Supplementary Information

Diversity, abundance, and multiple adaptations of archaea to the animal gut

Courtney M. Thomas[1,2], Elie Desmond-Le Quemener[3], Simonetta Gribaldo[1] and Guillaume Borrel[1,*]

* Corresponding author: guillaume.borrel@pasteur.fr

The file contains:

Supplementary Text
Supplementary Tables list
Supplementary Tables 2 to 6 (1 and 7 are provided in separate files)
Supplementary Figures 1 to 17
Supplementary References

Supplementary text

Influence of other factors on archaeal community structure and abundance

The geographic origin of the samples only explained a minor part of the variance (Table 1; Figure S14) and species sampled from different zoos generally have close archaeal composition profiles (Figure S15). Our dataset did not allow to determine if captivity significantly influences the structure of the archaeal community as it would need to compare wild and captive animals belonging to the same species.

Other factors such as the absolute abundance of archaea (copies of 16S per gram of feces) and archaea:bacteria ratio as determined by qPCR did significantly impact the structure of the mammalian archaeome, but at the margin (Table 1). Abundance of archaea and bacteria are correlated (Figure 2b). Animal body mass and stomach pH were not significantly related to the abundance of archaea in these animals.

Absolute abundance of archaeal lineages in host lineages

Using lineage specific qPCR primers, we determined that the Methanobacteriales were the most prevalent methanogens detected (36% of species), followed by the Methanomassiliicoccales (28.3%%), the *Methanimicrococcus* (26%) and finally the Methanomicrobiales (23.8%). The Thaumarcheota were detected in (46.1%) of species sampled. The Methanobacteriales and Methanomassiliicoccales were not detected by qPCR in the sampled Invertebrates (Insecta, Malacostraca, Mollusca, and Gastropoda), and the Actinopterygii with group specific primers, potentially because they were below the detection limit. The Thaumarchaeota, Methanomicrobiales, and *Methanimicrococcus*, were detected in at least one species from each animal class (Figure S7). We did not identify any significant relationship between the absolute abundance of archaea and the animal phylogeny through the Moran Index.

Methane metabolisms

Overall, 45.3% of animals hosted a methanogen community dominated ($\geq$60% of the reads) by CO_2-dependent hydrogenotrophic methanogens, while 22% of the hosts have a methanogen community dominated by methyl-dependent hydrogenotrophs. In several animal lineages, the intestinal methanogens are consistently dominated by either methyl-dependent hydrogenotrophic methanogens or CO_2-dependent hydrogenotrophic methanogens. In the Perissodactyla/Cetartiodactyla, CO_2-dependent hydrogenotrophic methanogens represent 86.9% of the reads on average. Methyl-reducing methanogens are the dominant or codominant methanogens in 15 out of the 24 Primates species. In the Lemuridae methyl-dependent hydrogenotrophic methanogens represent 68.7% of the reads on average. While in the *Hominidae*, CO_2-dependent hydrogenotrophic methanogens represent 76% of reads on average (Figure 3e).

Details on the dominant archaeal orders in the animal gut

Methanobacteriales represent the largest fraction of the reads (43.1 % of reads, 491 ASVs out of 1307; Figure 2a). The Methanobacteriales were the most common archaea in the *Hominidae*, Cetartiodactyla, Rodentia, and Lagomorpha and were found in high relative abundance in two large flightless birds– the Lesser Rhea and Greater Rhea (Figure 3d). The majority of the Methanobacteriales reads (79.9%) were annotated as *Methanobrevibacter*

(Figure 4). More than 80% of the Methanobacteriales reads were found in mammals and these reads formed several clades in the reference tree that are enriched in specific host orders or host diet (Figure 4). Other mammalian-derived *Methanobrevibacter* ASVs tend to form clades according to the taxonomic affiliation of their host, Perissodactyla (two clades), Cetartiodactyla (two clades), Perissodactyla/ Cetartiodactyla mixed (one clade), Primates (three clades), and Rodentia (one clade) (Figure 4). At the difference of other mammalian orders with more than five sampled species, there is no large monophyletic clades of *Methanobrevibacter* strongly enriched in sequences form Carnivora.

We identified fewer ASVs in other Methanobacteriales genera. Indeed only 14.8% and 4.5% of all the Methanobacteriales ASVs were annotated as *Methanosphaera* and *Methanobacterium*, respectively. Within the *Methanosphaera*, there are three distinct clades of ASVs from Cetartiodactyla/Perissodactyla, Primates, and Carnivora. *Methanobacterium* was also detected in mammals from the orders Carnivora, Pilosa, and Cingulata (Figure 4). A large proportion of the most prevalent Methanobacteriales ASVs (found in $\geq$5% of mammalian species) are highly specific of the ruminant Cetartiodactyla (Figure S12).

Methanomassiliicoccales represents the second largest fraction of the reads (19.1% of the reads, 206 ASVs; Figure 2a). Lemuridae have a high relative abundance of Methanomassiliicoccales compared to closely related animal groups (Figure 3d). Within the Methanomassiliicoccales included in our reference tree (those representing at least 1% of the archaeome in one host), most of the ASVs (93/98) cluster with the *Methanomethylophilaceae* previously referred as "host-associated clade" [Sup1–3], as do a majority of the associated reads (91.7%). The majority (69.4%) of the Methanomassiliicoccales ASVs were also mammal-derived (Figure S5). ASVs from the Perissodactyla cluster with sequences that were obtained from other studies in the closely related Cetartiodactyla. One clade is enriched in ASVs derived from Primates. ASVs from the two reptilian orders, Squamata (carnivorous) and Testudines also separate in the tree. Interestingly, a majority of reads attributed to reptiles came from the *Methanomassiliicoccaceae* / "free-living clade" (Figure S5). Like the Methanobacteriales, the Methanomassiliicoccales are largely absent in the bony fishes and invertebrates which is supported by qPCR data (Figure S7).

Nitrososphaerales/*Nitrososphaeraceae* (phylum Thaumarchaeota) is the third lineage that totalizes the largest number of reads (15.7% of the reads, 267 ASVs; Figure 2a). However, their absolute abundance is generally around 10^6 copies per gram of feces or below (Figure 2c). Thus, they generally dominate the archaeome of animals with a low concentration of archaea such as the invertebrates, some birds, and the order Carnivora (Figure 3b,d; Figure 2d). Nitrososphaerales ASVs represent 93% and 90% of Thaumarchaeota ASVs and reads, respectively. This order was previously identified in humans and apes [Sup4] (Figure S6). At the difference of methanogen orders, there are no clades enriched in one type of host in the Thaumarchaeota (Figure S6). But surprisingly, several ASVs are mostly enriched in both Carnivora and Gastropoda (Figure S16).

Methanomicrobiales are less common than the Methanobacteriales and Methanomassiliicoccales (13.7% of the reads, 177 ASVs; Figure 2a), but they generally occur at high absolute abundance and are the dominant archaea in a third of all reptile species and in almost all existing species of the three Perissodactyla families (Equidae, Tapiridae, Rhinocerotidae), the only exception being the Malayan Tapir (Figure 3d). We found that 99.7%

of ASVs from this archaeal order cluster with *Methanocorpusculum* reference sequences- and none with *Methanomicrobium* (Figure S4). ASVs from *Methanocorpusculum* gather into two main clades. One clade has ASVs from a variety of mammalian orders, Rodentia, Diprotodontia, Carnivora, Pilosa, and Primates as well as ASVs derived from amphibians, invertebrates, reptiles (Squamata), and birds. In this clade, three ASVs from Rodentia are well separated from the others and are only present in members of this animal order. The second large clade subdivides into several subclades each of them composed of sequences with a strong specificity for Testudines, Perissodactyla, and Cetartiodactyla (Figure S4). The branching of ASVs from Testudines among ASVs from Perissodactyla (and the associated paraphyly of the ASVs from Perissodactyla) can be due to phylogenetic biases.

Methanosarcinales are sparsely distributed throughout animal species, with a high relative abundance in some reptiles and insect-eating mammals. In terms of number of reads, a single genus – Methanimicrococcus was the fourth most abundant archaea (Figure 2a). This order of methanogens had the lowest level of diversity with just 81 ASVs. These are spilt between to genera, *Methanimicrococcus* (55 ASVs) and *Methanosarcina* (26 ASVs) (Figure S17). However, 92.7% of the Methanosarcinales reads were attributed to *Methanimicrococcus*, while *Methanosarcina* only accounted for 7.3%. The genus *Methanimicrococcus* has largely been associated with the digestive tract of termites and cockroaches [Sup5–11], but recently was shown to form two distinct clades -one of sequences obtained from insects and another from mammals [Sup12]. One ASV that clustered with insect-derived *Methanimicrococcus* reference sequences came from a frog and a toad possibly indicating these amphibians acquired this methanogen from their prey. Similarly, we found that several ASVs from other animals that feed on invertebrates such as the long-spine squirrel fish, common gull, Eurasian coot, Great spotted woodpecker all cluster within the same clade of the insect-enriched reference sequences of *Methanimicrococcus*. The largest number of reads we obtained for the *Methanimicrococcus* came from animals whose primary diet is invertebrates (Figure 3d). Most *Methanimicrococcus* mammal ASVs fall within the previously proposed mammal clade [Sup12] (Figure S17).

Cooccurrence analysis

It was not possible to run a cooccurrence analysis on all animals because of the sampling distribution and the number of ASVs (clustered into OTUs) shared between groups. However, it was possible to run them on several lineages of animal for which enough samples were collected and there was enough overlap between ASVs. This analysis was thus run on all mammal samples and independently on mammalian orders, Primates, Perissodactyla/Cetartiodactyla and Carnivora. As few studies have been completed targeting the ecology of intestinal archaea in birds and reptiles and we had enough samples for robust statistical analyses, we also run a cooccurrence analyses on each of these two classes of animals.

1. **Mammals**

Across all mammals we identified 620 OTUs that were present in ≥10% of all mammal species. There were six archaea-bacteria relationships identified in both co-occurrence algorithms, four of which being between methanogens and Clostridiales members (Table S7). In addition to its cooccurrence with an OTUs closely related to *Lachnospira pectinoschiza*, *Methanomethylophilaceae* OTUarc_11 was also found to be significantly correlated to the

presence of a bacterial-OTU belonging to the family *Muribaculaceae* (phylum Bacteroidetes) which was particularly abundant in the Cingulata. A BLAST of this bacterial OTU sequence did not result in any significant results. This family of bacteria was previously identified to be a dominant member of the rodent intestinal microbiome and have the capacity to degrade pectin [Sup13,14], likely forming methanol. However, the lack of a clear annotation of this bacterial OTU leaves the details of the relationship between these archaea and bacteria unclear. The *Ca. Nitrosocosmicus* OTUarc_45 (ASV4) - the most widespread Thaumarchaeota in our dataset - was significantly linked to *Solobacterium* which was isolated from human feces [Sup15]. The cooccurrence of these two OTUs is also observed when only including Primates samples in the analysis.

2. Ungulata: Perissodactyla and Cetartiodactyla

The highest number of archaeal - bacterial relationships was observed in the Ungulata (Perissodactyla and Cetartiodactyla; Table S3 & S7) with 17 positive associations and seven negative associations. These orders also have the highest concentration of archaea on average ($3.3x10^8$ 16S rRNA gene copies/gram of feces), as well as some of the highest diversity of archaea. Nine positive relationships we identified were between Archaea and Clostridiales members, four were between Archaea and Bacteroidetes members, and the other four were between Archaea and either Melainabacteria, Patescibacteria, or Tenericutes members (Table S7).

Two *Christensenellaceae* OTUs (OTUbac_67 and OTUbac_2376) were negatively correlated to methanogen OTUs: a *Methanobrevibacter* OTUarc_20 (corresponding to *M. ruminantium*) and *Methanomethylophilaceae* OTUarc_189l; Table S7). This is surprising as *Christensenellaceae* have previously been shown to be positively correlated with *Methanobrevibacter smithii* in the human intestine and some of its representatives support the growth of this methanogen [Sup16]. However, a third *Christensenellaceae* OTUs (OTUbac_503) was positively correlated to *Methanobrevibacter* OTUarc_33 (corresponding to *Methanobrevibacter wolinii*; Table S7). *Methanobrevibacter* OTUarc_20 and *Methanomethylophilaceae* OTUarc_189 were also negatively correlated with a *Ruminococcaceae* OTU and a Patescibacteria OTU, respectively.

3. Primates

All the archaea-bacteria associations we identified in primates were between archaea and Firmicutes. The positive association of a pectin-degrading bacterium *Lachnospira* (OTU_2345) and Methanomethylophilaceae OTUarc_11, found when considering all Mammals, was also identified when the microbial community of primates was analyzed independently. Although the edge stability was slightly below the cut off threshold of 0.5, all the other metrics indicate the significance of this relationship (edge stability = 0.45, p = 0.05, *rho* =0.42). Two CO_2-dependent hydrogenotrophic methanogens the *Methanobrevibacter* and *Methanocorpusculum* were also found to be positively correlated to members of uncharacterized *Ruminococcaceae* one of the most abundant family of bacteria in the gut of mammals [Sup17].

4. Carnivora

Carnivora members host a unique community of intestinal archaea among mammals – a low concentration of archaea, and a number of them have an archaeome dominated by

Thaumarchaeota - thus we also performed an individual cooccurrence analysis on this group of mammals. OTUs belonging to the Methanobacteriales and Nitrososphaerales are significantly correlated to several members of the Firmicutes and one Proteobacteria in the Carnivora (Table S7). *Methanobacterium* (OTUarc_325 – 95% similarity to *Methanobacterium formicicum*) and *Methanobrevibacter* (OTUarc_20 – 98% similarity to *M. olleyae/ruminantium*) - were positively associated to OTUs from the Firmicutes like *Lachnospiraceae* (*Blautia* and Tyzzerella_4, respectively). Further, we found that *Methanobrevibacter* was positively correlated to the presence of a *Lactobacillus* OTU in the gut of Carnivora (Table S7). A *Ca.* Nitrosocosmicus OTU (Thaumarchaeota) was positively linked to a *Staphylococcus* OTU (Firmicutes), but negatively linked to Enterobacteriaceae OTU corresponding to *Escherichia/Shigella* (Proteobacteria). The cooccurrence between a dominant intestinal bacterium and a newly identified genus of the intestinal archaea warrants further investigation.

5. Aves and Reptiles

No archaea-bacteria relationships were identified in both the cooccurrence approaches in Aves (Table S7). The SparCC approach alone, identified one significant archaea-bacteria relationship between a *Methanosphaera* OTU and a Clostridiales (Clostridium *sensu stricto* 1). The SPIEC-EASI approach identified 57 relationships between archaea and bacteria. Interestingly, *M. smithii* (OTUarc_1), is negatively correlated to several bacterial OTUs in Aves, including common gut bacterial groups such as Clostridia as well as with *Arthrobacter* which was previously characterized as being part of the goose core-gut microbiome [Sup18]. It is currently not clear as to why this apparently well-adapted intestinal archaea would be negatively related to common intestinal bacteria in birds. Other archaeal OTUs such as *Methanosphaera*, *Methanomethylophilaceae*, and *Nitrososphaeraceae* were positively correlated to various bacteria groups (Table S7). For example, one of the *Nitrososphaeraceae* OTUs was positively correlated with a *Blautia* OTU. The same bacterial OTU was also negatively correlated to OTUarc_1.

Two archaeal OTUs were linked to bacteria in reptiles. A *Methanocorpusculum* OTU is significantly positively correlated to a *Providencia* OTU (100% identical to *Providencia rettgeri*; Enterobacteriales) which is a common constituent of the human and reptile intestinal tract [Sup19]. Also, *Methanomassiliicoccus* is positively correlated to the Clostridiaceae_1 family (95% identical to *Clostridium cylindrosporum)*.

6. Cooccurrence between archaeal OTUs

Globally, these analyses also highlighted an absence of negative archaea-archaea relationship and several positive archaea-archaea relationships between OTUs of a same family. This was notably the case between *Methanobrevibacter* and *Methanosphaera* OTUs (*Methanobacteriaceae*) in Primates, Ungulata and overall mammals, but no association were observed between *Methanobrevibacter* OTUs or between *Methanosphaera* OTUs. A *Methanomethylophilaceae* OTUs was also associated to another *Methanomethylophilaceae* OTUs in Primates (Table S7). *Nitrososphaeraceae* OTUs were also positively associated to each other in mammals.

Supplementary tables

Table S1: Sample information, animal species metadata, archaeal and bacterial abundance, archaeal ASV richness, and relative abundance of archaeal lineages. Only samples for which DNA was successfully extracted and PCR amplifications obtained were shown. Provided in a separate .xls file.

Table S2: Percentage of ASVs (n = 1307) corresponding to characterized archaea (isolates/enriched). ASVs sequences were compared to 16S rRNA genes of type strain archaea in the SILVA Living Tree Project LTP, plus additional sequences of enriched archaea whose genome was sequenced.

Table S3: Summary of co-occurrence analyses. ASVs were pooled into OTUs at a 97% similarity-cut off and normalized using the Archaea – Bacteria ratio determined via qPCR. Only cooccurring OTUs matching between SparCC and SPIEC-EASI algorithms were counted. Well represented groups (>6 species per order) of mammals, birds, and reptiles were all analysed. Perissodactyla and Cetartiodactyla (Ungulata) samples were analysed together to increase the robustness of the analysis.

Table S4: Known metabolisms of archaeal lineages identified in the animal gut.

Table S5: Primers used for quantitative PCR.

Table S6: Primers used for amplicon sequencing.

Table S7: Archaeal and bacterial OTUs cooccurring in mammals, reptiles and birds. Provided in a separate .xls file.

Table S2: Percentage of ASVs (n = 1307) corresponding to characterized archaea (isolates/enriched). ASVs sequences were compared to 16S rRNA genes of type strain archaea in the SILVA Living Tree Project LTP, plus additional sequences of enriched archaea whose genome was sequenced.

	Similarity with characterized archaea at different levels			
	95	**97**	**98.7***	**99**
Custom database + Silva LTP	*84.9%*	*53.9%*	*31%*	*20.4%*
% of reads with BLAST hit	*(94.5%)*	*(75.5%)*	*(52.5%)*	*(46.7%)*
Silva - Living Tree Project	*66.1%*	*35.6%*	*18%*	*11.8%*
% of reads with BLAST hit	*(74.2%)*	*(51.3%)*	*(34.5%)*	*(31.3%)*

* Proposed species level similarity of 16S rRNA genes by Yarza et al 2014

Table S3: Summary of co-occurrence analyses. ASVs were pooled into OTUs at a 97% similarity-cut off and normalized using the Archaea – Bacteria ratio determined via qPCR. Only cooccurring OTUs matching between SparCC and SPIEC-EASI algorithms were counted. Well represented groups (>6 species per order) of mammals, birds, and reptiles were all analysed. Perissodactyla and Cetartiodactyla (Ungulata) samples were analysed together to increase the robustness of the analysis.

Group	OTUS-ASV in 10% Species (Archaea OTUs)	Sparcc/Spiec Matches (Arc-Bac connections)	SparCC/SPIEC-EASI - Filtered matches			
			Total	Bac-Bac	Arc-Bac	Arc-Arc
Mammals	682 (18)	1180 (8)	804	794	6	4
Carnivora	286 (19)	589 (22)	108	103	5	0
Primates	677 (12)	2585 (19)	501	494	5	2
Perissodactyla & Certiartiodactyla	994 (22)	4523 (182)	1042	1016	24	2
Aves	571 (9)	450 (0)	268	268	0	0
Reptiles	659 (17)	466 (4)	168	166	2	0

Table S4: Known metabolisms of archaeal lineages identified in the animal gut

Phylum	Class	Order	Family	Best annotation	Methane metabolism	Remark on methane metabolism	Other energetic metabolism	Relation to oxygen
Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanobacteriaceae*	*Methanobacterium*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanobacteriaceae*	*Methanobrevibacter*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanothermobacteraceae*	*Methanothermobacter*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanocellales	*Methanocellaceae*	*Methanocella*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanocellales	*Methanocellaceae*	*Methanocellaceae*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanocorpusculaceae*	*Methanocorpusculum*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanoregulaceae*	*Methanoregula*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanospirillaceae*	*Methanospirillum*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanomicrobiales	NA	*Methanomicrobiales*	Hydorgenotrophic CO2-reducing			Strict anaerobe
Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanobacteriaceae*	*Methanosphaera*	Hydorgenotrophic CH3-reducing	One isolate reduces methanol by only using ethanol as electron donor (not H_2)		Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanimicrococcus*	Hydorgenotrophic CH3-reducing			Strict anaerobe
Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomassiliicoccaceae*	*Methanomassiliicoccaceae*	Hydorgenotrophic CH3-reducing			Strict anaerobe
Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomassiliicoccaceae*	*Methanomassiliicoccus*	Hydorgenotrophic CH3-reducing			Strict anaerobe
Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomethylophilaceae*	*Candidatus Methanogranum*	Hydorgenotrophic CH3-reducing			Strict anaerobe
Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomethylophilaceae*	*Candidatus Methanomethylophilus*	Hydorgenotrophic CH3-reducing			Strict anaerobe
Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomethylophilaceae*	*Candidatus Methanoplasma*	Hydorgenotrophic CH3-reducing			Strict anaerobe
Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomethylophilaceae*	*Methanomethylophilaceae*	Hydorgenotrophic CH3-reducing			Strict anaerobe
Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	Methylotrophic / Acetoclastic / Hydrogenotrophic CH_3-reducing / Hydrogenotrophic CO_2-reducing	All species are methylotrophic and some have various combinaisons of the other methane metabolism		Strict anaerobe
Bathyarchaeia	Bathyarchaeia	Bathyarchaeia_subgroup-6	NA	*Candidatus* Termitimicrobium	None		Possibly homoacetogen / organotrophic	Strict anaerobe
Bathyarchaeia	Bathyarchaeia	Bathyarchaeia_subgroup-6	NA	*Candidatus* Termiticorpusculum	None		Organotrophic	Strict anaerobe
Euryarchaeota	Halobacteria	Halobacteriales	*Halococcaceae*	*Halococcus*	None		Organotrophic	Strict aerobe
Thaumarchaeota	Nitrososphaeria	Nitrosopumilales	*Nitrosopumilaceae*	*Candidatus* Nitrosopumilus	None		Ammonia oxydizer	Strict aerobe
Thaumarchaeota	Nitrososphaeria	Nitrosopumilales	*Nitrosopumilaceae*	*Candidatus* Nitrosotenuis	None		Ammonia oxydizer	Strict aerobe
Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	*Nitrososphaeraceae*	*Candidatus* Nitrocosmicus	None		Ammonia oxydizer	Strict aerobe
Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	*Nitrososphaeraceae*	*Candidatus* Nitrososphaera	None		Ammonia oxydizer	Strict aerobe
Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	*Nitrososphaeraceae*	*Nitrososphaeraceae*	None		Ammonia oxydizer	Strict aerobe
Thaumarchaeota	Nitrososphaeria	Nitrosotaleales	*Nitrosotaleaceae*	*Nitrosotaleaceae*	None		Ammonia oxydizer	Strict aerobe

Table S5: Primers used for quantitative PCR.

Primer Pair	Targeted Group	Sequence	Annealing temperature	Reference
Thaum494F/	Thaumarchaeota	GAATAAGGGGTGGGCAAGT/	61°C	Sup 20
806R		GGACTACVSGGGTATCTAA		Sup 21
MMB282F/	Methanomicrobiales	ATCGRTACGGGTTGTGGG/	61.6°C	Sup 22
MM832R		CACCTAACGCRCATHGTTTAC		
Mx765F/	Methanomassiliicoccales	GGGGTAGGGGTAAAATCCTG/	61.2°C	Sup 23
Mx887R		CGGGGTATCTAATCCCGTTT		
Methani388F/	*Methanimicrococcus*	ACAATGCAGGAAACTGTG/	60.7°C	This study
Methani564R		TAGACCMAATAAAAGCGGCTA		
MBT857F/	Methanobacteriales	CGWAGGGAAGCTGTTAAGT/	60°C	Sup 22
MBT1196R		TACCGTCGTCCACTCCTT		
A-934b-F/	All archaea	GAATTGGCGGGGGAGCA/	60°C	Sup 24 (modified)
A-1000R		GGCCATGCACYWCYTCTC		Sup 25
B-357F	All bacteria	CTCCTACGGGAGGCAGCAG	65°C	Sup 26
B-531R		CTNYGTMTTACCGCGGCTGC		Sup 27

Table S6: Primers used for amplicon sequencing.

Primer Pair	Targeted Group	Sequence	Annealing temperature	Cycles	References
515F/	All 16S rRNA genes	*GTGYCAGCMGCCGCGGTAA /	55.5°C	30 x	Sup 28
806bR		*GGACTACNVGGGTWTCTAAT			Sup 29
344F/	Archaea (step 1)	ACGGGGYGCAGCAGKCGCG/	66.6°C	20 x	Sup 30 (modified)
1041R		GGCCATGCACCWCCTCTC			Sup 25
519F/	Archaea (step 2)	*CAGCMGCCGCGGTAA/	55.5°C	25 x	Sup 31
915R		*GTGCTCCCCCGCCAATTCCT			Sup 32

* Illumina tags were present on 515F/519F as well as 806bR/915R
Illumina F: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
Illumina R: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG

Figure S1.

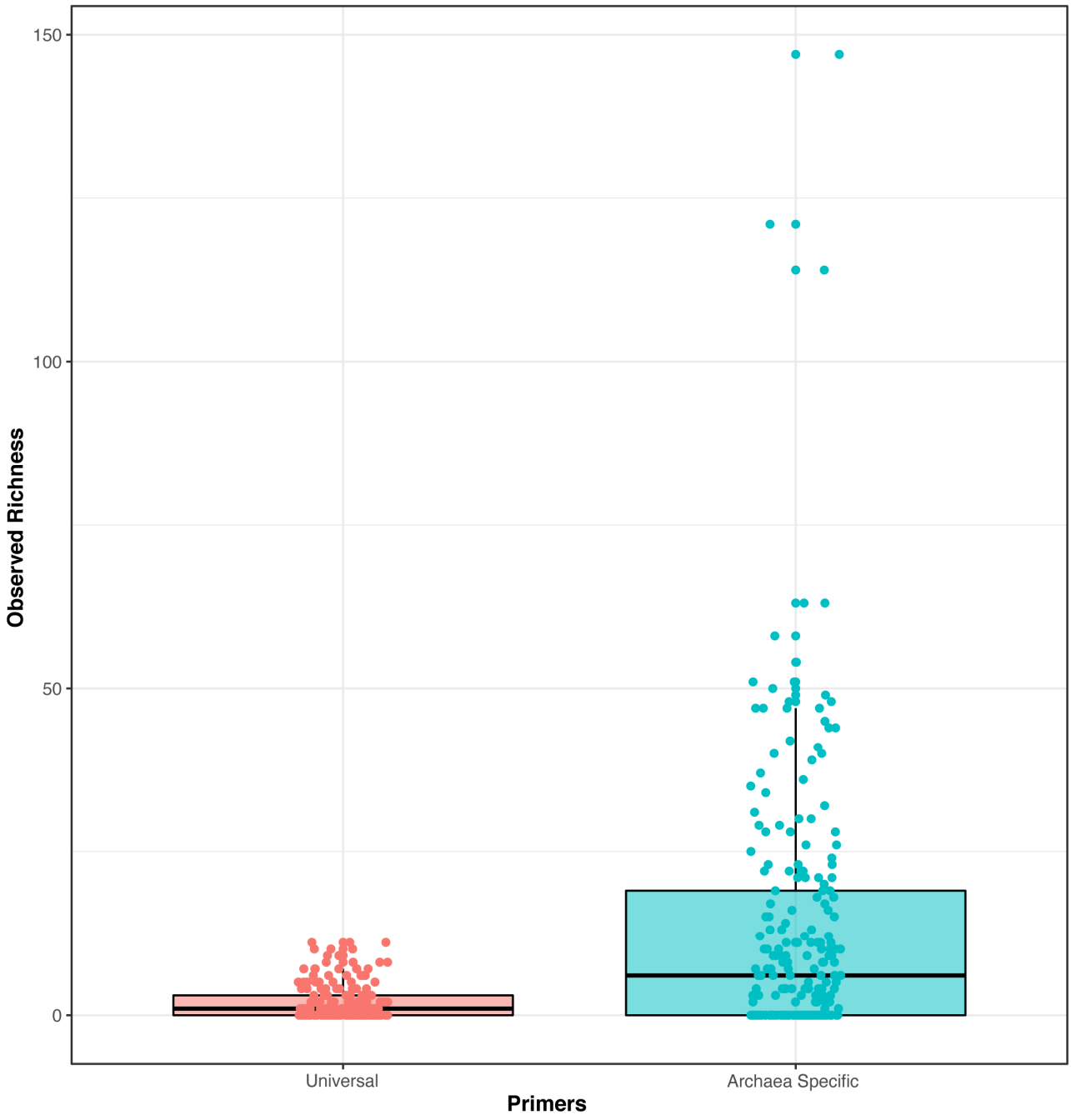

Figure S1: Comparison of the observed archaeal richness captured using universal and Archaea-specific 16S rRNA gene primer.

Figure S2.

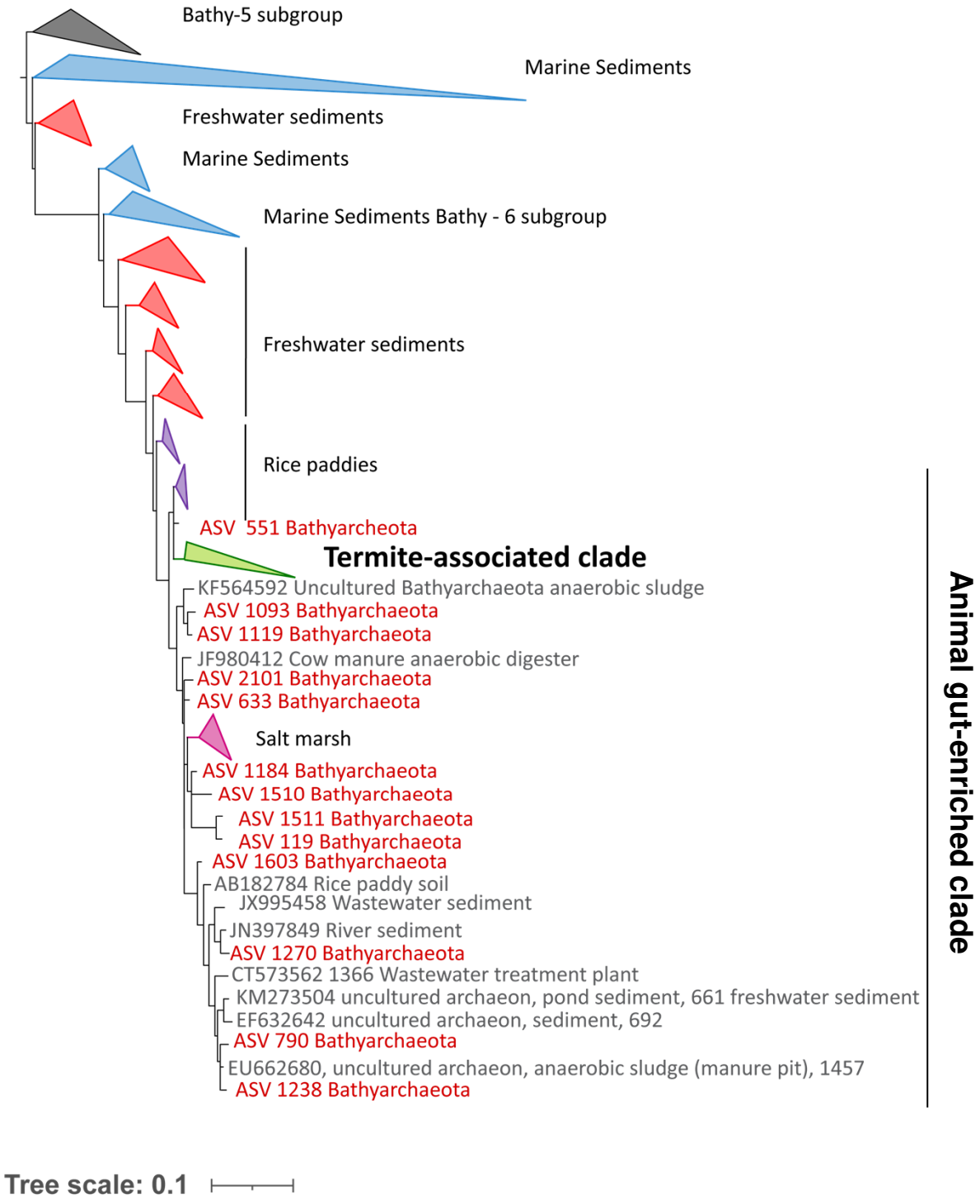

Figure S2: Close relationship between Bathyarchaeota sequences recovered from the animal gut. Sequences from this study are in red. They correspond to ASVs representing at least 1% of the gut archaeome of one animal species. Maximum-likelihood tree based on 16S rRNA gene sequences and built with GTR+G4 model.

Figure S3.

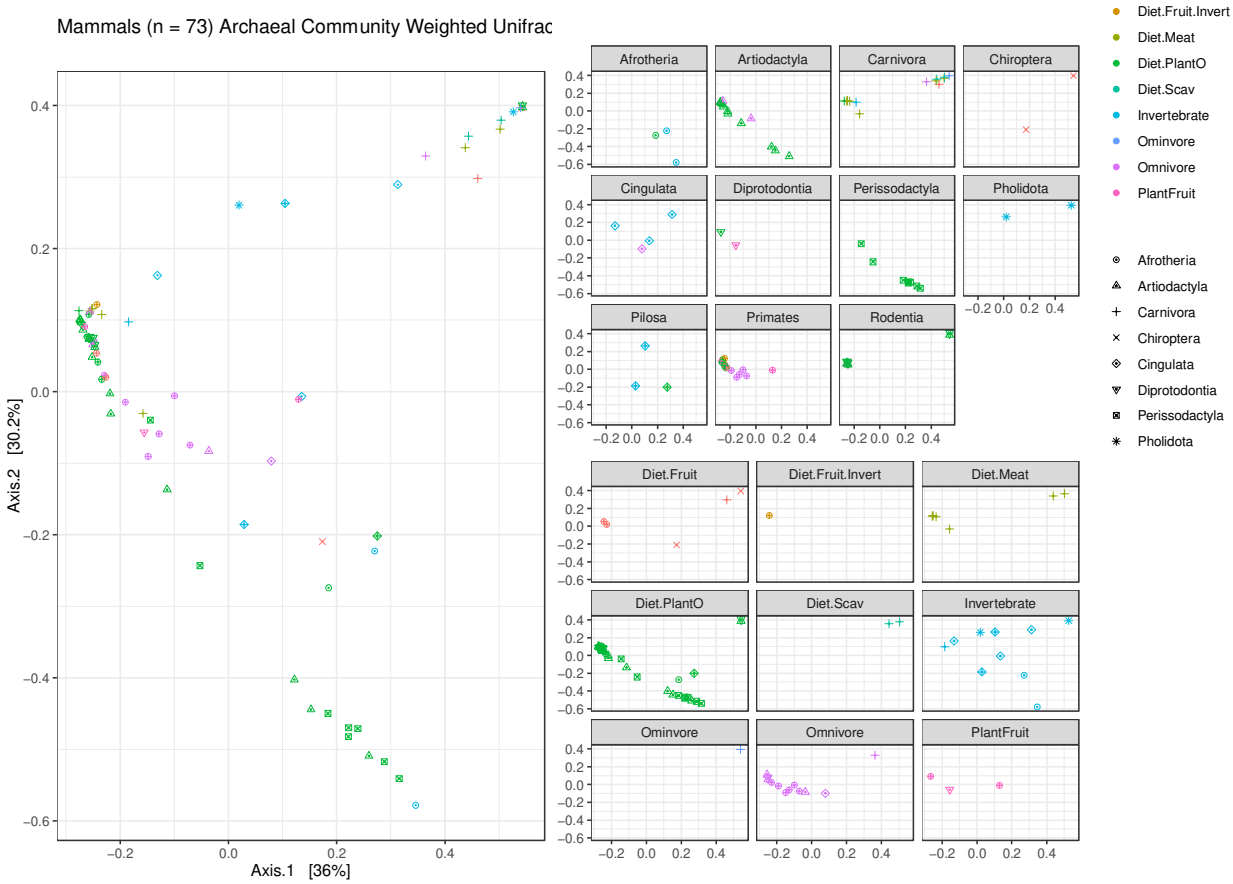

Figure S3: Weighted UniFrac analysis of the gut archaeome of Mammals, based on 73 rarefied samples with ≥3000 reads.

Figure S4

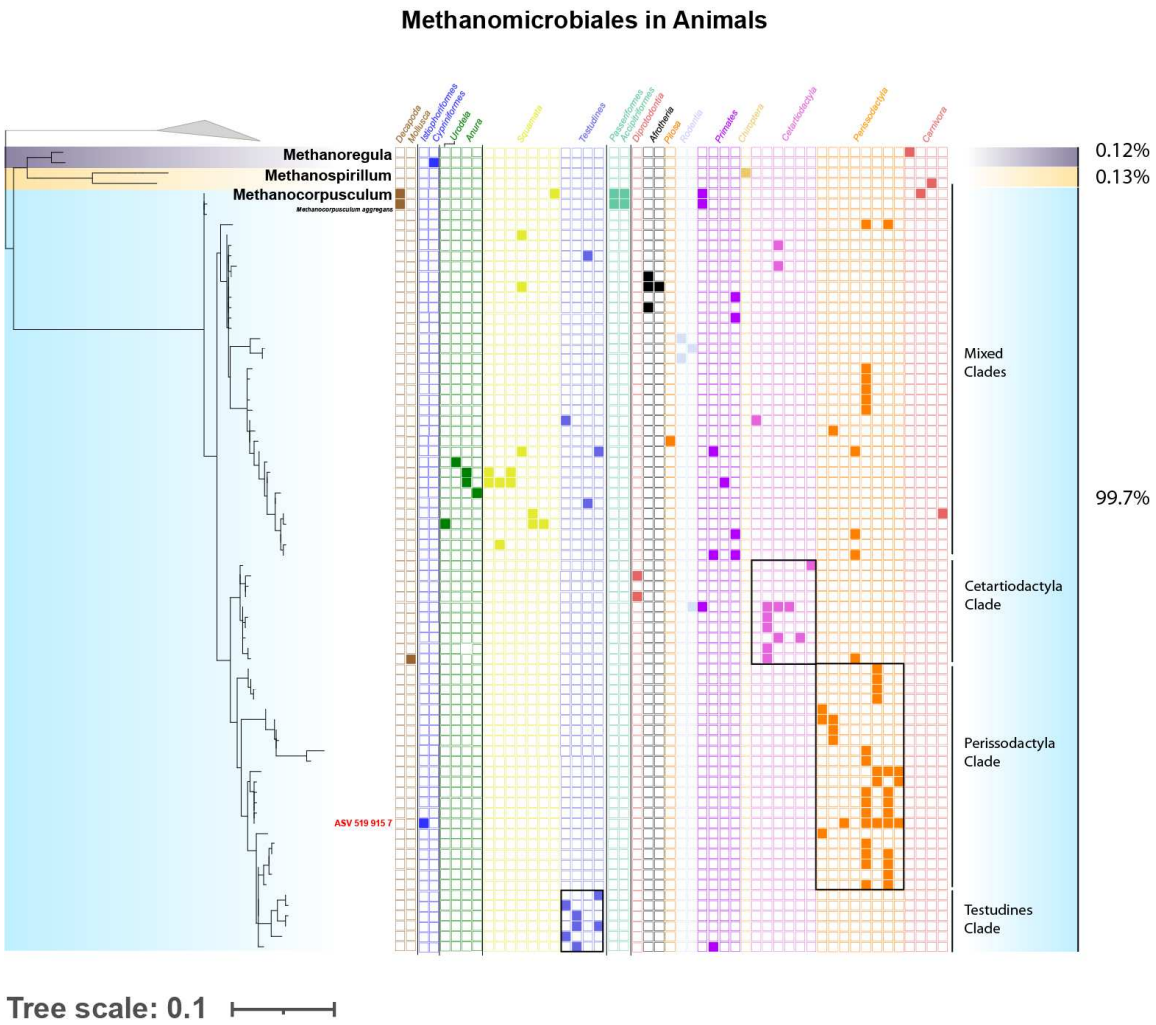

Figure S4: Distribution of Methanomicrobiales ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was built with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were then removed from the tree. Only ASV representing more than 1% of the sequences per sample were included. The percentages on the right represent the the proportion of reads from this order that were annotated as Methanocorpusculum, Methanospirillum and Methanoregula.

Figure S5.

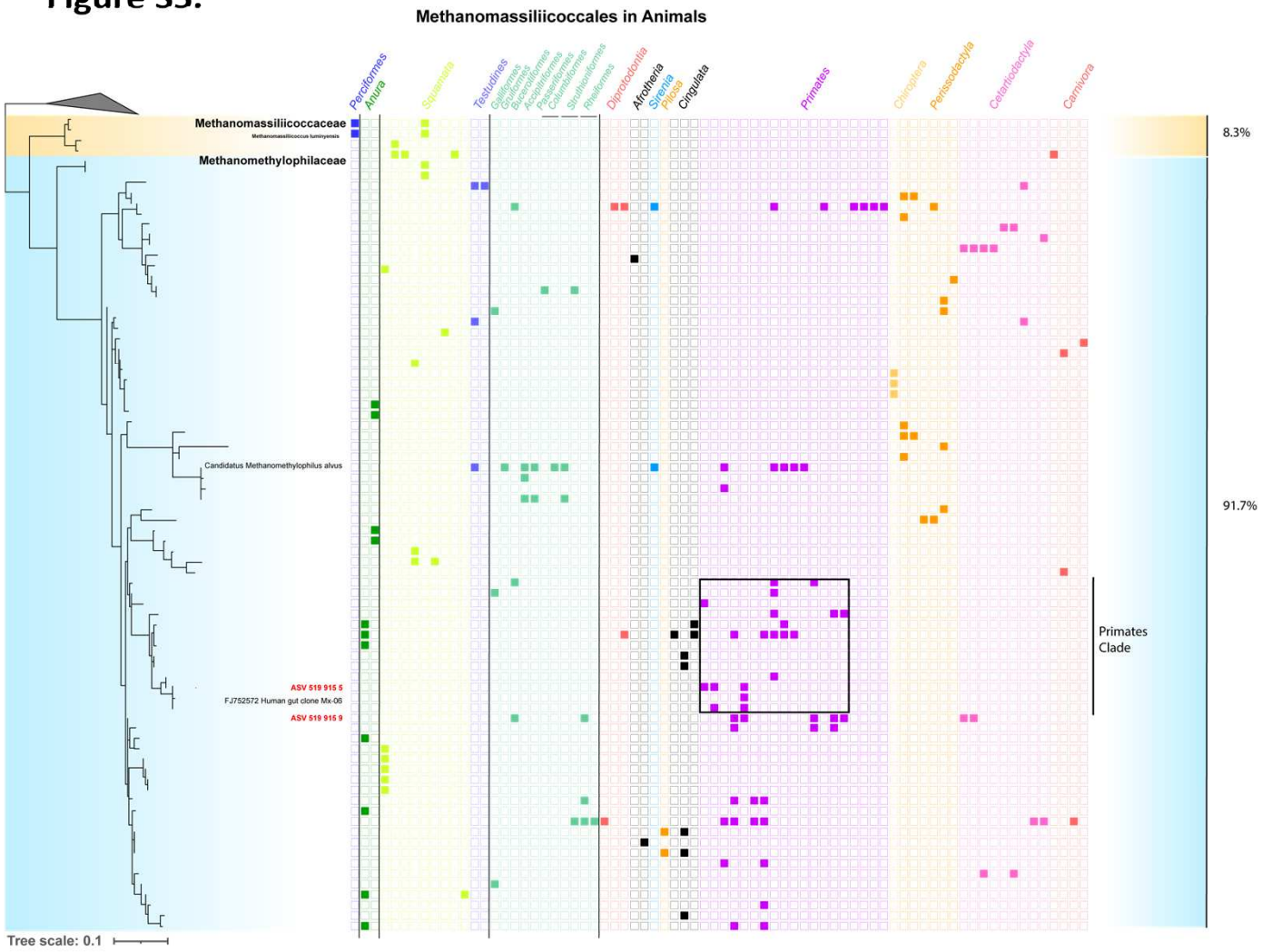

Figure S5: Distribution of Methanomassiliicoccales ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was constructed with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were then removed from the tree. Only ASV representing more than 1% of the sequences per sample were included. The percentages on the right represent the the proportion of reads from this order that were annotated as Methanomethylophilaceae and Methanomassiliicoccaceae.

Figure S6.

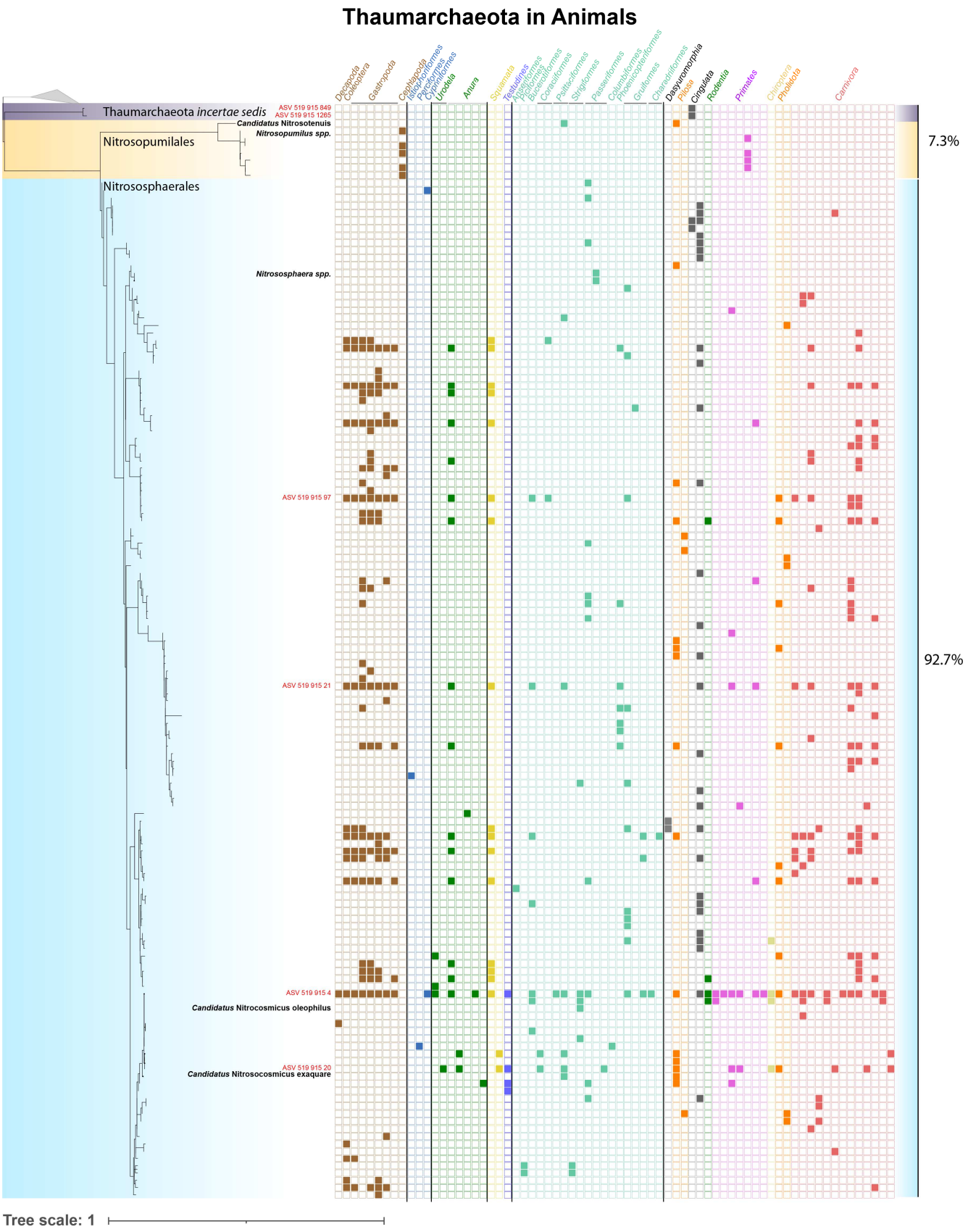

Figure S6: Distribution of Thaumarchaeota ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was constructed with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were then removed from the tree. Only ASV representing more than 1% of the sequences per sample were included. The percentages on the right represent the the proportion of reads from this order that were annotated as Nitrososphaerales or Nitrosopumilales.

Figure S7.

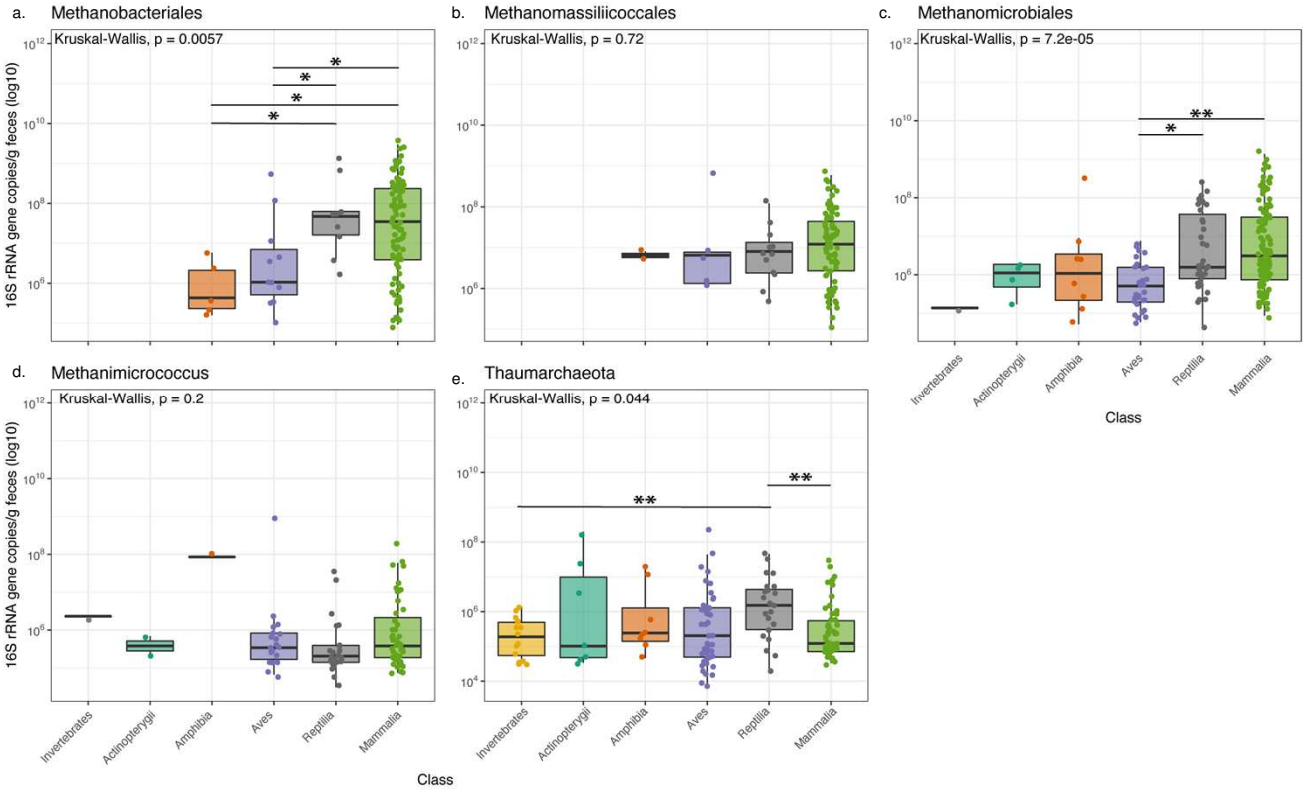

Figure S7: Absolute abundance of five archaeal lineages determined via qPCR. Significant differences across all groups were determined via the Kruskal-Wallis test, $p < 0.05$ is significant. a) Methanobacteriales (n = 121); b) Methanomassiliicoccales (n = 99); c) Methanomicrobiales (n = 204); d) Methanimicrococcus (n = 94); e) Thaumarcheota (n = 151). Wilcoxon rank sum test with continuity correction was used to determine differences between animal classes, *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$.

Figure S8.

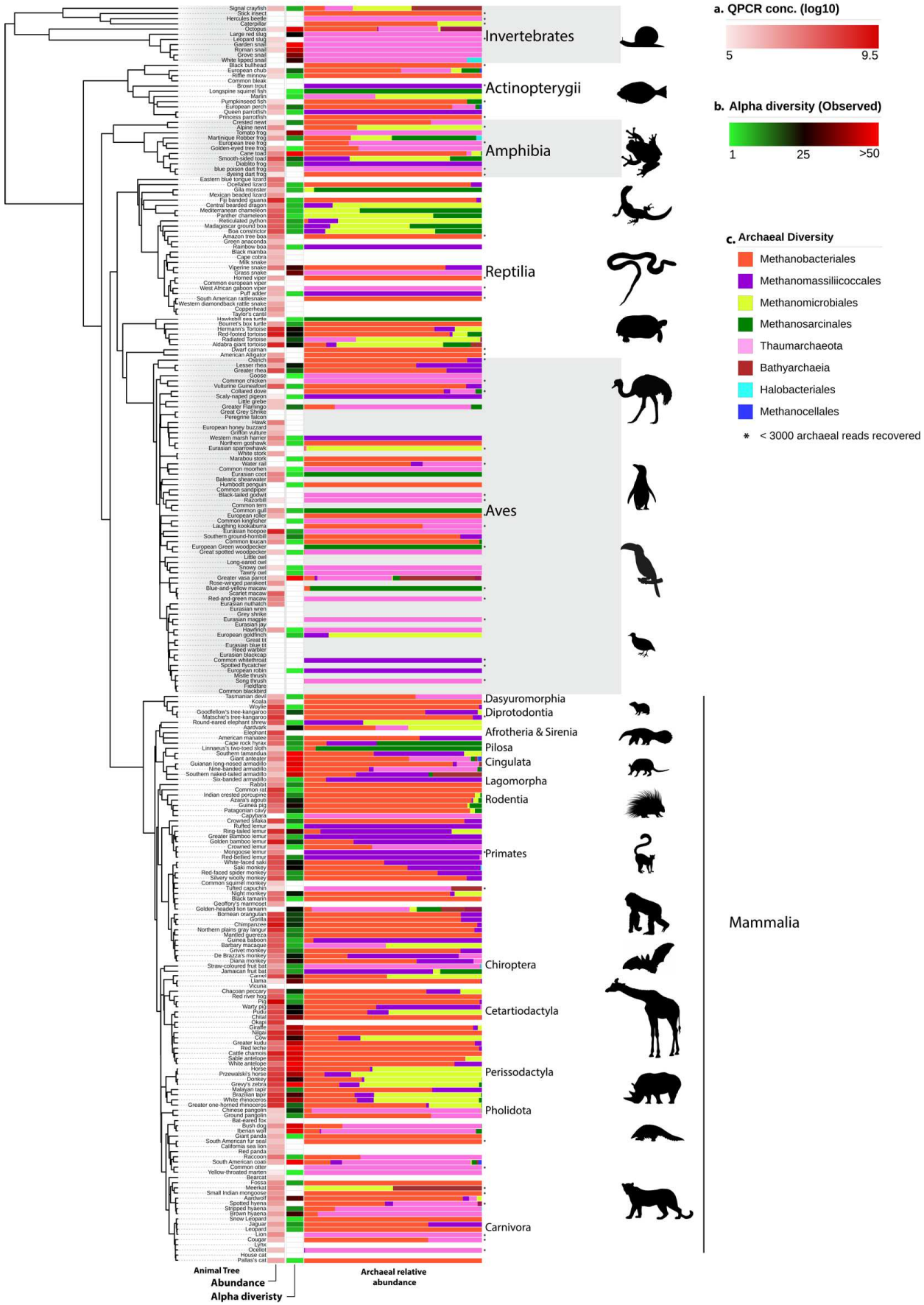

Figure S8: Archaeal diversity across animals including species for which a low number (< 3000 reads) of reads or no reads were obtained. Animal species with less 5 to 3000 reads are indicatde by a *.

Figure S9.

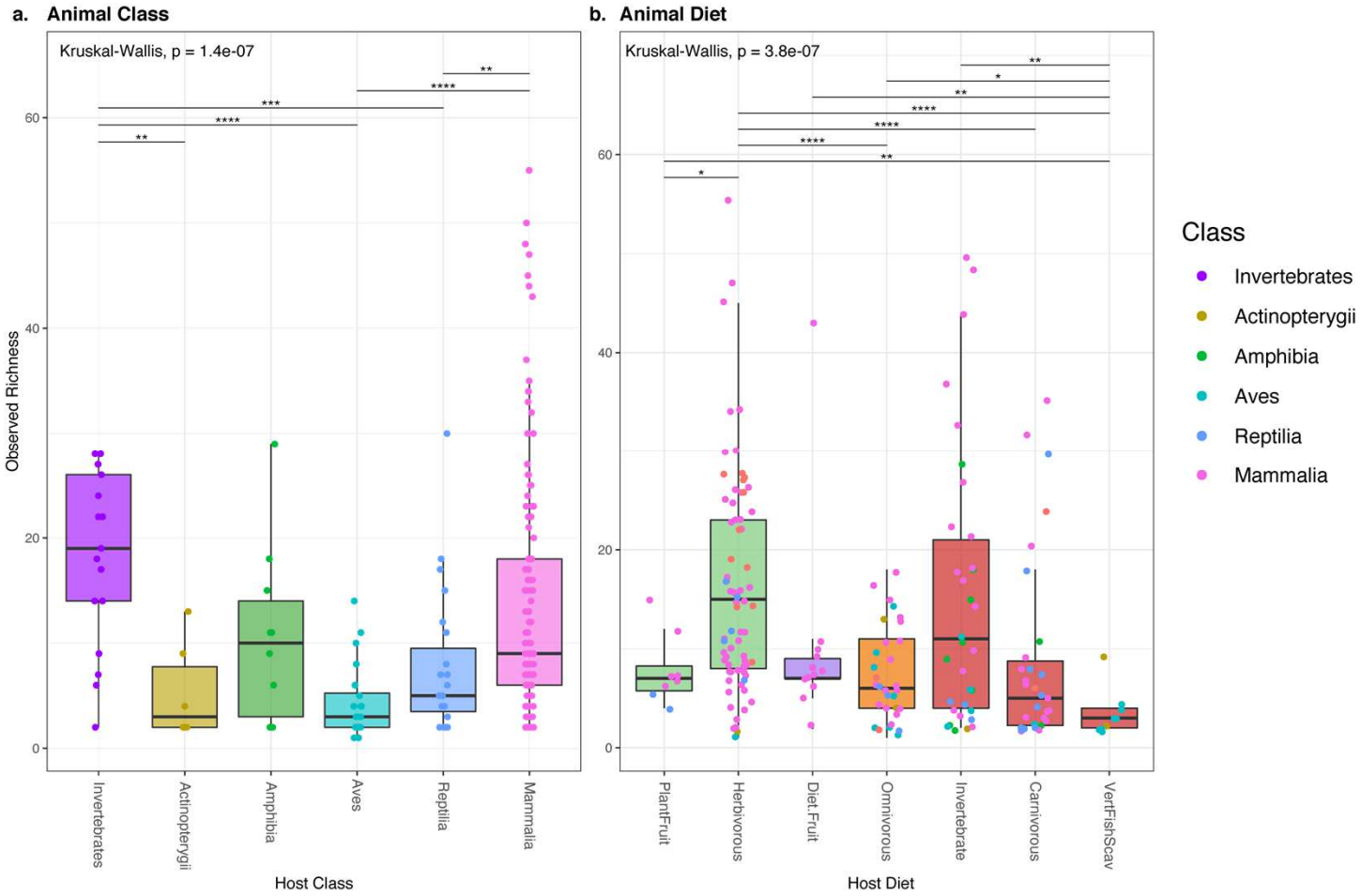

Figure S9: a) Observed richness of Archaea between animal classes (n = 218). b) Observed richness of Archaea between animal diets (n = 197), diet types with fewer than 3 representatives were removed. Wilcoxon rank sum test *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$.

Figure S10.

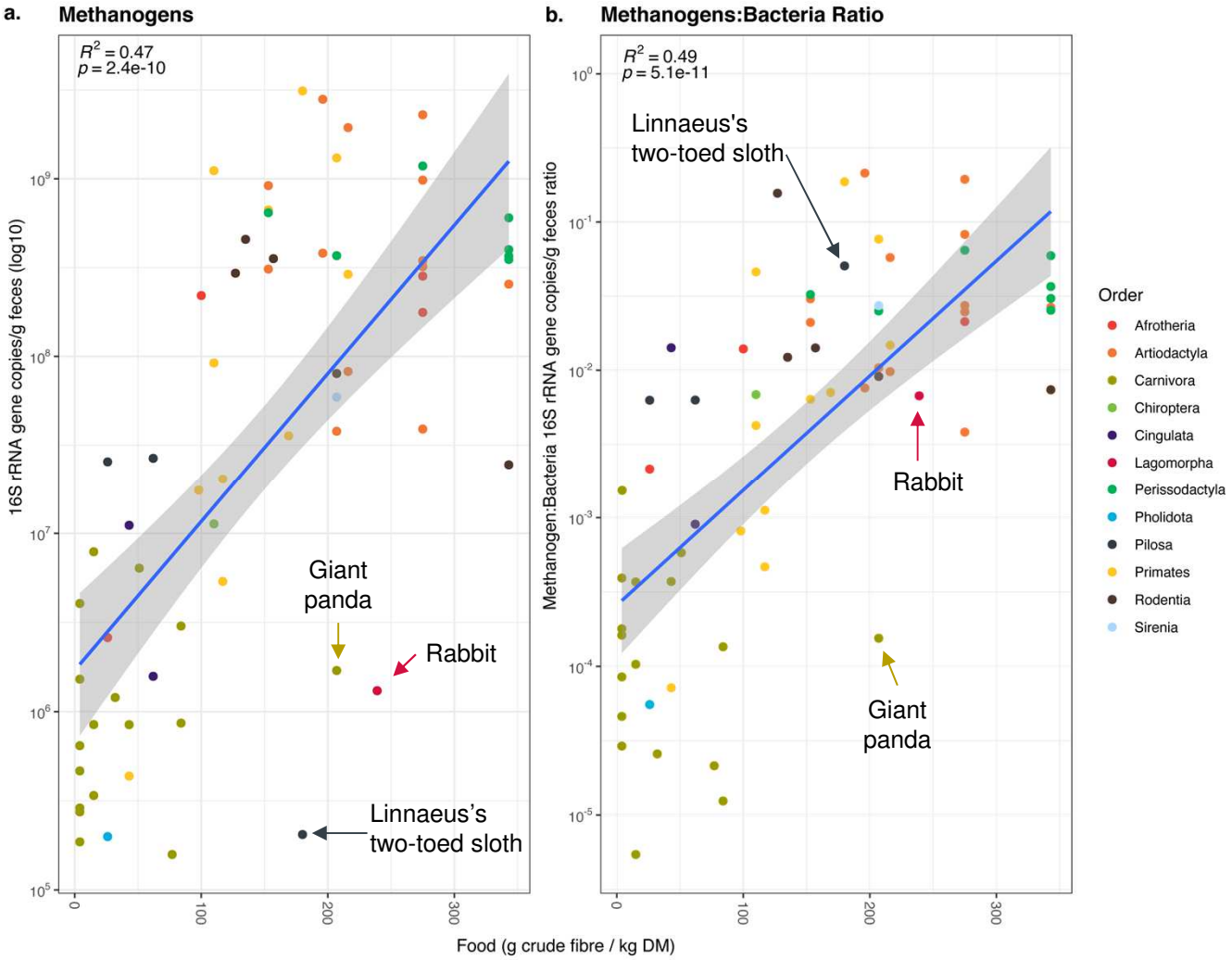

Figure S10: Correlation between diet-fibre content and a) absolute abundance and b) relative abundance of methanogens in mammals faeces (n = 110). The abundance methanogen is determined by the sum of the abundance of abundance of Methanobacteriales, Methanomassiliicoccales, Methanomicrobiales, and *Methanimicrococcus* determined by qPCR. The ratio of methanogens to bacteria are also based on qPCR measurements.

Figure S11.

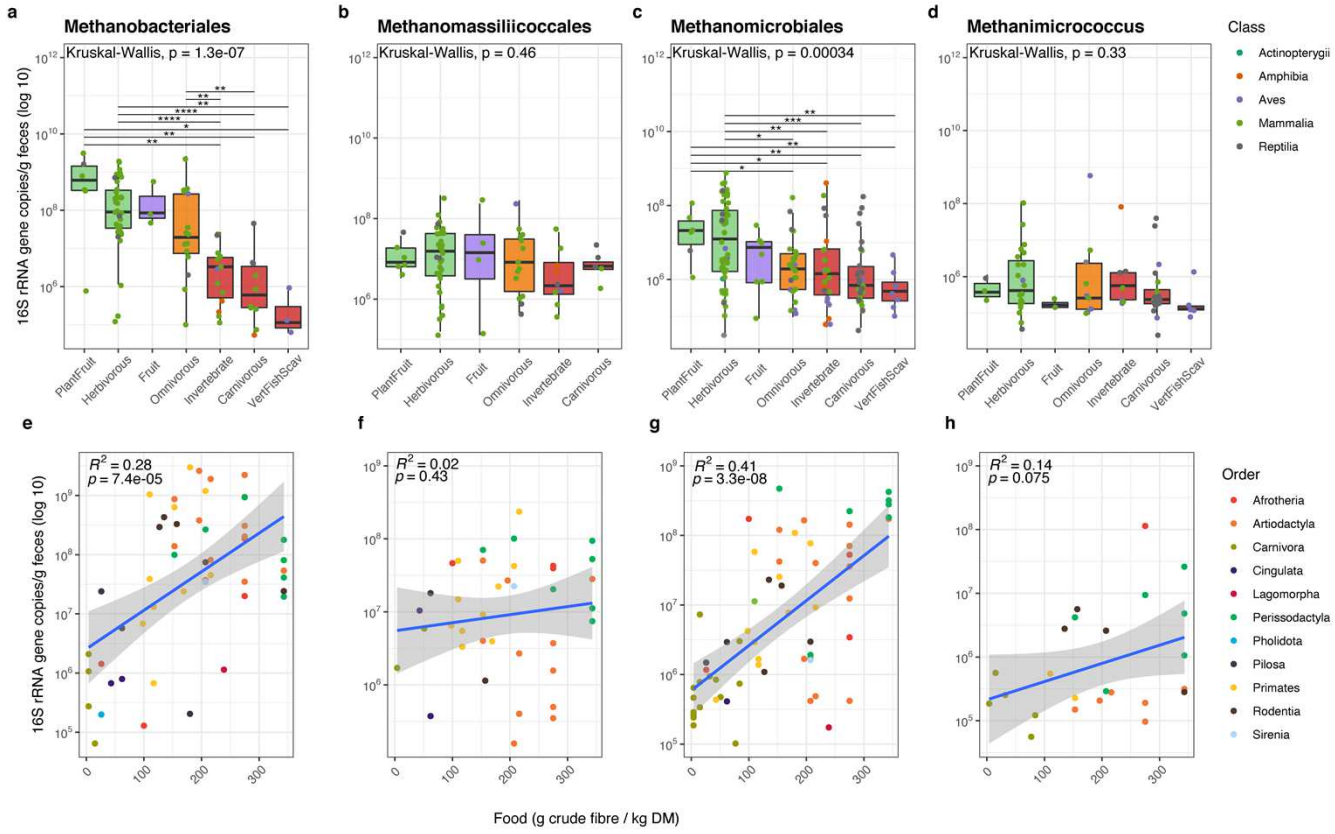

Figure S11: Influence of host diet-type and diet-fibre content on the absolute abundance of four methanogen lineages. Abundance of **a**) Methanobacteriales (n = 91), **b**) Methanomassiliicoccales (n = 73), **c**) Methanomicrobiales (n = 148) and **d**) Methanimicrococcus (n = 69) according to host diets-type. Significant differences across all groups were determined via the Kruskal-Wallis test, $p < 0.05$ is significant. Wilcoxon rank sum test with continuity correction was used to determine differences between diet types *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$. Correlation between diet-fibre content and absolute abundance of **e**) Methanobacteriales (n = 79), **f**) Methanomassiliicoccales (n = 63), **g**) Methanomicrobiales (n = 97) and **h**) Methanimicrococcus (n = 41) in mammal species.

Figure S12.

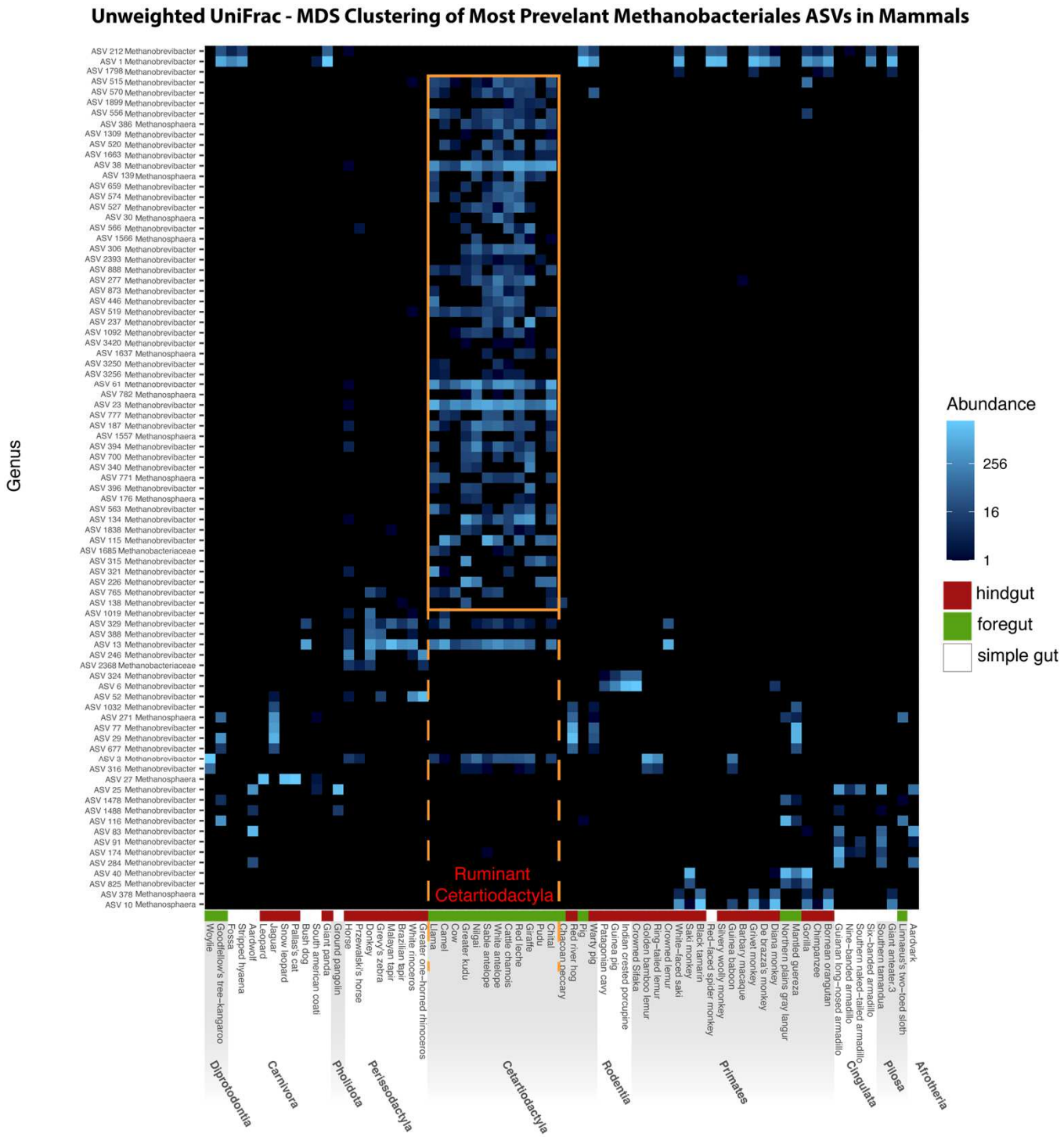

Figure S12: Unweighted UniFrac – MDS clustering of most prevelant Methanobacteriales ASVs in mammals. Methanobacteriales ASVs that are found in ≥ 5% of mammals (n = 67; rarefied 3000 reads/sample). Animals are organized according to mammal phylogeny, and ASVs on y-axis are clustered by the unweighted UniFrac dissimilarity measure. Certiartiodatcyla host a distinct community of Methanobacteriales (orange box).

Figure S13.

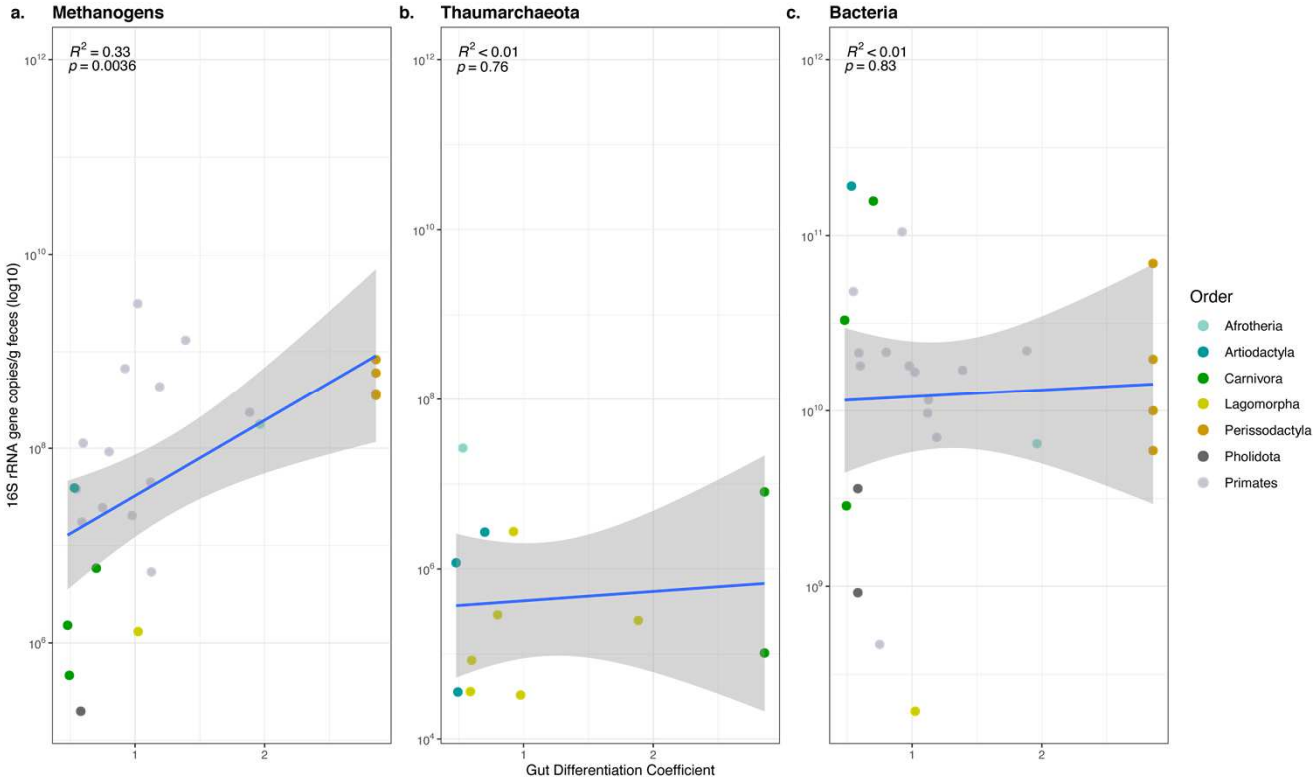

Figure S13: Correlation between gut differentiation coefficient and the absolute abundance of a) methanogens (n = 25), b) Thaumarchaeota (n = 12) and c) bacteria (n = 25) (c) in mammals. The coefficient of gut differentiation corresponds to the surface area of the stomach, caecum, and colon, relative to the small intestine and is related to the proportion of the intestinal tract dedicated to fermentation.

Figure S14.

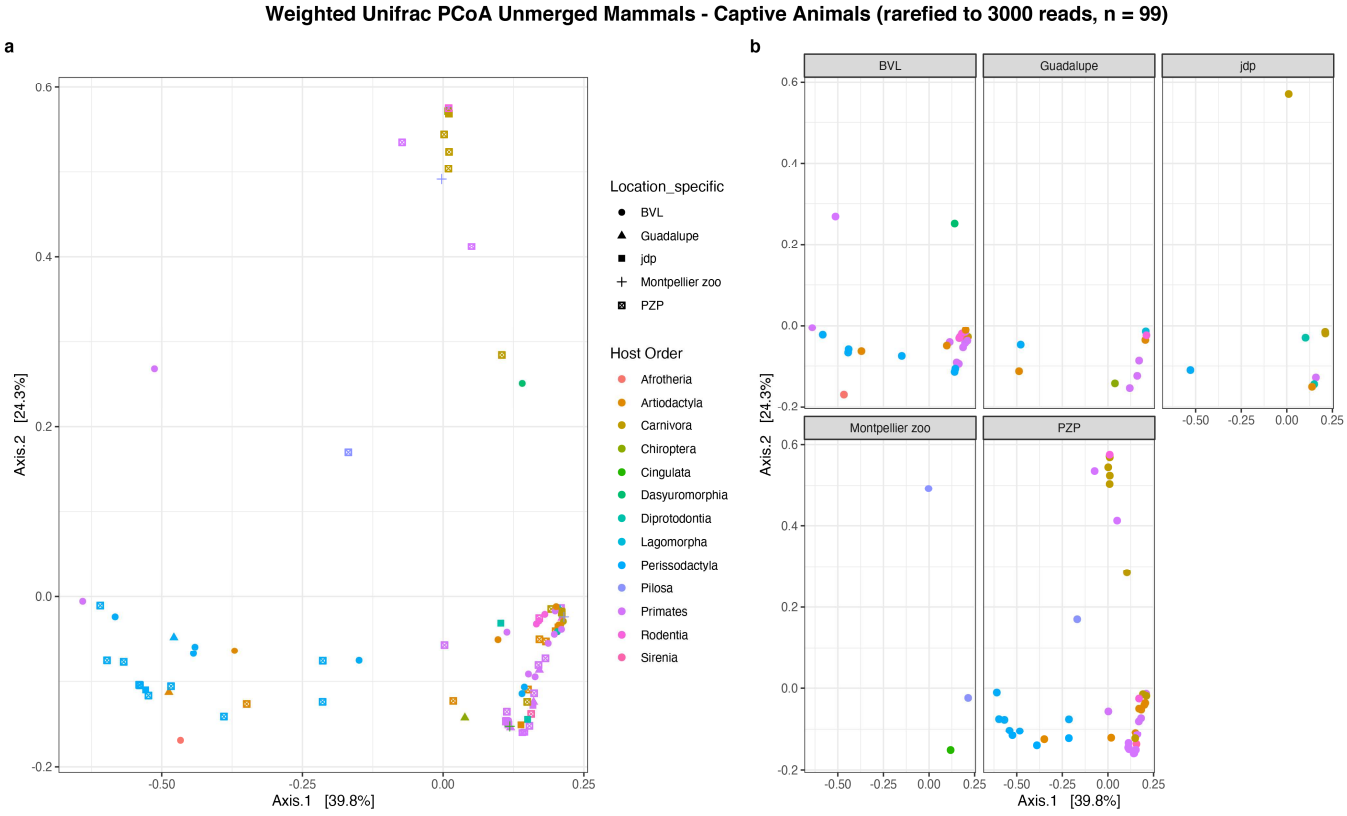

Figure S14: Weighted Unifrac analysis of individual Captive Mammals. Mammals with ≥3000 reads per sample (n = 99) were analyzed using a Weighted Unifrac analysis, location did not significantly explain the variance between the archaeal community in captive mammals ($p > 0.05$).

Figure S15.

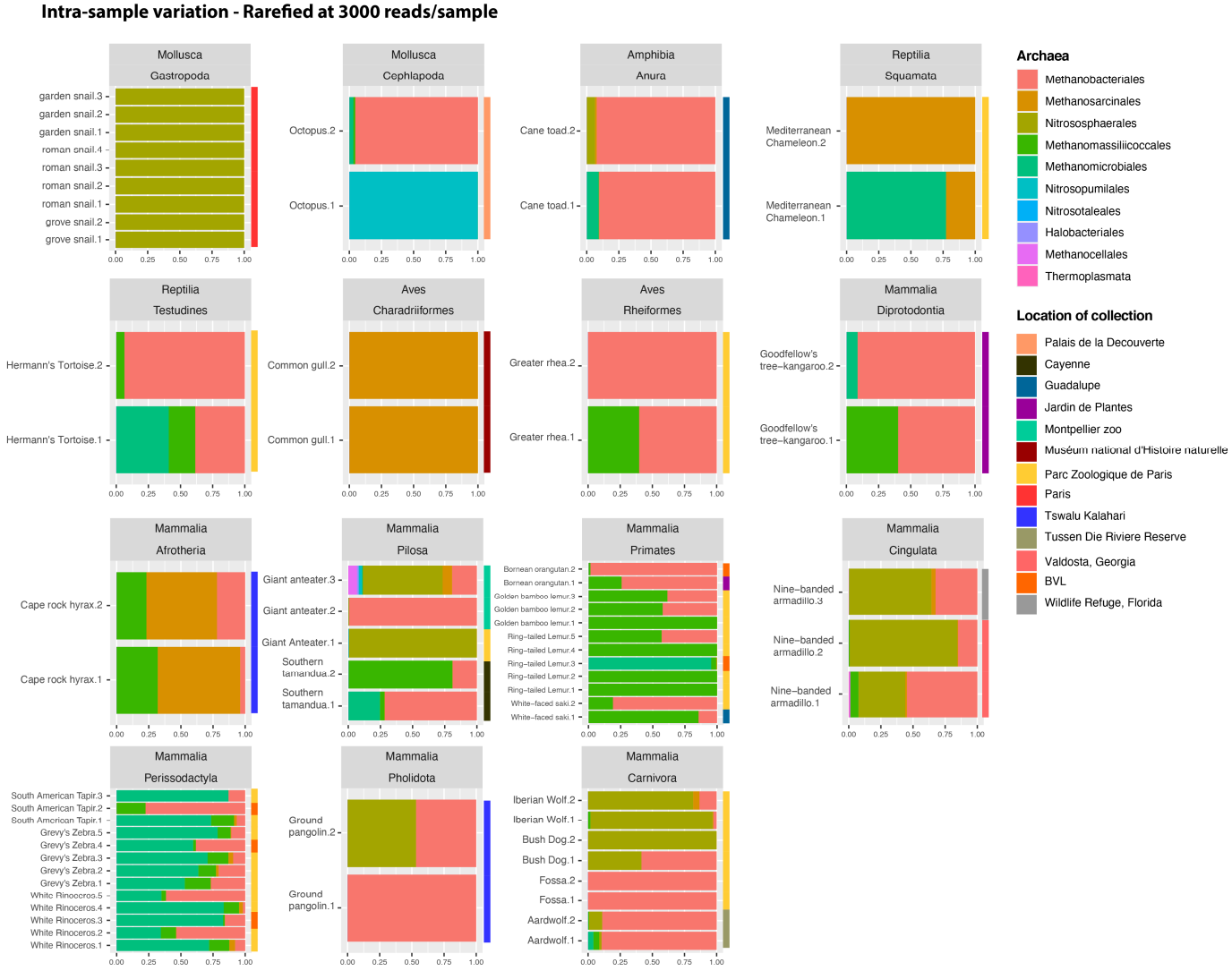

Figure S15: Intra-species variation rarefied at 3000 reads per sample. Animals for which multiple fecal samples were used the variation in relative abundance of archaeal ASVs between samples is low, even between samples from varying locations.

Figure S16.

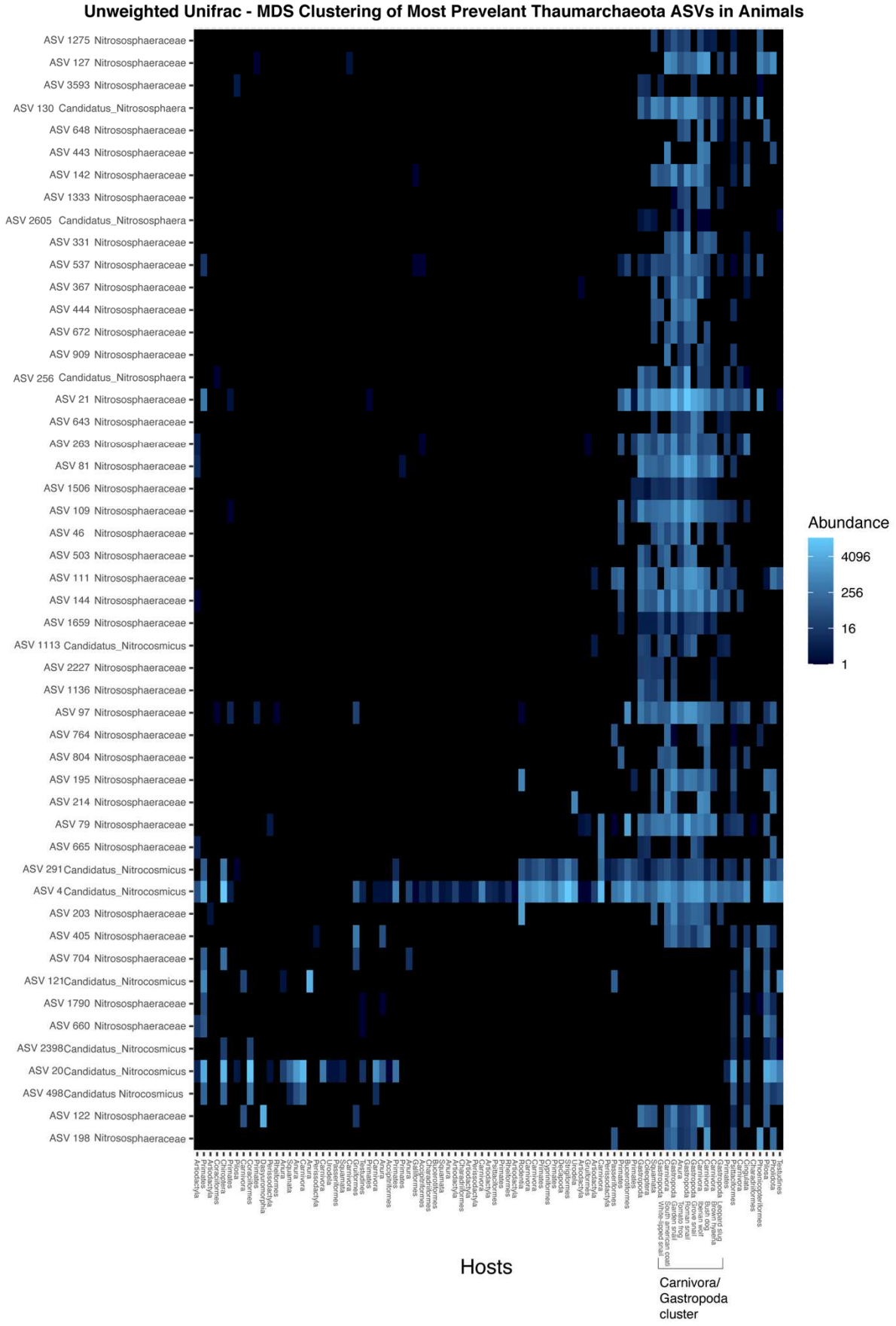

Figure S16: Unweighted UniFrac – MDS Thaumarchaeota of most prevelant Thaumarchaeota ASVs in all animals. Thaumarchaeota ASVs that are found in ≥ 5% of samples (n = 89). Animals and ASVs are clustered by the unweighted UnifFrac dissimilarity measure. Carnivora and Gastropoda host similar Thaumarchaeota communities.

Figure S17.

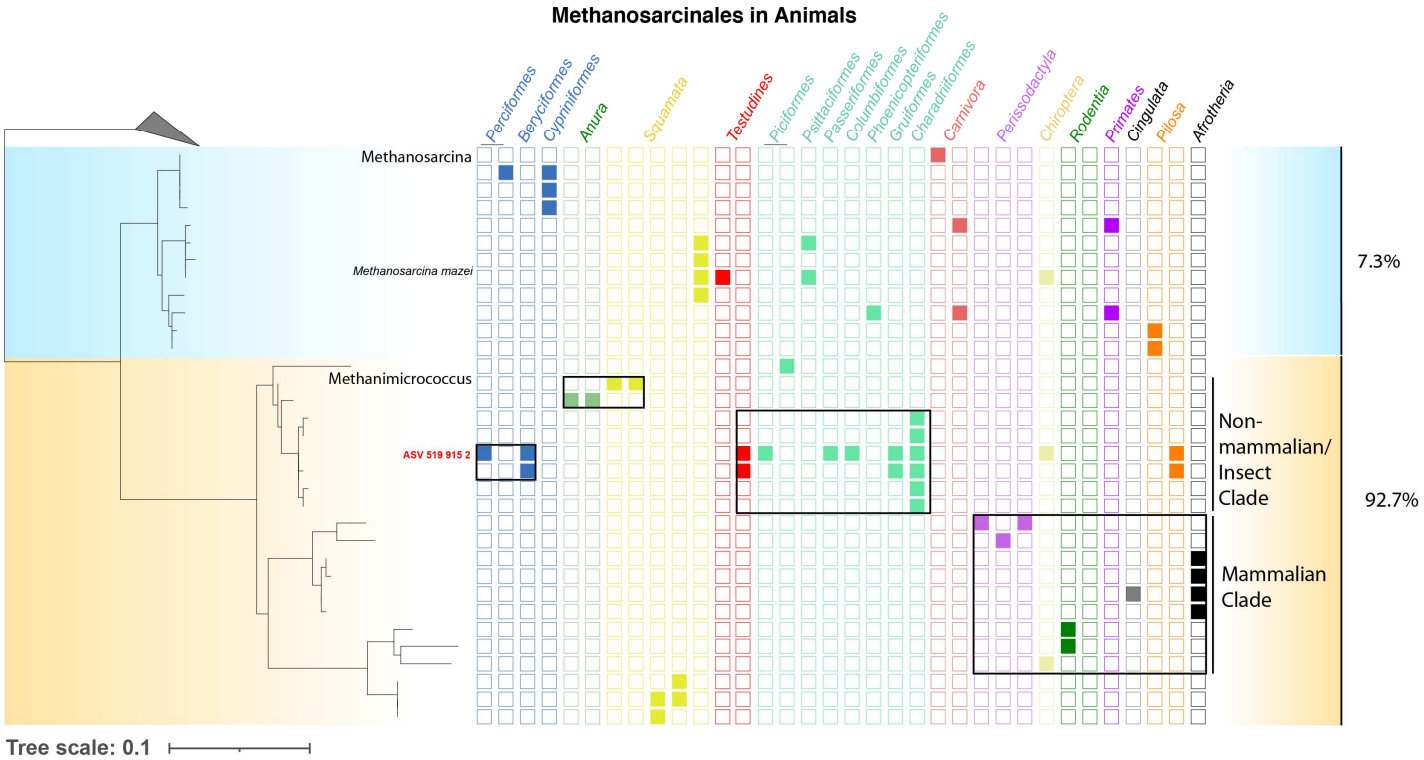

Figure S17: Distribution of Methanosarcinales ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was built with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were then removed from the tree. Only ASV representing more than 1% of the sequences per sample were included. The percentages on the right represent the the proportion of reads from this order that were annotated as Methanimicrococcus and Methanosarcina. Insect and mammalian clades were defined in Thomas et al 2021. Here the insect clade comprises mainly sequences from vertebrates feeding on invertebrates.

Supplementary references

1. Borrel, G. *et al.* Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis. *Genome Biol. Evol.* **5**, 1769–1780 (2013).
2. Söllinger, A. *et al.* Phylogenetic and genomic analysis of Methanomassiliicoccales in wetlands and animal intestinal tracts reveals clade-specific habitat preferences. *FEMS Microbiol. Ecol.* **92**, (2016).
3. Lang, K. *et al.* New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of 'Candidatus Methanoplasma termitum'. *Appl. Environ. Microbiol.* **81**, 1338–1352 (2015).
4. Raymann, K., Moeller, A. H., Goodman, A. L. & Ochman, H. Unexplored Archaeal Diversity in the Great Ape Gut Microbiome. *mSphere* **2**, e00026-17 (2017).
5. Leadbetter, J. R. & Breznak, J. A. Physiological ecology of Methanobrevibacter cuticularis sp. nov. and Methanobrevibacter curvatus sp. nov., isolated from the hindgut of the termite Reticulitermes flavipes. *Appl. Environ. Microbiol.* **62**, 3620–3631 (1996).
6. Leadbetter, J. R., Crosby, L. D. & Breznak, J. A. Methanobrevibacter filiformis sp. nov., a filamentous methanogen from termite hindguts. *Arch. Microbiol.* **169**, 287–292 (1998).
7. Hackstein, J. H. & Stumm, C. K. Methane production in terrestrial arthropods. *Proc. Natl. Acad. Sci.* **91**, 5441–5445 (1994).
8. Sprenger, W. W., Hackstein, J. H. P. & Keltjens, J. T. The energy metabolism of Methanomicrococcus blatticola: Physiological and biochemical aspects. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **87**, 289–299 (2005).
9. Sprenger, W. W., Hackstein, J. H. P. & Keltjens, J. T. The competitive success of Methanomicrococcus blatticola, a dominant methylotrophic methanogen in the cockroach hindgut, is supported by high substrate affinities and favorable thermodynamics. *FEMS Microbiol. Ecol.* **60**, 266–275 (2007).
10. Brune, A. & Dietrich, C. The Gut Microbiota of Termites: Digesting the Diversity in the Light of Ecology and Evolution. *Annu. Rev. Microbiol.* **69**, 145–166 (2015).
11. Sprenger, W. W., Van Belzen, M. C., Rosenberg, J., Hackstein, J. H. P. & Keltjens, J. T. Methanomicrococcus blatticola gen. nov., sp. nov., a methanol- and methylamine-reducing methanogen from the hindgut of the cockroach Periplaneta americana. *Int. J. Syst. Evol. Microbiol.* **50**, 1989–1999 (2000).
12. Thomas, C. M., Taib, N., Gribaldo, S. & Borrel, G. Comparative genomic analysis of Methanimicrococcus blatticola provides insights into host-adaptation in archaea and the evolution of methanogenesis. *ISME Commun.*
13. Lagkouvardos, I. *et al.* Sequence and cultivation study of Muribaculaceae reveals novel species, host preference, and functional potential of this yet undescribed family. *Microbiome* **7**, 1–15 (2019).
14. Ormerod, K. L. *et al.* Genomic characterization of the uncultured Bacteroidales family S24-7 inhabiting the guts of homeothermic animals. *Microbiome* **4**, 1–17 (2016).
15. Kageyama, A. & Benno, Y. Phylogenic and phenotypic characterization of some Eubacterium-like isolates from human feces: description of Solobacterium moorei Gen. Nov., Sp. Nov. *Microbiol. Immunol.* **44**, 223–227 (2000).
16. Ruaud, A. *et al.* Syntrophy via interspecies H2 transfer between christensenella and

methanobrevibacter underlies their global cooccurrence in the human gut. *MBio* **11**, e03235-19 (2020).
17. Biddle, A., Stewart, L., Blanchard, J. & Leschine, S. Untangling the genetic basis of fibrolytic specialization by lachnospiraceae and ruminococcaceae in diverse gut communities. *Diversity* **5**, 627–640 (2013).
18. Wang, W. *et al.* High-throughput sequencing reveals the core gut microbiome of Bar-headed goose (Anser indicus) in different wintering areas in Tibet. *Microbiologyopen* **5**, 287–295 (2016).
19. Manos, J. I. M. & Belas, R. The Genera *Proteus*, *Providencia*, and *Morganella*. 245–269 (2006).
20. Hong, J.-K., Kim, H.-J. & Cho, J.-C. Novel PCR primers for the archaeal phylum Thaumarchaeota designed based on the comparative analysis of 16S rRNA gene sequences. *PLoS One* **9**, e96197 (2014).
21. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **108**, 4516–4522 (2011).
22. Yu, Y., Lee, C., Kim, J. & Hwang, S. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol. Bioeng.* **89**, 670–679 (2005).
23. Gaci, N., Chaudhary, P. P., Tottey, W., Alric, M. & Brugère, J.-F. Functional amplification and preservation of human gut microbiota. *Microb. Ecol. Health Dis.* **28**, 1308070 (2017).
24. Amann, R. I., Ludwig, W. & Schleifer, K.-H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**, 143–169 (1995).
25. Gantner, S., Andersson, A. F., Alonso-Sáez, L. & Bertilsson, S. Novel primers for 16S rRNA-based archaeal community analyses in environmental samples. *J. Microbiol. Methods* **84**, 12–18 (2011).
26. Lane, D. J. 16S/23S rRNA sequencing. *Nucleic acid Tech. Bact. Syst.* 115–175 (1991).
27. Nossa, C. W. *et al.* Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. *World J. Gastroenterol. WJG* **16**, 4135 (2010).
28. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).
29. Apprill, A., McNally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* **75**, 129–137 (2015).
30. Casamayor, E. O. *et al.* Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern. *Environ. Microbiol.* **4**, 338–348 (2002).
31. Ovreås, L., Forney, L., Daae, F. L. & Torsvik, V. Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Appl. Environ. Microbiol.* **63**, 3367–3373 (1997).
32. Stahl, D. A. & Amann, R. I. Development and application of nucleic acid probes in bacterial systematics. in *Nucleic acid techniques in bacterial systematics* (eds. Stackebrandt, E. & Goodfellow, M.) 205–248 (John Wiley & Sons, 1991).