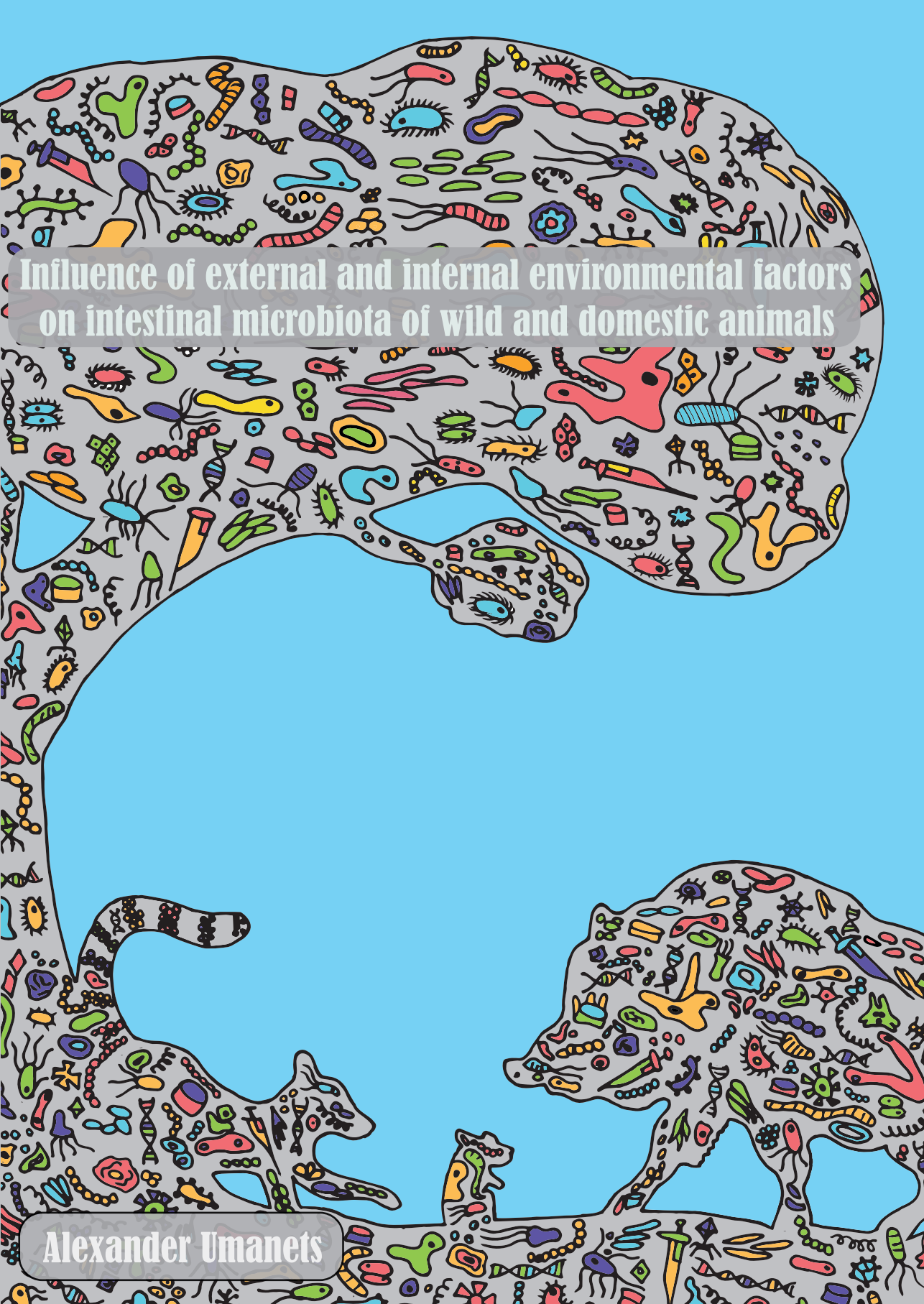

Influence of external and internal environmental factors on intestinal microbiota of wild and domestic animals

Alexander Umanets

Propositions

1. Intestinal microbiota and resistome composition of wild animals are mostly shaped by the animals' diet and lifestyle.
   (this thesis)

2. When other environmental factors are controlled, genetics of the host lead to species- or breed specific microbiota patterns.
   (this thesis)

3. Identifying the response of microbial communities to factors that only have a minor contribution to overall microbiota variation faces the same problems as the discovery of exoplanets.

4. Observational studies in microbial ecology using cultivation-independent methods should be considered only as a guide for further investigations that employ controlled experimental conditions and mechanistic studies of cause-effect relationships.

5. Public fear of genetic engineering and artificial intelligence is not helped by insufficient public education and misleading images created through mass- and social media.

6. Principles of positive (Darwinian) and negative selection govern the repertoire of techniques used within martial arts.

Propositions belonging to the thesis, entitled
Influence of external and internal environmental factors on intestinal microbiota of wild and domestic animals
Alexander Umanets
Wageningen, 17 October 2019

Influence of external and internal environmental factors on intestinal microbiota of wild and domestic animals

Alexander Umanets

Thesis committee

Promotor

Prof. Dr H. Smidt

Personal chair at the Laboratory of Microbiology

Wageningen University & Research

Other members

Prof. Dr Michiel Kleerebezem, Wageningen University & Research

Dr John Penders, Maastricht University

Dr Leo Lahti, University of Turku, Finland

Dr Nadia Everaert, University of Liège, Belgium

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences)

Influence of external and internal environmental factors on intestinal microbiota of wild and domestic animals

Alexander Umanets

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 17 October 2019
at 11 a.m. in the Aula.

Alexander Umanets
Influence of external and internal environmental factors on intestinal microbiota of wild and domestic animals,
208 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2019)
With references, with summary in English.

ISBN: 978-94-6395-093-0
DOI: https://doi.org/10.18174/498870

Table of contents

Chapter 1

General introduction and thesis outline

They were the first

Microbes. They have been around almost since the beginning of times on Earth - the first life form appeared on the face of our planet around 3.45 billion years ago [1], and only a billion years after planet Earth was formed from a hot cloud of gasses and space dust [2]. Much later, around 600 million years ago in the Late Precambrian period, the first multicellular organisms appeared [3], which means that single-celled organisms had seven times more time to evolve and adapt, and they didn't waste it. A tree of life published in 2016 by Laura A. Hug et al. [4] depicted 92 bacterial and 26 archaeal phyla, providing a good visual representation of the tremendous diversity of prokaryotic life. Phylogenetic diversity of prokaryotes goes hand to hand with their ability to colonise the widest range of natural and man-made environments and to utilise various chemical compounds to recover energy. Microbial life can be found practically everywhere: in hot springs near volcanoes [5], the Arctic sea [6], outside and inside of human and animal bodies [7], and even possibly on other planets [8].

Since its origin, eukaryotic and later also multicellular life has never been isolated from its prokaryotic ancestors and relatives. The life on Earth became more complex with time, including intricate relationships between all its domains. Looking at the modern time eukaryotic-prokaryotic relationships we can find fascinating examples where the entire life of an animal depends on a symbiosis with microorganisms. It is known that more than 10% of insects rely heavily in their metabolism on obligate bacterial symbionts [9]. A textbook example is the presence of a special type of cells (bacteriocytes) in aphids that host an obligatory bacterial symbiont, *Buchnera aphidicola*, that covers the aphid's needs in amino acids in exchange for shelter and sugar [10]. In the sea bacteria-host mutualism can go into extremes, as exemplified by some oligochaetes such as *Olavius crassitunicatus* that completely lost (or never developed) a digestive system and rely for the acquisition of nutrients completely on microbial communities residing within their body that produce organic nutrients via chemosynthesis [11]. When we are talking about mammals we will

not find such exotic examples of bacteria-host mutualism, however, the importance of microbes in their life and wellbeing can hardly be overestimated.

Microbes and humanity

Keen humanity has been using microorganisms unintentionally for their own good as long as civilization exists. One of the first domesticated microorganisms were probably yeasts that were used for beer brewing and bread making as early as 10,000 years ago, followed by *Streptococcus* and *Lactobacillus* species that came to play an important role together with domestication of milk producing animals and the discovery of savouriness and health benefits of fermented milk products [12]. Application of microorganisms became more and more complex alongside with technological progress, allowing humanity to tackle challenges that were not feasible before, such as production of valuable chemical compounds and pharmaceuticals using genetically modified microorganisms [13]. However, until only a few years back microbes were perceived almost exclusively as something inherently harmful when human and animal health were considered. The negative reputation of bacteria is not surprising and started with the postulation of microbial aetiology of diseases by Louis Pasteur and Robert Koch [14]. Unofficially, a war on microbes had been declared.

Only in recent years a notion about "beneficial microorganisms" entered the mass media and has ever since been a hot topic. However, more than a hundred years ago Ilya Mechnikov was already convinced that certain lactic-acid bacteria could be a reason for human longevity, and he was actively working on a theory about the interaction between host and intestinal microorganisms [15, 16]. A few years later Alfred Nissle isolated the first well defined and probably the most widely used probiotic strain, *Escherichia coli* Nissle 1917, with remarkable antagonistic properties against *Salmonella* [17]. We went a long way from the discoveries of Ilya Mechnikov and Alfred Nissle, and the modern age of high-throughput technologies and big data allowed us to take a look not only at individual members of microbial communities, but also consider microbial

consortia as a whole. In particular, bacteria dwelling along the gastrointestinal (GI) tract of humans and animals attracted a lot of attention and showed to be a fascinating topic of research. Currently the term "microbiota" is used to depict these and other microbial communities, but also the term "microbiome" is often used to describe microbial composition, function and environmental factors together.

The origin of the term microbiome itself is not clear, and often Joshua Lederberg is credited as a person who used it first in 2001. However, the term "microbiota" has been used since 1960 for handling of germ-free animals, and in 1988 the book "Mycoparasitism and plant disease control" provided a definition [18]. "Characteristic microbial community occupying a reasonably well defined habitat which has distinct physio-chemical properties" was that first definition given to microbiome [19]. It is clear that originally the term microbiome was defined as a microbial community living in any ecological system, which does not have to be the human body. Until now we can see that a human centric definition is widely used. For example, the Merriam-Webster dictionary makes an emphasis on "microorganisms living in or on the human body" in an article on the word "microbiota" (https://www.merriam-webster.com/dictionary/microbiome). However, this is not a surprising "mistake" - the vast majority of microbiome research is dedicated to microbial communities associated with the human body, with those residing in the GI tract leading among any others.

GI tract microbiota and health

The GI tract microbiota is regarded as an extra organ, due to its importance in maintenance of overall health of an individual [20]. More than 40 years ago, it was observed that germ-free animals (free of microorganisms) show underdevelopment of the intestinal lining and lymphoid tissue, absence of antigens in the blood and overall abnormal immune response, as well as decreased capacity for nutrient absorption [21]. Still today in 2019 we continue to discover new and exciting correlations between microbiota and various aspects of

human and animal health. A number of convincing studies showed correlations between gut microbiota and, e.g., metabolic syndrome, liver disease, colorectal cancer inflammatory bowel disease, obesity-related disease [20], stress, anxiety [22] and drug absorption and efficiency [23].

Mechanistically speaking the microbiota and its host have several ways of interaction, including, e.g., production of bioactive chemical compounds by microorganisms, direct exposure of the host to microbial cells, alteration of food consistency and bioavailability of nutrients. Also members of the GI tract microbiota can indirectly influence the host via interactions with intestinal pathogens.

Using ingested food- and host derived compounds, such as mucin and cell debris, the gut microbiota produces a large variety of bioactive compounds that play an important role in host metabolism. A large body of research is dedicated to the microbial production of short chain fatty acids (SCFAs) from dietary fibre and host derived compounds such as mucin [24]. In particular, three SCFAs have been well studied: butyrate has been shown to be a primary energy source for colonocytes and to be involved in inhibition of histone deacetylases [25]; acetate plays a role in maintenance of intestinal homeostasis via promotion of intestinal IgA responses [26] and affects the parasympathic nervous system [26]; and propionate that was reported to reduce cafncer cell proliferation in liver [27], increase insulin sensitivity, and regulate cholesterol levels and blood pressure [28, 29]. Also, it has been shown that members of the gut microbiota can produce free radical nitric oxide (NO), which is regarded as a signalling molecule that regulates blood flow in intestinal tissues, smooth muscles tonus, and immunological responses [30]. In addition, various polyphenolic compounds produced by gut microbiota showed to have an anti-inflammatory effect [31]. Finally, the production of very specific host binding compounds by specific bacterial taxa has been shown, such as aryl hydrocarbon receptor ligands that are produced by *Lactobacillus* spp. from tryptophan fermentation [32].

Besides production of metabolites bacterial cells themselves and their parts can stimulate an immune-response of a host. Host-microbe interaction is

happening via interaction between host enterocytes and dendritic cells with microbial flagella, fimbriae, secreted surface proteins, glycan ligands, cell wall-associated polysaccharides, lipoteichoic acids, peptidoglycans and lipopolysaccharides. To this end, it was shown that the probiotic effect of lactic-acid bacteria probably relies, among other factors, on their richly protein-decorated Gram-positive cell walls [33]. A well know example of immunomodulatory effects of gut bacteria is induction of Th17 cell maturation by close proximity of segmented filamentous bacteria to enterocytes in the developing mouse intestine [34].

Members of every microbial community interact with each other in various direct and indirect ways, forming complex ecological networks [35, 36]. When microbiota-pathogen interactions are considered, however, competition and antagonistic relationships are most interesting for researchers. Saprophytic members of the microbiota could inhibit growth and colonisation of intestinal pathogens via direct inhibition by excreted antimicrobial compounds, or indirectly via competition for substrates and living space [37]. Interactions between enterohaemorrhagic *Escherichia coli* serotype O157:H7, *Salmonella Typhimurium*, *Clostridioides difficile* and microbiota have been carefully studied for the past few years in order to provide mechanistic understanding of the interplay between above-listed pathogens and other members of the GI tract microbiota. This has led to the discovery of effective intervention strategies via modulation of gut microbiota, such as faecal transplantation in case of *C. difficile* overgrowth [38]. In turn, it was shown that disturbance of microbiota by antibiotics promotes colonization of mice by *Salmonella enterica* [39].

Methods for microbiota exploration

The tremendous complexity of GI tract microbial communities poses a number of methodological challenges towards elucidating the true composition as well as the complex relationships between community members. Recent acceleration of GI tract microbiota research and rapid advances in understanding

has relied almost exclusively on the development and implementation of new molecular methods, and most prominently, next generation sequencing technologies (NGS). Historically, microbiology has relied on classical cultivation methods, when a scientist targeting certain microorganisms would design a (hopefully) appropriate nutritious subtract and cultivation conditions. Despite the fundamental importance of the above-mentioned approaches in clinical and environmental microbiology, cultivation has several limitations. First and most importantly it is practically impossible to isolate and culture all, or even a small fraction of microbial species that represent a complex community typically comprising hundreds or thousands of different microbial taxa. Furthermore, a number of microorganisms has exquisitely fastidious growth requirements, or is only able to grow in the presence of a syntrophic partner. Also, a complex microbiome could be comprised of numerous ecological niches that will require completely different approaches for microbial isolation. Sometimes bacterial species that are present in a complex microbial community, but have never been cultured, are referred to as "microbial dark matter", making a direct analogy with cosmic dark matter that outweighs the visible universe by six-fold, but could not be observed directly, or detected yet (https://home.cern/science/physics/dark-matter). Interestingly, the fraction of cells belonging to uncultured bacterial species that has been estimated to comprise approximately 81% (or 7.3×10^{29} cells) of the total number of bacterial and archaeal cells on Earth [40] is the same as the ratio of dark to visible matter in the universe.

A breakthrough in understanding of microbial complexity came together with the introduction of DNA sequencing followed by NGS. The comparably cheap and rapid method to determine the nucleotide sequence in a DNA string described by Sanger and Coulson revolutionised the field of biological science in general and microbial ecology in particular [41]. Sanger sequencing allowed to determine bacterial composition without culturing by amplification of phylogenetic marker genes directly from an environmental DNA sample using PCR, cloning in *E. coli*, and sequencing of purified genes. The most important marker gene has, without doubt, been the 16S ribosomal RNA (rRNA)-encoding gene. This method tremendously improved our understanding of complex

microbial communities such as those found in soil and other environments [42]. However, the true bloom of ecological studies of microbial communities came together with the introduction of NGS technologies. The first NGS platform was 454 pyrosequencing that allowed to create libraries of DNA fragments that are attached to beads and amplified by water-in-oil emulsion PCR, creating beads covered in clonal DNA fragments. Using a picoliter reaction plate, fragments are then sequenced in parallel generating about a million of 400-500bp long reads [43]. A number of NGS platforms were developed after introduction of 454, however, the most prominent and widely used is Solexa that is currently known as Illumina. Solexa utilizes a hybridization between an adaptor at the end of a short DNA fragment and a complimentary oligonucleotide attached to the surface of the flow cell. The attached DNA molecule will be transformed into a clonal cluster of identical DNA fragments by bridge amplification and sequenced in a sequence-by-synthesis manner [44]. Currently the latest machine from Illumina is the NextSeq that allows to generate relatively cheaply an astonishing 400 million short reads (150bp) within a few days (https://www.illumina.com/systems/sequencing-platforms.html).

Such powerful sequencing technology provides an opportunity to capture the complete genetic makeup of a complex microbial community directly from an environmental sample, however, with great power came great challenges, and the biggest of them is how to deal with the vast amount of data that is generated. When it comes to analysis of a complex microbial community, amplicon or metagenomic sequencing are commonly used techniques. Regarding the frequently used amplicon sequencing, the process of data analysis can be divided roughly into two parts. The first part is dealing with raw sequencing reads and summarizes them into biologically/ecologically relevant information such as abundance tables of genes or microbial taxa. The second part then deals with data visualisation, statistical hypothesis testing and data mining. A number of pipelines are available for processing of data from 16S rRNA gene amplicon sequencing such as QIIME [45], MOTHUR [46] and NG-tax [47], the latter of which is an in-house developed pipeline that was used in the research described in this thesis. These pipelines allow in a more or less standardized way to get

from the raw reads to a taxa abundance table. The second part of microbiota analysis is less streamlined and requires a customized approach that will depend on research questions, study design and size of the dataset. To answer data questions in the different research chapters we used various methods of multivariate statistics, machine learning and advanced data visualisation using the R environment.

With the rapid decrease of sequencing costs, shotgun metagenomics became more popular. This technology allows to sequence all DNA present in a sample in contrast to amplicon sequencing where only specific target gene(s) are amplified by PCR and sequenced. The opportunity to sequence every strand of DNA in a metagenome allows to get inside into metabolic capacity of a microbial community, specific gene groups and the genetic makeup of uncultured microorganisms. The pipeline for analysis of shotgun metagenomics data will largely depend on the research question. In some cases tools that employ classification of short reads without prepossessing, such as DeepARG [48] and Kaiju [49], will allow to retrieve necessary information about abundance and/or taxonomy of specific genes of interest, such as, for example, antibiotic resistance genes.

Furthermore, researchers are often interested in the presence of specific pathways or individual genomes of community members. In this case, short reads should first be assembled into longer contigs, binned (if the goal would be to assemble individual draft genomes), and annotated using databases of interest [50]. Most popular tools for assembly of short reads into longer contigs currently include MEGAHIT [51], MetaVelvet [52] and metaSPAdes [53], to just name a few, and employ computationally efficient *De Bruijn Graph* approaches that allow to arrange reads into a continuous graph based on k-mer similarities within reads [54]. Assembled contigs can then be further annotated to reveal desired genomic features, which can be quantified and used in similar ways as an OTU table obtained after amplicon sequencing.

Factors influencing GI tract microbiota

Recognition of the roles that the GI tract microbiota plays in host homeostasis reinforces attempts to understand the factors that affect its composition and function. Factors affecting gut microbiota have a complex web of inter-correlations, however, one could roughly divide them into two main categories in relationship to a microbiota host - endogenous and exogenous factors. Factors related with genetics of the host can be considered endogenous, whereas the influence of food and the environment of the host are exogenous factors. There is a longstanding debate in the scientific community about the relative contribution of above-mentioned groups of factors. Despite our great progress in understanding of gut microbiota, stochastic relationships between influencing factors are not clear. It was shown that the human GI tract microbiota at a populational level is affected by a vast number of factors, with a number of surprising correlations showing that within a define human population faeces consistency (Bristol stool scale) plays a predominant role in shaping microbiota composition [55]. In addition, microbiota has been repeatedly shown to exhibit extensive variation between individuals within a group [56].

Most promising attempts to understand the influence of the genetic background on microbiota composition involve twin studies and laboratory animals with a homogeneous genetic background. The biggest twin study to date examined more than two thousand participants and showed that some taxa have a higher temporal stability in genetically close individuals and could be considered as heritable taxa. Furthermore, this study also showed an association between the lactase gene locus and *Bifidobacterium* [57]. An earlier study from the same authors showed that despite host genetics being a factor that strongly influences microbiota composition, certain taxa such as *Christensenella minuta* could affect the metabolic state even within pairs of monozygotic twins [58]. In contrast to twin studies the research of Rothschild et al. demonstrated a dominance of environmental factors in a shared household over ancestral genetic background in shaping the composition of gut microbiota [59]. When investigating the influence of host genetic background on microbiota composition,

laboratory and farm animals provide a superior model as compared to humans due to much higher genetic homogeneity within often well-defined genotypes. A number of studies focused on laboratory mouse strains and showed clear effects of genotype on microbiota composition, however, the genotype influence was fairly small when compared with diet induced microbiota variations [60, 61]. Relatively little attention was devoted to the influence of genotype on gut microbiota in farm animals, however, some studies evaluated the effect of genotype in cross-fostering models where piglets where raised by sows of a different genotype/race [62, 63].

Despite the established role of host genetics in shaping gut microbiota, environmental factors by far have the strongest influence on microbiota composition and functional profiles, with diet leading the chart. Diet of an individual directly modulates microbiota via provision of nutrients/substrate for microbial growth and creating an environment by altering the consistency of intestinal content. Diet has been shown to have a profound effect on the metabolic state of an intestinal microbial community, altering to a greater extent its metabolic functionality than its compositional make up [64]. It can be expected that such metabolic changes might also have an effect on host intestinal health. Although functional changes are more profound, a change in diet can also have an effect on overall microbiota composition within 24h, and drastic long term dietary changes were shown to inflict a switch of microbial enterotype [65]. Furthermore, it was shown that within four generations of laboratory mice a change in diet could not only alter relative abundance of core microbial taxa, but also lead to a complete extinction of certain bacterial lineages [66]. The influence of various components of human diet were extensively investigated in the context of human microbiota modulation, however, variability in dietary habits of an individual coupled with low control over environmental conditions makes it difficult to pinpoint specific diet-microbiota correlations [67]. In addition, large variability in response to a dietary intervention between individuals was observed, leading to the definition of responder- and non-responders groups of individuals within the same population [68]. In contrast, global patterns such as an increase in the production of SCFA in response to a high fibre diet, and a

compositional shift towards complex polysaccharide degrading species, were observed numerous times [69]. Even more apparent differences in overall composition were observed when studies go beyond humans and laboratory animals, comparing different host species with distinct diets. Ley et al. showed that despite apparent differences in intestinal microbiota of mammalian species, adaptations to a specific nutritional lifestyle, i.e. animals being carnivores, omnivores or herbivores, lead to similarities in microbial composition between phylogenetically distant hosts [7]. Despite a considerable effort to understand microbiota compositional trends in mammals, the response of microbiota to various diets within one or closely related mammalian species in natural habitats received relatively little attention.

Anthropogenic factors affecting microbiomes

In contrast with natural/pristine environment microbiota where microbial composition is primarily dictated by host ecological niches, lifestyle and available food, the anthropogenic environment adds a yet largely unknown dimension to "wild" microbiome challenges. These challenges are yet different from those affecting the human population, domesticated animals and wildlife.

Human lifestyle has a pronounced effect on neighbouring species – we are changing the surrounding environment rather than adapting to it. Anthropogenic influence on the environment can manifest in a range of different ways when considering wild animals and human impact, however, degradation of natural habitats could be considered as the most prominent. Replacement of native ecosystems with agricultural lands creates pressure on various species of plants and animals [70], and often leads to decrease of ecological diversity, which has become a pronounced problem around the globe [71]. Environmental changes at macroscale have clearly visible manifestation, however, smaller scale aspects of native species adaptations to new conditions are more subtle and not always receive the proper amount of attention. Only several studies looked into aspects of composition and function of microbiota of wild animals within

anthropogenically affected environments in comparison with pristine areas [70, 72]. A limited research interest for this subject is understandable, however, deciphering the changes that a species undergoes in the process of adaptation could give a better understanding of the general rules that govern evolution and development of microbiota.

Beside modification of the environment for our needs, we as humanity use and abuse technology to sustain our constantly growing population and personal quality of life. Medicine and in particular antibiotics, are main reasons why we are able to push the boundaries of life expectancy and produce a vast amount of agricultural products. As any ground-breaking technology, antibiotics besides their tremendous benefits have also hidden dangers. Antibiotics are known to have a prominent long term effect on gut microbiota composition in humans and animals [73, 74]. Several studies showed that infants treated with antibiotics have significant differences in microbiota composition in comparison with untreated groups [75]. It was furthermore shown that usage of antibiotics increases the concentration of antibiotic resistance genes within a host's gut community as well as in the environment.

Thesis aims and outline

The goal of this thesis is to explore factors affecting gut microbiota composition of wild and domestic animals, and to provide insight into the interplay between gut microbiota, host, environment and other members of the intestinal community (parasitic nematodes) in natural and controlled conditions. To this end we used NGS of 16S rRNA gene amplicons and shotgun metagenomics to gain insight into composition and functional attributes of the intestinal microbiota of studied subjects in combination with advanced statistics and data science to untangle complex relationships between environmental factors and microbiota.

Chapter II provides an insight into the natural variation of gut microbiota among congeneric species of lemurs dwelling across the island of Madagascar. This chapter provides convincing evidence for a predominant role of occupancy in shaping gut microbiota of wild lemurs, and shows that species identity of investigated lemurs has a markedly smaller contribution. We furthermore explore how occupancy of wild animals dictates their foraging behaviour and modulates gut microbiota.

Chapter III evaluates potential relationships between gut microbiota of wild lemurs and their intestinal parasites, and investigates seasonal variations of the intestinal community, and influence of anthropogenic factors. This chapter shows a long lasting effect of logging on the Madagascar forest ecosystem and as a consequence an impact on microbiota of lemurs dwelling within the area. Relationships between microbiota and parasites are more elusive.

Evidence for a significant effect of the genetic background of domestic pigs on gut microbiota composition is provided in **Chapter IV**. In particular, several microbial taxa that are known to have a prominent effect on intestinal health and overall wellbeing of mammals differ between breeds. Using machine learning approaches we distinguish between animals based on their microbiota composition. In addition, we show that gut microbial compositional differences are not equally extensive between different pig breeds, and microbiota of Landrace pigs has more dissimilarities from the microbiota of Duroc and Large Wight, than the latter two between each other. This chapter provides relevant information for the pig production sector, that could help to maintain intestinal health of pigs and facilitate breed choice decision making.

In **Chapter V** we investigate the resistome of free living Wight-Toothed Shrews and Wood Mice caught in the surroundings of several pig farms in the Netherlands using a combination of shotgun metagenomics and 16S rRNA gene amplicon sequencing. We show that the resistome composition has a direct correlation with the phylogenetic profile of gut microbiota and varies markedly between above-mentioned animal species. This finding suggests that the

potential of an animal to carry antibiotic resistant genes largely depends on its gut microbiota composition and the ecological niche.

Finally, **Chapter VI** of this thesis provides a synthesis and critical discussion of the results described in the experimental chapters, and closes with an outline of perspectives for future research.

Chapter 2

Occupancy strongly influences faecal microbial composition of wild lemurs

Alexander Umanets*, Iris de Winter*, Freek IJdema, Javier Ramiro-Garcia, Pim van Hooft, Ignas M.A. Heitkönig, Herbert H.T. Prins, Hauke Smidt

* Contributed equally

This chapter was published as Umanets A, de Winter I, et al. (2018). *FEMS Microbiol Ecol* 94 (3): fiy017

Abstract

The microbiota of the mammalian gut is a complex ecosystem, the composition of which is greatly influenced by host genetics and environmental factors. In this study, we aim to investigate the influence of occupancy (a geographical area of habitation), species, age and sex on intestinal microbiota composition of the three lemur species *Eulemur fulvus*, *E. rubriventer*, and *E. rufifrons*. Faecal samples were collected from a total of 138 wild lemurs across Madagascar, and microbial composition was determined using next generation sequencing of PCR-amplified 16S ribosomal RNA gene fragments. Consistent with reports from other primate species, the predominant phyla were *Firmicutes* (43±6.4% [s.d.]) and *Bacteroidetes* (30.3±5.3%). The microbial composition was strongly associated with occupancy in the *E. fulvus* population, with up to 19.9% of the total variation in microbial composition being explained by this factor. In turn, geographical differences observed in faecal microbiota of sympatric lemur species were less pronounced, as was the impact of the factors sex and age. Our findings showed that among the studied factors occupancy had the strongest influence on intestinal microbiota of congeneric lemur species. This suggests adaptation of microbiota to differences in forest composition, climate variations, and correspondingly available diet in different geographical locations of Madagascar.

Keywords: Microbiota, *Eulemur*, Madagascar, Gastro-intestinal tract, environment, multivariate statistics.

Introduction

The intestinal microbiota of mammals is an integral part of an animal's body that contributes significantly to the overall health of the host through modulation of its immune system, facilitation of food digestion, competition with pathogenic microorganisms and production of metabolites beneficial for the host [76, 77]. Expression of these beneficial properties directly correlates with microbial community diversity and composition [78]. Hence, identifying the factors and underlying processes that shape the intestinal microbiota is important for a better understanding of its contribution to host health. Previous studies in humans have shown that host genetics [79], lifestyle [80], and food preferences [81] contribute to shaping microbiota composition of an individual within a population. Intestinal microbiota composition can be distinguished between different mammalian species, suggesting co-evolution and adaptation of animals and their microbes [82, 83]. It is not clear, however, to what extent host genotype and environmental factors influence intestinal microbial composition under natural conditions among closely related animal species dwelling in different biogeographical regions.

Although wildlife microbiota has received less attention in comparison with that of humans, farm and rodent model animals, data collected from animals in wild conditions can provide complementary information that contributes to our understanding of processes that shape mammalian intestinal microbiota. For instance, studies that highlight similarities and differences in microbiota between humans and other *Homininae* species [84-86] provided new insight into evolution of microbiota, suggesting adaptation of human microbes to an animal protein-based diet. Studies on microbiota composition of primates that are evolutionarily more distant from humans, such as yellow baboons (*Papio cynocephalus*) [87], black howler monkey (*Alouatta pigra*) [88], black and white colobus (*Colubus guereza*), red colobus (*Piliocolobus tephrosceles*), and red-tailed guenon (*Cercopithecus ascanius*) [89] revealed that microbiota composition of these primates is highly variable, also intra-individually, and mostly depends on the available diet. Correspondingly, the diet of a wild animal directly depends on

suitable food availability, which depends on climate, flora and fauna of an area. This statement is also true for wild lemurs of Madagascar, for which several studies showed variation in feeding patterns and diet when comparing areas with different forest composition [90], as well as different seasons [91, 92]. Fogel (2015) compared the microbiota composition of sympatric wild *Lemur catta* and *Propithecus verreauxi* across dry and wet seasons and showed that microbiota of both lemur species is variable between individuals and dynamic over time. Researchers observed differences in microbial composition between wild *vs* captive *L. catta* as well as wild populations of *L. catta* and *P. verrauxi*, albeit only with respect to relative abundance of specific microbial groups rather than their presence or absence [93]. Wild rufous mouse lemurs (*Microcebus rufus*) showed an increase in gut microbial diversity with age and differences in microbiota richness and diversity between sampling.. Furthermore, microbial composition was affected by site, sex and year, whereas temporal trends within a year were weak [94].

The above mentioned studies of lemur microbiota were focused on a single lemur species [94], two sympatric lemur species dwelling in the same area [95], or captive lemurs of different species [93]. Taken together, these studies showed that lemurs harbour complex intestinal microbiota, the composition of which fluctuates over time among and within individuals, and is affected by season, captivity, age, site of sampling, and sex. In our study we focused on microbiota of three closely related *Eulemur* species (*E. fulvus, E. rufifrons* and *E. rubriventer).* To the best of our knowledge this is the first comparative study of intestinal microbiota composition of multiple wild *Eulemur* species across Madagascar exposed to large variations in climate conditions and biogeography.

In addition to an explorative assessment of the most important features of intestinal microbial composition in the studied species we addressed to what extent occupancy, host species, sex and age influence lemur intestinal microbial composition, and which of these factors contribute most strongly to intestinal microbiota differentiation in wild lemurs. To this end, we hypothesised that intestinal microbial composition is similar among congeneric lemur species and

that of the remaining factors occupancy, with habitat as a determining factor of food items availability and other climate related effects, has the strongest influence on intestinal microbiota differentiation.

Material and Methods

Study design

Faecal samples (N = 138) were selected from wild lemurs collected across Madagascar from April to July 2014 (Fig. 1). To investigate the effect of the occupancy on intestinal microbiota, *E. fulvus* samples from three climatic regions and *E. rufifrons* samples from two climatic regions were compared with each other. To assess the influence of different species, *E. rubriventer* samples collected in Ranomafana National Park (NP) were compared to *E. rufifrons* samples from the same area. The effect of age and sex was estimated based on *E. rufifrons* samples collected in Kirindy NP and Ranomafana NP (Table S1).

Study sites

Madagascar experiences a strong variation in climate conditions, resulting in different vegetation zones across the island [96]. The effect of environmental factors on lemur microbiota composition was investigated at five sites across Madagascar (Fig. 1). Kirindy Mitea NP (20˚07'S, 44˚67'E, 722 km^2) and Ankarafantsika NP (16˚25'S, 46˚80'E, 1350 km^2) consist of dry-deciduous forest and are located on the western- and north-western side of Madagascar, respectively [97]. Kirindy Mitea NP is characterised by pronounced seasonality. The area contains more than 200 species of trees with a mean canopy height of 12-18 meter, containing mostly deciduous trees with adaptations to water stress [98]. Ankarafantsika NP is a mosaic of floristically heterogeneous dry deciduous forests dissected by small valleys with abundant Raffia palms [99-102]

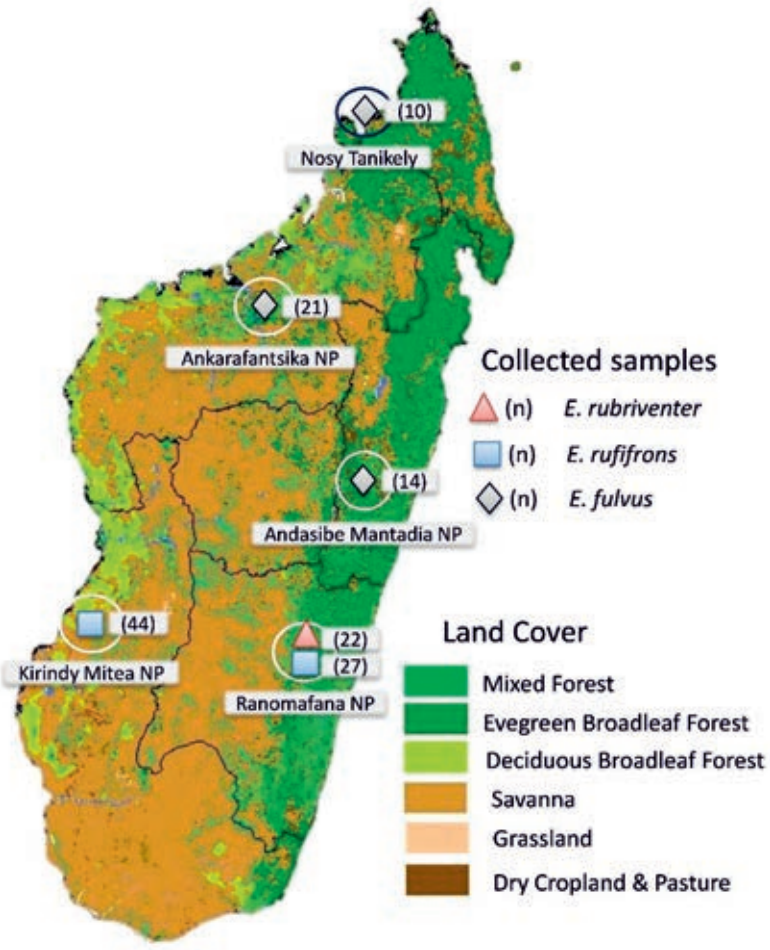

Figure 1. Lemur faecal sample collection areas across Madagascar. The map showing main types of land cover and vegetation was adapted from www.wildmadagascar.org, and was produced with data taken from the FAO Country Profiles and Mapping Information System (The United Nations Food and Agricultural Organization; © FAO 2004). Faecal samples were collected at five geographical locations across the island from *E. rufifrons* (two sites), *E. fulvus* (three sites) and *E. rubriventer* (one site). (n) = number of samples.

Ranomafana NP is located in southeastern Madagascar (21˚16'S, 47˚20'E) and encompasses approximately 435 km² of montane moist forest, ranging from altitudes of 500 m up to 1500 m, and receives an average of 3000 mm rainfall per year [103]. The rainfall in Ranomafana NP differs highly between the wet-warm season (December to March, 482-1170 mm per month) and dry-cold season (April

to November, 55-513 mm per month) [104]. Andasibe Mantadia NP (155 km^2) is located at the eastern side of Madagascar (18˚92'S, 48˚42'E), and is also characterized by relatively wet rain forests.

Nosy Tanikely (13˚28'S, 48˚14'E) is an island in the north-east of Madagascar covered with tropical vegetation. This island comprises less than 0.3 km^2 and is located between Nosy Be (8 km) and the mainland of Madagascars(13 km). Elevation ranges from 0 to 47 meters above sea level [105]. The island's vegetation consists of low forest with planted banana and mango trees, surrounded by a sandy shore with large rock formations (de Winter, personal observation)

Studied species

This study focused on three *Eulemur* species: the red-fronted lemur (*E. rufifrons*), the common brown lemur (*E. fulvus*) and the red-bellied lemur (*E. rubriventer*). These species are morphologically alike and are frugivorous, although they may include other food sources, such as leaves and invertebrates, in their diet [106-109]. The main difference in social organization between the different *Eulemur* species is their group size. *E. rufifrons* and *E. fulvus* live in multi-male, multi-female groups from 4 - 18 individuals [107, 109, 110], whereas *E. rubriventer* lives in small monogamous groups from two up to five individuals[111, 112].

Sampling and data collection

Immediately after defecation, fresh faecal samples (3-4 g) from individual lemurs were collected non-invasively. Within 12 h after collection, the samples were stored at ambient temperature in sterile plastic tubes that were prefilled with 5 ml of 70% ethanol until further analyses at the Laboratory of Microbiology, Wageningen University, The Netherlands. Species, age and sex

were recorded. All samples included in this paper were taken in compliance with the laws of the Government of Madagascar and no animal experimentation was involved.

DNA extraction

Samples collected in the Ranomafana NP were processed using a modified protocol based on method proposed by Yu [113] with modifications described previously [114]. For this method faecal material was air-dried for 15-20 min in a fume hood to remove ethanol from the samples. Subsequently, 0.1 - 0.17 g of dried samples were added into double autoclaved screw-cap tubes containing 0.3 g of 0.1 mm zirconia beads, three pieces of 2.5 mm glass beads and 700 µl of lysis buffer (500 nM NaCl, 50 mM Tris-HCI (pH=8), 50 mM EDTA, 4% SDS) in each. Samples were treated for 3x1 min at $5.5x10^3$ movements per minute in a Precellys 24 beadbeater (Bertin technologies, France). After homogenization, samples were incubated at 95 ˚C for 15 min in a shaking heating block (Vartemp 56, Labnet International, Edison, NJ, USA) at 100 rpm, then centrifuged at 4 ˚C for 5 min at 13,000 rpm. Clean supernatants were transferred into 2 ml tubes. 300 µl of fresh lysis buffer was added in the same tubes to the pellets, bead beating/incubation steps were repeated, and freshly collected supernatant was pooled with that previously collected. Subsequent steps were performed according to the original protocol [113].

Samples collected in Adnasibe NP, Kirindy Mitea NP, Ankarafantsika NP and Nosy Tanikely were extracted using an automatic system, the Maxwell® 16 Research Instrument (Promega, Madison, USA), and the corresponding RNA extraction kit according to manufacturer's instructions. To improve DNA yield, samples preserved in 70% ethanol were rehydrated through a series of ethanol solutions with decreasing proportions of ethanol in steps of 10%. For rehydration 1.5 ml of 70% ethanol with faecal particles was transferred into a fresh 2ml tube and centrifuged at 13,000 rpm for 5 min to separate solid fractions from the liquid. After centrifugation part of the supernatant was replaced with the same

amount of distilled water to decrease ethanol concentration by 10 percent point, vortexed and incubated for 10 min at RT. These steps were repeated until the ethanol was replaced by distilled water. Cell disruption and lysis was performed as described above, but instead of lysis buffer we used S.T.A.R buffer (Roche Molecular Systems, USA).

DNA quality and concentration was determined spectrophotometrically (Nanodrop Technologies, Wilmington, USA). Comparison of the two DNA extraction methods mentioned above, using human faecal samples, indicated that both methods delivered DNA of essentially equal quality, resulting in comparable results with respect to microbial composition based on analyses with the Human Intestinal Tract Chip (HITChip), a DNA oligonucleotide microarray targeting human intestinal microbiota (Heikamp-de Jong & Hartman, personal communication).

Amplification of 16S rRNA gene fragments and library preparation

After DNA extraction, regions V1-V2 of the 16S rRNA genes were amplified using an in house two-step PCR protocol. In the first step regions of interest were amplified using the following primers: 27F–DegS: GTTYGATYMTGGCTCAG [115] and an equimolar mix of 338R–I: GCWGCCTCCCGTAGGAGT [116] and 338R–II: GCWGCCACCCGTAGGTGT[117], with attached UniTag I (forward) and II (reverse) linkers (I – GAGCCGTAGCCAGTCTGC; II - GCCGTGACCGTGACATCG)[118]. The PCR mix for one reaction at step one contained 10 µl of 5x HF buffer, 1µl dNTPs (10 µM), 1U of Phusion Hot start II DNA polymerase (2U/ µl), 31.5 µl of nuclease free water, 2.5 µl of forward (10 µM) and 2.5 µl of reverse primers (10 µM), and 40 ng of DNA template. Amplification was performed in a LabCycler Gradient (SensoQuest, Germany) programmed for initial denaturation at 98 ^{o}C for 30 s and 25 cycles of denaturation at 98 ^{o}C for 10 s, annealing at 56 ^{o}C for 20 s and extension at 72 ^{o}C for 20 s, followed by final extension at 72 ^{o}C for 10 min. After amplification, the success of the PCR reaction

was checked visually by agarose gel electrophoresis, considering amount and size of the amplicon as quality parameters.

Amplicons were subsequently used as template for a second PCR for the introduction of sample-specific barcodes, using individual barcode primers for each sample. In total we used 48 pairs of forward and reverse barcode primers that target UniTag1 and UniTag2 sequences introduced during the first PCR, respectively, and that were appended with sample specific barcodes. Composition of PCR reagents and cycling conditions were as described for the first PCR, with 10 µl of PCR products from the first step as template. Reactions were performed in a final volume of 100 µl. PCR products were purified and concentrated using magnetic beads (MagBio, Switzerland) according to the HighPrep protocol with adaptation for 2 ml tubes. Purified products were quantified using the Qubit dsDNA BR Assay Kit (Life Technologies, USA) following the manufacturer's protocol. PCR products were pooled in equimolar amounts into libraries of 48 samples each, and sequenced on an Illumina MiSeq platform in 300bp paired end mode at GATC Biotech (Constance, Germany).

Data processing and statistical analysis

Initial analysis of raw 16S rRNA gene sequencing data was performed using NG-Tax pipeline [47]. Sequences were separated into sample-specific bins based on the barcodes, after initial filtering of paired-end libraries to contain only read pairs with perfectly matching barcodes. OTUs were defined using an open reference approach, and taxonomy was assigned using a SILVA 16S rRNA gene reference database [119]. Microbial composition plots were generated using a workflow based on Quantitative Insights Into Microbial Ecology (QIIME) v1.9.1 [45].

Reads assigned to OTUs of plant origin such as chloroplast and plant mitochondrial DNA were removed from the dataset used for downstream analyses. OTU counts were normalized using cumulative sum scaling (CSS)

[120]. To get an overview of species composition, a normalized OTU matrix was exported to Microsoft Excel, and the relative contribution based on normalized OTU numbers per taxa was calculated. Median values of taxa relative abundance in a group of samples were used to compare groups with each other (e.g. male vs female). The OTU matrix was filtered to exclude OTUs that were present only in a small number of samples. More specifically, for each dataset, OTUs were removed that were present in less than five samples (50% of the smallest group size).

Measures of alpha and beta diversity and initial multivariate analysis using principal coordinate analysis (PCoA) were performed on the rarefied matrix (depth – 1650 observations) using weighted and unweighted UniFrac as distance measures as implemented in QIIME. Significance of differences in relative abundances of OTUs between individual samples was determined using Kruskal-Wallis tests when comparing more than two groups, and nonparametric t-tests with 500 Monte Carlo permutations in case of comparisons of two groups, using normalized, summarized and filtered OTU tables. False discovery rate (FDR) correction of p-values was used to reduce the chance of type I statistical errors, when multiple statistical hypotheses were tested. To identify strength and statistical significance of sample groupings with weighted and unweighted UniFrac as distance measures we used the “adonis” test as implemented in the R package “vegan”. Canoco 5.0 was used for multivariate statistical analysis and visualization of correlations between microbial composition of samples and explanatory factors. Redundancy analysis (RDA) was performed as described previously [121]. As input dataset for RDA we used the taxonomy summary table at genus level after removal of taxa which were found in less than eight samples (applied for each dataset individually) with addition of the corresponding sample metadata. No transformation or normalization of the data was done. The significance of observed community variations was evaluated using a Monte Carlo Permutation test.

To identify microbial species most strongly correlated with investigated factors such as area or host species we used the LefSe (Linear discriminant

analysis of Effect Size) algorithm for biomarker identification [122]. Data were processed using tools developed by the Huttenhower laboratory implemented in the Galaxy environment (www.huttenhower.sph.harvard.edu/galaxy/). Preparation of input data and analysis were performed according to the standard workflow, using default settings (0.05 - alpha value for the factorial Kruskal-Wallis test among classes; threshold on the logarithmic LDA score for discriminative features was 2.0; the strategy for multi-class analysis was "All-against-all").

After raw data processing and initial analysis samples were organized into five sets (samples are used in several sets), allowing us to perform separate analyses and statistics while focusing on a particular research question: "All_samples" - initial set of all 138 samples; "E_fulvus" - samples obtained from *E. fulvus* (n = 45) from three different areas (Andasibe n = 14; Ankarafantsika n = 21; Nosy Tanikely n = 10); "Kirindy" - set of samples from *E. rufifrons* collected in Kirindy NP (n = 44); "Ranomafana" - samples from *E. rufifrons* (n = 27) and *E. rubriventer* (n = 22) collected in Ranomafana NP (total n = 49); "E_rufifrons" - samples from *E. rufifrons* collected from two different areas (Ranomafana NP n = 27 and Kirindy NP n = 44).

Availability of data and materials

Datasets generated in this study are available in the public read archive EBI, study name 'Area of habitation strongly influences faecal microbial composition of wild lemurs', with accession number PRJEB20007.

Results

In this study, we analysed the faecal microbiota of a total of 138 individuals belonging to three different *Eulemur* species, using Illumina MiSeq sequencing of PCR-amplified 16S ribosomal RNA gene fragments covering the V1-V2 variable region. In total we obtained 6,220,515 reads, ranging from 1652 to 178,522 reads per sample (r/s) with a median of 36,092 r/s. Obtained reads were assigned to 1053 Operational Taxonomic Units (OTUs) using NG-Tax, an in house developed pipeline [47]. Across all samples, OTUs belonged to 12 bacterial phyla, i.e., *Acidobacteria*, *Actinobacteria, Bacteroidetes, "Candidatus* Saccharibacteria*"* (TM7)*, Cyanobacteria, Firmicutes, Lentisphaerae, Proteobacteria, Spirochaetes, Synergistetes, Tenericutes* and *Verrucomicrobia*. The fraction of non-assigned to any taxonomic level OTUs varied from 2.6 – 16.7% with an average of 9.3±2.5% [s.d.] in all analyzed samples. Predominant phyla, regardless of lemur species and sampling location, were *Firmicutes* 43.3±6.4%, *Bacteroidetes* 30.3±5.3%, *Cyanobacteria* 5.2±3.3%, and *Proteobacteria* 7.4±3.1%. At the genus level a total of 59 taxa were identified, 15 of which had an average relative abundance across all samples of more than 1% and comprised more than 80% of all sequences. Overall, 34% of all OTUs could be assigned at genus level, and 63% at family level. Phylogenetic clustering based on relative abundance at the genus level showed that the most abundant genera could be clustered into two groups: one group consisting of the two most abundant genera (unidentified genus (UG_1 (*Clostridiales*) 24.9%±5.4%, UG_2 (*Bacteroidales*) 14.9±3.6%), and another group consisting of another five genera (UG_3 (*Prevotellaceae*) 7.7±2.3%, UG_4 (C*yanobacteria*) 5.2±3.3%, UG_5 (*Bacteroidaceae*) 4.6%±2.4%, UG_6 (*Lachnospiraceae*) 4.9%±2.8%, UG_7 (*Ruminococcaceae*) 4.2%±2%) (Fig. 2; Table S2).

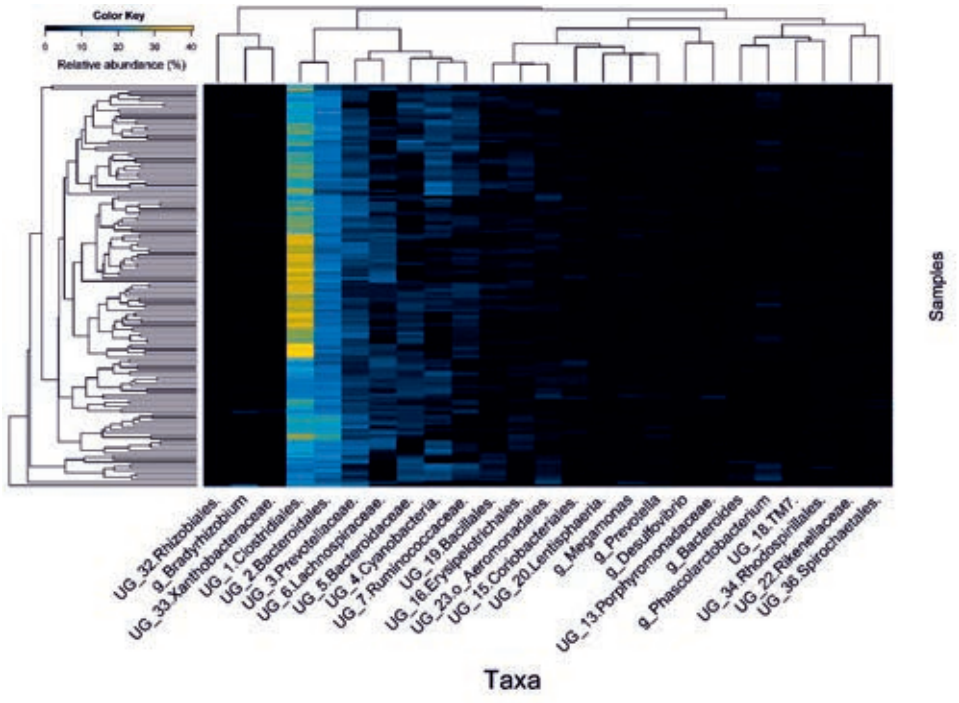

Figure 2. Heatmap of relative microbial abundance at genus level in the whole dataset, with samples placed on Y-axes, and genera with relative abundance more than 2.5% across all samples on X-axes. The red colour indicates high relative abundance values; dark green indicates low relative abundance values. Samples clustering and dendogram were produced using the "Bray" method as it implemented in the "vegan" R package.

Effect of lemur species on faecal microbiota composition

In order to address the influence of the host species on observed variation in microbial composition, two different datasets (see information regarding the composition of datasets in Experimental Procedures) were analysed: "All_samples", i.e., the entire dataset of 138 samples, and "Ranomafana", the latter of which allowed us to minimize the influence of explanatory variables other than host species. Analysis of "All_samples" revealed that samples from *E. fulvus* (n = 45) had significantly lower alpha diversity (p = 0.003) in comparison with samples taken from *E. rufifrons* (n = 22) (Fig. S1). The relative abundance of *Proteobacteria, Lentisphaerae, Synergistetes, Bacteroidetes, Actinobacteria, Firmicutes, Tenericutes,* and Candidate division TM7differed among studied

lemur species (FDR-corrected p<0.012, Fig. 3A). At genus level, 26 taxa differed in relative abundance (corrected p<0.036), with some present only within one lemur species. The genera *Anaeroplasma* and UG_8 (*Desulfovibrionaceae*) were found only in samples belonging to *E. rufifrons*, albeit with relative abundances below 1%. Eight genera were identified by LefSe analysis as potential biomarkers of the different lemur species: *Bacteroides* and *Phascolarctobacterium* were identified as microbial biomarkers of *E. rufifrons*; UG_6 (*Lachnospiraceae*), *Campylobacter,* UG_9 (*Synergistales*), UG_10 (*Clostridiales*) and UG_11 (*Xanthomonadales*) were biomarkers for *E. rubriventer*; and UG_12 (*Anaeroplasmatales*) was associated with *E. fulvus* (Fig. 3B). No clear grouping of samples by lemur species was observed based either on weighted or unweighted UniFrac distance. This was confirmed by the "adonis" test that revealed only a weak linear correlation between samples, with an R^2 of 0.11 and 0.13 for weighted and unweighted distances, respectively. RDA with lemur species as the only explanatory variable showed that this variable significantly (p = 0.008) contributed to the observed variation in faecal microbiota composition (Fig. 4A). Furthermore, when comparing the faecal microbiota of *E. rufifrons* and *E. rubriventer* in Ranomafana NP, RDA analysis showed that although 'lemur species' was a significant explanatory variable (p = 0.024), it only explained 6.8% of the observed variation in microbial community composition (Fig. 4B). All phyla observed in the full dataset were also present in Ranomafana NP, and only the phylum *Firmicutes* showed nearly significant differences in relative abundance between *E. rufifrons* (44.3±7.1%) and *E. rubriventer* (39.7±4.7%) (p = 0.008; corrected p = 0.1). At genus level UG_13 (*Porphyromonadaceae*), UG_5 (*Bacteroidaceae*), and UG_19 (*Bacillales*) differed in relative abundance when comparing microbial composition in both lemur species (corrected p<0.04). Similar to dataset "All_samples", no separation or grouping was observed among samples from the "Ranomafana" dataset in weighted or unweighted UniFrac matrix-based PCoA plots (R^2=0.06 and R^2=0.05 for unweighted and weighted distance matrices, respectively; data not shown).

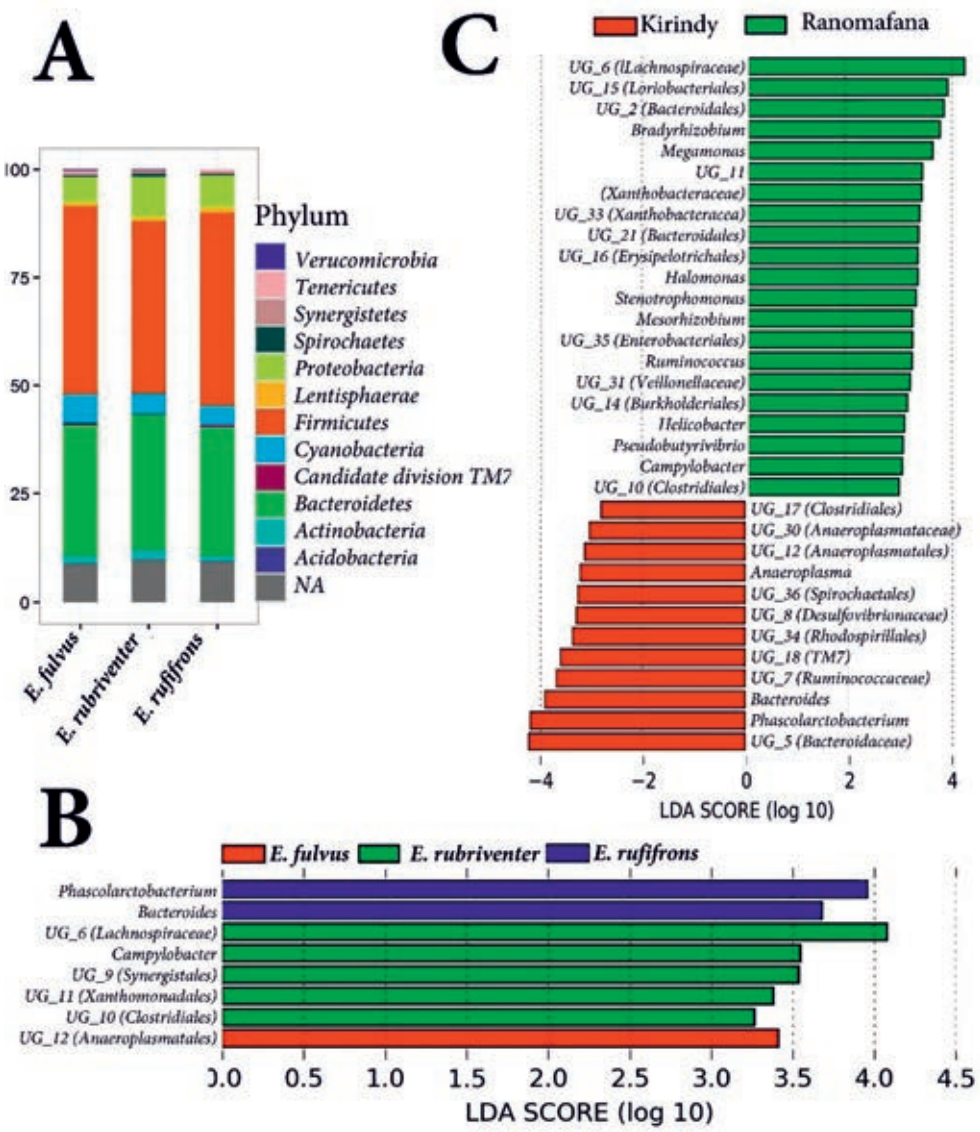

Figure 3. Differences in bacterial composition between lemur species (*E. fulvus, E. rubriventer, E. rufifrons*) (**A**) relative abundance at the phylum level of faecal microbiota of the different lemur species using dataset "All_samples"; (**B**) taxa identified by LefSe as potential biomarkers for the discrimination of studied *Eulemur* species. LDA - linear discriminant analysis; (**C**) taxa identified by LefSe as potential biomarkers for the discrimination of fecal samples taken in Ranomafana NP and Kirindy NP.

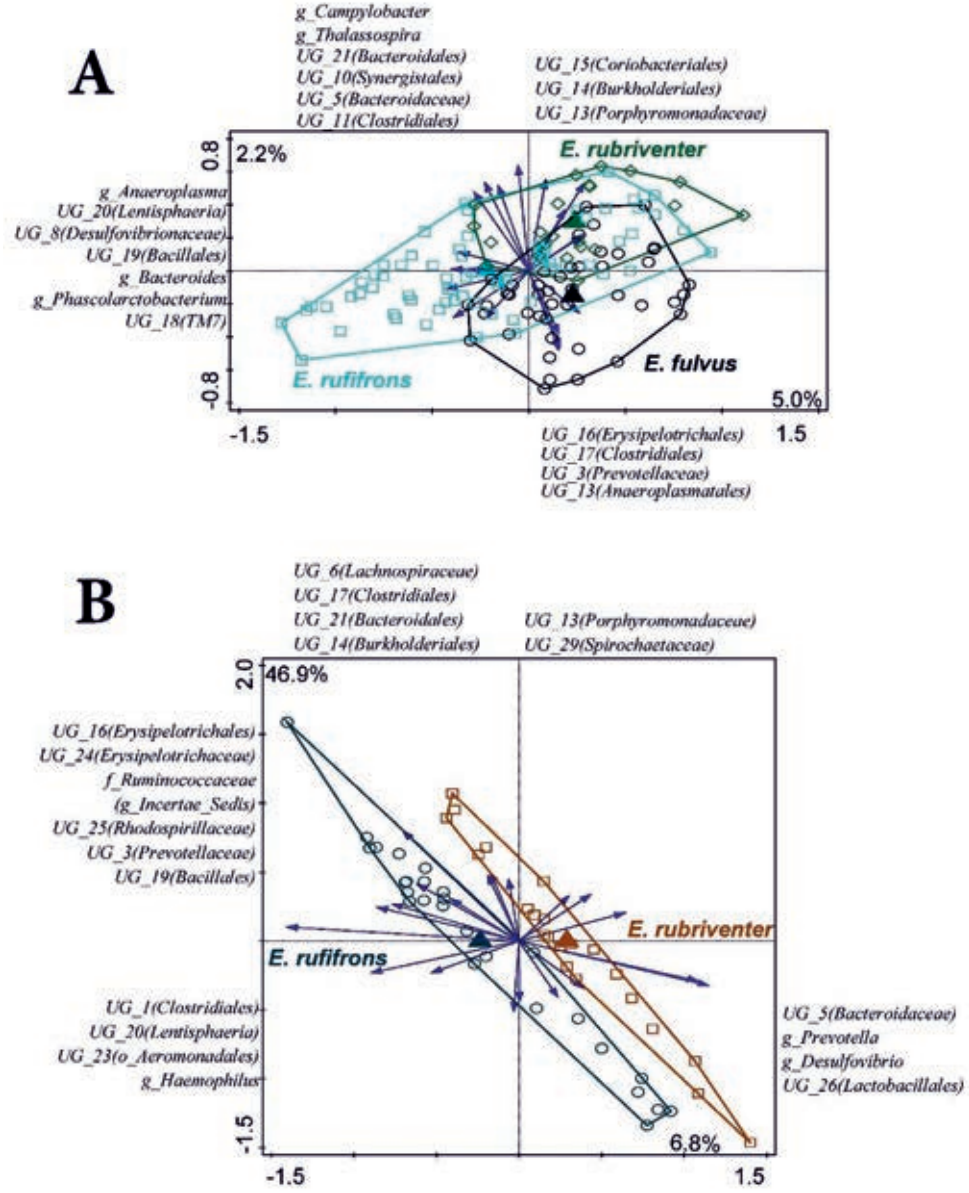

Figure 4. Ordination triplots based on RDA with lemur species as explanatory variables. (**A**) In dataset "All_samples" 7.2% of the variation is captured by the first two canonical axes; (**B**) in dataset "Ranomafana" 6.8% of variation is captured by the first canonical (constrain) axis, and both host species significantly ($p = 0.002$) contributed to explaining the observed variation in microbiota composition.

Variation of the lemur microbiota in contrasting regions of Madagascar

To determine to what extent occupancy (i.e., area of habitation) can explain the observed variation in faecal microbiota we analysed three datasets: "All_samples", "E_fulvus", and "E_rufifrons". Dataset "All_samples" allowed us to identify the influence of "area" among all other variables such as host species, sex and age. With dataset "E_fulvus" we focused on a single lemur species, *E. fulvus*, which was sampled at three different locations, allowing us to more specifically address variation in faecal microbiota composition found within one species exposed to different environmental conditions. We constructed dataset "E_rufifrons" for the same purpose as dataset "E_fulvus", taking into account that individuals of *E. rufifrons* were sampled in two distinct areas. Dataset "All_samples"(n = 138) showed that samples from Nosy Tanikely (n = 10) area had lower alpha diversity as compared to all other areas ($p = 0.01$, Fig. S2). The relative abundances of 10 out of 12 phyla, except for *Bacteroidetes* and *Acidobacteria*, were different (corrected $p<0.026$) between sampling sites. It should be noted that members of the phylum *Acidobacteria* were found in only a few samples (10 out of 138). Furthermore, at genus level 54 out of 59 taxa that were observed in more than five samples showed significant differences in relative abundance between areas (corrected $p<0.05$). Among these genera, some were found only within Kirindy NP, namely *Anaeroplasma, Rhizobium,* and UG_8 (*Desulfovibrionaceae*). Members of the genus *Bacteroides*, UG_13 (*Anaeroplasmatales*), UG_17 (*Clostridiales*) order and UG_30 (*Anaeroplasmataceae)* were found exclusively in samples from the relatively dry areas (Kirindy and Ankarafantsika). The most abundant genus (UG_1 (*Clostridiales*)) did not vary significant between areas. Samples showed a slight visual grouping according to area in PCoA plots based on weighted UniFrac distances ($R^2 = 0.29$), with better group separation being observed in the case of unweighted UniFrac ($R^2 = 0.34$) (Fig. 5A, B). Furthermore, constrained analysis (RDA) showed that all areas included as explanatory variable significantly ($p<0.05$) contributed to the observed variation in faecal microbiota composition (Fig. 6A).

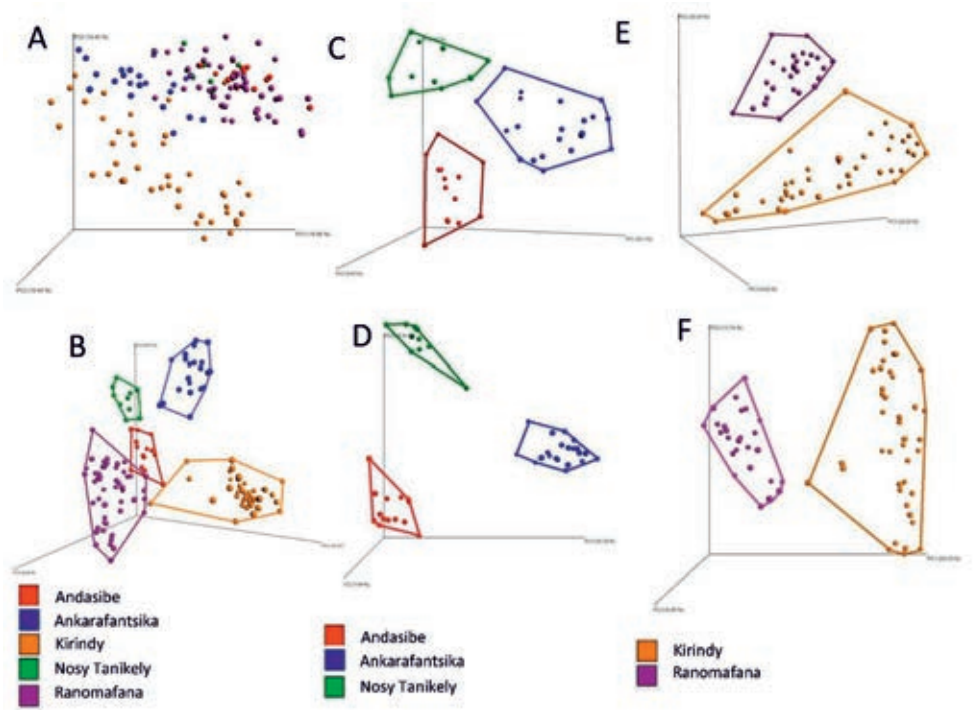

Figure 5. Principal coordinate analysis (PCoA) three dimensional (first tree PCoA axes) plots based on weighted (A, C, E) and unweighted (B, D, F) UniFrac distance matrices. Samples are represented by dots, colour-coded by sampling location. Plots contain all samples (A, B), or are species specific (*E. fulvus* - C, D;*E. rufifrons* - E, F).

In dataset "E_fulvus" eight out of 12 phyla (*Tenericutes, Cyanobacteria, Spirochaetes, Lentisphaerae, Firmicutes, Candidate division* TM7*, Proteobacteria, Actinobacteria*) differed in relative abundance between areas (corrected $p<0.01$). In line with the extensive differences observed at the phylum level, 34 out of 55 detected genera showed significant differences in relative abundance between areas. UG_17 (*Clostridiales*), UG_35 (*Enterobacteriales*), UG_29 (*Spirochaetaceae*), UG_12 (*Anaeroplasmatales*) and the genus *Bacteroides* were found only in the samples from Ankarafantsika NP. UG_31 (*Veillonellaceae*) was found exclusively in the samples from Andasibe NP, and *Mesorhizobium* only in Nosy Tanikely. Several genera were absent in one out of three areas: UG_19 (*Bacillales*), *Helicobacter* and *Thalassospira* were not detected in samples taken in Nosy Tanikely, whereas *Pseudobutyrivibrio* was not found in Andasibe.

Samples in the "E_fulvus" dataset clustered into three groups, correlating with the three different sampling sites based on the relative abundance of bacterial genera (Fig. 7). Furthermore, samples formed separated groups in PCoA plots based on weighted and unweighted UniFrac distances (R^2 = 0.35 and R^2 = 0.41, respectively; Fig.5C, D). RDA showed that among all factors only area significantly (p = 0.002) contributed to explaining observed differences in faecal microbial composition of *E. fulvus*, with Ankarafantsika having the highest explanatory value (19.9%) (Fig.6B).

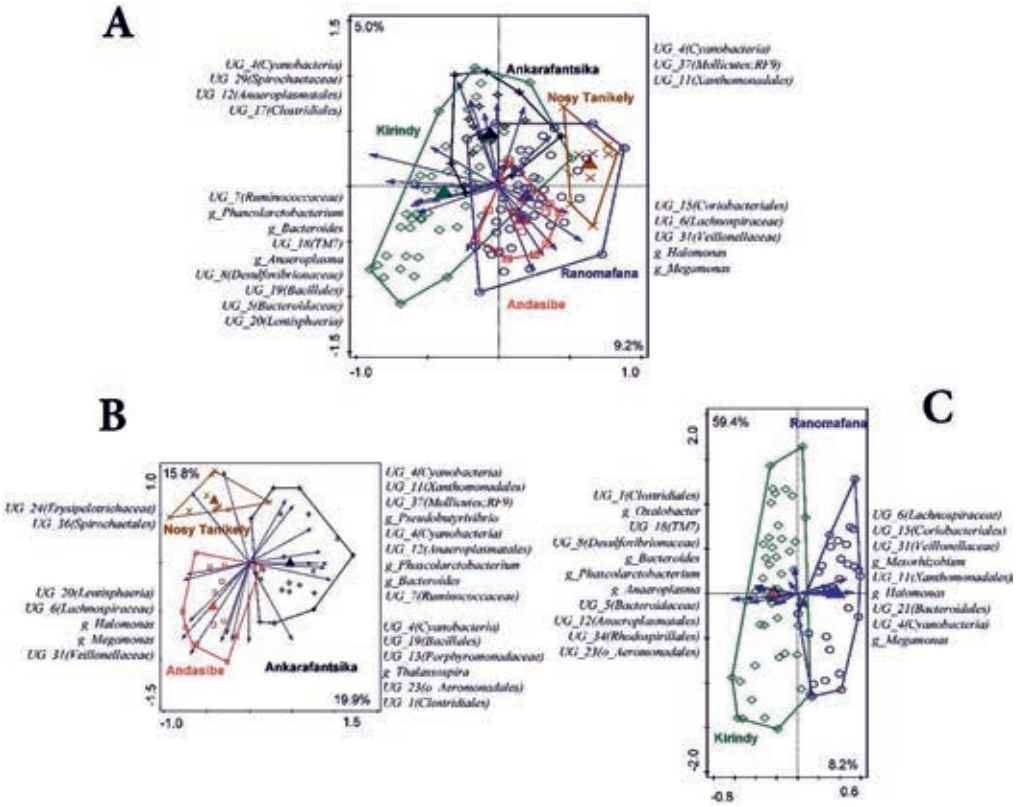

Figure 6. Ordination triplot based on RDA with areas of sampling as explanatory variables. (**A**) In dataset "All_samples" 14.2 % of the variation was captured by the first two canonical axes, and a statistically significant (p < 0.05) effect of areas as explanatory factors was observed; (**B**) in dataset "E_fulvus" 35.7% of variation was captured by the first two canonical axes, and all three areas significantly contributed to explaining the observed variation in microbiota composition (p =0.002); (**C**) in dataset "E_rufifrons" 8.2% of variation was captured by the first canonical (constrained) axis. Explanatory variable "areas" was statistically significant (p = 0.004) as a conditional effect.

In dataset "E. rufifrons", the relative abundance of the phyla *Actinobacteria, Candidate division* TM7*,* and *Proteobacteria* was different between the two locations where this lemur species was found, i.e., Kirindy NP and Ranomafana NP (corrected $p<0.009$). In total, 26 genera were different in relative abundance when comparing both locations (corrected $p<0.047$). Members of the genus *Bacteroides* were completely absent from the samples collected from Ranomafana NP, while their mean relative abundance was 1.4±1% among samples collected from Kirindy NP. All remaining genera that were found exclusively in samples taken in Kirindy NP (UG_8 (*Desulfovibrionaceae*), UG_12 (*Anaeroplasmatales*), *Anaeroplasma*) had relative abundances below 0.3%. The genus *Phascolarctobacterium* was found in all samples from Kirindy NP (mean abundance 3±1.3%), but only in 12 out of 27 samples (average abundance 0.2±0.2%) from Ranomafana NP. UG_21 (*Bacteroidales*, 0.2±0.4%), *Mesorhizobium* (0.1±0.2%), UG_33 (*Xanthobacteraceae*, 0.4±0.2%), *Stenotrophomonas* (0.1±0.2%), UG_32 (*Rhizobiales*, 0.3±0.9%), and UG_10 (*Clostridiales*, 0.1±0.1%) were found exclusively in Ramonofana NP, albeit not in all samples collected in that area, and at low relative abundance. Additional differences included UG_5 (*Bacteroidaceae*) (6.2±2.2% in Kirindy NP vs. 3.3±2% in Ranomafana NP) and UG_6 (*Lachnospiraceae*) (6.2±2.3% in Ranomafana NP vs. 2.9±1.9% in Kirindy NP). Twenty one genera for Ranomofana NP and 12 for Kirindy NP were identified by LefSe as microbial "biomarkers" (Fig. 3C). Multivariate analyses supported the separation of samples according to sampling location, with a clear grouping being observed in PCoA plots based on both weighted ($R^2=0.19$) and unweighted ($R^2=0.23$) UniFrac distance matrices (Fig. 5E, F). Furthermore, RDA showed that from all explanatory variables (area, age, sex) only variables in the group "area" significantly contributed to explaining the observed variation in microbial community composition ($p<0.004$), with both areas (Kirindy and Ranomafana) explaining 8.2% of variation (Fig. 6C).

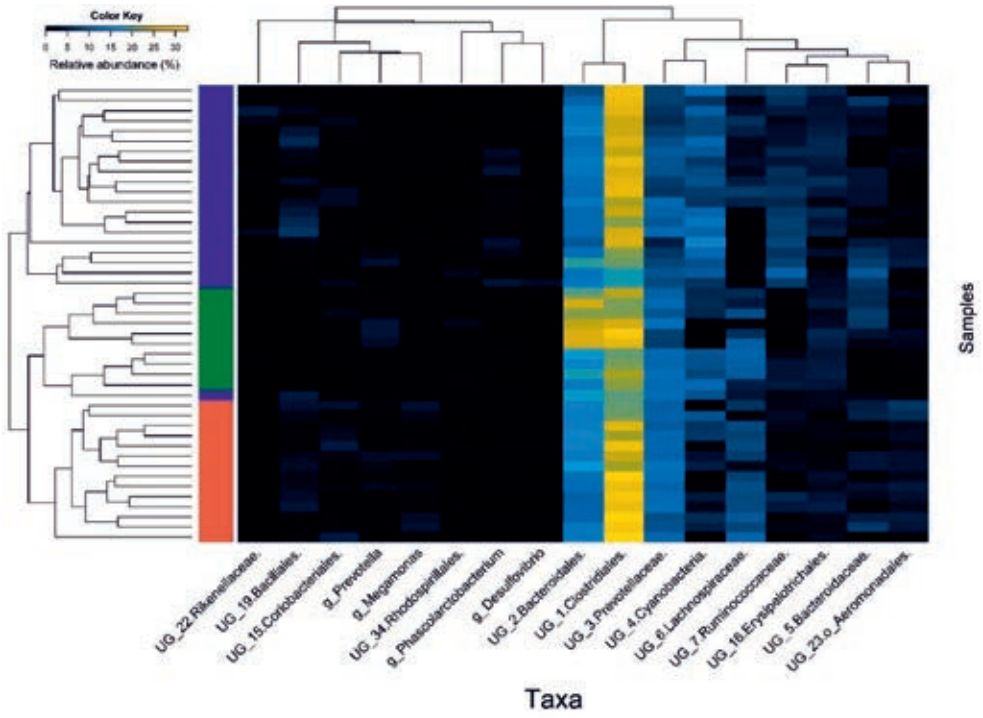

Figure 7. Heatmap of relative microbial abundance at genus level in the dataset "E_fulvus", with samples placed on Y-axes, and genera with relative abundance more than 2.5% across the dataset on X-axes. The yellow colour indicates high relative abundance values; dark blue indicates low relative abundance values. Samples clustering and dendogram were produced using the "Bray" method as it implemented in the "vegan" R package. Side bar (left) indicates area of sample collection: green – Nosy Tanikely, brown – Andasibe NP, blue – Ankarafantsika.

Variation in microbiota composition associated with sex and age of animals

Influence of sex and age on faecal microbiota composition was investigated using datasets "All_samples" and "Kirindy". No significant differences were observed in microbiota between samples collected from males and females in any of the datasets at phylum or genus level, and no grouping was observed with multivariate analysis using PCoA and RDA (data not shown). Similarly, no significant variation in relative abundance of detected phyla was observed between age groups, and only the relative abundance of the genus

Phascolarctobacterium (corrected p = 0.01) differed in age groups in dataset "All_samples". Multivariate analysis (RDA and PCoA) confirmed that age did not significantly contribute to explaining the variation in faecal microbiota composition.

Discussion

In the current study, we characterized the faecal microbiota of three frugivorous *Eulemur* species, and assessed to what extent the naturally occurring variation in intestinal microbiota composition is associated with occupancy, species, age and sex of individuals. Findings presented here showed that the gut microbial community of these animals is dominated by members of the phylum *Firmicutes* and to a lesser extent *Bacteroidetes*. It has previously been reported that predominance of *Firmicutes* or *Bacteroidetes* is different among animal species and mostly correlated with dietary mode and taxonomic lineage of a given species [82]. Our results confirmed the high proportion of *Firmicutes* that was previously observed in other species of frugivorous and omnivorous primates [123, 124], including humans, despite the fact that phylogenetically lemurs are one of the most distinct and ancient groups within the primates [125]. Notably, human studies showed that the *Firmicutes* to *Bacteroidetes* ratio is not static and can be largely influenced by the presence of carbohydrates in the diet, although it is not clear which of the two phyla has a leading role as key degrader of complex carbohydrates in the human intestine. For example, on one hand an increase in relative abundance of *Firmicutes* was correlated with consumption of whole grains and total carbohydrate intake [126], and several species belonging to this phylum are viewed as key degraders of resistant starch [127]. On the other hand it has been shown that the depletion of *Firmicutes* and increase in *Bacteroidetes* in African children from a rural area in comparison with European children was related to consumption of a traditional African diet rich in fibres and polysaccharides [128]. Such seemingly conflicting evidence might be related to the high phylogenetic and functional diversity within both phyla, including a

large number of fibre- and carbohydrate-degrading species. Consequently, one can speculate that specific aspects of the diet of lemurs will result in a shift of the *Firmicutes/Bacteroidetes* ratio, the direction of which might not be predictable based on general characteristics of the diet. Our study showed that a large fraction of *Firmicutes* associated sequences was assigned to a single genus-level taxon, UG_1 (*Clostridiales*), accounting for 24.9±5.7% of the total bacterial community. It is tempting to speculate that members of this genus have an important role in intestinal function of the three *Eulemur* species covered in our study, however, due to lack of physiological and ecological data, conclusions regarding their function and place in gut ecology remain speculative, awaiting isolation and/or (meta)omics analyses [129].

Notably, members of the *Proteobacteria* had relatively high abundance (7.4±3.1%) in all three studied *Eulemur* species. In humans, high relative abundance of this phylum (9.7% -14.9%) has been associated with gastric bypass, metabolic disorders, inflammation and cancer, whereas its relative abundance in healthy individuals amounts to only about 4.5% [130]. On the other hand, previous research showed host species related differences in abundance of *Proteobacteria* among primates. For instance, Bello González et al. (2015) observed that faecal samples of humans and chimpanzees had similar relative abundances of *Proteobacteria* (1% and 1.2%, respectively), whereas in gorilla samples this phylum reached a relative abundance of 7% [131]. Furthermore, a similar relative abundance of *Proteobacteria* (9.1%) was reported in faecal microbiota of *Lemur catta* [93]. Hence, we suggest that the high relative abundance observed in this study is not necessarily a sign of a health problem of the investigated population of lemurs, but rather a feature of the normal microbial composition of frugivorous lemurs.

We found that on average 5.2±3.3% of all reads were assigned to OTUs belonging to the *Cyanobacteria* phylum. Latest research shows that members of this phylum are indeed a genuine part of the human intestinal microbiota [132]. Furthermore, the presence of this phylum was observed in previous studies which characterized the intestinal microbial composition of other primates,

including *Lemur catta* [93]. Furthermore, we found that 66% of genus-level taxa could not be confidently classified to a particular genus in the Silva v111 database, including several of the most predominant taxa. This is in line with the limited attention that the intestinal microbiota of lemurs has received to date, and hence there is a lack of knowledge regarding specific taxa present in the intestine of these animals. Research on intestinal microbiota of other poorly studied animals showed similar findings. Roggenbuck et al. (2014) found that only 28% of the observed genera in the giraffe rumen could be assigned to known taxa [133]. Similar observations have even been made for less well characterized human populations. In a recent study, Schnorr et al. (2014) found that 22% of the total microbial community of central Tanzanian Hadza individuals could not be assigned at family and genus level, whereas this was not the case for the Italian control population [85].

To assess the role of different natural environmental factors in shaping the intestinal microbiota and how these factors relate to the influence of the different host species, we divided samples into subsets. This approach allowed us to gain a better insight into the effect of specific factors on microbiota composition, in addition to a more generic analysis of all factors at the same time in a relatively heterogeneous dataset. The value of this tiered approach was confirmed by the fact that for all factors of interest, a first insight into potential effects that could be obtained with the whole dataset, received additional, more robust, support from the analysis of specific subsets of samples. It should also be noted that, due to the nature of wildlife sampling under natural conditions, it remains challenging to obtain balanced sample sets with equal numbers of samples in each group.

We discovered that in samples analysed in our study, the most influential factor contributing to shaping microbiota composition was the area of sampling. When we applied PCoA based on either weighted or unweighted UniFrac distances, separation into areas was obvious in all datasets when included. Remarkably, clustering of samples was tighter with better separation of samples when unweighted UniFrac was used as a distance measure. This observation

suggests, taking into account the nature of the UniFrac distance calculation, that the faecal microbiota of lemurs from different areas is more distinct with respect to microbial species composition than in relative abundance of prevalent taxa. Constrained multivariate analysis (RDA) confirmed that occupancy is the most influential explanatory variable with respect to the observed variation in lemur intestinal microbiota composition. Madagascar is known to have different environmental conditions and biodiversity within relatively small areas [134]. Furthermore, sampling locations were positioned with considerable distance from each other, and were characterized by major climatic differences such as amount of precipitation and forest composition. Hence, it is likely that the availability of food items during the year, in particular fruits and flowers, is the driving force that leads to differences in the intestinal microbiota. These food items are scarce during the dry season in the dry-deciduous forest areas such as Kirindy NP [135, 136], whereas in the areas characterized by wet rainforests such as Ranomafana NP and Andasibe NP, as well as in Nosy Tanikely, these food items are abundantly available almost all year around. Surprisingly, relatively low microbial alpha diversity was observed in the tropical rainforest. It should be noted, however, that samples from this area only included *E. fulvus*. It is tempting to speculate that the lower alpha diversity in *E. fulvus* might be explained by adaptation of the microbiota to sugar-rich food that is prevalent in Nosy Tanikely, due to the abundant presence of mango trees.

We also observed differences in microbiota composition related to the species of lemurs, however, these differences were secondary to those observed between different areas of habitation. Host genetic differences are among the major driving forces shaping intestinal microbiota composition [137], including different animal species with similar dietary habits [86]. In our study this was not the case. One explanation for this finding could be that this study has been conducted on congeneric species which by definition are genetically close [138], and have almost identical digestive systems [139], leading to the moderate effect of genetics on gut microbiota composition observed here.

We did not find any evidence for an influence of sex or age on microbiota composition. It should be noted, however, that in this study different age-classes were not equally represented in the datasets, as we mostly sampled adult individuals (74.6% of all samples). Furthermore, all of the non-adults were at or beyond juvenile stage. Based on knowledge about microbiota development of human infants, the transformation of microbiota to an adult-like mature, composition occurs before reaching the juvenile stage [140, 141]. Surprisingly we did not find any differences in beta- and alpha diversities of microbiota between males and females. Many studies showed an influence of this factor in different species [142], including lemurs [94]. However, it was pointed out before that other factors, such as host genetics, can outweigh the influence of this factor [143]. Hence, it is tempting to speculate that in the present study the effect of sex on the faecal microbiota of the different *Eulemur* species might have been obscured by more influential factors as well as relatively large variation in microbial composition between individual animals.

In conclusion, we showed that intestinal microbiota in three genetically close species of lemurs was most strongly influenced by their occupancy, whereas the influence of genetic differences was minor, and influence of sex and age was not detectable. All three lemur species had similar bacterial composition in terms of predominant and prevalent bacterial taxa. The findings reported here contribute to our knowledge about the intestinal microbiota in non-human primates, and factors that shape the bacterial composition in wild lemur populations, which can be extrapolated into general rules of intestinal microbiota assembly. Furthermore, the high fraction of poorly assigned taxa reinforces the notion that microbiota of non-humanoid primates has so far received little attention, harbouring a broad range of potentially novel bacterial species and genera that deserve attention in future studies.

Funding

This work was supported by an "Nederlandse Organisatie voor Wetenschappelijk Onderzoek" and "The Graduate School for Production Ecology and Resource Conservation" grant [Grant nr. 1208, project nr. 5120873001], and the European Union Erasmus-Mundus PhD fellowship [EM Action 2-Partnership IAMONET- 20132520]. Other necessary funds to conduct the fieldwork of this study were provided by: the KNAW Academy Ecology Fund, the Dr. J.L. Dobberke Foundation, the Treub Foundation, the Dutch Fund for Research on Nature Conservation, the Dutch Royal Botanical Society, and the Dutch Royal Zoological Society.

Acknowledgments

For their advice in the field, the authors want to thank Dr. P. Wright, Dr. E. Larney, and John E. Cadle, and for their logistical support we want to thank the Centre Valbio staff (CVB), the Institute for the Conservation of Tropical Environments (ICTE) and the Madagascar Institute pour la Conservation des Ecosystèmes Tropicaux (MICET), especially B. Andriamihaja, who facilitated arrangements on necessary permits. Thanks go to the Malagasy government and Madagascar National Parks (MNP), Centre de Formation Professionelle Forestière (CFPF), and Mitsinjo Association for their formal permission. For their fieldwork assistance and indispensable knowledge, we want to thank all field technicians and guides that helped us at the various national parks.

Supporting materials

Table S1. Samples collected from three wild *Elemur* spp. across Madagascar. A – adult ; SA – sub-adult; J –juvenile.

Type of environment:		***Wet rain forests***						***Dry-deciduous forest***						***Tropical rain forests***		
Area ⇨ **Species**		*Andasibe Mantadia*			*Ranomafana*			*Ankarafantsika*			*Kirindy Mitea*			*Nosy Tanikely*		
⇩	**Age Sex**	A	SA	J	A	SA	J	A	SA	J	A	SA	J	A	SA	J
E. rubriventer	*Male*				9	4										
	Female				5	4										
E. rufifrons	*Male*				10	4					10	2	4			
	Female				11	2					17	2	9			
E. fulvus	*Male*	6						10						3	2	
	Female	8						11						3	2	

Table 2. Reference abbreviation (Genus ID) of not assigned genera used in the text of the paper.

New ID	**Original taxonomy**
UG_28(RB25)	k__Bacteria;p__Acidobacteria;c__RB25;o__o;f__f;g__g
UG_27(Coriobacteriaceae)	k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__g
UG_15(Coriobacteriales)	k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__f;g__g
g_Bacteroides	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides
UG_5(Bacteroidaceae)	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__g
UG_2(Bacteroidales)	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__f;g__g
UG_13(Porphyromonadaceae)	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__g
UG_3(Prevotellaceae)	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__g
g_Prevotella	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella

UG_22(Rikenellaceae)	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__g
UG_21(Bacteroidales)	k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__g
UG_18(TM7)	k__Bacteria;p__Candidate_division_TM7;c__c;o__o;f__f;g__g
UG_4(Cyanobacteria)	k__Bacteria;p__Cyanobacteria;c__4C0d-2;o__o;f__f;g__g
UG_19(Bacillales)	k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__f;g__g
UG_26(Lactobacillales)	k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__f;g__g
UG_1(Clostridiales)	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__f;g__g
UG_17(Clostridiales)	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Family_XIII_Incertae_Sedis;g__g
UG_6(Lachnospiraceae)	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__g
g_Pseudobutyrivibrio	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Pseudobutyrivibrio
UG_7(Ruminococcaceae)	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__g
f_Ruminococcaceae (g_Incertae_Sedis)	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Incertae_Sedis
g_Ruminococcus	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus
UG_10(Clostridiales)	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__uncultured;g__g
UG_31(Veillonellaceae)	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__g
g_Megamonas	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megamonas
g_Phascolarctobacterium	k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium
UG_24(Erysipelotrichaceae)	k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__g
UG_16(Erysipelotrichales)	k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__f;g__g
UG_20(Lentisphaeria)	k__Bacteria;p__Lentisphaerae;c__Lentisphaeria;o__RFP12_gut_group;f__f;g__g
UG_32(Rhizobiales)	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__alphaI_cluster;g__g

g_Bradyrhizobium	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Bradyrhizobiaceae;g__Bradyrhizobium
g_Mesorhizobium	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Phyllobacteriaceae;g__Mesorhizobium
g_Rhizobium	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;g__Rhizobium
UG_33(Xanthobacteraceae)	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Xanthobacteraceae;g__g
UG_34(Rhodospirillales)	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__f;g__g
UG_25(Rhodospirillaceae)	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__g
g_Thalassospira	k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__Thalassospira
UG_14(Burkholderiales)	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__f;g__g
g_Oxalobacter	k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Oxalobacter
g_Desulfovibrio	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio
UG_8(Desulfovibrionaceae)	k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__g
g_Campylobacter	k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Campylobacteraceae;g__Campylobacter
g_Helicobacter	k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Helicobacter
UG_23(o_Aeromonadales)	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__f;g__g
g_Enterobacter	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Enterobacter
UG_35(Enterobacteriales)	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__f;g__g
g_Halomonas	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales;f__Halomonadaceae;g__Halomonas
g_Haemophilus	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus
UG_11(Xanthomonadales)	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__f;g__g
g_Stenotrophomonas	k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Stenotrophomonas

UG_36(Spirochaetales)	k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__f;g__g
UG_29(Spirochaetaceae)	k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__g
UG_9(Synergistales)	k__Bacteria;p__Synergistetes;c__Synergistia;o__Synergistales;f__f;g__g
g_Anaeroplasma	k__Bacteria;p__Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__Anaeroplasma
UG_30(Anaeroplasmataceae)	k__Bacteria;p__Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__Anaeroplasmataceae;g__g
UG_12(Anaeroplasmatales)	k__Bacteria;p__Tenericutes;c__Mollicutes;o__Anaeroplasmatales;f__f;g__g
UG_37(Mollicutes;RF9)	k__Bacteria;p__Tenericutes;c__Mollicutes;o__RF9;f__f;g__g
UG_38(Verrucomicrobiales)	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__f;g__g
g_Akkermansia	k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Akkermansia
p_NA	k__NA;p__p;c__c;o__o;f__f;g__g

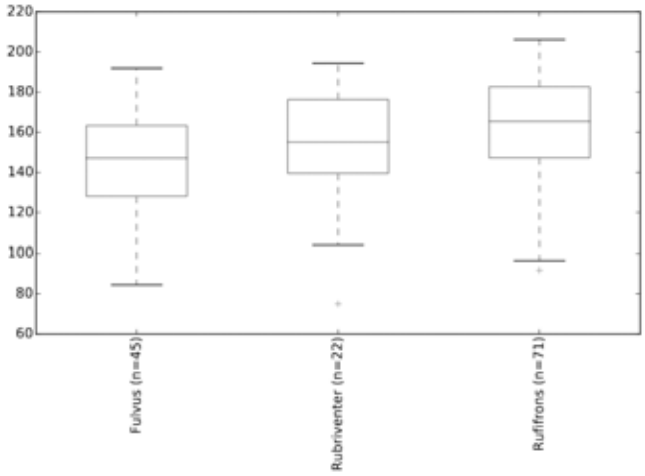

Figure S1. Differences in alpha diversity (chio1) between samples from different lemur species.

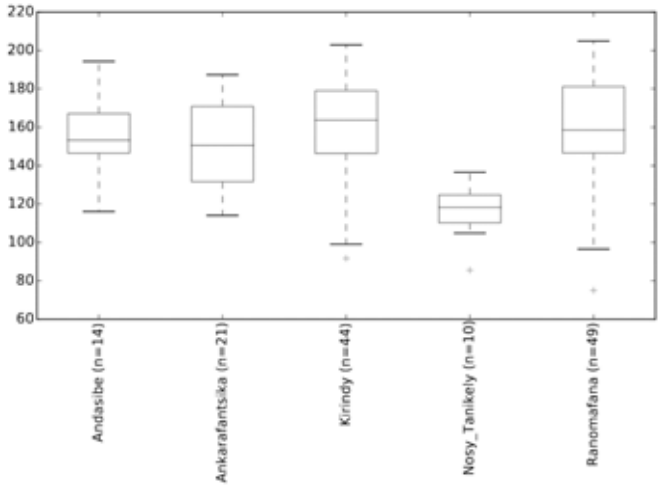

Figure S2. Differences in alpha diversity (chio1) between lemurs from the different areas

Chapter 3

Effects of seasonality and previous logging on faecal helminth-microbiota associations in wild lemurs

de Winter, I.I.*, Umanets, A.*, Gort, G., Nieuwland, W., van Hooft, P., Heitkönig, I.M.A., Kappeler, P.M., Prins, H.H.T.1, Smidt, H.

* Contributed equally

This chapter has been submitted for publication.

Abstract

Gastro-intestinal helminth-microbiota associations are shaped by various ecological processes. However, the effect of the ecological context of the host in terms of geographic location, seasonality (i.e., dry versus wet season), and anthropogenic effects (i.e., logging history) on both groups of gastro-intestinal inhabitants is unknown. We provide a first exploration thereof, and also examine the interactive effects between gastro-intestinal helminths and microbiota. Fresh faecal samples (N = 335) from eight wild *Eulemur* populations were collected over a 2-year period across Madagascar. We used 16S ribosomal RNA gene sequencing to characterise the bacterial microbiota composition, and faecal flotation to isolate and morphologically identify nematode eggs. Infections with nematodes of the genera *Callistoura* and *Lemuricola* occurred in all lemur populations. Seasonality significantly contributed to the observed variation in microbiota composition, especially in the dry deciduous forest. Microbial richness and *Lemuricola* spp. infection prevalence were highest in a previously intensely logged site, while *Callistoura* spp. showed no such pattern. In addition, we observed significant correlations between gastro-intestinal parasites and bacterial microbiota composition in these lemurs. With this study, we show that environmental conditions affect gastro-intestinal nematodes and bacterial interactions in ways that, as far as we know, have not previously been reported.

Introduction

The gastro-intestinal (GI) microbiota plays an important role in the physiology, health, and nutrition of its host [144]. In addition, the GI microbiota can prevent gut colonisation by pathogenic microorganisms [145]. A stable and diverse GI microbiota composition has been shown to be crucial for mammalian health [146, 147], and defining the mechanisms influencing its composition and diversity is considered important [148].

Next to the microbiota, GI macroparasites, including protozoa and nematodes, can be present within a host's digestive tract. They can spread through the faecal-oral route, which involves ingestion of contaminated soil or food [149]. Parasitism can impact the host's health, behaviour, and survival, thereby influencing evolutionary processes and population dynamics [150]. In addition, parasites are known to affect the host's reproduction directly through pathologic effects and mate choice as well as indirectly by impaired nutrition and energy deficits [151].

Faecal bacterial GI microbiota and macroparasites living in internal body surfaces are part of an animals' microbiome and are involved in key host functions [152]. As studying wild populations under natural conditions is rather complex, most studies on the determinants of the GI microbiota composition and parasite prevalence either comprise laboratory or clinical studies that focus on a single host species or infection with a single parasite species [153, 154]. While these studies have provided important insights, understanding of ecological processes that shape composition and functionality of GI microbiota and parasites in wild populations is limited [154].

The composition of the GI microbiota is known to be shaped by multiple factors, including host genetics, evolutionary history, physiology, sex, and age [70, 79]. Several recent studies showed that the microbial composition can remain stable over the host's lifespan [155, 156]. However, other studies found that extrinsic factors, including diet composition [65, 128, 157], pathogens [158], seasonality [159], habitat degradation [72], and geographical differences [160]

influence GI microbiota. For example, it has been shown that the microbial composition in black howler monkeys (*Alouatta pigra*) differs across seasons and is correlated with diet [88]. Also, the distribution of parasite infections in wild host populations is influenced by a number of factors, including host susceptibility and exposure [161]. The nematodes that are the focus of the present study, spend part of their life cycle outside the host and are therefore exposed to environmental conditions that shape temporal variations in parasite infections. Climatic seasonality has been identified as an important driver of this temporal variation in several wild primate species [162, 163]. However, studies investigating these links have yielded different outcomes [88, 164, 165]. It has also been shown that some nematodes have an accelerated development and increased reproduction and survival rates in wetter and warmer conditions [163, 166], and desiccate more frequently under dry circumstances [162]. Several studies found GI parasite richness, prevalence, and abundance to be higher in the warm wet season, compared to the cold dry season, e.g., in lemurs [167-169], chimpanzees [162], as well as howler and spider monkeys (*Ateles geoffroyi*) [170]. However, some helminth species (e.g., *Enterobius* spp.) seem to prefer relatively low temperatures [171]. Although the underlying processes remain unclear [172], these examples show that environmental factors are able to influence the microbial composition and parasite prevalence [70, 159], and require further study in wild mammals.

In addition to environmental factors, the impact of anthropogenic forest disturbance, including logging, on health and pathogens in both wildlife and humans may be far reaching [173]. Anthropogenic forest disturbance may lead to changes in host population densities and interaction patterns of wildlife with humans, domestic animals, and other wildlife species [166, 174]. Such disturbances can thereby enforce changes in the GI microbiota composition and parasite infections [70, 72, 175]. Microbiota diversity can be reduced in degraded areas, as has been shown in howler monkeys, red colobus monkeys (*Procolobus gordonorum*), and other primate species [70, 72, 176]. Furthermore, increased parasite prevalence, virulence, and transmission rates were found in such disturbed forests [175, 177, 178]. Although the exact mechanisms influencing the

microbial composition and parasite infections in disturbed forests is still unknown, nutritional stress is considered important [179]. Nutritional stress can alter the microbiome and lower an animal's immune status, resulting in a higher susceptibility to parasites [180]. Forest disturbance can also directly influence parasites that spend part of their life cycle outside of the host, as changes in forest structure lead to differences in light exposure, temperature, and humidity [181]. Despite the relevance of understanding parasite and microbiome ecology in wild primates living in natural versus human-modified forests, an integrated study on forest disturbance effects on both the parasites and the microbiome has, as far as we know, not been performed before.

Microbiota and parasites co-inhabit the GI-tract and have evolved in close association, suggesting that they have the potential to influence each other [182]. Research on this interplay between host, parasites, and the microbiome has increased over the last decade [183] and recent studies in humans showed associations between nematode infections and changes in the GI microbiota structure [184-186]. However, this observation is not consistent across human populations [187, 188]. Another study experimentally demonstrated that the gut bacterial composition in mice (*Mus musculus*) can change when exposed to a GI parasite (*Trichuris muris*) [189]. Associations between specific bacteria and the abundance of enteric nematodes were also found in wild wood mice (*Apodemus sylvaticus*) [159]. Most of these aforementioned studies focussed on mice, pigs (*Sus scrofa*), or humans. However, recent studies have begun to address the interaction between the microbiome and parasites in primates [190] and we aim to contribute with this study more comparative data on the interactive effect of parasite infections and microbiota composition of wild lemurs.

Specifically, we aim to assess the effects of seasonality (i.e., dry versus wet season), and forest disturbance on the interaction between GI parasites and bacterial microbiota composition in two lemur species. Recently, the microbial composition of lemurs has been studied in captive lemurs [155], in two sympatric wild lemur species [95], and in wild sifakas [191]. However, the processes leading to the natural variation of faecal microbiota in wild lemurs, and how its variation

is influenced by environmental conditions, need further study. Furthermore, only a few studies to date have used a metataxonomic 16S ribosomal RNA (rRNA) gene-targeted approach to address the association and interactive effects between parasites and the microbiome [182, 185, 187-189, 192-194]. In the present study, we focus on four congeneric lemur species at eight geographic locations: *Eulemur rufifrons, E. fulvus, E. macaco*, and *E. rubriventer*. The large heterogeneity in lemur habitats across Madagascar is created by an interaction of the east-west and north-south rainfall gradient [195]. The four lemur species belong to the genus *Eulemur* and are morphologically alike [196], are present in the distinct geographic regions of Madagascar, and inhabit both large intact forests and forests that have experienced past logging [103].

Given the major role of environmental factors in shaping seasonal variation in microbial community structure and parasite infections, we expected that (1) lemurs inhabiting the dry deciduous forests of western Madagascar with strong seasonal variation in rainfall and temperature show larger seasonal contrasts in both parasite infections and microbial composition compared to lemurs in the rainforests of eastern Madagascar with less seasonal variation. We further expected (2) that the microbiota composition is altered and parasite infection prevalence is increased in lemurs whose habitat is restricted to previously logged rainforests compared to lemurs living in less disturbed forests. Lastly, we explored (3) correlations between GI microbiota and natural parasite infections. Hence, in this study we determine how the GI microbiota and parasite infections vary with their geographic distribution spatially in wild lemurs along with seasonal variation and past logging. In addition, we explore the interactive effects between the parasites and microbiota present.

Materials and Methods

Study site

Our research was performed in eight geographically distinct sites (Fig. 1, Table S1). Kirindy Forest, Ankarafantsika National Park (NP), and Zombitse NP are located on the western, north-western, and south-western side of Madagascar, respectively. They consist of dry deciduous forest with pronounced seasonality [197]. These western regions have a higher annual mean temperature than the eastern rainforests, but receive less rainfall.

In contrast, Andasibe Mantadia NP and Ranomafana NP are located on the eastern side of Madagascar and are relatively wet rain forests with a less distinct dry season compared to the western areas [96]. Within Ranomafana NP, we distinguished two research sites, Talatakely (TALA) and Vatoharanana-Valohoaka (VATO-VALO) with different degrees of anthropogenic disturbance (Fig. 2) [198, 199]. Before the establishment of the national park in 1991, the forests in this area were used by local inhabitants, amongst others for slash-and-burn agriculture [96]. Now, more than 25 years after the last logging activities, Ranomafana NP shows a high heterogeneity in forest structure.

The islands Nosy Be, Nosy Komba, and Nosy Tanikely, are located in the north-west of Madagascar. The forests of Nosy Be (~320 km^2) are largely replaced by coffee, fruit, and ylang-ylang plantations, and by rice and sugar cane fields. Only Lokobe NP (~7 km^2) at the south-eastern part of the island still contains the island's original forest vegetation. Nosy Komba and Nosy Tanikely are located in between Nosy Be and the mainland. The vegetation on Nosy Komba (~25 km^2) is similar to Nosy Be. The vegetation at Nosy Tanikely (~0.06 km^2) mainly consists of low forest and bushy vegetation, including palm trees and planted banana and mango trees, surrounded by a sandy shore with large rock formations (de Winter, *pers. obs.;* [105]).

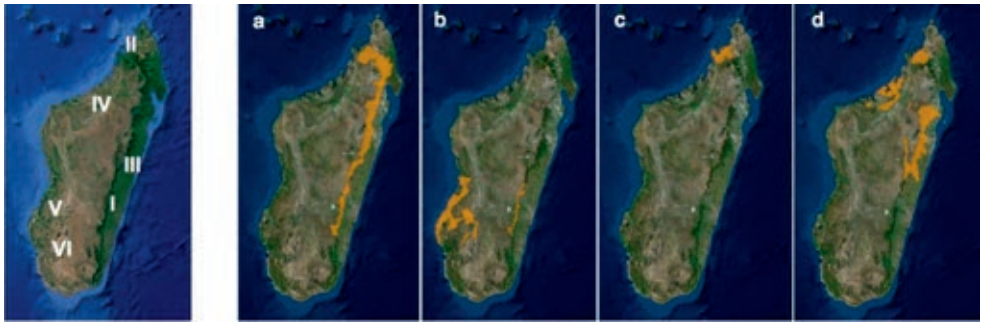

Figure 1. Study sites and the geographic ranges of the different *Eulemur* species (Google Maps, 2015). Left: Map of Madagascar with the study sites Ranomafana NP (I), Nosy Be, Nosy Komba, and Nosy Tanikely (II), Andasibe NP (III), Ankarafantsika NP (IV), Kirindy Forest Reserve (V), Zombitse NP (VI). Right: the geographic ranges of a) *E. rubriventer*, b) *E. rufifrons,* c) *E. macaco,* d) *E. fulvus*.

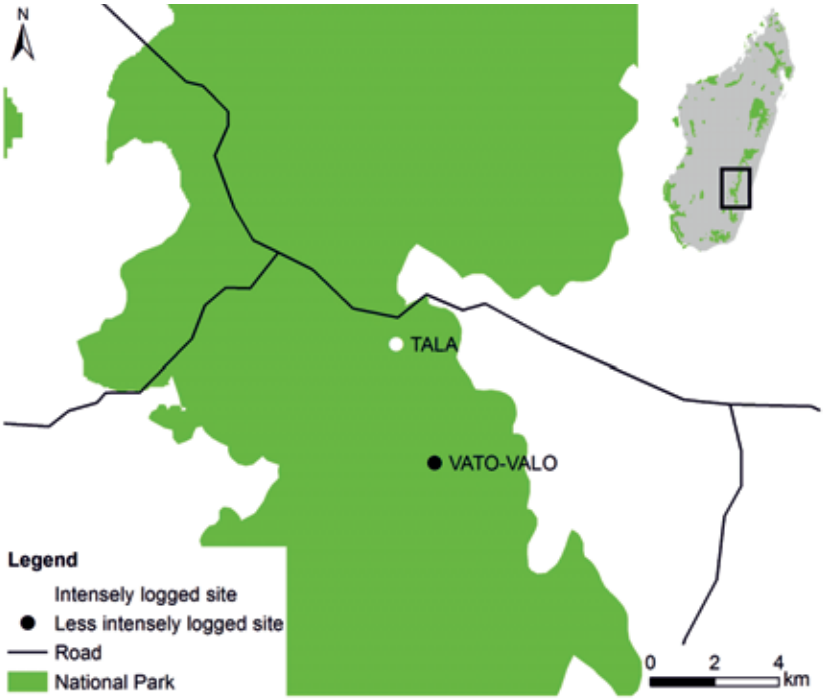

Figure 2. Map of Ranomafana National Park and the two forest sites that were surveyed in this study. Talatakely (white dot) experienced relatively intense logging in the past, while Vatoharanana- Valohoaka (black dot) experienced no such disturbances. This map was generated via ArcGIS version 10.5. Data was downloaded from UNEP-WCMC and IUCN (2016).

Study species

True lemurs (genus *Eulemur*, family Lemuridae) are medium-sized (body and tail length 30-50 cm, 2-4 kg) arboreal primates that occasionally move on four legs on the ground. They live in social groups ranging from two to fifteen individuals and their diet primarily consists of fruits, flowers, and leaves [196]. We studied four *Eulemur* species: *Eulemur rufifrons, E. fulvus, E. macaco*, and *E. rubriventer*. The main difference between the *Eulemur* species is their group size *E. rufifrons, E. macaco,* and *E. fulvus* live in multi-male, multi-female groups from four to 18 individuals [107, 110, 200], whereas *E. rubriventer* lives in small monogamous groups from two up to five individuals [112]. *E. rufifrons* lives in the southwest and east and the native range of *E. fulvus* is in the north of Madagascar, on both the east and west side [201]. This species has also been introduced to the northern island Nosy Tanikely. *E. macaco* is found on the mainland and several islands in the north-west, while *E. rubriventer* inhabits forests in eastern Madagascar (Fig. 1, Table S1). *E. rubriventer* and *E. rufifrons* live sympatrically in Ranomafana NP [92].

Faecal sample collection

We collected 338 faecal samples between October 2013 and February 2015, of which 138 were also used in a previous study [202]. Within Ranomafana NP we collected 103 samples; 38 samples from a previously logged site (Talatakely) and 65 from a less disturbed site in terms of its logging history (Vatoharanana-Valohoaka). Immediately after defecation, fresh faecal samples (3-4 g) were collected non-invasively. We noted visual characteristics, i.e., consistency, colour, presence of blood, mucus, or tapeworm proglottids. We also reported GPS coordinates, time, group size, group composition, age (sub-adult if ≤ 2 years old or adult if ≥ 3 years old), and sex. We allocated a body fur condition score to the individuals whose faeces were collected [203]. We aimed at sampling all adults within a social group and we did not resample the same individuals. As soon as we were not sure whether the faeces were from a new individual or whether we

already sampled the animal, we moved on to another group. As we worked mostly within National Parks or Reserves, the lemurs were all habituated to human observers, mainly due to the frequent visits by tourists or researchers, which facilitated the faecal collection. We found no abnormalities in the consistency and colour of the faeces and we did not find blood, mucus or tapeworm proglottids in any of the faecal samples. Within twelve hours after collection, each faecal sample was divided over two sterile tubes: one gram of faeces was stored in a tube filled with 5 ml of 70% ethanol and two grams of faeces was placed in a tube filled with 15 ml SAF fixative [179, 204]. Samples were analysed at the Laboratory of Microbiology, Wageningen University, and the Department of Infectious Diseases and Immunology, Utrecht University. All described methods were performed in accordance with the relevant guidelines and regulations and was approved by the trilateral commission (CAFF/CORE) in Madagascar (permits 297/13 and 143/14/MEF/SG/DGF/DCB.SAP/SCBSE).

DNA-based bacterial composition analyses

Faecal bacterial microbiota composition, determined by next generation sequencing of 16S rRNA gene fragments, was used as proxy for the intestinal microbial community. We extracted microbial DNA from the faecal samples collected in Ranomafana NP following a modified double bead-beating procedure using the QIAamp® DNA Stool Mini Kit (Qiagen, Valencia, CA, USA) [113]. For the sample processing, we used the modified protocol proposed by Yu & Morrison (2004) [114]. Prior to DNA extraction faecal material was air-dried during 15-20 min in a fume hood to remove ethanol from samples. We extracted DNA from samples collected at the other sites using the Maxwell® 16 Research Instrument (Promega, Madison, USA) in combination with the corresponding RNA extraction kit customised for faecal DNA extraction according to manufacturer's instructions. Prior to DNA extraction, samples were rehydrated through series of ethanol solutions with decreasing proportions of ethanol in steps of 10%. For rehydration, 1.5 ml of 70% ethanol with faecal particles was transferred into a

fresh 2 ml tube and centrifuged at 13,000 rpm for 5 min. After centrifugation, part of the supernatant was replaced with the same amount of distilled water to decrease ethanol concentration by 10%, vortexed, and incubated for 10 min at RT. These steps were repeated until the ethanol was completely replaced by distilled water. Cell disruption and lysis was performed as described above, but instead of lysis buffer we used S.T.A.R buffer (Roche Molecular Systems, Boston, USA). DNA quality and concentration were spectrophotometrically verified (Nanodrop Technologies, Wilmington, USA. For each sample, barcoded amplicons were amplified from 40 ng of extracted DNA using a two-step PCR method in a LabCycler Gradient (SensoQuest, Göttingen, Germany) and pooled afterwards as described previously [118]. Briefly, the V1 - V2 region of the 16S rRNA was first amplified by PCR (25 cycles of 95 ˚C (30 s), 52 ˚C (40 s), and 72 ˚C (90 s)), followed by post-elongation (72˚C, 7 min) using primer pair 27F–DegS: 5'–GTTYGATYMTGGCTCAG–3' [115] and 338R–I: 5'–GCWGCCTCCCGTAGGAGT–3' / 338R–II: 5'–GCWGCCACCCGTAGGTGT–3' [116, 117] that contained forward and reverse linkers UniTag I (5'–GAGCCGTAGCCAGTCTGC–3') and UniTag2 (5'–GCCGTGACCGTGACATCG–3'), respectively. Amplicons were then used as template for a second PCR in order to introduce sample-specific barcodes, using individual barcode primers targeting UniTag1 and UniTag2 sequences. The amount and size of the amplicons were checked visually by agarose gel electrophoresis. The PCR products were purified and concentrated using magnetic beads (MagBio, Lausanne, Switzerland) according to the HighPrep protocol, quantified using the Qubit dsDNA BR Assay Kit (Life Technologies, Austin, USA), and pooled in equimolar amounts into libraries of 48 samples, including two mock communities of defined composition, for paired-end sequencing (300 bp) on the Illumina Miseq platform at the European Genome and Diagnostics Centre (Konstanz, Germany; now part of Eurofins Genomics Germany GmbH). Mock communities, i.e. mixes of quantified and purified copies of bacterial 16S rRNA genes in known proportions, are routinely used in our laboratory to assess quality and reliability of a sequencing run, amplicon preparations, and quality of data processing, as was described previously [47].

The amplicon sequences were demultiplexed and the subsequent analysis of raw rRNA gene sequence data was performed using NG-Tax [47]. Reads assigned to OTUs of plant origin such as chloroplast and plant mitochondrial DNA were removed from the dataset used for downstream analysis. The raw data was ranked per individual sample based on the matching of reads to OTUs, allowing an error of one nucleotide.

Parasite isolation

The collected faecal samples were examined for the presence of GI nematodes with the use of the Centrifugation-Sedimentation-Flotation (CSF) method [205]. GI nematode species identification was based on morphological traits such as colour, shape, size, and content of eggs [206-208]. A rough estimation of the number of parasite eggs per gram of faeces (EGP) was obtained by simple counts. Since the number of eggs that end up in the faeces is not a reliable index of adult worm burden [209], the egg count cannot be regarded as a measurement of infection intensity, but rather as a measurement of infectivity.

Statistical analysis

After initial sequence data processing with NG-tax, we combined the OTU table, metadata, and phylogenetic tree into a "phyloseq" object, as implemented in the "phyloseq" R package (v.1.22.3) [210]. Further analyses were carried out in R statistical and programming environment (v 3.4.1). OTUs that were encountered in less than three samples, OTUs not assigned to any taxonomic level (NA) and OTUs identified as chloroplast and mitochondria were removed. In addition, samples with low numbers of reads (less than 1000 reads), missing metadata of interest, and one sample (i.e., 'NT9F', due to the low quality of the starting material) were removed from the data set. For beta diversity analysis, the weighted UniFrac distance matrix was calculated from the OTU table and

phylogenetic tree as implemented in the "phyloseq" package, with the phylogenetic tree rooted at midpoint (package "phangorn" [211]). Multidimensional scaling with weighted UniFrac as a distance matrix (PCoA) was applied (package "phyloseq") to obtain a first insight into the beta diversity of faecal microbial communities in the investigated lemur populations. We used dbRDA and an ANOVA like permutation test (anova.cca; permutations = 9999) to identify variables that significantly contribute to explaining the observed variation in microbial composition (package "vegan") [212]. Variable "Social Group" was excluded from the analysis due to extremely uneven sample distribution, with 28 out of a total of 92 social groups including only one sample. The degree to which individual factors could explain microbiota composition was estimated by partial dbRDA with control for variables that were not used as a constraint. R^2 values were used as estimator of variation explained by a constraint (package "vegan"). Phylogenetic diversity was used as a primary alpha diversity measure, and was calculated from the phyloseq object with the OTU table rarefied at a read depth of 1051, using a custom function (author Thomas W. Battaglia, https://github.com/twbattaglia). Statistical differences between alpha diversity of pre-defined sample groups was assessed by posthoc Kruskal Nemenyi-tests (package "PMCMR") [213]. The datasets generated during this study are available in the public read archive EBI (study name 'ena-STUDY-WAGENINGEN UNIVERSIT-03-04-2017-14:57', with accession number 'PRJEB20227' (link: https://www.ebi.ac.uk/ena/data/view/PRJEB20227)).

To analyse the effect of seasonality (early dry vs early wet) and location (western dry deciduous forests vs eastern rainforests) on the infection prevalence of *Callistoura* and *Lemuricola* spp. in *Eulemur* species, Generalized Linear Mixed Models (GLMMs) were used, assuming a binomial distribution and logit link function for data aggregated per social group. We included random effects for sites within location and observation-level random effects for social groups, and fixed effects for species, season, location, and the interaction between season and location. The observation-level random effects handle possible binomial overdispersion. The factor species entered the model as a control variable to avoid confounding of location effects with species effects. We focused specifically on the

interaction between location and season in order to test the seasonality hypothesis as formulated in the Introduction. To present estimated infection prevalence with 95% confidence intervals (CI) on the probability scale, we back-transformed the results (on the logit-scale) from the GLMMs first and next applied a shrinkage factor [214], which is needed for GLMMs, to obtain predicted population means instead of medians. To test whether infections by the two nematode genera occurred independently, we modified the GLMM for *Callistoura* spp., using unaggregated data, by adding an indicator variable for *Lemuricola* spp. as regressor to the model. In this way, we allowed the infection prevalence for *Callistoura* spp. to be different among lemurs with or without *Lemuricola* spp. infections.

In a subset of the data (Ranomafana NP; n = 103 individuals of *E. rubriventer* and *E. rufifrons* only), we analysed *Callistoura* and *Lemuricola* spp. infection prevalence comparing disturbed and less disturbed subsites. Again, we aggregated infection scores per social group, and used ordinary Generalised Linear Models (GLMs) assuming a binomial distribution for the number of infected animals per social group and logit link function. We entered effects for xthe control factor species, and for the main factors of interest: disturbance (less vs more disturbed subsites), season (early dry vs early wet), and their interaction into the model. In the analysis of *Callistoura* spp. prevalence, a smaller model was fitted due to the low numbers of cases (14 cases, with just 1 in the less disturbed site). Extra-binomial variation could not be ruled out, because individuals within social groups may have correlated responses. Because of different group sizes (range 1-7), we used Williams' method as available in the dispmod package of R [215]. If the overdispersion was not present, when judging the residual deviance, we used an ordinary binomial GLM. We calculated back-transformed predicted means presented with 95% CI for the previously disturbed and less disturbed sites.

Model assumptions were checked by inspection of residuals, leverages, and collinearity statistics, and model stability (for GLMM) and dfbetas (for GLM) were assessed. After model checking, comparisons of the full model (separately

for the analyses of *Callistoura* and *Lemuricola* spp. infections, seasonality, and disturbance analyses) with reduced models were made using likelihood ratio tests (LRT), followed by tests for individual factors in case of significant results. Regardless of the results from omnibus tests, we tested the specifically formulated hypotheses regarding seasonality and forest disturbance (see Introduction). Pseudo R^2 [216] for the full models were calculated (Table S3).

We performed the statistical analyses in base R [217], using the R packages lme4 for the GLMMs [218] and emmeans [219] for prediction of group means, with car [220] for variance inflation factors, DHARMAa [221] for residual checking in GLMMs, and MuMIn [222] for creating pseudo R^2 values.

Results

Seasonality

We found clear separation of samples by season in the bacterial microbiota composition of multiple lemur populations sampled across Madagascar, using principal coordinate analysis (PCoA) based on the weighted UniFrac distance matrix (early wet season n = 128, early dry season n = 196, R^2= 0.08, Adonis; p = 0.0001, Fig. 3). Distance-based redundancy analysis (dbRDA), identified the area of sample collection as the most influential variable followed by season, when considering all samples as a single dataset. We observed an increase in the percentage of explained variance in microbiota composition by seasonality when we focused on samples collected within one area and one lemur species (Fig. 4). Specifically, for *E. fulvus* populations from Ankarafantsika NP and Andasibe NP, and *E. rubriventer* and *E. rufifrons* populations from Ranomafana NP, the percentage of variation in microbiota composition explained by season increased from 5.7% for the entire dataset to 16.9%, 20.2%, 12.5% and 13.5%, respectively (Fig. 4a-d). Hence, these populations harboured a different microbial composition in the early dry season compared to the early wet season. With regards to alpha diversity, the *E. fulvus* population in Ankarafantsika showed a significantly

higher mean phylogenetic diversity (PD index, $p < 0.001$) in the early dry season ($n = 21$) compared to the early wet season ($n = 29$). No statistically significant differences in alpha diversity were observed for other subsets of samples as defined by area of habitation and lemur species, as we showed in a previous study [202].

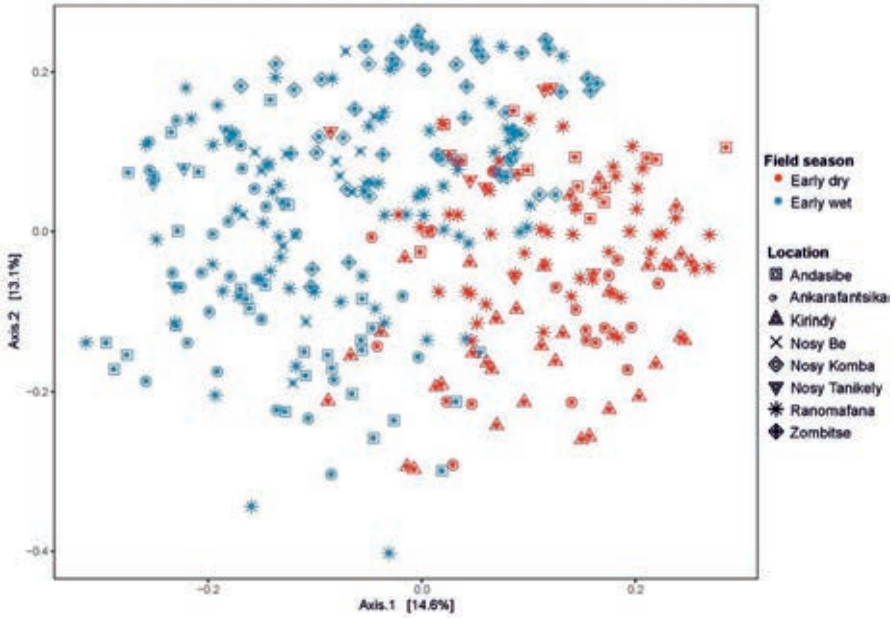

Figure 3. Lemur faecal microbiota composition across seasons and locations. Ordination of faecal microbial composition in multiple lemur populations across Madagascar sampled in different seasons (early dry and early wet) and at multiple locations. This figure shows the results of a principal coordinate analysis (PCoA) based on the weighted UniFrac distance matrix. Grouping strength of samples by season- R2= 0.09 (Adonis; p = 0.001).

Based on morphological analyses, nematode species of two genera, *Callistoura* and *Lemuricola,* were present in the GI tract of nearly all *Eulemur* individuals from eight geographically distinct populations (Fig. 5). Of all the sampled lemurs (N = 335), 188 (56.1%) were only infected with *Callistoura* spp., 17 (5.1%) were only infected with *Lemuricola* spp., 34 (10.1%) were infected with both nematode species, and 96 (28.7%) were not infected (Table S2). The observed co-occurrence (10.1%) is very close to the expected co-occurrence for independent infections (67.5% x 15.1%= 9.9%), suggesting that infections with both

Callistoura and *Lemuricola* spp. occured independently, and therefore, co-infection appeared to be independent.

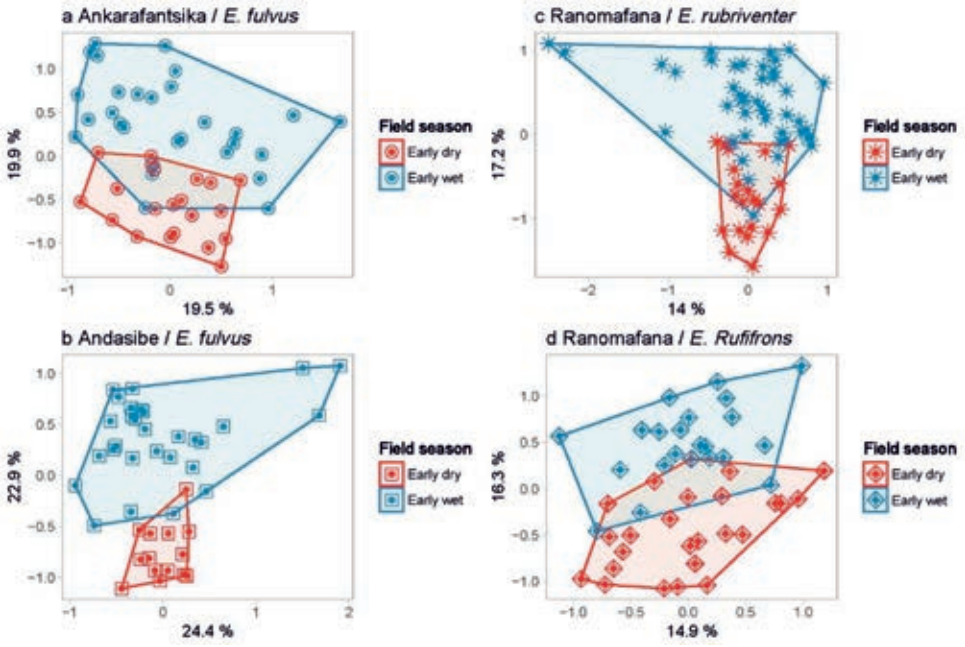

Figure 4. Lemur faecal microbiota composition across seasons and locations. dbRDA Analyses of the abundance-weighted phylogenetic composition at OTU level of individual lemurs across seasons (early dry and early wet) in different geographic areas visualised by ordination. Faecal microbiota significantly clustered by season. Results are given for the percentage of variation explained by the sum of the first two canonical axes, percentage explained by season with corresponding P-value. a) *Eulemur fulvus* in Ankarafantsika National Park, (39.8 %, 16.9%, p = 0.001). b) *E. fulvus* in Andasibe (46.5%, 20.2%, p = 0.001. c) *E. rufifrons* in Ranomafana NP (31.1%, 13.5%, p = 0.001). d) *E. rubriventer*, Ranomafana NP (31.3%, 12.5%, p = 0.001).

In the analysis of *Callistoura* spp. prevalence across seasons and locations using GLMMs, we did not find alarming problems regarding model diagnostics. A highly significant full model was found (LRT; $p < 0.001$, Table S3; model SC1 vs SC2). This result was solely attributed to the random part of the model (SC1 vs SC3), with a larger part being explained by variation among social groups, representing binomial overdispersion (SC1 vs SC5, p = 0.0004), compared to variation among sites (SC1 vs SC6; p = 0.021). We found no significant effect of species, location, season, and the interaction between location and season (SC1 vs

SC4; p = 0.20). The hypothesised interaction of location and season was not significantly different from zero (SC1 vs SC7; p = 0.32).

Also, in the analysis of *Lemuricola* spp. prevalence across seasons and locations using GLMMs, no alarming problems regarding model diagnostics were found (Table S3). The full model explained a significant amount of variation (Table S3; model SL1 vs SL2; p = 0.012). We did not find significant variation due to random effects for sites or social groups, and hence no binomial overdispersion (SL1 vs SL3; p = 0.50), but the fixed part of the model was significant (SL1 vs SL4; p = 0.026). The hypothesised interaction of location and season was not found (SL1 vs SL6; p = 0.84), but the location main effect was significant (SL1 vs SL7; p = 0.032). The *Lemuricola* spp. prevalence was estimated as 25% (95% Confidence Interval (CI): 16%-37%) in the dry Western areas compared to 10% (5%-20%) in the wet Eastern areas.

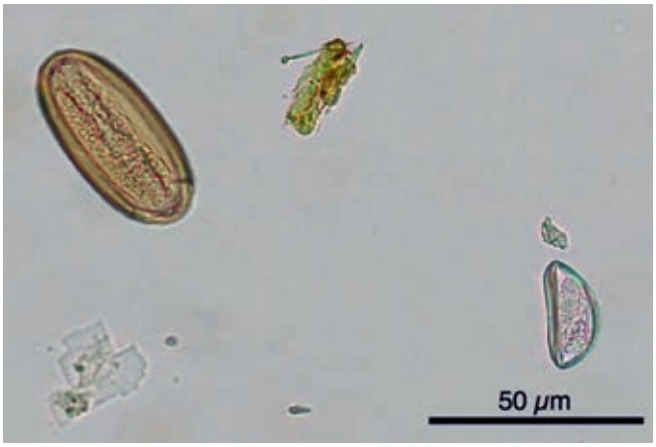

Figure 5. Detected parasite species. *Callistoura* sp. egg (left) and *Lemuricola* sp. egg (right), isolated from a faecal sample of *Eulemur rufifrons*, magnification 200x (picture taken by IdW).

No difference in infection prevalence of *Callistoura* spp. between animals with and without *Lemuricola* spp. infection was found (LRT; p = 0.37). Overall, we found that 188 out of 284 lemurs without *Lemuricola* spp. were infected with *Callistoura* spp. (66%), and 34 out of 51 *Lemuricola* spp. infected animals were infected with *Callistoura* spp. (67%).

Disturbance

A possible association between forest disturbance and parasite infection and faecal bacterial microbiota composition was examined in lemurs from Ranomafana NP. Bacterial richness was significantly higher at the previously logged site (Talatakely, n = 29), compared to the less disturbed site (Vatoharanana/Valohoaka n = 27) (PD index= 7.3±1.1 vs 5.8±1.7, p = 0.001). The dbRDA also showed that the microbial composition was grouped according to sites with a different disturbance history (p = 0.004, Fig. 6).

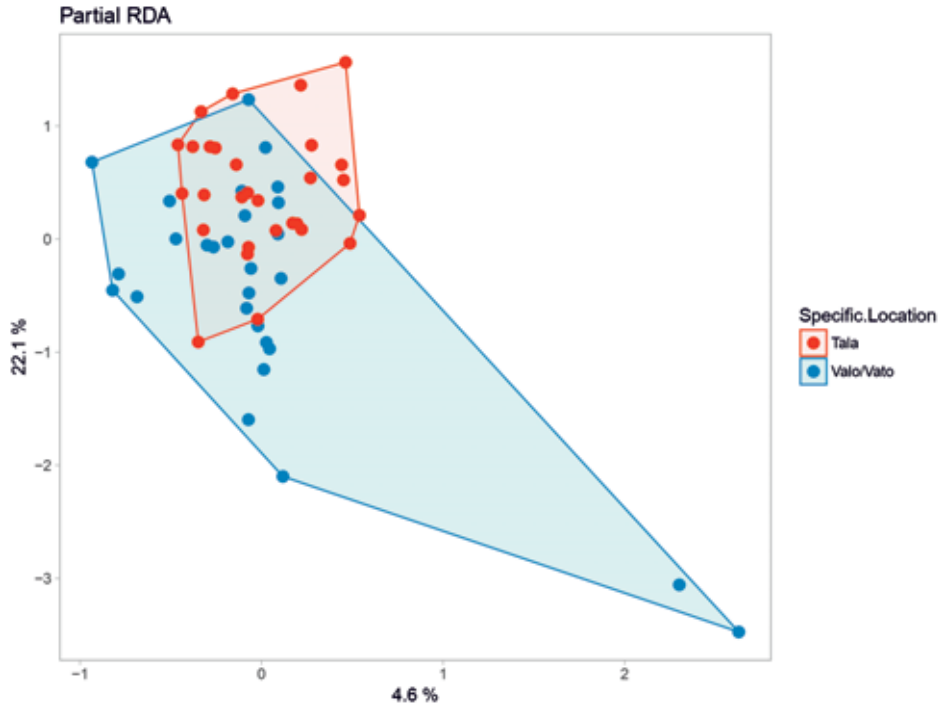

Figure 6. Faecal microbiota composition in disturbed and less disturbed sites. Ordination (RDA) of the microbial composition (OTU) across sites with a different disturbance history (disturbed vs. less disturbed) for *Eulemur rubriventer* and *E. rufifrons* in Ranomafana National Park, Madagascar. Cumulative variation explained by the first two axes was 26.7 % and the sampling location accounted for 3.8% of the total variation (p = 0.002).

In the analysis of prevalence of *Callistoura* spp. after model checking, we did not find an overall significant effect of species, disturbance, season or their interaction (Table S3; model DC1 vs DC2, $p = 0.185$). Focusing on the specific hypothesis on disturbance though, a significantly different prevalence between the two subsites was found (Table S3; model DC1 vs DC3, $p = 0.042$). The prevalence of *Callistoura* spp. in the non-disturbed subsite was 85% (95% CI: 72%-93%), and in the disturbed subsite 53% (36%-70%)(Fig. 7).

In the analysis of prevalence of *Lemuricola* spp., the omnibus test showed highly significant effects of species (Table S3; model DL1 vs DL2; $p < 0.001$), which could not be traced back to differences between subsites with different disturbance histories (Table S3; model DL1 vs DL4; $p < 0.001$). The prevalence in the less disturbed compared to the previously logged subsite was estimated as 1.2% (95% CI: 02%-09%) compared to 26% (15%-47%) in the less disturbed subsite. The infection rates of *Callistoura* spp. showed considerable extrabinomial variation, but the infection rates of *Lemuricola* spp. did not.

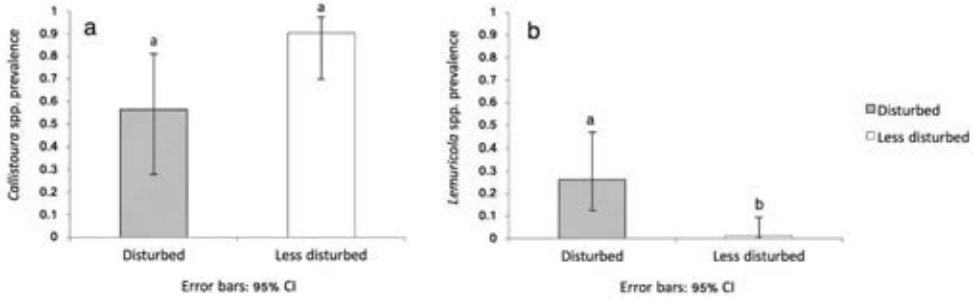

Figure 7. Parasite prevalence with disturbance. a) *Callistoura* spp. prevalence, b) *Lemuricola* spp. prevalence in *Eulemur rufifrons* and *E. rubriventer* populations in a previously disturbed and less disturbed site in Ranomafana NP, Madagascar. Mean with 95% confidence intervals and the letter coding above the bars indicate whether groups are significantly different.

Microbiota and parasites

Constrained ordination (dbRDA) showed that prevalence of *Callistoura* spp. accounted for 0.4% ($p = 0.024$) of the variation in microbiota composition found among all samples with available microbial and parasite infection data ($N = 324$), regardless of host species and habitation. However, we did not observe significant ($p = 0.49$) influence of *Lemuricola* spp. prevalence on microbiota composition. When samples from different seasons were analysed separately, we observed an increase of the relative weight of explained variations to 0.7% with maintained, albeit decreased (due to smaller sample size) significance ($p = 0.05$) in samples collected during early-wet season, and no significant influence during early-dry season.

When focussing on lemurs of one species from the same area and season, we could not find statistically significant correlations with *Callistoura* spp. prevalence. However, among the *E. rubriventer* population in Ranomafana NP in the early dry season, microbiota composition showed a nearly significant correlation with *Lemuricola* spp. prevalence ($p = 0.055$) with 9.2% of variation explained by this factor. Interestingly, a clear separation of samples could be observed in the corresponding dbRDA plots (data not shown), albeit without statistical support (all $p > 0.05$), probably due to the relatively low and unequal number of samples per group.

Discussion

We assessed the influence of environmental conditions on the faecal bacterial microbiota composition and parasite infections as well as the correlation between GI microbiota and parasites in wild lemurs. The two helminth genera *Callistoura* [223] and *Lemuricola* [224] were the only two nematode genera detected in all *Eulemur* populations. These microphagous pinworms belong to the family Oxyuridae and are directly transmitted [206]. They colonise distinct parts of the gut of their hosts: *Callistoura* spp. live in the ileum and colon and

Lemuricola spp. in the caecum and colon [206]. These parasite species were also found in most other lemur genera [206, 224], including other species from the genus *Eulemur*, i.e. in *E. flavifrons* [225], *E. macaco* [226], *E. fulvus* [227], and *E. albifrons* [228]. Hence, these nematode genera have a very broad distribution throughout Madagascar and do not show obvious specificity to a particular lemur host species [206].

We hypothesised that lemurs inhabiting dry deciduous forests, with strong seasonal variation in rainfall and temperature, would show larger seasonal contrasts in both parasite infections and microbial composition compared to lemurs in eastern rainforests with relatively low seasonal variation. Nevertheless, we found a strong seasonal contrast in the microbial composition at Organisational Taxonomic Unit (OTU) level across all lemur populations. Across Madagascar, lemurs are exposed to seasonality and have been observed to change their diet accordingly [103]. Diet was found to be an important driver of the GI microbial composition in many human studies (e.g., [128]). Although humans are assumed to have a stable microbiota over longer periods of time (> 10 days) [229], dietary changes can alter the relative abundance of specific members of the microbiota within 24 hours [65]. With respect to wildlife, e.g., wood mice (*Apodemus sylvaticus*) were shown to exhibit seasonal shifts in gut microbiota structure that coincide with their annual dietary changes [159]. Also, in wild Mexican black howler monkeys, temporal changes in the relative abundance of gut bacteria were strongly correlated with dietary variations [88]. In another study on *Eulemurs* we showed that differences in diet in geographically separated populations strongly influence intestinal microbiota [202]. Hence, seasonal diet shifts are likely to explain most of the variation in microbiota in lemurs across seasons observed in the current study.

In addition, the microbial diversity from lemurs in Ankarafantsika National Park [230] was higher in the early dry season compared to the early wet season. Over the dry season, lemurs experience conditions of relatively low temperatures and food and water restriction, especially in the dry western parts of Madagascar. This nutritional stress may result in a narrower diet and the

microbiota would be more specifically adapted to the food items available. This narrower diet during the dry season could therefore explain the gradual decrease in microbiota richness that we observed. Such dietary change might lead to an altered microbial composition, which potentially facilitates the digestion of specific food items. It is tempting to speculate that this could also lead to an increased caloric intake, which might contribute to an increased fitness of both the host and microbiota [159].

The presence of different fruit trees result in large dietary differences across populations [231, 232]. For example, the four most predominant food items consumed by *E. fulvus* in Ankarafantsika in the early and early wet season, were *Buddleja madagascariensis, Psychotria* sp., *Vitex perrieri* and *Diospyros tropophylla*)[231, 233], species that do not occur in Nosy Tanikely or Andasibe [232]. Furthermore, introduced mango trees (*Mangifera indica*) are only consumed at Nosy Tanikely. However, there is also some dietary overlap across populations, i.e., *Dichapetalum leucosia* and *Landolphia myrtifolia* were consumed by *E. fulvus* in both Ankarafantsika and Andasibe. Despite the overlap in some fruit species, the geographically separated lemur populations of this lemur species showed major dietary differences, probably leading to the observed major variations in microbiota composition in these populations.

We found a slight, indication that parasite infections in the dry regions of Madagascar showed larger seasonal contrasts compared to the eastern rainforest. Another study also found a higher parasite richness in areas with a large precipitation range throughout the year [234]. Many parasites require a certain temperature and humidity to complete their life cycles [234] or as microhabitats for their larva [235]. The drier conditions towards the end of the dry season can prevent egg development and can lead to desiccation of the fragile eggs [166]. However, some related nematode species are able to survive such short periods of drought by entering a state of hypobiosis, until humidity conditions improve to the point where free-living larval stages can survive [235]. In addition to these direct seasonal influences on parasites, the lemur host influences these infection patterns as well. The host's resource use and diet in general are considered as

major determinants of host exposure to parasites [166]. It was also experimentally established that host foraging ecology has important consequences for the exposure to and transmission of parasites [236]. Food scarcity for lemurs is relatively high towards the end of the dry season [111, 237] and the associated nutritional stress can have a repressive effect on the hosts' immune system, which may result in a higher susceptibility to parasite infection [175].

Seasonal changes in lemur reproductive status can also lead to changes in parasite infections patterns [207]. The early dry season coincides with the mating season of *Eulemur* spp. [238], and more frequent physical contact both within and between lemur groups during this period may enhance parasite infection [207]. Besides, androgen and glucocorticoid levels of the males and oestrogen levels of the females increase during the mating season, which can lead to a higher susceptibility to parasite infections due to their repressive effect on the immune system [239]. Furthermore, the early wet season coincides with the weaning season, a season that is energy demanding, especially for lactating females. These behavioural and physiological differences may thus lead to differences in parasites infection status across different seasons. It is likely that because of all these factors that influence parasite infections, we did not find a stronger effect of seasonality in areas with stronger seasonal contrasts.

We also did not find an interactive effect of the two nematode species as co-infection appears to be independent. *Lemuricola* and *Callistoura* spp. colonise distinct parts of the gastro-intestinal tract of their hosts, the caecum-colon and ileum-colon, respectively [206], which can explain the lack of interactions between these two species.

We hypothesised the microbiota composition to be altered and parasite infection prevalence to be increased in lemurs whose habitat is restricted to more intensely logged forests. For the microbial composition, we found statistically significant variation between samples taken at a previously logged and at a less disturbed site. Moreover, a higher richness of microbial consortia was observed in the logged area. Although only few studies have addressed the impact of

anthropogenic disturbance on gut microbiota of wild primates, most studies seem to contradict our findings. For example, habitat disturbance was reported to lead to reductions in *Alouatta* gut microbial diversity [72], and a similar pattern was found in Udzungwa red colobus monkeys [70]. These results may reflect a general pattern of habitat degradation and reduced diversity in the ecological pool of microbial taxa available to colonize hosts [72]. However, the number of studies in this field are very limited. In addition, the type and intensity of anthropogenic disturbance and the forests' regeneration time may be important as well [240]. Logging in our sites occurred nearly thirty years ago and sites have been regenerating since [241], which can explain the deviating patterns that were found in this study. Nevertheless, these forests still differ to a large extent in their structural characteristics, as well as tree species composition [198], which may explain the differences in microbiota composition we found.

Remarkably, we found a relatively high abundance of Cyanobacteria in the *Eulemur* population in the less disturbed compared to the previously logged site. Sequences identified as Cyanobacteria are most probably derived from their non-photosynthetic gut dwelling siblings [132]. Even though they are part of the normal gut microbiota of mammals, it is not clear what role they play in intestinal ecosystems.

Concerning parasites, the prevalence of *Lemuricola* spp. was significantly higher in the more intensely disturbed site compared to the less disturbed site, while *Callistoura* spp. prevalence showed no such pattern. Selective logging results in a suite of alterations that may increase infection risk and susceptibility to certain parasite infections in resident populations [174]. For example, studies on howler monkeys have reported higher GI parasite diversity and abundance in primates inhabiting degraded areas compared to those in less disturbed areas [242]. The depletion of the GI microbiota in degraded environments may explain these patterns. However, other studies show only minimal effects of disturbance on patterns of intestinal parasite infection [243]. As mentioned above, our logged forest site has been regenerating over decades, and it seems that lemurs have been able to adapt to differences in food availability and forest structural

differences accordingly [198]. As eggs of *Lemuricola* spp. are deposited in the perianal region of their host [206], body contact and grooming behaviour may be important factors in explaining the prevalence of this nematode within a population. Interaction rates and local lemur densities may be increased and home ranges more restricted in the more intensely logged forest, which has been shown to increase parasite infection risks [167, 244]. This may explain the higher *Lemuricola* spp. prevalence we found in these forests.

Several other studies observed a relationship between microbiota and GI parasites [159, 183-186, 189]. We found a small, but significant correlation of microbial composition with prevalence of *Callistoura* spp. In addition, the lemur population in Ankarafantsika had a significantly lower infection prevalence of *Callistoura* spp. compared to lemur populations in other areas and at the same time, this population showed the highest microbiota richness. Despite statistical significance of the correlations, interpretation of these correlations should be done with care. On one hand, GI parasites can have a direct influence on intestinal microbiota by damaging the host's intestinal epithelium, extracting nutrients in the GI tract [193], secreting antimicrobial products or inducing an inflammatory response [245]. On the other hand, observed correlations could not provide direct evidence for these mechanistic aspects. The microbiota is a dynamic ecosystem that has been shown to be affected by a broad range of environmental factors, however, the effect of factors with smaller relative weight is often masked by individual-specific factors like diet and genetic background [246], which could incorrectly reflect the true importance of such minor factors, particularly in wildlife studies where individual variation cannot easily be controlled.

Several studies found that the presence of some nematode species was linked to high microbiota diversity, with potential beneficial consequences for host health [182, 192, 194, 245, 247]. It is assumed that the immune system is regulated by the GI microbiota, but also that GI nematodes can alter the bacterial composition and structure, thereby creating conditions that can facilitate nematode infestations [247]. Although it has been shown that some

parasites change environmental conditions prevailing in the intestine, and thus affecting also microbial habitats, the exact relations between parasites and the microbiota remain unclear [248]. Most parasite species, and directly transmitted parasites in particular, co-evolve in association with only a few host species and adapt to the host gut environment and diet, resulting in host-driven diversification [249] that allowed to speculate about microbe-parasites evolutionary crosstalk. Understanding underlying mechanisms is critical for improving our knowledge on parasite–microbe interactions in wild primate populations. This can become achievable with a larger longitudinal sampling effort, genetic identification of the nematodes with molecular methods, and if possible *in vitro* and *in vivo* model experiments.

In conclusion, this study investigated the impact of seasonality and past logging on host-associated parasite infections, faecal bacterial communities, and correlative patterns between these GI inhabitants in geographically separated *Eulemur* populations. Our results show that seasonal differences and past logging events significantly contributed to explaining the observed temporal variations in parasite infections and microbial diversity. The variation in microbiota composition at the genus level showed a significant correlation with the presence of parasites, suggesting a relationship between GI parasites and microbiota composition under natural conditions. The factors that influence microbiota composition and presence of parasites may in turn affect host nutrition, behaviour, and health. These findings likely apply to other wild mammal communities as well. We believe it is important to consider the potential role of microbiome-parasite associations on the hosts' GI stability, health, and survival.

Acknowledgements

For their advice in the field, the authors want to thank Dr. P. Wright, Dr. E. Larney, and John E. Cadle, and for their logistical support we want to thank the Centre Valbio staff (CVB), the Institute for the Conservation of Tropical Environments (ICTE) and the Madagascar Institute pour la Conservation des Ecosystèmes Tropicaux (MICET), especially B. Andriamihaja, who facilitated arrangements on necessary permits. Thanks go to the Malagasy government and Madagascar National Parks (MNP), Centre de Formation Professionelle Forestière (CFPF), and Mitsinjo Association for their formal permission. For their fieldwork assistance and indispensable knowledge, we want to thank all field technicians and guides that helped as at the various national park, as well as the following MSc students: Anna Luijten, Jurian van Duursen, Nathalie Houtman, Jeroen Schütt, Jeroen Bentinck, Hilde de Laat, Elke van Gils, Freek IJdema, Maartje Papegaaij, and Rakotosolofo Andrianadrasanarivelo. We thank Harm Ploeger, Rolf Nijsse, and Herman Cremers (Utrecht University) for their aid with parasite identification and supervising students in the laboratory. Furthermore, we would like to thank J. Ramiro Garcia for assistance with statistical analysis. We would also like to thank the anonymous reviewers for their helpful and constructive comments, as well as Dr. R. Mundry (Max Planck Institute of Evolutionary Anthropology) for making available his R script for stability analyses in GLMM. This work is part of the PhD thesis of IdW and AU and was supported by an NWO/ALW and PE&RC grant (Grant nr. 1208, project nr. 5120873001). AU was supported by a PhD fellowship from the European Union Erasmus-Mundus programme (EM Action 2- Partnership IAMONET-20132520). Other necessary funds to conduct the fieldwork of this study were provided by: the KNAW Academy Ecology Fund, the Dr. J.L. Dobberke Foundation, the Treub Foundation, the Dutch Fund for Research on Nature Conservation, the Dutch Royal Botanical Society, and the Dutch Royal Zoological Society.

Supporting Materials

Table S1 Lemur study sites in Madagascar.

Study site	GPS coordinates (S, E)	Area (km^2)	Annual rainfall (mm)	Mean temperature (annual range, °C)	Altitude (range, m)	Lemur species	Samples (N)
Ranomafana NP	21.27, 47.33	435	3 000	11 – 25	500 – 1 500	*E. rufifrons*	48
						E. rubriventer	68
Nosy Be	13.33, 47.25	252	2 250	15 – 35	0 – 430	*E. macaco*	18
Nosy Komba	13.47, 48.35	25	2 250	15 – 35	0 – 620	*E. macaco*	23
Nosy Tanikely	13.47, 48.23	0.3	2 250	15 – 35	0 – 47	*E. fulvus*	17
Andasibe NP, Mitsinjo	18.92, 48.42	155	1 680	10 – 27	900 – 1 060	*E. fulvus*	43
Ankarafantsika NP	16.25, 46.80	1 350	1 300	17 – 28	80 – 330	*E. fulvus*	50
Kirindy	20.07, 44.67	722	767	19 – 31	20 – 90	*E. rufifrons*	40
Zombitse NP	22.87, 44.68	200	740	14 – 30	485 – 825	*E. rufifrons*	31

Table S2. *Callistoura* and *Lemuricola* spp. prevalence (0-1) in lemur populations at different areas across Madagascar.

	No nematodes	***Lemuricola* spp.**	***Callistoura* spp.**	**Both *Callistoura* and *Lemuricola* spp.**	**Total (N)**
Andasibe NP	0.37	0.02	0.56	0.05	43
Ankarafantsika NP	0.70	0.06	0.20	0.04	50
Kirindy Forest	0.05	0.05	0.73	0.16	37
Nosy Be	0.22	0.11	0.61	0.06	18
Nosy Komba	0.13	0.13	0.48	0.26	23
Nosy Tanikely	0.29	0.06	0.53	0.12	17
Ranomafana NP	0.25	0.03	0.62	0.10	116
Zombitse	0.07	0.03	0.77	0.13	31
Total	**0.29**	**0.05**	**0.56**	**0.10**	**335**

Table S3. Statistical modelling of *Callistoura* and *Lemuricola* prevalence

Seasonality ***Callistoura***

Model	Specification	Fixed part	Random part	LogLik	model df	Null hypothesis H_0	LRT	P-value	Distribution TS
SC1	Full model	ic+Sp+L+S+LxS	Site+Grp	-107.799	9				
SC2	Null model	ic		-137.852	1	no effect Sp,L,S,LxS, no variation sites, grps	60.106	<0.0001*	$\approx \chi^2_8$
SC3	Fixed only	ic+Sp+L+S+LxS		-119.533	7	no variation sites, groups	23.469	<0.0001*	$\approx \chi^2_2$
SC4	Random only		Site+Grp	-112.107	3	no effect Sp, L,S, LxS	8.616	0.196	χ^2_6
SC5	Fixed+Site	ic+Sp+L+S+LxS	Site	-113.374	8	no variation groups	11.150	0.0004	$0.5\chi^2_0 + 0.5\chi^2_1$
SC6	Fixed+Group	ic+Sp+L+S+LxS	Group	-109.851	8	no variation sites	4.104	0.021	$0.5\chi^2_0 + 0.5\chi^2_1$
SC7	Additive+Random	ic+Sp+L+S	Site+Grp	-108.285	8	no effect LxS	0.971	0.32	χ^2_1

Seasonality ***Lemuricola***

Model	Specification	Fixed part	Random part	LogLik	df	Null hypothesis H_0	LRT	P-value	Distribution TS
SL1	Full model	ic+Sp+L+S+LxS	Site+Grp	-74.997	9				
SL2	Null model	ic		-84.818	1	no effect Sp,S,L,SxL, no variation sites, grps	19.642	0.012*	$\approx \chi^2_8$
SL3	Fixed only	ic+Sp+L+S+LxS		-75.695	7	no variation sites, grps	1.396	0.50*	$\approx \chi^2_2$
SL4	Random only		Site+Grp	-82.184	3	no effect Sp, S, L, LxS	14.373	0.026	χ^2_6
SL5	Sp+random	ic+Sp	Site+Grp	-79.202	6	no effect S, L,LxS	8.410	0.038	χ^2_3
SL6	Additive+random	ic+Sp+L+S	Site+Grp	-75.017	8	no effect LxS	0.040	0.84	χ^2_1
SL7	Sp+season+ia+random	ic+Sp+S+LxS	Site+Grp	-77.309	8	no effect L	4.623	0.032	χ^2_1
SL8	Sp+loc+ia+random	ic+Sp+L+LxS	Site+Grp	-76.592	8	no effect S	3.189	0.074	χ^2_1

Disturbance ***Callistoura***

Model	Specification	Fixed part		LogLik	df	Null hypothesis H_0	LRT	P-value	Distribution TS
DC1	Full model	ic+Sp+D+S+DxS		-24.735	5				
DC2	Null model	ic		-27.835	1	no effect Sp,D,S,DxS	6.199	0.185	χ^2_4
DC3	Sp+season+ia	ic+Sp+S+DxS		-26.803	4	no effect D	4.136	0.042	χ^2_1

Disturbance ***Lemuricola***

Model	Specification	Fixed part		LogLik	df	Null hypothesis H_0	LRT	P-value	Distribution TS
DL1	Full model	ic+Sp+D+S		-19.176	4				
DL2	Null model	ic		-32.129	1	no effect Sp,D,S	25.905	<0.0001	χ^2_3
DL3	Sp+S	ic+Sp+S		-28.548	3	no effect D	18.438	<0.0001	χ^2_1
DL4	Sp+D	ic+Sp+D		-20.237	3	no effect S	2.122	0.15	χ^2_1

Chapter 4

Variation of caecal microbial composition among performance-tested Italian Landrace, Italian Large White, and Italian Duroc pig breeds demonstrate the effect of host genetic background on microbiota diversity

Alexander Umanets*, Diana Luise*, Luca Fontanesi, Paolo Bosi, Maurizio Gallo, Hauke Smidt, Paolo Trevisi

* Contributed equally

Abstract

Intestinal microbiota plays a crucial role in intestinal health, nutrient absorption and overall well-being of productive animals. Understanding the variation in microbiota composition across different domestic breeds of pigs could facilitate better health and nutrition management. We investigated caecal microbiota composition of Italian Landrace (ILA), Italian Duroc (IDU) and Italian Large White (ILW) that were kept in controlled conditions of a breeding facility. We found clear differences in microbial composition of ILA pigs in comparison with IDU and ILW animals, whereas caecal microbiota of the latter two breeds was very similar. Microbiota of ILA pigs showed significantly higher alpha diversity, as well as differences in overall composition and microbial network topology. We identified potential biomarker taxa that had a significantly higher relative abundance in pigs of the ILA breed group and were the strongest discriminant genera when ILA animals were compared with all other groups. These included genus-level taxa that are generally accepted as intestinal health indicators such as *Ruminococcaceae* UCG-005*, Lachnospiraceae* XPB1014*, Oscillospira* and *Faecalibacterium*.

Introduction

The terms gut or intestinal microbiota refer to a consortium of microorganisms dwelling along the gastro-intestinal tract. It is generally considered that gut microbiota is an important element of animal and human health homeostasis. A large number of studies have already shown that this "microbial organ" influences various aspects of human and animal health and behaviour. For example, interest in this field was fuelled by observed correlations between microbial composition and several pathologies, such as obesity [250], type 1 and 2 diabetes [251, 252], and inflammatory bowel disease [253]. Furthermore, a causal role of microbiota in intestinal and systemic health was supported by the efficiency of faecal transplantation for treatment of *Clostridioides difficile* infection [254] and the discovery of the molecular mechanisms behind induction of intestinal immunity by Segmented Filamentous Bacteria [34]. Research interests focusing on gut microbiota have recently exploded providing new opportunities and novel applications in human medicine and in the animal production sector [255]. However, gut microbiota of domestic and wild animals has received comparatively less attention than that of humans and rodent models, even for addressing basic concepts of gut microbial ecology and its potential role to improve animal health and to develop more sustainable feeding practices [256].

Among all livestock species, the impact of microbiota and intestinal health studies is particularly relevant in pigs. On one hand, being a monogastric species with diet and physiology similar to that of humans, this livestock species can be a unique animal model [257]. On the other hand, as the pig is one of the major sources of animal proteins worldwide (FAO 2014 http://www.fao.org/ag/againfo/themes/en/meat/background.html), modulation of intestinal microbiota can potentially improve the feed efficiency leading to more sustainable pig farming systems [258-261].

Gut microbiota studies in pigs have been mainly dedicated to the modulation of microbial composition and applications to decrease the use of

antibiotics through dietary supplementation with pre- and probiotic ingredients [262-264]. In contrast, studies focusing on the impact of the genetic background on microbial composition and resulting differences in fermentation capacity, immunomodulation properties and resilience to pathogens have been limited.

The relevance of host genetic background in shaping microbiota is still a controversial scientific topic. Studies in humans have reported that the genetic makeup of an individual is a prominent factor affecting its intestinal microbial community [57, 58, 137]. In contrast, other studies have shown that the effect of the host genetic background was negligible in comparison to other environmental factors or animal age [59, 265, 266]. In laboratory mice, however, genetic differences between populations have been shown to play a major role in the modulation of gut microbiota composition [267-269]. Few studies have investigated the influence of candidate genes, associated with intestinal diseases, on the microbiota profile in pigs [270, 271]. Other studies have investigated the influence of the host genetic background on microbial composition by comparing gut microbiota of different pig breeds kept in controlled environments [272-275]. However, most of these studies have a number of limitations, including: 1) a small sample size, 2) the inclusion of animals reared in different conditions, 3) the approximation of the gut microbiota composition from faecal samples, while the cecum which represents the part of intestinal tract where the microbial fermentation processes are taking place, can provide more accurate information [276].

In this study, we report on an extended analysis of cecum microbial composition in a large number of pigs belonging to three different breeds, i.e. Italian Landrace (ILA), Italian Duroc (IDU) and Italian Large White (ILW), performance-tested in a breeding facility and balanced for environmental factors. In this study we provide a comprehensive overview of diversity and structure of caecal microbiota in different breeds, identify breed specific microbial biomarkers and explore the possibility to identify the breed based on microbial composition.

Material and Methods

Ethical approval

All animals used in this study were kept according to the Italian and European legislation for pig production. All procedures described here followed the European Council Directive 98/58/EC concerning the protection of animals kept for farming purposes and the Council Directive 2008/120/EC laying down minimum standards for the protection of pigs.

Animal transport was performed according to Council Regulation (EC) No 1/2005 on the protection of animals during transport and related operations. Slaughter was performed at a commercial abattoir following Council Regulation (EC) 1099/2009 on the protection of animals at the time of killing and under the control of the Veterinary Service of the Italian Ministry of Health, as indicated in the Regulation (EU) 2017/625 of the European Parliament and of the Council on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products.

Animals and sampling

A total of 248 pigs, belonging to three different breeds [85 Italian Duroc (IDU), 71 Italian Landrace (ILA), and 92 Italian Large White (ILW) pigs] were included in this study. These animals were from the national selection sib testing programme managed by the National Pig Breeders Association (ANAS).

All pigs were registered to their breed herd books and were performance tested at the genetic station of ANAS (Emilia Romagna, Italy). Pigs were selected according to sex (IDU: 41 females and 44 castrated males; ILA: 38 females and 33 castrated males; ILW: 44 females and 48 castrated males) and balanced as much as possible for the date of slaughtering. A total of 19 slaughtering dates were followed throughout the year and balanced for season (winter: 8 slaughtering

dates including: 37 IDU, 14 ILA and 37 ILW; spring: 6 slaughtering dates including: 21 IDU, 23 ILA and 21 ILW; autumn: 6 slaughtering dates including 27 IDU, 34 ILA, 34 ILW). The slaughters performed during summer were not included in the trial in order to avoid differences due to the warm season.

Furthermore, pigs were chosen as representative for the Estimated Breeding Value (EBV) for average daily gain (ADG) and feed convention ratio (FCR) within breed (ADG: 54.87 ± 9.53 for IDU; 67.25 ± 21.24 for ILA and 53.28 ± 11.04 for ILW; FCR: -0.23 ± 0.06 for IDU; -0.24 ± 0.09 for ILA and -0.22 ± 0.05 for ILW). ADG was calculated by weighing animals every two weeks from the beginning to the end of the performance test of the animals (i.e. calculated from about 30 to 155 ± 5 kg of live weight). Feeding level was quasi ad libitum, i.e. 60% of the animals were able to ingest the entire supplied ration [277]. FCR was obtained from feed intake recorded daily and body weight measured bimonthly. Estimated breeding values for these traits were calculated using a Multiple-Trait BLUP Animal Model as previously described [277, 278].

At the end of the trial, the animals were moved to a commercial slaughterhouse and sacrificed according to standard procedures. The intestine was removed and the cecum was isolated. Cecum content was then collected in a sterile tube, immediately frozen on dry ice and stored at -80 °C until lyophilisation.

Each frozen cecum sample (approximately 32 g) was then freeze-dried using an Edwards Pirani 1001 freeze-dry system (BOC Edwards, USA). Each sample was weighed before and after freeze-drying, and the percent mass loss was calculated for each sample.

DNA extraction and library preparation

DNA was extracted from freeze-dried cecum content using a repeated bead beating protocol and the Maxwell® 16 Research Instrument (Promega, Madison, USA) for automated DNA extraction in combination with a custom RNA/DNA

extraction kit. In short, 0.1-0.15 g of a sample was added into a sterile screw-cap tube containing 0.5 g of 0.1 mm zirconia beads and five 2.5 mm glass beads. Then, 700 µL of S.T.A.R. buffer (Roche Molecular Systems, Boston, USA) was added, and tightly closed tubes were treated in a Precellys 24 beadbeater (Bertin Technologies, Nantes, France) for 3x1 min at 5.5x1000 movements per min. Bead beating treated samples were incubated at 95 ˚C for 15 min in a heating block (Vartemp 56, Labnet International, Edison, NJ, USA) at 100 rpm, then centrifuged at 4 ˚C for 5 min at 13,000 rpm. Two hundred fifty µL of supernatant was collected and stored on ice, and another 300 µL of fresh S.T.A.R. buffer was added in the same tube. Bead beating, incubation, centrifugation, and supernatant collection steps were repeated according the protocols described above, and supernatants from both rounds were pooled. Collected supernatants were used for DNA extraction by using the Maxwell automated system (Promega), following manufacturer's instructions. DNA was eluted in nuclease free water, and DNA quality was assessed by spectrophotometric analysis with a Nanodrop instrument (Nanodrop Technologies, Wilmington, USA).

The V4 region of the 16S rRNA gene was selected as phylogenetic marker. Primers (515f: 5′-GTGCCAGCMGCCGCGGTAA; and 806r: 5′-GGACTACHVGGGTWTCTAAT) recommended by the Earth Microbiome Project (http://www.earthmicrobiome.org/protocols-and-standards/16s/) with attached barcodes were used for the amplification. PCR reactions were performed in triplicate for each sample using a LabCycler Gradient (SensoQuest, Germany) thermal cycler programmed for initial denaturation at 98 ^{0}C for 30 s and 25 cycles of denaturation at 98 ^{0}C for 10 s, annealing at 56 ^{0}C for 10 s and extension at 72 ^{0}C for 10 s, followed by a final extension of 7 min at 72 ^{0}C. Each PCR reaction contained 10 µL of 5x HF buffer, 1µL dNTPs (10 µM), 1U of Phusion Hot start II DNA polymerase (Thermo Fisher Scientific; 2U/ µL,), 29.5 µl of nuclease free water, 2 µL each of forward and reverse primers (10 µM), and 100 ng of DNA template in 5 µL of nuclease free water.

Triplicate PCR products were pooled, and samples were purified individually using magnetic beads (MagBio, Switzerland) according to the

HighPrep protocol with adaptation for 2 mL tubes. Concentrations of purified PCR products were determined by Qubit dsDNA BR Assay Kit (Life Technologies, USA) following the manufacturer's protocol and sequenced using an Illumina HiSeq 2500 device at GATC Biotech (Konstanz, Germany; now part of Eurofins Genomics Germany GmbH).

Data analysis

Initial analysis of raw 16S rRNA gene sequencing data was performed using NG-Tax [47]. Sequences were separated into sample-specific bins based on barcodes, after initial filtering of paired-end libraries to contain only read pairs with perfectly matching barcodes. Operational taxonomic units (OTU) were defined using an open reference approach, and taxonomy was assigned using a SILVA 16S rRNA gene reference database v 123 [279].

Downstream data formatting and statistical analysis were carried out in R statistical and programming environment (v. 3.5.0) [217]. OTUs that were encountered in less than three samples were removed from the dataset and subsequent analyses. OTU count was normalized by rarefaction at 33211 reads as implemented in the "phyloseq" package [210] (function *rarefy_even_depth*).

Alpha diversity analyses

Several richness and evenness indices were calculated per sample. Phylogenetic diversity was calculated using the "picante" package [280], and number of observed species, Shannon, inverted Simpson and Chao1 diversity estimators were calculated as implemented in the "phyloseq" package. To test for statistical differences in alpha diversity between groups of animals we used Pairwise Test for Multiple Comparisons of Mean Rank Sums (Posthoc Kruskal and Wallis Nemenyi-Tests) with Chi-squared distances for p-value estimation as implemented in the "PMCMR" package [213]. Confidence intervals at 95% level (95% CI) and pseudo-median were calculated using the *wilcox.test* function (package "stats").

Beta-diversity analyses

To explore variation of microbial community composition between groups of investigated samples we used Principal Coordinate Analysis (PCoA) using Bray-Curtis, Jaccard, and Weighted and Unweighted UniFrac distances. PCoA results were visualised on two-dimensional scatter plots, where *x* and *y* axes captured the largest attainable variation. Distance calculation, ordination and visualisation were performed as implemented in the "phyloseq" package. We used the function *adonis* from the "vegan" package [212], with 9999 permutations, to investigate significance of microbiota variations in relation to independent variables (breed, sex, slaughter), calculate coefficients of determination (r^2), and significance of independent variable interactions.

Kruskal and Wallis test was used to evaluate differences in relative abundance of individual taxa between breed-groups. Prior to the analysis, OTUs were summarized at genus level (*tax_glom* function, "phyloseq"), and rarefied read counts were transformed into relative abundance (percentage). If a genus was not assigned a specific name in the reference database ("g__"), then the name of the corresponding higher taxonomic level was used with identifier affront (for example "f__" – family level), and a count number at the end, if several distinct genera within one family could not be assigned. Genera encountered in less than 30 samples (approximately 50% from the smallest group – ILA) were removed from the analysis, to avoid false positive results due to sparsity. After data preparation, we compared differences in relative abundance between breed groups using Kruskal and Wallis test (*kruskal.test* function, "stat"), and taxa with FDR adjusted p-values (q-value) less or equal 0.05 were selected for the next step. Selected genera were tested using Posthoc Kruskal and Wallis Nemenyi-Tests between groups differences as described before. Mean relative abundances of a given taxon and confidence interval were calculated using the "ci.mean" function as implemented in the "Publish" package [281].

Microbial network analysis

For the construction of microbial ecological association networks we used Sparse Inverse Covariance Estimation for Ecological Association Inference as implemented in the "SpiecEasi" package [282]. Neighbourhood selection ("MB" option for neighbourhood selection in SpiecEasi package) was used as a method for graphical model inference, the minimum sparsity/lambda parameter (*lambda.min.ratio*) was set at 1e-2, and the number of subsampling repetitions of StARS (Stability Approach to Regularization Selection) was set at 999. The resulting SpiecEasi object was converted into an "igraph" object (function *adj2igraph*) and plotted using the default graphic R function (*plot*). Prior to plotting and further analysis, all unconnected nodes were removed from the network. Obtained networks were subjected to community analysis (identification of node agglomerates within a network based on their connectivity) and estimation of modularity using *cluster_optimal* and *modularity* functions from the "igraph" package [283]. The Kendall correlation method as implemented in the "stat" package was used for analysis of between breed-group correlations based on the number of connections per node (genus). For every node in a breed, a specific network global hub score was calculated using the *hub.score* function from the "igraph" package.

Identification of breed-associated microbial biomarkers

Finally, we explored the possibility to classify a sample to a breed-group and to identify breed specific biomarkers by a Random Forest (RF) machine learning algorithm. Classification was performed using the "randomForest" package with 5000 classification trees per analysis masked under the *rfPermute* function from the "*rfPermute"* package [284]. We extracted data regarding the contribution of taxa to RF classification accuracy, further referred to as Mean Decrease Accuracy (MDA), from the trained model using function *importance* from the "randomForest" package [285]. Statistical significance of taxon contribution to RF classification was estimated by the *rfPemute* function and based on 999 permutations. The Area Under the Receiver Operating Characteristic Curve (AUC) was calculated for each RF model using function

prediction from the "ROCR" package [286] on the object resulting from the *predict* function (argument -"prob", package – "stat"). All obtained results were plotted using "ggplot2" package [287]. We used RF in a pair-wise fashion comparing only two breed-groups at a time. Also, every breed-group was compared with a group of samples consisting of randomly drawn samples from the other two breed-groups.

To have a visual representation of microbial composition at genus level across all samples, we visualized log transformed rarefied counts of reads plus one as a heat map using "ComplexHeatmap" package [288]. Clustering of rows and columns of a heat map was performed as implemented in this package with default settings.

Results

Two hundred forty-eight samples passed quality control and selection based on availability of metadata. After quality control, chimera removal, taxonomy assignment and initial OTU filtering, a total of 54,475,978 reads remained in the dataset. Reads were unevenly distributed among the samples, ranging from 33,211 to 909,441 reads and a median read count of 191,066 reads per sample. In total 713 OTUs were present in the complete dataset.

We started with assessing variations in microbiota richness and evenness between breed groups. We observed significantly ($p < 0.05$) higher alpha diversity in ILA animals regardless of the used metrics, and no differences between IDU and ILW breed-groups (Fig. 1).

When assessing beta-diversity based on different pairwise dissimilarity metrics, no clear visual separation was observed with PCoA (Supplementary Fig. S1). In contrast, PERMANOVA (*adonis* function) revealed a significant difference in overall microbial composition between breed-groups ($p < 0.003$). The coefficient of determination (r^2) varied from 0.028 to 0.040 depending on the used distance metric (0.037 for Bray-Curtis; 0.028 for Jaccard; 0.028 for Weighted UniFrac;

0.040 for Unweighted UniFrac). We could not find any significant interactions between any pairs of employed explanatory variables, or combinations of them. It is important to note that Slaughter date significantly contributed to explaining the observed variation in microbial composition ($p < 0.0001$) with highest coefficient of determination (r^2 from 0.12 to 0.17), however, as mentioned above no significant interactions were found between Slaughter and Breed.

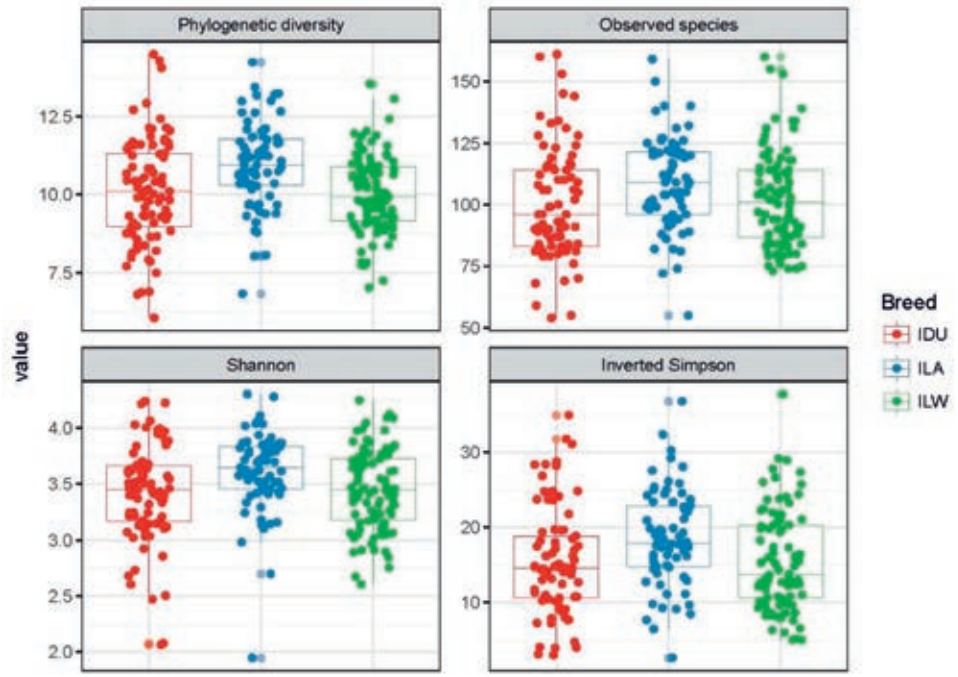

Figure 1. Comparison of microbial alpha diversity between breed-groups, as measured at OTU level. Samples are grouped and coloured by the corresponding breed (see legend, IDU – Italian Duroc, ILA - Italian Landrace, ILW - Italian Large White). Microbial alpha diversity metrics as specified in the head of plots. Each point represents the diversity score of an individual sampled animal. Between group variations were tested by Posthoc Kruskal and Wallis Nemenyi-Tests. ILA breed group has significantly higher ($p < 0.05$) diversity compared with ILW and IDU breed-groups.

Table 1. Genera that differ in relative abundance between breed-groups. Significance of differences was assessed by Posthoc Kruskal and Wallis Nemenyi-Tests. Mean decrease accuracy coefficient determined by RF; number of connections in network constructed by SpeacEsae. Green fill of q-value cells indicates outcomes with q-value lower than 0.05; means and confidence intervals highlighted in bold reflect significantly higher abundance when a breed group was compared with another; intensity of the red colour in the Network section corresponds to the number of connections (darker colour – more connections). Genera are arranged in blocks by results of relative abundance comparison and within blocks by decreasing relative abundance.

	Relative Abundance						Pairwise comparison (q - value)			Mean decrease accuracy						Network		
	IDU		ILA		ILW		IDU vs ILA	IDU vs ILW	ILW vs ILA	ILA vs IDU	ILA vs ILW	ILA vs IDU/ ILW	IDU vs ILA/ ILW	ILW vs IDU/ ILA	IDU vs ILW	Connections		
	mean	95% CI	mean	95% CI	mean	95% CI										IDU	ILA	ILW
g__Ruminococcaceae_UCG-005	1.18	0.94-1.42	**1.97**	**1.71-2.22**	1.22	1.04-1.4	0	0.877	0	25.7	29.6	21.9	17.5	11		2	8	5
g__Lachnospiraceae_XPB1014_group	0.34	0.21-0.46	**0.92**	**0.69-1.15**	0.37	0.26-0.48	0	0.858	0	21.8	16.2	11.2		7		7	5	10
f__Bacteroidales_S24-7_group	0.35	0.21-0.48	**0.86**	**0.59-1.13**	0.45	0.33-0.56	0	0.152	0.007	18.1			11.8		6.2	1	1	0
g__Parabacteroides	0.24	0.13-0.35	**0.58**	**0.32-0.84**	0.25	0.17-0.34	0	0.77	0	20.1	8.9	11.7	11.5			6	6	4
c__WCHB1-41	0.22	0.14-0.29	**0.58**	**0.38-0.79**	0.13	0.07-0.19	0.004	0.116	0	6	17.7	10.1		14.4	13.2	3	2	4
g__Akkermansia	0.19	0.01-0.36	**0.52**	**0.17-0.88**	0.08	0-0.15	0.003	0.883	0		11.9	6.5		5.1		5	4	1
g__Leeia	0.2	0.13-0.27	**0.52**	**0.36-0.69**	0.2	0.05-0.36	0.01	0.152	0	7	18.1	13.7		11.9		0	3	0
g__Oscillospira	0.09	0.06-0.13	**0.25**	**0.2-0.31**	0.07	0.04-0.11	0	0.649	0	17	26.6	21.7		19.5		0	4	2
f__Ruminococcaceae	0.1	0.06-0.13	**0.22**	**0.16-0.29**	0.07	0.05-0.09	0.003	0.683	0		7.1			5.2		1	1	3
g__Christensenellaceae_R-7_group	0.05	0.02-0.08	**0.21**	**0.08-0.34**	0.09	0.04-0.13	0.005	0.852	0.02	11.4			8.1		6.3	1	4	4
g__Faecalibacterium	0.48	0.38-0.57	**0.17**	**0.11-0.23**	0.48	0.38-0.58	0	0.935	0	26.9	24.5	25.2	5.9	11.4		7	5	9
g__Ruminococcaceae_UCG-008	0.08	0.05-0.11	**0.17**	**0.13-0.21**	0.08	0.05-0.1	0	0.99	0	20.5	14.1	11.1	6.3	11		1	0	0
g__Ruminococcaceae_UCG-002	0.02	0.01-0.03	**0.08**	**0.05-0.11**	0.04	0.02-0.06	0.001	0.409	0.037							1	2	1
g__Quinella	0.02	0-0.03	**0.05**	**0.03-0.07**	0.01	0-0.03	0.004	0.492	0							3	1	2
g__Oscillibacter	0	0-0.01	**0.05**	**0.03-0.07**	0.01	0-0.02	0	0.579	0.001							1	0	0
g__Turicibacter	**18.07**	**15.28-20.86**	13.17	10.64-15.69	**18**	**15.64-20.36**	0.041	0.936	0.014	13.1	11.7	12.9		6.2		3	2	2

g__Prevotella_9	**15.58**	**13.61-17.55**	10.52	8.89-12.15	**14.31**	**12.71-15.9**	0.001	0.841	0.007	7.7	9.1	6.3		5.1		8	7	10
g__Anaerovibrio	**1.92**	**1.6-2.25**	1.23	1.02-1.43	**1.71**	**1.48-1.94**	0.001	0.789	0.007	13.2	9.3		5.3			4	3	5
g__Alloprevotella	**1.53**	**1.3-1.75**	1.18	0.89-1.47	**1.6**	**1.36-1.84**	0.025	0.995	0.017					9.1		1	2	2
g__Prevotella_2	**1.28**	**1.07-1.49**	0.9	0.44-1.37	**1.21**	**1.02-1.4**	0	0.918	0.001	12.6	10.3	11.6				5	5	6
g__Blautia	**0.88**	**0.74-1.02**	0.58	0.44-0.71	**0.88**	**0.75-1**	0.003	0.976	0.001	6.9	10.8	7				5	10	6
g__Subdoligranulum	**0.55**	**0.47-0.63**	0.32	0.24-0.41	**0.49**	**0.41-0.56**	0	0.603	0.004	16.8	9.4	9				4	5	4
f__Prevotellaceae-2	**0.64**	**0.51-0.77**	0.31	0.22-0.4	**0.61**	**0.48-0.73**	0	0.771	0.003							6	6	5
g__Sarcina	**0.42**	**0.24-0.6**	0.19	0.06-0.33	**0.72**	**0.36-1.08**	0.028	0.965	0.012					6.4		2	1	1
g__Intestinibacter	**0.16**	**0.12-0.2**	0.08	0.05-0.12	**0.16**	**0.12-0.2**	0.034	0.987	0.02							2	2	0
g__Holdemanella	**0.13**	**0.09-0.17**	0.03	0.01-0.05	**0.1**	**0.07-0.13**	0	0.739	0.005	12.4	9.5	5.7				4	5	5
g__Fusicatenibacter	**0.06**	**0.03-0.08**	0.01	0-0.02	**0.07**	**0.04-0.09**	0.008	0.9	0.002	6	7.9	7.1		5.7		5	4	2
g__Clostridium_sensu_stricto_1	10.51	8.42-12.61	**14.37**	**12.11-16.63**	12.81	10.77-14.85	0.019	0.255	0.438	7.4	5.6		11.8	7.2		5	5	5
f__Prevotellaceae-1	0.34	0.22-0.45	**0.63**	**0.4-0.86**	0.39	0.22-0.57	0.042	0.982	0.058							1	2	1
g__Prevotellaceae_UCG-003	0.44	0.32-0.56	**0.61**	**0.47-0.75**	0.41	0.32-0.49	0.014	0.7	0.096	7.4					6.3	1	0	0
f__Bacteroidales_RF16_group	0.17	0.11-0.24	**0.36**	**0.25-0.46**	0.19	0.13-0.25	0.007	0.64	0.073	5.5						7	4	3
g__Lachnospiraceae_UCG-007	0.13	0.08-0.18	**0.25**	**0.19-0.32**	0.16	0.12-0.2	0.002	0.308	0.097	5.2	6.8	5.9				1	4	2
g__Bacteroides	0.06	0.01-0.1	**0.11**	**0.06-0.17**	0.04	0.02-0.06	0.018	0.826	0.071							3	2	3
g__Peptococcus	0.02	0-0.03	**0.07**	**0.04-0.1**	0.03	0.02-0.05	0.003	0.343	0.122	7.2		5.3						
g__Dialister	**0.06**	**0.04-0.09**	0.02	0-0.03	0.05	0.03-0.08	0.012	0.735	0.073							5	5	6
g__Lachnospiraceae_NK4A136_group	0.57	0.5-0.64	0.74	0.61-0.88	**0.82**	**0.72-0.92**	0.243	0.001	0.18		8.7		11.4	8.3	21.8	2	0	3
g__Lactobacillus	8.93	7-10.86	**11.35**	**9.03-13.67**	7.8	5.82-9.79	0.211	0.447	0.011		7.9	11.4				0	0	2
g__Streptococcus	4.67	3.68-5.65	**5.76**	**4.5-7.01**	3.62	2.84-4.41	0.492	0.165	0.01		5.9	6.6		9.3				
g__Treponema_2	0.95	0.66-1.24	**0.93**	**0.7-1.16**	0.61	0.41-0.8	0.213	0.464	0.012	9.8	5.9				6	6	6	7
g__Corynebacterium_1	0.32	0.03-0.61	**0.29**	**0.14-0.44**	0.1	0.03-0.18	0.075	0.514	0.003			5.3				1	3	1
f__Enterobacteriaceae	0.21	0.14-0.29	**0.26**	**0.15-0.37**	0.14	0.09-0.19	0.323	0.324	0.013		5.3					1	1	1
g__Rikenellaceae_RC9_gut_group	0.15	0.1-0.2	**0.21**	**0.16-0.27**	0.11	0.07-0.14	0.085	0.811	0.016					7		3	2	1
g__Lachnospiraceae_ND3007_group	0.11	0.08-0.14	0.08	0.05-0.1	**0.14**	**0.11-0.17**	0.243	0.239	0.004		6.9			6.3		1	3	3
g__Coprococcus_2	0.04	0.02-0.05	0.02	0-0.03	**0.06**	**0.04-0.08**	0.284	0.401	0.015							2	1	1

In total, 129 genus-level taxa were identified in the complete dataset, with 82 of them being considered as prevalent taxa (present in more than 30 samples). A total of 44 genera had significantly different relative abundances between the investigated breed-groups (Table 1). The genus-level taxon *Lachnospiraceae* NK4A136 was significantly more abundant in the ILW breed-group than in the IDU breed-group. Twenty-seven genera displayed significant differences in relative abundance when the ILA breed-group was compared either with IDU or ILW breed-groups (Table 1).

To explore the possibility to allocate a sample to one of the three investigated breeds based on microbial composition and to identify potential breed-specific microbial biomarkers, we employed a Random Forest based approach. Accuracy of classification was used as a proxy for between-group variations and was expressed via the Out Of Bag (OOB) error, i.e. the percentage of misclassified samples. Higher accuracy of classification and correspondingly lower OOB error correspond to more pronounced and more consistent differences between two breed-groups. A higher precision of classification of the sample was obtained by comparing the ILA sample group with ILW, IDU or mixed IDU+ILW sample groups using OTUs, or genera as discriminant features with OOB errors of 23.31 %, 25.64, 23.72% and 23.93%, 26.92%, 30.38%. respectively. A consistent decrease of classification accuracy was observed when higher taxonomic levels were used for discrimination, with few exceptions (Fig. 2). A low discrimination capacity (OOB above 32.5%) was observed for the comparison of the following breed-groups: IDU vs ILW, IDU vs ILA/ILW, and ILW vs IDU/ILA, regardless of the used taxonomic level. We then identified taxonomic features that contributed most strongly to the accuracy and robustness of the identification (Fig. 3). At OTU level, OTUs from genus-level taxa *Ruminococcaceae* UCG−005 and *Ruminococcaceae* UCG−008 were the top discriminant features for IDU vs ILA, IDU+ILW vs ILA, and ILA vs ILW comparison pairs (MDA 23.8 – 17.6), followed by OTUs belonging to genera *Faecalibacterium*, *Bacteroidales* S24−7 and *Prevotella* 2/57, which showed a considerable discriminant power (MDA 20.3 – 13.3). In the IDU vs ILW comparison, an OTU from *Lachnospiraceae* NK4A136 had the highest MDA score (i.e. 22.4) followed by *Roseburia* (MDA = 19.0), while

for IDU vs ILA+ILW, *Phascolarctobacterium* (MDA 19.0) and *Clostridium* sensu stricto 1 (MDA 15.0) had the highest scores.

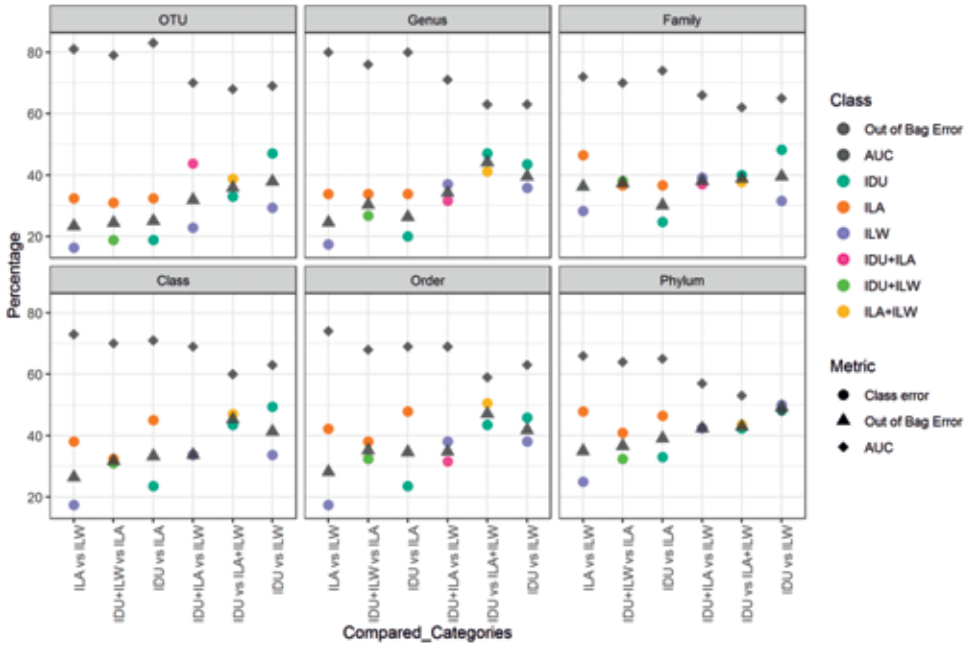

Figure 2. Scatter plot of Class errors (I, II, III), Out Of Bag Error and Area Under the Curve (AUC) in percentage extrapolated from Random Forest analysis. Metric types are encoded by different symbols, and the breed-groups by colour. Breed-groups are abbreviated as follows: IDU – Italian Duroc, ILW – Italian Large Wight, ILA – Italian Landrace, IDU+ILW – random subset of samples from IDU and ILW groups, ILA+ILW – random subset of samples from ILA and ILW groups, IDU+ILA - random subset of samples from IDU and ILA groups.

We further explored potentially conserved patterns in caecal microbial communities in the different breed-groups using network analysis. Large variations in microbial network topology were observed among breed-groups (Fig. 4A, 4B, 4C and Fig. S2). Only four connections were shared by all networks, fourteen connections were shared by IDU and ILW breed-groups, and nine and eight connections were shared by ILA and IDU and by ILA and ILW breed-groups, respectively (Fig. 4D).

Using data inferred from networks, we calculated the number of connections per node (genus) within a breed-group. The number of connections in

corresponding nodes was then compared between breed-groups. The highest correlation was observed between IDU and ILW breed-groups (tau = 0.45, $p < 0.0001$), whereas comparison between IDU and ILW groups with ILA resulted in lower correlation coefficients (tau = 0.33 and tau = 0.35, respectively, $p < 0.0001$).

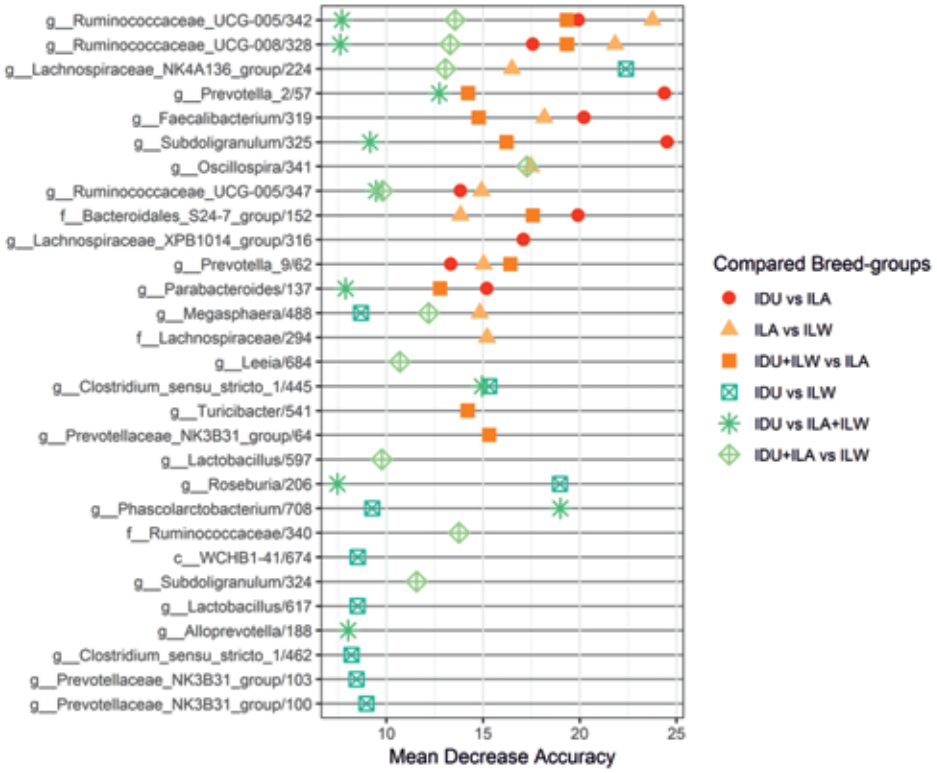

Figure 3. Mean Decrease Accuracy (MDA) of the Random Forest (RF) analysis when a given classification feature (microbial taxon) was removed from the respective RF model. Ten most influential OTUs with highest MDA values per comparison are displayed; OTU names are given as follows: lowest identified taxonomic level / assigned OTU number; shape and colour of symbols reflect different compared groups. Breed groups are abbreviated as follows: IDU – Italian Duroc, ILW – Italian Large Wight, ILA – Italian Landrace, IDU+ILW – random subset of samples from IDU and ILW groups, ILA+ILW – random subset of samples from ILA and ILW groups, IDU+ILA - random subset of samples from IDU and ILA groups.

We found a number of sub-communities within every breed-group network with differences in modularity, connectivity and composition (Fig. 4A, 4B, 4C). ILW had highest network modularity among investigated breed-groups (0.68), followed by IDU (0.67), and ILA (0.65) breed-groups. Nine sub-communities with more than three members were identified in the ILW breed-group, ten in the IDU breed-group, and eight in the ILA breed-group (Fig. 5). We found that sub-communities with highest modularity from IDU and ILW shared four common genera, IDU and ILA shared one, ILW and ILA shared two, and only one genus was shared in all three breed-groups (Table S1).

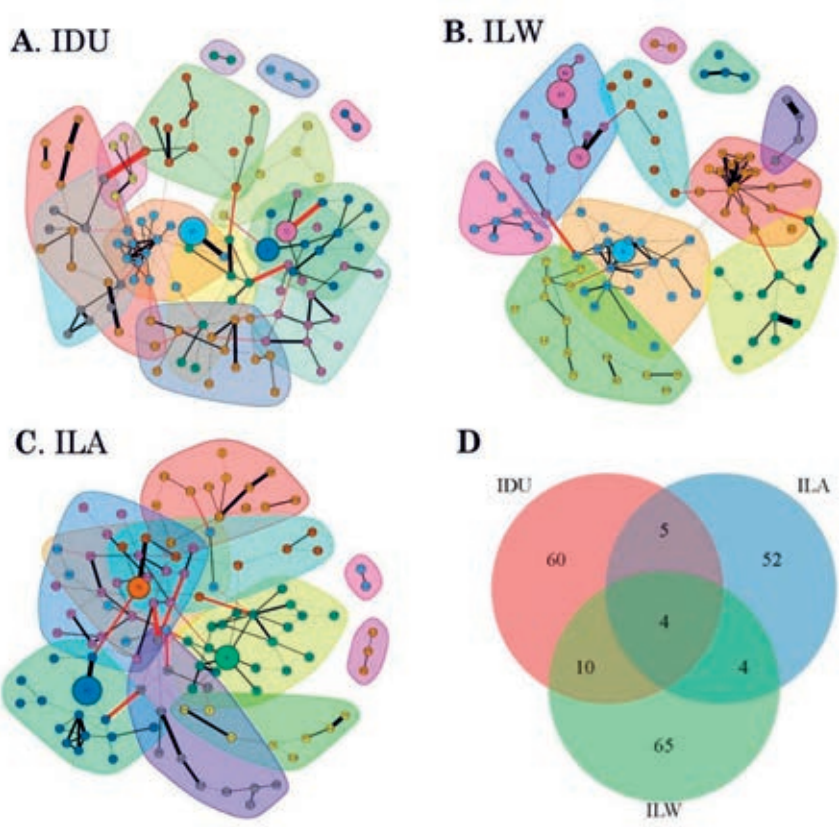

Figure 4. A, B, C) Sub-communities identified in microbial networks constructed using Sparse Inverse Covariance Estimation for Ecological Association Inference (SpiecEasi). Sub-communities are highlighted by differently coloured clouds. Colours of clouds do not correspond to taxonomic composition of sub-communities, and were assigned independently for each network. **D)** Venn diagram of node to node connections.

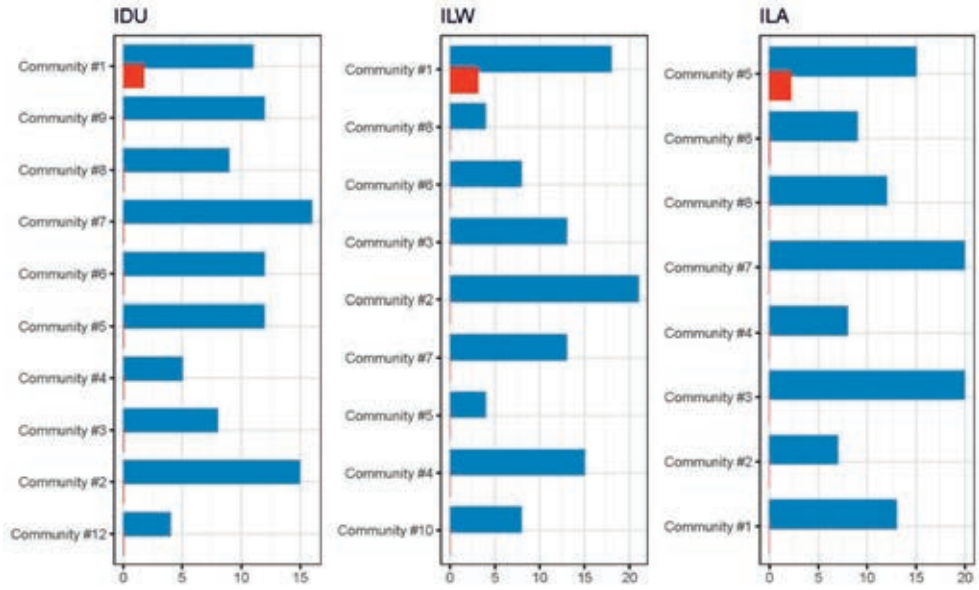

Figure 5. Bar plot of node number (blue) and sum of hub scores of all nodes (red) within a community. Hub scores were calculated for nodes based on complete breed specific networks. Sub-communities were ordered by modularity score from highest (on top) to lowest (on the bottom)

Discussion

In the current study, we set out to assess potential differences in caecal microbial composition among pigs belonging to three different breeds commonly reared in Italy. To this end, our results showed a clear difference in cecum microbial composition of the ILA breed-group in comparison with both ILW and IDU breed-groups. We observed higher microbial alpha diversity in animals belonging to the ILA breed-group with different diversity metrics. High diversity of an ecosystem has for a long time been associated with high stability and resilience to perturbations [289]. This general ecological rule has been also applied to microbial communities. A few studies have shown that low gut microbial diversity is associated with high risk of colorectal cancer [290], obesity [291], attention deficit hyperactivity disorder (ADHD) [292], necrotizing enterocolitis in infants [293] and recurrent *Clostridioides difficile* infection [294].

High diversity of gut microbiota has also been interpreted as a sign of intestinal health and high resilience to perturbations by changing environmental conditions [295]. In turn, large variations in microbial diversity unrelated with the health status were also observed between different mammalian species, due to differences in their diets [82]. The higher alpha diversity observed in our study in the ILA breed-group is difficult to explain. It could be speculatively correlated only with physiological features of ILA since the other environmental factors (rearing conditions) were controlled and standardized. The effect of the host genetic background on microbial diversity was previously observed in mouse lines [143] and in pig breeds [272-275]. Our finding that the ILA breed-group had a higher microbial diversity in the cecum compared to IDU and ILW is in line with the results reported by Pajarillo et al. (2014), who showed that Landrace pigs had the most diverse faecal bacterial community composition compared to Duroc and Yorkshire pigs. Future research needs to set out to identify what drives higher microbial diversity of Landrace pigs, if their microbiota is more stable and resilient to perturbations compared to that of other domestic pig breeds, and whether there are any functional consequences of the observed differences between breed-groups.

To assess the degree of dissimilarity in overall microbial composition within and between breed-groups, beta diversity analyses were performed. This method allowed to measure dissimilarities between samples and take overall microbiota composition into consideration to identify differences in community structure [296]. We could not find clear visual separation when samples were ordinated using PCoA based on different dissimilarity metrics, however, a PERMANOVA test of statistical significance and strength of grouping (*adonis*) showed that breed and slaughter date were the two main factors significantly affecting the microbial composition. Absence of a clear visual separation could be explained by relatively small between breed-group variation in comparison to within breed-group variation. We found that slaughter date had the strongest influence on microbiota composition followed by breed, whereas sex was not significant. The strong influence of slaughter date could be ascribed to potential differences in environmental conditions affecting the animals around the time of

slaughter. Indeed, even considering that animals were chosen in order to balance the slaughter date among breeds, and were raised on the same farm, under the same management conditions and were fed the same diet, environmental factors such as temperature or transport time and conditions could not clearly be distinguished and controlled. Breed was a significant factor that affected the microbial community structure. However, we observed a relatively low coefficient of determination. It is not surprizing taking into account that in general microbial composition has high variability between individuals and could be influenced by a large number of other factors [55].

We applied a network analysis to allow for a more complete evaluation of microbe-microbe relationships within breed-groups. This approach has previously proven to be useful in microbial ecology by providing an insight into microbial relationships based on co- and anti-occurrence of microbial features, and allowed to find possible modes of interactions and identify interconnected groups of microbial taxa (sub-communities) as well as central, or hub-specific taxa [35].

Despite the large differences in overall connectivity between breed-group networks, the IDU and ILW breed-groups had the largest numbers of conserved taxa to taxa connections in comparison with the ILA breed-group. A lower number of connected taxa and a lower modularity in the ILA network could be explained by higher inter-individual variation in microbial composition within this breed-group. It will be interesting to evaluate whether a larger genetic variation in the ILA breed as compared to the other two pig breeds could be one of the possible reasons of the observed higher inter-individual microbiota variations in this breed.

Random forest is one of the most popular machine learning algorithms for classification and regression in ecology due to its accuracy and robustness even if applied on complex data with nonlinear relationships [297]. RF is a “white box” machine learning approach which means that the contribution of a single feature to a given classification model can be evaluated. The opportunity to obtain information about the discriminative power of specific features (i.e. microbial

taxa) allowed to use RF for identification of biomarkers [298] from complex microbiota compositional datasets [299, 300].

RF classification allowed us to obtain an additional insight into between breed-group microbiota variation by assessing the accuracy of sample classification and by identifying microbial taxa with high discriminative power. Pairwise comparison of breed-groups proved to be the most informative mode of RF classifier application. Best classification was achieved when ILA was compared with ILW, IDU, or the merged ILW and IDU breed-groups. This finding is in line with observed differences between breed-groups in alpha and beta diversity. Decomposition of the OOB errors into class errors revealed that both ILW and IDU had lower classification errors when compared with the ILA breed-group. A possible explanation could be the lower variation of microbiota composition between samples within the ILW and IDU breed-groups, as a consequence of a more homogeneous genetic background. Kovacs et al. (2011) described that the host genetic background has a pronounced influence on mouse gut microbiota. Furthermore, human twins have more similar gut microbiota than two unrelated individuals within a population [57]. Furthermore, since feeding behaviour can vary according to the breed [301], this could have affected the intestinal transit and thus the microbial profile. Interestingly, fattening Italian pigs obtained from ANAS Large White sows and Duroc boars tended to present a higher feed digestibility than the more improved Danbred Duroc, both reared up to 130 kg slaughter weight [302]. This indirectly indicates that there can be genotype-dependent variation in the nutrient flux of fattening pigs, which may in turn be related to differences in gut microbiota composition and function.

The identification of microbial biomarkers related to productive traits, heath and robustness of animals is an attractive target for microbiota research in production animals [303]. Microbiome biomarkers showed diagnostic potential in identification of colorectal cancer [304, 305], multiple sclerosis [306], ulcerative colitis [307], obesity and inflammatory bowel disease [308].

In our study we did not look for health/disease state biomarkers, but rather taxa that could be representative of a given host population and that could

help to understand differences between them. We focused only on taxa that where identified by RF to have high MDA [309] and that showed significant differences in relative abundance between breed-groups (i.e. 44 genera). It should be noted that the precise functional role of taxa identified as biomarkers could not be assessed in the correlation study reported here.

Most of differentially abundant taxa were observed when ILA was compared with ILW or IDU. Only the genus *Lachnospiraceae* NK4A136 was significantly different in relative abundance when comparing ILW vs IDU breed-groups. Several studies reported modified levels of *Lachnospiraceae* NK4A136 in experimentally induced dysbiosis [310] and colitis in mice [311-313]. These studies showed negative correlation between *Lachnospiraceae* NK4A136 and *Akkermansia* that we also observed in our study. The function of *Lachnospiraceae* NK4A136 in the gut ecology of pigs is not completely clear, however, Chen et al. (2018) showed that this genus is one of a few genera that reacted to an addition of alpha-ketoglutarate to a low protein diet [314].

Ruminococcaceae UCG-005*, Lachnospiraceae* XPB1014*, Oscillospira,* and *Faecalibacterium* were significantly more abundant in animals of the ILA breed-group and were the strongest discriminant genera when the ILA group was compared with all other groups. All these genera are common members of a mammalian intestinal microbiome. *Ruminococcaceae* UCG-005 was indicated to contribute to human intestinal health [315] and to be a butyrate producer [316]. *Oscillospira* is generally acknowledged as a key player in maintenance of intestinal homeostasis [317], however, no representatives of this genus have been isolated in pure culture until now. *Faecalibacterium* is considered as beneficial microbe with prominent probiotic properties [318]. All above-mentioned taxa were shown to be common members of pig microbiota and have been correlated with changes in production of various metabolites in relation with diet [257], exposure to toxins [319], as well as the physiological and health status of the host [320-323].

Overall, we found clear differences in microbial composition, richness, and network topologies in ILA versus ILW and IDU breed-groups. Most of the

microbial biomarkers that allowed to differentiate between breed-groups have previously been implicated as intestinal health markers. These microbial markers, in general, had a higher relative abundance in the ILA breed-group. Other studies are needed to better understand the role of the host genetic background on microbiota diversity in pigs. Potential pig genes that could be relevant to explain this variability might be those involved in intestinal immunity and digestion regulation related processes.

Supporting information

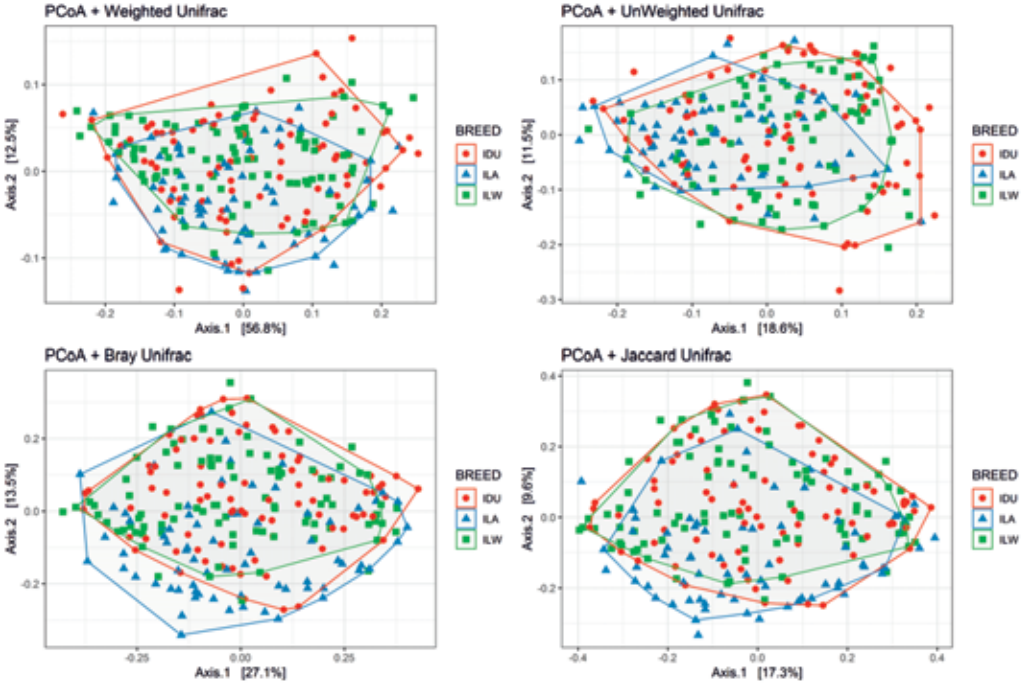

Figure S1. Ordination plots. Headers of plots indicate ordination method and used distance metric. Colour and shape of the symbols correspond to the different breed groups. Envelopes were drawn around 85% of the samples in the tightest group by breed identity.

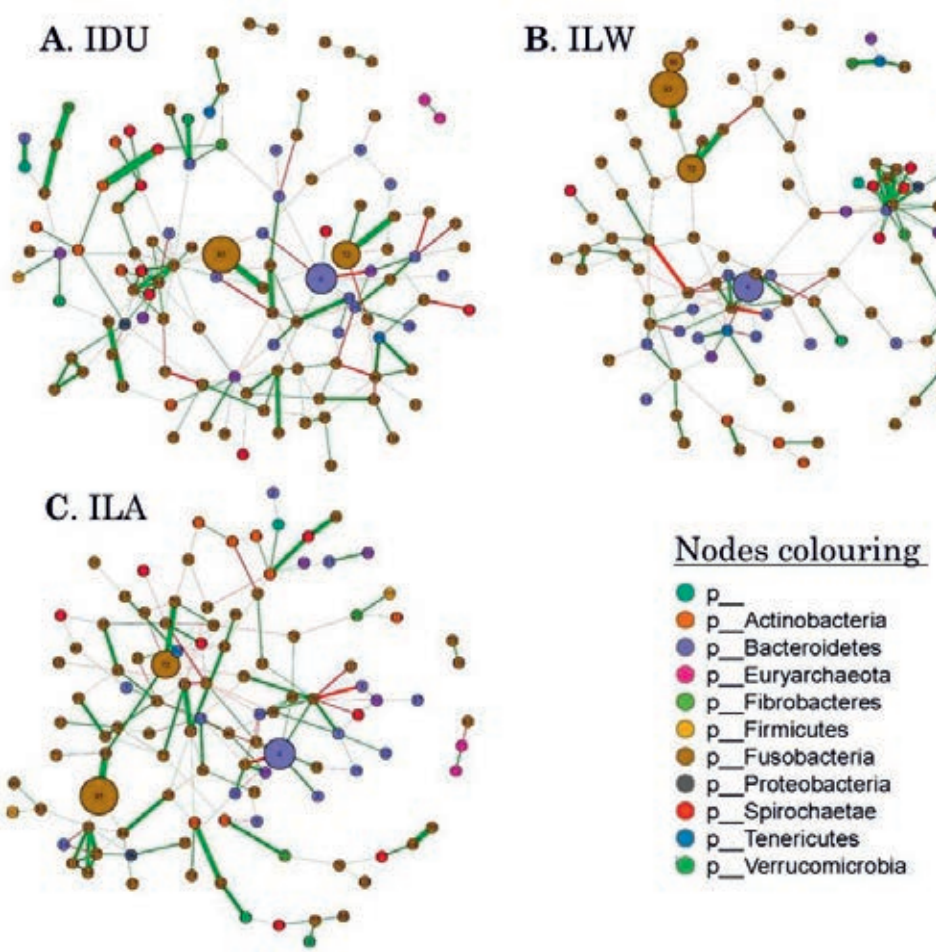

Figure S2 . Networks constructed with connections between genera identified by Sparse Inverse Covariance Estimation for Ecological Association Inference (SpiecEasi). Size of the nodes represents mean relative abundance of a genus in a breed-group dataset; genera with relative abundance ≤ 5% were displayed in the same size. Colours of the nodes correspond to different genus phylum. Positive interactions between nodes are coloured in green and negative interactions are coloured in red, with thickness representing the strength of the interactions (β value * 25). A) Italian Duroc B) Italian Large White C) Italian Landrace.

Table S1. Taxonomic composition of sub-communities with highest modularity score as identified by network analysis. **Comm. ID** stands for Community Identifier and corresponds to Figure 5 community identifiers. **Genus/Node ID** – The lowest identified taxonomic level is used as a genus name and in case of two or more unidentified genus-level groups were found within same higher taxonomic level, a numeric identifier was added; Node ID corresponds to nodes identifier in Figure 4 and Figure S2. Cells are colour coded as follows: red – present in every breed group; green - present in IDU and ILW breed groups; blue – present in IDU and ILA groups; purple – present in ILW and ILA.

IDU		ILW		ILA	
Genus/Node ID	**Comm. ID**	**Genus/Node ID**	**Comm. ID**	**Genus/Node ID**	**Comm. ID**
c__WCHB1-41.1/101	#1	*c__WCHB1-41.1/101*	#1	*f__Clostridiales_vadinBB60_group/118*	#5
f__Bacteroidales_S24-7_group/2	#1	*c__WCHB1-41.2/102*	#1	*f__Erysipelotrichaceae.1/84*	#5
f__Clostridiales_vadinBB60_group/118	#1	*f__Clostridiales_vadinBB60_group/118*	#1	*f__Fibrobacteraceae/103*	#5
f__Coriobacteriaceae.1/124	#1	*f__Peptococcaceae/66*	#1	*f__Peptostreptococcaceae/68*	#5
g__Anaerotruncus/62	#1	*f__Veillonellaceae/76*	#1	*g__[Eubacterium]_eligens_group/34*	#5
g__Family_XIII_AD3011_group/80	#1	*g__dgA-11_gut_group/13*	#1	*g__Clostridium_sensu_stricto_6/60*	#5
g__Fibrobacter/99	#1	*g__Epulopiscium/22*	#1	*g__Coprococcus_2/37*	#5
g__Methanobrevibacter/20	#1	*g__Family_XIII_AD3011_group/80*	#1	*g__Epulopiscium/22*	#5
g__Ruminococcaceae_UCG-014/83	#1	*g__Fibrobacter/99*	#1	*g__Fusobacterium/96*	#5
o__Mollicutes_RF9/123	#1	*g__Fusobacterium/96*	#1	*g__Lachnospira/35*	#5
p__/1	#1	*g__Helicobacter/104*	#1	*g__Lachnospiraceae_UCG-003/27*	#5
		g__Lachnospiraceae_NK3A20_group/44	#1	*g__Rikenellaceae_RC9_gut_group/14*	#5
		g__Methanosphaera/21	#1	*g__Ruminococcaceae_UCG-014/83*	#5
		g__Ruminobacter/120	#1	*g__Sarcina/71*	#5
		g__Ruminococcus_2/61	#1	*g__Turicibacter/85*	#5
		g__Succinivibrionaceae_UCG-001/106	#1		
		g__Succinivibrionaceae_UCG-002/97	#1		
		p__/1	#1		

Chapter 5

Differences in intestinal microbiota composition between Great White-Toothed Shrews and Wood Mice shapes differences in their resistome composition

Alexander Umanets, Fred de Boer, Pim van Hooft, Joost Lahr, Kevin D. Matson, Dik J. Mevius, Hauke Smidt, Nico van den Brink

Abstract

Increased resistance of bacteria to antibiotics holds a serious threat for human health and food security, and better understanding of forces that shape resistome composition is needed. We investigated the resistome of Greater White-Toothed Shrews (WTS) and Wood Mice (WM) from six different farm areas across The Netherlands. Our goals were to investigate differences in resistome composition between both animal species, and to assess correlations between gut microbiota and antibiotic resistance gene (ARG) composition. In total 121 faecal samples were collected from WTS (n = 71) and WM (n = 50) that were caught around six pig farms. For microbial composition profiling amplicon sequencing of the V4 region of the prokaryotic 16S rRNA gene was used on individual samples. Resistome composition was assessed using shotgun metagenomics of equimolar DNA pools per area and species (n = 12). We found that reads with ARG motifs accounted for 0.26%-0.66% (95%, CI[0.38, 0.5]) of the total number of reads. When samples from WTS and WM were compared we observed a nearly significant ($p = 0.07$) difference in relative abundance of the total number of reads with ARG motifs. Hierarchical clustering based on log transformed weighted ARG composition showed perfect grouping by animal species. Differences in ARG diversity between species groups was not significant, however, we found a high negative correlation between ARG diversity and abundance (kendal; tau -0.74; $p = 0.001$). To investigate correlations between gut microbiota composition and ARG composition we compared Principal Coordinate (PCoA) ordinations of both and found a high level of symmetry when using Procrustes rotation (0.83, $p < 0.0001$). In conclusion, this work for the first time provides an overview of the resistome of free living WTS and WM, as well as showed a clear correlation between gut microbiota and ARG composition.

Introduction

Resistance of bacteria to antibiotics is one of their fundamental survival mechanisms [324]. Along their evolutionary history, microorganisms in general did not develop in isolation, but rather in diverse and complex microbial communities, where they engage a broad spectrum of both positive and negative interactions. These range from mutualistic interactions to competition for resources and even amensalistic interactions, where a ruthless chemical warfare is the name of a game. Many microorganisms, including bacteria, produce bioactive compounds that may kill or inhibit growth of other microorganisms [325]. One of the most well-known examples of such antagonistic relationships, namely that between *Penicillium notatum* and various bacteria, was discovered by Alexander Fleming almost one hundred years ago, and a few decades later in the 1950s the first antibiotics were purified for medical use [326, 327].

It is difficult to imagine modern human medicine or agricultural practices without antibiotics, but their widespread use also resulted in the development and spread of resistance potential in a broad range of different microorganisms. It is estimated that worldwide antibiotic consumption in the human population is at ten doses per person (based on consumption in 2010) [328]. In countries where antibiotic use in agriculture is not limited by law such as the USA, up to 80%, or around 13 thousand tonnes of all antibiotics used are applied for livestock [329]. It was estimated that in 2010 around 63 thousand tonnes of antibiotics were used globally as veterinary drugs or feed additives for livestock [330]. Intensive use of antimicrobials in animal husbandry creates a selective pressure leading to survival of antibiotic resistant microorganisms. This comes with a plausible risk of antibiotic resistant zoonotic, or facultative pathogenic bacteria transmission from animals to humans, and enrichment of the microbial metagenome with mobile antibiotic resistant genes (ARG), which could horizontally transfer to other bacteria, including human pathogens [331].

Horizontal gene transfer (HGT) is a naturally occurring phenomenon, that provides the potential for exchange of genetic information between

microorganisms, even across taxonomically distinct species. For prokaryotes HGT is one of the main ways to acquire new properties and adaptation mechanisms including, among others, resistance to antibiotics [332]. Transfer of ARG was initially associated only with the mechanism of conjugation, whereby bacteria exchange DNA via direct contact with each other by pili or adhesins. However, a growing body of evidence showed that uptake of extracellular DNA from the environment (transformation), or introduction together with phage genetic material (transduction) could be also important routes [333]. Hence, it is easily conceivable that bacteria carrying ARGs on their genome or on mobile genetic elements such as plasmids can spread resistance to other members after entering a naive microbial community. In the environment, different factors may drive the spread of ARGs, including wildlife [334, 335]. In countries like The Netherlands, despite high standards of animal husbandry, small wild animals could be a vector of ARG transmission due to the difficulty to control their access to farmlands. Some studies have provided insights that small mammals may carry ARGs, although little is known on the mechanisms that underlie the observed differences in ARG prevalence and composition among species. To address this, we investigated in the present study the resistome, i.e. the diversity and abundance of ARGs, in Wood Mice (WM, *Apodemus sylvaticus*) and Greater White-Toothed Shrews (WTS, *Crocidura russula*) dwelling around farm areas. These species have different feeding habits, with WM being omnivorous and WTS being insectivorous, and they also differ in their spatial behaviour. Because of these differences, possible links between resistome and animal gut microbiota composition were evaluated. Small mammals have long been recognized as important vectors for spreading a number of zoonotic diseases [336], and it is generally accepted that rats were a driving force of several epidemics of bubonic plague [337] "Spill-over" of pathogens from domestic to wild animals is a well-known phenomenon that could not only endanger a population of wild animals but also create a stable infection-reservoir within the population [338]. The potential for ARG "spill-over", however, has not been investigated before. Here we focused on WTS and WM populations dwelling around pig production farms with known low use of antibiotic. In this way effects of a high use of antibiotics,

which may overwhelm the effects of species-specific traits on prevalence of ARG in the two species was prevented, although animals still can be exposed to antibiotics and ARGs in a farm environment.

We used culture independent methods to study microbiota and resistome composition. To this end, we applied 16S ribosomal RNA (rRNA) gene amplicon sequencing in combination with shotgun metagenomics to assess the presence of ARGs in the faecal microbiota of WM and WTS [339]. We used a tiered approach, whereby metagenomic sequencing was done for samples pooled within defined animal groups (by species and location; Table1), whereas microbiota composition was measured for all individuals separately. We used a number of bioinformatic tools that allowed to map the vast number of short reads obtained by shotgun metagenomics against various ARG databases, giving an overview of a microbial community's resistome, that could be coupled to microbiota profiles assessed by amplicon sequencing.

To our knowledge this is the first study focusing on ARG diversity and abundance in WTS and WM. We hypothesized that differences in microbiota composition related to their drastic differences in lifestyle and diet has a direct influence on ARG abundance and diversity.

Material and Methods

Sample collection

At six different farms in the southern part of the Netherlands faecal samples of small mammals were collected. All farms were pig rearing, all with low antibiotic use. Animals were caught at night in Longworth's life traps. Life traps were baited with peanut butter (to attract WM) and cat food (To attract WM and WTS). Early in the morning the traps were checked, and of all animals, species and sex was determined. No further biometry was taken, and the animals were released directly. Generally, animals defecated in the trap, and the faeces were collected (Table 1). At return to the lab the faeces were stored at -80 °C prior to analysis.

Table 1. Number of animal samples per location

Species	Location					
	1 (n)	*2 (n)*	*3 (n)*	*4 (n)*	*5 (n)*	*6 (n)*
Wood Mice	5	3	4	13	22	3
Greater White-Toothed Shrews	16	8	22	7	10	8

DNA extraction and sequence analysis

DNA was extracted using a repetitive bead beating protocol combined with an automated DNA extraction system (Maxwell® 16 Research Instrument, Promega, Madison, USA) with a custom RNA/DNA extraction kit. In short, three to five droppings from each shrew and three from each mouse were used for DNA extraction. Samples were processed as follows. Faecal droppings were added into a sterile screw-cap tube containing 0.5 g of 0.1 mm zirconia beads and five pieces of 2.5 mm glass beads. Then, 700 µL of S.T.A.R. buffer (Roche Molecular Systems, Boston, USA) was added and tubes were treated in a Precellys 24 beadbeater (Bertin Technologies, Nantes, France) for 3x1 min at 5.5x1000 movements. Next, samples were incubated at 95 ˚C for 15 min in a heating block (Vartemp 56, Labnet International, Edison, NJ, USA), and then centrifuged at 4 ˚C for 5 min at 13,000 rpm. Two hundred fifty µL of supernatant was collected after the first round of bead beating and stored on ice. Then another 300 µL of fresh S.T.A.R. buffer was added in the same tube. Bead beating, incubation, centrifugation, and supernatant collection steps were repeated, and 250 µL of supernatant from the second round was combined with the corresponding supernatant collected after the first round. Combined supernatants were used for DNA extraction by using the Maxwell automated system (Promega) following the manufacturer's instructions. DNA was eluted in nuclease free water, and DNA quality was assessed by spectrophotometric analysis with a Nanodrop instrument (Nanodrop Technologies, Wilmington, USA).

Primers (515f: 5′-GTGCCAGCMGCCGCGGTAA; and 806r: 5′-GGACTACHVGGGTWTCTAAT) recommended by the Earth Microbiome Project

(http://www.earthmicrobiome.org/protocols-and-standards/16s/) with attached barcodes were used for the amplification of the V4 region of the 16S rRNA gene. PCR amplifications were performed using a LabCycler Gradient (SensoQuest, Germany) thermal cycler programmed for initial denaturation at 98 ˚C for 30 s and 25 cycles of denaturation at 98 ˚C for 10 s, annealing at 56 ˚C for 10 s and extension at 72 ˚C for 10 s, followed by a final extension of 7 min at 72 ˚C. Each PCR reaction contained 10 µL of 5x HF buffer, 1µL dNTPs (10 µM), 1U of Phusion Hot start II DNA polymerase (Thermo Fisher Scientific; 2U/ µL,), 29.5 µl of nuclease free water, 2 µL of forward (10 µM) and 2 µL of reverse primers (10 µM), and 100 ng of DNA template in 5 µL of nuclease free water. To minimize amplification bias, each PCR reaction for each sample was performed in triplicate, and resulting PCR products of a sample were pooled. Pooled PCR products were purified individually using magnetic beads (MagBio, Switzerland, Lausanne) according to the HighPrep protocol with adaptation for 2 mL tubes. Concentration of purified PCR products was determined by Qubit dsDNA BR Assay Kit (Life Technologies, USA).

In total 121 samples pooled into three amplicon libraries were sequenced at GATC Biotech (Konstanz, Germany; now part of Eurofins Genomics Germany GmbH). Processing of raw reads and identification of unique sequences (amplified sequence variants, in the following referred to as operational taxonomic units, OTUs) was performed using NG-Tax [47], and taxonomy was assigned against the SILVA 16S rRNA gene reference database version 126 [119]. The resulting OTU count table together with taxonomic table and phylogenetic tree build from OTU representative sequences was imported in R 3.5.0 into a phyloseq object ("phyloseq") [210]. OTUs with less than three reads or presence in less than three samples were removed from the dataset.

For shotgun metagenomics we pooled DNA from individual animals by location and species into equimolar mixes (Table 1). DNA pools were sent to GATC Biotech for sequencing on the Illumina HiSeq 2500 platform. Nucleotides with a quality score below 20 were removed, and sequences with sequencing artefacts or a length below 70 bp were removed using "bbduk" tools

(http://jgi.doe.gov/data-and-tools/bb-tools/). Cleaned sequences were taxonomically assigned using the “Kaiju” tool with the recommended full prokaryote database [49]. ARG sequences were identified using the “DeepARG” tool and the recommended DeepARG database [48]. Resulting taxonomic and ARG abundance tables were imported into R 3.5.0. For more convenient data uploading and manipulation R packages “readr” [340], "dplyr" [341] and "reshape2" [342] were used.

Confidence intervals were calculated at alpha 0.05 using the "gmodels" [343] package, unless stated otherwise.

We estimated the diversity of ARGs within sample pools using the inverted Simpson diversity index and using raw count data as input (“vegan” [212], *diversity*). Correlation between ARG diversity and abundance of reads with ARG motifs (ARGM) was assessed by Kendall correlation analysis (“stat”, *cor.test*). Heatmaps were build using the “ComplexHeatmap” [288] package, and rows and/or columns were clustered using hierarchical clustering as implemented in the package.

Compositional correlations between reads with ARGM and taxonomic assignments were assessed by comparison of ordination results. Firstly, read count data for taxonomy and ARG tables was independently normalized using Hellinger transformation (*decostand*, “vegan”), and Principal Coordinate Ordinations (PCoA) based on Bray-Curtis dissimilarity scores were constructed (*capscale*, “vegan”). Ordinations were compared with each other by Procrustes Rotation Test with 9999 permutations (*protsest*, “vegan”), and Correlation in a Symmetric Procrustes Rotation (CSPR) with significance were reported.

Due to differences in data processing methodology between data obtained from shotgun metagenomics and amplicon sequencing, direct comparison between the datasets is not possible. To this end, we employed indirect comparison using Procrustes correlation of ordinations as described above. Prior to Procrustes correlation analysis, amplicon sequencing data from the same samples that were used for shotgun sequencing was combined with the exception

of samples with a read count below 5000 (sample 32 in pool S2). Read counts per sample were normalized by rarefaction at an even depth (20113 reads). Strength (R^2) and significance of sample grouping based on dissimilarity distances were estimated using the adonis statistical test (*adonis*, "vegan") with 9999 permutations.

Correlations between relative abundance of microbial genera and ARG classes were identified in corresponding sample sets using Kendall correlation (*cor.test*, "stat"). As input data for correlation tests we used relative abundance of ARG classes and microbial profiling data from 16S rRNA gene amplicon sequencing. The OTU abundance table was rarefied and pooled to match corresponding metagenomics sample pools as described above.

To find possible sample clusters we used the complete linkage method with manhattan dissimilarity distances on the log transformed genus count matrix (*dist*, *hclust*, "stat").

Results

In total 43±6 million reads were obtained by shotgun metagenomics sequencing per sample pool after removal of sequences with low quality (Figure S1). Around half of all reads (48.4-%-53.1%, 95% CI) could be assigned at least at the phylum level, however, if only reads containing ARGMs were considered, 75.3-83.5% (CI, 95%) were assigned at the genus level, while 95.6-97% (CI, 95%) could be assigned at phylum level.

The percentage of reads with ARGMs varied from 0.26% to 0.66% (95%, CI[0.38, 0.5]) of the total number of retained reads per sample (Fig. 1A). A nearly significant difference ($p = 0.07$) in relative abundance of reads with ARGMs could be observed when samples were grouped by animal species (Fig. 1C). ARG diversity didn't significantly differ between WM and WTS sample groups ($p = 0.27$), however, noticeable differences were observed with respect to the variation in diversity within species groups, with WM samples showing high consistence in

ARG diversity (5.8 – 6.1) and WTS samples having a wide spread (4.0 – 6.7) (Fig. 1B). In addition, we found a strong negative correlation (Kendall; tau -0.74; p = 0.001) between the relative abundance of reads with ARGMs and ARG diversity (data not shown).

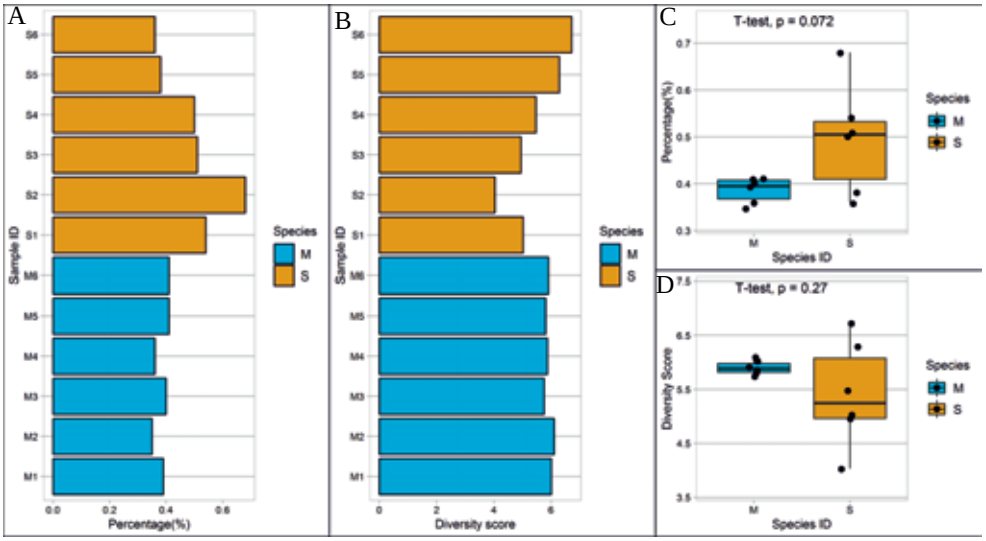

Figure 1. Relative abundance of reads with identified antibiotic resistance gene motifs (A, C) and their diversity based on the inverted Simpson index (B, D) by sample (A, B) and by species group (C, D). Dark gold colour and "M", Wood Mice; blue colour and "S", Great White-toothed Shrew.

Reads with ARGMs showed clear clustering by species (Fig. 2). Macrolide and multidrug ARGs were the most abundant classes across all samples, however, in WM samples, macrolide ARGM-containing reads were more abundant than those with multidrug ARGMs (95%, CI, [0.096, 0.158] vs. [0.061, 0.127]), whereas the opposite was observed in WTS samples (95% CI, [0.074, 0.116] vs. CI[0.094, 0.249]). Overall 13 out of 23 identified classes of ARGs showed statistically significant differences in relative abundance between species-groups (Fig. 3, Table S1).

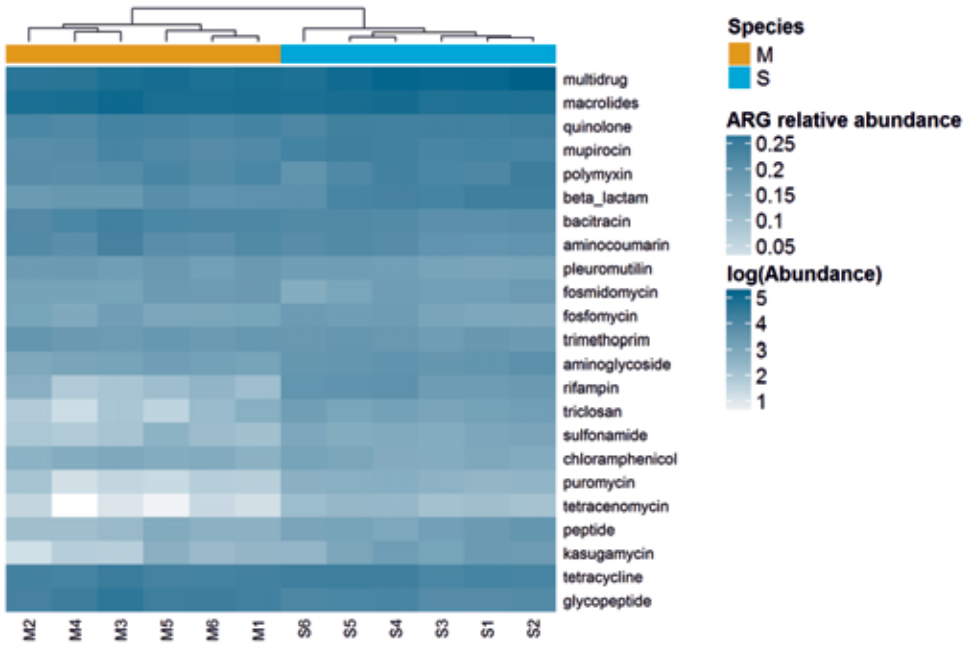

Figure 2. Heatmap based on log transformed absolute abundance of reads with ARGMs. Rows and columns are clustered using hierarchical clustering. Dark gold colour and "M", Wood Mice; blue colour and "S", Great White-toothed Shrew.

Using PCoA, a clear separation of sample pools by species groups could be observed based on shotgun metagenome-derived phylogenetic (Fig. 4A) and ARGM-containing read composition (Fig. 4B) as well as 16S rRNA gene-derived composition (Fig. 4C). Comparison of the ordinations based on metagenomic sequence data showed high (CSPR = 0.95) and significant ($p < 0.0001$) correlation in a symmetric Procrustes rotation (Procrustes correlation). Also, results of phylogenetic profiling based on shotgun metagenomics and 16S rRNA gene amplicon sequencing were comparable with Procrustes correlation (CSPR) of 0.89 ($p < 0.0001$). Expectedly, Procrustes correlation of 16S rRNA gene sequencing derived ordination with the ordination based on reads with ARGMs was lower (CSPR = 0.83, $p < 0.0001$) than correlation between ARGM abundance and phylogenetic composition derived from shotgun metagenomics (CSPR 0.95; $p < 0.0001$).

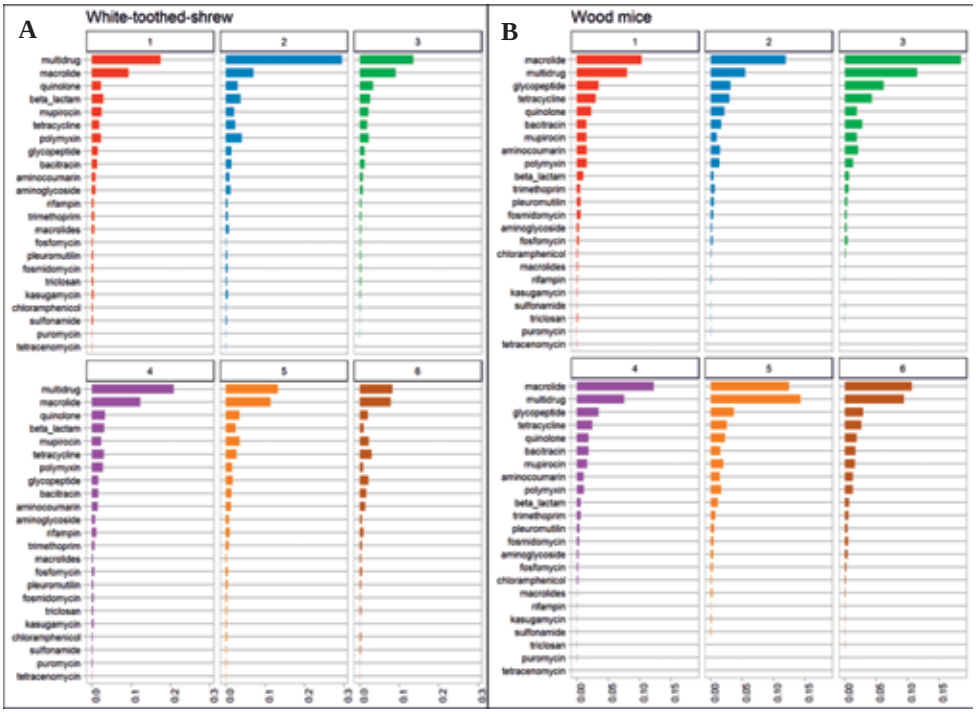

Figure 3. Relative abundance of ARGs by location in samples from White-toothed-shrews (A) and Wood Mice (B). Every graph corresponds to one of six locations where samples were collected as indicated above each plot.

When samples were ordinated individually based on 16S rRNA gene derived phylogenetic profiles (Fig. 4D) the clear grouping by animal species was retained (adonis; R^2=0.37 ; $p < 0.0001$), whereas grouping by sampling location was much less obvious, albeit also statistically significant (adonis; R^2=0.14 ; $p < 0.0001$).

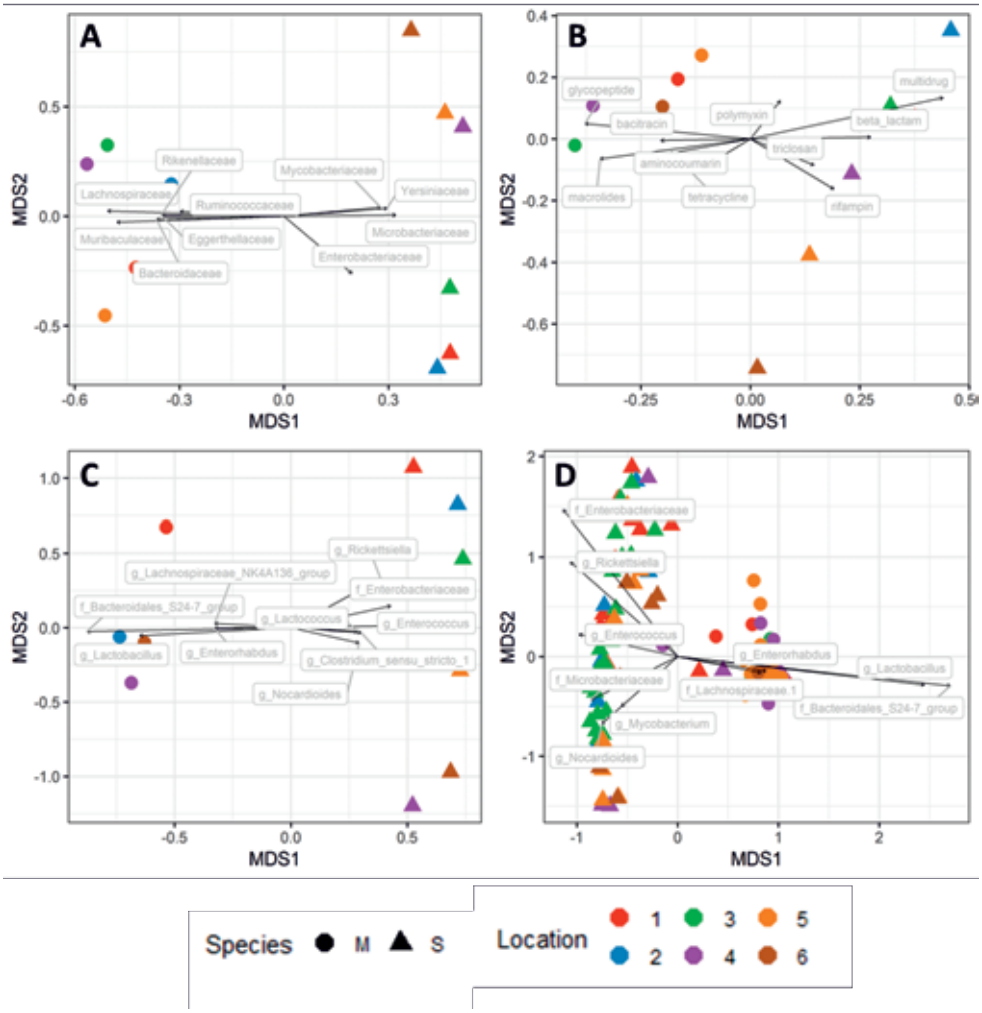

Figure 4. Principal coordinate plots based on Bray-Curtis distances. A) Phylogenetic composition at genus level derived from shotgun metagenomics; B) Hellinger transformed counts of reads with ARGMs; C) phylogenetic composition at genus level derived from 16S rRNA gene amplicon sequencing with individual samples collapsed by location and host species; D) phylogenetic composition at genus level derived from 16S rRNA gene amplicon sequencing by individual samples. The top ten influential features are depictured by arrows with their length corresponding to the relative contribution of a feature to the ordination. Animal species are depicted as follows: letter "M" and circle – Wood Mice; letter "S" and triangle - White-Toothed Shrew. Colours of points correspond to location number as shown in the plot legend.

As the next step we explored correlations between individual microbial genera and ARG classes. We correlated all available samples with disregard to species group and found a number of correlations when employing a dual threshold of i) FDR (False Discovery Rate) adjusted p-values at 0.1 and ii) Tau values only above 0.5 or below -0.5 (Fig. 5). The number of correlated microbial genera varied greatly per class of antibiotics, ranging from one to 16 significant correlations (Fig. 6) if both negative and positive correlations are considered. In general, ARG classes could be divided into three groups based on the number and nature of correlations with microbial genera. The first group comprised ARG classes without a clear tendency to either negative or positive correlations. This group included the ARG classes with the highest number of negative and positive correlations (peptide, n = 16; glycopeptide, n = 12; sulfonamide, n = 12 and several others triclosan, n = 8; pleuromutilin, n = 7; rifampin, n = 7, quinolone, n = 6). The second group included ARG classes with predominantly positive correlations (mupirocin, n = 10; aminoglycoside, n = 7; kasugamycin, n = 7; macrolides, n = 5; beta_lactam, n = 5), whereas a third group included ARG classes with mostly negative correlations (puromycin, n = 7; tetracenomycin, n = 7; bacitracin, n = 7, aminocoumarin, n = 5). With respect to microbial taxa we also observed a homogeneous distribution of positive and negative correlations, however, some genera (*Sanguibacter*, Unidentified genus (UG) within the *Intrasporangiaceae* and *Clostridium* sensu stricto 1) displayed only positive correlations, whereas others (UG within the *Bacteroidales* S24−7_group, UG within the *Lachnospiraceae* 1 and *Streptococcus*) showed more negative than positive correlations (Fig. 5).

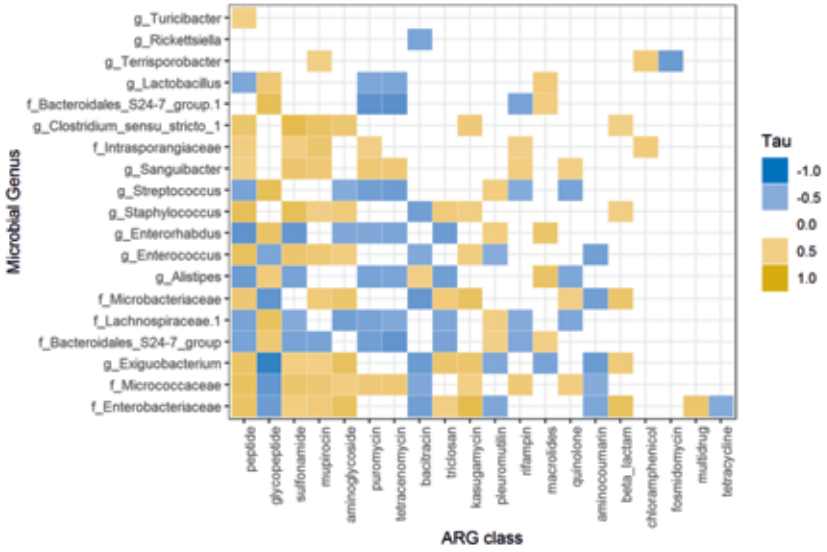

Figure 5. Heatmap of correlations between reads with ARGMs and relative abundance of microbial genera based on 16S rRNA gene profiling; only correlations below an FDR adjusted p-value of 0.1 and having a tau value > 0.5 or < -0.5 are displayed.

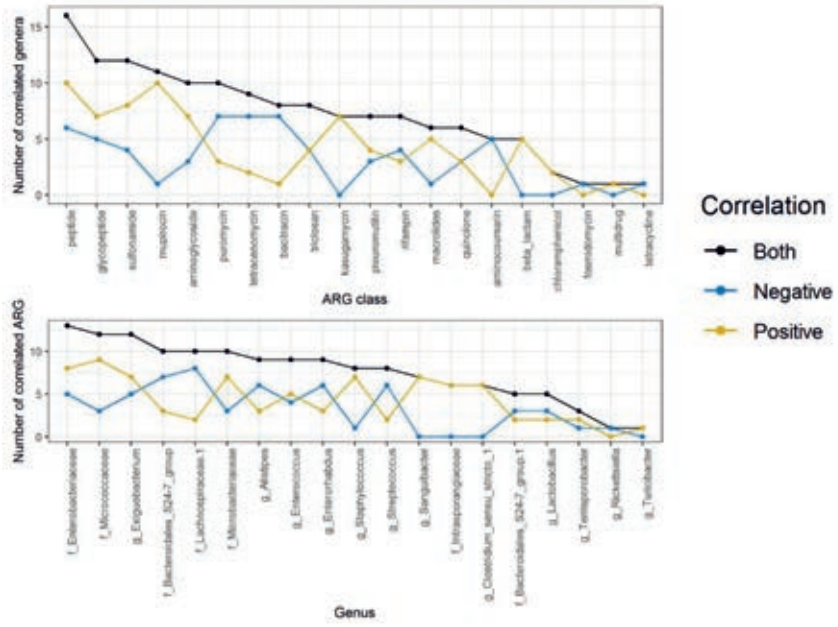

Figure 6. Number of total, negative and positive correlations per ARG class (upper figure) and microbial genus (lower figure).

Lastly we investigated the distribution among samples of microbial taxa correlated with ARG classes. We found clear differences in relative abundance between samples from WM and WTS (Fig. 7) and almost perfect clustering by host species that is reflecting clustering analysis when all genera were taken into account (Fig. 8).

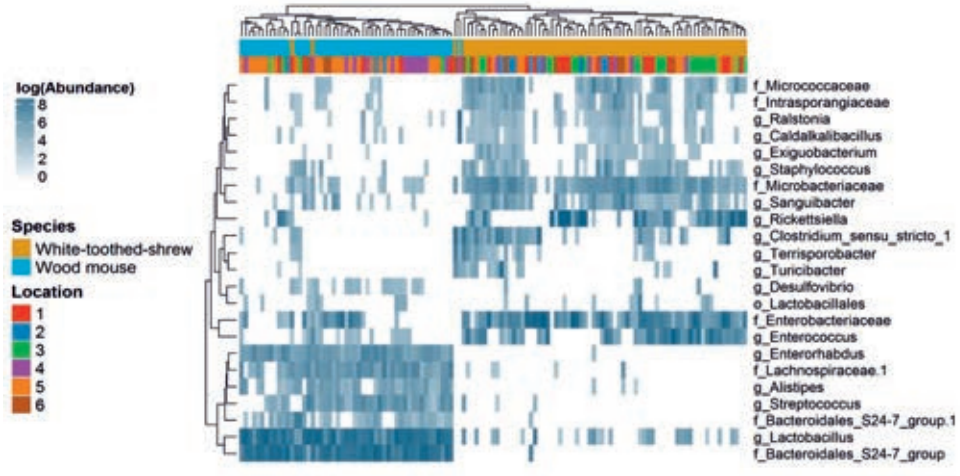

Figure 7. Heatmap based on relative abundance of genera that are considered to have correlations with ARGM reads across all samples.

Discussion

Diverse resistomes were observed in faecal samples from both WTS and WM. This was also expected, as antibiotic resistance is a naturally occurring phenomenon and can be expected in any complex microbial community such as the gut microbiota [344, 345]. It was shown before that rodents carry a number of bacteria resistant to a variety of antibiotics even in areas without recent animal production practice [346]. Nevertheless, only a selected number of studies found a correlation between anthropogenic influence and concentration of ARGs in the environment. For example, it has been shown that soil at pig production sites with extensive use of antimicrobial growth promoters was characterized by as much as a two- to three orders of magnitude higher abundance of ARGs in comparison with pristine forest soil [347].

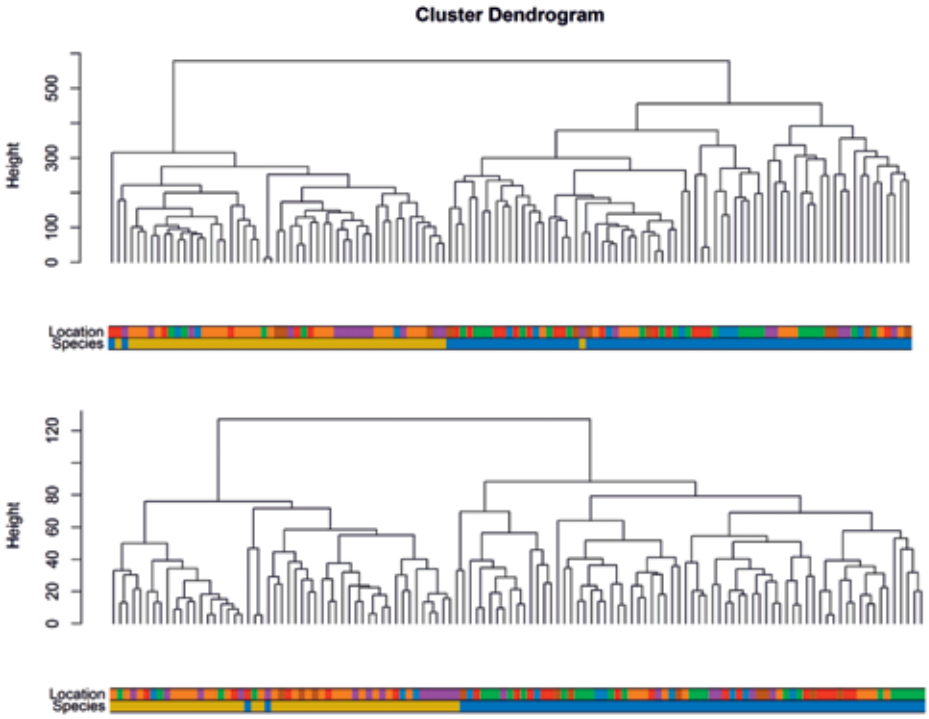

Figure 8. Hierarchical clustering of samples based on log transformed relative abundance of all microbial genera identified by 16S rRNA gene sequencing (upper figure) and only taxa correlated with ARG classes that passed the threshold as stated in figure A (lower figure).

In our study, half of all obtained reads could be taxonomically classified using direct classification with the Kaiju tool, which is comparable with results of the tool developers when it was applied to field (not simulated) data [49]. The Kaiju tool bases its taxonomic classification on microbial genomes available in the NCBI database, which could explain the high percentage of taxonomic assignment among reads with ARGMs detected by DeepARG. The DeepARG employs a neural network approach that provides a flexible and precise tool to mine metagenomic reads for presence of ARGMs, and uses training datasets that comprise information about ARGs from CARD, ARDB and UNIPROT databases, in which most of the presented genes are derived from previously described genomes that are also available in the NCBI database [48, 348-350].

We found that overall 0.26% to 0.66% of all metagenomic reads contained ARGMs. Abundance of ARGs within a microbial community largely depends on microbial community origin and structure. A metagenomics analysis of ARG abundance in various natural and anthropogenic environments showed several fold differences between them, with human and livestock faecal samples leading the chart [339], displaying up to 3.1 ARGs per 16S rRNA gene copy. To the best of our knowledge no previous metagenomics based study focused on ARG presence within the gut microbiota of wild living WTS and WM, and thus direct comparison with published data is not possible. Nevertheless, despite differences in approaches and used technology, the percentage of identified reads with ARGMs in our study is in line with data from other gut environments [351].

Multidrug and macrolide resistance genes were the most abundant classes of ARGs across all samples. Multidrug-resistance efflux pumps play various roles in bacterial life besides antibiotic resistance [352], and it is hard to judge if high relative abundance of multidrug ARGMs could be associated with environmental factors in addition to natural differences in microbiota composition. Macrolides, however, have been shown to have a prominent and specific effect on human gut microbiota composition [353]. Furthermore, high abundance of macrolide-associated ARGs were observed in microbial communities of water sediments and were shown to be directly correlated with anthropogenic pollution [354]. In all samples from WTS, relative abundance of multidrug ARGMs was highest followed by that of macrolide-associated ARGMs. This order was reversed in faecal samples of WM, with the exception of sample M5 (Fig. 5). In addition to ARGMs related to macrolides and multidrug resistance, also other classes of ARGMs differed in relative abundance between WM an WTS. In total we found 13 differentially abundant ARGM groups, and due to the fact that samples from different species were collected at the same locations, our data suggest that these ARGMs have a predominantly species-specific pattern in addition to an environmentally driven abundance pattern.

Overall we observed a very high correlation between phylogenetic and resistome composition of faecal microbiota. It has been observed before that the resistome has a tight correlation with microbial composition in biogas reactors [355]. However, meta-analysis of metagenomes from the human gut and soil showed that presence of ARGs was largely defined by the type of samples, with ß-lactamase and tetracycline resistance genes being the strongest discriminators between soil and human samples, rather than phylogenetic composition of the corresponding microbial communities [356]. The host species investigated in our study have drastic differences in ecology, physiology and dietary habits. WTS are actively hunting insectivorous animals with a diet consisting of various invertebrates and e.g. earthworms [357], whereas WM are typical rodents feeding on various seeds, green plants with occasional addition of animal pray [358]. In addition, WTS have an accelerated metabolism which leads to a very short digesta retention time of 2.15 ± 0.202 h, which is almost five times shorter than in rodents of comparable size [359]. Such behavioural and physiological differences result in a different level of environmental pressure for gut microbial communities, resulting in compositional and functional variations in gut microbiota of investigated mammalian species. It is difficult to disentangle the relative contribution of the differences in intestinal environment and phylogenetic composition of microbiota as factors in shaping resistomes of WTS and WM, because of the direct correlation between intestinal environment and microbiota composition.

Unfortunately due to the small number of metagenome samples collected within one location (n = 2) we could not statistically investigate the influence of occupancy on resistome composition beyond descriptive investigation of the plots. However, the observed correlation between microbiota and resistome composition allowed us to perform an indirect comparison between sampling location based 16S rRNA gene sequencing of individual samples, and observed significant differences in microbiota composition between animals from different locations. It has been shown previously that location and type of habitat have a significant effect on gut microbiota composition of wild mice [360]. The overall tight correlation between microbiota composition and resistome allowed us to

speculate about existence of location specific ARG signatures alongside with compositional differences in investigated microbial communities. Which factors in particular drive such location-specific distinctions in microbiota composition, however, is not clear.

Using Kendall correlation analysis we found a number of positive as well as negative associations between specific microbial taxa and ARG classes. Positive correlations could be explained by a higher number of ARGs within certain microbial taxa, whereas negative correlations could correspond to competition and/or amensalistic interactions between microbial taxa. Overall, peptide ARGMs had the highest number of correlated taxa (n = 16), which could be explained by a broad definition of this group, as well as presence of natural antimicrobial peptides produced by mammalian cells which affect a broad range of bacteria [361, 362]. Second by number of correlations was the glycopeptide ARG class, which includes antibiotics such as vancomycin and teicoplanin. However, currently these antibiotics are not widely used due to their high toxicity to the host [363]. Use of vancomycin in humans has been associated with a decrease of microbial diversity, and decrease of Gram-positive bacterial relative abundance [364].

In turn, microbial taxa also showed a different number of associations with ARGMs, ranging from 13 for an UG within the family of Enterobacteriaceae to a single correlation for *Turicibacter* and *Rickettsiella*, that could suggest that different microbial taxa have different capacities to carry ARGs. When we performed clustering analysis of only ARGM correlated taxa we observed a practically identical picture to clustering based on all detected microbial taxa with clear separation by host species (Fig. 8). This finding could be interpreted as another conformation of host specific gut microbiota capacity to accumulate ARGs.

With this work we demonstrated the presence of a diverse resistome in small wild mammals delving around farming areas. We showed clear differences in resistome composition between WTS and WM, and tight correlation of these differences with overall gut microbiota composition.

Supporting information

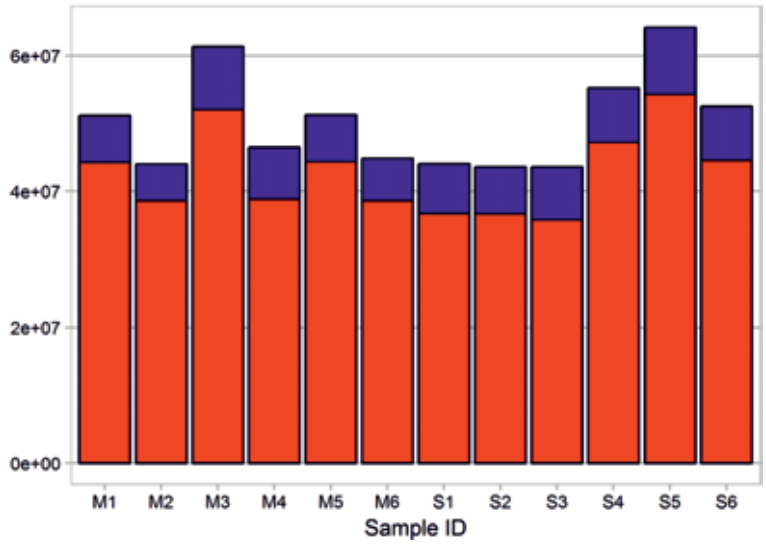

Figure S1. Number of reads (y-axis) obtained from shotgun metagenomics sequencing. Red colour indicates number of reads remained after quality control and blue indicates the portion of reads that didn't meet quality standards.

Table S1. Statistical comparison of relative abundance of antibiotic resistance gene (ARG) classes between White-toothed Shrews (WTS) and Wood mice (WM). ARG classes marked in bold have significantly different relative abundance between WTS and WM.

	Wilcoxon test		95% CI	
ARG class	p - value	q - value	WTS	WM
aminoglycoside	0.00216	0.00711	0.00596 - 0.01027	0.00212 - 0.00394
mupirocin	0.00216	0.00711	0.01943 - 0.03039	0.012 - 0.01998
peptide	0.00216	0.00711	0.00193 - 0.00632	0.00024 - 0.0017
puromycin	0.00216	0.00711	0.00083 - 0.00152	0.00006 - 0.00022
rifampin	0.00216	0.00711	0.00374 - 0.01038	0.00023 - 0.00094
sulfonamide	0.00216	0.00711	0.00145 - 0.00258	0.00021 - 0.0007
triclosan	0.00216	0.00711	0.00272 - 0.00357	0 - 0.00086
glycopeptide	0.005	0.01277	0.01259 - 0.01956	0.02538 - 0.05053
tetracenomycin	0.005	0.01277	0.00038 - 0.00066	0.00001 - 0.00009
beta_lactam	0.00866	0.0181	0.01661 - 0.036	0.00454 - 0.01029
kasugamycin	0.00866	0.0181	0.00142 - 0.00484	0.00005 - 0.00103
pleuromutilin	0.01515	0.02681	0.00307 - 0.0043	0.00388 - 0.00551
quinolone	0.01515	0.02681	0.02294 - 0.03561	0.01861 - 0.0226
bacitracin	0.04113	0.05912	0.01265 - 0.01621	0.01331 - 0.02331
macrolides	0.04113	0.05912	0.07395 - 0.11599	0.09562 - 0.15834
multidrug	0.04113	0.05912	0.09399 - 0.24898	0.06072 - 0.12679
aminocoumarin	0.06494	0.08297	0.00814 - 0.01403	0.01104 - 0.01851
chloramphenicol	0.06494	0.08297	0.00146 - 0.00289	0.00113 - 0.00179
polymyxin	0.09307	0.11267	0.01171 - 0.03454	0.0118 - 0.01586
fosmidomycin	0.39394	0.45303	0.00231 - 0.00443	0.00297 - 0.00549
tetracycline	0.48485	0.50689	0.01917 - 0.03032	0.02253 - 0.03713
trimethoprim	0.48485	0.50689	0.00437 - 0.00646	0.00512 - 0.00674
fosfomycin	0.58874	0.58874	0.00187 - 0.00568	0.00185 - 0.00371

Chapter 5

Chapter 6

General discussion and future perspectives

In the framework of the research I conducted during the past four years, I was lucky to work with various microbiomes from different research angles, and a part of this work is summarized in this thesis. It is a product of numerous, seemingly disconnected collaborations that were, however, all devoted to my central goal, namely to uncover driving forces that influence gut microbiota in host species other than humans, and to extrapolate obtained knowledge into a broader picture of the mammalian microbiota landscape. Microbiota composition has been shown to be affected by various aspects of genotype, lifestyle and other environmental factors, and is furthermore known to significantly fluctuate within and between individuals [80]. Nevertheless, at the population level environmental factors can create stable compositional patterns and affect host health and well-being [55]. In this chapter I will discuss findings and observations made in the course of my research, and will provide perspectives for future research.

The fantastic microbiomes and where to find them

Research described in this thesis, besides answering specific research questions formulated in research chapters, provided an account of microbiota composition of four lemur species (**Chapter 3** and **Chapter 3**), three pure breeds of domestic pigs (**Chapter 4**) and free living Great White-Toothed Shrews and Wood Mice (**Chapter 5**). Every investigated microbial community showed unique compositional profiles strongly correlated with the host species, maintaining to a certain extent conservation of compositional patterns within a given species group. It has previously been shown that different mammals harbour microbiota that differs significantly in composition and metabolic capacity, clustering mostly by host dietary habits [7].

The number of published studies focusing on wildlife microbiota remains limited. This can probably at least in part be explained by the far from trivial task of sampling. Obtaining (faecal) samples of a wild animal in natural conditions requires an inventive approach open to flexibility and improvisation,

further complicated by a number of technical limitations, such as difficulty in collection, proper metadata recording and transportation. **Chapter 2** and **Chapter 3** describe the composition of microbiota of wild lemurs in their natural habitat, which means samples for the studies could be collected only from the island of Madagascar, as this harbours the only natural territories inhabited by various lemur species [365]. Madagascar is a unique location by itself; it was separated from the African continent around 121 million years ago by the Mozambique Channel [366], and its isolation from the African mainland has allowed for the development of unique ecosystems. Most of the large mammals of Madagascar are endemic and cannot be found anywhere else in the world, which makes it a hotspot for evolutionary and natural history research. Unfortunately, Madagascar is not only a hotspot of scientific interest, but also a country with raging poverty, serious problems with overpopulation and consequently tremendous pressure on wildlife and natural ecosystems in general [367]. Our sampling efforts were limited to national parks with a strong involvement of international wildlife conservation organizations, where natural habitats are preserved in a more or less native state. Nevertheless, my collaborator Iris de Winter encountered a number of practical issues with organization of sampling events and tracking of lemurs. Besides such more organizational difficulties, sampling of wild animals in remote locations holds a number of challenges such as recording of adequate metadata, as well as collection and preservation of contamination-free samples. It has previously been pointed out by various studies that the method of sample preservation and storage can have a dramatic effect on microbiota composition profiling. The gold standard of sample preservation is immediate freezing and storage at -80^{0}C [368], however, this method is not available when sampling in such remote areas as rainforests on Madagascar. For our study we decided to use 70 - 100% ethanol for sample preservation and storage. Ethanol preservation showed to be adequate in terms of long term microbiota composition preservation [369, 370], and allows to fix samples immediately after collection. In addition this method fixates faecal samples, reducing chances of pathogen transmission, which is particularly important when

samples come from wild animals and countries with a high risk of emerging diseases.

It is not always necessary to travel across the globe for sampling of exotic animals to find differently shaped gut communities. In the research described in **Chapter 5** we focused on small wild mammals living in surroundings of Dutch farm lands. The mainland of the Netherlands is one of the most densely populated areas of Europe. With a land surface of 41,543 km^2 the country has 17.08 million inhabitants, with no pristine territories. Nevertheless, this small country is a homeland for a number of wild animals (https://www.nederlandsesoorten.nl), which are dwelling in close proximity to humans in highly anthropogenically influenced areas. Sampling in such locations is a relatively easy task with the opportunity to preserve samples in appropriate conditions. To this end, we used live-traps to catch small mammals around the farm areas, allowing to collect good quality faecal samples without harming animals and use the most suitable preservation method such as freezing on the spot in dry ice.

Microbiota, due to its complexity and diversity, is an endless field for exploration even in very well studied host species such as humans and domestic animals. For instance, research described in the PhD thesis of Tom van den Bogert showed how much novel insights could be gathered when focussing on the human small intestine rather than the more commonly studied faecal microbiota [371]. In order to add new perspectives on a well-studied model one should not always reach for the low hanging fruits, but rather aim to be inventive in obtaining yet unexplored and often more relevant samples, or finding a new angle for research. In **Chapter 4** we describe similarities and differences in caecal microbiota of domestic pigs belonging to three different breeds, including Landrace, Large White and Duroc. Porcine gut microbiota is a subject of intensive research due to the importance of pig production for food security and provision of high quality animal proteins [372]. A literature search for the term "pig microbiota" (Google Scholar, March 28th, 2019) yielded around 16,200 results, confirming pigs as the most popular subject for microbiota research among other farm animals (chicken – 15,800; cows – 14,600; sheep – 14,100).

Despite the wealth of research about pig microbiota, little attention has been given to date to investigations of microbiota variation between different breeds, in particular regarding the microbial communities residing in the upper parts of the intestine. Together with our collaborators from the University of Bologna, Italy, we were able to obtain cecum content from pedigree pigs that were living in a controlled environment with a defined diet. Our attention in this study was focused on caecal microbiota, which so far has received much less attention in comparison with faecal microbiota despite the cecum being the major fermentation compartment of the intestinal tract of pigs [373]. The research on well-defined animal models is extremely important and provides opportunities to investigate the often minor variations in microbiota composition and function due to e.g. genotype without them been overshadowed by large effectors such as diet. In addition, findings could be relevant for the pig industry through showing the path for better animal health management if a model employs production animals. For example we found that caecal microbiota of Landrace pigs was characterized by a higher average microbial diversity, as well as higher relative abundance of members of genera *Akkermansia, Ruminococcaceae* UCG-005, *Oscillospira* and *Faecalibacterium*. Previous research showed that these features of microbiota are associated with intestinal health [78, 315-318], suggesting that Landrace could be viewed as an attractive breed for farming in various environments.

The wild, the domesticated and the human

Modern day people separate themselves from the natural environment and other species. We are drawing a clear line between humans and animals with disregard of the fact that all parts of the biosphere are intertwined with each other. Humanity has particularly tight connections with other mammalian species. For the largest part of our history hunting of large mammals was essential for survival, providing food rich in proteins and fats, material for building shelters and clothing to survive in colder climates, and tendons and bones for crafting of weapons and tools. With the development of technology,

agriculture and society, humans discovered more sophisticated and reliable ways beyond hunting to sustain an ever-growing population. Nevertheless, our bonds with other mammalian species did not become any weaker. Hunting for food was replaced by animal husbandry, and it has been estimated that more than 1.3 billion pigs and 250 million cows are raised yearly worldwide (www.ciwf.org.uk). By definition animal husbandry implies very tight relationships between farmed animals and humans. In some areas, such as production of tools and clothing, technology has provided synthetic or plant based alternatives for animal derived materials, however, with the development of medical sciences animals gained new importance. Animal models are essential for drug development, production of antigens, and as a source of cell lines and other compounds that are widely used in science and medicine. Microbiota research is not an exception, as it employs a number of laboratory animal models to investigate the influence of microbial communities on body mass index (BMI) [374], metabolism [375] and insulin resistance [376], to just give a few examples. On the other side of the spectrum of microbiota research are investigations of wild animals with the goal of discerning general ecological patterns. Ruth Ley and co-workers showed that the dietary habits of animals as well as their phylogeny have a strong influence on intestinal microbiota, however, mammals as a group are fairly similar with respect to the composition of their intestinal microbiota, which is clearly distinguishable from microbial communities inhabiting other complex environments [7]. Primates, among other wild animals, attracted most attention due to their phylogenetic closeness to humans [377]. It has been shown that the microbiota of humans and other hominids share the same origin and evolved and diverged in parallel with primate evolution [83]. Moeller et al. discovered that divergence of human microbiota from that of wild Chimpanzees, Bonobos and Gorillas strongly depends on diet and lifestyle of a human population. Members of tribes that live with the traditional rural lifestyle of for example Malawi and Venezuela have been shown to be much closer to other hominids with respect to their intestinal microbiota, in contrast to people from western countries, and more specifically from the USA [86].

In **Chapters 2** and **Chapters 3** of this thesis we investigated faecal microbiota of wild lemurs, focusing on effects of environmental (external) and internal factors on overall microbiota composition. Lemurs are one of the most diverse groups of primates comprising a total of 33 species with an impressive divergence within the group. Lemurs are phylogenetically distant from other primates and probably closely resemble the last common ancestor of primates [377, 378]. It is impossible to directly extrapolate findings with respect to lemur gut microbiota to humans or other primates, however, knowledge about within-species variation of gut microbiota and the effect of different environmental factors is necessary to understand the dynamic landscape of intestinal microbial communities. It has previously been shown that even the microbiota of non-human primates starts to resemble that of humans when animals are kept in captivity and being fed diets different from those consumed in the wild [379].

Even though the gut microbiota of lemurs holds a number of specific features, it overall showed high similarity to that of other primate species, including humans [7, 89, 380]. The resolution of studies described in **Chapters 2** and **3** was high enough to catch not only variations induced by commonly accepted effectors of microbiota such as diet and genetics, but also interactions with more subtle factors like intestinal parasites. Detection of the microbiota response to such factors in wild animals showed its responsiveness to variations in environmental conditions, however, we could not extrapolate and make direct correlations between changes in microbiota and animal health or wellbeing.

Pigs were domesticated around 7000 years ago, and currently one of the five most common domestic livestock animals [381]. As indicated above, pig production holds a key role in food security for the growing world population. In addition, *Sus scrofa* has been shown to be a valuable model for pharmacology, toxicology, transplantation sciences and other biomedical fields due to similarity in anatomy and physiology with humans [382, 383]. Pigs are particularly interesting as a model to study evolution and domestication due to their genetic closeness with wild boars (*Sus scrofa*). Wild boars are cousins and ancestor of all domestic pigs [384] and are wide-spread around the world. Currently, the population of wild boars is thriving in Europe, and it has been estimated that in

2012 around 2.2 million animals were killed by hunters across 18 European countries [385]. Such abundance of wild *Sus scrofa* gives a wide range of opportunities for sample collection from living and killed animals for various research purposes. Comparison of microbiota between wild and domesticated pigs can help discerning how domestication and change of lifestyle has affected the gut ecosystem. In **Chapter 4** we focused on microbiota of three breeds of domestic pigs and found a number of microbial taxa that were statistically significantly different in relative abundance between breeds. This comparison didn't include samples from more genetically distant wild boars due to the scope of the research on fine variations of microbiota between animals that were kept in highly controlled conditions to negate effects other than those elicited by breed factors. However, as a part of my project in collaboration with Iratxe Diez Delgado and Christian Gortázar Schmidt (Health & Biotechnology Group, Instituto de Investigación en Recursos Cinegéticos, Ciudad Real, Spain) I was able to obtain a number of samples from different intestinal sections of wild boars during sampling events in Spain (Ciudad Real). Microbiota of animals was profiled using Illumina HiSeq sequencing of PCR-amplified and barcoded 16S ribosomal RNA (rRNA) gene fragments covering the V4 variable region (see Chapter 4 for experimental details of DNA extraction, PCR, sequencing and sequence processing), and we discovered that a number of OTUs could not be assigned to any known genus (Fig. 1), in contrast to microbial composition analysis of domestic pigs. The major reason for discovering OTUs with ambiguous assignment at e.g. genus level is the absence of a cultured/described member of the genus in the reference database used for taxonomic classification. Discovery of potentially novel microbial taxa in a well-studied model such as human, laboratory- or farm animals is not very common when a conventional population is considered. Nevertheless, if the sampled population has a different lifestyle from previously studied individuals, then discovery of population-specific microbiota signatures is more common. As an example, Schnorr et al. found a number of OTUs in the gut microbiota of members of the Hadza that were not assigned to any genus or even higher phylogenetic levels. Overall microbial composition was shown to be very different from that of a western population,

providing evidence of microbiota adaptation to specific dietary and lifestyle conditions [386].

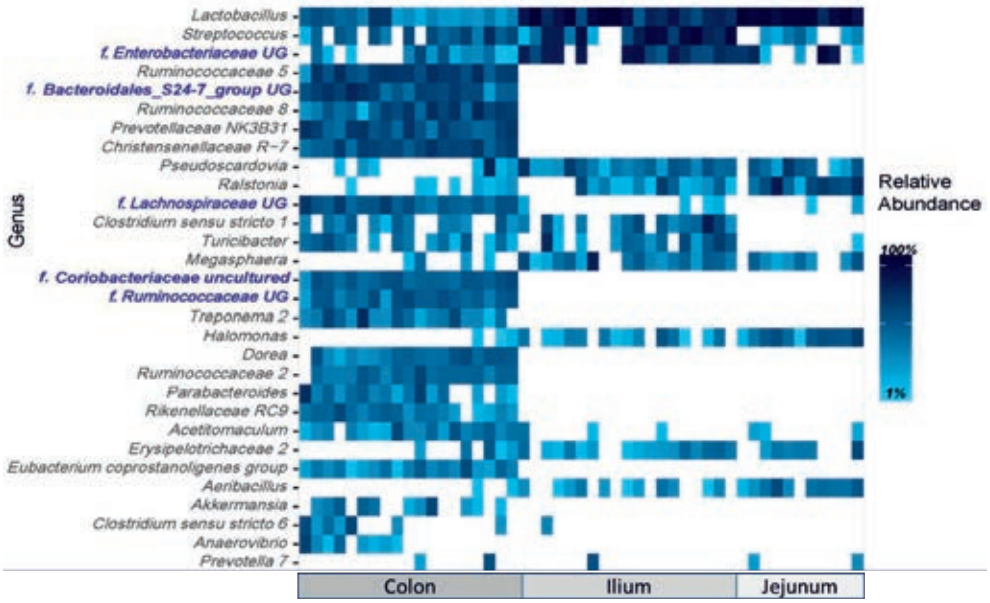

Figure 1. Compositional heatmap of the most abundant genera (> 5% abundance per sample) along the intestinal tract of wild boars as measured based on 16S rRNA gene V4 sequences. Yet unidentified genera are shown in blue and bold type. Only samples that passed quality control were displayed (see Chapter 4 for details). UG, unidentified genus.

Microbiota of wild animals is very sensitive to changes in environmental conditions. We found significant variations in microbiota composition between wild boars sampled in private (n = 11) and public (n = 11) hunting grounds of Ciudad Real province (Fig. 2). In this case differences in environmental conditions probably have a more determinative effect on gut microbiota as compared to differences in genetic makeup of the population. Sampling areas were only around 40 km apart from each other and do not have any obvious natural or artificial barriers that would prevent migration of animals and intermixing. Even though it is difficult to pinpoint the particular environmental factors that affect microbiota, it can be speculated that differences in land management and anthropogenic influence are factors underlining microbiota variations. Private hunting grounds are preserved in a more pristine state and surrounded by relatively large areas of mountain forest. Even though animals

are fed throughout the year by land owners, the amount of supplied food is limited and aimed to supplement the natural diet of the wild boars rather than to replace it. In contrast, public hunting areas are situated in proximity to farm lands and human settlements. Seasonally, wild boars of public hunting lands have access to crops as well as to a high number of feeding lots organized by farmers and hunters.

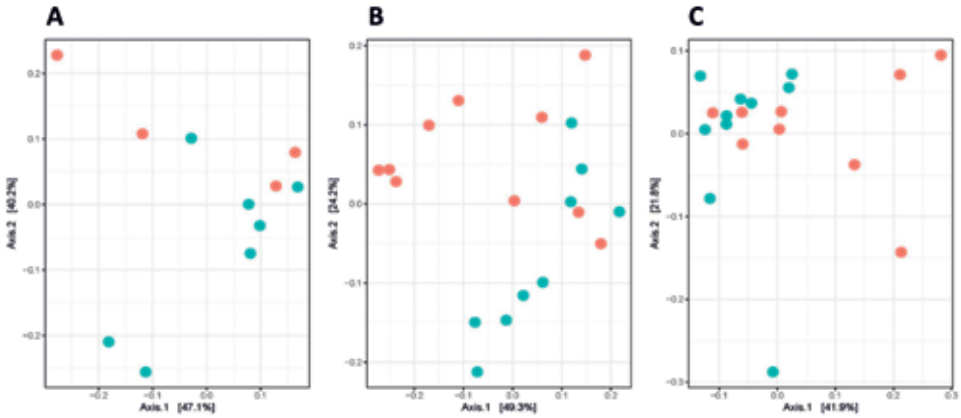

Figure 2. Multidimensional scaling of wild boar microbial beta diversity with weighted unifrac as the distance matrix. Different plots focus on samples taken either from jejunum (A), ileum (B) or colon (C). Samples are coloured by public (green) vs. private (red) hunting grounds. Only samples that passed quality control were used.

Exploration of wildlife ecology in populated areas is an important task for promotion of sustainable civilisation growth and peaceful wildlife-human co-existence. Small mammals inhabiting farmlands could be good indicators of overall environmental health. To this end, primary focus of **Chapter 5** was the correlation between intestinal microbiota composition and resistome in small mammals on Dutch farmlands.

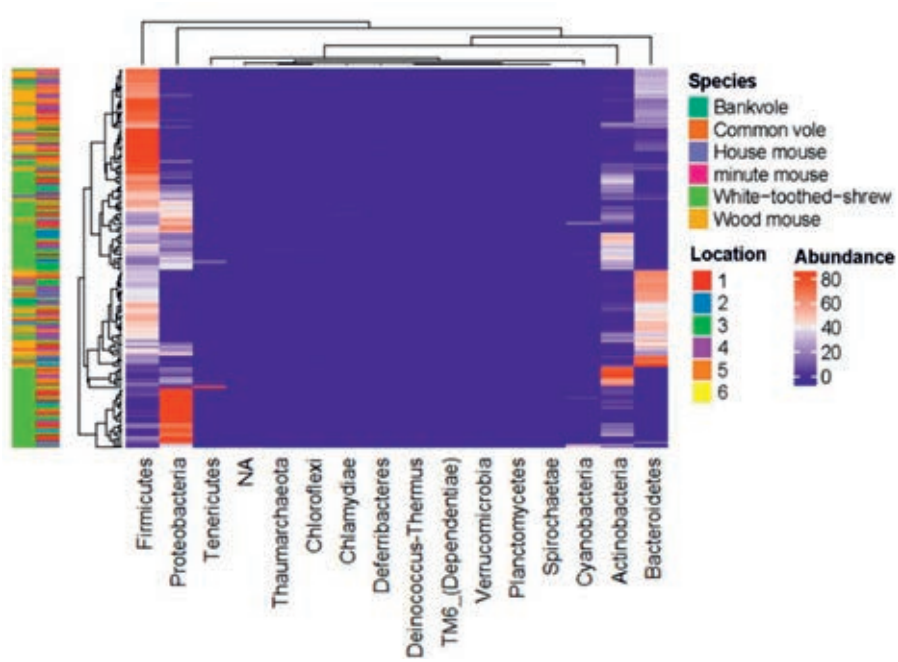

Figure 3. Heatmap of relative abundance of microbial phyla in faecal microbiota of six small mammals as based on 16S rRNA gene sequences (see Chapter 5 for experimental details regarding sampling, DNA extraction, PCR, sequencing and sequence processing and analysis). Intensity of the colour shows the relative abundance (%) as indicated in the right side legend. Coloured sidebars at the left side indicate individual sample location and species identity, respectively, with colours corresponding to the legend on the right side. Rows and columns were arranged and dendrograms were built using hierarchical clustering.

In total six species were sampled, however, due to the varying number of available samples and experimental design only samples from Great White-Toothed shrews and Wood mice were presented in **Chapter 5**. Even though samples from Bank vole (*Myodes glareolus*), Common vole (*Microtus arvalis*), House mouse (*Mus musculus*) and Minute mouse were not included in **Chapter 5** their microbiota composition was determined. We found drastic differences in microbiota composition between shrews and rodents already at the phylum level (Fig. 3). In contrast to faecal microbiota profiles of rodents that were predominated by *Firmicutes* and *Bacteroidetes*, samples taken from shrews were predominated by *Proteobacteria* and *Actinobacteria*, in line with findings of Knowles et al. [10].

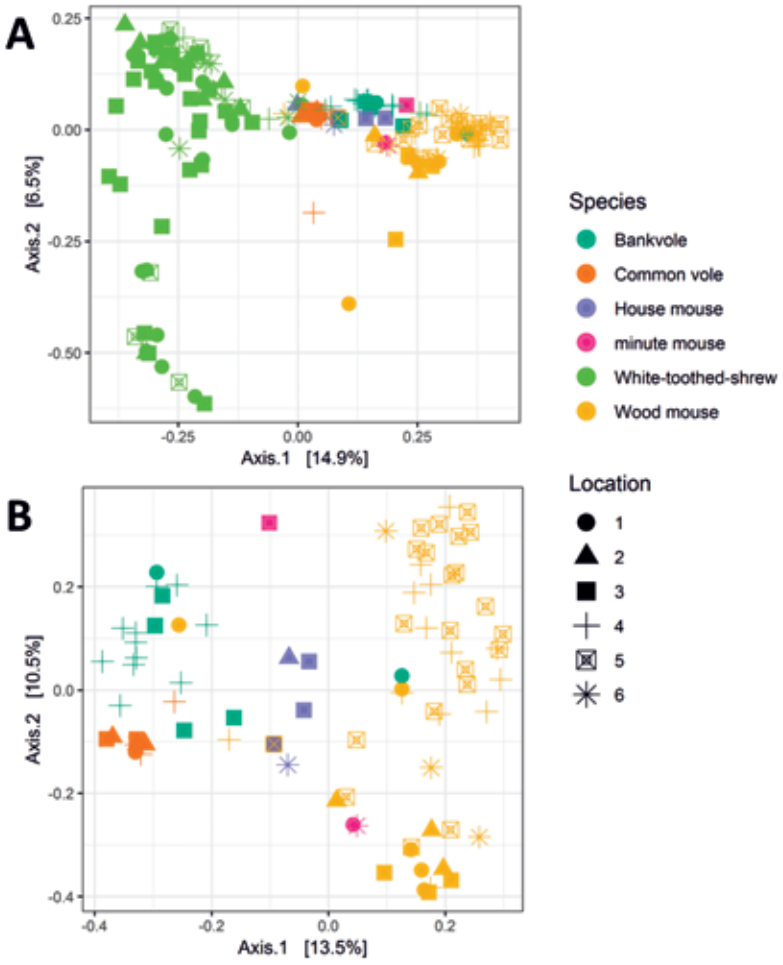

Figure 4. Principal coordinate ordination plots based on Bray-Curtis distances. A) Samples from Bankvoles, Common voles, House mice, Minute mice, Wood mice and Great White-Toothed shrew; B) Samples from Bankvoles, Common voles, House mice, Minute mice and Wood mice. Colour of the symbols represents species identity and shape corresponds to sampling location.

It is safe to assume that the similarity in diets of investigated rodent species in contrast to shrews' drastically different dietary habits (as discussed in detail in Chapter 5) play an important role in shaping microbiota, even at higher taxonomic levels. In depth understanding of wildlife intestinal microbiota in anthropogenic environments such as those found in The Netherlands, could help to understand factors that shape their ecological niche and adaptability. Even

species that are very similar in their appearance often carry distinct microbial features, reflecting minor differences in their lifestyle. To this end, besides the more pronounced differences in microbial composition between investigated representatives of *Eulipotyphla* (shrews) and *Rodentia* (mice and voles) orders, we also observed noticeable differences between the different members of the *Rodentia*, correlating with their phylogenetic identity and sampling location (Fig. 4). It is important to point out that the number of samples per species group within the *Rodentia*, as well as strong linkage ($p < 0.001$) between species identity and sampling location did not allow to draw a conclusive picture of microbiota differences and similarities in animals with similar appearance.

The sequencing, or from a sample to conclusions and back again

Culture independent methods have been dominating the field of microbial ecology for the past two decades. In the General Introduction of my thesis I discussed how the development particularly of next generation sequencing (NGS) technologies has provided unprecedented power for the investigation of complex microbial communities. NGS is an extremely flexible and powerful tool in modern day biological and medical sciences, however, wrong application coupled with blind believe in complete accuracy of results could harm the outcome of the research and lead to wrong or at least biased conclusions. In this part of the discussion I would like to share my experience of NGS amplicon sequencing for profiling of animal gut microbiota.

NGS is a powerful tool when applied properly and can help to understand important aspects of microbial ecology. Nevertheless, it is pivotal to carefully consider the application of NGS in research. Very often, generated data is seen as an absolute truth without appreciation of possible biases that could skew the data and produce biased results. Earlier in this chapter I discussed the importance of the preservation protocol and storage conditions for recovery of a representative DNA sample. It is difficult to underestimate the importance of the

DNA quality on final results. Various commercial kits are available for DNA extraction for a wide range of sample types, and there have been a number of good attempts to benchmark DNA extraction protocols with respect to their impact on microbiota profiling. Yuan et al. tested a number of methods for DNA extraction and showed that the best results in terms of composition recovery were delivered by methods that include bead beating steps [387]. Most of the DNA extraction methods that have been suggested for intestinal microbiota research have been optimized for human faecal samples, and it cannot be excluded that such methods could be less optimal for DNA extraction from animal faecal or other intestinal samples such as those taken from the small intestine. For the majority of samples analysed in this thesis we used an in-house DNA extraction protocol that combines repetitive bead beating [113] with an automated extraction and purification system (Promega Maxwell). The employed two-step bead beating allowed to disrupt fibrous material that is often present in faecal matter of herbivorous animals and cell walls of Gram negative and Gram positive bacteria. This DNA purification protocol showed to be superior or on par with other methods. Nevertheless, it was not always possible to obtain good quality DNA from all animal samples. In particular, samples from the small intestine of wild boars and insectivorous animals showed to be difficult to handle (data not shown). Samples from the small intestine contain relatively small numbers of microbial cells and a high amount of host associated compounds [388] resulting in a small amount or poor quality of extracted microbial DNA.. A low recovery of microbial DNA fragments could result in inadequate coverage of microbiota when metagenomics studies are considered, and insufficient PCR amplification for amplicon sequencing studies.

Table 1. Genera that are commonly found in samples used as negative controls in my projects with a short description of origin, or specific properties. A star (*) behind a genus name indicates that this genus was also identified as a common contaminant of commercially available kits (Salter, 2014 [389]).

Genera	Mean	SD	Origin
*Ralstonia**	47.5	19.7	water/soil
Halomonas	14.2	9.3	halophilic proteobacteria (up to 25% of NaCl)
Shewanella	4.3	1.9	marine bacteria fish surface, model for anaerobic respiration
*Cupriavidus**	3.7	1.7	soil, resistant to metals
Caldalkalibacillus	3.0	3.2	isolated from hot spring
Rothia	2.6	2.3	human associated (mouth)/pathogens
*Bacillus**	1.8	2.7	water/soil
*Sphingomonas**	0.3	0.2	water/soil

In my work samples from the small intestine were included only in the project that dealt with wild boar microbiota, which is not included in this thesis as a separate research chapter. Nevertheless, I experienced extensive complications associated with this type of samples, ranging from difficulties with DNA extraction and PCR amplification to misleading results. After rounds of the fine-tuning of DNA extraction and PCR procedures I was able to obtain sequencing data, however, the microbial composition of the samples with low quality/quantity of starting DNA showed considerable differences from the rest of the intestinal samples (Fig. 5, cluster III). We found that microbial composition of the samples with low quality DNA mostly represented microbial species that are commonly found in other environments such as water and soil (Fig. 6, Table 1). To ensure quality of the results it is important to have a clear understanding of contaminant composition. It was previously shown that even high quality laboratory reagents such as DNA extraction kits contain fragments of amplifiable bacterial DNA that could be sequenced and create bias [389]. In my experience the best way to tackle the issue of amplification and sequencing bias is to prepare DNA of adequate quality and include a negative control in every sequencing run. However, different researchers could have a different understanding of what is a negative control when it comes to sequencing. In our experience the best negative

control should include either samples of DNA/RNA free water or a growth medium used in the experiment processed alongside with actual samples. Meta-analysis of negative control samples from several separate experiments showed consistent composition of microbial taxa that could be recovered from a blank sample (Fig. 6). Knowledge of the microbial composition of amplicons consistently retrieved from blank extractions is extremely useful for ruling out non-biological variations in microbiota composition in particular when an actual biological sample has a low amount of bacterial DNA.

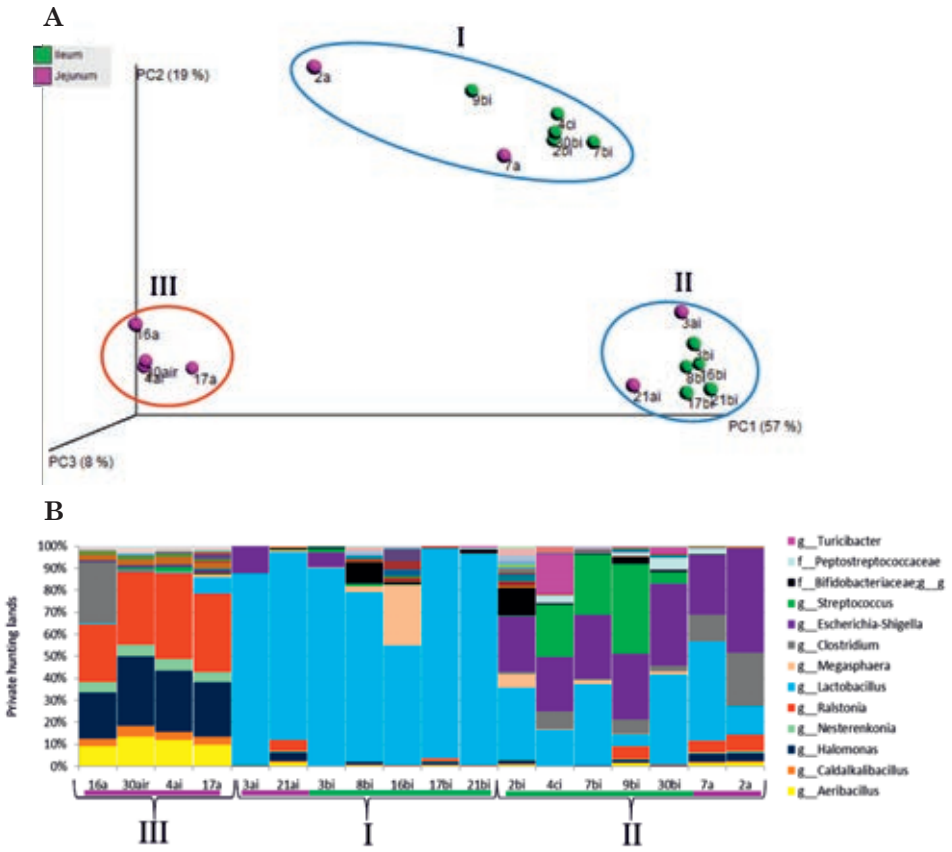

Figure 5. A) PCoA ordination based on microbiota composition of small intestinal samples from wild boars. Envelopes were drawn around samples showing visible clustering. Cluster III corresponds to samples with low DNA quality. B) Bar chart showing microbial composition at genus level per sample. Brackets at the bottom of the graph indicate clusters based on the PCoA plot and coloured bars indicate the intestinal part (purple for Jejunum and green for Ileum).

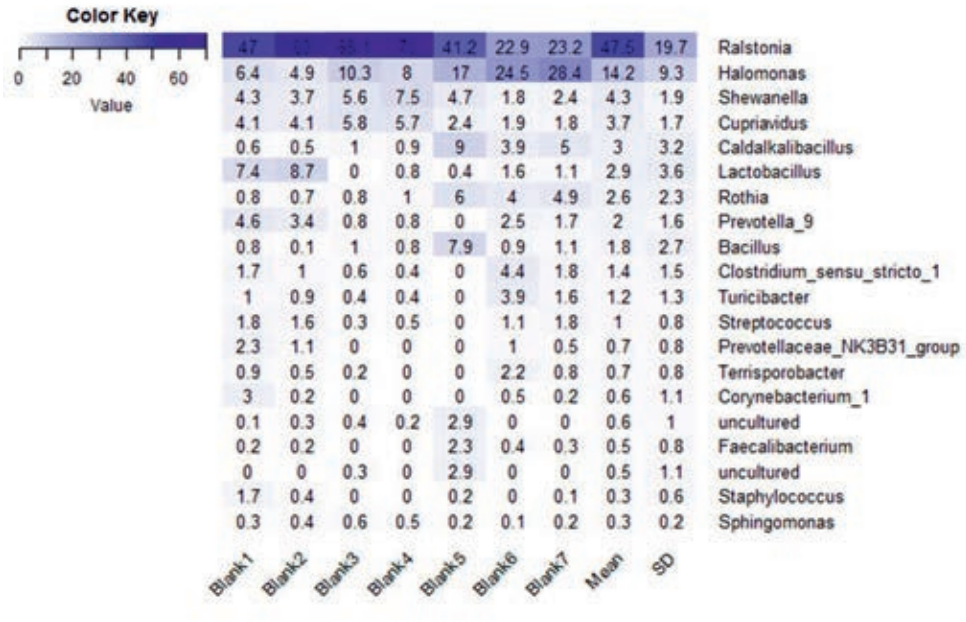

Figure 6. Heatmap showing relative abundance of top 20 genera found in no-template control samples sequenced alongside with other samples used in my projects.

Preparation of high quality DNA and proper sequencing are tremendously important in obtaining high quality data, however, following bioinformatics and statistical analysis should be taken with great care as well. A number of pipelines such as QIIME and mothur are available for streamlined analysis that will bring researchers from raw Illumina output to an abundance table. It was shown before that above-mentioned pipelines have similar performance with respect to amplicon sequencing of gut microbiota [390]. In the studies described in this thesis, we used NG-tax, a sequence processing pipeline developed in house, that gives a highly accurate estimation of relative abundance distributions and was tested using a number of synthetic MOCK communities. Recovery of the MOCK community composition was better than with QIIME with default parameters [47]. The path from raw data to OTU/abundance tables is quite strait forward, however, following statistical analyses will largely depend on experimental design and research goals. The number of options for advanced uni- and multivariate statistical analyses of gut microbiota composition data is almost limitless with a wide variety of ecological methods for community analysis that

employ multivariate statistics and diversity measurements. It is easy to get lost in the jungles of ecological statistics. For future experiments I would advocate an approach where researchers propose a plan of statistical analyses for hypothesis testing prior to the execution of the experiment in order to not get lost in the jungle of the ever increasing amounts of data that can easily be generated, and that are often prone to merely descriptive data analysis and interpretation. In addition, a researcher should clearly understand what kind of questions could be answered using any of the broad range of currently available ~omics techniques.

Back to the Future of microbiota research

This thesis gives a glimpse on the vast amount of unexplored mysteries related to the gut microbiota of wild and domestic animals. It is fascinating to explore differences in microbiota composition driven by different factors with extrapolation of the knowledge to a broader picture of the dynamic microbiota landscape. Nevertheless, it should be clear that mechanistic studies and isolation of bacterial strains by culturing techniques could provide much greater reward in terms of application to animal and human health. Isolation of the bacterial members from wild relatives of modern livestock could provide a range of probiotic species that once resided in the intestinal tract but went extinct in the process of domestication. A collection of the strains from wild animals will allow to study in detail metabolism of exotic food compounds and to design probiotics specifically for animals. Microbiota of wild and endangered animal species could be a unique collection of microorganisms with potential health modulating properties that could be lost due to extinction of the host species, and thus efforts for the conservation of microbial biodiversity and heritage are needed. Besides traditional culturing methods, more advanced molecular methods such as deep metagenomic sequencing should be applied to obtain a better understanding of microbial metabolism and genetic makeup of uncultured members of intestinal microbial communities [275, 391, 392]. Animals have a wide range of specific diet adaptations and physiology that could be completely different from those of humans or domestic animals, potentially resulting in the development of distinct

patterns of host-microbe interactions, or food degradation pathways that could be studied by application of deep metagenomics and isolation of individual members of the various microbiomes.

Investigation of gut microbiota of different domestic animal breeds and their wild relatives could contribute to promotion of intestinal health, increased feed efficiency and improved sustainability of animal husbandry overall. In the research described in this thesis we showed that even closely related breeds have significant differences in their microbiota, however, it will be hugely beneficial to more precisely understand the mechanisms underlying such microbiota modulation. One way of approaching this challenge could be to merge detailed information about host genetic makeup and microbiota compositional as well as functional profiles at the strain level. Detailed understanding of microbe-host genetic correlations will create opportunities for design of breed specific probiotics and facilitate informed decisions in breed selection.

Appendices

References

1. Allwood AC, Walter MR, Kamber BS, Marshall CP, Burch IW: **Stromatolite reef from the Early Archaean era of Australia.** *Nature* 2006, **441:**714.
2. Jacobsen SB: **How old is planet Earth?** *Science* 2003, **300:**1513-1514.
3. Stanley SM: **An ecological theory for the sudden origin of multicellular life in the late Precambrian.** *Proceedings of the National Academy of Sciences* 1973, **70:**1486-1489.
4. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K: **A new view of the tree of life.** *Nature Microbiology* 2016, **1:**16048.
5. Stetter KO: **Extremophiles and their adaptation to hot environments.** *FEBS Letters* 1999, **452:**22-25.
6. Mock T, Thomas DN: **Recent advances in sea-ice microbiology.** *Environmental Microbiology* 2005, **7:**605-619.
7. Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI: **Worlds within worlds: evolution of the vertebrate gut microbiota.** *Nature Reviews Microbiology* 2008, **6:**776.
8. Nixon S, Cousins CR, Cockell C: **Plausible microbial metabolisms on Mars.** *Astronomy & Geophysics* 2013.
9. Wernegreen JJ: **Genome evolution in bacterial endosymbionts of insects.** *Nature Reviews Genetics* 2002, **3:**850.
10. Munson MA, Baumann P, Kinsey MG: **Buchnera gen. nov. and Buchnera aphidicola sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids.** *International Journal of Systematic and Evolutionary Microbiology* 1991, **41:**566-568.
11. Blazejak A, Erséus C, Amann R, Dubilier N: **Coexistence of bacterial sulfide oxidizers, sulfate reducers, and spirochetes in a gutless worm (Oligochaeta) from the Peru margin.** *Applied and Environmental Microbiology* 2005, **71:**1553-1561.
12. Gogineni VK, Morrow LE, Gregory PJ, Malesker MA: **Probiotics: history and evolution.** *Journal of Ancient Diseases & Preventive Remedies* 2013.
13. Demain AL, Vandamme EJ, Collins J, Buchholz K: **History of industrial biotechnology.** *Industrial Biotechnology: Microorganisms* 2017, **1:**1-84.

14. Ullmann A: **Pasteur-Koch: Distinctive ways of thinking about infectious diseases.** *Microbe-American Society for Microbiology* 2007, **2:**383.
15. Metchnikoff E: **Essais optimistes. Paris. The prolongation of life. Optimistic studies. Translated and edited by P. Chalmers Mitchell.** London: Heinemann; 1907.
16. Zalkind S: *Ilya Mechnikov: his life and work.* The Minerva Group, Inc.; 2001.
17. Sonnenborn U: **Escherichia coli strain Nissle 1917—from bench to bedside and back: history of a special Escherichia coli strain with probiotic properties.** *FEMS Microbiology Letters* 2016, **363**.
18. Prescott SL: **History of medicine: Origin of the term microbiome and why it matters.** *Human Microbiome Journal* 2017, **4:**24-25.
19. Eisen J: **What does the term microbiome mean? And where did it come from? A bit of a surprise.** 2015.
20. Marchesi JR, Adams DH, Fava F, Hermes GD, Hirschfield GM, Hold G, Quraishi MN, Kinross J, Smidt H, Tuohy KM: **The gut microbiota and host health: a new clinical frontier.** *Gut* 2016, **65:**330-339.
21. Thompson G, Trexler P: **Gastrointestinal structure and function in germ-free or gnotobiotic animals.** *Gut* 1971, **12:**230.
22. Rieder R, Wisniewski PJ, Alderman BL, Campbell SC: **Microbes and mental health: a review.** *Brain, Behavior, and Immunity* 2017, **66:**9-17.
23. Wilson ID, Nicholson JK: **Gut microbiome interactions with drug metabolism, efficacy, and toxicity.** *Translational Research* 2017, **179:**204-222.
24. den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud D-J, Bakker BM: **The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism.** *Journal of Lipid Research* 2013, **54:**2325-2340.
25. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F: **From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites.** *Cell* 2016, **165:**1332-1345.
26. Wu W, Sun M, Chen F, Cao AT, Liu H, Zhao Y, Huang X, Xiao Y, Yao S, Zhao Q: **Microbiota metabolite short-chain fatty acid acetate promotes intestinal IgA response to microbiota which is mediated by GPR43.** *Mucosal Immunology* 2017, **10:**946.
27. Bindels LB, Porporato P, Dewulf E, Verrax J, Neyrinck AM, Martin J, Scott K, Calderon PB, Feron O, Muccioli G: **Gut microbiota-derived propionate**

reduces cancer cell proliferation in the liver. *British Journal of Cancer* 2012, **107:**1337.

28. Hosseini E, Grootaert C, Verstraete W, Van de Wiele T: **Propionate as a health-promoting microbial metabolite in the human gut.** *Nutrition Reviews* 2011, **69:**245-258.
29. Pluznick J: **A novel SCFA receptor, the microbiota, and blood pressure regulation.** *Gut Microbes* 2014, **5:**202-207.
30. Vermeiren J, Van de Wiele T, Verstraete W, Boeckx P, Boon N: **Nitric oxide production by the human intestinal microbiota by dissimilatory nitrate reduction to ammonium.** *BioMed Research International* 2009, **2009**.
31. Larrosa M, Luceri C, Vivoli E, Pagliuca C, Lodovici M, Moneti G, Dolara P: **Polyphenol metabolites from colonic microbiota exert anti-inflammatory activity on different inflammation models.** *Molecular Nutrition & Food Research* 2009, **53:**1044-1054.
32. Rooks MG, Garrett WS: **Gut microbiota, metabolites and host immunity.** *Nature Reviews Immunology* 2016, **16:**341.
33. Lebeer S, Vanderleyden J, De Keersmaecker SC: **Host interactions of probiotic bacterial surface molecules: comparison with commensals and pathogens.** *Nature Reviews Microbiology* 2010, **8:**171.
34. Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei D, Goldfarb KC, Santee CA, Lynch SV: **Induction of intestinal Th17 cells by segmented filamentous bacteria.** *Cell* 2009, **139:**485-498.
35. Faust K, Raes J: **Microbial interactions: from networks to models.** *Nature Reviews Microbiology* 2012, **10:**538.
36. Perez-Garcia O, Lear G, Singhal N: **Metabolic network modeling of microbial interactions in natural and engineered environmental systems.** *Frontiers in Microbiology* 2016, **7:**673.
37. Stecher B, Hardt W-D: **Mechanisms controlling pathogen colonization of the gut.** *Current Opinion in Microbiology* 2011, **14:**82-91.
38. Bäumler AJ, Sperandio V: **Interactions between the microbiota and pathogenic bacteria in the gut.** *Nature* 2016, **535:**85.
39. Sekirov I, Tam NM, Jogova M, Robertson ML, Li Y, Lupp C, Finlay BB: **Antibiotic-induced perturbations of the intestinal microbiota alter host susceptibility to enteric infection.** *Infection and Immunity* 2008, **76:**4726-4736.

40. Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L: **Phylogenetically novel uncultured microbial cells dominate Earth microbiomes.** *MSystems* 2018, **3:**e00055-00018.

41. Schuster SC: **Next-generation sequencing transforms today's biology.** *Nature Methods* 2007, **5:**16.

42. Janssen PH: **Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes.** *Appl Environ Microbiol* 2006, **72:**1719-1728.

43. Rothberg JM, Leamon JH: **The development and impact of 454 sequencing.** *Nature Biotechnology* 2008, **26:**1117.

44. Heather JM, Chain B: **The sequence of sequencers: The history of sequencing DNA.** *Genomics* 2016, **107:**1-8.

45. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI: **QIIME allows analysis of high-throughput community sequencing data.** *Nature Methods* 2010, **7:**335.

46. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities.** *Appl Environ Microbiol* 2009, **75:**7537-7541.

47. Ramiro-Garcia J, Hermes GD, Giatsis C, Sipkema D, Zoetendal EG, Schaap PJ, Smidt H: **NG-Tax, a highly accurate and validated pipeline for analysis of 16S rRNA amplicons from complex biomes.** *F1000Research* 2018, **5**.

48. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L: **DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data.** *Microbiome* 2018, **6:**23.

49. Menzel P, Ng KL, Krogh A: **Fast and sensitive taxonomic classification for metagenomics with Kaiju.** *Nature Communications* 2016, **7:**11257.

50. Roumpeka DD, Wallace RJ, Escalettes F, Fotheringham I, Watson M: **A review of bioinformatics tools for bio-prospecting from metagenomic sequence data.** *Frontiers in Genetics* 2017, **8:**23.

51. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W: **MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph.** *Bioinformatics* 2015, **31:**1674-1676.

52. Namiki T, Hachiya T, Tanaka H, Sakakibara Y: **MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads.** *Nucleic Acids Research* 2012, **40:**e155-e155.

53. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA: **metaSPAdes: a new versatile metagenomic assembler.** *Genome Research* 2017, **27:**824-834.

54. Compeau PE, Pevzner PA, Tesler G: **How to apply de Bruijn graphs to genome assembly.** *Nature Biotechnology* 2011, **29:**987.

55. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D: **Population-level analysis of gut microbiome variation.** *Science* 2016, **352:**560-564.

56. Spor A, Koren O, Ley R: **Unravelling the effects of the environment and host genotype on the gut microbiome.** *Nature Reviews Microbiology* 2011, **9:**279.

57. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, Spector TD, Bell JT, Clark AG, Ley RE: **Genetic determinants of the gut microbiome in UK twins.** *Cell Host & Microbe* 2016, **19:**731-743.

58. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, Beaumont M, Van Treuren W, Knight R, Bell JT: **Human genetics shape the gut microbiome.** *Cell* 2014, **159:**789-799.

59. Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N: **Environment dominates over host genetics in shaping human gut microbiota.** *Nature* 2018, **555:**210.

60. Zhang C, Zhang M, Wang S, Han R, Cao Y, Hua W, Mao Y, Zhang X, Pang X, Wei C: **Interactions between gut microbiota, host genetics and diet relevant to development of metabolic syndromes in mice.** *The ISME Journal* 2010, **4:**232.

61. Carmody RN, Gerber GK, Luevano Jr JM, Gatti DM, Somes L, Svenson KL, Turnbaugh PJ: **Diet dominates host genotype in shaping the murine gut microbiota.** *Cell host & microbe* 2015, **17:**72-84.

62. Bian G, Ma S, Zhu Z, Su Y, Zoetendal EG, Mackie R, Liu J, Mu C, Huang R, Smidt H: **Age, introduction of solid feed and weaning are more important determinants of gut bacterial succession in piglets than breed and nursing mother as revealed by a reciprocal cross-fostering model.** *Environmental Microbiology* 2016, **18:**1566-1577.

63. Mu C, Bian G, Su Y, Zhu W: **Differential effects of breed and nursing on early-life colonic microbiota and immune status as revealed in a cross-fostering piglet model.** *Appl Environ Microbiol* 2019**:**AEM. 02510-02518.

64. Daniel H, Gholami AM, Berry D, Desmarchelier C, Hahne H, Loh G, Mondot S, Lepage P, Rothballer M, Walker A: **High-fat diet alters gut microbiota physiology in mice.** *The ISME journal* 2014, **8:**295.

65. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R: **Linking long-term dietary patterns with gut microbial enterotypes.** *Science* 2011, **334:**105-108.

66. Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL: **Diet-induced extinctions in the gut microbiota compound over generations.** *Nature* 2016, **529:**212.

67. Graf D, Di Cagno R, Fåk F, Flint HJ, Nyman M, Saarela M, Watzl B: **Contribution of diet to the composition of the human gut microbiota.** *Microbial Ecology in Health and Disease* 2015, **26:**26164.

68. Salonen A, Lahti L, Salojärvi J, Holtrop G, Korpela K, Duncan SH, Date P, Farquharson F, Johnstone AM, Lobley GE: **Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men.** *The ISME Journal* 2014, **8:**2218.

69. Simpson HL, Campbell BJ: **Dietary fibre–microbiota interactions.** *Alimentary Pharmacology & Therapeutics* 2015, **42:**158-179.

70. Barelli C, Albanese D, Donati C, Pindo M, Dallago C, Rovero F, Cavalieri D, Tuohy KM, Hauffe HC, De Filippo C: **Habitat fragmentation is associated to gut microbiota diversity of an endangered primate: implications for conservation.** *Scientific Reports* 2015, **5:**14862.

71. Tilman D, Fargione J, Wolff B, D'antonio C, Dobson A, Howarth R, Schindler D, Schlesinger WH, Simberloff D, Swackhamer D: **Forecasting agriculturally driven global environmental change.** *Science* 2001, **292:**281-284.

72. Amato KR, Yeoman CJ, Kent A, Righini N, Carbonero F, Estrada A, Gaskins HR, Stumpf RM, Yildirim S, Torralba M: **Habitat degradation impacts black howler monkey (Alouatta pigra) gastrointestinal microbiomes.** *The ISME Journal* 2013, **7:**1344.

73. Modi SR, Collins JJ, Relman DA: **Antibiotics and the gut microbiota.** *The Journal of Clinical Investigation* 2014, **124:**4212-4218.

74. Looft T, Johnson TA, Allen HK, Bayles DO, Alt DP, Stedtfeld RD, Sul WJ, Stedtfeld TM, Chai B, Cole JR: **In-feed antibiotic effects on the swine intestinal microbiome.** *Proceedings of the National Academy of Sciences* 2012, **109:**1691-1696.

75. Arboleya S, Sánchez B, Solís G, Fernández N, Suárez M, Hernández-Barranco A, Milani C, Margolles A, de los Reyes-Gavilán C, Ventura M: **Impact of prematurity and perinatal antibiotics on the developing intestinal microbiota: a functional inference study.** *International Journal of Molecular Sciences* 2016, **17:**649.

76. Patterson E, Cryan J, Fitzgerald G, Ross R, Dinan T, Stanton C: **Gut microbiota, the pharmabiotics they produce and host health.** *The Proceedings of the Nutrition Society* 2014**:**1-13.

77. Kabat AM, Srinivasan N, Maloy KJ: **Modulation of immune development and function by intestinal microbiota.** *Trends in Immunology* 2014.

78. Clemente JC, Ursell LK, Parfrey LW, Knight R: **The impact of the gut microbiota on human health: an integrative view.** *Cell* 2012, **148:**1258-1270.

79. Hansen J, Gulati A, Sartor RB: **The role of mucosal immunity and host genetics in defining intestinal commensal bacteria.** *Current Opinion in Gastroenterology* 2010, **26:**564.

80. David LA, Materna AC, Friedman J, Campos-Baptista MI, Blackburn MC, Perrotta A, Erdman SE, Alm EJ: **Host lifestyle affects human microbiota on daily timescales.** *Genome Biology* 2014, **15:**R89.

81. Maukonen J, Saarela M: **Human gut microbiota: does diet matter?** *The Proceedings of the Nutrition Society* 2014**:**1-14.

82. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R: **Evolution of mammals and their gut microbes.** *Science* 2008, **320:**1647-1651.

83. Moeller AH, Caro-Quintero A, Mjungu D, Georgiev AV, Lonsdorf EV, Muller MN, Pusey AE, Peeters M, Hahn BH, Ochman H: **Cospeciation of gut microbiota with hominids.** *Science* 2016, **353:**380-382.

84. Ellis RJ, Bruce KD, Jenkins C, Stothard JR, Ajarova L, Mugisha L, Viney ME: **Comparison of the distal gut microbiota from people and animals in Africa.** *PloS One* 2013, **8**.

85. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turroni S, Biagi E, Peano C, Severgnini M: **Gut microbiome of the Hadza hunter-gatherers.** *Nature Communications* 2014, **5**.

86. Moeller AH, Li Y, Ngole EM, Ahuka-Mundeke S, Lonsdorf EV, Pusey AE, Peeters M, Hahn BH, Ochman H: **Rapid changes in the gut microbiome during human evolution.** *Proceedings of the National Academy of Sciences* 2014, **111:**16431-16435.

87. Ren T, Grieneisen LE, Alberts SC, Archie EA, Wu M: **Development, diet and dynamism: longitudinal and cross-sectional predictors of gut microbial communities in wild baboons.** *Environmental Microbiology* 2015.

88. Amato KR, Leigh SR, Kent A, Mackie RI, Yeoman CJ, Stumpf RM, Wilson BA, Nelson KE, White BA, Garber PA: **The Gut Microbiota Appears to Compensate for Seasonal Diet Variation in the Wild Black Howler Monkey (Alouatta pigra).** *Microbial Ecology* 2015, **69:**434-443.

89. Yildirim S, Yeoman CJ, Sipos M, Torralba M, Wilson BA, Goldberg TL, Stumpf RM, Leigh SR, White BA, Nelson KE: **Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities.** *PloS One* 2010, **5:**e13963.

90. Balko EA, Brian Underwood H: **Effects of forest structure and composition on food availability for Varecia variegata at Ranomafana National Park, Madagascar.** *American Journal of Primatology* 2005, **66:**45-70.

91. Overdorff DJ, Strait SG, Telo A: **Seasonal variation in activity and diet in a small-bodied folivorous primate, Hapalemur griseus, in southeastern Madagascar.** *American Journal of Primatology* 1997, **43:**211-223.

92. Overdorff DJ: **Similarities, differences, and seasonal patterns in the diets ofEulemur rubriventer andEulemur fulvus rufus in the Ranomafana National Park, Madagascar.** *International Journal of Primatology* 1993, **14:**721-753.

93. McKenney EA, Rodrigo A, Yoder AD: **Patterns of Gut Bacterial Colonization in Three Primate Species.** 2015.

94. Aivelo T, Laakkonen J, Jernvall J: **Population-and Individual-Level Dynamics of the Intestinal Microbiota of a Small Primate.** *Applied and environmental microbiology* 2016, **82:**3537-3545.

95. Fogel AT: **The gut microbiome of wild lemurs: a comparison of sympatric Lemur catta and Propithecus verreauxi.** *Folia Primatologica* 2015, **86:**85-95.

96. Irwin MT, Wright PC, Birkinshaw C, Fisher BL, Gardner CJ, Glos J, Goodman SM, Loiselle P, Rabeson P, Raharison J-L: **Patterns of species change in anthropogenically disturbed forests of Madagascar.** *Biological Conservation* 2010, **143:**2351-2362.

97. Goodman SM, Benstead JP: **Updated estimates of biotic diversity and endemism for Madagascar.** *Oryx* 2005, **39:**73-77.

98. Lewis RJ, Bannar-Martin KH: **The impact of cyclone Fanele on a tropical dry forest in Madagascar.** *Biotropica* 2012, **44:**135-140.

99. García G, Goodman SM: **Hunting of protected animals in the Parc National d'Ankarafantsika, north-western Madagascar.** *Oryx* 2003, **37:**115-118.

100. Sorg J, Ganzhorn J, Kappeler P: **Forestry and research in the Kirindy Forest/Centre de Formation Professionnelle Forestière.** *The Natural History of Madagascar* 2003**:**1512-1519.

101. Ganzhorn JU, Schmid J: **Different population dynamics of Microcebus murinus in primary and secondary deciduous dry forests of Madagascar.** *International Journal of Primatology* 1998, **19:**785-796.

102. Rakotonirina: **Composition and structure of a dry forest on sandy soils near Morondava.** *Primate Rep* 1996, **:**81–87.

103. Wright PC, Erhart EM, Tecot S, Baden AL, Arrigo-Nelson SJ, Herrera J, Morelli TL, Blanco MB, Deppe A, Atsalis S: **Long-term lemur research at Centre ValBio, Ranomafana National Park, Madagascar.** In *Long-term Field Studies of Primates.* Springer; 2012: 67-100

104. Atsalis S: **Spatial distribution and population composition of the brown mouse lemur (Microcebus rufus) in Ranomafana National Park, Madagascar, and its implications for social organization.** *American Journal of Primatology* 2000, **51:**61-78.

105. Vences JK-FG-M, Köhler J, Glaw F: **First record of Mabuya comorensis (Reptilia: Scincidae) for the Madagascan fauna, with notes on the reptile fauna of the offshore island Nosy Tanikely.** *Museo Regionale di Scienze Naturali* 1997**:**15-N11.

106. Sato H: **Seasonal fruiting and seed dispersal by the brown lemur in a tropical dry forest, north-western Madagascar.** *Journal of Tropical Ecology* 2013, **29:**61-69.

107. Pyritz LW, Kappeler PM, Fichtel C: **Coordination of group movements in wild red-fronted lemurs (Eulemur rufifrons): processes and influence of**

ecological and reproductive seasonality. *International Journal of Primatology* 2011, **32:**1325-1347.

108. Sussman R: **Ecological distinctions in sympatric species of Lemur.** *Prosimian Biology* 1974**:**75-108.

109. Overdorff DJ: **Ecological correlates to activity and habitat use of two prosimian primates: Eulemur rubriventer and Eulemur fulvus rufus in Madagascar.** *American Journal of Primatology* 1996, **40:**327-342.

110. Johnson SE, Gordon AD, Stumpf RM, Overdorff DJ, Wright PC: **Morphological variation in populations of Eulemur albocollaris and E. fulvus rufus.** *International Journal of Primatology* 2005, **26:**1399-1416.

111. Tecot SR: **Seasonality and predictability: the hormonal and behavioral responses of the red-bellied lemur, Eulemur rubriventer, in southeastern Madagascar.** 2008.

112. Tecot SR: **It's all in the timing: birth seasonality and infant survival in Eulemur rubriventer.** *International Journal of Primatology* 2010, **31:**715-735.

113. Yu Z, Morrison M: **Improved extraction of PCR-quality community DNA from digesta and fecal samples.** *Biotechniques* 2004, **36:**808-812.

114. Salonen A, Nikkilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M, Kekkonen RA, Palva A, de Vos WM: **Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis.** *Journal of Microbiological Methods* 2010, **81:**127-134.

115. van den Bogert B, de Vos WM, Zoetendal EG, Kleerebezem M: **Microarray analysis and barcoded pyrosequencing provide consistent microbial profiles depending on the source of human intestinal samples.** *Applied and Environmental Microbiology* 2011, **77:**2071-2080.

116. Daims H, Brühl A, Amann R, Schleifer K-H, Wagner M: **The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set.** *Systematic and Applied Microbiology* 1999, **22:**434-444.

117. van den Bogert B, Erkus O, Boekhorst J, de Goffau M, Smid EJ, Zoetendal EG, Kleerebezem M: **Diversity of human small intestinal Streptococcus and Veillonella populations.** *FEMS Microbiology Ecology* 2013, **85:**376-388.

118. Tian L, Scholte J, Borewicz K, van den Bogert B, Smidt H, Scheurink AJ, Gruppen H, Schols HA: **Effects of pectin supplementation on the**

fermentation patterns of different structural carbohydrates in rats. *Molecular Nutrition & Food Research* 2016, **60:**2256-2266.

119. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO: **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.** *Nucleic Acids Research* 2013, **41:**D590-D596.
120. Paulson JN, Stine OC, Bravo HC, Pop M: **Differential abundance analysis for microbial marker-gene surveys.** *Nature Methods* 2013, **10:**1200-1202.
121. Šmilauer P, Lepš J: *Multivariate analysis of ecological data using CANOCO 5*. Cambridge university press; 2014.
122. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C: **Metagenomic biomarker discovery and explanation.** *Genome Biol* 2011, **12:**R60.
123. Gomez A, Petrzelkova K, Yeoman CJ, Vlckova K, Mrázek J, Koppova I, Carbonero F, Ulanov A, Modry D, Todd A: **Gut microbiome composition and metabolomic profiles of wild western lowland gorillas (Gorilla gorilla gorilla) reflect host ecology.** *Molecular Ecology* 2015, **24:**2551-2565.
124. Ochman H, Worobey M, Kuo C-H, Ndjango J-BN, Peeters M, Hahn BH, Hugenholtz P: **Evolutionary relationships of wild hominids recapitulated by gut microbial communities.** *PLoS Biology* 2010, **8:**2638.
125. Ni X, Gebo DL, Dagosto M, Meng J, Tafforeau P, Flynn JJ, Beard KC: **The oldest known primate skeleton and early haplorhine evolution.** *Nature* 2013, **498:**60-64.
126. Martínez I, Lattimer JM, Hubach KL, Case JA, Yang J, Weber CG, Louk JA, Rose DJ, Kyureghian G, Peterson DA: **Gut microbiome composition is linked to whole grain-induced immunological improvements.** *The ISME Journal* 2013, **7:**269-280.
127. Ze X, Duncan SH, Louis P, Flint HJ: **Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon.** *The ISME Journal* 2012, **6:**1535-1543.
128. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, Lionetti P: **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* 2010, **107:**14691-14696.

129. Gutleben J, Chaib De Mares M, van Elsas JD, Smidt H, Overmann J, Sipkema D: **The multi-omics promise in context: from sequence to microbial isolate.** *Critical Reviews in Microbiology* 2017**:**1-18.

130. Shin N-R, Whon TW, Bae J-W: **Proteobacteria: microbial signature of dysbiosis in gut microbiota.** *Trends in Biotechnology* 2015.

131. Bello González T, van Passel M, Tims S, Fuentes S, De Vos W, Smidt H, Belzer C: **Application of the Human Intestinal Tract Chip to the non-human primate gut microbiota.** *Beneficial Microbes* 2015, **6:**271-276.

132. Di Rienzi SC, Sharon I, Wrighton KC, Koren O, Hug LA, Thomas BC, Goodrich JK, Bell JT, Spector TD, Banfield JF: **The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria.** *Elife* 2013, **2:**e01102.

133. Roggenbuck M, Sauer C, Poulsen M, Bertelsen MF, Sørensen SJ: **The giraffe (Giraffa camelopardalis) rumen microbiome.** *FEMS Microbiology Ecology* 2014, **90:**237-246.

134. Wilmé L, Goodman SM, Ganzhorn JU: **Biogeographic evolution of Madagascar's microendemic biota.** *Science* 2006, **312:**1063-1065.

135. Sorg J, Rohner U: **Climate and tree phenology of the dry deciduous forest of the Kirindy Forest.** *Primate Report* 1996, **46:**57-80.

136. Ganzhorn JU: **Distribution of a folivorous lemur in relation to seasonally varying food resources: integrating quantitative and qualitative aspects of food characteristics.** *Oecologia* 2002, **131:**427-435.

137. Khachatryan ZA, Ktsoyan ZA, Manukyan GP, Kelly D, Ghazaryan KA, Aminov RI: **Predominant role of host genetics in controlling the composition of gut microbiota.** *PloS One* 2008, **3:**e3064.

138. Markolf M, Kappeler PM: **Phylogeographic analysis of the true lemurs (genus Eulemur) underlines the role of river catchments for the evolution of micro-endemism in Madagascar.** *Frontiers in Zoology* 2013, **10:**70.

139. Tattersall I: **Speciation and morphological differentiation in the genus Lemur.** In *Species, Species Concepts and Primate Evolution.* Springer; 1993: 163-176

140. Wopereis H, Oozeer R, Knipping K, Belzer C, Knol J: **The first thousand days–intestinal microbiology of early life: establishing a symbiosis.** *Pediatric Allergy and Immunology* 2014, **25:**428-438.

141. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE: **Succession of microbial consortia in the developing infant gut microbiome.** *Proceedings of the National Academy of Sciences* 2011, **108:**4578-4585.

142. Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Org E, Parks B, Lusis AJ, Knight R, Caporaso JG, Svanbäck R: **Individual diet has sex-dependent effects on vertebrate gut microbiota.** *Nature Communications* 2014, **5**.

143. Kovacs A, Ben-Jacob N, Tayem H, Halperin E, Iraqi FA, Gophna U: **Genotype is a stronger determinant than sex of the mouse gut microbiota.** *Microbial Ecology* 2011, **61:**423-428.

144. Pfeiffer JK, Virgin HW: **Transkingdom control of viral infection and immunity in the mammalian intestine.** *Science* 2016, **351:**aad5872.

145. Kabat AM, Srinivasan N, Maloy KJ: **Modulation of immune development and function by intestinal microbiota.** *Trends in immunology* 2014, **35:**507-517.

146. Sekirov I, Russell SL, Antunes LCM, Finlay BB: **Gut microbiota in health and disease.** *Physiological Reviews* 2010, **90:**859-904.

147. de Vos WM, de Vos EA: **Role of the intestinal microbiome in health and disease: from correlation to causation.** *Nutrition reviews* 2012, **70:**S45-S56.

148. Patterson E, Cryan JF, Fitzgerald GF, Ross RP, Dinan TG, Stanton C: **Gut microbiota, the pharmabiotics they produce and host health.** *Proceedings of the Nutrition Society* 2014, **73:**477-489.

149. Nunn CL, Thrall PH, Leendertz FH, Boesch C: **The spread of fecally transmitted parasites in socially-structured populations.** *PloS One* 2011, **6:**e21677.

150. Ramanan D, Bowcutt R, Lee SC, San Tang M, Kurtz ZD, Ding Y, Honda K, Gause WC, Blaser MJ, Bonneau RA: **Helminth infection promotes colonization resistance via type 2 immunity.** *Science* 2016, **352:**608-612.

151. Leclaire S, Faulkner CT: **Gastrointestinal parasites in relation to host traits and group factors in wild meerkats Suricata suricatta.** *Parasitology* 2014, **141:**925-933.

152. Bennett G, Malone M, Sauther ML, Cuozzo FP, White B, Nelson KE, Stumpf RM, Knight R, Leigh SR, Amato KR: **Host age, social group, and habitat type influence the gut microbiota of wild ring-tailed lemurs (Lemur catta).** *American Journal of Primatology* 2016, **78:**883-892.

153. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA: **Diversity of the human intestinal microbial flora.** *Science* 2005, **308:**1635-1638.

154. Tompkins DM, Dunn AM, Smith MJ, Telfer S: **Wildlife diseases: from individuals to ecosystems.** *Journal of Animal Ecology* 2011, **80:**19-38.

155. McKenney EA, Rodrigo A, Yoder AD: **Patterns of gut bacterial colonization in three primate species.** *PloS one* 2015, **10:**e0124618.

156. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R: **Diversity, stability and resilience of the human gut microbiota.** *Nature* 2012, **489:**220.

157. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, Henrissat B, Knight R, Gordon JI: **Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans.** *Science* 2011, **332:**970-974.

158. Boutin S, Bernatchez L, Audet C, Derôme N: **Network analysis highlights complex interactions between pathogen, host and commensal microbiota.** *PLoS One* 2013, **8:**e84772.

159. Maurice CF, Knowles SC, Ladau J, Pollard KS, Fenton A, Pedersen AB, Turnbaugh PJ: **Marked seasonal variation in the wild mouse gut microbiota.** *The ISME Journal* 2015, **9:**2423.

160. Dishaw LJ, Flores-Torres J, Lax S, Gemayel K, Leigh B, Melillo D, Mueller MG, Natale L, Zucchetti I, De Santis R: **The gut of geographically disparate Ciona intestinalis harbors a core microbiota.** *PLoS One* 2014, **9:**e93386.

161. Moore SL, Wilson K: **Parasites as a viability cost of sexual selection in natural populations of mammals.** *Science* 2002, **297:**2015-2018.

162. Huffman MA, Gotoh S, Turner LA, Hamai M, Yoshida K: **Seasonal trends in intestinal nematode infection and medicinal plant use among chimpanzees in the Mahale Mountains, Tanzania.** *Primates* 1997, **38:**111-125.

163. Benavides JA, Huchard E, Pettorelli N, King AJ, Brown ME, Archer CE, Appleton CC, Raymond M, Cowlishaw G: **From parasite encounter to infection: Multiple-scale drivers of parasite richness in a wild social primate population.** *American Journal of Physical Anthropology* 2012, **147:**52-63.

164. Barrett MA, Brown JL, Junge RE, Yoder AD: **Climate change, predictive modeling and lemur health: assessing impacts of changing climate on**

health and conservation in Madagascar. *Biological Conservation* 2013, **157:**409-422.

165. Aivelo T, Laakkonen J, Jernvall J: **Population-and individual-level dynamics of the intestinal microbiota of a small primate.** *Appl Environ Microbiol* 2016, **82:**3537-3545.

166. Nunn C, Altizer S, Altizer SM: *Infectious diseases in primates: behavior, ecology and evolution*. Oxford University Press; 2006.

167. Raharivololona BM, Ganzhorn JU: **Seasonal variations in gastrointestinal parasites excreted by the gray mouse lemur Microcebus murinus in Madagascar.** *Endangered Species Research* 2010, **11:**113-122.

168. Huffman MA, Chapman CA: *Primate parasite ecology: the dynamics and study of host-parasite relationships*. 2009.

169. Setchell JM, Bedjabaga I-B, Goossens B, Reed P, Wickings EJ, Knapp LA: **Parasite prevalence, abundance, and diversity in a semi-free-ranging colony of Mandrillus sphinx.** *International Journal of Primatology* 2007, **28:**1345-1362.

170. Maldonado-López S, Maldonado-Lopez Y, Ch AG-T, Cuevas-Reyes P, Stoner KE: **Patterns of infection by intestinal parasites in sympatric howler monkey (Alouatta palliata) and spider monkey (Ateles geoffroyi) populations in a tropical dry forest in Costa Rica.** *Primates* 2014, **55:**383-392.

171. Caldwell J: **Pinworms (enterobius vermicularis).** *Canadian Family Physician* 1982, **28:**306.

172. Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JB: **Beyond biogeographic patterns: processes shaping the microbial landscape.** *Nature Reviews Microbiology* 2012, **10:**497.

173. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain LV, Liegeois F: **Chimpanzee reservoirs of pandemic and nonpandemic HIV-1.** *Science* 2006, **313:**523-526.

174. Gillespie TR, Chapman CA, Greiner EC: **Effects of logging on gastrointestinal parasite infections and infection risk in African primates.** *Journal of Applied Ecology* 2005, **42:**699-707.

175. Chapman CA, Gillespie TR, Goldberg TL: **Primates and the ecology of their infectious diseases: how will anthropogenic change affect host-parasite interactions?** *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews* 2005, **14:**134-144.

176. McCord AI, Chapman CA, Weny G, Tumukunde A, Hyeroba D, Klotz K, Koblings AS, Mbora DN, Cregger M, White BA: **Fecal microbiomes of non-human primates in Western Uganda reveal species-specific communities largely resistant to habitat perturbation.** *American Journal of Primatology* 2014, **76:**347-354.

177. Chapman CA, Speirs ML, Gillespie TR, Holland T, Austad KM: **Life on the edge: gastrointestinal parasites from the forest edge and interior primate groups.** *American Journal of Primatology: Official Journal of the American Society of Primatologists* 2006, **68:**397-409.

178. Kowalewski MM, Salzer JS, Deutsch JC, Raño M, Kuhlenschmidt MS, Gillespie TR: **Black and gold howler monkeys (Alouatta caraya) as sentinels of ecosystem health: patterns of zoonotic protozoa infection relative to degree of human–primate contact.** *American Journal of Primatology* 2011, **73:**75-83.

179. Chapman CA, Wasserman MD, Gillespie TR, Speirs ML, Lawes MJ, Saj TL, Ziegler TE: **Do food availability, parasitism, and stress have synergistic effects on red colobus populations living in forest fragments?** *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 2006, **131:**525-534.

180. Hughes S, Kelly P: **Interactions of malnutrition and immune impairment, with specific reference to immunity against parasites.** *Parasite Immunology* 2006, **28:**577.

181. Angelstam P, Roberge J-M, Lõhmus A, Bergmanis M, Brazaitis G, Dönz-Breuss M, Edenius L, Kosinski Z, Kurlavicius P, Lārmanis V: **Habitat modelling as a tool for landscape-scale conservation: a review of parameters for focal forest birds.** *Ecological Bulletins* 2004**:**427-453.

182. Kreisinger J, Bastien G, Hauffe HC, Marchesi J, Perkins SE: **Interactions between multiple helminths and the gut microbiota in wild rodents.** *Philosophical Transactions of the Royal Society B: Biological Sciences* 2015, **370:**20140295.

183. Mutapi F: **The gut microbiome in the helminth infected host.** *Trends in Parasitology* 2015, **31:**405-406.

184. Kay GL, Millard A, Sergeant MJ, Midzi N, Gwisai R, Mduluza T, Ivens A, Nausch N, Mutapi F, Pallen M: **Differences in the faecal microbiome in**

Schistosoma haematobium infected children vs. uninfected children. *PLoS Neglected Tropical Diseases* 2015, **9:**e0003861.

185. Lee SC, San Tang M, Lim YA, Choy SH, Kurtz ZD, Cox LM, Gundra UM, Cho I, Bonneau R, Blaser MJ: **Helminth colonization is associated with increased diversity of the gut microbiota.** *PLoS Neglected Tropical Diseases* 2014, **8:**e2880.

186. Morton E, Lynch J, Froment A, Lafosse S, Heyer E, Przeworski M, Blekhman R, Ségurel L: **Variation in rural African gut microbiomes is strongly shaped by parasitism and diet.** *bioRxiv* 2015**:**016949.

187. Cooper P, Walker AW, Reyes J, Chico M, Salter SJ, Vaca M, Parkhill J: **Patent human infections with the whipworm, Trichuris trichiura, are not associated with alterations in the faecal microbiota.** *PloS One* 2013, **8:**e76573.

188. Cantacessi C, Giacomin P, Croese J, Zakrzewski M, Sotillo J, McCann L, Nolan MJ, Mitreva M, Krause L, Loukas A: **Impact of experimental hookworm infection on the human gut microbiota.** *The Journal of Infectious Diseases* 2014, **210:**1431-1434.

189. Houlden A, Hayes KS, Bancroft AJ, Worthington JJ, Wang P, Grencis RK, Roberts IS: **Chronic Trichuris muris infection in C57BL/6 mice causes significant changes in host microbiota and metabolome: effects reversed by pathogen clearance.** *PloS One* 2015, **10:**e0125945.

190. McKenney EA, Greene LK, Drea CM, Yoder AD: **Down for the count: Cryptosporidium infection depletes the gut microbiome in Coquerel's sifakas.** *Microbial Ecology in Health and Disease* 2017, **28:**1335165.

191. Springer A, Fichtel C, Al-Ghalith GA, Koch F, Amato KR, Clayton JB, Knights D, Kappeler PM: **Patterns of seasonality and group membership characterize the gut microbiota in a longitudinal study of wild Verreaux's sifakas (Propithecus verreauxi).** *Ecology and Evolution* 2017, **7:**5732-5745.

192. Walk ST, Blum AM, Ewing SA-S, Weinstock JV, Young VB: **Alteration of the murine gut microbiota during infection with the parasitic helminth Heligmosomoides polygyrus.** *Inflammatory Bowel Diseases* 2010, **16:**1841-1849.

193. Li RW, Wu S, Li W, Navarro K, Couch RD, Hill D, Urban JF: **Alterations in the porcine colon microbiota induced by the gastrointestinal nematode Trichuris suis.** *Infection and Immunity* 2012, **80:**2150-2157.

194. Rausch S, Held J, Fischer A, Heimesaat MM, Kühl AA, Bereswill S, Hartmann S: **Small intestinal nematode infection of mice is associated with increased enterobacterial loads alongside the intestinal tract.** *PloS One* 2013, **8:**e74026.

195. Irwin MT, Johnson SE, Wright PC: **The state of lemur conservation in south-eastern Madagascar: population and habitat assessments for diurnal and cathemeral lemurs using surveys, satellite imagery and GIS.** *Oryx* 2005, **39:**204-218.

196. Markolf M, Rakotonirina H, Fichtel C, von Grumbkow P, Brameier M, Kappeler PM: **True lemurs… true species-species delimitation using multiple data sources in the brown lemur complex.** *BMC Evolutionary Biology* 2013, **13:**233.

197. Goodman SM, Andriafidison D, Andrianaivoarivelo R, Cardiff SG, Ifticene E, Jenkins RK, Kofoky A, Mbohoahy T, Rakotondravony D, Ranivo J: **The distribution and conservation of bats in the dry regions of Madagascar.** In *Animal Conservation Forum*. Cambridge University Press; 2005: 153-165.

198. Herrera JP, Wright PC, Lauterbur E, Ratovonjanahary L, Taylor LL: **The effects of habitat disturbance on lemurs at Ranomafana National Park, Madagascar.** *International Journal of Primatology* 2011, **32:**1091-1108.

199. Balko EA, Brian Underwood H: **Effects of forest structure and composition on food availability for Varecia variegata at Ranomafana National Park, Madagascar.** *American Journal of Primatology: Official Journal of the American Society of Primatologists* 2005, **66:**45-70.

200. Erhart EM, Overdorff DJ: **Population demography and social structure changes in Eulemur fulvus rufus from 1988 to 2003.** *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* 2008, **136:**183-193.

201. Mittermeier RA, Ganzhorn JU, Konstant WR, Glander K, Tattersall I, Groves CP, Rylands AB, Hapke A, Ratsimbazafy J, Mayor MI: **Lemur diversity in Madagascar.** *International Journal of Primatology* 2008, **29:**1607-1656.

202. Umanets A, de Winter I, IJdema F, Ramiro-Garcia J, van Hooft P, Heitkönig IM, Prins HH, Smidt H: **Occupancy strongly influences faecal microbial composition of wild lemurs.** *FEMS Microbiology Ecology* 2018, **94:**fiy017.

203. Berg W, Jolly A, Rambeloarivony H, Andrianome V, Rasamimanana H: **A scoring system for coat and tail condition in ringtailed lemurs, Lemur**

catta. *American Journal of Primatology: Official Journal of the American Society of Primatologists* 2009, **71:**183-190.

204. Van Gool T, Weijts R, Lommerse E, Mank T: **Triple faeces test: an effective tool for detection of intestinal parasites in routine clinical practice.** *European Journal of Clinical Microbiology and Infectious Diseases* 2003, **22:**284-290.

205. Dryden MW, Payne PA, Ridley R, Smith V: **Comparison of common fecal flotation techniques for the recovery of parasite eggs and oocysts.** *Vet Ther* 2005, **6:**15-28.

206. Irwin MT, Raharison J-L: **A review of the endoparasites of the lemurs of Madagascar.** *Malagasy Nat* 2009, **2:**66-93.

207. Clough D: **Gastro-intestinal parasites of red-fronted lemurs in Kirindy Forest, Western Madagascar.** *Journal of Parasitology* 2010, **96:**245-252.

208. Gillespie TR: **Noninvasive assessment of gastrointestinal parasite infections in free-ranging primates.** *International Journal of Primatology* 2006, **27:**1129.

209. Gillespie TR, Chapman CA: **Prediction of parasite infection dynamics in primate metapopulations based on attributes of forest fragmentation.** *Conservation Biology* 2006, **20:**441-448.

210. McMurdie PJ, Holmes S: **phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data.** *PloS One* 2013, **8:**e61217.

211. Schliep K, Potts AJ, Morrison DA, Grimm GW: **Intertwining phylogenetic trees and networks.** *Methods in Ecology and Evolution* 2017, **8:**1212-1220.

212. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, et al: **vegan: Community Ecology Package.** *https://CRANR-projectorg/package=vegan* 2018, **10:**631-637.

213. Pohlert T: **The pairwise multiple comparison of mean ranks package (PMCMR).** *R package* 2014, **27**.

214. Zeger SL, Liang K-Y, Albert PS: **Models for longitudinal data: a generalized estimating equation approach.** *Biometrics* 1988**:**1049-1060.

215. Scrucca L: **dispmod: Dispersion models. R package version 1.1.** 2012.

216. Nakagawa S, Schielzeth H: **A general and simple method for obtaining R2 from generalized linear mixed-effects models.** *Methods in Ecology and Evolution* 2013, **4:**133-142.

217. Team RC: **R: A language and environment for statistical computing.** 2013.

218. Bates D, Mächler M, Bolker B, Walker S: **Fitting linear mixed-effects models using lme4.** *arXiv preprint arXiv:14065823* 2014.

219. Lenth RV: **Least-squares means: the R package lsmeans.** *Journal of Statistical Software* 2016, **69:**1-33.

220. Fox J, Friendly M, Weisberg S: **Hypothesis tests for multivariate linear models using the car package.** *The R Journal* 2013, **5:**39-52.

221. Hartig F: **DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models.** *R package version 01* 2017, **5**.

222. Barton K: **MuMIn: multi-model inference. R package version 1. 0. 0.** *http://r-forge r-project org/projects/mumin/* 2009.

223. Chabaud A, Petter AJ: **Les nématodes parasites de lémuriens malgaches. II. Un nouvel oxyure: Lemuricola contagiosus.** *Mémoires de l'Institut Scientifique de Madagasacar* 1959, **13:**127-132.

224. Chabaud A-G, Brygoo E-R, Petter A-J: **Les Nématodes parasites de Lémuriens malgaches-VI.—Description de six espèces nouvelles et conclusions générales.** *Annales de Parasitologie Humaine et Comparée* 1965, **40:**181-214.

225. Schwitzer N, Clough D, Zahner H, Kaumanns W, Kappeler P, Schwitzer C: **Parasite prevalence in blue-eyed black lemurs Eulemur flavifrons in differently degraded forest fragments.** *Endangered Species Research* 2010, **12:**215-225.

226. Junge RE, Louis EE: **Biomedical evaluation of black lemurs (Eulemur macaco macaco) in Lokobe Reserve, Madagascar.** *Journal of Zoo and Wildlife Medicine* 2007, **38:**67-77.

227. Nègre A, Tarnaud L, Roblot J, Gantier J-C, Guillot J: **Plants consumed by Eulemur fulvus in Comoros Islands (Mayotte) and potential effects on intestinal parasites.** *International Journal of Primatology* 2006, **27:**1495-1517.

228. Junge RE, Dutton CJ, Knightly F, Williams CV, Rasambainarivo FT, Louis EE: **Comparison of biomedical evaluation for white-fronted brown lemurs (Eulemur fulvus albifrons) from four sites in Madagascar.** *Journal of Zoo and Wildlife Medicine* 2008, **39:**567-576.

229. Rajilić-Stojanović M, Heilig HG, Tims S, Zoetendal EG, de Vos WM: **Long-term monitoring of the human intestinal microbiota composition.** *Environmental Microbiology* 2013, **15:**1146-1159.

230. Crowley BE, McGoogan KC, Lehman SM: **Edge effects on foliar stable isotope values in a Madagascan tropical dry forest.** *PLoS One* 2012, **7:**e44538.

231. Sato H, Ichino S, Hanya G: **Dietary modification by common brown lemurs (Eulemur fulvus) during seasonal drought conditions in western Madagascar.** *Primates* 2014, **55:**219-230.

232. Styger E, Rakotoarimanana J, Rabevohitra R, Fernandes E: **Indigenous fruit trees of Madagascar: potential components of agroforestry systems to improve human nutrition and restore biological diversity.** *Agroforestry Systems* 1999, **46:**289-310.

233. Sato H: **Habitat shifting by the common brown lemur (Eulemur fulvus fulvus): a response to food scarcity.** *Primates* 2013, **54:**229-235.

234. Guernier V, Hochberg ME, Guégan J-F: **Ecology drives the worldwide distribution of human diseases.** *PLoS Biology* 2004, **2:**e141.

235. Froeschke G, Harf R, Sommer S, Matthee S: **Effects of precipitation on parasite burden along a natural climatic gradient in southern Africa–implications for possible shifts in infestation patterns due to global changes.** *Oikos* 2010, **119:**1029-1039.

236. Luong LT, Grear DA, Hudson PJ: **Manipulation of host-resource dynamics impacts transmission of trophic parasites.** *International Journal for Parasitology* 2014, **44:**737-742.

237. Wright PC, Razafindratsita VR, Pochron ST, Jernvall J: **The key to Madagascar frugivores.** In *Tropical Fruits and Frugivores.* Springer; 2005: 121-138

238. Krause D: **The Natural History of Madagascar, eds Goodman SM, Benstead JP.** Univ of Chicago Press, Chicago; 2003.

239. Ostner J, Kappeler P, Heistermann M: **Androgen and glucocorticoid levels reflect seasonally occurring social challenges in male redfronted lemurs (Eulemur fulvus rufus).** *Behavioral Ecology and Sociobiology* 2008, **62:**627-638.

240. Johns AD, Skorupa JP: **Responses of rain-forest primates to habitat disturbance: a review.** *International Journal of Primatology* 1987, **8:**157.

241. Terborgh J, van Schaik C, Davenport L, Rao M: *Making parks work: strategies for preserving tropical nature.* Island Press; 2002.

242. Vitazkova SK, Wade SE: **Effects of ecology on the gastrointestinal parasites of Alouatta pigra.** *International Journal of Primatology* 2007, **28:**1327-1343.

243. Martinez Mota R: **The effects of habitat disturbance, host traits, and host physiology on patterns of gastrointestinal parasite infection in black howler monkeys (Alouatta pigra).** University of Illinois at Urbana-Champaign, 2015.

244. Arneberg P: **Host population density and body mass as determinants of species richness in parasite communities: comparative analyses of directly transmitted nematodes of mammals.** *Ecography* 2002, **25:**88-94.

245. Reynolds LA, Smith KA, Filbey KJ, Harcus Y, Hewitson JP, Redpath SA, Valdez Y, Yebra MJ, Finlay BB, Maizels RM: **Commensal-pathogen interactions in the intestinal tract: lactobacilli promote infection with, and are promoted by, helminth parasites.** *Gut Microbes* 2014, **5:**522-532.

246. Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, Mujagic Z, Vila AV, Falony G, Vieira-Silva S: **Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity.** *Science* 2016, **352:**565-569.

247. Hayes K, Bancroft A, Goldrick M, Portsmouth C, Roberts I, Grencis R: **Exploitation of the intestinal microflora by the parasitic nematode Trichuris muris.** *Science* 2010, **328:**1391-1394.

248. Reynolds LA, Finlay BB, Maizels RM: **Cohabitation in the intestine: interactions among helminth parasites, bacterial microbiota, and host immunity.** *The Journal of Immunology* 2015, **195:**4059-4066.

249. Pedersen AB, Altizer S, Poss M, Cunningham AA, Nunn CL: **Patterns of host specificity and transmission among parasites of wild primates.** *International Journal for Parasitology* 2005, **35:**647-657.

250. Ley RE, Turnbaugh PJ, Klein S, Gordon JI: **Microbial ecology: human gut microbes associated with obesity.** *Nature* 2006, **444:**1022.

251. Wen L, Ley RE, Volchkov PY, Stranges PB, Avanesyan L, Stonebraker AC, Hu C, Wong FS, Szot GL, Bluestone JA: **Innate immunity and intestinal microbiota in the development of Type 1 diabetes.** *Nature* 2008, **455:**1109.

252. Larsen N, Vogensen FK, Van Den Berg FW, Nielsen DS, Andreasen AS, Pedersen BK, Al-Soud WA, Sørensen SJ, Hansen LH, Jakobsen M: **Gut microbiota in**

human adults with type 2 diabetes differs from non-diabetic adults. *PloS One* 2010, **5:**e9085.

253. Sokol H, Seksik P, Rigottier-Gois L, Lay C, Lepage P, Podglajen I, Marteau P, Doré J: **Specificities of the fecal microbiota in inflammatory bowel disease.** *Inflammatory Bowel Diseases* 2006, **12:**106-111.

254. Kassam Z, Lee CH, Yuan Y, Hunt RH: **Fecal microbiota transplantation for Clostridium difficile infection: systematic review and meta-analysis.** *The American Journal of Gastroenterology* 2013, **108:**500.

255. Tremlett H, Bauer KC, Appel-Cresswell S, Finlay BB, Waubant E: **The gut microbiome in human neurological disease: a review.** *Annals of Neurology* 2017, **81:**369-382.

256. Pascoe EL, Hauffe HC, Marchesi JR, Perkins SE: **Network analysis of gut microbiota literature: an overview of the research landscape in non-human animal studies.** *The ISME Journal* 2017, **11:**2644.

257. Heinritz SN, Mosenthin R, Weiss E: **Use of pigs as a potential model for research into dietary modulation of the human gut microbiota.** *Nutrition Research Reviews* 2013, **26:**191-209.

258. Leser TD, Mølbak L: **Better living through microbial action: the benefits of the mammalian gastrointestinal microbiota on the host.** *Environmental Microbiology* 2009, **11:**2194-2206.

259. Rojo-Gimeno C, Postma M, Dewulf J, Hogeveen H, Lauwers L, Wauters E: **Farm-economic analysis of reducing antimicrobial use whilst adopting improved management strategies on farrow-to-finish pig farms.** *Preventive Veterinary Medicine* 2016, **129:**74-87.

260. McCormack UM, Curião T, Buzoianu SG, Prieto ML, Ryan T, Varley P, Crispie F, Magowan E, Metzler-Zebeli BU, Berry D: **Exploring a possible link between the intestinal microbiota and feed efficiency in pigs.** *Appl Environ Microbiol* 2017, **83:**e00380-00317.

261. Yang H, Huang X, Fang S, He M, Zhao Y, Wu Z, Yang M, Zhang Z, Chen C, Huang L: **Unraveling the fecal microbiota and metagenomic functional capacity associated with feed efficiency in pigs.** *Frontiers in Microbiology* 2017, **8:**1555.

262. Roselli M, Pieper R, Rogel-Gaillard C, de Vries H, Bailey M, Smidt H, Lauridsen C: **Immunomodulating effects of probiotics for microbiota modulation,**

gut health and disease resistance in pigs. *Animal Feed Science and Technology* 2017, **233:**104-119.

263. Kritas SK: **Probiotics and Prebiotics for the Health of Pigs and Horses.** In *Probiotics and Prebiotics in Animal Health and Food Safety*. Springer; 2018: 109-126

264. Wang C, Shi C, Zhang Y, Song D, Lu Z, Wang Y: **Microbiota in fermented feed and swine gut.** *Applied Microbiology and Biotechnology* 2018, **102:**2941-2948.

265. Motta V, Luise D, Bosi P, Trevisi P: **Faecal microbiota shift during weaning transition in piglets and evaluation of AO blood types as shaping factor for the bacterial community profile.** *PloS One* 2019, **14:**e0217001.

266. Chen L, Xu Y, Chen X, Fang C, Zhao L, Chen F: **The maturing development of gut microbiota in commercial piglets during the weaning transition.** *Frontiers in Microbiology* 2017, **8:**1688.

267. Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, Zhang M, Oh PL, Nehrenberg D, Hua K: **Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors.** *Proceedings of the National Academy of Sciences* 2010, **107:**18933-18938.

268. Hufeldt MR, Nielsen DS, Vogensen FK, Midtvedt T, Hansen AK: **Variation in the gut microbiota of laboratory mice is related to both genetic and environmental factors.** *Comparative Medicine* 2010, **60:**336-347.

269. McKnite AM, Perez-Munoz ME, Lu L, Williams EG, Brewer S, Andreux PA, Bastiaansen JW, Wang X, Kachman SD, Auwerx J: **Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits.** *PloS One* 2012, **7:**e39191.

270. Luise D, Motta V, Bertocchi M, Salvarani C, Clavenzani P, Fanelli F, Pagotto U, Bosi P, Trevisi P: **Effect of Mucine 4 and Fucosyltransferase 1 genetic variants on gut homoeostasis of growing healthy pigs.** *Journal of Animal Physiology and Animal Nutrition* 2019.

271. Poulsen A-SR, Luise D, Curtasu MV, Sugiharto S, Canibe N, Trevisi P, Lauridsen C: **Effects of alpha-(1, 2)-fucosyltransferase genotype variants on plasma metabolome, immune responses and gastrointestinal bacterial enumeration of pigs pre-and post-weaning.** *PloS One* 2018, **13:**e0202970.

272. Pajarillo EAB, Chae J-P, Balolong MP, Kim HB, Kang D-K: **Assessment of fecal bacterial diversity among healthy piglets during the weaning transition.** *The Journal of General and Applied Microbiology* 2014, **60:**140-146.

273. Xiao Y, Li K, Xiang Y, Zhou W, Gui G, Yang H: **The fecal microbiota composition of boar Duroc, Yorkshire, Landrace and Hampshire pigs.** *Asian-Australasian Journal of Animal Sciences* 2017, **30:**1456.

274. Kanengoni AT, Chimonyo M, Tasara T, Cormican P, Chapwanya A, Ndimba BK, Dzama K: **A comparison of faecal microbial populations of South African Windsnyer-type indigenous pigs (SAWIPs) and Large White× Landrace (LW× LR) crosses fed diets containing ensiled maize cobs.** *FEMS Microbiology Letters* 2015, **362:**fnv100.

275. Xiao L, Estelle J, Kiilerich P, Ramayo-Caldas Y, Xia Z, Feng Q, Liang S, Pedersen AØ, Kjeldsen NJ, Liu C: **A reference gene catalogue of the pig gut microbiome.** *Nature Microbiology* 2016, **1:**16161.

276. Panasevich MR, Wankhade UD, Chintapalli SV, Shankar K, Rector RS: **Cecal versus fecal microbiota in Ossabaw swine and implications for obesity.** *Physiological Genomics* 2018, **50:**355-368.

277. Fontanesi L, Schiavo G, Galimberti G, Calo D, Russo V: **A genomewide association study for average daily gain in Italian Large White pigs.** *Journal of Animal Science* 2014, **92:**1385-1394.

278. Bertolini F, Schiavo G, Galimberti G, Bovo S, D'Andrea M, Gallo M, Buttazzoni L, Rothschild M, Fontanesi L: **Genome-wide association studies for seven production traits highlight genomic regions useful to dissect dry-cured ham quality and production traits in Duroc heavy pigs.** *Animal* 2018, **12:**1777-1784.

279. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO: **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.** *Nucleic Acids Research* 2012, **41:**D590-D596.

280. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO: **Picante: R tools for integrating phylogenies and ecology.** *Bioinformatics* 2010, **26:**1463-1464.

281. Gerds TA, Ozenne B: **Publish: Format Output of Various Routines in a Suitable Way for Reports and Publication.** R package version; 2017.

282. Kurtz ZD, Müller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA: **Sparse and compositionally robust inference of microbial ecological networks.** *PLoS Computational Biology* 2015, **11:**e1004226.

283. Csardi G, Nepusz T: **The igraph software package for complex network research.** *InterJournal, Complex Systems* 2006, **1695:**1-9.

284. Archer E: **rfPermute: estimate permutation p-values for Random Forest importance metrics.** *R package (Zenodo), Version* 2016, **2**.

285. Liaw A, Wiener M: **Classification and regression by randomForest.** *R News* 2002, **2:**18-22.

286. Sing T, Sander O, Beerenwinkel N, Lengauer T: **ROCR: visualizing classifier performance in R.** *Bioinformatics* 2005, **21:**3940-3941.

287. Wickham H: *ggplot2: elegant graphics for data analysis.* Springer; 2016.

288. Gu Z, Eils R, Schlesner M: **Complex heatmaps reveal patterns and correlations in multidimensional genomic data.** *Bioinformatics* 2016, **32:**2847-2849.

289. McCann KS: **The diversity–stability debate.** *Nature* 2000, **405:**228.

290. Ahn J, Sinha R, Pei Z, Dominianni C, Wu J, Shi J, Goedert JJ, Hayes RB, Yang L: **Human gut microbiome and risk for colorectal cancer.** *Journal of the National Cancer Institute* 2013, **105:**1907-1911.

291. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP: **A core gut microbiome in obese and lean twins.** *Nature* 2009, **457:**480.

292. Prehn-Kristensen A, Zimmermann A, Tittmann L, Lieb W, Schreiber S, Baving L, Fischer A: **Reduced microbiome alpha diversity in young patients with ADHD.** *PloS One* 2018, **13:**e0200728.

293. Dobbler PT, Procianoy RS, Mai V, Silveira RC, Corso AL, Rojas BS, Roesch LF: **Low microbial diversity and abnormal microbial succession is associated with necrotizing enterocolitis in preterm infants.** *Frontiers in Microbiology* 2017, **8:**2243.

294. Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, Schmidt TM, Young VB: **Decreased diversity of the fecal microbiome in recurrent Clostridium difficile—associated diarrhea.** *The Journal of Infectious Diseases* 2008, **197:**435-438.

295. Sommer F, Anderson JM, Bharti R, Raes J, Rosenstiel P: **The resilience of the intestinal microbiota influences health and disease.** *Nature Reviews Microbiology* 2017, **15:**630.

296. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez A, Kosciolek T, McCall L-I, McDonald D: **Best practices for analysing microbiomes.** *Nature Reviews Microbiology* 2018, **16:**410.

297. Cutler DR, Edwards Jr TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ: **Random forests for classification in ecology.** *Ecology* 2007, **88:**2783-2792.

298. Qi Y: **Random forest for bioinformatics.** In *Ensemble Machine Learning*. Springer; 2012: 307-323

299. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP: **Human gut microbiome viewed across age and geography.** *Nature* 2012, **486:**222.

300. Mach N, Berri M, Estellé J, Levenez F, Lemonnier G, Denis C, Leplat JJ, Chevaleyre C, Billon Y, Doré J: **Early-life establishment of the swine gut microbiome and impact on host phenotypes.** *Environmental Microbiology Reports* 2015, **7:**554-569.

301. Do DN, Strathe AB, Jensen J, Mark T, Kadarmideen HN: **Genetic parameters for different measures of feed efficiency and related traits in boars of three pig breeds.** *Journal of Animal Science* 2013, **91:**4069-4079.

302. Galassi G, Malagutti L, Colombini S, Rapetti L, Gallo L, Schiavon S, Tagliapietra F, Crovetto GM: **Nitrogen and energy partitioning in two genetic groups of pigs fed low-protein diets at 130 kg body weight.** *Italian Journal of Animal Science* 2015, **14:**4012.

303. Dietert RR, Silbergeld EK: **Biomarkers for the 21st century: listening to the microbiome.** *Toxicological Sciences* 2015, **144:**208-216.

304. Narayanan V, Peppelenbosch MP, Konstantinov SR: **Human Fecal Microbiome–Based Biomarkers for Colorectal Cancer.** *Cancer Prevention Research* 2014, **7:**1108-1111.

305. Yu J, Feng Q, Wong SH, Zhang D, yi Liang Q, Qin Y, Tang L, Zhao H, Stenvang J, Li Y: **Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer.** *Gut* 2017, **66:**70-78.

306. Farrokhi V, Nemati R, Nichols FC, Yao X, Anstadt E, Fujiwara M, Grady J, Wakefield D, Castro W, Donaldson J: **Bacterial lipodipeptide, Lipid 654, is a**

microbiome-associated biomarker for multiple sclerosis. *Clinical & Translational Immunology* 2013, **2:**e8.

307. Duranti S, Gaiani F, Mancabelli L, Milani C, Grandi A, Bolchi A, Santoni A, Lugli GA, Ferrario C, Mangifesta M: **Elucidating the gut microbiome of ulcerative colitis: bifidobacteria as novel microbial biomarkers.** *FEMS Microbiology Ecology* 2016, **92:**fiw191.
308. Greenblum S, Turnbaugh PJ, Borenstein E: **Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease.** *Proceedings of the National Academy of Sciences* 2012, **109:**594-599.
309. Strobl C, Boulesteix A-L, Zeileis A, Hothorn T: **Bias in random forest variable importance measures: Illustrations, sources and a solution.** *BMC Bioinformatics* 2007, **8:**25.
310. Caparrós-Martín JA, Lareu RR, Ramsay JP, Peplies J, Reen FJ, Headlam HA, Ward NC, Croft KD, Newsholme P, Hughes JD: **Statin therapy causes gut dysbiosis in mice through a PXR-dependent mechanism.** *Microbiome* 2017, **5:**95.
311. Shang Q, Sun W, Shan X, Jiang H, Cai C, Hao J, Li G, Yu G: **Carrageenan-induced colitis is associated with decreased population of anti-inflammatory bacterium, Akkermansia muciniphila, in the gut microbiota of C57BL/6J mice.** *Toxicology Letters* 2017, **279:**87-95.
312. Jing Y, Li A, Liu Z, Yang P, Wei J, Chen X, Zhao T, Bai Y, Zha L, Zhang C: **Absorption of Codonopsis pilosula Saponins by Coexisting Polysaccharides Alleviates Gut Microbial Dysbiosis with Dextran Sulfate Sodium-Induced Colitis in Model Mice.** *BioMed Research International* 2018, **2018**.
313. Zheng J, Yuan X, Cheng G, Jiao S, Feng C, Zhao X, Yin H, Du Y, Liu H: **Chitosan oligosaccharides improve the disturbance in glucose metabolism and reverse the dysbiosis of gut microbiota in diabetic mice.** *Carbohydrate Polymers* 2018, **190:**77-86.
314. Chen J, Kang B, Jiang Q, Han M, Zhao Y, Long L, Fu C, Yao K: **Alpha-Ketoglutarate in low-protein diets for growing pigs: effects on cecal microbial communities and parameters of microbial metabolism.** *Frontiers in Microbiology* 2018, **9**.

315. Mancabelli L, Milani C, Lugli GA, Turroni F, Cocconi D, van Sinderen D, Ventura M: **Identification of universal gut microbial biomarkers of common human intestinal diseases by meta-analysis.** *FEMS Microbiology Ecology* 2017, **93:**fix153.

316. Flint HJ, Scott KP, Louis P, Duncan SH: **The role of the gut microbiota in nutrition and health.** *Nature Reviews Gastroenterology & Hepatology* 2012, **9:**577.

317. Konikoff T, Gophna U: **Oscillospira: a central, enigmatic component of the human gut microbiota.** *Trends in Microbiology* 2016, **24:**523-524.

318. Miquel S, Martin R, Rossi O, Bermudez-Humaran L, Chatel J, Sokol H, Thomas M, Wells J, Langella P: **Faecalibacterium prausnitzii and human intestinal health.** *Current Opinion in Microbiology* 2013, **16:**255-261.

319. Mateos I, Combes S, Pascal G, Cauquil L, Barilly C, Cossalter A-M, Laffitte J, Botti S, Pinton P, Oswald I: **Fumonisin-exposure impairs age-related ecological succession of bacterial species in weaned pig gut microbiota.** *Toxins* 2018, **10:**230.

320. Al K, Sarr O, Dunlop K, Gloor GB, Reid G, Burton J, Regnault TR: **Impact of birth weight and postnatal diet on the gut microbiota of young adult guinea pigs.** *PeerJ* 2017, **5:**e2840.

321. Yang H, Xiao Y, Wang J, Xiang Y, Gong Y, Wen X, Li D: **Core gut microbiota in jinhua pigs and its correlation with strain, farm and weaning age.** *Journal of Microbiology* 2018, **56:**346-355.

322. Kim HB, Borewicz K, White BA, Singer RS, Sreevatsan S, Tu ZJ, Isaacson RE: **Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs.** *Veterinary Microbiology* 2011, **153:**124-133.

323. De Rodas B, Youmans BP, Danzeisen JL, Tran H, Johnson TJ: **Microbiome profiling of commercial pigs from farrow to finish.** *Journal of Animal Science* 2018, **96:**1778-1794.

324. Wright GD, Poinar H: **Antibiotic resistance is ancient: implications for drug discovery.** *Trends in Microbiology* 2012, **20:**157-159.

325. Aminov RI: **The role of antibiotics and antibiotic resistance in nature.** *Environmental Microbiology* 2009, **11:**2970-2988.

326. Bush K: **The coming of age of antibiotics: discovery and therapeutic value.** *Annals of the New York Academy of Sciences* 2010, **1213:**1-4.

327. Clardy J, Fischbach MA, Currie CR: **The natural history of antibiotics.** *Current Biology* 2009, **19:**R437-R441.

328. Blaser MJ: **Antibiotic use and its consequences for the normal microbiome.** *Science* 2016, **352:**544-545.

329. Hollis A, Ahmed Z: **Preserving antibiotics, rationally.** *New England Journal of Medicine* 2013, **369:**2474-2476.

330. Van Boeckel TP, Gandra S, Ashok A, Caudron Q, Grenfell BT, Levin SA, Laxminarayan R: **Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data.** *The Lancet Infectious Diseases* 2014, **14:**742-750.

331. Chang Q, Wang W, Regev-Yochay G, Lipsitch M, Hanage WP: **Antibiotics in agriculture and the risk to human health: how worried should we be?** *Evolutionary Applications* 2015, **8:**240-247.

332. Soucy SM, Huang J, Gogarten JP: **Horizontal gene transfer: building the web of life.** *Nature Reviews Genetics* 2015, **16:**472.

333. von Wintersdorff CJ, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, Savelkoul PH, Wolffs PF: **Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer.** *Frontiers in Microbiology* 2016, **7:**173.

334. Arnold KE, Williams NJ, Bennett M: **'Disperse abroad in the land': the role of wildlife in the dissemination of antimicrobial resistance.** *Biology Letters* 2016, **12:**20160137.

335. Greig J, Rajić A, Young I, Mascarenhas M, Waddell L, LeJeune J: **A scoping review of the role of wildlife in the transmission of bacterial pathogens and antimicrobial resistance to the food chain.** *Zoonoses and Public Health* 2015, **62:**269-284.

336. Davis S, Calvet E: **Fluctuating rodent populations and risk to humans from rodent-borne zoonoses.** *Vector-Borne & Zoonotic Diseases* 2005, **5:**305-314.

337. Gage KL, Kosoy MY: **Natural history of plague: perspectives from more than a century of research.** *Annu Rev Entomol* 2005, **50:**505-528.

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

339. Li B, Yang Y, Ma L, Ju F, Guo F, Tiedje JM, Zhang T: **Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes.** *The ISME Journal* 2015, **9:**2490.

340. Wickham H, Hester J, Francois R: **Readr: Read rectangular text data.** *R package version* 2017, **1**.

341. Wickham H, Francois R, Henry L: **Müller K. dplyr: a grammar of data manipulation. 2017.** 2018.

342. Wickham H: **Reshaping data with the reshape package.** *Journal of Statistical Software* 2007, **21:**1-20.

343. Warnes GR, Bolker B, Lumley T, Warnes MGR, Imports M: **Package 'gmodels'.** 2018.

344. Baquero F, Alvarez-Ortega C, Martinez J: **Ecology and evolution of antibiotic resistance.** *Environmental Microbiology Reports* 2009, **1:**469-476.

345. D'Costa VM, King CE, Kalan L, Morar M, Sung WW, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R: **Antibiotic resistance is ancient.** *Nature* 2011, **477:**457.

346. Gilliver MA, Bennett M, Begon M, Hazel SM, Hart CA: **Enterobacteria: antibiotic resistance found in wild rodents.** *Nature* 1999, **401:**233.

347. Zhu Y-G, Johnson TA, Su J-Q, Qiao M, Guo G-X, Stedtfeld RD, Hashsham SA, Tiedje JM: **Diverse and abundant antibiotic resistance genes in Chinese swine farms.** *Proceedings of the National Academy of Sciences* 2013**:**201222743.

348. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L: **The comprehensive antibiotic resistance database.** *Antimicrobial Agents and Chemotherapy* 2013**:**AAC. 00419-00413.

349. Liu B, Pop M: **ARDB—antibiotic resistance genes database.** *Nucleic acids research* 2008, **37:**D443-D447.

350. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M: **UniProt: the universal protein knowledgebase.** *Nucleic acids research* 2004, **32:**D115-D119.

351. Nesme J, Cécillon S, Delmont TO, Monier J-M, Vogel TM, Simonet P: **Large-scale metagenomic-based study of antibiotic resistance in the environment.** *Current Biology* 2014, **24:**1096-1100.

352. Piddock LJ: **Multidrug-resistance efflux pumps? not just for resistance.** *Nature Reviews Microbiology* 2006, **4:**629.

353. Korpela K, Salonen A, Virta LJ, Kekkonen RA, Forslund K, Bork P, De Vos WM: **Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children.** *Nature Communications* 2016, **7:**10410.

354. Chen B, Yang Y, Liang X, Yu K, Zhang T, Li X: **Metagenomic profiles of antibiotic resistance genes (ARGs) between human impacted estuary and deep ocean sediments.** *Environmental science & technology* 2013, **47:**12753-12760.

355. Luo G, Li B, Li L-G, Zhang T, Angelidaki I: **Antibiotic resistance genes and correlations with microbial community and metal resistance genes in full-scale biogas reactors as revealed by metagenomic analysis.** *Environmental Science & Technology* 2017, **51:**4069-4080.

356. Gibson MK, Forsberg KJ, Dantas G: **Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology.** *The ISME Journal* 2015, **9:**207.

357. Pernetta J: **Anatomical and behavioural specialisations of shrews in relation to their diet.** *Canadian Journal of Zoology* 1977, **55:**1442-1453.

358. Montgomery S, Montgomery W: **Intrapopulation variation in the diet of the wood mouse Apodemus sylvaticus.** *Journal of Zoology* 1990, **222:**641-651.

359. YUAN L-l, PENG X, WANG Y, SUN X-g, GUAN G-h, YANG M: **Mean Retention Time of Digesta Through the Gut in Lesser White-Toothed Shrew and Striped Hamster [J].** *Journal of Shenyang Normal University (Natural Science Edition)* 2008, **2:**028.

360. Weldon L, Abolins S, Lenzi L, Bourne C, Riley EM, Viney M: **The gut microbiota of wild mice.** *PLoS One* 2015, **10:**e0134643.

361. Yeaman MR, Yount NY: **Mechanisms of antimicrobial peptide action and resistance.** *Pharmacological Reviews* 2003, **55:**27-55.

362. Peschel A, Sahl H-G: **The co-evolution of host cationic antimicrobial peptides and microbial resistance.** *Nature Reviews Microbiology* 2006, **4:**529.

363. Pootoolal J, Neu J, Wright GD: **Glycopeptide antibiotic resistance.** *Annual Review of Pharmacology and Toxicology* 2002, **42:**381-408.

364. Vrieze A, Out C, Fuentes S, Jonker L, Reuling I, Kootte RS, van Nood E, Holleman F, Knaapen M, Romijn JA: **Impact of oral vancomycin on gut microbiota, bile acid metabolism, and insulin sensitivity.** *Journal of Hepatology* 2014, **60:**824-831.

365. Horvath JE, Willard HF: **Primate comparative genomics: lemur biology and evolution.** *Trends in Genetics* 2007, **23:**173-182.

366. Rabinowitz PD, Coffin MF, Falvey D: **The separation of Madagascar and Africa.** *Science* 1983, **220:**67-69.

367. Sarrasin B: **Ecotourism, poverty and resources management in Ranomafana, Madagascar.** *Tourism Geographies* 2013, **15:**3-24.

368. Choo JM, Leong LE, Rogers GB: **Sample storage conditions significantly influence faecal microbiome profiles.** *Scientific Reports* 2015, **5:**16350.

369. Hale VL, Tan CL, Knight R, Amato KR: **Effect of preservation method on spider monkey (Ateles geoffroyi) fecal microbiota over 8 weeks.** *Journal of Microbiological Methods* 2015, **113:**16-26.

370. Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, Knight R: **Preservation methods differ in fecal microbiome stability, affecting suitability for field studies.** *MSystems* 2016, **1:**e00021-00016.

371. van den Bogert B: *Community and genomic analysis of the human small intestine microbiota.* 2013.

372. Roguet C, Duflot B, Rieu M: **Changes in Pig Farming Models in Europe and Impact on Technical and Economic Performance.** *Économie Rurale* 2017**:**73-86.

373. Yasuda K, Maiorano R, Welch RM, Miller DD, Lei XG: **Cecum is the major degradation site of ingested inulin in young pigs.** *The Journal of Nutrition* 2007, **137:**2399-2404.

374. Million M, Lagier J-C, Yahav D, Paul M: **Gut bacterial microbiota and obesity.** *Clinical Microbiology and Infection* 2013, **19:**305-313.

375. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR: **Gut microbiota from twins discordant for obesity modulate metabolism in mice.** *Science* 2013, **341:**1241214.

376. Caricilli AM, Picardi PK, de Abreu LL, Ueno M, Prada PO, Ropelle ER, Hirabara SM, Castoldi Â, Vieira P, Camara NO: **Gut microbiota is a key modulator of insulin resistance in TLR 2 knockout mice.** *PLoS Biology* 2011, **9:**e1001212.

377. Perelman P, Johnson WE, Roos C, Seuánez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpler Y: **A molecular phylogeny of living primates.** *PLoS Genetics* 2011, **7:**e1001342.

378. Martin RD: **Origins, diversity and relationships of lemurs.** *International Journal of Primatology* 2000, **21:**1021-1049.

379. Clayton JB, Vangay P, Huang H, Ward T, Hillmann BM, Al-Ghalith GA, Travis DA, Long HT, Van Tuan B, Van Minh V: **Captivity humanizes the primate microbiome.** *Proceedings of the National Academy of Sciences* 2016, **113:**10376-10381.

380. Uenishi G, Fujita S, Ohashi G, Kato A, Yamauchi S, Matsuzawa T, Ushida K: **Molecular analyses of the intestinal microbiota of chimpanzees in the wild and in captivity.** *American Journal of Primatology: Official Journal of the American Society of Primatologists* 2007, **69:**367-376.

381. Gentry A, Clutton-Brock J, Groves CP: **The naming of wild animal species and their domestic derivatives.** *Journal of Archaeological Science* 2004, **31:**645-651.

382. Monteiro-Riviere NA, Riviere J: **The pig as a model for cutaneous pharmacology and toxicology research.** In *Advances in Swine in Biomedical Research*. Springer; 1996: 425-458

383. Cooper DK, Gollackner B, Sachs DH: **Will the pig solve the transplantation backlog?** *Annual Review of Medicine* 2002, **53:**133-147.

384. Groenen MA: **A decade of pig genome sequencing: a window on pig domestication and evolution.** *Genetics Selection Evolution* 2016, **48:**23.

385. Massei G, Kindberg J, Licoppe A, Gačić D, Šprem N, Kamler J, Baubet E, Hohmann U, Monaco A, Ozoliņš J: **Wild boar populations up, numbers of hunters down? A review of trends and implications for Europe.** *Pest Management Science* 2015, **71:**492-500.

386. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turroni S, Biagi E, Peano C, Severgnini M: **Gut microbiome of the Hadza hunter-gatherers.** *Nature communications* 2014, **5:**3654.

387. Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ: **Evaluation of methods for the extraction and purification of DNA from the human microbiome.** *PloS One* 2012, **7:**e33865.

388. El Aidy S, Van Den Bogert B, Kleerebezem M: **The small intestine microbiota, nutritional modulation and relevance for health.** *Current Opinion in Biotechnology* 2015, **32:**14-20.

389. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW: **Reagent and laboratory contamination can critically impact sequence-based microbiome analyses.** *BMC Biology* 2014, **12:**87.

390. Plummer E, Twin J, Bulach DM, Garland SM, Tabrizi SN: **A Comparison of Three Bioinformatics Pipelines for the Analysis of Preterm Gut Microbiota using 16S rRNA Gene Sequencing Data.** *Journal of Proteomics & Bioinformatics* 2015, **Volume 8:** 283-291.

391. Jovel J, Patterson J, Wang W, Hotte N, O'Keefe S, Mitchel T, Perry T, Kao D, Mason AL, Madsen KL: **Characterization of the gut microbiome using 16S or shotgun metagenomics.** *Frontiers in Microbiology* 2016, **7:**459.

392. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N: **Shotgun metagenomics, from sampling to analysis.** *Nature Biotechnology* 2017, **35:**833.

Summary

The intestinal microbiota has a profound effect on health and well-being of an individual. Microbiota composition has been shown to be associated with many aspects of a host's intestinal health, however, at the same time lifestyle and genetic background strongly influence gut microbiota composition. The aim of this thesis was to contribute to understanding of forces that shape microbiota composition and its properties in wild and domestic animals. In the scope of this work we investigated composition and influence of various factors on intestinal microbiota of seven wild animal species and three breeds of domestic pigs.

Chapter 1 provided state-of-the-art background information about microbial communities, their influence on health and a long history of prokaryote - eukaryote co-existence. This chapter also gave an overview of methods that are used for investigation of complex microbial communities.

In **Chapter 2** we investigated influence of occupancy and consequent differences in diet on gut microbiota of congeneric lemur species dwelling in ecologically distinct areas of Madagascar. Using amplicon sequencing of bacterial 16S rRNA genes we profiled faecal microbiota composition of free living red-fronted lemurs (*E. rufifrons*), common brown lemurs (*E. fulvus*) and red-bellied lemurs (*E. rubriventer*). We showed that environmental conditions of the habitation area were strongest modulators of microbiota, overshadowing the effects of genetic background, gender and social group identity. Lemurs of the same species but delving in different areas, and those exposed to different environmental conditions, showed to have distinct microbiota profiles, reflecting the importance of the habitation area in microbial consortium modulation.

In **Chapter 3** we build on knowledge obtained in the course of research described in Chapter 2 and greatly extended the sampling effort covering a larger number of lemur species and locations across Madagascar. In this chapter we more specifically addressed the influence of seasonality and logging on gut ecology of lemurs. Besides profiling of faecal microbiota we quantified the amount of eggs of nematodes from genera *Callistoura* and *Lemuricola*. We found that seasonality had a profound effect on microbiota composition and could be explained by variations in environmental conditions between wet and dry

seasons. We furthermore observed a significantly higher bacterial diversity and prevalence of *Lemuricola* spp. in previously logged forest areas. In addition, we discovered a small but significant influence of nematodes' egg load on overall microbiota composition.

Research described in **Chapter 4** of this thesis investigated relationships between genetic background and gut microbiota composition in pedigree pigs. To our knowledge this is the first large study that investigates variations in cecum microbiota of Italian Landrace (ILA), Italian Duroc (IDU) and Italian Large White (ILW) pigs that were kept in a controlled environment and have a defined genetic background. We demonstrated that microbiota correlated significantly with the pigs' breed identity. We showed that a number of microbial genus-level taxa traditionally associated with health homeostasis, such as *Ruminococcaceae* UCG-005, *Lachnospiraceae* XPB1014, *Oscillospira* and *Faecalibacterium* were differentially abundant between breed groups.

Resistance to antibiotics is a widespread phenomenon in the microbial world, and its development and spread across different environments is receiving increasing attention. In **Chapter 5** we explored the correlation between the composition of faecal microbiota and its resistome in wild Great White-Toothed Shrews and Wood Mice delving around pig farms. Using a combination of 16S rRNA gene amplicon and shotgun metagenomic sequencing we profiled microbiota phylogenetic and resistome composition. We found that abundance and diversity of antibiotic resistance genes strongly correlated with the phylogenetic composition of microbiota and differed between investigated species.

Finally, **Chapter 6** summarised and discussed findings presented in this thesis and provided an outlook towards future research perspectives and needs. In addition, the chapter provided data that was not included in the previous chapters but was used to further illustrate discussed points.

Co-Author affiliations

Hauke Smidt[1]

Iris de Winter [2]

Javier Ramiro-Garcia [1]

Ignas M.A. Heitkönig [2]

Gerrit Gort [3]

Peter M. Kappeler [4]

Luca Fontanesi [5]

Maurizio Gallo [6]

Fred de Boer [7]

Kevin D. Matson [2]

Nico van den Brink [8]

Freek IJdema [1,2]

Pim van Hooft [2]

Herbert H.T. Prins [2]

Nieuwland, W [1]

Diana Luise [5]

Paolo Bosi [5]

Paolo Trevisi [5]

Joost Lahr [9]

Dik J. Mevius [7]

1 Laboratory of Microbiology, Wageningen University & Research, Stippeneng 4, 6708 WE Wageningen, The Netherlands

2 Resource Ecology Group, Wageningen University& Research, Droevendaalsesteeg 3a, 6708 PB Wageningen, The Netherlands

3 Biometris, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

4 Behavioral Ecology and Sociobiology Unit, German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany

5 Department of Agricultural and Food Sciences (DISTAL), University of Bologna, Bologna, Italy.

6 Associazione Nazionale Allevatori Suini (ANAS), Roma, Italy

7 Wageningen Bioveterinary Research, Lelystad, The Netherlands

8 Laboratory of Toxicology, Wageningen University, Wageningen, The Netherlands

9 Wageningen Environmental Research, Wageningen University, Wageningen, The Netherlands

Acknowledgments

First of all I would like to thank my parents – Anna and Oleg Umanets who were a constant support for me not only during my PhD, but all my life. Without your love I would not have been able to get through it. Same goes to my sweet heart Annemiek Blasweiler, who gave me energy to keep going, and brought joy back to my life.

I would like to thank Hauke Smidt, who trusted in me and gave me an opportunity to pursue this PhD in one of the best microbiological labs in the world. His personal qualities and outstanding scientific expertise made this journey possible.

A special thanks goes to two of my friends, who also agreed to be my paranymphs. Dear Bastian, you are a great friend and able to withstand all my complaining about everything. Thank you for being such a terrific person. Dear Irene, your energy and optimism makes people around you feel energized. I am glad for all the fun moments that we shared, and wish you, Peer and Elisa oceans of happiness.

I was privileged to be a part of the largest and the most diverse group in MIB – MolEco. Thank you Detmer, Clara, Erwin and Hauke for making the group smooth running and "pleasant to be part of". Also, I would be most definitely lost and buried somewhere in the lab without help of technicians; thank you Ineke, Steven, Tom, Wilma, Philippe, Sjon, Wim and Bart (who was the most helpful when I was trying to abuse the micro4 server).

To be a PhD candidate at MIB means to have as your colleagues wonderful people from all around the world.

Johanna, we had numerous hangouts together and each time is precious to me - from staring at Yellowstone geysers to nights near camp fires in Droov. On this note I would like to thank all people who put up with me during our trips together. Thank you Javier for all your help and nice company during sample collection. Thank you Yuan for your great company in France, and everyday lab life. Thank you all participants of both PhD trips and after-trips.

Thank you Sven for being intense and a great help whenever it was needed. Thank you Monika for being a good friend and a fire starter at parties.

Thank you Sudarshan and Gerben for your microbiota and drinks advice. Lara, thank you for the fun times and pushing me to once in a while go into a museum instead of a party. Emmy, thank you for being active and trying to integrate stubborn foreigners into Dutch culture. Jueely, Cata, Beka Daan, Jorun, Aleks, Nico(Dutch and Greek-Dutch), Ioannis, Jannie – you are a lovely company to have around either for a lunch chat, trip or a party. Ioannis thank you – whenever I needed to film some stupid video you were always ready to help.

I am very thankful for all my officemates. With many of you I have become good friends and I hope to keep in touch. Thank you Klaudyna for your company and good life advise. Hugo – thank you for turning the office into a small mushroom farm. Thank you Yue, Floor, Tikka, Carry, Marta for making office live more interesting. And thank you Indra for being such a nice person.

My life will be definitely less entertaining without the company of the fun girls and guys from the neighbouring SSB department. Benua and Rob – I have no idea why people don't like the French and British – you guys are really nice people to have around. Thank you for your company and friendship. Maarten, our most famous athlete and bioinformatian – you are like olives - it definitely takes time to get used to, but after a while you cannot imagine a party without some. Erica and Wen (I know that you are not from SSB) you are two fun people who always manage to make me smile. Ruben and Dorret – it was really fun to have DnD sessions together and lunch chats. Maria, I really enjoyed our conversations, thank you.

This thesis and my PhD in general would not have been possible without my scientific collaborators. I was very lucky in this regard. Iris, thank you for introducing me into world of lemurs, and the productive collaboration in general. Many thanks for my collaborators from Bologna University (Italy) - Pablo, Diana and Maurizio our collaboration started with the COST action, but did not end there and after almost three years of joined projects I view you more as friends that collaborators. Another acknowledgment goes to Iratxe, who was a wonderful host during my visits for sampling in Spain. Also, a big thank you for Coline and Niels - I really enjoy to work with you and to have some projects that are a bit outside of "business as usual" routine.

Also, no PhD candidate can do without master students. I was lucky to have a number of you. In particularly, I have to give thanks to Freek, Anna, Christina and Nathalie. I think I learned from you just as much as you did from me (I hope).

I would like to thank the Erasmus Mundus foundation for financing my PhD project and Ewa Wietsma who helped me enormously during my stay at Wageningen.

Special thanks goes for Lim Chong-Wong, who introduced me into world of international science by adopting me as a master student at his lab in Chonbuk National University.

I would like to thank Primorskaya State Academy of Agriculture and the great people who are working there. This was my starting place and without the opportunities provided there this thesis would not be possible.

I really have to thank my current colleagues at WBVR and particularly Ad and the whole IBIS group for the support during the last stretch of my PhD.

A great thanks for the WAC and its members. This place played a very important part in my life. I found good friends and much more at this family-like place. Marleen, thank you for keeping it going! On the same note I thank all my coaches and sparing partners for the Jiu-Jitsu and Judo sessions.

I would like to thank all members of the Umanets and Blasweiler family who were always kind and welcoming to me.

I would like to finish these acknowledgments by apologizing for people that I didn't list here. It doesn't mean that I didn't appreciate you, it is just the limited format!

About the author

Alexander Umanets was born in 1988 in the Kraunowka village which is situated in the Primorsiky kray at the Far East of Russia. After his graduation from High School with silver distinction he joined the Doctor of Veterinary Medicine program at Primorskaya State Academy of Agriculture. There he actively participated in the internal scientific program and simultaneously undertook the joined program in English language for professional communications. He graduated from Primorskaya State Academy of Agriculture in 2011 with distinction. The same year he was awarded with a scholarship to join the Master of Veterinary Pathology program in Chonbuk National University under supervision of Prof. Dr. Lim Chong-Wong. After receiving the degree of Master of Veterinary Pathology in 2013 he went back in Primorskaya State Academy of Agriculture and took a position of lecturer. In 2014 he received Erasmus Mundus fellowship for Doctorates and moved in The Netherlands to pursue a PhD in Microbiology at The Laboratory of Microbiology (Wageningen University & Research) under supervision of Prof. Dr. Hauke Smidt. His PhD research was focused on external and internal factors that shape the intestinal microbiota of wild and domestic animals. From November 2018 he joined Wageningen Bioveterinary Research as a researcher with a focus on intracellular pathogens and their interplay with intestinal microbiota.

List of publications

- **Umanets A***, de Winter I*, IJdema F, Ramiro-Garcia J, van Hooft P, Heitkönig IMA, Prins HHT, Smidt H (2018). Occupancy strongly influences faecal microbial composition of wild lemurs. FEMS Microbiology Ecology
- Gerritsen J, **Umanets A**, Staneva I, Hornung B, Ritari J, Paulin L, Rijkers GT, de Vos WM, Smidt H (2018). *Romboutsia hominis sp.* nov., the first human gut-derived representative of the genus *Romboutsia*, isolated from ileostoma effluent. International Journal of Systematic and Evolutionary Microbiology (IJSEM)
- Verhulst NO, **Umanets A**, Weldegergis BT, Maas JPA, Visser TM, Dicke M, Smidt H, Takken W (2018). Do apes smell like humans? The role of skin bacteria and volatiles of primates in mosquito host selection. Journal of Experimental Biology,
- **Umanets Alexander***, Chae-Wong Lim,*, Bumseok Kim, Eui-Ju Hong, Hyeon-Cheol Kim, Bae-Keun Park (2018). Morphological and Molecular Characterization of Toxocara tanuki (Nematoda: Ascaridae) from Korean Raccoon Dog, Nyctereutes procyonoides koreensis. The Korean Journal of Parasitology
- Nasitrema attenuata (*Digenia: Nasitrematidae*) Infection of Long-beaked Common Dolphin (*Delphius capensis*) in the East Sea, Korea. Chae-Wong Lim, Seajin Han, Bumseok Kim, **Umanets Alexander** *at al.* Journal of Veterinary Clinics, 2016
- A case of nonserotypable *Escherichia coli* infection in a Korean rabbit farm. Gerry A. Camer, Yoon-Seok Roh, Ara Cho, Jong-Won Kim, **Alexander Umanets** *et al.* Korean J Vet Serv, 2012
- Effects of seasonality and previous logging on faecal helminth-microbiota associations in wild lemurs. de Winter, I.I.*, **Umanets, A.***, Gort, G., Nieuwland, W., van Hooft, P., Heitkönig, I.M.A., Kappeler, P.M., Prins, H.H.T.1, Smidt, H. Submitted to Scientific Reports

* Contributed equally

In preparation

- Variation of caecal microbial composition among performance-tested Italian Landrace, Italian Large White, and Italian Duroc pig breeds demonstrate the effect of host genetic background on microbiota diversity. **Alexander Umanets***, Diana Luise*, Luca Fontanesi, Paolo Bosi, Maurizio Gallo, Hauke Smidt, Paolo Trevisi
  **Authors contributed equally*
- Differences in intestinal microbiota composition between Great White-Toothed Shrews and Wood Mice shapes differences in their resistome composition. **Alexander Umanets**, Fred de Boer, Pim van Hooft, Joost Lahr, Kevin D. Matson, Dik J. Mevius, Hauke Smidt, Nico van den Brink
- Metatranscriptome analysis of microbial activity and function in the pig GI tract Floor Hugenholtz, **Alexander Umanets** et al.

Overview of completed training activities

The intestinal Microbiome and Diet in Human and Animals Health	VLAG, Wageningen, NL	2014
Systems Biology: "Statistical analysis of -omics data"	WUR, Wageningen, NL	2014
Advanced proteomics	VLAG, Wageningen, NL	2015
Gut day	Gut Flora Foundation, Rotterdam, NL	2015
Gut day	Gut Flora Foundation, Venlo, NL	2016
PiGutNet meeting	COST action, Barcelona , ES	2016
Scientific Spring Meeting KNVM & NVMM 2016	KNVM, Papendal, NL	2016
"GUT MICROBIOLOGY: twenty years and counting ..." - poster	INRA-Rowett Institure, Clermont-Ferrand , FR	2016
Gut Microbiota for Health Summit	GMFH Conference, Barcelona, ES	2015
Scientific Spring Meeting KNVM & NVMM 2017 - two posters	KNVM, Arnhem, NL	2017
KNVM Fall Meeting 2016 (section Microbial Ecology) - oral	KNVM, Nijmegen, NL	2016
Analysis of porcine metagenomic datasets	INRA – COST, Paris, FR	2017
Microbiology Centennial Symposium - pitch	WUR, Wageningen, NL	2017
VLAG PhD week	VLAG, Baarlo, NL	2015
Basic statistics	PE&RC, Wageningen, NL	2014
Career Perspectives	WGS,	2017

	Wageningen, NL	
Mobilising your Scientific Network	WGS, Wageningen, NL	2017
Entrepreneurship in and outside Science	WGS, Wageningen, NL	2017
Optionals		
Preparation of research proposal	WU-Microbiology, Wageningen, NL	2014
PhD/postdoc meetings	WU-Microbiology, Wageningen, NL	2014-2018
Laboratory of Microbiology group meetings	WU-Microbiology, Wageningen, NL	2014-2018
Strategy Days	WU-Microbiology, Wageningen, NL	2014-2018
2 weeks PhD trip to USA	WU-Microbiology, West-cost, USA	2015
2 weeks PhD trip to Europe	WU-Microbiology, Northen Europe, EU	2017
Microbial ecology (Master course)	WUR, Wageningen, NL	2014

Colophon

Printing of this thesis was financially supported by The Laboratory of Microbiology (Wageningen University).
The PhD project was financed by Erasmus Mundus foundation and The Laboratory of Microbiology (Wageningen University).

Cover design by Annemiek Blasweiler
Layout design by Alexander Umanets
Thesis printed by DigiForce