$). In this case, seven clusters were calculated including one individual time series each. Different time series of the same individual, no matter if wildebeest or tiger were classified in different clusters by both algorithms in both clustering algorithms.

3.3 | Identification of bacterial features that explain species differences

Using the LASSO algorithm, bacterial families leading to the specific cluster solution were identified. First, we compared all wildebeests and tigers to identify the taxa that led to a species-specific cluster solution using the known true labels for each species. Compared to the wildebeests, the tiger-specific bacterial families were Clostridiaceae and Fusobacteriaceae (Figure 3a). Both families are core bacteria in all individuals, meaning that they appear in each consecutive sample (Caporaso et al., 2011). On average, Clostridiaceae constituted between 5.75 ± 3.23% and 38.09 ± 14.95% to the microbiota of the seven tiger datasets. Within the microbiota of three individuals (Ind1_Zoo2, Ind2_Zoo2, Ind2.2_Zoo3) this family had a share of more than 24% on average. Fusobacteriaceae constituted between 2.12 ± 1.89% and 25.26 ± 6.64% on average to the tigers' microbiota. This family either occurred in larger proportions of more than 20% of the average microbiota or is represented by very small proportions (<5%). Considering daily time intervals, both species-specific bacterial families seemed to be subject to larger fluctuations as can be seen in Ind2.1_Zoo3. Here, the average proportion of Clostridiaceae increased from 11.81 ± 6.68% to 38.09 ± 14.95% and that of Fusobacteriaceae decreased from 20.46 ± 9.38% to 3.21 ± 3.24% within 2 days.

In contrast to these results, wildebeest-specific bacteria accounted for a much lower average proportion of the individual microbiota (Figure 3b). The major wildebeest-specific bacteria were Methanobacteriaceae and the phylum Spirochaetes. Methanobacteriaceae as a core bacterial family in all individuals occurred on average between 3.28 ± 1.13% and 13.19 ± 4.28% in each dataset. Spirochaetes constituted from 0.95 ± 0.20% to 4.24 ± 1.07% to the wildebeest microbiota and were also a core bacterial phylum in all individuals except for one. Dysgomonadaceae and Clostridiaceae both occurred in <0.2% of all datasets. Nevertheless, Clostridiaceae were a persistent family within three individuals, being present in at least two consecutive sampling days. Dysgomonadaceae, as a core bacterial family in Ind1_Zoo2, constituted on average only 0.15 ± 0.05% to the respective microbiota. Finally, the M2PB4–65 termite group only occurred in minor proportions in both individuals of Zoo1 (Ind1 = 1.13 ± 0.45%, Ind2 = 0.69 ± 0.34%) but act as a core taxon in both of them.

3.4 | Identification of bacterial features that explain individual differences

Second, the same approach was used to identify bacteria that are responsible for the clustering of tiger individuals. The feature selection revealed seven bacterial families: Atopobiaceae, Bacteroidaceae, Clostridiaceae, Enterobacteriaceae, Family XI, Methanobacteriaceae, and Prevotellaceae. Even though

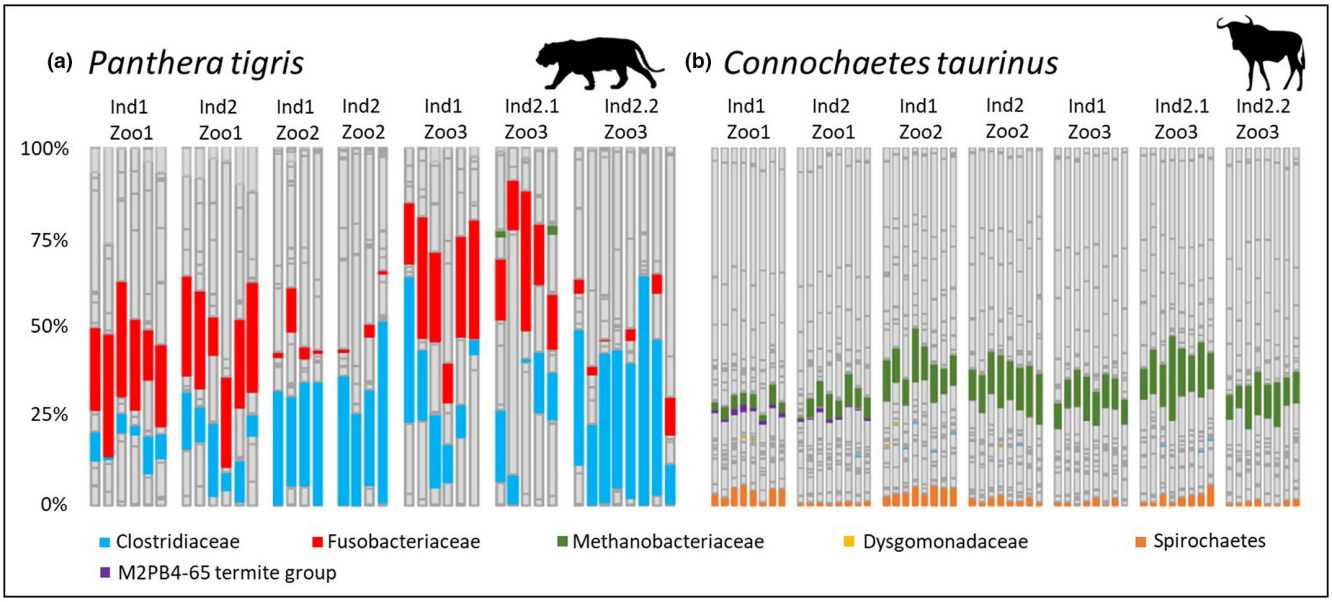

FIGURE 3 Comparison of the species-specific bacterial families that were identified with the LASSO algorithm. (a) Tiger-specific bacterial families as proportion of the total microbiota. (b) Wildebeest-specific bacterial families as proportion of the total microbiota.

Bacteroidaceae, Clostridiaceae, and Enterobacteriaceae occurred daily in almost all individuals, individuals could be distinguished based on specific bacterial combinations and fluctuations. The microbiota composition of individuals from the same zoo is often very similar (Figure 4a). For example, Prevotellaceae only appeared in greater proportions as a core taxon in both individuals from Zoo1 (Ind1_Zoo1: 16.46 ± 9.25%, Ind2_Zoo1: 20.70 ± 15.49%) and otherwise only occurred in smaller proportions within individuals from Zoo2 (Ind1_Zoo2: 0.35 ± 0.48%, Ind2_Zoo1: 0.93 ± 1.45%). Additionally, the individuals from Zoo1 showed a low average proportion of Enterobacteriaceae being persistent members (Ind1_Zoo1: 1.75 ± 1.71%, Ind2_Zoo1: 0.52 ± 0.48%) while this family was a core member in all other individuals. Both individuals from Zoo2 also showed a lower proportion of Enterobacteriaceae (Ind1_Zoo2: 1.28 ± 1.82%, Ind2_Zoo2: 1.17 ± 1.95%), but can be distinguished from the former individuals because of a 10 times lower average amount of Bacteroidaceae. The greatest proportion of Enterobacteriaceae was found in Ind1_Zoo3 (8.28 ± 7.70%) and Ind2.1_Zoo3 (22.84 ± 22.58%) whereas this family is less abundant in Ind2.2_Zoo3 (3.50 ± 3.83%). Even if the examined tiger individuals from the same habitat seemed to be similar (Figure A2), the time series clustering via the community detection algorithm showed that individuals can be identified via individual abundance patterns in their microbiota composition.

For individual discrimination within the wildebeests, the LASSO algorithm identified 12 bacterial taxa: Bacteroidales UCG-001, Muribaculaceae, p-251-o5, p-2534–18S gut group, Acidaminococcaceae, Saccharimonadaceae, Burkholderiaceae, Moraxellaceae, Spirochaetaceae, COB P4-1 termite group, Dysgomonadaceae and Tannerellaceae (Figure 4b). Even if most of these were core bacteria in many individuals, their proportion on the total microbiota is much lower compared to the tiger-specific bacteria. For example, Acidaminococcaceae on average only occurred between 1.46 ± 0.40% in Ind2.1_Zoo2 and 2.60 ± 0.49% in Ind3.2_Zoo2. Two individuals could not be discriminated from each other and were classified in the same cluster (Ind1_Zoo1, Ind1_Zoo2). These two individuals were the only ones in which p-2534–18S was found in the fecal microbiota. Because of the wildebeest microbial composition being very similar across all animals (Figure A2), individual clustering was based on low-frequent but steadily occurring bacterial families. Nevertheless, it is possible to identify individual microbiota within the wildebeests using a dynamic time-warping approach followed by a community detection clustering algorithm.

4 | DISCUSSION

4.1 | Evaluation of clustering algorithms

The aim of this study was to characterize abundance patterns in species-specific as well as in individual-specific fecal microbiota in an herbivore and a carnivore species over eight consecutive sampling days. Therefore, we chose two hierarchical agglomerative clustering algorithms that are able to perform reliably on smaller and biological datasets (Girvan & Newman, 2002; Terada, 2013; Yang et al., 2016). The ability of Ward's linkage to correctly assign species was limited when applying the single clustering pipeline. Possible limitations of this cluster algorithm include unequal distribution of sample sizes, the occurrence of outlier samples, and an elliptical rather than circular distribution of samples (Everitt et al., 2010). The former does not apply to the dataset used, but the latter two points could be a reason for the inadequate results of this algorithm since outlier samples, in particular, are well possible due to daily fluctuations in the microbiota. The limitations described above may also result in the inability to unambiguously classify species using the community detection algorithm in single clustering, resulting in three cluster solutions instead of two.

In order to improve clustering results, we synchronized the time series samples per individual with dynamic time warping to balance uneven sampling points. This approach has recently become the focus of meaningful interpretation of longitudinal data sets. By aligning different time series, data sets become comparable and temporal effects on the microbiota between or within individuals can be evaluated (Armoni & Borenstein, 2022; Lugo-Martinez et al., 2019; Muinck & Trosvik, 2018). The alignment of the individual time series to each other also led to a significant improvement in the cluster solutions in our study. Regarding the time series clustering, Ward's linkage resulted in a correct species-specific clustering when comparing the labels to the true species labels. The community detection algorithm also resulted in a correct assignment of wildebeests but failed to correctly identify all tiger samples. Nevertheless, this result confirms our assumption and previous results due to the greater variability in the microbiota within this species (Figure A2) (Karmacharya et al., 2019; Ning et al., 2020; Zoelzer et al., 2021).

4.2 | Individual-specific clustering based on longitudinal data

Both species differed significantly in alpha diversity as well as in microbial composition, which was to be expected and has already been shown in several studies (Nishida & Ochman, 2018; Zhu et al., 2018). To avoid biasing the cluster algorithms with these fundamental differences, we split the dataset into two species-specific datasets to test for individual variation in the microbiota. Similar to species-specific clustering, the high variability in the microbiota of tigers was probably responsible for the fact that no individuals can be assigned via single clustering. Both algorithms formed non-specific clusters, represented through a very low NMI when each sample is considered individually. Nevertheless, we managed to considerably improve the results of individual clustering by applying dynamic time warping to the individual samples and thus synchronizing the individual time series. The combination of up to eight samples per time series provided the cluster algorithm with more information

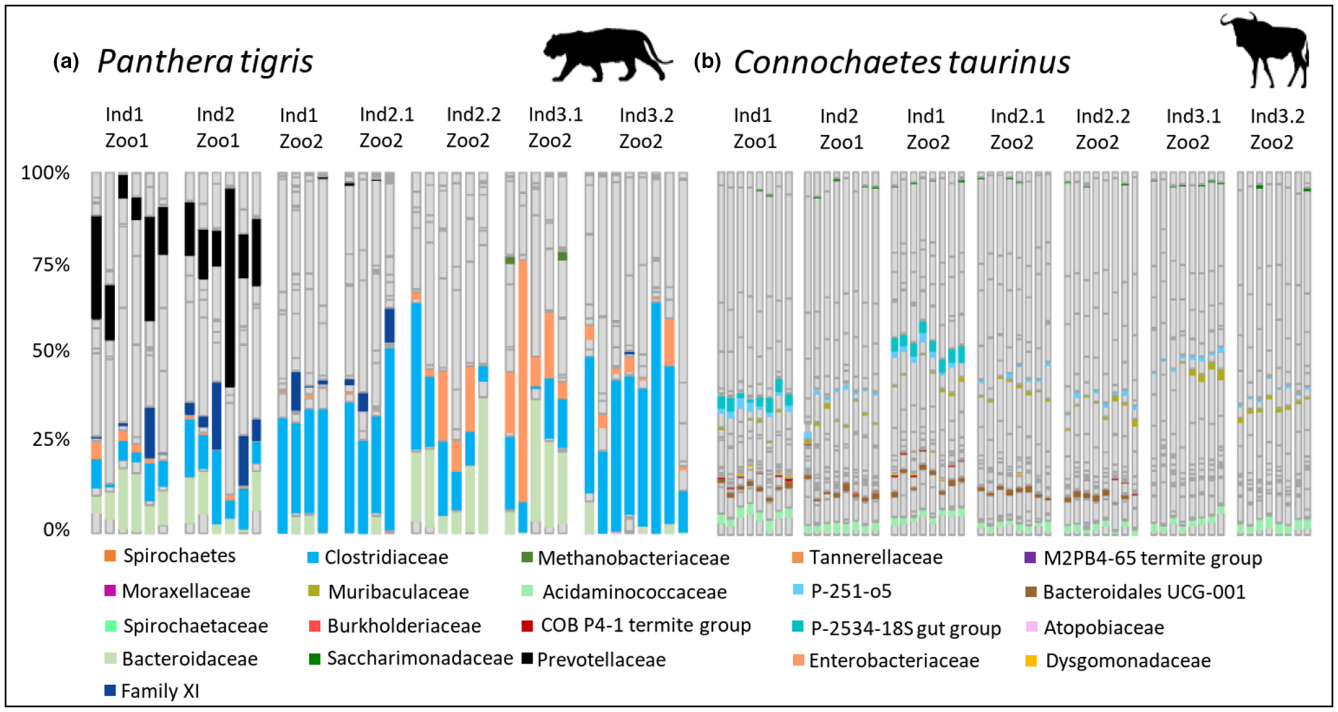

FIGURE 4 Comparison of the individual-specific bacterial families that were identified with the LASSO algorithm. (a) Individual-specific bacterial families identified within all tiger individuals as proportion of the total microbiota. (b) Individual-specific bacterial families identified within all wildebeest individuals as proportion of the total microbiota.

to process and thus led to more reliable clustering. In addition, outlier samples which negatively influence the algorithms (Everitt et al., 2010), could be correctly classified into the natural microbial oscillations of an individual, which was not possible when considering each sample individually.

Limitations to successful individual assignment arose from comparing different time periods of the same individual. While eight consecutive samples were sufficient to characterize the individual microbiota for exactly this period, more samples are needed to close the gap between longer sampling intervals. A reason for that might be the temporal dynamics of the fecal microbiota. Even if we found individual abundance patterns in the microbiota, the actual rhythm could be longer than 8 days and would not be fully captured in this study. Furthermore, other influencing factors that are known to shape the microbiota as diet, habitat, or seasonal shifts can lead to ongoing changes in the bacterial composition. In zoos, wildebeests are typically sustained on a diet of hay, alfalfa, or grass, which remains fairly consistent throughout the year. In contrast, the diet of tigers is more variable and can undergo daily changes in meat origin or preparation, such as whole-body or sheared meat. Studies have shown that felids' fecal microbiota is significantly affected by changes in diet and its composition, particularly alterations in the ratio of carbohydrates to protein (Bermingham et al., 2017; Butowski et al., 2019; Wernimont et al., 2020). To control for the habitat as an influencing factor, we also compared the clustering results with the true zoo-related labels. However, we did not obtain meaningful results and therefore excluded the habitat factor from further analysis. Another reason for the distinct time series clustering of the same individual might be seasonal adjustments of the microbiota. Seasonal shifts in microbiota composition have been reported previously in bisons (*Bison bison*) (Bergmann et al., 2015), musk deer (*Moschus* spp) (Jiang et al., 2021), primates (Baniel et al., 2021; Sawada et al., 2022), and also giant pandas (*Ailuropoda melanoleuca*) (Xue et al., 2015). Given all these variables, it is clear that even in a controlled habitat such as a zoo, changes or variations in feeding schedules or enclosure design must be well documented and taken into account when comparing longitudinal data sets.

4.3 | LASSO identified features for cluster

Once we assigned an individual microbiota to tigers and wildebeests, the next step was to identify the bacteria being responsible for this distinction. Species-specific clustering was mainly influenced by Fusobacteriaceae, Clostridiaceae, Methanobacteriaceae, and Spirochaetes with the first two being tiger-specific and the last mentioned being wildebeest-specific bacteria. Both Fusobacteriaceae and Clostridiaceae occurred within each individual and in each sample. This is in line with the results of our previous comparative study in which we were able to show that exactly these two families are the major families in many carnivore species (Zoelzer et al., 2021). Both taxa are involved in protein metabolism and the production of short-chain fatty acids (Basson et al., 2016; Bermingham et al., 2017; Vital et al., 2014) and are more abundant in species with a high-fat diet such as different predators (Bragg et al., 2020; Milani et al., 2020; Vital

et al., 2015). In contrast, Methanobacteriaceae and Spirochaetes were core bacteria in nearly all wildebeest individuals. Herbivores depend on microbial fermentation for carbohydrate digestion and, for example, Spirochaetes are capable of producing short-chain fatty acids from polysaccharide intake (Angelakis et al., 2019; de Filippo et al., 2010). Accordingly, this phylum has been found less in carnivores but more often in many herbivore species (Thingholm et al., 2021; Zoelzer et al., 2021). Methanobacteriaceae play an important role especially in ruminants as wildebeests, as they use the end products of microbial fermentation, CO_2 or H_2, as substrates to produce methane. This avoids an excessive increase in H_2 partial pressure in the rumen and the ambient factors for microbial digestive enzymes can be kept constant (Balch et al., 1979; Delzenne & Cani, 2011; Hook et al., 2010; Morgavi et al., 2010; Patra et al., 2017). All things considered, taxa that are responsible for species-specific clustering are mostly core bacteria that are involved in the specific herbivore or carnivore digestion process.

Additionally, we identified bacterial families that are necessary to characterize the individual microbiota of tigers and wildebeests. On the one hand, tiger individuals could be distinguished by core bacterial families such as Clostridiaceae, Enterobacteriaceae, and Bacteroidaceae. These are either involved in protein digestion or are members of the normal carnivore fecal microbiota (Kerr et al., 2013; Panasevich et al., 2015; Schwab & Gänzle, 2011; Xue et al., 2015). Even though these occurred daily, they showed host-specific fluctuations and can thus be used to determine individual differences. On the other hand, not only variations in the core bacteria determined the individual tiger microbiota but also individual-specific bacterial families. These accounted for either a large (e.g. Prevotellaceae) or small proportion (e.g., Atopobiaceae) of the total microbiota and are known to undergo fluctuations in the carnivore fecal microbiota (Guo et al., 2020; Ley et al., 2008).

In contrast to tigers, wildebeests showed a very uniform microbiota but nevertheless individual clustering and thus an assignment to an individual microbiota based on specific bacteria was possible. Accordingly, variation of high abundant bacterial families could not be responsible for individual differences as in tigers, but the distinctions were in low abundant bacterial families. Nevertheless, the pattern remained the same. General rumen-specific bacteria such as Acidaminococcaceae or the phylum Spirochaetes (Savin et al., 2022; Snelling et al., 2019) and bacterial families contributing to milk production in cattle such as Muribaculaceae and p-251-o5 (Boggio et al., 2021; Kodithuwakku et al., 2022) were considered to distinguish individual wildebeests. In contrast, bacterial families that only occurred in single individuals or that varied greatly among individuals were found by the LASSO algorithm to cluster individual microbiota (e.g. Burkholderiaceae, Moraxellaceae, p-2534–18B5 gut group). Overall, it can be concluded that there are individual abundance patterns of bacterial families, both in animal species with a highly variable as well as in species with a very constant microbiota. These are a combination of core bacteria of the respective species and other individual or zoo-specific families in varying abundance.

5 | CONCLUSION

To the best of our knowledge, this is the first study focusing on the characterization of individual oscillations in the microbiota of two species, applying different clustering algorithms to the sequencing data. With this research, we were able to show that two species with completely different diets exhibit both species-specific and individual abundance patterns in the fecal microbiota over the period of 1 week. Thus, we confirm our first hypothesis that a species-specific microbiota can be detected by the applied clustering pipeline. In addition, we showed that these results were considerably improved if time series data are considered and evaluated via dynamic time warping and community detection algorithm. Our second hypothesis that individual abundance patterns are more reliably detected in tigers than in wildebeest due to the more variable microbiota, cannot be clearly confirmed. Although the individual identification of the tigers resulted in a slightly higher NMI value when compared to the true labels, individual differences can also be detected within the very constant fecal microbiota of the wildebeests, with only marginally inferior cluster solutions compared to the tigers. Bacterial families that are responsible for individual clustering follow a similar pattern in both species. Individual abundance patterns are subject to a combination of species-specific core and individual-specific highly-fluctuating bacterial families.

From a methodological point of view, it can be implied for further studies that the use and interpretation of single individual samples or collective group samples is critical. Our results show that the microbiota of wildebeest and tiger are also subject to fluctuations that can only be captured through time series data. As one of the main tasks of zoos is to continuously improve animal welfare, they can benefit from an individualized microbial profile of some animal species that show e.g. special dietary requirements or increased stress susceptibility. The non-invasive sampling is easy to integrate into the daily routine and the evaluation is cost-effective. Accordingly, a fecal microbial profile is an easy-to-use method to continuously monitor individuals and, if necessary, to perform individual treatments or adjustments in the feeding schedule.

AUTHOR CONTRIBUTIONS

Franziska Zoelzer: Conceptualization (lead); formal analysis (lead); methodology (lead); writing – original draft (lead). **Sebastian Schneider:** Conceptualization (supporting); software (lead); writing – review and editing (equal). **Paul W. Dierkes:** Funding acquisition (lead); supervision (lead); writing – review and editing (equal).

ACKNOWLEDGMENTS

Open-access funding enabled and organized by Open-Access publication fond of the Goethe University. F. Zoelzer, S. Schneider, and P.W. Dierkes were supported by Opel-Zoo Foundation Professorship in Zoo Biology from the "von Opel Hessische Zoostiftung" and by the "Vereinigung von Freunden und Förderern der Goethe-Universität e.V.". We gratefully acknowledge support from the directors, curators, and especially keepers of Opel-Zoo Kronberg, Zoo Neuwied,

Zoo Landau in der Pfalz, Zoo Frankfurt, and Tierpark Berlin who enabled this project. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

DNA sequences: NCBI SRA BioProject PRJNA912880 (all samples) and PRJNA716130 (BioSamples: SAMN18396313, SAMN18396314, SAMN18396324–SAMN18396328, SAMN18396336, SAMN18396346, SAMN18396347, SAMN18396350, SAMN18396351, SAMN18396498, SAMN18396525, SAMN18396526, SAMN18396532, SAMN18396533, SAMN18396536–SAMN18396539, SAMN18396542, SAMN18396543, SAMN18396652–SAMN18396654, SAMN18396657–SAMN18396659, SAMN18396666–SAMN18396668, SAMN18396681–SAMN18396683, SAMN18396686–SAMN18396688, SAMN18396691–SAMN18396693, SAMN18396696–SAMN18396698, SAMN18396701–SAMN18396703).

ORCID

Franziska Zoelzer https://orcid.org/0000-0002-9659-6464

REFERENCES

Amin, N., & Seifert, J. (2021). Dynamic progression of the calf's microbiome and its influence on host health. *Computational and Structural Biotechnology Journal*, *19*, 989–1001. https://doi.org/10.1016/j.csbj.2021.01.035

Angelakis, E., Bachar, D., Yasir, M., Musso, D., Djossou, F., Gaborit, B., Brah, S., Diallo, A., Ndombe, G. M., Mediannikov, O., Robert, C., Azhar, E. I., Bibi, F., Nsana, N. S., Parra, H. J., Akiana, J., Sokhna, C., Davoust, B., Dutour, A., & Raoult, D. (2019). Treponema species enrich the gut microbiota of traditional rural populations but are absent from urban individuals. *New Microbes and New Infections*, *27*, 14–21. https://doi.org/10.1016/j.nmni.2018.10.009

Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129–137. https://doi.org/10.3354/ame01753

Armoni, R., & Borenstein, E. (2022). Temporal alignment of longitudinal microbiome data. *Frontiers in Microbiology*, *13*, 909313. https://doi.org/10.3389/fmicb.2022.909313

Ayoub, C., Arroyo, L. G., MacNicol, J. L., Renaud, D., Weese, J. S., & Gomez, D. E. (2022). Fecal microbiota of horses with colitis and its association with laminitis and survival during hospitalization. *Journal of Veterinary Internal Medicine*, *36*, 2213–2223. https://doi.org/10.1111/jvim.16562

Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R., & Wolfe, R. S. (1979). Methanogens: Reevaluation of a unique biological group. *Microbiological Reviews*, *43*(2), 260–296. https://doi.org/10.1128/mr.43.2.260-296.1979

Baniel, A., Amato, K. R., Beehner, J. C., Bergman, T. J., Mercer, A., Perlman, R. F., Petrullo, L., Reitsema, L., Sams, S., Lu, A., & Snyder-Mackler, N. (2021). Seasonal shifts in the gut microbiome indicate plastic responses to diet in wild geladas. *Microbiome*, *9*(1), 26. https://doi.org/10.1186/s40168-020-00977-9

Basson, A., Trotter, A., Rodriguez-Palacios, A., & Cominelli, F. (2016). Mucosal interactions between genetics, diet, and microbiome in inflammatory bowel disease. *Frontiers in Immunology*, *7*, 290. https://doi.org/10.3389/fimmu.2016.00290

Bergmann, G. T., Craine, J. M., Robeson, M. S., & Fierer, N. (2015). Seasonal shifts in diet and gut microbiota of the American bison (Bison bison). *PLoS One*, *10*(11), e0142409. https://doi.org/10.1371/journal.pone.0142409

Bermingham, E. N., Maclean, P., Thomas, D. G., Cave, N. J., & Young, W. (2017). Key bacterial families (Clostridiaceae, Erysipelotrichaceae and Bacteroidaceae) are related to the digestion of protein and energy in dogs. *PeerJ*, *5*, e3019. https://doi.org/10.7717/peerj.3019

Bhar, A., Gierse, L. C., Meene, A., Wang, H., Karte, C., Schwaiger, T., Schröder, C., Mettenleiter, T. C., Urich, T., Riedel, K., & Kaderali, L. (2022). Application of a maximal-clique based community detection algorithm to gut microbiome data reveals driver microbes during influenza a virus infection. *Frontiers in Microbiology*, *13*, 979320. https://doi.org/10.3389/fmicb.2022.979320

Björk, J. R., Dasari, M., Grieneisen, L., & Archie, E. A. (2019). Primate microbiomes over time: Longitudinal answers to standing questions in microbiome research. *American Journal of Primatology*, *81*(10–11), e22970. https://doi.org/10.1002/ajp.22970

Björk, J. R., Dasari, M. R., Roche, K., Grieneisen, L., Gould, T. J., Grenier, J.-C., Yotova, V., Gottel, N., Jansen, D., Gesquiere, L. R., Gordon, J. B., Learn, N. H., Wango, T. L., Mututua, R. S., Kinyua Warutere, J., Siodi, L.'i., Mukherjee, S., Barreiro, L. B., Alberts, S. C., … Archie, E. A. (2022). Synchrony and idiosyncrasy in the gut microbiome of wild baboons. *Nature Ecology & Evolution*, *6*(7), 955–964. https://doi.org/10.1038/s41559-022-01773-4

Boggio, M., Meynadier, A., Daunis-I-Estadella, P., & Marie-Etancelin, C. (2021). Compositional analysis of ruminal bacteria from ewes selected for somatic cell score and milk persistency. *PLoS One*, *16*(7), e0254874. https://doi.org/10.1371/journal.pone.0254874

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., & Knight, R. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*, *6*(1), 90. https://doi.org/10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bragg, M., Freeman, E. W., Lim, H. C., Songsasen, N., & Muletz-Wolz, C. R. (2020). Gut microbiomes differ among dietary types and stool consistency in the captive red wolf (*Canis rufus*). *Frontiers in Microbiology*, *11*, 590212. https://doi.org/10.3389/fmicb.2020.590212

Butowski, C. F., Thomas, D. G., Young, W., Cave, N. J., McKenzie, C. M., Rosendale, D. I., & Bermingham, E. N. (2019). Addition of plant dietary fibre to a raw red meat high protein, high fat diet, alters the faecal bacteriome and organic acid profiles of the domestic cat (*Felis catus*). *PloS One*, *14*(5), e0216072. https://doi.org/10.1371/journal.pone.0216072

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Caporaso, J. G., Lauber, C. L., Costello, E. K., Berg-Lyons, D., Gonzalez, A., Stombaugh, J., Knights, D., Gajer, P., Ravel, J., Fierer, N., Gordon, J. I., & Knight, R. (2011). Moving pictures of the human microbiome. *Genome Biology*, *12*(5), R50. https://doi.org/10.1186/gb-2011-12-5-r50

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., & Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, *6*(8), 1621–1624. https://doi.org/10.1038/ismej.2012.8

de Filippo, C., Cavalieri, D., di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., Collini, S., Pieraccini, G., & Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(33), 14691–14696. https://doi.org/10.1073/pnas.1005963107

Delzenne, N. M., & Cani, P. D. (2011). Interaction between obesity and the gut microbiota: Relevance in nutrition. *Annual Review of Nutrition*, *31*, 15–31. https://doi.org/10.1146/annurev-nutr-072610-145146

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2010). *Cluster analysis* (5th ed.). Wiley (Wiley Series in Probability and Statistics, v. 886). https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=4029211

Faith, J. J., Guruge, J. L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A. L., Clemente, J. C., Knight, R., Heath, A. C., Leibel, R. L., Rosenbaum, M., & Gordon, J. I. (2013). The long-term stability of the human gut microbiota. *Science*, *341*(6141), 1237439. https://doi.org/10.1126/science.1237439

Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On power-law relationships of the internet topology. *SIGCOMM Computer Communication Review*, *29*(4), 251–262. https://doi.org/10.1145/316194.316229

Fell, D. A., & Wagner, A. (2000). The small world of metabolism. *Nature Biotechnology*, *18*(11), 1121–1122. https://doi.org/10.1038/81025

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3–5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Friedman, J. H., Bentley, J. L., & Finkel, R. A. (1977). An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, *3*(3), 209–226. https://doi.org/10.1145/355744.355745

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 7821–7826. https://doi.org/10.1073/pnas.122653799

Guevarra, R. B., Lee, J. H., Lee, S. H., Seok, M.-J., Kim, D. W., Kang, B. N., Johnson, T. J., Isaacson, R. E., & Kim, H. B. (2019). Piglet gut microbial shifts early in life: Causes and effects. *Journal of Animal Science and Biotechnology*, *10*, 1. https://doi.org/10.1186/s40104-018-0308-3

Guo, X., Lei, H., Zhang, K., Ke, F., & Song, C. (2020). Diversification of animal gut microbes and NRPS gene clusters in some carnivores, herbivores and omnivores. *Biotechnology & Biotechnological Equipment*, *34*(1), 1280–1287. https://doi.org/10.1080/13102818.2020.1835536

Hildebrand, F., Gossmann, T. I., Frioux, C., Özkurt, E., Myers, P. N., Ferretti, P., Kuhn, M., Bahram, M., Nielsen, H. B., & Bork, P. (2021). Dispersal strategies shape persistence and evolution of human gut bacteria. *Cell Host & Microbe*, *29*(7), 1167–1176. https://doi.org/10.1016/j.chom.2021.05.008

Hoffman, M., Steinley, D., Gates, K. M., Prinstein, M. J., & Brusco, M. J. (2018). Detecting clusters/communities in social networks. *Multivariate Behavioral Research*, *53*(1), 57–73. https://doi.org/10.1080/00273171.2017.1391682

Hook, S. E., Wright, A.-D. G., & McBride, B. W. (2010). Methanogens: Methane producers of the rumen and mitigation strategies. *Archaea*, *2010*, 945785. https://doi.org/10.1155/2010/945785

Jiang, F., Gao, H., Qin, W., Song, P., Wang, H., Zhang, J., Liu, D., Wang, D., & Zhang, T. (2021). Marked seasonal variation in structure and function of gut microbiota in Forest and alpine musk deer. *Frontiers in Microbiology*, *12*, 699797. https://doi.org/10.3389/fmicb.2021.699797

Karmacharya, D., Manandhar, P., Manandhar, S., Sherchan, A. M., Sharma, A. N., Joshi, J., Bista, M., Bajracharya, S., Awasthi, N. P., Sharma, N., Llewellyn, B., Waits, L. P., Thapa, K., Kelly, M. J., Vuyisich, M., Starkenburg, S. R., Hero, J.-M., Hughes, J., Wultsch, C., … Sinha, A. K. (2019). Gut microbiota and their putative metabolic functions in fragmented Bengal tiger population of Nepal. *PLoS One*, *14*(8), e0221868. https://doi.org/10.1371/journal.pone.0221868

Katoh, K., Misawa, K., Kuma, K.-i., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. https://doi.org/10.1093/nar/gkf436

Kerr, K. R., Forster, G., Dowd, S. E., Ryan, E. P., & Swanson, K. S. (2013). Effects of dietary cooked navy bean on the fecal microbiome of healthy companion dogs. *PloS One*, *8*(9), e74998. https://doi.org/10.1371/journal.pone.0074998

Kodithuwakku, H., Maruyama, D., Owada, H., Watabe, Y., Miura, H., Suzuki, Y., Hirano, K., Kobayashi, Y., & Koike, S. (2022). Alterations in rumen microbiota via oral fiber administration during early life in dairy cows. *Scientific Reports*, *12*(1), 10798. https://doi.org/10.1038/s41598-022-15155-0

Koh, A., Vadder, F., Kovatcheva-Datchary, P., & Bäckhed, F. (2016). From dietary fiber to host physiology: Short-chain fatty acids as key bacterial metabolites. *Cell*, *165*(6), 1332–1345. https://doi.org/10.1016/j.cell.2016.05.041

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., Bircher, J. S., Schlegel, M. L., Tucker, T. A., Schrenzel, M. D., Knight, R., & Gordon, J. I. (2008). Evolution of mammals and their gut microbes. *Science*, *320*(5883), 1647–1651. https://doi.org/10.1126/science.1155725

Lugo-Martinez, J., Ruiz-Perez, D., Narasimhan, G., & Bar-Joseph, Z. (2019). Dynamic interaction network inference from longitudinal microbiome data. *Microbiome*, *7*(1), 54. https://doi.org/10.1186/s40168-019-0660-3

Lyu, T., Liu, G., Zhang, H., Wang, L., Zhou, S., Dou, H., Pang, B., Sha, W., & Zhang, H. (2018). Changes in feeding habits promoted the differentiation of the composition and function of gut microbiotas between domestic dogs (*Canis lupus* familiaris) and gray wolves (*Canis lupus*). *AMB Express*, *8*(1), 123. https://doi.org/10.1186/s13568-018-0652-x

Mamun, M. A. A., Sandeman, M., Rayment, P., Brook-Carter, P., Scholes, E., Kasinadhuni, N., Piedrafita, D., & Greenhill, A. R. (2020). Variation in gut bacterial composition is associated with Haemonchus contortus parasite infection of sheep. *Animal Microbiome*, *2*(1), 3. https://doi.org/10.1186/s42523-020-0021-3

Martínez, I., Muller, C. E., & Walter, J. (2013). Long-term temporal analysis of the human fecal microbiota revealed a stable core of dominant bacterial species. *PLoS One*, *8*(7), e69621. https://doi.org/10.1371/journal.pone.0069621

Milani, C., Alessandri, G., Mancabelli, L., Mangifesta, M., Lugli, G. A., Viappiani, A., Longhi, G., Anzalone, R., Duranti, S., Turroni, F., Ossiprandi, M. C., van Sinderen, D., & Ventura, M. (2020). Multi-omics approaches to decipher the impact of diet and host physiology on the mammalian gut microbiome. *Applied and Environmental Microbiology*, *86*(23), e01864. https://doi.org/10.1128/AEM.01864-20

Morgavi, D. P., Forano, E., Martin, C., & Newbold, C. J. (2010). Microbial ecosystem and methanogenesis in ruminants. *Animal: An International Journal of Animal Bioscience*, *4*(7), 1024–1036. https://doi.org/10.1017/S1751731110000546

Muinck, E. J., & Trosvik, P. (2018). Individuality and convergence of the infant gut microbiota during the first year of life. *Nature Communications*, *9*(1), 2233. https://doi.org/10.1038/s41467-018-04641-7

Nelson, T. M., Apprill, A., Mann, J., Rogers, T. L., & Brown, M. V. (2015). The marine mammal microbiome: Current knowledge and future directions. *Microbiology Australia*, *36*(1), 8. https://doi.org/10.1071/MA15004

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, *69*, 66133. https://doi.org/10.1103/PhysRevE.69.066133

Ning, Y., Qi, J., Dobbins, M. T., Liang, X., Wang, J., Chen, S., Ma, J., & Jiang, G. (2020). Comparative analysis of microbial community structure and function in the gut of wild and captive Amur Tiger. *Frontiers in Microbiology*, *11*, 1665. https://doi.org/10.3389/fmicb.2020.01665

Nishida, A. H., & Ochman, H. (2018). Rates of gut microbiome divergence in mammals. *Molecular Ecology*, *27*(8), 1884–1897. https://doi.org/10.1111/mec.14473

Ogle, D. H., Doll, J. C., Wheeler, P., & Dinno, A. (2022). FSA: Fisheries stock analysis. Version 0.9.3. https://github.com/fishR-Core-Team/FSA

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G., Solymos, P., Stevens, M., & Wagner, H. (2019). vegan: Community ecology package. https://CRAN.R-project.org/package=vegan

Olsson, L. M., Boulund, F., Nilsson, S., Khan, M. T., Gummesson, A., Fagerberg, L., Engstrand, L., Perkins, R., Uhlén, M., Bergström, G., Tremaroli, V., & Bäckhed, F. (2022). Dynamics of the normal gut microbiota: A longitudinal one-year population study in Sweden. *Cell Host & Microbe*, *30*(5), 726–739.e3. https://doi.org/10.1016/j.chom.2022.03.002

Paliwal, K. K., Agarwal, A., & Sinha, S. S. (1982). A modification over Sakoe and Chiba's dynamic time warping algorithm for isolated word recognition. *Signal Processing*, *4*(4), 329–333. https://doi.org/10.1016/0165-1684(82)90009-3

Panasevich, M. R., Kerr, K. R., Dilger, R. N., Fahey, G. C., Guérin-Deremaux, L., Lynch, G. L., Wils, D., Suchodolski, J. S., Steer, J. M., Dowd, S. E., & Swanson, K. S. (2015). Modulation of the faecal microbiome of healthy adult dogs by inclusion of potato fibre in the diet. *The British Journal of Nutrition*, *113*(1), 125–133. https://doi.org/10.1017/S0007114514003274

Patra, A., Park, T., Kim, M., & Yu, Z. (2017). Rumen methanogens and mitigation of methane emission by anti-methanogenic compounds and substances. *Journal of Animal Science and Biotechnology*, *8*, 13. https://doi.org/10.1186/s40104-017-0145-9

Petri, R. M., Schwaiger, T., Penner, G. B., Beauchemin, K. A., Forster, R. J., McKinnon, J. J., & McAllister, T. A. (2013). Characterization of the core rumen microbiome in cattle during transition from forage to concentrate as well as during and after an acidotic challenge. *PLoS One*, *8*(12), e83424. https://doi.org/10.1371/journal.pone.0083424

Ponziani, F. R., de Luca, A., Picca, A., Marzetti, E., Petito, V., del Chierico, F., Reddel, S., Paroni Sterbini, F., Sanguinetti, M., Putignani, L., Gasbarrini, A., & Pompili, M. (2022). Gut Dysbiosis and fecal calprotectin predict response to immune checkpoint inhibitors in patients with hepatocellular carcinoma. *Hepatology Communications*, *6*(6), 1492–1501. https://doi.org/10.1002/hep4.1905

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*, D590–D596. https://doi.org/10.1093/nar/gks1219

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Risely, A., Wilhelm, K., Clutton-Brock, T., Manser, M. B., & Sommer, S. (2021). Diurnal oscillations in gut bacterial load and composition eclipse seasonal and lifetime dynamics in wild meerkats. *Nature Communications*, *12*(1), 6017. https://doi.org/10.1038/s41467-021-26298-5

Rojas, C. A., Holekamp, K. E., Jasso M, V., Souza, V., Eisen, J. A., & Theis, K. R. (2023). Taxonomic, genomic, and functional variation in the gut microbiomes of wild spotted hyenas across 2 decades of study. *mSystems*, *8*(1), e0096522. https://doi.org/10.1128/msystems.00965-22

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(1), 43–49. https://doi.org/10.1109/TASSP.1978.1163055

Sanna, S., van Zuydam, N. R., Mahajan, A., Kurilshikov, A., Vich Vila, A., Võsa, U., Mujagic, Z., Masclee, A. A. M., Jonkers, D. M. A. E., Oosting, M., Joosten, L. A. B., Netea, M. G., Franke, L., Zhernakova, A., Fu, J., Wijmenga, C., & McCarthy, M. I. (2019). Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nature Genetics*, *51*(4), 600–605. https://doi.org/10.1038/s41588-019-0350-x

Savin, K. W., Moate, P. J., Williams, S. R. O., Bath, C., Hemsworth, J., Wang, J., Ram, D., Zawadzki, J., Rochfort, S., & Cocks, B. G. (2022). Dietary wheat and reduced methane yield are linked to rumen microbiome changes in dairy cows. *PLoS One*, *17*(5), e0268157. https://doi.org/10.1371/journal.pone.0268157

Sawada, A., Hayakawa, T., Kurihara, Y., Lee, W., & Hanya, G. (2022). Seasonal responses and host uniqueness of gut microbiome of Japanese macaques in lowland Yakushima. *Animal Microbiome*, *4*(1), 54. https://doi.org/10.1186/s42523-022-00205-9

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D. R., Kultima, J. R., Martin, J., Kota, K., Sunyaev, S. R., Weinstock, G. M., & Bork, P. (2013). Genomic variation landscape of the human gut microbiome. *Nature*, *493*(7430), 45–50. https://doi.org/10.1038/nature11711

Schneider, S., Hammerschmidt, K., & Dierkes, P. W. (2022). Introducing the software CASE (cluster and analyze sound events) by comparing different clustering methods and audio transformation techniques using animal vocalizations. *Animals: An Open Access Journal from MDPI*, *12*(16), 2020. https://doi.org/10.3390/ani12162020

Schwab, C., & Gänzle, M. (2011). Comparative analysis of fecal microbiota and intestinal microbial metabolic activity in captive polar bears. *Canadian Journal of Microbiology*, *57*(3), 177–185. https://doi.org/10.1139/W10-113

Shaani, Y., Zehavi, T., Eyal, S., Miron, J., & Mizrahi, I. (2018). Microbiome niche modification drives diurnal rumen community assembly, overpowering individual variability and diet effects. *The ISME Journal*, *12*(10), 2446–2457. https://doi.org/10.1038/s41396-018-0203-0

Snelling, T. J., Auffret, M. D., Duthie, C.-A., Stewart, R. D., Watson, M., Dewhurst, R. J., Roehe, R., & Walker, A. W. (2019). Temporal stability of the rumen microbiota in beef cattle, and response to diet and supplements. *Animal Microbiome*, *1*(1), 16. https://doi.org/10.1186/s42523-019-0018-y

Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., & Nishijima, M. (2014). Development of a prokaryotic universal primer for simultaneous analysis of bacteria and archaea using next-generation sequencing. *PloS One*, *9*(8), e105592. https://doi.org/10.1371/journal.pone.0105592

Terada, Y. (2013). Clustering for high-dimension, low-sample size data using distance vectors. https://arxiv.org/pdf/1312.3386

The MathWorks Inc. (2020). MathWorks MATLAB 2020a. Version 9.11.

Thingholm, L. B., Bang, C., Rühlemann, M. C., Starke, A., Sicks, F., Kaspari, V., Jandowsky, A., Frölich, K., Ismer, G., Bernhard, A., Bombis, C., Struve, B., Rausch, P., & Franke, A. (2021). Ecology impacts the decrease of Spirochaetes and Prevotella in the fecal gut microbiota of urban humans. *BMC Microbiology*, *21*(1), 276. https://doi.org/10.1186/s12866-021-02337-5

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Vandeputte, D., de Commer, L., Tito, R. Y., Kathagen, G., Sabino, J., Vermeire, S., Faust, K., & Raes, J. (2021). Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nature Communications*, *12*(1), 6740. https://doi.org/10.1038/s41467-021-27098-7

Vital, M., Gao, J., Rizzo, M., Harrison, T., & Tiedje, J. M. (2015). Diet is a major factor governing the fecal butyrate-producing community

structure across Mammalia, Aves and Reptilia. *The ISME Journal*, *9*(4), 832–843. https://doi.org/10.1038/ismej.2014.179

Vital, M., Howe, A. C., & Tiedje, J. M. (2014). Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *mBio*, *5*(2), e00889. https://doi.org/10.1128/mBio.00889-14

Wang, X., Tsai, T., Deng, F., Wei, X., Chai, J., Knapp, J., Apple, J., Maxwell, C. V., Lee, J. A., Li, Y., & Zhao, J. (2019). Longitudinal investigation of the swine gut microbiome from birth to market reveals stage and growth performance associated bacteria. *Microbiome*, *7*(1), 109. https://doi.org/10.1186/s40168-019-0721-7

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Wernimont, S. M., Radosevich, J., Jackson, M. I., Ephraim, E., Badri, D. V., MacLeay, J. M., Jewell, D. E., & Suchodolski, J. S. (2020). The effects of nutrition on the gastrointestinal microbiome of cats and dogs: Impact on health and disease. *Frontiers in Microbiology*, *11*, 1266. https://doi.org/10.3389/fmicb.2020.01266

Xue, Z., Zhang, W., Wang, L., Hou, R., Zhang, M., Fei, L., Zhang, X., Huang, H., Bridgewater, L. C., Jiang, Y., Jiang, C., Zhao, L., Pang, X., & Zhang, Z. (2015). The bamboo-eating giant panda harbors a carnivore-like gut microbiota, with excessive seasonal variations. *mBio*, *6*(3), e00022–e00015. https://doi.org/10.1128/mBio.00022-15

Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports*, *6*, 30750. https://doi.org/10.1038/srep30750

Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., Ley, R. E., & Farnleitner, A. H. (2019). Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nature Communications*, *10*(1), 2200. https://doi.org/10.1038/s41467-019-10191-3

Zhu, L., Wu, Q., Deng, C., Zhang, M., Zhang, C., Chen, H., Lu, G., & Wei, F. (2018). Adaptive evolution to a high purine and fat diet of carnivorans revealed by gut microbiomes and host genomes. *Environmental Microbiology*, *20*(5), 1711–1722. https://doi.org/10.1111/1462-2920.14096

Zoelzer, F., Burger, A. L., & Dierkes, P. W. (2021). Unraveling differences in fecal microbiota stability in mammals: From high variable carnivores and consistently stable herbivores. *Animal Microbiome*, *3*(1), 77. https://doi.org/10.1186/s42523-021-00141-0

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zoelzer, F., Schneider, S., & Dierkes, P. W. (2023). Time series cluster analysis reveals individual assignment of microbiota in captive tiger (*Panthera tigris*) and wildebeest (*Connochaetes taurinus*). *Ecology and Evolution*, *13*, e10066. https://doi.org/10.1002/ece3.10066

APPENDIX A

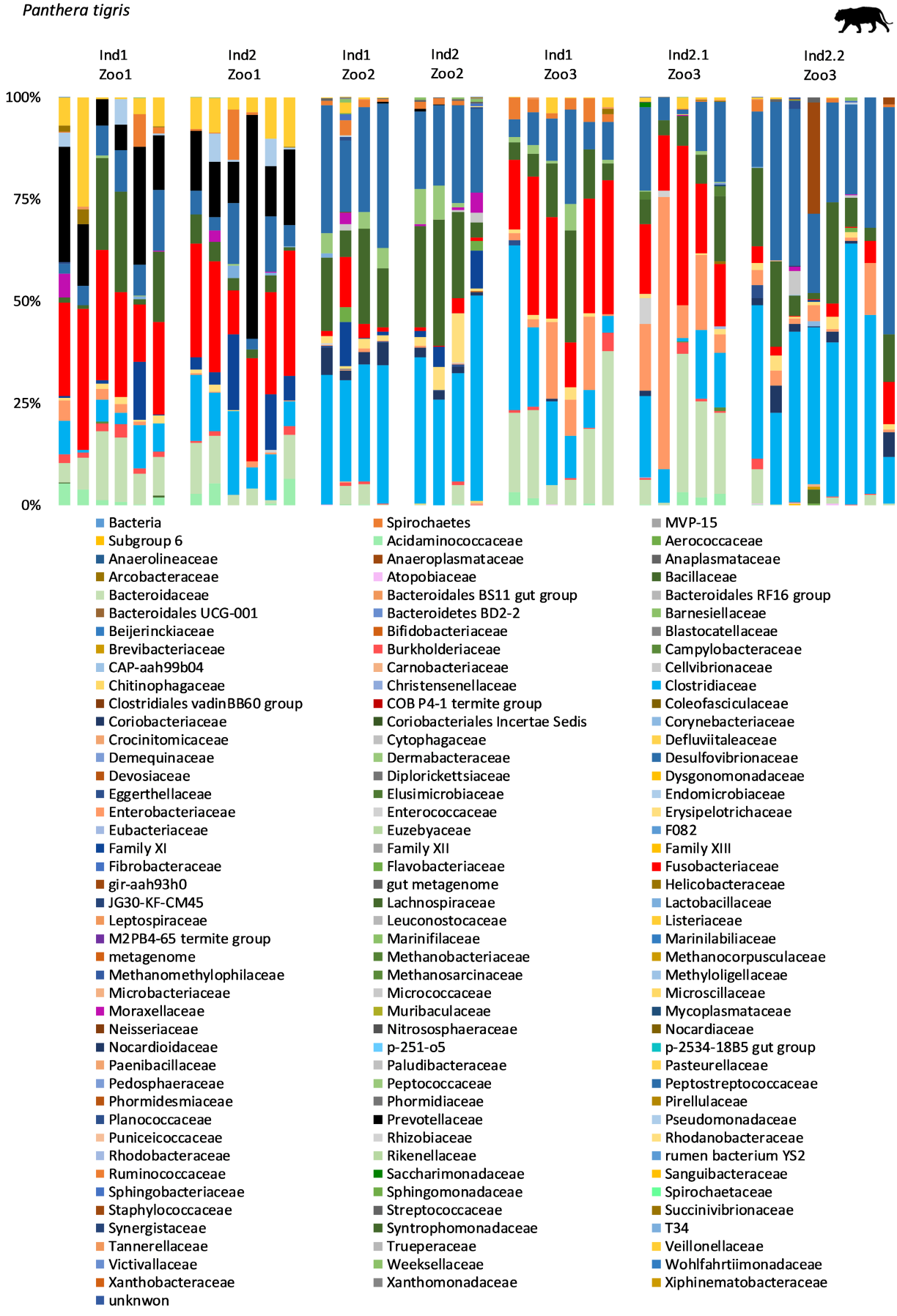

FIGURE A1 Visualization of the cluster solutions for both approaches via t-SNE in two dimensions.

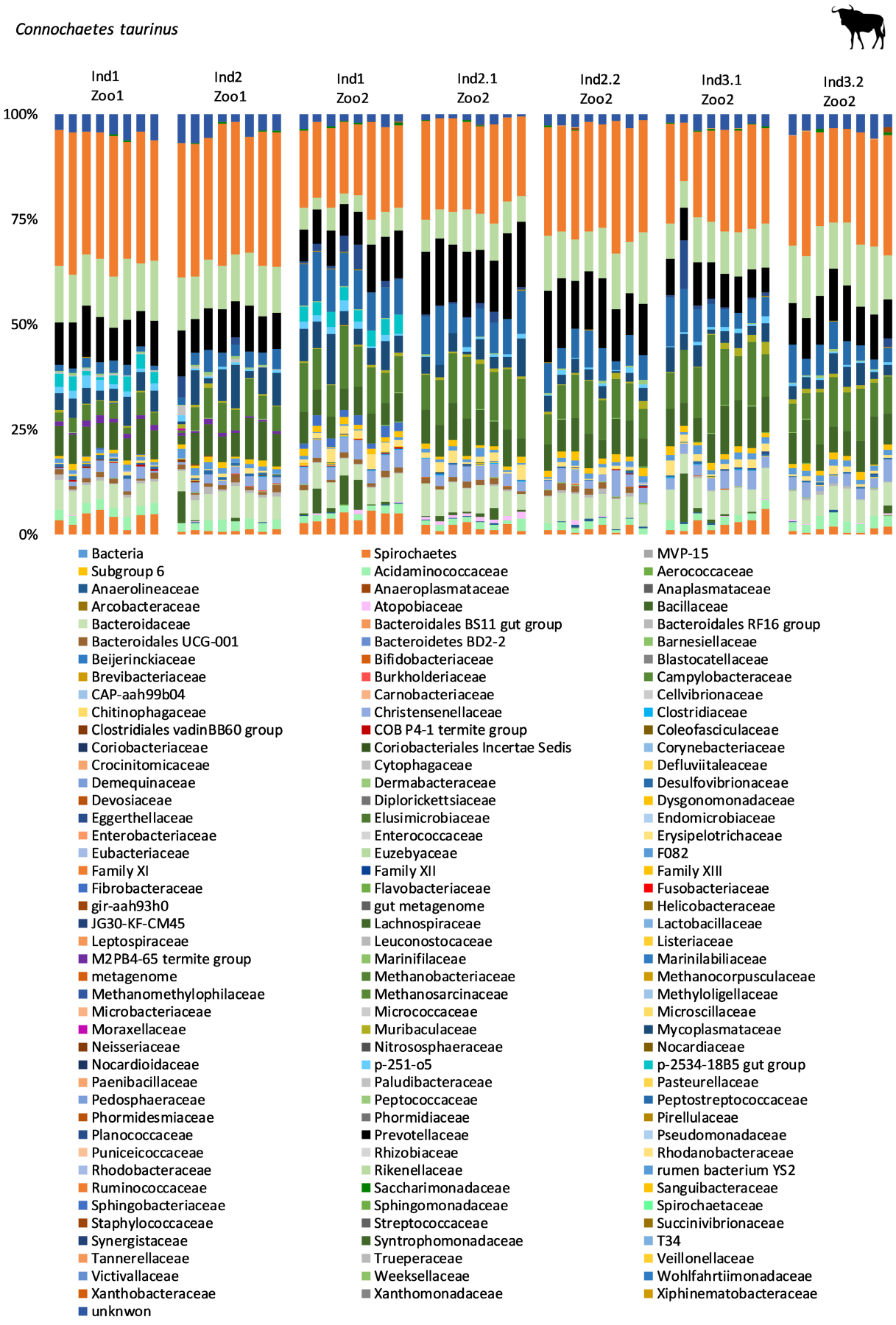

FIGURE A2 Fecal microbiota composition of all sampled tiger and wildebeest individuals.

TABLE A1 Metadata of all samples analyzed in this study including the species, sex, age and sampling date.

Individual_ID	Species	Sampling_Day	Sex	Year	Month	Age
Ind1_Zoo1	*Connochaetes_taurinus*	1	Female	2020	2	2
Ind1_Zoo1	*Connochaetes_taurinus*	2	Female	2020	2	2
Ind1_Zoo1	*Connochaetes_taurinus*	3	Female	2020	2	2
Ind1_Zoo1	*Connochaetes_taurinus*	4	Female	2020	2	2
Ind1_Zoo1	*Connochaetes_taurinus*	5	Female	2020	2	2
Ind1_Zoo1	*Connochaetes_taurinus*	6	Female	2020	2	2
Ind1_Zoo1	*Connochaetes_taurinus*	7	Female	2020	2	2
Ind1_Zoo1	*Connochaetes_taurinus*	8	Female	2020	2	2
Ind1_Zoo2	*Connochaetes_taurinus*	1	Male	2020	7	18
Ind1_Zoo2	*Connochaetes_taurinus*	2	Male	2020	7	18
Ind1_Zoo2	*Connochaetes_taurinus*	3	Male	2020	7	18
Ind1_Zoo2	*Connochaetes_taurinus*	4	Male	2020	7	18
Ind1_Zoo2	*Connochaetes_taurinus*	5	Male	2020	7	18
Ind1_Zoo2	*Connochaetes_taurinus*	6	Male	2020	8	18
Ind1_Zoo2	*Connochaetes_taurinus*	7	Male	2020	8	18
Ind1_Zoo2	*Connochaetes_taurinus*	8	Male	2020	8	18
Ind2.1_Zoo2	*Connochaetes_taurinus*	1	Female	2020	7	17
Ind2.1_Zoo2	*Connochaetes_taurinus*	2	Female	2020	7	17
Ind2.1_Zoo2	*Connochaetes_taurinus*	3	Female	2020	7	17
Ind2.1_Zoo2	*Connochaetes_taurinus*	4	Female	2020	7	17
Ind2.1_Zoo2	*Connochaetes_taurinus*	5	Female	2020	7	17
Ind2.1_Zoo2	*Connochaetes_taurinus*	6	Female	2020	8	17
Ind2.1_Zoo2	*Connochaetes_taurinus*	7	Female	2020	8	17
Ind2.1_Zoo2	*Connochaetes_taurinus*	8	Female	2020	8	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	1	Female	2020	12	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	2	Female	2020	12	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	3	Female	2020	12	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	4	Female	2020	12	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	5	Female	2020	12	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	6	Female	2020	12	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	7	Female	2020	12	17
Ind2.2_Zoo2	*Connochaetes_taurinus*	8	Female	2020	12	17
Ind2_Zoo1	*Connochaetes_taurinus*	1	Male	2020	2	3
Ind2_Zoo1	*Connochaetes_taurinus*	2	Male	2020	2	3
Ind2_Zoo1	*Connochaetes_taurinus*	3	Male	2020	2	3
Ind2_Zoo1	*Connochaetes_taurinus*	4	Male	2020	2	3
Ind2_Zoo1	*Connochaetes_taurinus*	5	Male	2020	2	3
Ind2_Zoo1	*Connochaetes_taurinus*	6	Male	2020	2	3
Ind2_Zoo1	*Connochaetes_taurinus*	7	Male	2020	2	3
Ind2_Zoo1	*Connochaetes_taurinus*	8	Male	2020	2	3
Ind3.1_Zoo2	*Connochaetes_taurinus*	1	Female	2020	7	1
Ind3.1_Zoo2	*Connochaetes_taurinus*	2	Female	2020	7	1
Ind3.1_Zoo2	*Connochaetes_taurinus*	3	Female	2020	7	1
Ind3.1_Zoo2	*Connochaetes_taurinus*	4	Female	2020	7	1
Ind3.1_Zoo2	*Connochaetes_taurinus*	5	Female	2020	7	1

TABLE A1 (Continued)

Individual_ID	Species	Sampling_Day	Sex	Year	Month	Age
Ind3.1_Zoo2	*Connochaetes_taurinus*	6	Female	2020	8	1
Ind3.1_Zoo2	*Connochaetes_taurinus*	7	Female	2020	8	1
Ind3.1_Zoo2	*Connochaetes_taurinus*	8	Female	2020	8	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	1	Female	2020	12	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	2	Female	2020	12	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	3	Female	2020	12	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	4	Female	2020	12	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	5	Female	2020	12	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	6	Female	2020	12	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	7	Female	2020	12	1
Ind3.2_Zoo2	*Connochaetes_taurinus*	8	Female	2020	12	1
Ind1_Zoo1	*Panthera_tigris*	1	Female	2018	5	10
Ind1_Zoo1	*Panthera_tigris*	2	Female	2018	5	10
Ind1_Zoo1	*Panthera_tigris*	3	Female	2018	5	10
Ind1_Zoo1	*Panthera_tigris*	5	Female	2018	5	10
Ind1_Zoo1	*Panthera_tigris*	6	Female	2018	5	10
Ind1_Zoo1	*Panthera_tigris*	8	Female	2018	5	10
Ind1_Zoo2	*Panthera_tigris*	2	Male	2020	12	8
Ind1_Zoo2	*Panthera_tigris*	3	Male	2020	12	8
Ind1_Zoo2	*Panthera_tigris*	4	Male	2020	12	8
Ind1_Zoo2	*Panthera_tigris*	5	Male	2020	12	8
Ind1_Zoo3	*Panthera_tigris*	1	Female	2018	6	17
Ind1_Zoo3	*Panthera_tigris*	2	Female	2018	6	17
Ind1_Zoo3	*Panthera_tigris*	3	Female	2018	6	17
Ind1_Zoo3	*Panthera_tigris*	4	Female	2018	6	17
Ind1_Zoo3	*Panthera_tigris*	5	Female	2018	6	17
Ind1_Zoo3	*Panthera_tigris*	8	Female	2018	6	17
Ind2.1_Zoo3	*Panthera_tigris*	1	Male	2018	6	8
Ind2.1_Zoo3	*Panthera_tigris*	2	Male	2018	6	8
Ind2.1_Zoo3	*Panthera_tigris*	3	Male	2018	6	8
Ind2.1_Zoo3	*Panthera_tigris*	4	Male	2018	6	8
Ind2.1_Zoo3	*Panthera_tigris*	5	Male	2018	6	8
Ind2.2_Zoo3	*Panthera_tigris*	1	Male	2020	11	10
Ind2.2_Zoo3	*Panthera_tigris*	2	Male	2020	11	10
Ind2.2_Zoo3	*Panthera_tigris*	3	Male	2020	11	10
Ind2.2_Zoo3	*Panthera_tigris*	4	Male	2020	11	10
Ind2.2_Zoo3	*Panthera_tigris*	5	Male	2020	11	10
Ind2.2_Zoo3	*Panthera_tigris*	6	Male	2020	11	10
Ind2.2_Zoo3	*Panthera_tigris*	7	Male	2020	11	10
Ind2.2_Zoo3	*Panthera_tigris*	8	Male	2020	11	10
Ind2_Zoo1	*Panthera_tigris*	1	Male	2018	5	7
Ind2_Zoo1	*Panthera_tigris*	2	Male	2018	5	7
Ind2_Zoo1	*Panthera_tigris*	3	Male	2018	5	7
Ind2_Zoo1	*Panthera_tigris*	5	Male	2018	5	7
Ind2_Zoo1	*Panthera_tigris*	7	Male	2018	5	7

(Continues)

TABLE A1 (Continued)

Individual_ID	Species	Sampling_Day	Sex	Year	Month	Age
Ind2_Zoo1	*Panthera_tigris*	8	Male	2018	5	7
Ind2_Zoo2	*Panthera_tigris*	1	Female	2020	12	7
Ind2_Zoo2	*Panthera_tigris*	3	Female	2020	12	7
Ind2_Zoo2	*Panthera_tigris*	5	Female	2020	12	7
Ind2_Zoo2	*Panthera_tigris*	6	Female	2020	12	7

Manuscript C

Declaration of author contributions to the manuscript Development and evaluation of an ensemble model to identify host-related metadata from fecal microbiota of zoo-housed mammals:

Status: *submitted*

Contributing authors:

Franziska Zölzer (FZ, PhD candidate)
Daniel Monteiro (DM)
Paul Wilhelm Dierkes (PWD)

What are the contributions of the doctoral candidate and his co-authors?

(1) Concept and design
Doctoral candidate FZ: 90
PWD: 10

(2) Conducting tests and experiments
Doctoral candidate FZ: 100

(3) Compilation of data sets and figures
Doctoral candidate FZ: 100

(4) Analysis and interpretation of data
Doctoral candidate FZ: 70
DM: 25
PWD: 5

(5) Drafting of manuscript
Doctoral candidate FZ: 80
DM: 10
PWD: 10

I hereby certify that the information above is correct.

_____ _____

Date and place Signature doctoral candidate

_____ _____

Date and place Signature supervisor

Development and evaluation of an ensemble model to identify host-related metadata from fecal microbiota of zoo-housed mammals

Franziska Zoelzer[1*], Daniel Monteiro[1], Paul Wilhelm Dierkes[1]

[1]Bioscience Education and Zoo Biology, Goethe University, Frankfurt am Main, Germany

*** Correspondence:**
Franziska Zoelzer
zoelzer@bio.uni-frankfurt.de

Keywords: 16S rRNA gene, microbiota, logistic regression, phylosymbiosis

Abstract

Much research has been conducted to describe the factors that determine the fecal microbiome, with diet and host phylogeny as the main drivers. The influence of diet has been described at different levels. Firstly, there are major differences in the microbiomes of herbivorous and carnivorous species and secondly the morphology of the digestive system also determines the composition and diversity of the microbiota. In this study, we aim to describe the influence of the three factors – diet, digestive system and host - on the microbiota in order to develop a model that is able to characterize host-specific metadata from an unknown fecal sample.

We therefore analyzed the 16s rRNA from 525 fecal samples of 14 zoo-housed species belonging to different phylogenetic groups including herbivores, carnivores and omnivores. We found significant differences in the bacterial taxa correlated with these groups. While herbivores show positive correlations with a large number of bacterial taxa, we found fewer taxa correlating with carnivores or omnivores. We also detected considerable differences in the microbiota of the ruminant, hindgut fermenting and simple digestive system. Based on these results, we developed a logistic ensemble model, that predicts the diet and based on these findings either the herbivorous digestive system or the carnivorous host-family from a given fecal microbiota composition. This model is able to effectively discriminate herbivores, omnivores and carnivores. It also excels at predicting the herbivore-specific digestive system with 98% accuracy, further reinforcing the strong link between microbiota and digestive system morphology. Carnivorous host-family identification achieves an overall accuracy of 79%, although this performance varies between families.

We provide this trained model as a tool to enable users to generate host-specific information from their microbiome data. In future research, tools such as the one presented here could lead to a combined approach of microbiome and host-specific analyses which would be a great advantage in non-invasive wildlife monitoring.

1 Introduction

In recent years, both extensive research efforts in the field of microbiome science as well as advances in DNA sequencing technology highlighted the importance of gastrointestinal microorganisms. These microorganisms have been extensively characterized, with a particular focus on their ability to produce short-chain fatty acids (SCFAs) as an essential energy source for the host organism (Koh et al., 2016;

Ríos-Covián et al., 2016; Sanna et al., 2019). In addition, a large number of studies have been carried out investigating composition, diversity, and intricate relationships that govern the gastrointestinal microbiome in different animal species as well as within different phylogenetic and dietary groups (Ley et al., 2008a; Nelson et al., 2015; Youngblut et al., 2019; Milani et al., 2020). Contemporary research raised compelling questions beyond the mere description of species-specific microbiomes, especially on influencing factors. Diet and phylogeny are now widely recognized as the primary drivers that shape the composition of the fecal microbiota.

For dietary studies on animals, the morphology of the respective digestive system must be taken into account. Carnivorous species are characterized by a simple digestive system with a short intestine and colon as well as a small cecum. In general, carnivores show little adaptations to microbial fermentation, as they rely on an easily digestible protein-based diet and have lower glucose requirements (Stevens and Hume, 1995; Mackie, 2002). In contrast, herbivores depend on microbial fermentation to break down cellulose and hemicellulose. Within herbivores, two digestive systems are predominant. On the one hand, hindgut fermenters have an enlarged large intestine to increase food retention time, and an enlarged cecum, which serves as the primary site for microbial fermentation. Ruminants, on the other hand, have a segmented stomach consisting of the rumen, reticulum, omasum and abomasum. Unlike hindgut fermenters, ruminants are foregut fermenters, with the majority of microbial fermentation occurring in the rumen. Although the small and large intestines of both digestive systems are similar in size, ruminants have a reduced cecum (Mitchell, 1905; Douglas, 2018). Several studies have shown that each of these digestive systems hosts its own microbiota due to the different morphological adaptations and specializations (Ley et al., 2008a; Muegge et al., 2011; Nishida and Ochman, 2018; Zoelzer et al., 2021).

Host phylogeny is the second key factor that shapes the fecal microbiota composition. The interaction between host and microorganisms is defined as phylosymbiosis (Lim and Bordenstein, 2020). Phylosymbiosis is described on different scales. Firstly, regarding a great phylogenetic context where numerous animal orders are compared with each other, the microbiota similarity increases with an increasing degree of kinship among the host species (Kartzinel et al., 2019; Rojas et al., 2021; Wu et al., 2022). Secondly, on a smaller phylogenetic range, closely related groups of species (Li et al., 2018; Fu et al., 2021) or species within the same genera (Knowles et al., 2019) were assessed and their results show that even closely related species have a distinct microbiota.

Due to the strong influence of the host on the microbiota, we aim to reverse this approach by developing a computational model that predicts the host from a given fecal microbiota sample. In a first step, we test whether different bacterial taxa correlate with the host's characteristics: the host diet group (herbivore, carnivore, omnivore), the host digestive system (ruminant, hindgut fermenter, simple) and the host-family applying a correlational analysis. In a second step, we develop a model that is able to predict those metadata from a given microbiota composition. Various modelling approaches gained widespread recognition and are now being applied in the fields of ecology and evolution, as they provide a versatile approach to effectively deal with complex data structures (Bolker et al., 2009). In microbiome research, machine learning techniques are mainly used to predict disease susceptibility patterns in the human microbiome (Yazdani et al.; Korpela et al., 2014; Duvallet et al., 2017; Espinoza, 2018). Recently, some studies developed models to predict host-specific factors from the fecal microbiota such as age or sex (Pannoni et al., 2022; Sweeny et al., 2023). Based on these findings and on the strong influence of host phylogeny, we hypothesize that it is possible to develop a model that identifies different host-families from an undisclosed fecal sample. This approach would open up new possibilities for microbiome analysis in non-invasive wildlife monitoring, i.e. using a single fecal sample of undisclosed origin and only one DNA sequencing workflow for an analysis of both

microbiome and host-specific information. Compared to common microsatellite methods, this is more cost- and labor-efficient in sampling, sequencing and analysis efforts.

2 Materials and Methods

2.1 Sample collection

Between May 2018 and November 2020, a total of 525 fecal samples were collected from 14 species belonging to different dietary groups of 17 zoos across Germany (Table S1). To ensure adequate representation, a minimum of 20 samples were analyzed for each species. The collection method employed is non-invasive, with samples primarily obtained during the daily cleaning routines by the keepers. After collection, the fecal samples were promptly transferred to sterile cryotubes and stored in liquid nitrogen until further processing. The care and use of animals during the research adhered to the guidelines set by the European Association of Zoos and Aquaria (EAZA). StarSEQ GmbH, located in Mainz, Germany, conducted the subsequent sample preparation. Initially, the samples underwent homogenization using the Precellys® Evolution Homogenizer (Bertin Instruments, Rockville, USA). Next, DNA extraction was carried out using the QIAamp® PowerFecal DNA Kit (Qiagen, Hilden, Germany). To measure the DNA concentration in the extracted samples, a NanoDrop spectrophotometer (Thermofisher, Massachusetts, USA) was utilized.

2.2 16S rRNA gene sequencing and data processing

At StarSEQ GmbH, sequencing of the V3-V4 region of the 16S rRNA gene was performed using a dual-index strategy based on the protocol of Caporaso et al., 2012 with minor modifications. To generate amplicons, a single step PCR of 33 cycles was performed using the primer combination 341f and 806bR as described by (Apprill et al., 2015) and (Takahashi et al., 2014). The resulting library was then sequenced on the Illumina MiSeq platform in paired-end mode (300 nt), including a 25% PhiX control library.

Subsequent data analysis was carried out using the QIIME 2 platform (Bolyen et al., 2019), and amplicon sequence variants (ASVs) were determined using DADA2 (Callahan et al., 2016). A phylogenetic tree was constructed for all sequences using MAFFT sequence alignment (Katoh et al., 2002) and low abundant features that are covered by less than 10 sequences, chloroplast and mitochondrial sequences were removed from the dataset. For taxonomic assignment of ASVs, a pre-trained Naive Bayes classifier (Bokulich et al., 2018) based on the SILVA 138 full-length database (Quast et al., 2013) was employed. The following statistical analyses were performed in R version 4.3 (R Core Team, 2022) as well as in Python version 3.10 (van Rossum and Drake, 2009). To identify bacterial families that are linked to either dietary, morphologically similar digestive systems or phylogenetic groups, we calculated the Pearson correlation coefficient with a minimum of $R>0.3$. Taxa that showed correlation coefficients from 0.3 to 0.5 are referred to as moderately correlating, taxa above 0.5 as strongly correlating. Based on those results, we performed a general linear model (glm) to test whether the correlating taxa are statistically significant factors in explaining whether a given microbiome composition belongs to one of the given categories or not.

Next, we developed a model to predict the diet, digestive system and host-family of an undisclosed sample. The preprocessing steps included:

1. Ensuring no missing values were present, which attests to the thoroughness of our data collection process.

2. Normalizing the data by converting raw counts to relative abundances, thus facilitating comparison across samples.

3. Removing sequences that were low in abundance or non-bacterial, refining the dataset to emphasize bacterial profiles relevant to the animals' diets.

These steps, essential for preparing the data for machine learning, complement the methodical sample collection and sequencing methods described in our documentation. We acknowledge that due to the unstandardized collection procedure of the samples and the time component among other factors our dataset does not strongly support the assumption of IID. However, we have taken several steps to ensure that the data we used was thoroughly checked and prepared for further analysis and the use in machine learning models.

2.3 Model training and evaluation framework

The dataset was divided into training, development, and test sets with a ratio of 64%, 16%, and 20%, respectively. This split ensured enough data for training the models, allowed for hyperparameter optimization on the development set, and provided an unbiased evaluation on the test set. The training of the models was systematically conducted through grid search, meticulously iterating over a range of hyperparameter combinations to identify the optimal settings based on F1 scores. The F1 score was prioritized as our evaluation metric due to its balance between precision and recall, which is crucial when dealing with imbalanced classes.

We enhanced the predictive accuracy by employing an ensemble of logistic regression models, adhering to the hierarchical classification procedure of diet, herbivorous digestion types, and carnivorous host-family. Ensemble modelling combines the predictions from multiple models to increase the robustness of the results, leveraging the strength of each individual model without weakness of general model.

1. Hierarchical Model Training and Validation: We trained individual logistic regression models at each level of the hierarchy—first on diet, then on digestion types for herbivores, and on the host-family for carnivores and omnivores. Each model was validated using the development set, ensuring that we could fine-tune the hyperparameters effectively. This hierarchical approach allowed the ensemble to build upon the structure and dependencies inherent in the data, using the development set to guide the selection and combination of models without biasing the final evaluation.

2. Aggregation of Model Probabilities: Aggregation of Model Probabilities: The ensemble method was implemented using a sequential, multi-stage approach. In each stage, multiple models predict the label for the respective category, and the label with the highest calculated probability is selected. This approach ensures that the most precise prediction of each model is considered, with the final decision based on the most probable forecast. This method allows for a differentiated and context-dependent integration of predictions from various models, taking into account both the accuracy and the consensus of individual models at different classification stages

3. Performance Evaluation: The final evaluation of the ensemble model's performance was conducted on the test set. This step is crucial, as it provides an unbiased estimate of how well the ensemble model generalizes to new, unseen data. The metrics from this evaluation were used to assess the success of the modelling approach.

The emphasis throughout the modelling process was on the refinement of the ensemble based on development set results rather than test data, maintaining the integrity of the test set for a genuine evaluation of model generalization. The ensemble's performance was quantified using the F1 score, which harmonizes precision and recall, to ensure the model's efficacy across both prevalent and rare classes within our dataset.

3 Results

A total of 525 fecal samples comprising 14 species were sequenced on Illumina MiSeq platform in paired-end mode, targeting the V3-V4 region of the 16S rRNA gene. After preprocessing, the dataset contained 27,188,318 sequences, ranging from 5,836 to 303,032 sequences per sample, with an average of 51,299 sequences per sample.

3.1 Correlation of bacterial taxa with diet

Applying the Pearson's correlation to the data set with respect to diet groups, the strongest correlation values are found for herbivores. Within the $R>0.3$ limit, ten positively correlating bacterial families are identified (Figure 1A). Within these, Prevotellaceae ($R=0.45$), Methanocorpusculareae ($R=0.32$) and Fibrobacteraceae ($R=0.42$) show a moderately positive correlation, while Methanobacteriaceae ($R=0.52$), F082 ($R=0.56$), Christensenellaceae ($R=0.54$), Akkermansiaceae ($R=0.54$), Spirochaetaceae ($R=0.60$), Ruminococcaceae ($R=0.71$) and Rikenellaceae ($R=0.73$) show a strong positive correlation with herbivores. The following generalized linear model (glm) confirms these findings, indicating that the occurrence of Methanobacteriaceae ($t=9.84$, $p<0.01$), Spirochaetaceae ($t=5.55$, $p<0.01$), Akkermansiaceae ($t=3.89$, $p<0.01$), Ruminococcaceae ($t=3.55$, $p<0.01$) and Rikenellaceae ($t=3.47$, $p<0.01$), as well as the absence of Peptostreptococcaceae ($t=-10.26$, $p<0.01$) and Fusobacteriaceae ($t=-8.64$, $p<0.01$), are significant factors in distinguishing a herbivorous from carnivorous or omnivorous host. Nevertheless, these bacterial taxa contribute in different proportions to the herbivore microbiota. While Ruminococcaceae (Average ± standard deviation: 16.30%±10.20%), Spirochaetaceae (9.06%±8.74%) and Rikenellaceae (7.55%±4.78%) are on average more abundant, Akkermansiaceae (2.53%±2.85%) and Methanobacteriaceae (4.31%±5.37%) are less abundant (Figure 2). Carnivorous species show a negative correlation with most of the bacterial families that correlate with herbivores. Only three positively correlating bacterial taxa are identified: Fusobacteriaceae with a strong correlation ($R=0.55$, $t=9.03$, $p<0.001$) as well as Peptostreptococcaceae ($R=0.31$, $t=4.38$, $p<0.001$) and Burkholderiaceae ($R=0.31$, $t=0.99$, $p=0.32$) with a moderate correlation.
As shown in Figure 2, Fusobacteriaceae (19.93%±15.37%) and Peptostreptococcaceae (10.84%±9.54%) combined make up approximately one third of the carnivore microbiota. In addition, the glm identified the absence of Methanobacteriaceae ($R=-0.35$, $t=-2.62$, $p<0.001$) and Rikenellaceae ($R=-0.49$, $t=-1.98$, $p<0.05$) as a significant factor to discriminate a carnivore from a herbivore or omnivore. Contrary to herbivores and carnivores, the omnivorous hosts assessed in this study display no strong correlation to any microbial family. Only Enterococcaceae ($R=0.36$) and Clostridiaceae ($R=0.37$) correlate moderately with this diet group. Thus, Clostridiaceae form a major part of the omnivore microbiota (32.48%±15.21%) and show no positive correlation with herbivores or carnivores. The latter is also true for Enterococcaceae, but this family only contributes on average to 5.65%±8.51% to the microbiota composition of omnivores.

3.2 Correlation of bacterial taxa with digestive system

With regard to the digestive system, herbivores in this study are divided into foregut ruminants and hindgut colon fermenters. Ruminants show strong positive correlations with Ruminococcaceae ($R=0.77$, $t=12.31$, $p<0.001$), Methanobacteriaceae ($R=0.64$, $t=17.58$, $p<0.001$), Prevotellaceae

(R=0.51, t=4.69, p<0.001) and Akkermansiaceae (R=0.55, t=3.28, p<0.01). With the exception of Rikenellaceae (R=0.63, t=-1.38, p=0.19), which is shared with the hindgut fermenters, these bacterial taxa are exclusively positively correlated with ruminants. Ruminococcaceae are highly abundant in this dietary group with an average of 22.62%±8.58%, followed by Prevotellaceae (9.68%±4.51%), Rikenellaceae (9.24%±4.68%) and Methanobacteriaceae (6.75%±6.06%) (Figure 2). Other moderately positive correlations occur with Desulfovibrionaceae (R=0.34, t=5.06, p<0.001), Christensenellaceae (R=0.46, t=-4.46, p<0.001) and Anaerolineaceae (R=0.32, t=0.62, p=53). No strong negative correlations are observed, but the absence of Clostridiaceae (R=-0.36, t=3.46, p<0.001) is another indicator of a ruminant host species.

The correlation analysis reveals differences between ruminants and hindgut fermenters. While Ruminococcaceae are high-abundant in the microbiota of ruminants, this taxon does not correlate with hindgut fermenters (Figure 1). In contrast, Spirochaetaceae show the strongest positive correlation (R=0.86, t=15.52, p<0.001) and on average make up 18.81%±6.22% of the hindgut fermenters microbiota. Similar to ruminants, there is an archaeal taxon strongly positively correlating with hindgut fermenters, namely Methanocorpusculareeae (R=0.57, t=4.06, p<0.001). Other hindgut-specific bacterial families make up less than 5% on the average microbiota as shown in Figure 2. These include Fibrobacteraceae (R=0.64, t=6.61, p<0.01), Eubacteriaceae (R=0.64, t=7.70, p<0.001), T34 (R=0.39, t=5.18, p<0.001) and Synergistaceae (R=0.56, t= 4.30, p<0.001).

The simple carnivore digestive system is characterized by a different bacterial composition. In general, bacterial families are less positively correlated with this digestive system, which is reflected in the correlation values as well as in the lower t-values of the glm. Therefore, the five positively correlating bacterial families occur in higher proportions. Strong correlational values are found for Clostridiaceae (R=0.56, t=3.41, p<0.001) and Fusobacteriaceae (R=0.52, t=5.58, p<0.001), which contribute to 19.50%±17.52% and 15.71%±16.08% to the microbiota composition. Peptostreptococcaceae (R=0.48, t=3.41, p<0.001) and Enterobacteriaceae (R=0.37, t=3.28, p<0.001) correlate moderately with this digestive system, but also make up 11.14%±9.97% and 9.65%±15.15% of the average microbiota. The glm also identifies negatively correlating taxa as indicators to distinguish this digestive system from the others. These include Ruminococcaceae (R=-0.81, t=-10.43, p<0.001), Rikenellaceae (R=-0.83, t=-8.03, p<0.001), Spirochaetaceae (R=-0.68, t=-14.59, p<0.001) and Methanobacteriaceae (R=-0.58, t=-22.79, p<0.001), which were previously identified as herbivore-specific taxa.

3.3 Correlation of bacterial taxa with host-family

The calculation of correlations for bacterial taxa with host-families reveals only moderate correlations for all carnivores and omnivores but in strong correlations for herbivores. Within the carnivores, the glm for Ailuridae (AIC=-327.37) and Herpestidae (AIC=-292.04) is suitable to distinguish them from the other animals. Ailuridae specific bacterial taxa are Enterobacteriaceae (R=0.31, t=5.75, p<0.001), Erysipelotrichaceae (R=0.33, t=9.30, p<0.001), Planococcaceae (R=0.41, t=7.99, p<0.001), Sphingobacteriaceae (R=0.33, t=7.08, p<0.001) and Paenibacillaceae (R=0.36, t=3.68, p<0.001). Among them, Enterobacteriaceae have the largest average proportion on the microbiota (21.79%±19.37%), followed by Planococcaceae (12.45%±12.24%) and Erysipelotrichaceae (9.02%±10.68%). For Herpestidae, there is only one moderate but significant positive correlation with Listeriaceae (R=0.31, t=7.38, p<0.001). Canidae (AIC=185.02), Felidae (AIC=384.02) and Ursidae (AIC=3.42) do not discriminate well with the glm. Significant correlations occur with Fusobacteriaceae (R=0.45, t=7.62, p<0.001) for Canidae, Veillonellaceae (R=0.38, t=7.46, p<0.001) for Felidae and Clostridiaceae (R=0.45, t=5.91, p<0.001) for Ursidae (Figure 1).

For herbivores, the glm works best for Equidae (AIC=-495.45). Equidae display strong positive correlations with Spirochaetaceae (R=0.86, t=15.52, p<0.001), Fibrobacteraceae (R=0.64, t=6.61,

p<0.001), F082 (R=0.61, t=3.00, p<0.01) and Synergistaceae (R=0.56, t=4.30, p<0.001). In particular, the high average abundance of Spirochaetaceae (18.81%±6.22%) and the low abundance of Ruminococcaceae differentiate the Equidae from Giraffidae and Bovidae as is shown in Figure 2. In contrast, the other identified taxa represent less than 5% of the average Equidae microbiota. Giraffidae can also be identified by the glm (AIC=-376.33). Here, only two strongly positive correlating families, PeH15 (R=0.50, t=9.67, p<0.001) and Methanomethylophilaceae (R=0.50, t=8.64, p<0.001) can be used to differentiate this ruminant family. However, both families contribute on average less than 1% to the microbiota composition. In addition, Ruminococcaceae (R=0.37, t=4.30, p<0.001) is another important taxon for the discrimination and occurs on average in 23.34%±4.49% within the Giraffidae. The second ruminant family, Bovidae, shows the strongest positive correlation with Methanobacteriaceae (R=0.69, t=17.67, p<0.001), Ruminococcaceae (R=0.63, t=6.82, p<0.01) and Akkermansiaceae (R=0.54, t= 6.23, p<0.01). Compared to the Giraffidae, the Bovidae show a comparable average proportion of Ruminococcaceae (22.43%±9.39%), but a higher average proportion of Methanobacteriaceae (8.06%±6.13%). Nevertheless, the glm for Bovidae is not as good as for the other two herbivore species (AIC=-63.48).

3.4 Developing an ensemble model to identify host-specific information

As we expect herbivores, carnivores and omnivores to differ in their microbiota composition, we developed an ensemble model that can accurately identify the diet type (herbivore, carnivore, omnivore) based on the microbiota composition of a fecal sample. The model has an accuracy of 88%. The F1 score for identifying diet type varies between 0.73 for omnivores, 0.87 for carnivores, and 0.93 for herbivores (Table 1). In a second step, the model predicts the digestive system of herbivores to be simple, ruminant, or hindgut fermenter with an overall accuracy of 98%. The hindgut fermenters show the best results (F1=1.00), followed by the ruminants (F1=0.98) and the simple digestive system (F1=0.92). Due to the limited number of samples after this step and the resulting poor results of the model in predicting the host-family, we omitted this step for the herbivores. Nevertheless, since the Equidae is the only family present in the hindgut fermenters in this study, the F1 score applies to this family as well. Another distinguishable host-family due to its unique feature of a simple digestive system is the Ailuridae (F1=0.92). Since all carnivores have the same simple digestive system, a discrimination at this level is not necessary here. Instead, the model achieves an accuracy of 79% in distinguishing between host-families. Canidae are distinguished from the other carnivores with a F1 score of 0.93, while Felidae (F1=0.82) and Ursidae (F1=0.79) are detected with slightly less reliably. However, this step of the model fails to identify the Herpestidae because of an insufficient sample size (F1=0.00).

4 Discussion

4.1 The influence of diet, digestive system and host phylogeny on the microbiota

We found clear differences in bacterial families correlating with either herbivorous, carnivorous or omnivorous hosts. The greatest divergence between these groups is the number of either positively or negatively correlating bacterial taxa. While herbivores show positive correlations with many taxa, carnivores express the opposite pattern. One reason for this may be the overall higher microbial diversity in the feces of herbivores, which has been demonstrated repeatedly (Ley et al., 2008a; Vital et al., 2014; Youngblut et al., 2019; Guo et al., 2020; Zoelzer et al., 2021). The dependence of herbivores on fermentative bacteria is further supported by the fact that the most strongly correlating taxa are involved in fiber digestion. For example Rikenellaceae and Ruminococcaceae are known as herbivore-specific bacterial families (Milani et al., 2020; Zoelzer et al., 2021) and play an important

role in cellulose degradation and fiber digestion (Obregon-Tito et al., 2015; La Reau and Suen, 2018). Another important aspect of fiber digestion by fermentative bacteria is the production of short-chain fatty acids (SCFAs), which serve as an energy source for the host. Spirochaetaceae, which correlate strongly with herbivores, are able to produce the SCFAs butyrate and acetate from different polysaccharide intakes (Pascale et al., 2018; van den Abbeele et al., 2022). In contrast, less positively correlating taxa are found in carnivores. Only Fusobacteriaceae show strong correlation with carnivore hosts. This family is known to be dominant in meat-based diets and is involved in protein digestion (Vital et al., 2014; An et al., 2017; Badri et al., 2021; Martínez-López et al., 2021). A different pattern is observed within omnivores. Here only two bacterial families are found to correlate moderately, Enterococcaceae and Clostridiaceae. Nevertheless, both taxa belong to the phylum Firmicutes, which is known to be highly represented in omnivore species (Sommer et al., 2016; Trujillo et al., 2022a, 2022b). Overall, these results suggest that there are considerable differences in the microbiota, especially between herbivores and carnivores/omnivores.

Furthermore, we analyzed microbial taxa that correlate with the morphology of the digestive system. We found clear differences here as well (Figure 1). Both herbivore digestive systems rely on microbial degradation of cellulose and production of SCFA's, but the site of fermentation differs. As the name indicates, the fermentation in ruminants takes place in the highly comparted forestomach, prior to enzymatic digestion. We identified Ruminococcaceae, Prevotellaceae and Methanobacteriaceae as strongly correlating with this digestive system. The two ruminants assessed in this study, namely Giraffidae and Bovidae, can be distinguished in terms of the methanogenic archaea. While Bovidae correlate strongly with Methanobacteriaceae, Giraffidae show a strong correlation with Methanomethylophilaceae. The low proportion of Methanobacteriaceae in Giraffidae has also been demonstrated in several studies (Roggenbuck et al., 2014; Zoelzer et al., 2021). The differences between these two ruminants may be due to adaptation to either browsing (Giraffidae) or grazing (Bovidae). Compared to grazing ruminants, browsers have a smaller rumen and larger intestines to increase retention time in the digestive system (Giesecke and van Gylswyk, 1975; Hofmann, 1989; Woodall and Skinner, 1993; Ginnett and Demment, 1997; Clauss et al., 2003; Mitchell, 2021). As methanogenic archaea are responsible for balancing the pH-value in the rumen, as mentioned above, the different archaeal taxa may be adapted to different rumen sizes and compartments.

Hindgut fermenting species are adapted to a microbial fermentation that occurs after the enzymatic digestion. Although the site of fermentation is not the rumen but the large intestine, the fermentation process is quite similar. In this study, we identified Spirochaetaceae as significantly correlating with this digestive system and, as it is the only family assessed, with Equidae. This is consistent with previous findings that Spirochaetaceae is a predominant core member in the equid microbiota (Yatsunenko et al., 2012; Obregon-Tito et al., 2015; Edwards et al., 2020). Another important taxon in the Equidae is Fibrobacteraceae (Figure 1,2). This family consists of only one genus, *Fibrobacter*. Taxa within this genus are particularly known for their ability to degrade cellulose in low-oxygen environments, and therefore contain many genes encoding glycoside hydrolases (Montgomery et al., 1988; Abdul Rahman et al., 2015). In terms of archaeal taxa, the Equidae are distinct from the Giraffidae and Bovidae. Here, we identified Methanocorpusculareae as another class of methanogenic archaea within the fecal microbiota of herbivores. This taxon has also been identified in several Equidae species (Edwards et al., 2020) suggesting that the archaeal microbiota is also host- or at least digestion type specific.

Regarding the simple digestive system, strongly correlating taxa such as Fusobacteriaceae and Clostridiaceae, are mainly involved in protein degradation and digestion of high-fat diets, which again strengthens the influence of diet on the microbiota. We found no strong correlations with microbial

taxa for carnivore and omnivore host species, which may be a limitation for the model prediction. One reason for this may be that this digestive system does not show a specific adaptation to microbial fermentation, as is the case for herbivores. An exception within the simple digestive system is the Ailuridae, which belongs to the order of Carnivora but in its natural habitat the diet of this family consists mainly of *Arundinaria spp*. as well as seasonal fruits (Pradhan et al., 2001; Panthi et al., 2015). Across the assessed zoos, the diet of this family was very close to this. Bamboo was provided *ad libitum* and a daily variation of fresh fruit was added to the feeding protocol. This family shows a large proportion of Enterobacteriaceae. Enterobacteriaceae express many genes encoding thiosulfate sulfurtransferases that is able to break down cyanide compounds. These metabolites are found in bamboo species such as *Arundinaria spp*. and therefore this bacterial taxon may play a role in detoxification of the Ailuridae (McKenney et al., 2018; Zhu et al., 2018).

Further indications for a possible differentiation of carnivore and omnivore host species can be morphological differences in the simple digestive system. For example, both *Ailurus fulgens* and *Ursus arctos* have no caecum so microbial fermentation takes place in the enlarged large intestine (Roberts and Gittleman, 1984; Stevens and Hume, 1995). This increased intestinal surface might be an adaptation for microbial fermentation of plant material.

Overall, the different digestive systems have a strong impact on the respective microbiota and the microbial taxa necessary for the digestive process are important to distinguish between the digestive system or host-families.

4.2 Developing a model to identify host-specific information from the microbiota

The analytical challenge we addressed is a classification problem, with the aim to predict categorical outcomes such as diet, digestion type or host-family, based on the microbiota composition of various species. For such classification issues, logistic regression and decision trees are commonly employed models. Logistic regression is advantageous when the probability of a class membership is a linear function of the features, and it is particularly effective in binary and multinomial scenarios. Decision trees are favored for their interpretability and ease of handling non-linear relationships and have been used in microbiome studies as well (Roguet et al., 2018). In the present study, however, we decided to use logistic regression instead of random forest models. This improved efficacy can be attributed to logistic regression's resilience against overfitting, especially when the underlying relationship between the predictor variables and the log-odds is linear. Decision trees, while powerful, can sometimes overfit the training data, particularly when the feature space is large, and the model is complex. Given these considerations, logistic regression was selected as the primary model moving forward.

A logistic ensemble model was developed to extract as many host-specific information as possible from an unknown fecal sample. Mainly, the three factors diet, digestive system and host species, that shape the microbiota composition, are of interest. The model accurately distinguishes between the diet type and the herbivore digestive systems with a high model accuracy and nearly perfect F1 scores. This indicates that the model is sound from a computational perspective and furthermore supports the hypothesis that diet and the morphology of the digestive system have a strong influence on the composition of the microbiota (Ley et al., 2008a; Ley et al., 2008b; O' Donnell et al., 2017; Zoelzer et al., 2021). For carnivores, which all have a simple digestive system, we tried to identify the host-family. The overall model accuracy is 79%, but the F1 scores vary between the host-families. The best results are obtained for Canidae (F1=0.93) and Felidae (F1=0.82) samples. To the best of our knowledge, this is one of the first studies that clearly distinguishes between canid and felid microbiota sample. Many studies focusing on the differences between carnivores, herbivores and omnivores found no distinct

clustering of samples from these two groups. In contrast, canids and felids mostly fall into the same cluster and even express a high variability within the respective cluster (Vital et al., 2014; Zhu et al., 2018; Guo et al., 2020). The fact that it is possible to distinguish between these two families using the modelling approach shown here, demonstrates the potential utility of the method in the field of microbiome research. Differences in the microbiota of canids and felids occur e.g. in the respective correlating taxa (Figure 1). While canids correlate with Bacteroidaceae, felids show a correlation with Peptostreptococcaceae that occur in greater proportion in this family. This is in line with the fact that canids are able to cope with higher amounts of dietary carbohydrates and a fiber-rich diet (Deng and Swanson, 2015). Bacteroidaceae are often linked and affected by an increased amount of carbohydrates (Kerr et al., 2013; Panasevich et al., 2015; Vázquez-Baeza et al., 2016) while Peptostreptococcaceae are linked to the dietary protein metabolism (Schulz et al., 2014; Fan et al., 2017). Another aspect in the distinction of the canid and felid microbiota may be phylosymbiosis. Since the Canidae evolved in the Eocene about 40-43 Mya and the Felidae in the Oligocene about 34-24 Mya each microbiota may have evolved with the respective host-family (Wang, 2008; Tedford et al., 2009; Berta, 2011). The classification of the Ursidae results in a F1 score of 0.79, slightly less accurate than the Canidae and Felidae. Overall a distinction of the Ursidae was to be expected. Many descriptive studies on the microbiome of different mammals show that the Ursidae cluster slightly separated from other carnivore species, being grouped together with panda species (Ley et al., 2008a; Zoelzer et al., 2021). During this third step of host-family identification limitations occur in identifying the Herpestidae. Reasons for this may be the small sample size of 20, which is the minimum number that was being applied to the model. Therefore, it is recommended for future studies that for host-family or even species detection, the sample size should be increased.

With an increased sample size, host-family detection would also be possible within the different herbivore digestive systems. The model reaches a very good accuracy for the herbivore digestive systems. Furthermore, Figures 1 and 2 suggest that the microbiota of hindgut fermenters such as the Equidae is distinct from other herbivores, so it makes sense that the model is able to identify this host-family. Furthermore, the model reaches a F1 score of 0.92 for the herbivores with a simple digestive system, namely the Ailuridae. This was also expected. As previously discussed, this family is characterized by a herbivore diet, simple digestive system and belongs to the Carnivora. Because of this unique combination of host-specific factors, and the high model accuracy in discriminating diet and digestive system, the Ailuridae can be identified as host-family.

Overall, we have developed a model that is able to predict the host's diet and digestive system very reliably from a given fecal microbiota composition. Limitations arise in the prediction of host-families which works very well for the carnivores, but more effort is needed to train the model on herbivores. To improve the accuracy of the model, the number of samples per family should be increased considerably. With this enlarged dataset the model is even able to identify host-species without complications. As a guideline, we would use a number of at least 50 samples per species, as this worked well to identify canid and felid host-families. Furthermore, it would be interesting to implement time series data for different individuals per species. As we have shown in a previous study, it is possible to identify individuals, that show a very unique microbiota, based on such datasets (Zoelzer et al., 2023). This was possible because of low-abundant bacterial taxa that varied temporarily within and between individuals. Not only could this improve the model accuracy because of more input data per species, it would also open up a new level beside species recognition – individual identification. Some studies have already successfully identified host-specific factors such as age (Biagi et al., 2012; Yatsunenko et al., 2012; Björk et al., 2019; Low et al., 2022), diet (Ley et al., 2005), health status (Greenblum et al., 2012; Tuddenham and Sears, 2015; Gupta et al., 2020) or even kinship from the microbiota

composition, nevertheless this studies were mostly performed on one species. With an ensemble model approach as presented in this study, such host-specific information could be gained from a variety of species, originated from undisclosed fecal samples.

Conclusion

We were able to show that diet, digestive system and host-family are influencing factors for an animal's microbiota. Based on these findings, we developed a logistic ensemble model that is able to identify the host's diet and digestive system very accurately. Furthermore, the identification of host-families works very well for carnivores, but limitations arose due to small sample sizes. Future studies will overcome these limitations by increasing sample size and including more host species in the model.

Such modelling approaches are very promising in the field of microbiome research as they provide new opportunities to combine the analysis of microbial data with host metadata. This has a wide range of applications, particularly in field studies. Fecal samples can be easily collected from wildlife hotspots such as waterholes or grazing areas, but host information is usually rare or difficult to collect non-invasively. Currently, this is usually done by fecal microsatellite analysis. Unfortunately, this method is cost-intensive because microsatellite loci need to be described for a species or at least a group of animals, and multiple microsatellites are required for species identification (Kurose et al., 2005; Miller et al., 2016; Walker et al., 2019). Another limitation is the need for high quality host DNA, which is possible but difficult to obtain from fecal samples. Therefore, the modelling approach to gain as much host specific information as possible from the fecal microbial composition is promising as it is more cost- and labor efficient than traditional approaches and opens up a new field in microbiome research.

Figure 1: Pearson correlation between different host-specific groups and microbial taxa. **(A)** shows the correlation of diet, **(B)** the correlation of the digestive system and **(c)** the correlation of the host-family with microbial families. The coloring scheme ranges from highly negatively correlating (dark blue) to highly positively correlating (dark red) taxa. For this analysis a threshold of the correlation coefficients was set to a minimum of $R>0.3$.

Figure 2: Average taxonomic microbial composition of different host-specific groups. Only the respective correlating taxa ($R>0.3$) are colored to show the differences within diet type, digestive system and host-family. Furthermore, the F1 score of the logistic regression model is shown for each category.

Table 1: Logistic ensemble model to characterize an unknown fecal sample. The different levels are shown in bold in the first column including the respective categories. The overall model accuracy, F1 scores and support are given in the following columns.

	F1 score	Support	Accuracy
Diet			0.88
herbivore	0.93	48	
carnivore	0.87	45	
omnivore	0.71	12	
Digestive system			0.98
ruminant	0.98	25	
hindgut	1.00	17	
simple	0.92	6	
Host-family			0.79
Canidae	0.93	14	
Felidae	0.82	27	
Herpestidae	0.00	4	
Ursidae	0.79	12	

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

Franziska Zoelzer: Conceptualization (lead); formal analysis (lead); methodology (equal); writing – original draft (lead). Daniel Monteiro: Conceptualization (supporting), software (lead), methodology (equal), writing – review and editing (equal). Paul W. Dierkes: Funding acquisition (lead), supervision (lead), writing – review and editing (equal).

7 Funding

Open-access funding enabled and organized by Open-Access publication fond of the Goethe University. F. Zoelzer and P.W. Dierkes were supported by Opel-Zoo Foundation Professorship in Zoo Biology from the “von Opel Hessische Zoostiftung” and by the “Vereinigung von Freunden und Förderern der Goethe-Universität e.V.”.

8 Acknowledgments

We gratefully acknowledge support from the directors, curators and especially keepers of Berlin zoo, Tierpark Berlin, Vivarium Darmstadt, Zoo Dortmund, Zoo Duisburg, Zoo Frankfurt, Zoom Erlebniswelt Gelsenkirchen, Zoo Heidelberg, Kölner Zoo, Zoo Krefeld, Zoo Landau in der Pfalz, Zoo Neuwied, Tiergarten Nürnberg, Opel-Zoo, Zoo Schwerin, Wilhelma Stuttgart and Wuppertal Zoo, who enabled this project.

9 Data Availability Statement

The dataset generated and analyzed for this study can be found in the NCBI SRA under the BioProjects PRJNA1068843 (all samples), PRJNA912880 (all samples) and PRJNA716130 (BioSamples: SAMN18396855-SAMN18396844, SAMN18396842-SAMN18396841, SAMN18396837-SAMN18396824, SAMN18396821-SAMN18396812, SAMN18396808-SAMN18396806, SAMN18396799-SAMN18396795, SAMN18396791-SAMN1839679, SAMN18396783-SAMN18396762, SAMN18396755-SAMN18396754, SAMN18396750-SAMN18396741, SAMN18396732-SAMN18396730, SAMN18396727-SAMN18396722, SAMN18396705-SAMN18396678, SAMN18396676-SAMN18396652, SAMN18396649, SAMN18396644-SAMN18396643, SAMN18396638-SAMN18396636, SAMN18396628-SAMN18396623, SAMN18396620-SAMN18396618, SAMN18396614-SAMN18396612, SAMN18396608, SAMN18396602-SAMN18396585, SAMN18396583, SAMN18396581-SAMN18396577, SAMN18396573-SAMN18396566, SAMN18396560-SAMN18396555, SAMN18396551-SAMN18396523, SAMN18396520-SAMN18396512, SAMN18396505-SAMN18396492, SAMN18396485-SAMN18396480, SAMN18396477-SAMN18396476, SAMN18396468-SAMN18396467, SAMN18396464-SAMN18396459, SAMN18396456-SAMN18396455, SAMN18396452-SAMN18396450, SAMN18396445-SAMN18396433, SAMN18396422-SAMN18396421, SAMN18396411-SAMN18396397, SAMN18396392-SAMN18396386, SAMN18396384-SAMN18396383, SAMN18396377-SAMN18396372, SAMN18396364-SAMN18396363, SAMN18396359-SAMN18396353, SAMN18396351-SAMN18396346, SAMN18396344-SAMN18396343, SAMN18396337-SAMN18396325, SAMN18396323-SAMN18396313, SAMN18396309-SAMN18396308, SAMN18396303-SAMN18396273).

The logistic model is available on github: https://github.com/frankfurt-zoobiology/microbiota_ensemble_model

REFERENCES

Abdul Rahman, N., Parks, D. H., Vanwonterghem, I., Morrison, M., Tyson, G. W., and Hugenholtz, P. (2015). A Phylogenomic Analysis of the Bacterial Phylum Fibrobacteres. *Front Microbiol* 6, 1469. doi: 10.3389/fmicb.2015.01469

An, C., Okamoto, Y., Xu, S., Eo, K. Y., Kimura, J., and Yamamoto, N. (2017). Comparison of fecal microbiota of three captive carnivore species inhabiting Korea. *J Vet Med Sci* 79, 542–546. doi: 10.1292/jvms.16-0472

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* 75, 129–137. doi: 10.3354/ame01753

536 Badri, D. V., Jackson, M. I., and Jewell, D. E. (2021). Dietary Protein and Carbohydrate Levels
537 Affect the Gut Microbiota and Clinical Assessment in Healthy Adult Cats. *J Nutr* 151, 3637–3650.
538 doi: 10.1093/jn/nxab308

539 Berta, A. (2011). Carnivoran Evolution: New Views on Phylogeny, Form and Function. *Systematic*
540 *Biology* 60, 241–243. doi: 10.1093/sysbio/syq087

541 Biagi, E., Candela, M., Fairweather-Tait, S., Franceschi, C., and Brigidi, P. (2012). Aging of the
542 human metaorganism: the microbial counterpart. *Age (Dordr)* 34, 247–267. doi: 10.1007/s11357-
543 011-9217-5

544 Björk, J. R., Dasari, M., Grieneisen, L., and Archie, E. A. (2019). Primate microbiomes over time:
545 Longitudinal answers to standing questions in microbiome research. *Am J Primatol* 81, e22970.
546 doi: 10.1002/ajp.22970

547 Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018).
548 Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-
549 feature-classifier plugin. *Microbiome* 6, 90. doi: 10.1186/s40168-018-0470-z

550 Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., et al.
551 (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol*
552 *Evol* 24, 127–135. doi: 10.1016/j.tree.2008.10.008

553 Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al.
554 (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2.
555 *Nat Biotechnol* 37, 852–857. doi: 10.1038/s41587-019-0209-9

556 Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P.
557 (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13,
558 581–583. doi: 10.1038/nmeth.3869

559 Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012).
560 Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms.
561 *ISME J* 6, 1621–1624. doi: 10.1038/ismej.2012.8

562 Clauss, M., Lechner-Doll, M., and Streich, W. J. (2003). Ruminant diversification as an adaptation to
563 the physicomechanical characteristics of forage. *Oikos* 102, 253–262. doi: 10.1034/j.1600-
564 0706.2003.12406.x

565 Deng, P., and Swanson, K. S. (2015). Gut microbiota of humans, dogs and cats: current knowledge
566 and future opportunities and challenges. *Br J Nutr* 113 Suppl, S6-17. doi:
567 10.1017/S0007114514002943

568 Douglas, A. E. (2018). *Fundamentals of Microbiome Science*. Princeton University Press.

569 Duvallet, C., Gibbons, S. M., Gurry, T., Irizarry, R. A., and Alm, E. J. (2017). Meta-analysis of gut
570 microbiome studies identifies disease-specific and shared responses. *Nat Commun* 8, 1784. doi:
571 10.1038/s41467-017-01973-8

572 Edwards, J. E., Shetty, S. A., van den Berg, P., Burden, F., van Doorn, D. A., Pellikaan, W. F., et al.
573 (2020). Multi-kingdom characterization of the core equine fecal microbiota based on multiple
574 equine (sub)species. *Anim Microbiome* 2, 6. doi: 10.1186/s42523-020-0023-1

575 Espinoza, J. L. (2018). Machine learning for tackling microbiota data and infection complications in
576 immunocompromised patients with cancer. *J Intern Med*. doi: 10.1111/joim.12746

577 Fan, P., Liu, P., Song, P., Chen, X., and Ma, X. (2017). Moderate dietary protein restriction alters the
578 composition of gut microbiota and improves ileal barrier function in adult pig model. *Sci Rep* 7,
579 43412. doi: 10.1038/srep43412

580 Fu, H., Zhang, L., Fan, C., Liu, C., Li, W., Cheng, Q., et al. (2021). Environment and host species
581 identity shape gut microbiota diversity in sympatric herbivorous mammals. *Microb Biotechnol* 14,
582 1300–1315. doi: 10.1111/1751-7915.13687

583 Giesecke, D., and van Gylswyk, N. O. (1975). A study of feeding types and certain rumen functions
584 in six species of South African wild ruminants. *J. Agric. Sci.* 85, 75–83. doi:
585 10.1017/S0021859600053430

586 Ginnett, T. F., and Demment, M. W. (1997). Sex differences in giraffe foraging behavior at two
587 spatial scales. *Oecologia* 110, 291–300. doi: 10.1007/s004420050162

588 Greenblum, S., Turnbaugh, P. J., and Borenstein, E. (2012). Metagenomic systems biology of the
589 human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel
590 disease. *Proc Natl Acad Sci U S A* 109, 594–599. doi: 10.1073/pnas.1116053109

591 Guo, X., Lei, H., Zhang, K., Ke, F., and Song, C. (2020). Diversification of animal gut microbes and
592 NRPS gene clusters in some carnivores, herbivores and omnivores. *Biotechnology &*
593 *Biotechnological Equipment* 34, 1280–1287. doi: 10.1080/13102818.2020.1835536

594 Gupta, V. K., Kim, M., Bakshi, U., Cunningham, K. Y., Davis, J. M., Lazaridis, K. N., et al. (2020).
595 A predictive index for health status using species-level gut microbiome profiling. *Nat Commun* 11,
596 4635. doi: 10.1038/s41467-020-18476-8

597 Hofmann, R. R. (1989). Evolutionary steps of ecophysiological adaptation and diversification of
598 ruminants: a comparative view of their digestive system. *Oecologia* 78, 443–457. doi:
599 10.1007/BF00378733

600 Kartzinel, T. R., Hsing, J. C., Musili, P. M., Brown, B. R. P., and Pringle, R. M. (2019). Covariation
601 of diet and gut microbiome in African megafauna. *Proc Natl Acad Sci U S A* 116, 23588–23593.
602 doi: 10.1073/pnas.1905666116

603 Katoh, K., Misawa, K., Kuma, K.-i., and Miyata, T. (2002). MAFFT: a novel method for rapid
604 multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30, 3059–3066.
605 doi: 10.1093/nar/gkf436

606 Kerr, K. R., Forster, G., Dowd, S. E., Ryan, E. P., and Swanson, K. S. (2013). Effects of dietary
607 cooked navy bean on the fecal microbiome of healthy companion dogs. *PLoS One* 8, e74998. doi:
608 10.1371/journal.pone.0074998

609 Knowles, S. C. L., Eccles, R. M., and Baltrūnaitė, L. (2019). Species identity dominates over
610 environment in shaping the microbiota of small mammals. *Ecol Lett* 22, 826–837. doi:
611 10.1111/ele.13240

612 Koh, A., Vadder, F. de, Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From Dietary Fiber to
613 Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* 165, 1332–1345. doi:
614 10.1016/j.cell.2016.05.041

615 Korpela, K., Flint, H. J., Johnstone, A. M., Lappi, J., Poutanen, K., Dewulf, E., et al. (2014). Gut
616 microbiota signatures predict host and microbiota responses to dietary interventions in obese
617 individuals. *PLoS One* 9, e90702. doi: 10.1371/journal.pone.0090702

618 Kurose, N., Masuda, R., and Tatara, M. (2005). Fecal DNA analysis for identifying species and sex
619 of sympatric carnivores: a noninvasive method for conservation on the Tsushima Islands, Japan. *J*
620 *Hered* 96, 688–697. doi: 10.1093/jhered/esi124

621 La Reau, A. J., and Suen, G. (2018). The Ruminococci: key symbionts of the gut ecosystem. *J*
622 *Microbiol* 56, 199–208. doi: 10.1007/s12275-018-8024-4

623 Ley, R. E., Bäckhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., and Gordon, J. I. (2005).
624 Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* 102, 11070–11075. doi:
625 10.1073/pnas.0504978102

626 Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., et al. (2008a).
627 Evolution of mammals and their gut microbes. *Science* 320, 1647–1651. doi:
628 10.1126/science.1155725

629 Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R., and Gordon, J. I. (2008b). Worlds within
630 worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* 6, 776–788. doi:
631 10.1038/nrmicro1978

632 Li, J., Zhan, S., Liu, X., Lin, Q., Jiang, J., and Li, X. (2018). Divergence of Fecal Microbiota and
633 Their Associations With Host Phylogeny in Cervinae. *Front Microbiol* 9, 1823. doi:
634 10.3389/fmicb.2018.01823

635 Lim, S. J., and Bordenstein, S. R. (2020). An introduction to phylosymbiosis. *Proc Biol Sci* 287. doi:
636 10.1098/rspb.2019.2900

637 Low, A., Soh, M., Miyake, S., and Seedorf, H. (2022). Host Age Prediction from Fecal Microbiota
638 Composition in Male C57BL/6J Mice. *Microbiol Spectr* 10, e0073522. doi:
639 10.1128/spectrum.00735-22

640 Mackie, R. I. (2002). Mutualistic fermentative digestion in the gastrointestinal tract: diversity and
641 evolution. *Integr Comp Biol* 42, 319–326. doi: 10.1093/icb/42.2.319

642 Martínez-López, L. M., Pepper, A., Pilla, R., Woodward, A. P., Suchodolski, J. S., and Mansfield, C.
643 (2021). Effect of sequentially fed high protein, hydrolyzed protein, and high fiber diets on the fecal
644 microbiota of healthy dogs: a cross-over study. *Anim Microbiome* 3, 42. doi: 10.1186/s42523-021-
645 00101-8

646 McKenney, E. A., Maslanka, M., Rodrigo, A., and Yoder, A. D. (2018). Bamboo Specialists from
647 Two Mammalian Orders (Primates, Carnivora) Share a High Number of Low-Abundance Gut
648 Microbes. *Microb Ecol* 76, 272–284. doi: 10.1007/s00248-017-1114-8

649 Milani, C., Alessandri, G., Mancabelli, L., Mangifesta, M., Lugli, G. A., Viappiani, A., et al. (2020).
650 Multi-omics Approaches To Decipher the Impact of Diet and Host Physiology on the Mammalian
651 Gut Microbiome. *Appl Environ Microbiol* 86. doi: 10.1128/AEM.01864-20

652 Miller, S. M., Clarke, A. B., Bloomer, P., Guthrie, A. J., and Harper, C. K. (2016). Evaluation of
653 microsatellites for common ungulates in the South African wildlife industry. *Conservation Genet*
654 *Resour* 8, 329–341. doi: 10.1007/s12686-016-0554-7

655 Mitchell, G. (2021). "The Machinery of Digestion," in *How giraffes work*, ed. G. Mitchell (New
656 York: Oxford University Press), 147–186.

657 Mitchell, P. C. (1905). On the Intestinal Tract of Mammals. *The Transactions of the Zoological*
658 *Society of London* 17, 437–536. doi: 10.1111/j.1096-3642.1905.tb00036.x

659 Montgomery, L., Flesher, B., and Stahl, D. (1988). Transfer of Bacteroides succinogenes (Hungate)
660 to Fibrobacter gen. nov. as Fibrobacter succinogenes comb. nov. and Description of Fibrobacter
661 intestinalis sp. nov. *International Journal of Systematic Bacteriology* 38, 430–435. doi:
662 10.1099/00207713-38-4-430

663 Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., González, A., Fontana, L., et al. (2011).
664 Diet drives convergence in gut microbiome functions across mammalian phylogeny and within
665 humans. *Science* 332, 970–974. doi: 10.1126/science.1198719

666 Nelson, T. M., Apprill, A., Mann, J., Rogers, T. L., and Brown, M. V. (2015). The marine mammal
667 microbiome: current knowledge and future directions. *Microbiol. Aust.* 36, 8. doi:
668 10.1071/MA15004

669 Nishida, A. H., and Ochman, H. (2018). Rates of gut microbiome divergence in mammals. *Mol Ecol*
670 27, 1884–1897. doi: 10.1111/mec.14473

671 O' Donnell, M. M., Harris, H. M. B., Ross, R. P., and O'Toole, P. W. (2017). Core fecal microbiota
672 of domesticated herbivorous ruminant, hindgut fermenters, and monogastric animals.
673 *Microbiologyopen* 6. doi: 10.1002/mbo3.509

674 Obregon-Tito, A. J., Tito, R. Y., Metcalf, J., Sankaranarayanan, K., Clemente, J. C., Ursell, L. K., et
675 al. (2015). Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat Commun*
676 6, 6505. doi: 10.1038/ncomms7505

677 Panasevich, M. R., Kerr, K. R., Dilger, R. N., Fahey, G. C., Guérin-Deremaux, L., Lynch, G. L., et
678 al. (2015). Modulation of the faecal microbiome of healthy adult dogs by inclusion of potato fibre
679 in the diet. *Br J Nutr* 113, 125–133. doi: 10.1017/S0007114514003274

680 Pannoni, S. B., Proffitt, K. M., and Holben, W. E. (2022). Non-invasive monitoring of multiple
681 wildlife health factors by fecal microbiome analysis. *Ecol Evol* 12, e8564. doi: 10.1002/ece3.8564

682 Panthi, S., Coogan, S. C. P., Aryal, A., and Raubenheimer, D. (2015). Diet and nutrient balance of
683 red panda in Nepal. *Naturwissenschaften* 102, 54. doi: 10.1007/s00114-015-1307-2

684 Pascale, A., Marchesi, N., Marelli, C., Coppola, A., Luzi, L., Govoni, S., et al. (2018). Microbiota
685 and metabolic diseases. *Endocrine* 61, 357–371. doi: 10.1007/s12020-018-1605-5

686 Pradhan, S., Saha, G. K., and Khan, J. A. (2001). Ecology of the red panda Ailurus fulgens in the
687 Singhalila National Park, Darjeeling, India. *Biological Conservation* 98, 11–18. doi:
688 10.1016/S0006-3207(00)00079-3

689 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA
690 ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic*
691 *Acids Res* 41, D590-6. doi: 10.1093/nar/gks1219

692 R Core Team (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R
693 Foundation for Statistical Computing.

694 Ríos-Covián, D., Ruas-Madiedo, P., Margolles, A., Gueimonde, M., Los Reyes-Gavilán, C. G. de,
695 and Salazar, N. (2016). Intestinal Short Chain Fatty Acids and their Link with Diet and Human
696 Health. *Front Microbiol* 7, 185. doi: 10.3389/fmicb.2016.00185

697 Roberts, M. S., and Gittleman, J. L. (1984). Ailurus fulgens. *Mammalian Species*, 1. doi:
698 10.2307/3503840

699 Roggenbuck, M., Sauer, C., Poulsen, M., Bertelsen, M. F., and Sørensen, S. J. (2014). The giraffe
700 (Giraffa camelopardalis) rumen microbiome. *FEMS Microbiol Ecol* 90, 237–246. doi:
701 10.1111/1574-6941.12402

702 Roguet, A., Eren, A. M., Newton, R. J., and McLellan, S. L. (2018). Fecal source identification using
703 random forest. *Microbiome* 6, 185. doi: 10.1186/s40168-018-0568-3

704 Rojas, C. A., Ramírez-Barahona, S., Holekamp, K. E., and Theis, K. R. (2021). Host phylogeny and
705 host ecology structure the mammalian gut microbiota at different taxonomic scales. *Anim*
706 *Microbiome* 3, 33. doi: 10.1186/s42523-021-00094-4

707 Sanna, S., van Zuydam, N. R., Mahajan, A., Kurilshikov, A., Vich Vila, A., Võsa, U., et al. (2019).
708 Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases.
709 *Nat Genet* 51, 600–605. doi: 10.1038/s41588-019-0350-x

710 Schulz, M. D., Atay, C., Heringer, J., Romrig, F. K., Schwitalla, S., Aydin, B., et al. (2014). High-
711 fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity. *Nature*
712 514, 508–512. doi: 10.1038/nature13398

713 Sommer, F., Ståhlman, M., Ilkayeva, O., Arnemo, J. M., Kindberg, J., Josefsson, J., et al. (2016). The
714 Gut Microbiota Modulates Energy Metabolism in the Hibernating Brown Bear Ursus arctos. *Cell*
715 *Rep* 14, 1655–1661. doi: 10.1016/j.celrep.2016.01.026

716 Stevens, C. E., and Hume, I. D. (1995). *Comparative physiology of the vertebrate digestive system*.
717 Cambridge: Cambridge University Press.

718 Sweeny, A. R., Lemon, H., Ibrahim, A., Watt, K. A., Wilson, K., Childs, D. Z., et al. (2023). A
719 mixed-model approach for estimating drivers of microbiota community composition and
720 differential taxonomic abundance. *mSystems* 8, e0004023. doi: 10.1128/msystems.00040-23

721 Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., and Nishijima, M. (2014). Development of a
722 prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-
723 generation sequencing. *PLoS One* 9, e105592. doi: 10.1371/journal.pone.0105592

724 Tedford, R. H., Wang, X., and Taylor, B. E. (2009). Phylogenetic Systematics of the North American
725 Fossil Caninae (Carnivora: Canidae). *Bulletin of the American Museum of Natural History* 325, 1–
726 218. doi: 10.1206/574.1

727 Trujillo, S. M., McKenney, E. A., Hilderbrand, G. V., Mangipane, L. S., Rogers, M. C., Joly, K., et
728 al. (2022a). Correlating gut microbial membership to brown bear health metrics. *Sci Rep* 12,
729 15415. doi: 10.1038/s41598-022-19527-4

730 Trujillo, S. M., McKenney, E. A., Hilderbrand, G. V., Mangipane, L. S., Rogers, M. C., Joly, K., et
731 al. (2022b). Intrinsic and extrinsic factors influence on an omnivore's gut microbiome. *PLoS One*
732 17, e0266698. doi: 10.1371/journal.pone.0266698

733 Tuddenham, S., and Sears, C. L. (2015). The Intestinal Microbiome and Health. *Curr Opin Infect Dis*
734 28, 464–470. doi: 10.1097/QCO.0000000000000196

735 van den Abbeele, P., Ghyselinck, J., Marzorati, M., Koch, A.-M., Lambert, W., Michiels, J., et al.
736 (2022). The Effect of Amino Acids on Production of SCFA and bCFA by Members of the Porcine
737 Colonic Microbiota. *Microorganisms* 10. doi: 10.3390/microorganisms10040762

738 van Rossum, G., and Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA:
739 CreateSpace.

740 Vázquez-Baeza, Y., Hyde, E. R., Suchodolski, J. S., and Knight, R. (2016). Dog and human
741 inflammatory bowel disease rely on overlapping yet distinct dysbiosis networks. *Nat Microbiol* 1,
742 16177. doi: 10.1038/nmicrobiol.2016.177

743 Vital, M., Gao, J., Rizzo, M., Harrison, T., and Tiedje, J. M. (2014). Diet is a major factor governing
744 the fecal butyrate-producing community structure across Mammalia, Aves and Reptilia. *ISME J* 9,
745 832–843. doi: 10.1038/ismej.2014.179

746 Walker, F. M., Tobin, A., Simmons, N. B., Sobek, C. J., Sanchez, D. E., Chambers, C. L., et al.
747 (2019). A fecal sequel: Testing the limits of a genetic assay for bat species identification. *PLoS*
748 *One* 14, e0224969. doi: 10.1371/journal.pone.0224969

749 Wang, X. (2008). *Dogs: Their Fossil Relatives and Evolutionary History*. New York: Columbia
750 University Press.

751 Woodall, P. F., and Skinner, J. D. (1993). Dimensions of the intestine, diet and faecal water loss in
752 some African antelope. *Journal of Zoology* 229, 457–471. doi: 10.1111/j.1469-
753 7998.1993.tb02648.x

754 Wu, X., Wei, Q., Wang, X., Shang, Y., and Zhang, H. (2022). Evolutionary and dietary relationships
755 of wild mammals based on the gut microbiome. *Gene* 808, 145999. doi:
756 10.1016/j.gene.2021.145999

757 Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., et al.
758 (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227. doi:
759 10.1038/nature11053

760 Yazdani, M., Taylor, B. C., Debelius, J. W., Li, W., Knight, R., and Smarr, L. "Using machine
761 learning to identify major shifts in human gut microbiome protein family abundance in disease," in
762 *2016*, 1272–1280.

763 Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., et al. (2019).
764 Host diet and evolutionary history explain different aspects of gut microbiome diversity among
765 vertebrate clades. *Nat Commun* 10, 2200. doi: 10.1038/s41467-019-10191-3

766 Zhu, L., Yang, Z., Yao, R., Xu, L., Chen, H., Gu, X., et al. (2018). Potential Mechanism of
767 Detoxification of Cyanide Compounds by Gut Microbiomes of Bamboo-Eating Pandas. *mSphere*
768 3. doi: 10.1128/mSphere.00229-18

769 Zoelzer, F., Burger, A. L., and Dierkes, P. W. (2021). Unraveling differences in fecal microbiota
770 stability in mammals: from high variable carnivores and consistently stable herbivores. *Anim*
771 *Microbiome* 3, 77. doi: 10.1186/s42523-021-00141-0

772 Zoelzer, F., Schneider, S., and Dierkes, P. W. (2023). Time series cluster analysis reveals individual
773 assignment of microbiota in captive tiger (Panthera tigris) and wildebeest (Connochaetes taurinus).
774 *Ecol Evol* 13, e10066. doi: 10.1002/ece3.10066