Archaea du microbiome intestinal animal : abondance, diversité et adaptation

Courtney Thomas

Université de Paris

École doctorale Bio Sorbonne Paris Cité (BioSPC) 562

Laboratoire Evolutionary Biology of the Microbial Cell

Archaea of the animal intestinal microbiome: abundance, diversity, and adaptation

Courtney M. Thomas

Thèse de doctorat de Microbiologie

Dirigée par Simonetta Gribaldo

Présentée et soutenue publiquement le 2 Décembre 2021

Rapporteurs : [Schleper Christa / Professor / University of Vienna] & [Brune Andreas / Professor / Max Planck Institute for Terrestrial Microbiology]

Examinateurs : [Kish Adrienne / Maitre de conférences / MNHN, Université de Paris] & [Brugère Jean-François / Maitre de conférences (HDR) / Université Clermont Auvergne]

Directeur de thèse : [Gribaldo Simonetta / Professor / Institut Pasteur]

Membre invités : [Borrel Guillaume / Chargé de Recherche / Institut Pasteur]

Archaea of the Animal Intestinal Microbiome: abundance, diversity, and adaptation

Résumé :

Les archées représentent une fraction peu étudiée du microbiote intestinal. Des études menées sur quelques groupes de mammifères et d'insectes (primates, ruminants et termites) ont mis en évidence que les archées intestinales sont en grande partie constituée de méthanogènes affiliées aux Methanobacteriales et aux Methanomassiliicoccales. Cependant, en raison de l'absence d'une étude incluant de nombreux animaux, plusieurs aspects importants restent mal compris : 1) quelle est la diversité globale des archées intestinales ; 2) quels facteurs dictent leur distribution et leur abondance ; et 3) quelles adaptations ont conduit à la colonisation du milieu digestif. Ce travail vise à répondre à ces questions fondamentales.

Pour caractériser les communautés d'archées associées à l'intestin, j'ai quantifié et séquencé les gènes de l'ARNr 16S de la communauté microbienne en utilisant des amorces universelles et spécifiques des archées dans près de 400 échantillons fécaux provenant de 269 espèces couvrant six grands groupes d'animaux - Aves, Amphibia, Actinopterygii, Mammalia, Reptilia, et divers invertébrés.

Nos données montrent que les archées sont communes dans le microbiote intestinal des animaux. Les *Methanobrevibacter, Methanosphaera, Methanomethylophilaceae, Methanimicrococcus et Methanocorpusculum* composent la majeure partie de l'archéome intestinal. Ceci indiquant notamment quatre à cinq événements majeurs d'adaptation à l'intestin chez les Archaea. Bien que moins abondantes, les *Nitrososphaeraceae* oxydant l'ammoniac sont aussi largement distribuées dans nos échantillons et correspondent aux phylotypes d'archées dominants du sol. Leur statut (résident ou en transit) dans le milieu digestif demeure incertain. J'ai déterminé que la phylogénie de l'hôte et le régime alimentaire jouent un rôle majeur dans l'abondance et la structuration des d'archées intestinales chez les mammifères. L'abondance des méthanogènes est fortement corrélée au contenu en fibre du régime alimentaire. Les bactéries dégradant la pectine (produisant du méthanol) cooccurrent avec les méthanogènes dépendants du méthanol et sont liées à la consommation de fruits. Les reptiles et les mammifères hébergent des populations d'archées plus importantes que les autres animaux échantillonnés. Enfin, certaines archées sont dominantes chez des ordres spécifiques d'animaux, ce qui suggère des adaptations spécialisées à ces hôtes.

Afin de mieux comprendre l'adaptation des archées à l'intestin, j'ai utilisé des approches de génomique comparative. Dans un premier temps je me suis concentré sur le génome de *Methanimicrococcus blatticola* PA, affilié à une lignée encore peu connue dans l'intestin (Methanosarcinales). Mes analyses ont révélé que cette espèce présente la plus grande réduction génomique au sein des Methanosarcinales. Les pertes de gènes ont affecté de nombreuses protéines impliquées dans la perception de l'environnement. Cette transition a également entraîné un changement de la principale voie de méthanogenèse et une modification substantielle des protéines de surface. Plusieurs de ces adaptations correspondent à celles précédemment observées chez des archées et des bactéries intestinales phylogénétiquement éloignées, suggérant des mécanismes convergents à grande échelle. Afin de compléter ces approches, j'ai réalisé un séquençage métagénomique en shotgun sur 58 espèces et reconstruit des génomes d'archées à partir de ces métagénomes. L'analyse de ces données devrait révéler d'autres informations clés sur la façon dont les archées se sont adaptées à l'intestin des animaux.

Ce travail est l'une des premières études à grande échelle sur la communauté d'archée intestinale d'un large éventail d'animaux. Mes résultats ont permis d'identifier les principales archées intestinales et les dynamiques évolutives associées à leur adaptation à cet environnent, ainsi que les facteurs contrôlant leur abondance et diversité.

Mots clefs : Archaea, microbiote intestinal, génomique comparative, méthanogènes, archaeome instestinal

Archaea of the Animal Intestinal Microbiome: abundance, diversity, and adaptation

Abstract:

The Archaea represent an understudied fraction of the animal intestinal microbiota. Studies conducted on few groups of mammals and insects (mainly primates, ruminants, and termites), have resulted in the current understanding that the intestinal archaeal community largely consists of methane-producing lineages belonging to the Methanobacteriales and Methanomassiliicoccales. However, due to the lack of a large-scale archaea-centric animal intestinal microbiome study, several fundamental aspects of the intestinal archaeal community remain poorly understood: 1) how diverse are intestinal archaea; 2) what factors dictate their abundance/distribution/community structure; and 3) what genomic adaptations have driven intestinal colonization. This work aimed to answer these fundamental questions.

To characterize gut-associated archaeal communities, I quantified and sequenced the 16S rRNA genes of the intestinal microbial community using both universal and archaeal-specific primers in almost 400 fecal samples from 269 species covering six major animal groups - Aves, Amphibia, Actinopterygii, Mammalia, Reptilia, and invertebrates. Our data shows that archaea are common constituents of the animal intestinal microbiome. Four genera and one family of methanogens (*Methanobrevibacter, Methanosphaera, Methanomethylophilaceae, Methanimicrococcus and Methanocorpusculum*) compose the bulk of the intestinal archaeome. I also identified four-to-five independent major adaptation events to the gut in the archaeal domain, that occurred at the base of these methanogen lineages. While less abundant, ammonia-oxidizing *Nitrososphaeraceae* are also widely distributed in our samples and correspond to the dominant archaeal phylotypes from soil, suggesting that they could either be transient in the gut, or, could be resident gut microbes, which may aid in their dispersal in the environment. These data also show that host diet and phylogeny are major factors influencing the structure and abundance of the archaeal community in mammals. More specifically, dietary fibre content is a major factor determining methanogen abundance. Also, pectin-degrading (methanol-producing) bacteria cooccur with methanol-dependent methanogens and are related with fruit consumption. Reptiles and mammals host higher concentrations of archaea than other animals. Finally, the dominance of some archaea in specific animal orders suggests specialized adaptations to these hosts.

In order to better understand the genomic adaptations of archaea to the gut, I used comparative genomics approaches. First, I focused on the genome of *Methanimicrococcus blatticola* PA, affiliated to a poorly understood gut lineage, the Methanosarcinales. My analyses revealed that this species has undergone the greatest genome reduction within the order Methanosarcinales as a result of intestinal colonization. Gene losses affected many proteins involved in sensing the environment. This transition also led to a change in the main methanogenesis pathway and substantial modification of cell surface proteins. Several of these adaptations parallel those previously observed in phylogenetically distant archaea and bacteria from the animal microbiome, suggesting large-scale convergent mechanisms.
To obtain further insights into the animal archaeome, I performed shotgun metagenomic sequencing on 58 species and reconstructed archaeal genomes from these metagenomes. Ongoing comparative analysis of these data is expected to reveal additional information on genomic adaptations of diverse intestinal archaea.

This work is one of the first large-scale studies of the gut archaeal community in a wide range of animals. My results provide key information concerning the diversity of intestinal archaea and the evolutionary dynamics associated with their adaptation to this environment.

Keywords: Archaea, gut microbiota, comparative genomics, methanogens, gut archaeome

This work is dedicated to my late grandfather, Robert Leeper.

Acknowledgements

This thesis would not have been possible without the overwhelming support of my friends, family, and colleagues.

I would first like to thank my supervisors, Guillaume Borrel and Simonetta Gribaldo, whose patience and expertise were fundamental to the success of this work. Your extensive knowledge of the field pushed me to develop as a researcher and consistently elevated my work. I cannot express my gratitude and respect for you both enough.

I would also like to acknowledge my colleagues who quickly became close friends. Without their steadfast support, I do not believe I would have succeeded in this academic endeavor. Thank you all for always reminding me to take a break. Najwa Taib, Nika Pende, Daniela Megrian, and Pierre Garcia, the last three years has been filled with laughter and growth largely due to all of you. In addition to my colleagues, I want to acknowledge all of my friends abroad, especially Chandler Countryman and Maria Letourneau, – I am truly grateful for your continued long-distance emotional support.

I would like to thank my loving parents, grandparents, and siblings. Your support in my decision to move abroad and understanding of all the complexities that came with it, has made all of this possible. I am truly blessed to call you all family.

Finally, I would like to thank my partner, Alexey Vorobev. Without you, I would never have started this adventure, and I am forever grateful to have had your love and support.

Table of Contents

Table of Figures and Tables

CHAPTER 2

Disclaimer

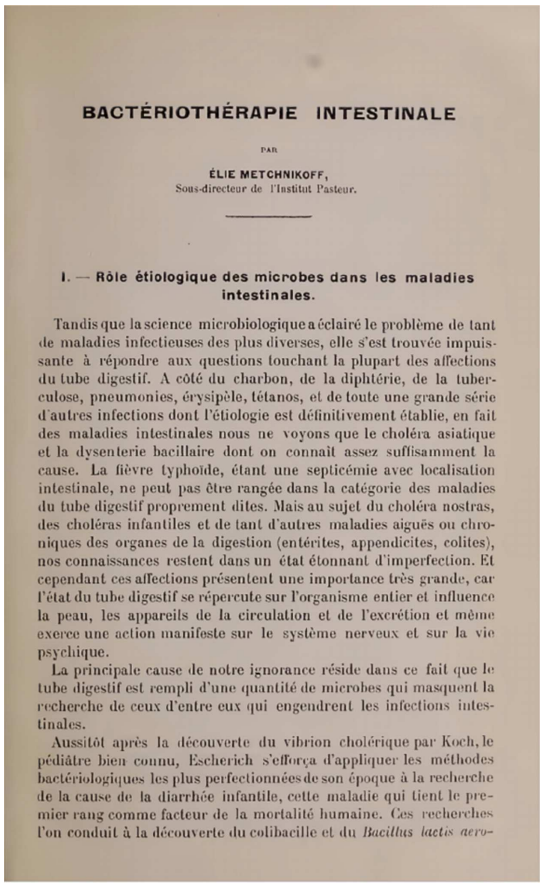
I declare that the work in this thesis is my own. All the material obtained from the literature or from other sources have been indicated as such. I have provided a full bibliography of all the texts referenced.

Je déclare que le travail de cette thèse est le mien. Tout le matériel obtenu de la littérature ou d'autres sources a été indiqué comme tel. J'ai fourni une bibliographie complète de tous les textes référencés.

Introduction

The Intestinal Microbiota

As early as the 1900s, researchers like Élie Metchnikoff predicted that the microbes inhabiting the intestinal tract were important for human health and preventing disease (Metchnikoff *et al.*, 1909; Mowat, 2021). In his essay, *Bactériothérapie intestinale* **(Introduction Figure 1)**, Metchnikoff put forth the hypothesis that a large part of intestinal microbes are actually beneficial for the organism they inhabit and that these microbes could be used to treat intestinal diseases (Metchnikoff *et al.*, 1909), an early suggestion of the usefulness of probiotics. Of course, since then, the understanding of the complex consortium of microbes - bacteria, archaea, fungi, and viruses - that make up the intestinal microbiota has drastically changed (Turnbaugh et al. 2007; Arumugam et al. 2011; Barko et al. 2018; Borrel et al. 2020).



BACTÉRIOTHÉRAPIE INTESTINALE

PAR

ÉLIE METCHNIKOFF,
Sous-directeur de l'Institut Pasteur.

I. — Rôle étiologique des microbes dans les maladies intestinales.

Tandis que la science microbiologique a éclairé le problème de tant de maladies infectieuses des plus diverses, elle s'est trouvée impuissante à répondre aux questions touchant la plupart des affections du tube digestif. A côté du charbon, de la diphtérie, de la tuberculose, pneumonies, érysipèle, tétanos, et de toute une grande série d'autres infections dont l'étiologie est définitivement établie, en fait des maladies intestinales nous ne voyons que le choléra asiatique et la dysenterie bacillaire dont on connait assez suffisamment la cause. La fièvre typhoïde, étant une septicémie avec localisation intestinale, ne peut pas être rangée dans la catégorie des maladies du tube digestif proprement dites. Mais au sujet du choléra nostras, des choléras infantiles et de tant d'autres maladies aiguës ou chroniques des organes de la digestion (entérites, appendicites, colites), nos connaissances restent dans un état étonnant d'imperfection. Et cependant ces affections présentent une importance très grande, car l'état du tube digestif se répercute sur l'organisme entier et influence la peau, les appareils de la circulation et de l'excrétion et même exerce une action manifeste sur le système nerveux et sur la vie psychique.

La principale cause de notre ignorance réside dans ce fait que le tube digestif est rempli d'une quantité de microbes qui masquent la recherche de ceux d'entre eux qui engendrent les infections intestinales.

Aussitôt après la découverte du vibrion cholérique par Koch, le pédiâtre bien connu, Escherich s'efforça d'appliquer les méthodes bactériologiques les plus perfectionnées de son époque à la recherche de la cause de la diarrhée infantile, cette maladie qui tient le premier rang comme facteur de la mortalité humaine. Ces recherches l'on conduit à la découverte du colibacille et du *Bacillus lactis aero-*

Figure 1. The title page of the Metchnikoff essay, reviewed in Mowat 2021.

Because bacteria compose the majority of the microbial community found in the intestinal tract, the most is known about them. For example, intestinal bacteria produce critical vitamins and short chain fatty acids (e.g., butyrate), aid in the development of the immune system, and may even influence the development of organs in their hosts (Pryde *et al.*, 2002; Scott *et al.*, 2006; Hooper *et al.*, 2012; Kamada *et al.*, 2013; Koh *et al.*, 2020). Intestinal microbes have also been shown to impact the central nervous system of their hosts by modulating the release of Short Chain Fatty Acids (SCFA) and tryptophan, altering their host's mood and cognition (Cryan *et al.*, 2012). Further, researchers have started to identify host-related factors that determine the nature of the bacterial community in different animals. More specifically, there have been individual events of convergence of bacterial communities in flying animals (Song *et al.*, 2020), insect-eating animals (Delsuc *et al.*, 2014), as well as blood-eating vertebrates (Song *et al.*, 2019). Furthermore, it has been found that some lineages of bacteria have inhabited the intestinal tract long enough to have co-speciated with their host (Moeller *et al.*, 2016; Youngblut *et al.*, 2019; Groussin *et al.*, 2020). For example, predominant gut bacteria belonging to the *Bacteroidaceae* and the *Bifidobacteriaceae* have coevolved with their host Hominidae species over several million years (Moeller *et al.*, 2016). Over evolutionary time, the entire microbial community can reflect the host phylogeny, a phenomenon known as phylosymbiosis (Lim *et al.*, 2020). Within the Mammalia, phylosymbiosis has been detected using beta-diversity measures of the bacterial community, illustrating that host evolution significantly impacts the diversity of bacteria in the intestinal tract (Youngblut *et al.*, 2019; Groussin *et al.*, 2020).

Some intestinal bacterial lineages have become genetically distinct from their free-living relatives, suggesting specialization to the gut environment. When these lineages are placed in a phylogeny, they form unique 'host-associated' clades. Examples of this are seen in the phyla Kiritimatiellaeota (Spring *et al.*, 2016) and Melainabacteria (Di Rienzi *et al.*, 2013), as well as in the *Muribaculaceae* (phylum Bacteroidetes) (Lagkouvardos *et al.*, 2019). While the genomic content and metabolic capacity of Kiritimatiellaeota are still poorly understood, they have been identified as being a part of the 'core' intestinal microbiome in equine species (Edwards *et al.*, 2020), and thus likely play an important role in the microbial community of these animals. The Melainabacteria are predicted to produce ethanol and formate, and are also thought to be important H_2 producing microbes due to the presence of genes encoding for

trimeric confurcating Fe-Fe hydrogenases in their genomes (Di Rienzi *et al.*, 2013). Finally, the *Muribaculaceae* are unique because they include distinct phylogenetic clades specialized on different types of glycans that they can degrade (Lagkouvardos *et al.*, 2019).

Because bacteria have been the main focus of most intestinal microbiota studies, other constituents of this complex community have been largely overlooked. This is notably the case of the Archaea.

Archaea – the third domain of life

Archaea were recognized as evolutionarily distinct from Bacteria over 40 years ago (Woese *et al.*, 1977), and originally believed to be microorganisms specifically adapted to inhabit environments that pushed our understanding of where life is possible - places of extreme temperature, salinity, and pH (Rothschild *et al.*, 2001; Gribaldo *et al.*, 2006; Valentine, 2007). Today, we know that Archaea thrive in a huge range of environments, from freshwater to soil (Delong, 1998; Schleper *et al.*, 2005; Adam *et al.*, 2017). At the molecular level, Archaea share properties with both Eukaryotes and Bacteria. For example, they have eukaryote-like RNA and DNA polymerases, but their genomes contain operons like Bacteria **(Introduction Figure 2)**. Archaea are also unique as they are the only microbes with the ability to produce methane gas **(Introduction Figure 2)**.

Feature	Bacteria	Archaea	Eukaryotes
Nucleus	No	No	Yes
Organelles	No	No	Yes
Spliceosomal introns	No	No	Yes
Chromosome shape	Circular and linear	Circular	Linear
Operons	Yes	Yes	Rare
RNA polymerase	Bacteria-like	Eukaryote-like	Eukaryote-like
DNA polymerase	Bacteria-like	Eukaryote-like	Eukaryote-like
Ribosome type	70S	70S	80S
Translation start (amino acid)	Formylmethionine	Methionine	Methionine
Histones	No	Yes	Yes
Peptidoglycan (PG)	Yes	No; pseudo-PG in some	No
Motility	Bacteria-type flagellum	Archaea-type flagellum (archaellum)	Eukarya-type flagellum
Lipopolysaccharide	Yes	No	No
Membrane lipids	Ester links (glycerol-1-phosphate backbone)	Ether links (glycerol-3-phosphate backbone)	Ester links (glycerol-1-phosphate backbone)
Methanogenesis	No	Yes	No
Oxygenic photosynthesis	Yes	No	Yes
Spores	Yes	No	Yes
Human pathogenicity	Yes	No	Yes

Figure 2. Unique and shared characteristics of Archaea. Adapted from Borrel *et al* 2020.

I. Methanogenesis, an ancient Archaea-specific metabolism

Methanogenic archaea (methanogens) represent an ecologically and evolutionarily important group and are also the most commonly encountered archaea in the intestinal microbial community (Wasburn *et al.*, 1937; Oremland, 1979; Hackstein *et al.*, 1994; Gaci *et al.*, 2014; Borrel *et al.*, 2020). Evolutionarily, methanogenesis is believed to be one of the oldest microbial metabolisms - perhaps as old as 3.5 billion years ago according to carbon isotope analyses of fluid inclusions containing methane (Ueno *et al.*, 2006). Furthermore, the last common ancestor of the entire domain Archaea may have been a methanogen (Raymann *et al.*, 2015; Borrel *et al.*, 2019; Wang *et al.*, 2021). Methane also represents an ecologically important biogas as it is a potent greenhouse gas, and economically, methanogens in livestock reduce energy available to be absorbed by the animal host (Liu *et al.*, 2008; Enzmann *et al.*, 2018).

Methanogens use a limited number of substrates for their metabolism and historically have been defined into three categories based on the electron donor for methanogenesis: (i) hydrogenotrophic, (ii) methylotrophic, and (iii) acetoclastic (Liu *et al.*, 2008; Borrel *et al.*, 2019). Until recently, "hydrogenotrophic methanogenesis" was exclusively used to refer to the reduction of CO_2 with H_2 to generate methane gas. However, the discovery of a large range of hydrogenotrophic methyl-reducing methanogens has called for a clarification of this term. Thus, there is now a distinction between two types of hydrogenotrophic methanogens: the CO_2-reducing hydrogenotrophic and the methyl-reducing hydrogenotrophic methanogens **(Introduction Figure 3)**. Methanogenic archaea are easily categorized into metabolic groups by taxonomic annotation with the exception of the *Methanosarcina* who can perform all of these **(Introduction Figure 4)**.

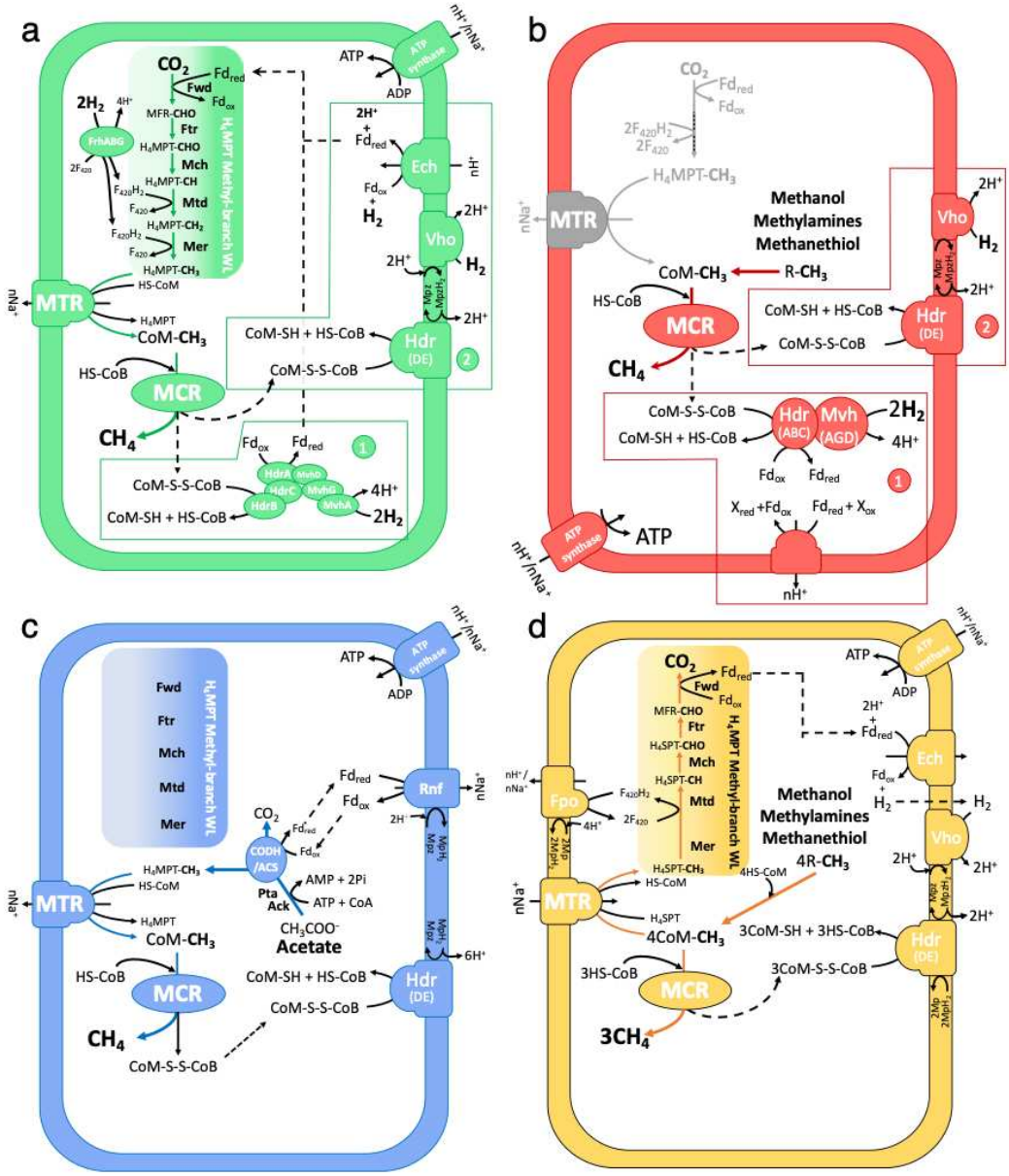


Figure 3. Main methanogenesis pathways. **a)** CO_2-reducing hydrogenotrophic methanogenesis in most methanogens without (1) and with (2) cytochromes. **b)** Methyl-reducing hydrogenotrophic methanogenesis in most methanogens with (1) and with (2) cytochromes. Enzymes in grey can be fully present, partially present or totally absent among methanogens with this metabolism. **c)** Acetoclastic methanogenesis in *Methanosarcina acetivorans*. **d)** Methylotrophic methanogenesis in *Methanosarcina spp.*

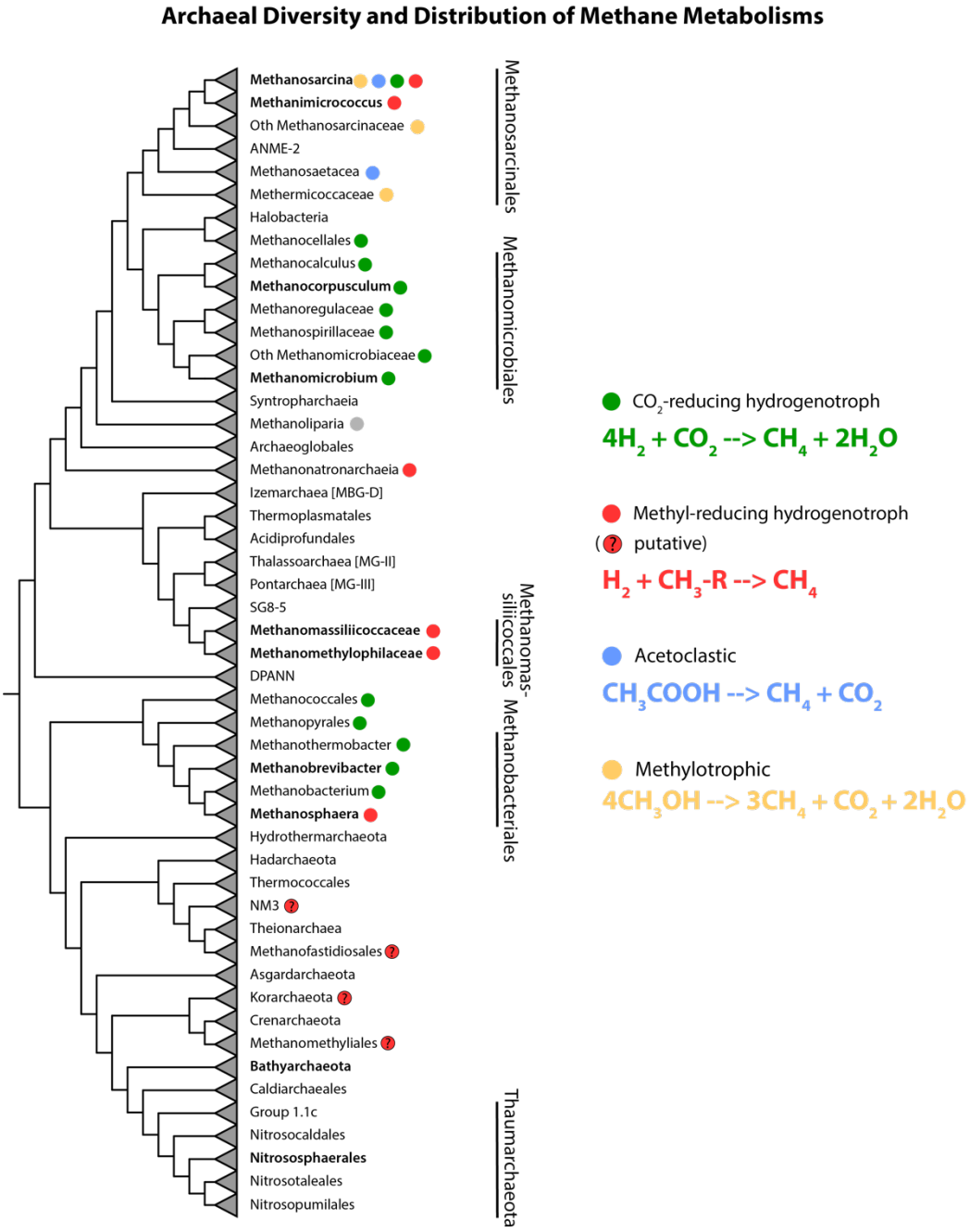


Figure 4. Distribution of Methane Metabolisms in the Archaea. The main types of methanogenesis are indicated by colored dots and detailed on the right. Methanoliparia (grey dot) likely use multi-carbon alkanes as substrates for methanogenesis. Names in bold indicate lineages previously identified in the intestinal tract of animals. The '?' indicates lineages where methanogenesis was predicted via genomic approaches only.

II. Archaea in the intestine

The presence of methanogens in the animal gut was known a century before the recognition of the domain Archaea. Indeed, methane was first detected in the rumen in 1875 and its

formation was attributed to microorganisms (see Washburn and Brody 1937 and references within). Much later, methane gas detection approaches were applied to human flatus and fecal samples (Calloway *et al.*, 1966). Then, methane gas measurement through breath tests was used to detect the presence of methane-producing microbes in the intestinal tract of humans. This breath test was used to investigate how various factors such as age, carbohydrate malabsorption, diseases, and genetics influence methanogens and if methane production can help identifying some of these physiological states (Bond *et al.*, 1971; Pimentel *et al.*, 2006; Polag *et al.*, 2014).

Eventually, methane gas detection was used in the 1990s to explore the distribution of methane-producing microbes in faeces/intestines of a variety of animals (vertebrates and invertebrates) (Hackstein and Stumm, 1994; Hackstein and van Alen, 1996). This led to the hypothesis that symbiosis between animals and methanogens is mainly driven by evolution rather than dietary or physiological constraints. A classic example provided in this study was the mammalian order Carnivora, which was determined to be a lineage of mammals unable to host intestinal methanogens **(Introduction Figure 5)**. The hypothesis stated that once the methanogen-hosting trait was lost, it could never be re-acquired. Thus, even Carnivora species that shifted their dietary niche to plants such as the *Ailuropoda* (panda) were considered 'non-producers' in terms of methane gas (Hackstein *et al.*, 1994, 1996). To-date, the data gathered from methane-gas detection studies have largely shaped our understanding of the distribution of methanogenic archaea in the animal intestine.

Of course, other "classical" approaches have also been used to better understand intestinal methanogens. For example, anaerobic culture methods have led to the identification of numerous intestinal methanogens from the Methanobacteriales order (Nottingham *et al.*, 1968; Miller *et al.*, 1982), among others (Biavati *et al.*, 1988; Sprenger *et al.*, 2000). Microscopy has also provided information concerning methanogens in the insect hindgut, showing that they are often closely associated to the host-intestinal wall (Hackstein *et al.*, 1994; Leadbetter *et al.*, 1996; Sprenger *et al.*, 2000; Bang *et al.*, 2014). All of these methods continue to aid in understanding the physiology/niche of intestinal archaea but remain extremely industrious. Moreover, these techniques usually focused investigation on a single archaeal species.

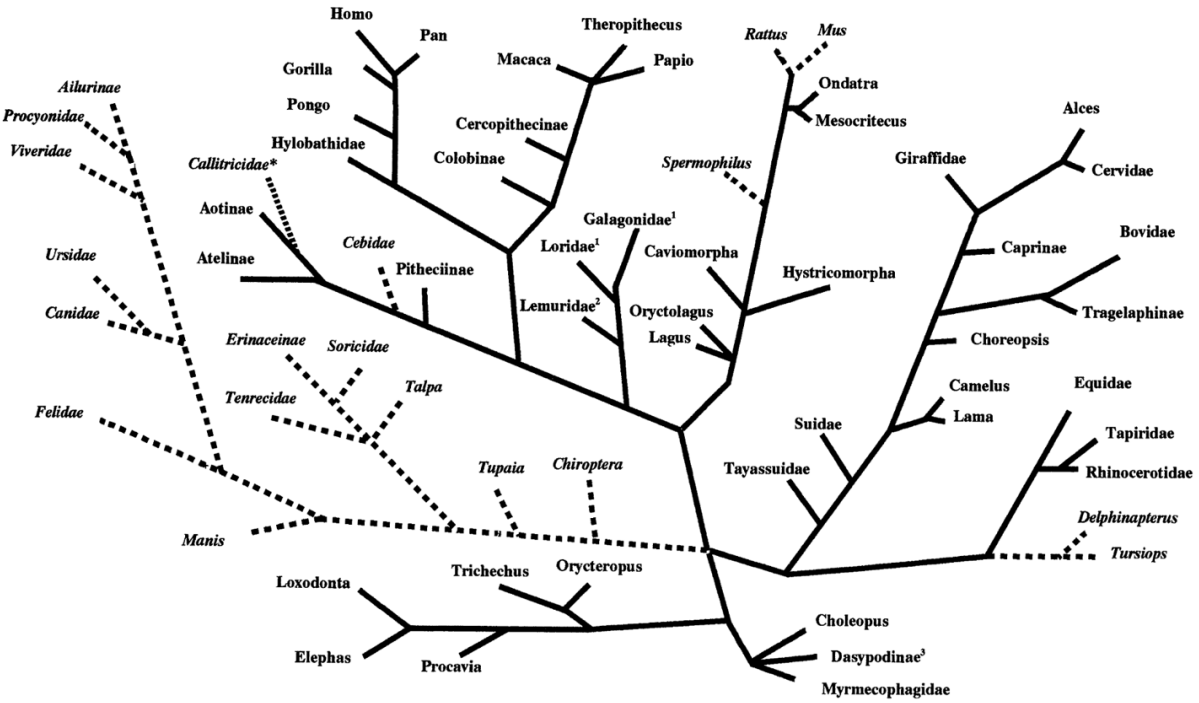


Figure 5. Distribution of methane producers within Mammalia. Dash lines indicate groups of mammals that were considered 'non-producers'. Adapted from Hackstein and van Alen 1996.

II.a Sequenced-based techniques accelerate the study of intestinal archaea

Studies on the intestinal microbiome generally rely on gene amplicon sequencing (metataxonomic) and/or shotgun sequencing (metagenomic). Gene amplicon sequencing targets one gene of interest (e.g., 16S rRNA gene) and results in short sequences (reads) that provide taxonomic information about the microorganisms present in the community. Using amplicon sequencing approaches, information on archaea in the gut microbiome is either obtained with primer sets targeting the overall prokaryotic community (Bacteria and Archaea) or only the domain Archaea. However, some of the so-called universal primers can be biased against archaea and lead to their under-representation in this type of study, as shown for example in a study on the primate microbiome (Raymann *et al.*, 2017). In other cases, archaeal reads were simply discarded from the analysis (Chung *et al.*, 2020; Guo *et al.*, 2021; Steiner *et al.*, 2021).

Studies using archaeal-specific primers have consistently found methanogens in the digestive tract, oral cavity, and vagina of humans (Gaci *et al.*, 2014; Borrel *et al.*, 2020), as well as in the

intestinal tract of ruminants, primates, and some insects (St-Pierre *et al.*, 2013; Brune, 2018; Borrel *et al.*, 2020). Shotgun metagenomic sequencing results in reads from all of the genomes present in the microbial community. These reads are assembled into longer sequences (contigs) that can reach more than 1 Mbp. Data from metagenomic approaches have been used in two ways to study gut archaea: *i*) to compare the relative abundance of specific methanogenic pathways between different animals or conditions (e.g., feeding), and *ii*) to reconstruct archaeal genomes (called metagenome-assembled genomes or MAGs) to characterize their metabolic potential and identify adaptations to the gut.

Sequencing approaches have provided key insights into the lifestyle and diversity of methanogens in the intestinal tract of some animals. They have notably revealed an important fraction of uncultured archaea in this environment, leading to the discovery of new lineages such as the Methanomassiliicoccales, and to the identification of non-methanogenic archaea such as the Bathyarchaeota and the Thaumarchaeota (Dridi, Fardeau, *et al.*, 2012; Paul *et al.*, 2012; Borrel, O'Toole, *et al.*, 2013; Poulsen *et al.*, 2013; Borrel *et al.*, 2012; Paul *et al.*, 2012; Hervé *et al.*, 2020; Loh *et al.*, 2021). These groups are discussed in more detail later.

III. Limited scope of sequenced-based archaeal studies

An overwhelming number of archaea-focused intestinal microbiome studies have been conducted in humans and ruminants (Vanderhaeghen *et al.*, 2015; Raymann *et al.*, 2017; Mizrahi *et al.*, 2021). Studies carried out on insects have revealed an increased diversity of species hosting archaea (Leadbetter *et al.*, 1998; Sprenger *et al.*, 2000; Samuel *et al.*, 2007; Paul *et al.*, 2012). However, information concerning archaea in many lineages of animals is currently very limited. This is also reflected by the source of the genomes from host-associated archaea currently available in public databases, which are derived primarily from humans, ruminants, and insects **(Introduction Figure 6)**.

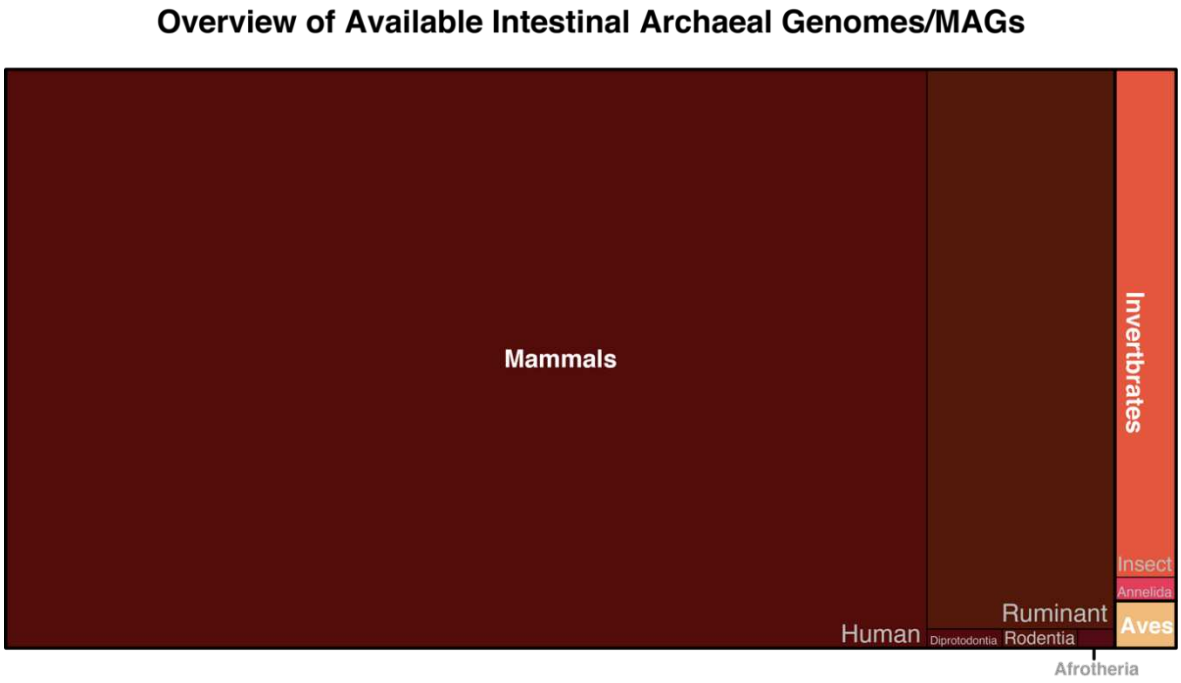


Figure 6. Origin of Genomes (isolates and metagenome-assembled-genomes) obtained from different animal hosts that were available as of summer 2020. Represented as a treemap, where the total area of the map represents the total number of genomes/MAGs and each section represents the proportion of genome/MAGs coming from each host. Built using the R package 'treemap' (v 2.4-2) (Tennekes, 2017).

III.a Human

Of course, a large amount of effort has been put into understanding human intestinal methanogens. In humans, archaea generally account for less than 1% of the total intestinal microbial population (Gaci *et al.*, 2014; Borrel *et al.*, 2020). In European and American populations, one adult out of two is positive for methanogens when using the methane breath test (Levitt *et al.*, 2006). Interestingly, this is not uniform between different populations, as in Japan, for example, there are fewer methane-positive individuals while in African populations there are more (Levitt *et al.*, 2006). Moreover, the prevalence and relative abundance of methanogens has been shown to increase with age (Polag *et al.*, 2014; Vanderhaeghen *et al.*, 2015), possibly because of slowed digestion time (Gaci *et al.*, 2014; Triantafyllou *et al.*, 2014; Jeffery *et al.*, 2016). Methanogens have also been associated to various health/disease states in humans. For example, methanogens that can use trimethylamines (TMA) are thought to be beneficial for health (Brugère *et al.*, 2013), as this compound is involved in the development of cardiovascular and kidney diseases (Koeth *et al.*, 2013; Tang *et al.*, 2014). Conversely, methanogens have been detected in higher-than-normal concentrations in inflammatory

bowel disease, periodontal disease and various abscesses (Sogodogo *et al.*, 2019; Borrel *et al.*, 2020). However, it is important to note that there is currently no experimental evidence that identifies methanogens or any other archaea as pathogenic.

III.b Ruminant

Ruminants host higher concentrations of methanogens than humans (3-5% of the total microbial community (Frey *et al.*, 2010)), and have also been subject to intensive archaea-centric investigations as the production of methane gas has important ecological and economic implications. Ecologically, ruminants release large amounts of methane which is a potent greenhouse gas (Liu *et al.*, 2008). Economically, methanogens negatively affect the amount of energy that livestock obtain from their feed (Nkrumah *et al.*, 2006). Thus, numerous studies have been focused on understanding methanogens in ruminants with the hope of generating methanogen 'vaccines' or other ways to engineer the rumen microbiome in a way that reduces methane emissions from these animals (Leahy *et al.*, 2010; Shibata *et al.*, 2010; Wedlock *et al.*, 2013; Hristov *et al.*, 2015). Methanogens are also extremely important in the rumen as their consumption of hydrogen allows other bacterial metabolisms to continue in a thermodynamically favorable fashion (Mizrahi *et al.*, 2021). However, recent studies have suggested that certain groups of hydrogenotrophic bacteria may be able to fulfill this role when methanogens are not present (Greening *et al.*, 2019).

III.c Termite

Termites represent a unique case among insects, as they have a well-developed hindgut capable of hosting methanogenic archaea and even release methane at the same order of magnitude as ruminants per body mass (Brune, 2018). This means that termites are significant contributors to global methane emissions (Sugimoto *et al.*, 1998). Indeed, termites have been shown to maintain relatively abundant communities of methanogens (1-5% of the total microbial community (Nonoh, 2013)). Generally speaking, these insects are grouped into two broad classifications – higher and lower termites. The methanogens that inhabit these two groups have found unique ways to do so, as methanogens in the lower termites are often found in close association with anaerobic flagellates (which provide them with H_2 and carbon substrates), while in higher termites they rely on fermentative bacteria for substrates for methanogenesis. Moreover, methanogenic communities are usually more diverse in higher

termites (Brune, 2018), and are found in proximity of the hindgut wall (Leadbetter *et al.*, 1998), where they are protected from washout and have access to ideal concentrations of hydrogen (Brune, 2018).

III.d Other animals

Beyond humans, ruminants, and termites, there exist a limited number of single-species studies in birds, herbivorous reptiles, primates, marsupials, and other insects, where intestinal methanogens have been detected (Evans *et al.*, 2009; Wright *et al.*, 2009; St-Pierre *et al.*, 2013). Thus, while it is clear that intestinal archaea inhabit the intestinal tract of a diverse range of animals, their overall distribution remains poorly understood and could be biased by the different approaches and methods used in these single-species studies. Furthermore, long standing hypotheses proposed over thirty years ago (Hackstein *et al.*, 1996) have not been systematically examined in most animals. Ultimately, these limited sampling efforts and varied approaches have also resulted in large gaps in understanding concerning the diversity, abundance, and adaptations of archaea to the intestinal tract.

The next paragraphs summarize the present-day understanding of intestinal archaea in terms of diversity and genomic adaptations. The final sections discuss what is known concerning the ecological niche of intestinal archaea – i.e., how do they interact with intestinal bacteria.

Diversity of Intestinal Archaea - Who's there?

The Methanobacteriales were among the first cultured archaea in the intestinal microbiome of animals (Smith *et al.*, 1958), and they were since then identified in some insects, herbivorous mammals, human, birds, and reptiles (Nottingham *et al.*, 1968; Miller *et al.*, 1982; Leadbetter *et al.*, 1998; Dridi *et al.*, 2011; St-Pierre *et al.*, 2013; Gaci *et al.*, 2014). Other less well characterized lineages of archaea have also been identified in the intestinal microbiome of various hosts, and include members of the Methanosarcinales in termites and cockroaches (Beijer, 1952; Leadbetter *et al.*, 1998; Sprenger *et al.*, 2000; Hong *et al.*, 2011; St-Pierre *et al.*, 2013), and of the Methanomicrobiales in ruminants (Paynter *et al.*, 1968; St-Pierre *et al.*, 2013). The more recent discovery of the Methanomassiliicoccales has revealed an additional lineage of relatively abundant intestinal methanogens (Mihajlovski *et al.*, 2008; Paul *et al.*,

2012; Borrel, Harris, *et al.*, 2013; Poulsen *et al.*, 2013). A handful of non-methanogenic lineages have also been detected in humans, apes, and termites (Nam *et al.*, 2008; Raymann *et al.*, 2017; Hervé *et al.*, 2020). The following section presents the known diversity of these dominant host-associated archaeal lineages.

I. Methanobacteriales

Within this order host-associated archaea belong to two genera, *Methanobrevibacter* and *Methanosphaera* within the *Methanobacteriaceae* family. A third genus from this family, *Methanobacterium*, is almost exclusively found outside of the gut (*e.g.,* wetland, sediments). Members of the *Methanobacteriaceae* are CO_2-reducing hydrogenotrophic methanogens, except for *Methanosphaera* spp. which perform methyl-reducing hydrogenotrophic methanogenesis (Miller *et al.*, 1985; Liu *et al.*, 2008; Facey *et al.*, 2012).

The genus *Methanobrevibacter* has been consistently identified in primates, humans, ruminants, cows, reptiles, and insects (Nottingham *et al.*, 1968; Miller *et al.*, 1982; Leadbetter *et al.*, 1998; Dridi, Henry, *et al.*, 2012; St-Pierre *et al.*, 2013; Gaci *et al.*, 2014). In humans, *Methanobrevibacter smithii* **(Introduction Figure 7a)** is the most commonly encountered methanogen (Miller *et al.*, 1986), and *Methanobrevibacter ruminantium* is the dominant methanogen in the rumen (Smith *et al.*, 1958). *Methanosphaera* have been detected in humans (*Methanosphaera stadtmanae* **(Introduction Figure 7b)** DSM 3091, (Miller *et al.*, 1985)), kangaroos (*Methanosphaera stadtmanae* WGK6, (Hoedt *et al.*, 2016)), cattle (St-Pierre *et al.*, 2013), and rabbits (*Methanosphaera cuniculi,* (Biavati *et al.*, 1988)). Some *Methanobacterium* sequences have also been identified in the rumen (Smith *et al.*, 1958; Kelly *et al.*, 2014).

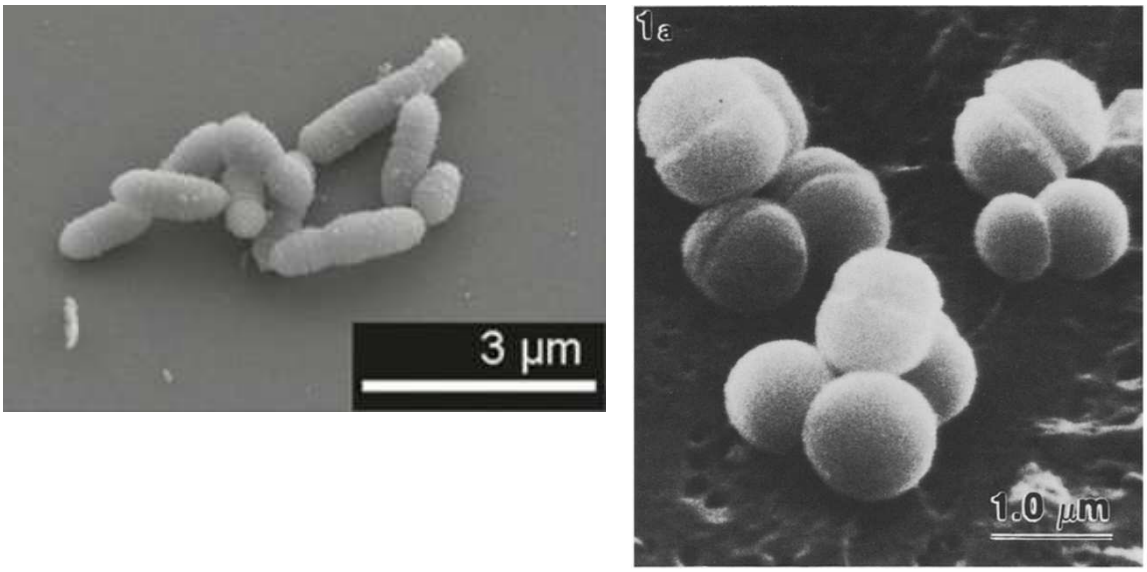


Figure 7. a) Scanning Electron micrograph of *M. smithii*, adapted from Ruaud et al 2020. **b)** Electron microscopy of *Methanosphaera stadtmanae*. adapted from Miller and Wolin 1985.

II. Methanomassiliicoccales

Another commonly encountered host-associated methanogenic order is the Methanomassiliicoccales (Mihajlovski *et al.*, 2008; Borrel, Harris, *et al.*, 2013; Söllinger *et al.*, 2016). These archaea are methyl-reducing hydrogenotrophic methanogens - meaning that they depend on methyl-compounds (e.g., methanol or trimethylamines) and H_2 to perform methanogenesis. The currently available genomes completely lack genes responsible for the oxidative steps of methylotrophic methanogenesis (Borrel, O'Toole, *et al.*, 2013; Lang *et al.*, 2015). The absence of these genes explains why, unlike most previously known methanogens using methyl-compounds, Methanomassiliicoccales must be supplied with H_2. Within the Methanomassiliicoccales there are two families, the mostly environmental *Methanomassiliicoccaceae* and the host-enriched "*Ca*. Methanomethylophilaceae" **(Introduction Figure 8)**(Paul *et al.*, 2012; Borrel, Harris, *et al.*, 2013; Söllinger *et al.*, 2016; Borrel *et al.*, 2017).

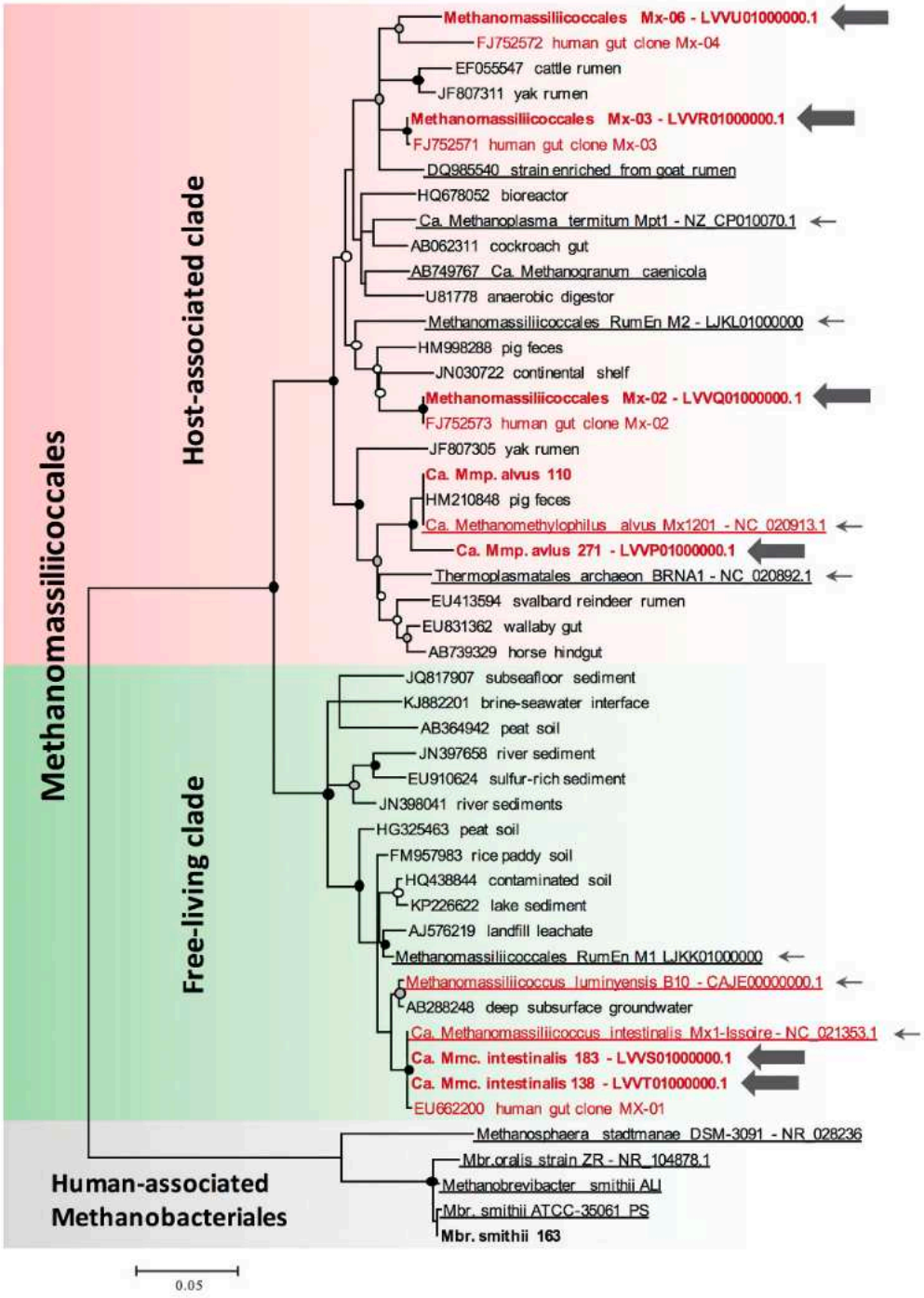


Figure 8. Phylogenetic tree of the Methanomassiliicoccales highlighting the host-associated and free-living clades of this methanogen, adapted from Borrel *et al.* 2017. Red sequences are human-associated Methanomassiliicoccales. Sequences in red are from human-associated Methanomassiliicoccales and sequences in bold are obtained by Borrel *et al.*; large arrows point to draft genomes obtained by Borrel *et al.* 2017 and small arrows point to complete/draft genomes obtained in previous studies. Underlined sequences correspond to cultured Methanomassiliicoccales.

Currently, there is a single cultured representative of this order, *Methanomassiliicoccus luminyensis* **(Introduction Figure 9)**, which was isolated from human faeces (Dridi, Fardeau, *et al.*, 2012). Interestingly, this microbe is part of the environmental clade of the Methanomassiliicoccales rather than the host-enriched clade (Borrel *et al.*, 2014; Söllinger *et al.*, 2016). Another uncultured species from the environmental clade, "*Ca.* Methanomassiliicoccus intestinalis" was also identified in the human gut (Borrel, Harris, *et al.*, 2013). However, this methanogen was almost only encountered in elderly people in long-stay residential care (Dridi, Fardeau, *et al.*, 2012; Borrel, O'Toole, *et al.*, 2013).

No isolated representatives of the host-enriched clade of the Methanomassiliicoccales are currently available. However, candidate genera have been detected in a variety of animals. For example, "*Ca.* Methanomethylophilus", is commonly detected in humans and ruminants, more specifically "*Ca.* Methanomethylophilus alvus", which is among the most commonly encountered Methanomassiliicoccales in the human intestinal tract (Borrel *et al.*, 2017). Other members of the host-enriched clade of the Methanomassiliicoccales include the "*Ca.* Methanoplasma" **(Introduction Figure 9)** which were first found in the termite hindgut (Lang *et al.*, 2015). These candidate genera have also been found in primates, wallabies, and herbivorous reptiles (Evans *et al.*, 2009; Paul *et al.*, 2012; Söllinger *et al.*, 2016; Raymann *et al.*, 2017).

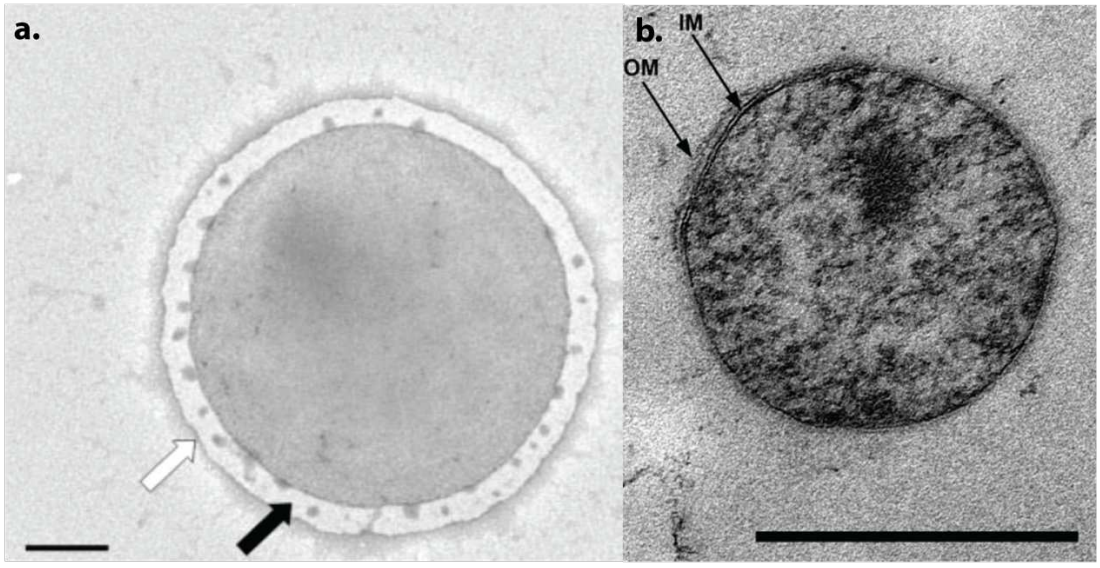


Figure 9. a) Transmission electron micrograph of *Methanomassiliicoccus luminyensis*, adapted from Dridi et al 2012. Scale bar = 0.3 µm. **b)** "*Ca*. Methanoplasma termitum", adapted from (Lang *et al.*, 2015). Scale bar = 500 nm. Arrows indicate the putative inner and outer membranes of *M. luminyensis* and *M. termitum*. Images were obtained from enrichment cultures.

III. Methanosarcinales

There are two cultured representatives of Methanosarcinales that were isolated from the rumen (*Methanosarcina barkeri* CM1 (Jarvis *et al.*, 2000)) and the hindgut of a cockroach (*Methanimicrococcus blatticola*; **(Introduction Figure 10)**(Sprenger *et al.*, 2000). Interestingly, the *Methanosarcina* are less frequently detected in the intestinal tract, and host-associated *Methanosarcina* are genomically very similar to the environmental species, *Methanosarcina barkeri fusaro* (Lambie *et al.*, 2015). This suggests that members of this genus of methanogens have only recently adapted to the intestine or may be transient.

Conversely, *the Methanimicrococcus* has been detected in many different animal hosts such as termites, horses, and ruminants (Sprenger *et al.*, 2000, 2005, 2007; Fernandes *et al.*, 2014; Tapio *et al.*, 2017). Like the Methanomassiliicoccales, the *Methanimicrococcus* are methyl-reducing hydrogenotrophic methanogens.

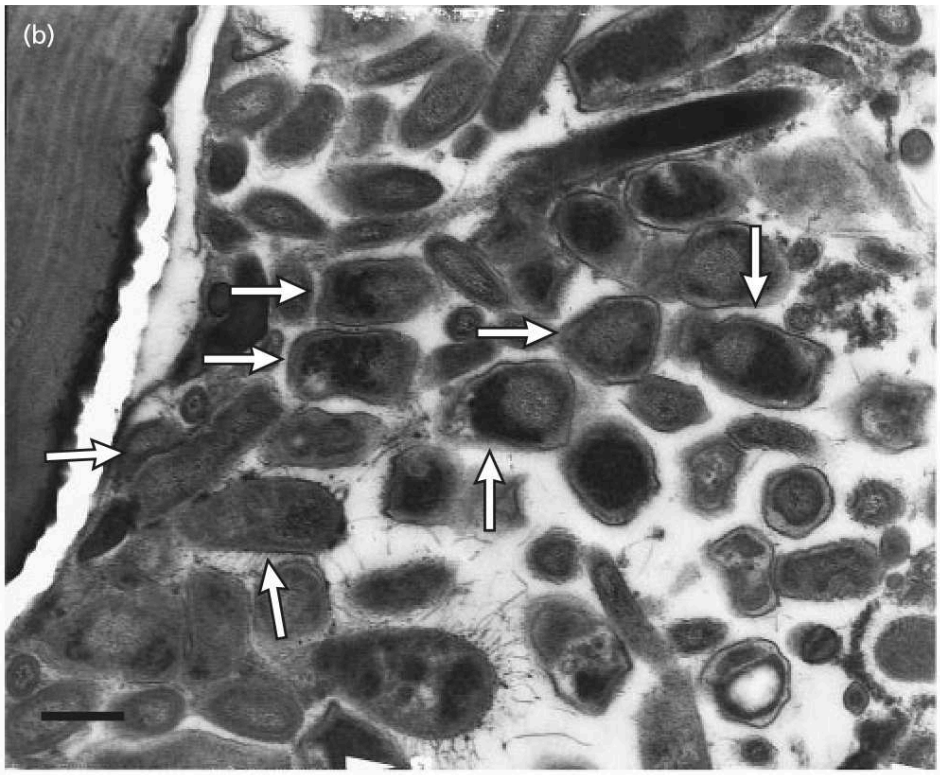


Figure 10. Transmission electron micrograph of *Methanimicrococcus blatticola* strain PA, indicated with white arrows, in the hindgut of *Periplaneta americana*, scale bar = 1 µm. Adapted from Sprenger et al 2000.

IV. Methanomicrobiales

Currently, the only representative of the Methanomicrobiales isolated from an intestinal environment is *Methanomicrobium mobile,* which originates from the rumen of a cow (Paynter *et al.*, 1968). Recent studies have detected sequences from the genus *Methanocorpusculum* **(Introduction Figure 11)** in a variety of animals such as amphibians (giant salamander; (Wu *et al.*, 2019)), reptiles (giant tortoises, (Hong *et al.*, 2011)), mammals (horses, (St-Pierre *et al.*, 2013); wombats (Shiffman *et al.*, 2017); rhinos (Gibson *et al.*, 2019); sika deer (Li *et al.*, 2019)), and even in anaerobic ciliates (Lewis *et al.*, 2018). The widespread distribution of this poorly understood genus suggests that it could be a common although overlooked intestinal methanogen.

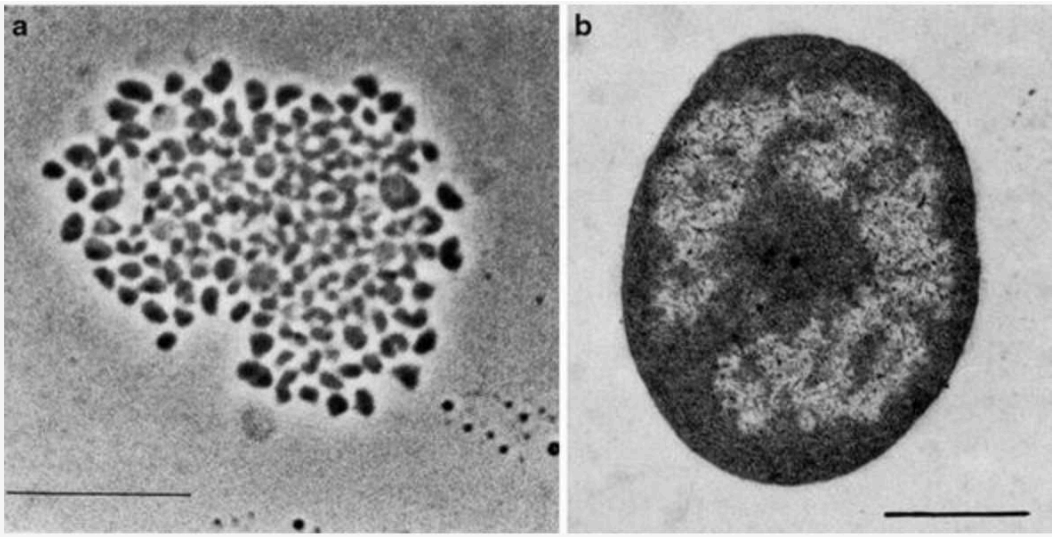


Figure 11. Images of *Methanocorpusculum aggregans*, isolated from an anaerobic digestor. **a)** Phase-contrast image of an aggregate of *Methanocorpusculum aggregans* cells. Scale bar = 10 µm. **b)** Electron micrograph of Methanocorpusculum *aggregans*. Scale bar = 0.5 µm. It is important to note that there is currently no isolate of *Methanocorpusculum* from the intestinal tract. These images are from species isolated in the environment. Adapted from (Oren, 2014).

V. Non-methanogenic Archaea

Non-methanogenic archaea have been sporadically identified in intestinal studies. Many of these lineages are poorly studied due to a limited number of genomes and cultured representatives.

V.a Halobacteria

Halobacteria were first detected in fecal samples from Korean people in 2008 (Nam *et al.*, 2008). Since then, two additional species have been isolated from human faeces: *Haloferax massiliensis and Haloferax alexandrinus* (Khelaifia *et al.*, 2016, 2017, 2018). It is still unclear if the *Halobacteria* represent a resident constituent of the intestinal microbiome or if they are simply transient microbes passing through the gut (for example through the ingestion of table salt), as *H. massiliensis* is unable to grow in anaerobic or even microaerophilic conditions (Khelaifia *et al.*, 2018). While O_2 concentrations vary along the intestinal tract, the accepted "healthy" concentration of oxygen in the human colon is believed to be < 10 mm Hg, well below the growth requirements of *H. massiliensis* (Zheng *et al.*, 2015). Moreover, *Halobacteria* are usually detected in individuals with diets extremely high in salt, which could be a mean of

entry into the human gastrointestinal tract (Nam *et al.*, 2008; Khelaifia *et al.*, 2018). A recent study on a large cohort of individuals from South Korea has detected Haloarchaea in ~43% of the samples via 16S rRNA gene sequencing, real-time PCR, and fluorescence in situ hybridization (Kim *et al.*, 2020). Taken together this provides strong support for the residency of *Halobacteria* in the intestinal microbiota of some humans.

V.b Thaumarcheota

Three studies in humans, apes, and termite, have detected members of the phylum Thaumarchaeota (order Nitrososphaerales; (Hoffmann *et al.*, 2013; Nonoh, 2013; Raymann *et al.*, 2017)). However, beyond this sporadic detection, the potential role of the Thaumarchaeota in the intestinal tract of these animals has not been examined. Outside of the gut, member of the *Nitrosopumilales* order (phylum Thaumarchaeota) are most commonly found in association with invertebrates such as tunicates and sponges (Preston *et al.*, 1996; López-Legentil *et al.*, 2010; Simister *et al.*, 2011; Zhang *et al.*, 2019), and have been shown to aid in nitrite/nitrate chemical cycling within these hosts (Moitinho-Silva *et al.*, 2017).

V.c Bathyarchaeota

Bathyarchaeota have also been identified in termites (Friedrich *et al.*, 2001; Shi *et al.*, 2015; Hervé *et al.*, 2020; Loh *et al.*, 2021). Analysis of Bathyarchaeota MAGs from the termite hindgut showed that these archaea may be capable of performing reductive homoacetogenesis – suggesting that they would compete with methanogens in the intestine for H_2 (Loh *et al.*, 2021). Although the exact pathway of this metabolism in these archaea has yet to be elucidated, it is interesting to note that reductive homoacetogenesis has been shown in some cases to be more prevalent in termite hindguts than methanogenesis (Tholen *et al.*, 2007; Brune, 2018; Loh *et al.*, 2021).

Host-associated genomic adaptations

Because the intestinal tract of animals presents a unique system that offers both a rich and relatively consistent influx of organic nutrients as well as a diverse set of stressors (e.g., removal, high rates of competition (Ley *et al.*, 2006; Quigley, 2011; Douglas, 2015)) the genomic content of archaea that inhabit these niches is likely distinct from that of their

environmental relatives. The low oxygen concentrations and lack of sunlight implies that the intestinal tract has fewer niches for microbes to colonize than external environments, despite being nutrient rich (reviewed in (Ley *et al.*, 2006)). This increases the competition between microbes and limits the number of those that can colonize the intestinal tract. Furthermore, the consistent peristaltic movement of the intestinal tract exerts an additional selective pressure on intestinal microbes (Quigley, 2011; Douglas, 2015).

Adding to this, different hosts have different digestive environments that may have also led to distinct functional traits of their microbiota. As discussed previously, within the intestinal bacteria and archaea, there are distinct clades of microbes that are almost only found in the gastrointestinal tract of animals. This is a clear indication that host-associated adaptations occurred at the divergence of each of these clades. However, the lack of diverse genomes and standardized comparative genomic analyses across archaeal lineages has made it difficult to determine exactly what genomic adaptations these archaea have developed as a result of intestine colonization. The following section will summarize what has been characterized concerning host-associated adaptations in intestinal archaea.

I. Genome reduction

In Bacteria, genome reduction has been closely linked to host-association. This is because, as functions become redundant in a nutrient rich environment, they become less energetically favorable to maintain and are thus lost (Moya *et al.*, 2008; Moran *et al.*, 2009; Fisher *et al.*, 2017). A similar phenomenon has been observed in some intestinal archaea (Borrel *et al.*, 2014; Lang *et al.*, 2015; Söllinger *et al.*, 2016; de la Cuesta-Zuluaga *et al.*, 2021). The Methanomassiliicoccales that fall within the 'host-enriched' clade (e.g. "*Ca*. Methanomethylophilus alvus*"* and "*Ca*. Methanoplasma termitum") have among the smallest genomes among the Methanomassiliicoccales (Borrel *et al.*, 2014; Söllinger *et al.*, 2016; de la Cuesta-Zuluaga *et al.*, 2021). Likewise, an endosymbiotic species of *Methanocorpusculum* living in anaerobic ciliates has undergone significant reduction in genome size (Lind *et al.*, 2018).

Examples of functions that have become redundant in intestinal archaea and are either absent or present in a reduced capacity include: cobalamin (vitamin B_{12}) synthesis - absent in endosymbiotic *Methanocorpusculum* (Lind *et al.*, 2018), and chemotaxis - reduced/absent in *Methanobrevibacter*, *Methanosphaera* (Samuel *et al.*, 2007), Methanomassiliicoccales (Borrel *et al.*, 2014), and endosymbiotic *Methanocorpusculum* (Lind *et al.*, 2018).

The Methanomassiliicoccales, "*Ca.* Methanoplasma termitum", may have even reduced the number of substrates it can use for methanogenesis, as it has been suggested that it may only be able to use monomethylamine and methanol for methanogenesis, as opposed to other members of this order that have the genetic potential to use mono-, di-, and tri-methylamines for methanogenesis (Lang *et al.*, 2015).

Paradoxically, a diverse range of host-associated methanogens do not have ability to synthesize coenzyme M which is used in the last step of methanogenesis. Indeed, *Methanimicrococcus blatticola* (Sprenger *et al.*, 2000), *Methanobrevibacter ruminantium* (Leahy *et al.*, 2010), endosymbiont *Methanobrevibacter* sp. NOE/*Methanocorpusculum* sp. MC (Lind *et al.*, 2018); Methanomassiliicoccales ISO4-H5/ISO4-G1 (Kelly, Pacheco, *et al.*, 2016; Li *et al.*, 2016), and "*Ca.* Methanoplasma termitum" (Lang *et al.*, 2015) must obtain this coenzyme externally in order to perform methanogenesis. However, the mode in which this acquisition occurs has yet to be identified.

Further investigation into which genes are absent from host-associated archaea is needed to reveal additional widespread losses of functions that are not energetically favorable to maintain in the intestinal environment.

II. Cell Surface Adaptations

The cell envelope and associated proteins are in direct interaction with the surrounding environment and perform a number of important functions including protection, adhesion, and nutrient acquisition. Therefore, it is not surprising that there have been several host-associated genomic adaptations linked to genes responsible for modifying the cell surface. Methanobacteriales members such as *M. smithii*, *M. stadtmanae*, *M. ruminantium*, *M.*

millerae and *M. olleyae* have genomes enriched in genes responsible for altering their cell surface (Samuel *et al.*, 2007; Lurie-Weinberger *et al.*, 2012; Kelly, Li, *et al.*, 2016; Kelly, Pacheco, *et al.*, 2016). More specifically, *M. smithii* and *M. stadtmanae* have an increased number of specific glycosyltransferases (GTs) involved in cell surface modification. Interestingly, a large portion of these GTs appear to have been acquired via horizontal gene transfer acquired from Firmicutes and Bacteroides, the two dominant bacterial lineages in the human gut (Lurie-Weinberger *et al.*, 2012; Lurie-Weinberger *et al.*, 2012). The level of expression of these GTs is likely regulated in *M. smithii*, since its cell surface is decorated with various glycans when retrieved from gnotobiotic mice ceca, while it is relatively bare when grown in a glass tube **(Introduction Figure 12)**(Samuel *et al.*, 2007). Moreover, *M. smithii* has a complete set of genes for the synthesis of a pseudaminic acid which was also detected on its cell surface (Lewis *et al.*, 2009). Two of the three genes involved in the synthesis of pseudaminic acid were also identified in another intestinal archaeon*, M. stadtmanae* (Kandiba *et al.*, 2013) – although the ability to produce this compound has not yet been experimentally investigated. As this sugar is found in the human intestinal epithelial cell surface, it was suggested that it allows *M. smithii* to mimic host cells and therefore avoid detection by the immune system (Samuel *et al.*, 2007). Interestingly, the genes necessary to make pseudaminic acid were detected in a diverse range of free-living archaea, but are absent in most intestinal methanogens (Kandiba *et al.*, 2013).

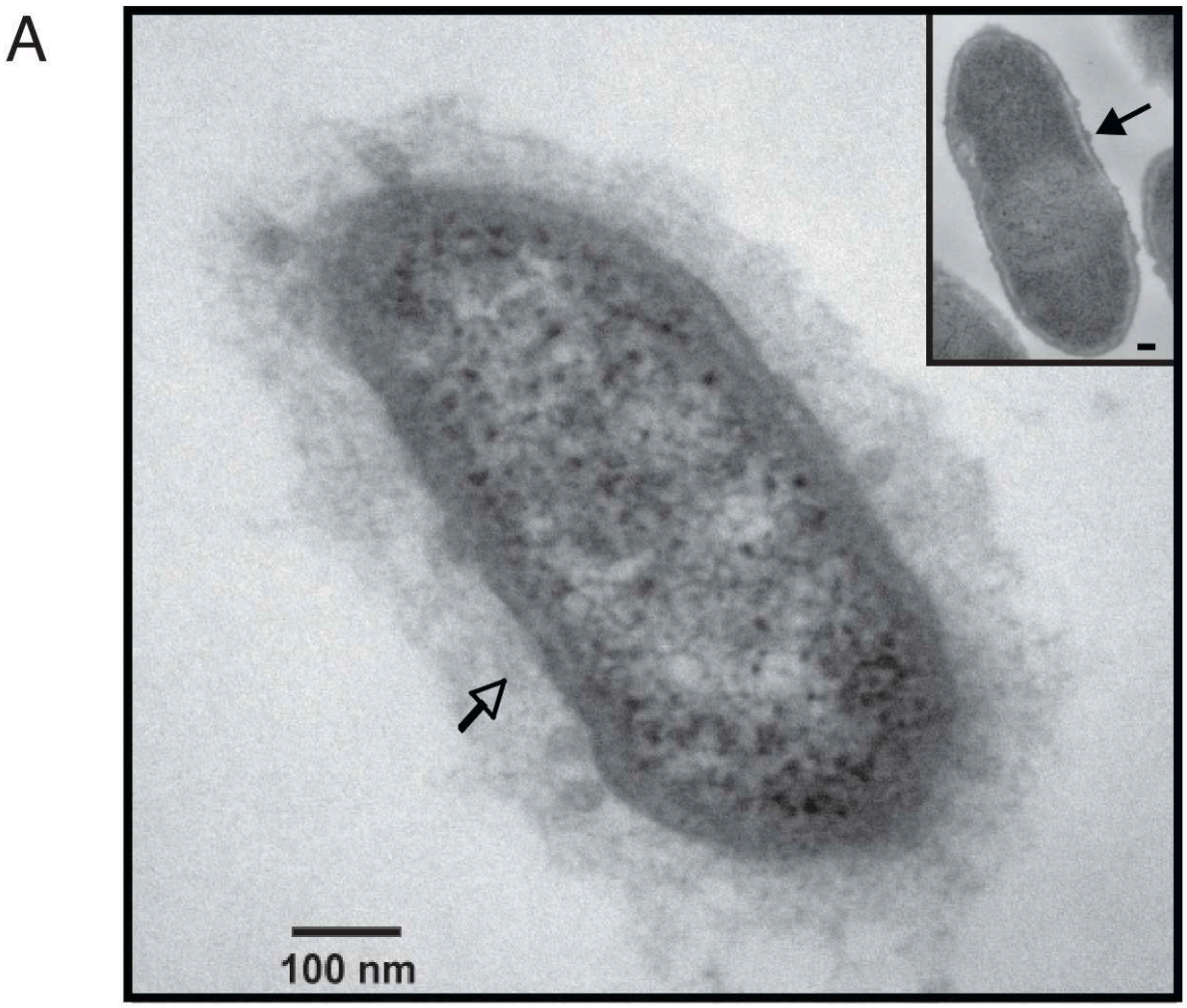


Figure 12. Transmission electron microscopy (TEM) of *M. smithii* collected from the ceca of mice. The inset shows *M. smithii* collected from a batch fermenter. Arrows are pointing to the differences in the cell capsule between the two different culture conditions. Adapted from Samuel et al 2007.

Another subset of cell-surface proteins that have been identified as enriched in intestinal archaea are adhesin-like proteins (ALP) (Samuel *et al.*, 2007; Leahy *et al.*, 2010; Lurie-Weinberger, Peeri and Gophna, 2012; Kelly, Li, *et al.*, 2016; Kelly, Pacheco, *et al.*, 2016). Adhesins are surface proteins that allow microbes to adhere to other cell surfaces – either of other microorganisms or host tissues (Klemm *et al.*, 2000). Below I'll describe more in detail such cases concerning the Methanobacteriales and the Methanomassiliicoccales.

II.a ALPs in the Methanobacteriales

In the Methanobacteriales, there is evidence that ALPs are actively transcribed and facilitate bacteria-archaea and archaea-host interactions. *M. smithii* strains isolated from human faeces have a unique repertoire of adhesin-like proteins that change expression when formate concentrations vary (Hansen *et al.*, 2011). These modifications in ALP expression level are considered as a way for *M. smithii* to interact with different bacterial partners generating different concentrations of formate. This may allow *M. smithii* to decrease intra-strain competition because strains can create separate bacterial-archaeal interactions depending on the amount of formate that bacteria generate (Hansen *et al.*, 2011). ALPs may also allow *M. smithii* to form biofilms in some situations (Bang *et al.*, 2015).

Co-culture experiments have revealed that methanogens sometimes form aggregates with their bacterial partners – possibly via ALPs. Indeed, *M. smithii* and *Christensenella* aggregate when grown together (Ruaud *et al.*, 2020). The rumen methanogen *M. ruminantium* M1 also aggregates with an H_2 producing intestinal bacterium, *Butyrivibrio proteoclasticus* (phylum Firmicutes) (Leahy *et al.*, 2010). Transcriptomic analyses have found that *M. ruminantium/B. proteoclasticus* co-cultures result in the upregulation of a suite of ALP genes in *M. ruminantium* (Ng *et al.*, 2016). The same study also showed that ALPs likely facilitate both bacterial-archaeal and eukaryotic-archaeal interactions, as they also bind to species of protozoa found in the rumen (Ng *et al.*, 2016).

Other Methanobacteriales species also display behaviors suggesting they have acquired unique cell surface proteins involved in attachment. *Methanobrevibacter cuticularis*, *Methanobrevibacter curvatus*, and *Methanobrevibacter filiformis* are commonly found associated to the hindgut epithelial tissue of their insect hosts (Hackstein *et al.*, 1994; Leadbetter *et al.*, 1996, 1998). It is currently unclear if this association is mediated by ALPs or other cell surface appendages. Further investigation into the genomic content of these three species showed that they are capable of growing and producing methane when small amounts of O_2 are present and when they have active catalase enzymes (Leadbetter *et al.*, 1996, 1998; Tholen *et al.*, 2007), which explains how they manage to inhabit this unique micro-oxic niche in the hindgut. Similar oxygen-coping adaptations have been identified in insect-associated

Methanimicrococcus blatticola (Sprenger *et al.*, 2007) and in other *Methanobrevibacter* spp. (Poehlein *et al.*, 2018).

A large number of ALP genes identified in *M. smithii, M. stadtmanae*, and *M. ruminantium* were acquired from intestinal Firmicutes (Leahy *et al.*, 2010; Lurie-Weinberger, Peeri and Gophna, 2012; Lurie-Weinberger, Peeri, Tuller, *et al.*, 2012). The widespread transfers of ALP genes from intestinal bacteria to the Methanobacteriales is particularly interesting as it highlights important cross-domain adaptation events that have occurred in the intestinal tract. However, because of the limited number of host-associated archaeal genomes, it is not currently clear if similar large transfer events of cell surface genes have also occurred in other lineages of intestinal archaea.

II.b Unique adhesin proteins in the Methanomassiliicoccales

The Methanomassiliicoccales have a distinct composition of adhesin-like proteins, suggesting that these methanogens have developed a different approach of adaptation than the *Methanobacteriales*. Indeed, the genomes of host-associated Methanomassiliicoccales were the first shown to include genes encoding adhesins containing proteins with Flg_new domains (Borrel *et al.*, 2017; de la Cuesta-Zuluaga *et al.*, 2021). The enrichment of Flg_new proteins in host-associated Methanomassiliicoccales was recently confirmed in a large-scale comparative genomic analysis of this order (de la Cuesta-Zuluaga *et al.*, 2021).

The absence of Flg_new proteins in other currently analyzed archaeal genomes is interesting and strongly supports the hypothesis that these proteins have played an important role in the adaptation of Methanomassiliicoccales to the intestinal tract.

III. Other adaptations

Intestinal archaea have colonized a diverse range of animal intestinal tracts and therefore have likely acquired other specialized adaptations that allowed them to successfully colonize various types of hosts. The following section highlights what is known about these unique adaptations.

The Methanomassiliicoccales have a distinct repertoire of proteins involved in facilitating inter-cellular interactions, such as tetratricopeptide repeat (TPR) domain containing proteins (Borrel *et al.*, 2014, 2017). More specifically, intestinal Methanomassiliicoccales have genomes enriched in a specific subfamily of TPR called Sel1_repeat proteins (Borrel *et al.*, 2017; de la Cuesta-Zuluaga *et al.*, 2021). Sel1_repeat domains are poorly characterized, but they have been shown to be present in proteins involved in signal transduction during interactions between eukaryote and bacteria (Mittl *et al.*, 2007).

Methanobrevibacter wolinii SH, isolated from cow rumen, has both the genes coding for alcohol and aldehyde dehydrogenases (*walCD*) as well as those for methanol: coenzyme M methyltransferase (*mtaABC*) (Poehlein *et al.*, 2018), indicating it may have the ability to use ethanol/methanol for methanogenesis. Likewise, *Methanobrevibacter sp.* AbM4 (isolated from sheep rumen) also has the *walCD* genes (Leahy *et al.*, 2013), as does *Methanobrevibacter ruminantium* M1 (Leahy *et al.*, 2010). Even though the ability to produce methane from ethanol alone has not been observed in any host associated *Methanobrevibacter* in a laboratory setting, the presence of these genes supports the hypothesis that some methanogens have altered their methanogenesis metabolism to successfully compete within the densely populated intestinal microbial community. For example, *Methanosphaera* sp. WGK6 -isolated from a kangaroo- can use ethanol to reduce methanol into methane - without the addition of H_2 (Hoedt *et al.*, 2016). This would make it more competitive in various intestinal environments where H_2 is in high demand. Another strategy to cope with H_2-competition has been identified in *M. blatticola*. This methanogen produces enzymes with affinities that are much higher for H_2 than its relatives (Sprenger *et al.*, 2005; Feldewert *et al.*, 2021), meaning it has access to H_2 at lower concentrations than many other microbes in the intestine.

Genes encoding for functions that may protect the methanogens from stressors in the intestine environment have also been observed. *M. millerae* SM9 is predicted to produce extracellular tannases enzymes capable of degrading tannin, a methanogenesis inhibiting compound which is found in many plants (Kelly, Pacheco, *et al.*, 2016). Also, the genomes of Methanomassiliicoccales in the host-enriched clade, as well as host-associated *Methanobrevibacter* and *Methanobacterium* contain bile salt resistance genes that may allow

them to survive the low pH bile acids found in the intestinal tract of animals (Jones *et al.*, 2008; Borrel *et al.*, 2014, 2020; Gaci *et al.*, 2014; Ouboter *et al.*, 2020; de la Cuesta-Zuluaga *et al.*, 2021).

IV. Recently adapted/generalists Archaea in the intestinal microbiota

Even though *Methanomassiliicoccus luminyensis* and "*Ca.* Methanomassiliicoccus intestinalis" were all initially detected in the intestinal tract, they both branch in the 16S rRNA gene tree within the clade of 'environmental' Methanomassiliicoccales (Borrel *et al.*, 2014; Söllinger *et al.*, 2016). Further analysis of the distribution and genomic content of these archaea resulted in the suggestion that "*Ca.* M. intestinalis" has recently adapted to the intestinal environment and that *M. luminyensis* is a generalist microbe, not specifically adapted to the intestinal tract (Borrel *et al.*, 2017; Cozannet *et al.*, 2021). Especially because *M. luminyensis* is rarely detected in intestinal microbiota studies (Borrel *et al.*, 2017), and it is the most commonly encountered Methanomassiliicoccales in environmental samples (Cozannet *et al.*, 2021). Conversely, "*Ca.* M. intestinalis" is not found in environmental samples and it can be highly prevalent in elderly individuals in long term care housing facilities (Borrel *et al.*, 2017). Interestingly, "*Ca.* M. intestinalis" has up to 38 genes encoding for the Flg_new domain ALPs discussed above, but these genes are absent in *M. luminyensis* (Borrel *et al.*, 2017). A variety of other genes responsible for environmental adaptations have been identified in *M. luminyensis* (including for example methionine synthesis (Lang *et al.*, 2015), osmolyte transport or N_2 fixation (Borrel *et al.*, 2014; Söllinger *et al.*, 2016), and oxidative-stress resistance (Borrel *et al.*, 2014; Söllinger *et al.*, 2016). Taken together, these results provide evidence that this methanogen has not specialized significantly to life in the gut environment.

Similarly, within the Methanobacteriales, *Methanobacterium formicium* also has several genomic indicators that it may not be specifically adapted to the intestinal tract. These include for example an increased number of genes for dealing with oxidative stress and osmolyte production (Gutiérrez, 2012). *Methanosarcina barkeri* CM1 (order Methanosarcinales) may be a generalist like *Methanomassiliicoccus luminyensis*. *Methanosarcina barkeri* CM1 has the largest genome among methanogens of the intestinal tract (4,837,408 bp (Maeder *et al.*, 2006)) and it is strikingly similar to its environmental relative *M. barkeri fusaro* (Lambie *et al.*,

2015). This has cast doubt on the status of *Methanosarcina barkeri* CM1 as truly adapted intestinal methanogen (Lambie *et al.*, 2015).

The collection of genomic adaptations discussed above provides strong evidence for the hypothesis of different levels of adaptation to the intestinal microbiome among Archaea. Indeed it appears that some archaea are better suited to colonize the intestinal tract (e.g., "*Ca*. Methanomethylophilaceae" (Borrel *et al.*, 2014)), while some may be 'generalists' that are able to colonize a diverse range of environments (e.g. *Methanosarcina* and *Methanobacterium*), and others still may have only recently adapted to the intestinal tract (e.g., *"Ca.* M. intestinalis"). However, it is important to keep in mind that most of these findings come from comparisons between a very limited number of genomes and from the two most well studied orders of host associated archaea, the Methanobacteriales and Methanomassiliicoccales, and no thorough comparative genomic analyses have been carried out in the Methanosarcinales and Methanomicrobiales. Thus, it is currently not clear if the adaptations to the gut that have been identified so far are universally present in other intestinal archaeal lineages.

Gut bacteria and methanogens - How do they interact?

Few studies have investigated the role of host-associated archaea in the context of the larger intestinal microbial community. It is well established that archaea are important H_2 consumers in some intestinal environments, and that their consumption of H_2 allows for the complete degradation of organic matter in anaerobic conditions (Borrel *et al.*, 2020). However, there is limited information concerning specific microbial interactions in the intestinal environment. Most of what is known concerning archaeal-bacterial interactions comes from studies of *M. smithii* (Nkamga *et al.*, 2016; Traore *et al.*, 2019; Ruaud *et al.*, 2020). More recently, co-occurrence analyses have also identified putative relationships between the Methanomassiliicoccales and intestinal bacteria (de la Cuesta-Zuluaga *et al.*, 2021). The following section discusses the role of archaea in anaerobic organic matter degradation, as well as the groups of bacteria that have been identified to interact or compete with methanogens in the intestinal tract.

I. Role of Methanogens in organic matter degradation under anoxic conditions

Hydrogenotrophic microbes like the methanogens are foundational members of the microbial community as they reduce H_2 concentration to levels that are low enough to allow fermentative bacteria to use metabolites in the degradation of organic matter in anoxic conditions (Worm *et al.*, 2010; Borrel *et al.*, 2020; Conrad, 2020; Ungerfeld, 2020). The degradation of organic matter in intestinal microbial communities can be separated into three trophic-like levels: i) polymer degradation when organic matter that the host can digest is broken down by intestinal microbes; ii) monomer metabolism, primary fermentative bacteria use these sugars as a source of energy, generating H_2, CO_2, acetate, SCFA, and alcohols; iii) secondary fermentative bacteria use SCFA and alcohols in combination with hydrogenotrophic microbes like the methanogens **(Introduction Figure 13)**(Ungerfeld, 2020; Mizrahi *et al.*, 2021). Because the consumption of H_2 allows the rest of the microbial community to perform their metabolisms in a thermodynamically efficient manner, it is likely that hydrogenotrophs like the methanogens in the intestinal microbiota are extremely important to both the microbial community and the hosts they live in.

However, it is worth noting that most of what is known concerning microbial degradation of organic matter in the animal digestive tract has been mostly gathered from studies in ruminants and other large herbivorous mammals (St-Pierre *et al.*, 2013; Ungerfeld, 2020; Mizrahi *et al.*, 2021; Yang *et al.*, 2021). This means that multi-domain interactions and microbial trophic food chains in reptiles, birds, and carnivorous animal gut remain poorly understood, but it is possible that methanogens play important roles in these communities.

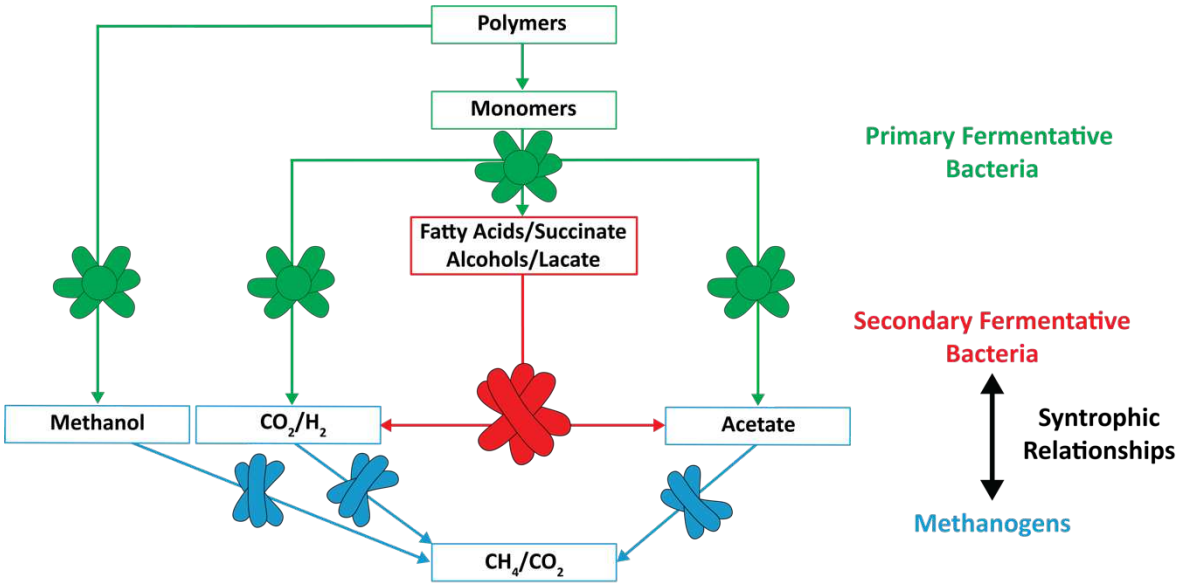


Figure 13. A simplified scheme of organic matter degradation in anoxic conditions, like the gut environment. Methanogens, represented in green, are major hydrogen-consuming microbes that drive the last step of this microbial food chain. Adapted from (Conrad, 2020).

II. Positive hydrogen-based interactions

When fibre is degraded in the intestinal microbiota, large amounts of hydrogen are released, thus one important group of bacteria to study in the context of intestinal methanogens is the fibre degraders. The main types of fibre degraders in the intestine are the cellulolytic, hemicellulolytic, and amylolytic bacteria (e.g. *Pseudobutyrivibrio* spp.; *Ruminococcus* spp.; and *Enterococcus* spp.*)* (Robert *et al.*, 2003; Kopeýný *et al.*, 2004; Flint *et al.*, 2012). Even though all these bacteria can break down fibrous material, they do not all establish relationships with methanogens. Indeed, in humans the type of fibre-degrading bacteria inhabiting the intestine influences methane production. People whose primary cellulolytic bacteria belong to the phylum Bacteroidetes do not excrete methane, while people who host cellulolytic bacteria from the Firmicutes do (Chassard *et al.*, 2010). These phyla, Firmicutes and Bacteroidetes, are the most abundant phyla in the intestinal tract of mammals (Qin *et al.*, 2010; Clavel *et al.*, 2016; Lagkouvardos *et al.*, 2019). The difference between these lineages lies in the metabolites they produce during fibre fermentation: Firmicutes produce acetate, succinate, and large amounts of hydrogen - which would be beneficial for nearby methanogens - when degrading cellulose; Bacteroidetes produce much smaller amounts of H_2 on a similar substrate.

A preference for the Firmicutes by methanogens has been observed during co-culture experiments. In co-culture, *Christensenella minuta* (Firmicutes) supported the growth of *M. smithii* and these species form flocs **(on Figure 14)**. These flocks were not observed in co-cultures of *M. smithii* with *Bacteroides thetaiotaomicron* (Bacteroidetes), and growth of *M. smithii* was weaker than with *C. minuta*. These results are in line with the co-occurrence pattern of *M. smithii* with *C. minuta* and the absence of co-occurrence between *M. smithii* and *B. thetaiotaomicron* (Nkamga *et al.*, 2016; Traore *et al.*, 2019; Ruaud *et al.*, 2020). Moreover, this evidence supports the idea of a hydrogen-based syntrophic relationship between the most abundant methanogen in the human gut and members of the Firmicutes **(on Figure 14)** (Ruaud *et al.*, 2020). These specific methanogens-Firmicutes interactions seem to extend beyond humans, as in a number of Perissodactyla and Certiartiodactyla methanogenic archaea are significantly correlated to the presence of cellulolytic bacteria from this phylum such as *Ruminococcus* spp., *Pseudobutyrivibrio* spp., *Enterococcus* spp., and *Clostridium* spp. (Morvan *et al.*, 1996; Mosoni *et al.*, 2007; Zhang *et al.*, 2017; Li *et al.*, 2019; Moraïs *et al.*, 2019).

Other positive H_2-based relationships between archaea and intestinal bacteria have been suggested from functional annotations of uncultured bacteria, such as the recently discovered lineage of hemicellulose degrading bacteria, the Melainabacteria. Members of this phylum are predicted to produce ethanol and formate - and are also thought to be important H_2 producing microbes due to the presence of genes encoding for trimeric confurcating Fe-Fe hydrogenases in their genomes (Di Rienzi *et al.*, 2013). Indeed, bacteria with these unique hydrogenases have been shown to produce higher concentrations of H_2 than other intestinal microbes (Sieber *et al.*, 2012), leading to the suggestion that they might be syntrophic partners with methanogens in the intestine (Di Rienzi *et al.*, 2013).

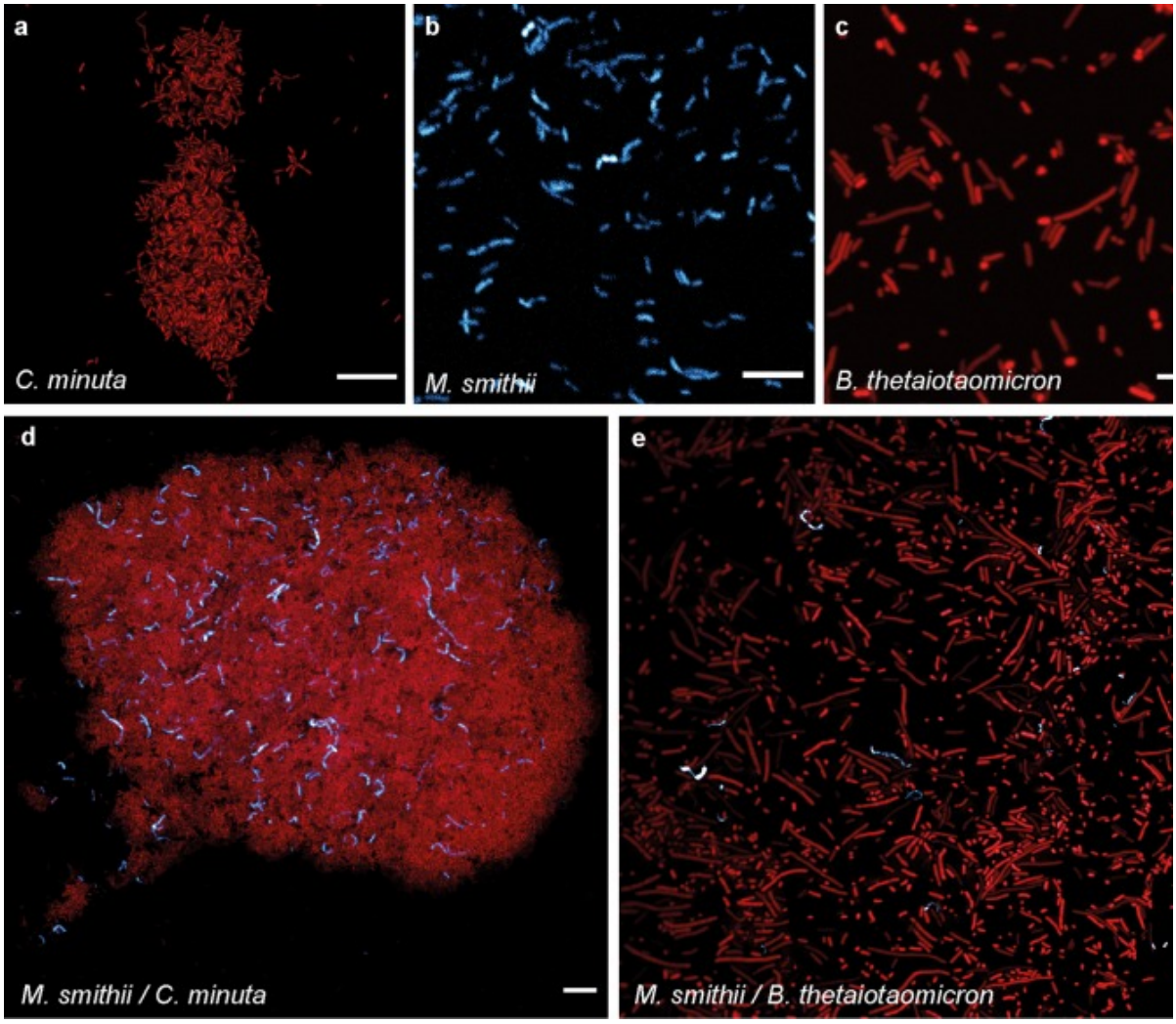


Figure 14. Confocal micrographs of *M.smithii, C.minuta,* and *B.thetaiotaomicron* alone and in co-culture. Scale bars = 10 µm. Adapted from Ruaud et al 2020.

III. Methyl-based relationships

Hydrogen may not be the only element involved in archaea-bacteria relationships in the intestinal tract, as different types of methanogenesis have different requirements for hydrogen (Feldewert *et al.*, 2021). For example, an increased availability of methyl-compounds could be particularly beneficial for methyl-reducing hydrogenotrophic methanogens like the *Methanomassiliicoccales* and *Methanosphaera* (order Methanobacteriales), which need less hydrogen than their CO_2-reducing hydrogenotrophic counterparts (Feldewert *et al.*, 2021).

Currently, there is limited research into how the methyl-reducing methanogens relate to the bacterial community in the intestine, but some studies have found that indeed, these methanogens positively co-occur with methyl-compound producing bacteria. In ruminants, the *Methanomassiliicoccales* positively co-occur with several methanol producing bacteria like *Succinivibrio* spp. (phylum Proteobacteria; (Henderson *et al.*, 2015)). Likewise, human-associated *Methanomassiliicoccales* positively correlate with the number of TMA-producing pathways encoded by gut bacteria, and negatively to TMA levels in faeces (Borrel *et al.*, 2017). Further, "*Ca*. Methanomethylophilus" and *Methanomassiliicoccus* co-occur with TMA-producing bacteria in the intestine such as *Providencia* and *Bacteroides* (de la Cuesta-Zuluaga *et al.*, 2021).

Taken together, the currently available data suggest that in the intestine, hydrogenotrophic methanogens have established positive relationships with bacteria producing either H_2 or methyl-compounds. However, positive associations between other methanogens and bacteria like those observed between *M. smithii* and *C. minuta* likely exist in the intestinal microbiota but have yet to be characterized.

IV. Competition for hydrogen between gut Archaea and Bacteria

Methanogens are not the only microbes using H_2 in the intestinal tract. Two other major functional groups of bacteria rely on H_2 for their metabolism that are traditionally recognized in this environment – sulfate-reducing bacteria and homoacetogenic bacteria (Ozuolmez *et al.*, 2015; Smith *et al.*, 2019). Energetically speaking, sulfate-reducing bacteria are the most competitive in this system, followed by methanogens (Smith *et al.*, 2019). Because methanogenesis and sulfate reduction pathways are active at lower H_2 level than homoacetogenesis (Smith *et al.*, 2019), they are thought to be the two dominant hydrogenotrophic metabolisms in the intestinal tract – and this has been observed in humans (Gibson *et al.*, 1993; Nakamura *et al.*, 2009). Several studies have sought to characterize the competitiveness of methanogens in the intestinal tract and as how they coexist with sulfate-reducing bacteria.

Initial data suggested that methanogens and sulfate-reducing bacteria are mutually exclusive

(Christl *et al.*, 1992). However, more recent studies have found that sulfate-reducing bacteria and methanogens actually co-occur in the human gut (Dore *et al.*, 2006; Smith *et al.*, 2019). The most common sulfate-reducing bacteria in the intestine of humans belong to the genus *Desulfovibrio* (order Desulfovibrionales; phylum Proteobacteria); (Gibson *et al.*, 1993; Dordević *et al.*, 2021). In primates, sulfate-reducing bacteria and methanogens have been detected at similar concentrations (Nakamura *et al.*, 2009), and in a variety of mammals from the orders Cetartiodactyla and Perissodactyla both hydrogenotrophic groups were detected within the same sample (Morvan *et al.*, 1996). Overall, these results suggest that methanogens and sulfate reducing bacteria are likely not mutually exclusive. However, the nature of their relationship and likely competition for H_2 has not been fully characterized.

Currently, wood-eating termites are the only group of animals in which homoacetogens are the dominant hydrogen consumers in the intestine (reviewed in Brune, 2018). In the termite hindgut, there is a distinct spatial separation in terms of hydrogen concentration. In the center of the hindgut, called the lumen, H_2 concentrations are very high, while closer to the hindgut wall H_2 concentrations decrease. This gradient creates spatially separated niches and removes the possibility of any H_2 competition between homoacetogens and the methanogens. Further, the types of homoacetogenic bacteria in these termites are unique. Indeed, while the Spirochetes are motile bacteria that are not typically found in the intestinal tract of animals (Leadbetter *et al.*, 1999), homoacetogenic Spirochetes colonize the hydrogen-rich lumen of the hindgut, while methanogens attach to the hydrogen-poor hindgut wall where they are more competitive for H_2 and are safe from removal by peristaltic movement (reviewed in Brune, 2018).

Recent metatranscriptomic studies in the rumen have highlighted that hydrogenotrophy may be a more widespread trait in intestinal microbes than previously thought (Greening *et al.*, 2019). Indeed, this work found that fumarate and nitrate reducing bacteria are as active as other hydrogenotrophic groups like the methanogens and sulfate reducers in the rumen. In fact, in sheep that were characterized as 'low-methane producers', marker genes of these alternative hydrogenotrophic pathways are transcribed at higher levels than that of methanogens (Greening *et al.*, 2019). This suggests that fumarate and nitrate reducing bacteria could be used to mitigate methane emissions in some ruminants. Further studies are

needed to understand how these previously overlooked groups of bacteria can out-compete methanogens for hydrogen in low-methane yield sheep but not in other rumen systems (e.g. high-methane yield sheep(Mizrahi *et al.*, 2021)).

As with the intestinal archaea, studies targeting the hydrogen-consuming microbes in the intestinal tract have been largely focused on humans, ruminants, and termites. This has resulted in a limited understanding of the mechanistic details of how hydrogenotrophic microbes compete for H_2 in the intestine of various animals. Standardized, large-scale quantification of the sulfate reducing bacteria, methanogens, homoacetogens, and other hydrogenotrophic lineages (i.e., fumarate and nitrate reducers) will likely provide key insights into the dynamics of these important microbes across the animal kingdom.

Objectives

Despite their obvious importance, the Archaea have been largely overlooked in most intestinal microbiome studies. Therefore, it is currently not clear if the ability to colonize the intestine is a widespread trait or if it is limited to the above discussed lineages. Data concerning the diversity of archaea in a wide range of animals has recently started to be gathered (Youngblut *et al.*, 2020). However, the narrow sampling of studied animals has made it difficult to confirm a number of hypotheses, such as for example that were put forth by Hackstein *et al.* (1996) that archaea are tightly linked to the evolutionary history of their host. Even though Hackstein's work has provided pioneering insights into the distribution of intestinal methanogens across animals, it relied on methane gas detection which has technical limitations including: (*i*) some intestinal archaea present may not be methanogens (Nam *et al.*, 2008; Raymann *et al.*, 2017; Hervé *et al.*, 2020); (*ii*) the methanogens that are present may not be producing enough methane to be detected; and (*iii*) gas detection provides no taxonomic information on the methanogens present.

The use of archaea-targeting 16S rRNA primers would be able to overcome these challenges and provide a clearer picture of their distribution and diversity in the animal intestinal microbiome. Further, quantitative studies have been rarely performed in intestinal microbiome studies and information about the abundance of methanogens in different animal

intestinal tracts would provide key insights into the host factors that select for archaea - e.g., diet and gastrointestinal tract physiology. Finally, the number of available genomes from intestinal archaea is still extremely limited both in terms of taxa and origin (i.e., most available genomes are from humans, ruminants (cows and sheep), and termites **(Introduction Figure 6)**. Without a more diverse range of intestinal archaeal genomes, our understanding of how archaea have adapted to the intestinal microbiota remains limited.

With this thesis, I aim to address these concerns by focusing on three main objectives:

- Determine the diversity and abundance of host-associated archaea and some of the host-factors influencing them throughout the animal kingdom. To do this, I used archaea-targeting, 16S rRNA gene amplicon sequencing and quantitative PCR in 269 animal species from 6 major lineages (invertebrates, fish, amphibians, reptiles, birds, and mammals).
- Identify which bacteria consistently co-occur with archaea within the host intestinal tract. I used universal 16S rRNA gene primers to obtain the taxonomic profile of the bacterial community in these fecal samples and performed qPCR normalized co-occurrence network analyses to characterize the bacteria-archaea associations in various animal groups.
- Ascertain the adaptations in archaea linked to colonization of the animal intestinal tract. I used comparative genomic approaches to infer the genes that were gained and lost in *Methanimicrococcus blatticola* in association to its colonization of the intestinal tract of animals.

The first two objectives are addressed in the first chapter of my thesis, while results from the third objective are presented in the second chapter.

Chapter 1.A: Factors shaping the abundance and diversity of archaea in the animal gut

Because of the important role of intestinal microbes in host health and disease, there have been extensive research efforts into understanding which microbes are found in different animals and what determines the make-up of the microbial community in the intestine. However, most studies have focused sole on the bacterial components of these ecosystems. Thus, a systematic large-scale, sequence-based analysis of the intestinal archaea in animals has never been performed.

The work presented in the first part of this chapter discusses the data I gathered from quantitative PCR and 16S rRNA gene amplicon sequencing approaches on a large array of fecal samples covering animal diversity. The use of archaea-specific primers in both approaches revealed that archaea widespread throughout the animal kingdom, inhabiting the intestinal tract of invertebrates, amphibians, birds, reptiles, and mammals. Archaea are less abundant than the bacteria in all animals, and largely consist of four genera of methanogens and one phylum of aerobic ammonia oxidizing archaea. Animal diet and evolutionary history plays a large role in determine which archaea are present and how abundant they are.

This work is currently under revision at Nature Communications.

In the second part of this chapter, I present data collected concerning the abundance of sulfate-reducing bacteria in the same fecal samples, which were previously suggested to be in direct competition with the methanogens in the intestine. Interestingly, there was no relationship between sulfate-reducing bacteria and CO_2-reducing hydrogentrophic methanogens, or fibre consumption. However, there was a positive correlation between the Methanomassiliicoccales and sulfate-reducing bacteria in mammals. This suggests that these methyl-reducing methanogens are not directly competing with sulfate-reducing bacteria for hydrogen.

Finally, an analysis of publicly available microbiome datasets largely confirms the trends identified in this work – e.g. the high relative abundance of Methanomicrobiales in the Perissodactyla and the Nitrososphaerales in the Chiroptera. It also revealed new information such as the presence of Nitrosopumilales in marine invertebrates and the Methanomassiliicoccales in whales. These results clearly highlight the need to continue to further expand the understanding of the intestinal archaeome in marine life as well as in wild animal samples.

Factors shaping the abundance and diversity of archaea in the animal gut

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

[1]Institut Pasteur, Université de Paris, UMR CNRS2001, Unit Evolutionary Biology of the Microbial Cell, 75015 Paris, France

[2]Sorbonne Université, Collège doctoral, F-75005 Paris, France

[3]INRAE, Université Montpellier, LBE, Narbonne, France

Abstract

The intestinal microbiota has been the focus of intensive research for decades, however because most large-scale studies are concerned with only the largest portion of the microbiome the bacteria. Thus, a thorough analysis of the diversity and abundance of intestinal archaea in a wide range of animals is lacking. We examined the intestinal archaeome of 269 species of animals from eight classes. Archaea inhabit the intestine of all the lineages of animals we examined and are largely represented by four genera and one family of methanogens, as well as members of the Thaumarchaeota. Five major adaptation events to the intestine were identified in the Archaea. In mammals, host taxonomy, diet, and intestinal tract physiology majorly influence the structure and abundance of the archaeome. The abundance of methanogens is positively correlated with diet fibre content in mammals. Our results provide unprecedented insights into the intestinal archaeome and pave the way for further studies on these microbiomes in animals.

Introduction

The intestinal microbiome plays key roles in host health (Stevens *et al.*, 1998; Blaut *et al.*, 2007; Gordon *et al.*, 2012; Lozupone *et al.*, 2012; Arrieta *et al.*, 2014; Barko *et al.*, 2018; Kapitan *et al.*, 2018; Borrel *et al.*, 2020). It is composed of bacteria, archaea, microbial eukaryotes, and viruses/phages. Research on the microbiome of many animals has unveiled features that influence the overall structure of the intestinal microbiome such as diet and the ability to fly (Delsuc *et al.*, 2014; Groussin *et al.*, 2017; Youngblut *et al.*, 2019; Song *et al.*, 2020). However, most of these studies have only targeted the bacterial intestinal community. It is known that host-associated archaeal methanogens produce a significant amount of methane gas in ruminants, which makes them ecologically and environmentally important (Liu *et al.*, 2008). In humans, archaea have been linked to various conditions of health and disease (Borrel *et al.*, 2020). Nevertheless, archaea-centric intestinal microbiome studies have been generally conducted in a narrow group of animals such as termites, primates, humans, and ruminants (Hackstein *et al.*, 1994; Deevong *et al.*, 2004; Samuel *et al.*, 2007; St-Pierre *et al.*, 2013; Poulsen *et al.*, 2013; Borrel *et al.*, 2017, 2020; Danielsson *et al.*, 2017; Koskinen *et al.*, 2017; Raymann *et al.*, 2017; Brune, 2018).

The gut archaeome of other animals such as rats, hoatzin, pigs, seals, wallabies, kangaroos, iguanas, fish, horses, and even in the tissue of sponges was examined by independent studies using different molecular and cultural approaches (Van Der Maarel *et al.*, 1998; Evans *et al.*, 2009; Glad *et al.*, 2010; Hong *et al.*, 2011; Jackson *et al.*, 2013; St-Pierre *et al.*, 2013; Saengkerdsub *et al.*, 2014; Thomas *et al.*, 2016). Overall, these studies reported that the most common methanogens in the gut are members of the Methanobacteriales and Methanomassiliicoccales, and that the Methanosarcinales and Methanomicrobiales are also present, although less frequently (Cruzen *et al.*, 1986; Liu *et al.*, 2008; Borrel, O'Toole, *et al.*, 2013; St-Pierre *et al.*, 2013). Only one study addressed the distribution of intestinal methanogens in a wide variety of animals, but using methane gas detection tests (Hackstein, 1997). This study detected methanogens in a wide range of animals. It also suggested that they have been acquired early in animal evolution and were completely lost in some lineages such as the Carnivora. However, the methodology used in this study has several limitations, as it does not provide taxonomic information and cannot detect methanogenic populations with low concentrations in faeces or non-methanogenic archaea.

Here, we carried out a sequence-based analysis of the gut archaeome based on nearly 400 samples from 269 species covering a broad spectrum of animal diversity. We investigated the host range of archaea in eight animal classes, identified the major gut archaeal lineages and predicted the dominant methane metabolisms using both sequencing and quantitative approaches. We discussed the number of events of adaptation to the gut in the Archaea, including in ammonia-oxidizing Thaumarchaeota and in Bathyarchaeota, both previously rarely identified in this environment. By using a large range of metadata from the literature we define key factors structuring the abundance and composition of the gut archaeome across the animal kingdom.

Results and discussion

Archaea are present in the gut microbiome throughout the animal kingdom

We collected faeces from 269 species of animals (*n* samples = 391) ranging from Invertebrates to Mammals – the majority of which, except for birds, fish, and gastropods, came from captive specimens (Table S1). We used three approaches to characterize the archaeal community of these samples: *i*) quantitative PCR (qPCR) targeting total Archaea, total Bacteria, and five archaeal lineages previously found in the animal intestine (Methanobacteriales, Methanomassiliicoccales, Methanomicrobiales, *Methanimicrococcus,* and Thaumarchaeota), *ii*) 16S rRNA gene amplicon sequencing of the Archaea only, and *iii*) of the entire microbial community. We detected the presence of archaea in the gut microbiome of 175 species belonging to all eight classes of animals investigated, including 14 orders of mammals (Figure 1; Table S1).

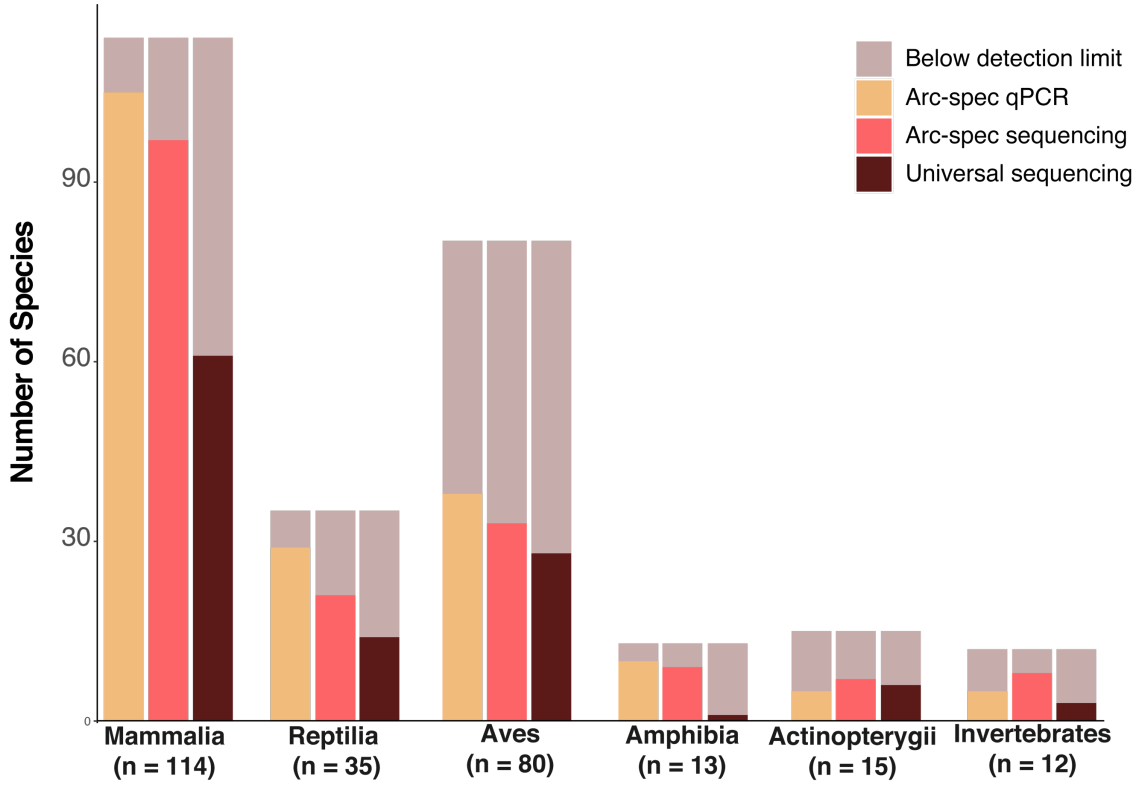


Figure 1. Detection of archaea in animal species with three approaches. Invertebrates gather 3 classes (Insecta, Mollusca, and Malacostraca).

Archaea were detected in a higher proportion of the samples when using archaea-specific primers for qPCR (78%) or amplicon sequencing (76%) compared to the universal prokaryote primers for amplicon sequencing (44%). This difference was observed in most animal classes (Figure 1). In addition, universal prokaryote primers also captured a lower number of ASVs (1.9 $\pm$ 2.6 ASVs per sample) compared to the archaea-specific primers (13.6 + 20.3 ASVs per sample) (Kruskal-Wallis p = 1.65e^{-8} (Figure S1), n = 218). With ~10,000 prokaryotic reads per samples, the archaeal species/ASVs that represent less than 0.01% of the microbial community are likely missed, which may explain both lower proportion of archaea-positive animals and the lower archaeal alpha-diversity in the approach relying on prokaryote universal primers.

Five major events of adaptation to the gut in Archaea

The broad taxonomic coverage of the animal hosts and the use of archaeal specific primers allowed us to identify archaeal ASVs belonging to 19 described families, 10 orders, 6 classes, and 3 phyla. 84.9% of these ASVs (94.5% of the reads) share more than 95% identity with

species in the Living Tree Project (LTP, v138) database (Yilmaz *et al.*, 2014) amended with characterized candidate species, and half of the reads share more than 99% identity with known species (Table S2). Consistently, the vast majority (93.7%) of the reads are affiliated to only six genera or families (Figure 2a): *Methanobrevibacter*, *Methanosphaera* (Methanobacteriales), *Methanomethylophilaceae* (Methanomassiliicoccales), *Methanocorpusculum* (Methanomicrobiales), *Methanimicrococcus* (Methanosarcinales), *Nitrososphaeraceae* (Nitrososphaerales/Thaumarchaeota group 1.1b). These lineages also constitute more than 50% of the gut archaeome in 92% of the sampled animals and can be qualified as "dominant gut archaea". Among them, *Methanobrevibacter*, *Methanosphaera* and *Methanomethylophilaceae* had already been extensively reported in the gut microbiota of ruminants, human and termites (St-Pierre *et al.*, 2013; Brune, 2018; Borrel *et al.*, 2020). Our results show that they are also the main methanogenic lineages in a much wider range of animals. *Methanobrevibacter* members are by far the most dominant methanogens in our dataset – composing over a third of the total number of reads, followed by *Methanomethylophilaceae* members which accounted for 17.5% of the total reads (Figure 2a). The two others methanogen lineages that are the most prevalent in our dataset, *Methanocorpusculum* and *Methanimicrococcus,* have been less often reported in previous studies on the animal digestive tract (St-Pierre *et al.*, 2013; Thomas *et al.*, 2021).

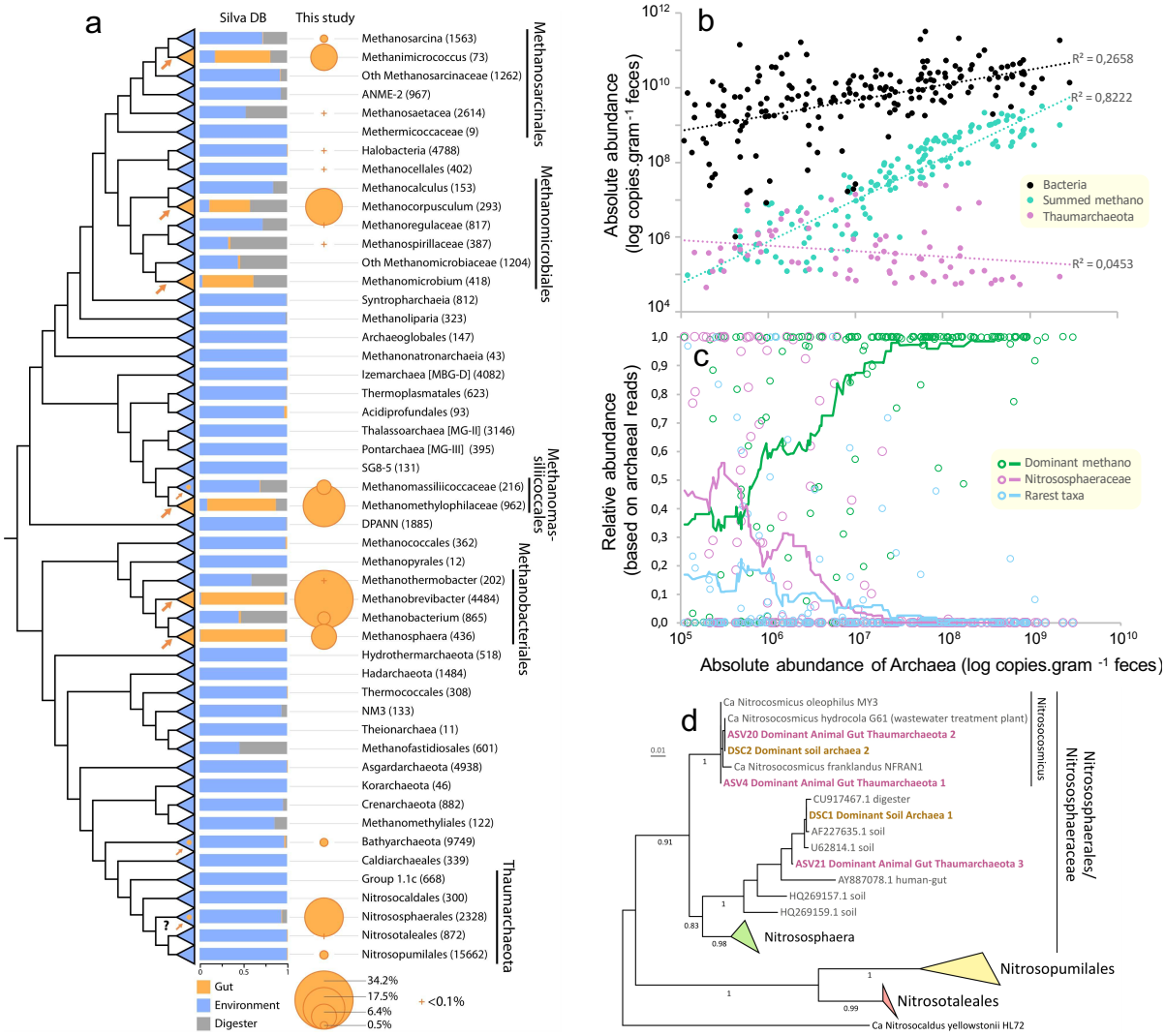


Figure 2. Proposed independent events of adaptation to the gut in the domain archaea. **a)** Distribution of archaeal 16S rRNA gene sequences in the gut and other environments based on sequences obtained from the Silva database and this study. The archaeal tree is based on Borrel et al., (Borrel *et al.*, 2019) enriched with DPANN lineages. Orange arrows on the tree indicate proposed events of adaptations to the gut environment, either at the base or within displayed lineages. The histogram shows the proportion of sequences from a given lineage present in either animal digestive tract ("Gut"), open natural environment ("Environment") or built environment ("Digester"). Circle surface area represents the percentage of reads attributed to each taxon in our study including only gut-related samples. **b)** Correlation between the absolute abundance of Archaea and the absolute abundance (16S rRNA copies/gram of faeces) of Bacteria (black), summed methanogen lineages (Methanobacteriales, Methanomassiliicoccales, Methanomicrobiales, *Methanimicrococcus*; green) and Thaumarchaeota (purple), all determined by qPCR using lineage specific primers. The scale of the absolute abundance of Archaea is on panel **c)** Plotted samples correspond to those with amplified Archaea in Miseq, presented in panel c. c) Proportion of archaea corresponding to the dominant methanogen lineages (green), *Nitrososphaeraceae* (purple) and rarer taxa (light blue) in samples, based on Miseq sequencing with archaeal specific primers, according to absolute abundance of archaea in the sample (qPCR). Dots indicate the relative

abundance of these three categories/lineages of archaea in each sample. Lines indicate the moving averages with a subset size of 25 samples. The dominant methanogen lineages category contains *Methanobrevibacter*, *Methanosphaera*, *Methanomethylophilaceae*, *Methanocorpusculum*, *Methanimicrococcus*. The rarest taxa category contains *Methanobacterium*, *Methanothermobacter*, *Methanomassiliicoccaceae*, *Methanosarcina*, *Methanoregulaceae*, *Methanospirillaceae*, *Methanosaeta*, Methanocellales, *Nitrosopumilaceae*, *Nitrosotaleaceae*, Bathyarchaeota, Halobacteriales. **d)** Phylogenetic position of dominant gut Thaumarchaeota (this study, ASV4, ASV20 and ASV21, purple) and dominant soil archaea (Bates *et al.*, 2011) (DSC1 and DSC2, brown). ASV4/ASV20 are similar to DSC2 representative sequence (only 1 indel in a 4/5Gs homopolymer region, which may be due to a 454-sequencing error in DSC2 (Huse *et al.*, 2007). ASV21 shares 99.2% identity with the DSC1 representative sequence.

The five dominant methanogen lineages in the gut have been rarely reported in open environments, as revealed by a meta-analysis of the archaeal 16S rRNA gene sequences present in the Silva database (Figure 2a). They are also often dominant in samples with the highest archaeal absolute abundance (Figure 2c). In contrast, *Nitrososphaeraceae* and rarer archaeal lineages (e.g., *Methanobacterium*, *Methanosarcina*, *Methanomassiliicoccaceae*) are more often dominating in samples with the overall lowest archaeal absolute abundance (Figure 2b and c). Sequences from the rarer archaeal lineages have been generally reported from non-gut environments such as sediments and wetland soils (Borrel, O'Toole, *et al.*, 2013; Söllinger *et al.*, 2016) (Figure 2a). Therefore, these lineages likely contain free-living archaea that have weaker capacity to develop in the gut and are only detectable when the most adapted archaea are absent or occur in low abundance. These lineages that are weakly associated with the gut microbiome are often the closest phylogenetic relatives to the ones that are highly abundant in the gut (Figure 2a). In addition, the dominant gut genera/families belong to orders that are often present in digestors/bioreactors, suggesting a more ancient adaptation to high resource availability in these lineages (Figure 2a). This suggests that some of the traits favouring development in the gut were already present in the last common ancestor of *Methanobacteriales*, *Methanomassiliicoccales* or *Methanosarcina+Methanimicroccus*, which may have facilitated the emergence of gut-specialized archaea, particularly in these lineages. Considering that the dominant gut methanogen lineages are rarely detected in open environments, and that closely related lineages are rarely present in the gut, a strong specialization to the gut microbiome likely occurred at the divergence of each of the five dominant gut methanogen lineages, suggesting

at least five major events of adaptation to the gut in the Archaea (Figure 2a). A few other methanogen lineages may have developed specific adaptations to certain gut conditions or specific hosts. For example, "*Candidatus* Methanomassiliicoccus intestinalis" was also detected in high absolute abundance in the gut of elderly people being in long-term residential care and having a disturbed gut microbiota (Borrel *et al.*, 2017). *Methanomicrobium mobile* was not detected in our samples (Figure 2a), but it can represent a large proportion of archaea in the rumen and it is almost never detected in open environments (Henderson *et al.*, 2015), suggesting a high specialization on this gut compartment, which was not examined here.

Non-methanogenic lineages are components of the animal gut archaeome

The family *Nitrososphaeraceae* (Thaumarchaeota), which gathers 15% of the total reads in our study, has rarely been reported in previous gut microbiomes studies (great apes and humans) and only when archaeal specific primers coupled with high-throughput sequencing (or nested PCR) were used (Rieu-Lesme *et al.*, 2005; Hoffmann *et al.*, 2013; Raymann *et al.*, 2017). Conversely, this lineage is common in terrestrial environments such as soils (Figure 2a) (Bates *et al.*, 2011; Pester *et al.*, 2011; Kerou *et al.*, 2015; Lehtovirta-Morley *et al.*, 2016). The high abundance of these obligate aerobes in the gut is surprising, even though oxygen is available in some gut sections and in proximity of the gut wall in other sections dedicated to fermentation (Sprenger *et al.*, 2007) and some animals host more aerobic communities than others (Sherrill-Mix *et al.*, 2018). The three thaumarcheotal ASVs that gather the largest number of reads in our dataset (ASV4/ASV20/ASV21) are also widely distributed among animal species, ASV4 being the most widespread archaeon in our samples (present in 65 animal species from 8 classes). These three ASVs correspond to the most prevalent and abundant archaeal phylotypes (named DSC1 and DSC2) among 146 soils from various biomes (Bates *et al.*, 2011) (Figure 2d). Because most animals live (eat, sleep, groom...) on soil, these dominant soil archaea may be ingested by chance, which could explain why they are present in a wide range of animals at a low absolute abundance (Figure 2d). However, a sequence closely related to ASV21 and DSC1 has also been found in the human gut (Rieu-Lesme *et al.*, 2005) (Figure 2d). Moreover, while Thaumarchaeota group I.1c are among the dominant archaea in soil (Bates *et al.*, 2011; Tripathi *et al.*, 2015), we only found one ASV belonging to this lineage in one sample. ASV4/ASV20 and DSC2 are closely related to several "*Candidatus* Nitrosocosmicus" species (Jung *et al.*, 2016; Lehtovirta-Morley *et al.*, 2016; Sauder *et al.*,

2017); Figure 2d). These species can grow at higher ammonia concentrations ((>20mM) prevailing in the gut (Hungate, 1966; Vester *et al.*, 2008; Wang *et al.*, 2012)) than other ammonia-oxidizing Thaumarchaeota (Lehtovirta-Morley *et al.*, 2016). Also, one of them was cultured from a wastewater treatment plant (Sauder *et al.*, 2017), an environment that shares some characteristics with the gut. Therefore, an alternative hypothesis is that the dominant Thaumarchaeota in the animal gut can maintain in this environment, which may be beneficial for their dispersion in soils, being spread by animal faeces. The degree of adaptation and role of *Nitrososphaeraceae* in the gut remains to be elucidated.

Finally, although Bathyarchaeota are not common in our samples, most of the sequences we retrieved are closely related to a clade formed by "*Ca.* Termiticorpusculum" and "*Ca.* Termitimicrobium" (>95% id to termite sequences), two lineages recently identified in the termite gut (Loh *et al.*, 2021). Together with sequences from anaerobic digestors and sediments, our sequences from mammal, birds, reptiles, and crayfish, form a sister clade to termite sequences (Figure S2), suggesting that some general traits needed to maintain in the gut are shared by these Bathyarchaeota.

Specific associations between archaea and their hosts

In mammals, the main factors that significantly affect the beta-diversity of archaea have with the following level of influence: host phylogeny > coefficient of gut differentiation > host diet > digestive tract type, regardless of the diversity measurement used – i.e., Weighted/Unweighted UniFrac, Bray-Curtis, and Jaccard (Table 1; Figure S3). Other factors such as the geographic origin of the samples and the body mass have little influence on the archaeal community structure (Supplementary text).

Beta diversity measure		**Host order df = 10**	**Gut diff Coeff n = 23 df = 1**	**Diet df = 7**	**GIT type df = 4**	**Body mass df = 1**	**qPCR archaeal abundance df = 1**	**Arc:bac ratio (qpcr) df = 1**	**Stomach pH n = 18 df = 1**	**Mean retention time df =1**	**Origin***
Weighted unifrac	R^2 =	0.40	0.32	0.20	0.10	0.06	0.04	0.04	*ns*	*ns*	*ns*
	p =	0.001	0.0009	0.002	0.02	0.002	0.01	0.03			
Unweighted unifrac	R^2 =	0.33	0.17	0.19	0. 13	0.03	*ns*	*ns*	*ns*	*ns*	0.05
	p =	0.001	0.0009	0.001	0.001	0.003					0.035

Bray-Curtis	R^2 =	0.25	0.13	0.16	0.14	0.02	0.02	0.02	*ns*	*ns*	0.05
	p =	0.001	0.0009	0.001	0.001	0.01	0.03	0.01			0.018
Jaccard	R^2 =	0.21	0.13	0.14	0.11	0.02	0.02	0.02	*ns*	*ns*	0.05
	p =	0.001	0.0009	0.001	0.001	0.02	0.02	0.02			0.018

Table 1. Factors that influence the Beta Diversity of archaea in Mammals. Mammals with > 2 species per order (n = 73, unless otherwise indicated) rarefied to 3000 reads per sample were subject to beta diversity analyses. * including only zoo from which more than three samples were collected, and samples from the same species were treated separately (n = 99; df = 11). Signficant differences were tested for between beta diversity metrics using a permutational analysis of variance (PERMANOVA), $p < 0.05$ was considered signifcant.

The level of beta-diversity variance explained by host phylogeny is as high (or higher) as the one previously reported for bacterial communities in mammals (Ley *et al.*, 2008; McKenzie *et al.*, 2017; Nishida *et al.*, 2017; Sherrill-Mix *et al.*, 2018; Katherine R Amato *et al.*, 2019; Youngblut *et al.*, 2019; Song *et al.*, 2020). Specific associations between archaeal and animal host lineages are visible through the dominance of gut archaeome by Methanobacteriales/*Methanobrevibacter* in Rodentia and most Cetartiodactyla, *Methanomassiliicoccales*/*Methanomethylophilaceae* in *Lemuridae*, Methanomicrobiales/*Methanocorpusculum* in Perissodactyla and several Reptiles or Thaumarchaeota/*Nitrososphaeraceae* in Gastropoda (Figure 3d). A link between archaea and their host is also evident from the existence of archaeal clades associated with specific host

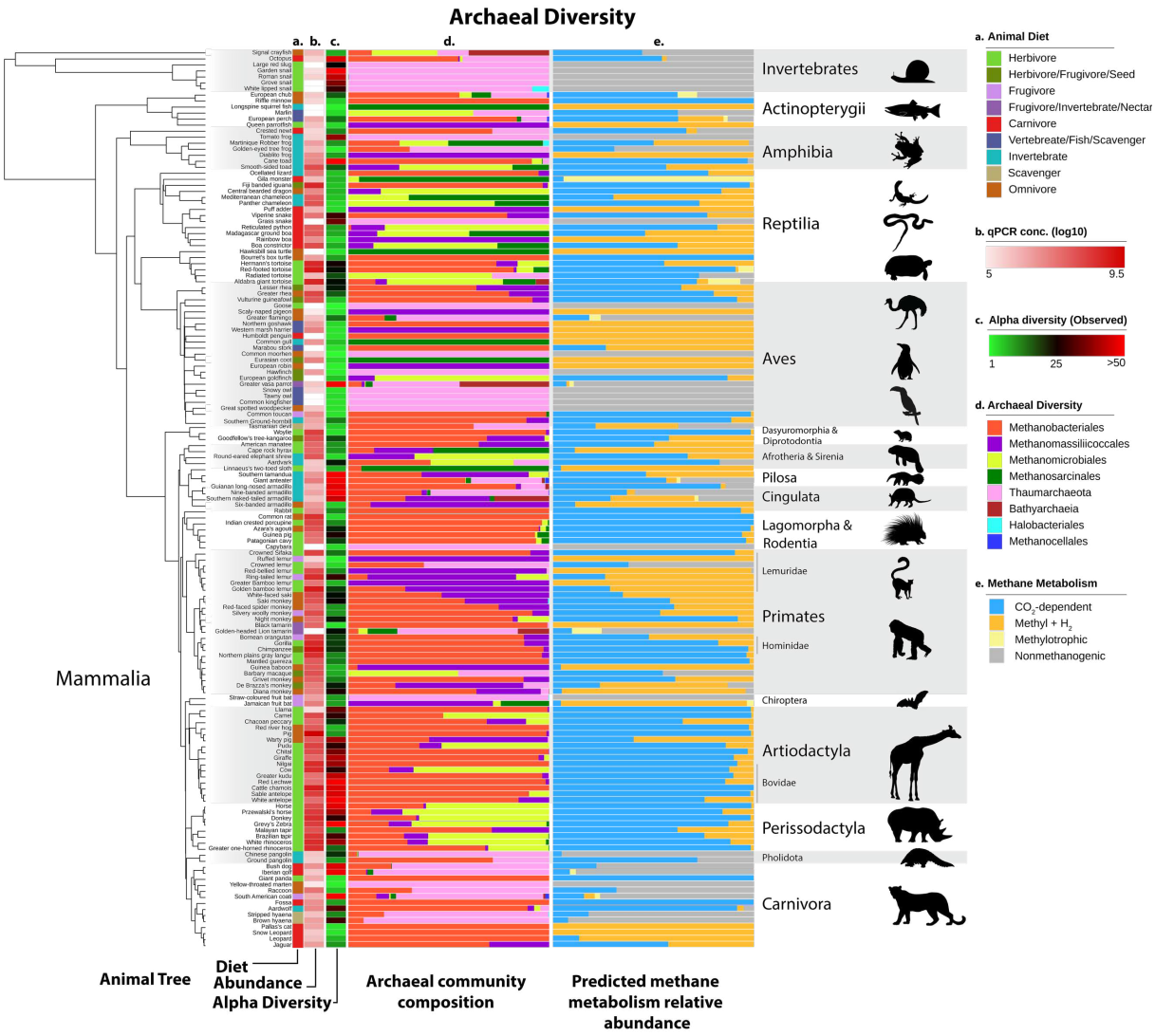


Figure 3. Archaeal taxonomic diversity and abundance in the animal gut. **a)** Information on animal primary diet gathered using the Elton Trait database, the Animal Diversity Website database, or from specialists who provided fecal samples. Primary diet was considered food material that made up ≥70% of the animal's diet**. b)** Absolute abundance of archaea as determined by qPCR with archaea-targeting primers on a log scale. **c)** Observed richness (number of different ASV) of archaea. **d)** Taxonomic diversity of archaea in the animal intestinal microbiome. Samples were rarefied to 3000 archaeal reads. **e)** Predicted methane metabolism, assigned to ASVs based on taxonomic annotation (Table S4). The Animal Tree was generated using Timetree.org (Kumar *et al.*, 2017).

orders. *Methanobrevibacter* is the most structured genus with respect to these specific associations, several *Methanobrevibacter* clades being enriched in one type of mammal host (Primates, Cetartiodactyla, Perissodactyla Rodentia) (Figure 4). These results are consistent with the previous report of *Methanobrevibacter* OTUs showing phylogenetic association with hosts (Youngblut *et al.*, 2019). In *Methanocorpusculum*, a large Perissodactyla-associated

clade is sister to a Cetartiodactyla-associated clade (Figure S4), suggesting that the ancestor of these two archaeal clades was present in the ancestor of the Ungulata. Close relationships between Perissodactyla and Cetartiodactyla ASVs are also visible in *Methanobrevibacter* and *Methanosphaera* (Figure 4).

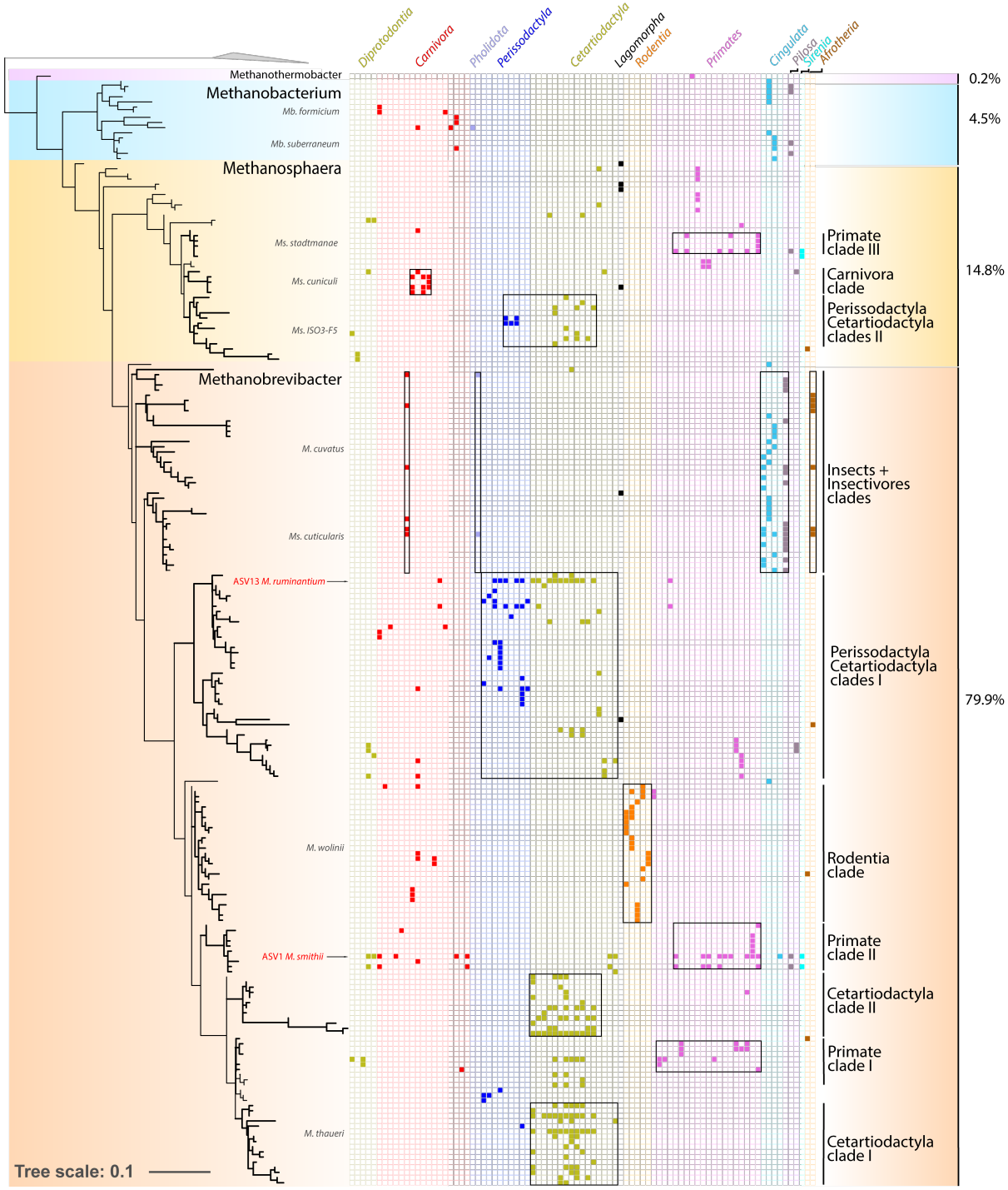


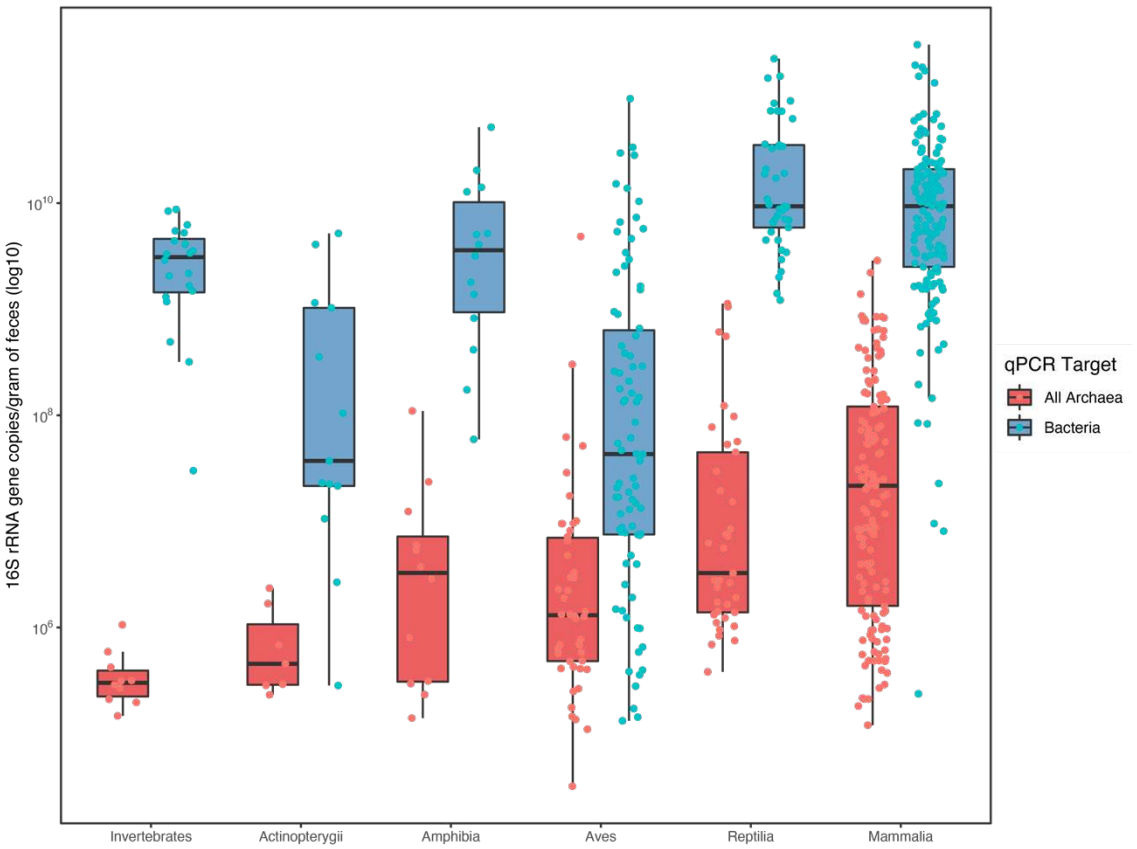
Figure 4. Distribution of Methanobacteriales ASVs among Mammals. The phylogenetic tree (maximum-likelihood, GTR+G4) was built with nearly full length 16S rRNA genes sequences from literature and the ASVs

sequences from this study. For display purposes, the shown tree includes only the ASVs representing more than 1% of the sequences per sample. The percentages on the right indicate the proportion of reads from *Methanobacteriales* that were annotated as *Methanobrevibacter*, *Methanosphaera*, *Methanobacterium* or *Methanothermobacter*. The squares in front of each leaf are colored according to Mammalian order, and represent the presence/absence of each ASV across orders. Black boxes highlight order specific clades and correspond to the clade labels on the right-hand side of the graph.

In the *Methanomethylophilaceae*, there is a large Primates-associated clade containing several of the typical human-associated species (Mx-03, Mx06; (Borrel *et al.*, 2017); Figure S5). Fewer host-specific clades are observed outside mammals, except for reptile-specific clades in *Methanocorpusculum.* These clades complement the previously reported insect/termites-specific ones within *Methanobrevibacter*, *Methanomethylophilaceae* and *Methanimicrococcus* (Brune, 2018; Thomas *et al.*, 2021)) and support the hypothesis that archaea developped adaptations for specific host lineages with which they may have been associated for a long evolutionary time. In contrast, no clear host-associations are visible in Nitrososphaerales (Figure S6) which points at the absence of specialization to individual/unique animal guts and supports the hypothesis of their low level of adaptation to the gut, as discussed above.

Host phylogeny also influences the absolute abundance and alpha diversity of archaea. Indeed, mammals have the highest absolute abundance of archaea, followed by reptiles and amphibians (Figure 5a). Birds, bony fishes, and invertebrates have instead the lowest concentrations. This global trend for archaea is mostly driven by variation in the methanogen lineages, particularly concerning the *Methanobacteriales* (Figure S7; Supplementary text).

a. Absolute Abundance of Archaea and Bacteria



b.

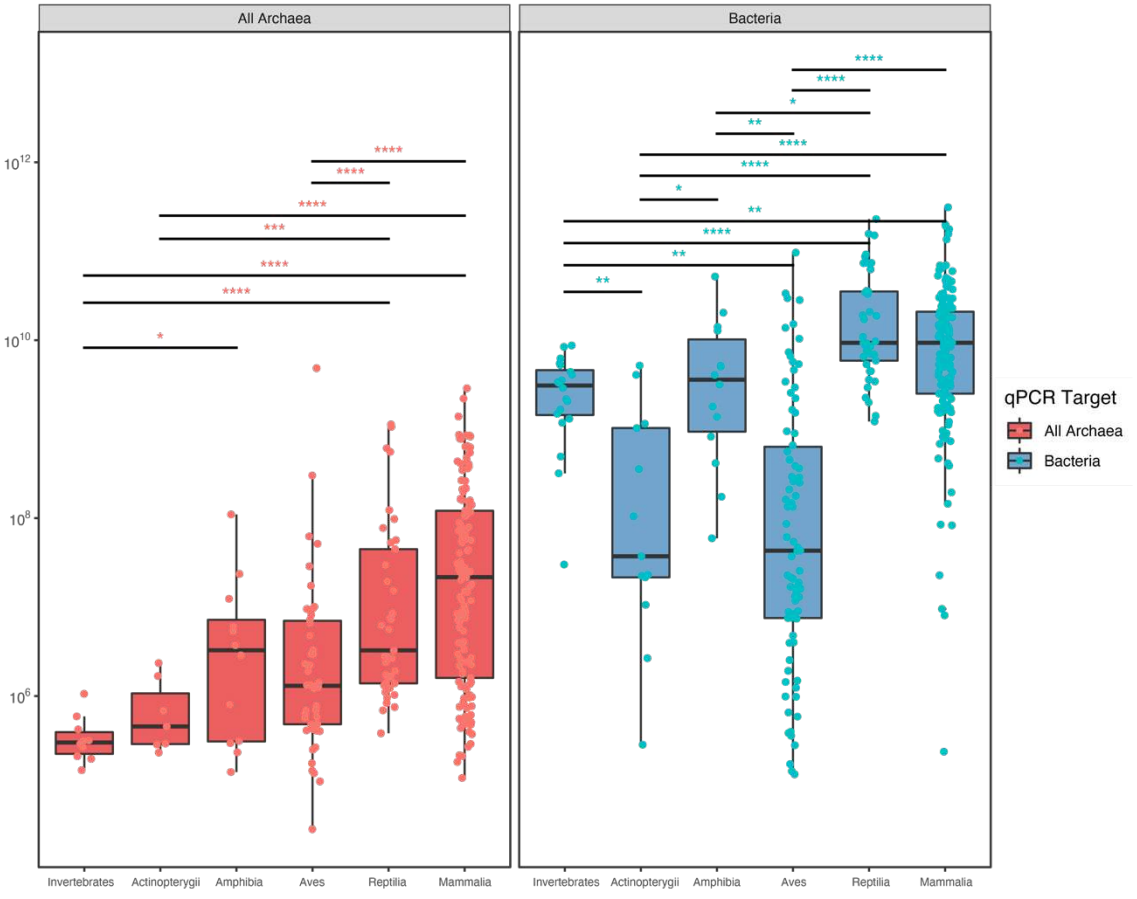


Figure 5: Absolute Abundance of Archaea (red) and Bacteria (blue) determined via qPCR in **a)** animal classes (n = 286) and **b)** pairwise comparisons of the absolute abundance between animal classes. Animal lineages with significantly different archaeal/bacterial abundances are labeled. Wilcoxon rank sum *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$.

Archaea were also detected in a lower proportion of bird and fish species than in other animal classes (Figure 1; Figure S8). In Mammalia, samples belonging to Perissodactyla, Cetartiodactyla, Primates, Diprotodontia, and Rodentia have the highest absolute concentration of archaea, whereas those belonging to Carnivora and Pholidota have the lowest one (Figure 3b). Conversely, the abundance of bacteria is more uniform across mammalian orders (Figure 3b). Closely related groups of animals also tend to have similar levels of archaeal alpha diversity, as supported by the Moran index (I=0.08, p= 0.001, n = 150). For example, the archaeal richness is high in the members of Gastropoda and in most members of the Cingulata, Equidae (order Perissodactyla), and Bovidae (order Cetartiodactyla) within mammals (Figure 3b; Table S1). Conversely, we found comparably low levels of archaeal richness within the Aves and Actinopterygii.

Strong influence of diet on methanogen abundance and composition

Diet is another important factor affecting the gut archaeome, as demonstrated by alpha-diversity, beta-diversity (Table 1; Figure S9), and absolute abundance (Figure 6).

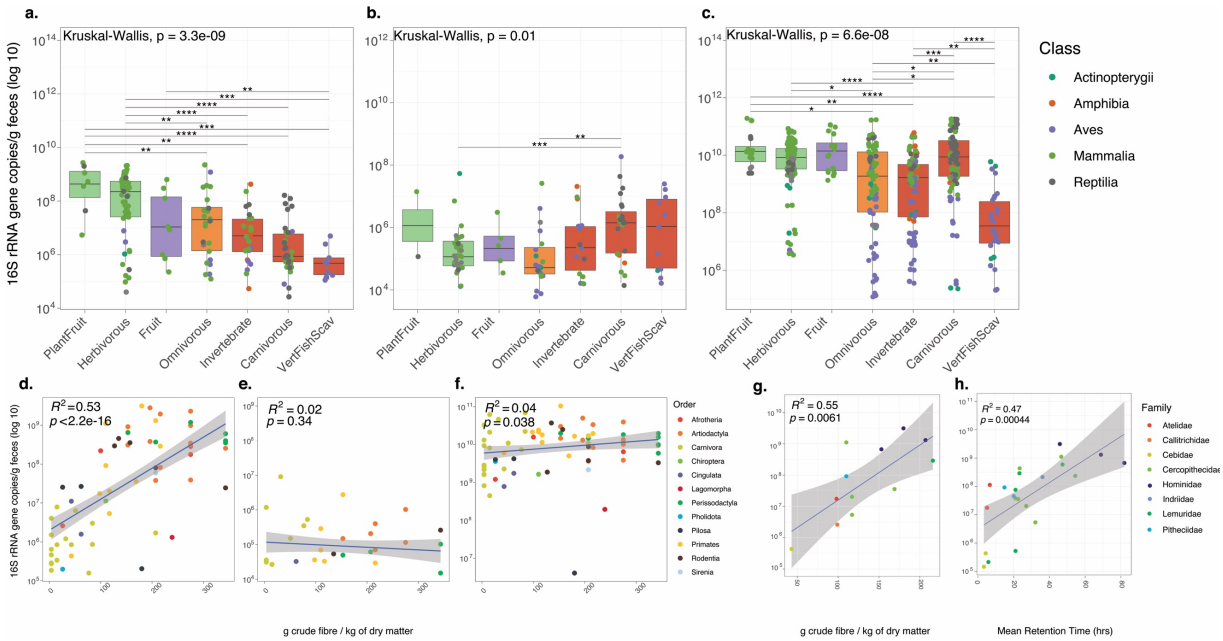


Figure 6. Influence of host diet-type, diet-fibre content, and mean retention time on the absolute abundance of total methanogens, Thaumarchaeota and Bacteria. Abundance of **a)** total methanogens (n = 161), **b)** Thaumarchaeota (n = 116) and **c)** Bacteria (n = 223) according to host diets-type. Significant differences across all groups were determined via the Kruskal-Wallis test, with $p < 0.05$ set as 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 **d)** methanogens (n = 103), **e)** Thaumarchaeota (n = 55) and **f)** Bacteria (n = 107) in mammal species. **g)** Dietary fibre versus averaged absolute abundance of methanogens in Primates (n = 12). Mean retention time is significantly related to the abundance of methanogens in primates. Statistical analyses and representation of the absolute/relative abundance of methanogens were carried out on species where archaea have been detected. h) Correlation between digesta mean retention time vs averaged absolute abundance of methanogens in Primates (n = 24). Fibre consumption is significantly correlated with the abundance of Methanogens in primates.

Indeed, herbivorous animals have a higher number of archaeal ASVs than carnivorous and omnivorous animals (Figure S9). Moreover, the absolute and relative abundance of methanogens is higher in animals with a plant-based diet (e.g., leaves, fruits) than in animals feeding on meat or insects, and their abundance is intermediate in omnivorous animals (Figure 6a). This link between methanogen abundance and diet-type is further supported by the positive correlation of both the absolute and relative abundances of methanogens (but not of Thaumarchaeota, and very weakly for bacteria) with the fibre content of the diet (Figure 6d-f; Figure S10). The increase in methanogen absolute/relative abundance reaches a plateau at around 200 g of crude fibre/kg of dry matter (Figure 6d; Figure S10). At lower host taxonomic levels, the positive correlation also holds for Primates, for which we sampled species with contrasted average fibre intake (Figure 6g). An increased diet fibre content was previously reported to be associated with a higher expression level of methanogenesis genes in humans (Tap *et al.*, 2015) and a greater methane production in pigs (Jensen *et al.*, 1994) and ruminants (Shibata *et al.*, 2010). As the vast majority of intestinal methanogens are hydrogenotrophic, these relationships can be explained by the higher production of hydrogen from fibre/carbohydrates-rich diets (plant) than from protein/fat-rich diets (meat) (Alibardi *et al.*, 2016).

However, the level of H_2 produced from fibre degradation also depends on which bacteria are involved, Clostridiales are known to produce more H_2 than Bacteroides when degrading fibre (Chassard *et al.*, 2010). Thus, other than the host's diet, methanogens are also

influenced by the composition of the bacteria degrading it. In humans, cellulolytic *Ruminococcaceae* (Clostridiales, Firmicutes) spp. have been reported to be present in the gut of methane producers, while cellulolytic *Bacteroides* spp. prevail in non-methane producers (Chassard *et al.*, 2010). Further, methanogens are generally enriched in individuals with the Firmicutes/*Ruminococcaceae* enterotype (Arumugam *et al.*, 2011). We found that eight *Ruminococcaceae* OTUs (including six from uncharacterized genera) co-occur with methanogens, and -more generally- 19 out of the 30 bacterial OTUs positively associated with methanogens belong to Clostridiales and only four to Bacteroidales (Table S3, Supplementary text). Other than benefiting from fibre degradation, methanogens can also favor it by stimulating microbes involved in its degradation. Indeed, the presence of methanogens in co-cultures has been shown to increase the level of extracellular polysaccharide-degrading enzymes of *Ruminococcus flavefaciens* (Williams *et al.*, 1994).

The abundance of hydrogenotrophic methyl-reducing methanogen lineages (i.e., Methanomassiliicoccales and *Methanimicrococcus*) is less influenced by fibre content than lineages that include hydrogenotrophic CO_2-reducing methanogens (i.e., Methanobacteriales and Methanomicrobiales; Figure S11). Moreover, hydrogenotrophic methyl-reducing methanogens represent a smaller proportion of the methanogens in herbivorous animals than in animals having another type of diet ($p = 0.003$). As methyl-reducing methanogens depend on different methyl-compounds (e.g. methanol, methylamines) for their energy metabolism and because they can utilize hydrogen at lower concentration than CO_2-reducing methanogens (Feldewert *et al.*, 2021), their distribution may be more affected by the availability of methyl-compounds than by fibre content. One of these methyl-compounds, methanol, is produced by the bacterial degradation of pectin (Schink *et al.*, 1980). This metabolism was shown to occur in the animal gut (e.g., human, pigs, lemurs, ruminants) as revealed by the identification of bacteria with a methylesterase activity (Dehority, 1969; Kelly *et al.*, 2019) and by the increase in methanol concentrations in response to pectin consumption (Lindinger *et al.*, 1997; Drochner *et al.*, 2004; McKenney *et al.*, 2018). Our data show that the ratio of hydrogenotrophic methyl-reducing to CO_2-reducing methanogens is higher in frugivorous species than in herbivorous ones ($p = 0.005$), which is likely related to large amounts of pectin in fruits. This supports a previous hypothesis that the high relative abundance of *Methanosphaera stadtmanae* (an obligate methanol-reducing methanogen) in orangutans is related to their high fruit consumption (Facey *et al.*, 2012).

We also found a high relative abundance of hydrogenotrophic methyl-reducing methanogens in most of the sampled Primates (Figure 3e), and particularly in *Lemuridae*, which may be related to the presence of fruits in their diet (Table S1). This relationship is further substantiated by the association between an archaeal OTU closely related to *Methanomethylophilaceae sp.* Mx06 (Borrel *et al.*, 2017) and a bacterial OTU closely related to *Lachnospira pectinoschiza* (OTUarc_11; OTUbac_2345; Table S3). This bacterium grows mainly on pectin, producing methanol as a by-product of its degradation (Cornick *et al.*, 1994), and *Methanomethylophilaceae sp.* Mx06 has the genetic potential to grow by reducing methanol and methylamines with H_2 (Borrel *et al.*, 2017). A similar link may exist in humans, as both *Lachnospira pectinoschiza* and *Methanomassiliicoccales* abundances increase with age (Gaci *et al.*, 2014; Odamaki *et al.*, 2016). Moreover, *Methanomethylophilaceae sp.* Mx06 is the dominant archaeon in the gut of Yanomami Amerindians (Clemente *et al.*, 2015; Borrel *et al.*, 2017), whose diet is largely composed of fruits (Milliken *et al.*, 1999).

As *Methanomassiliicoccales* and *Methanimicrococcus* may also grow on other methyl-compounds than methanol, such as trimethylamine, they might be influenced by other types of diet. *Methanomethylophilaceae* OTUarc_11 is also correlated with an OTU closely related to *Sarcina sp.* (OTUbac_4310; Table S3) that can produce trimethylamine (Fennema *et al.*, 2016). A similar correlation between *Methanomethylophilaceae* and *Sarcina* was previously reported in the human gut (de la Cuesta-Zuluaga *et al.*, 2021). Precursors of trimethylamine (i.e. glycine-betaine, carnitine and choline) are present in various diets (De Zwart *et al.*, 2003; Wang *et al.*, 2011; Koeth *et al.*, 2013) and pectin is not limited to fruit but is also a constituent of the plant cell wall (Voragen *et al.*, 2009). Thus, the presence of hydrogenotrophic methyl-reducing methanogens is not limit to just frugivorous animals. In our dataset, hydrogenotrophic methyl-reducing methanogens constitute almost 40% of the overall methanogen reads (Figure 7a; Table S4) and represent a large fraction of the methanogens in many animals (Figure 3e; Supplementary text). This contrasts with many non-host environments (e.g. sediments, peat bogs), where hydrogenotrophic methyl-reducing methanogens constitute a minor fraction of the overall methanogens (Söllinger *et al.*, 2016; Cozannet *et al.*, 2021).

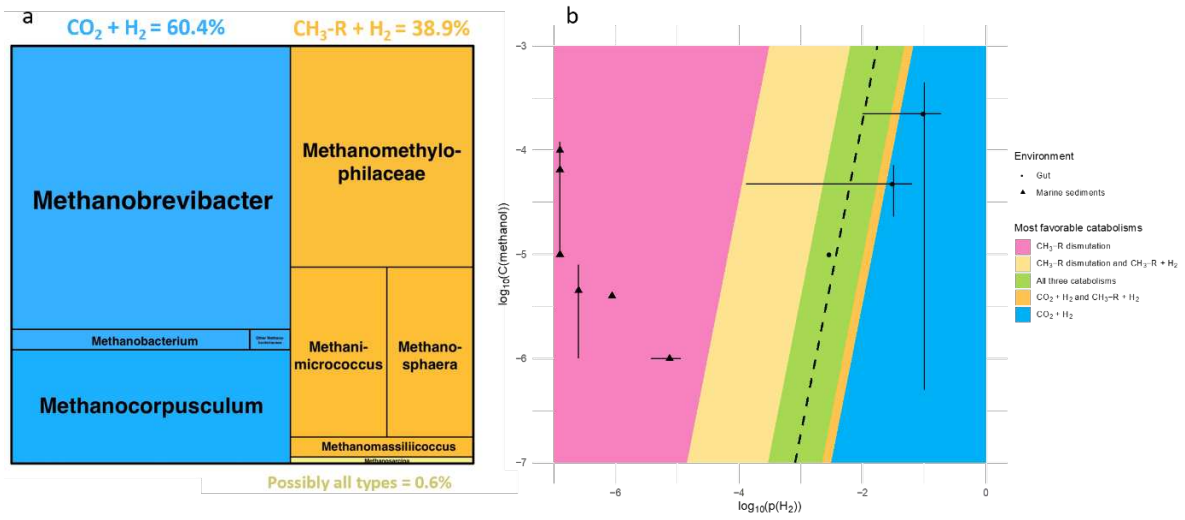


Figure 7. Main methanogenesis pathways in the animal gut. **a)** Proportion of the total archaeal reads that are assigned to taxa with a predicted CO_2-dependent hydrogenotrophic methanogenesis (H_2 + CO_2; blue) or methyl-dependent hydrogenotrophic methanogenesis (CH_3-R + H_2; orange) pathway. *Methanosarcina* spp. can have diverse methanogenesis pathways (the two above-mentioned pathways and the methyl-dismutation (or methylotrophic) and acetoclastic pathways. **b)** Diagram of the most favourable methanogenic metabolisms depending on methanol concentration (C(methanol) in mol/l) and hydrogen partial pressure ($p(H_2)$ in bar). Coloured area in the map indicate ranges of C(methanol) and $p(H_2)$ for which either CH_3-R dismutation, CH_3-R + CO_2 or CO_2 + H_2 is the most favourable pathway, i.e. concentrations and pressure ranges for which the associated ΔG_{cat} expressed in kJ/mol CH_4 is the most negative. ΔG_{cat} values were calculated for T = 298 K, pH = 7 and $p(CO_2) = p(CH_4) = 10^{-1}$ bar, when the difference in ΔG_{cat} between two or three catabolisms was less than 10 kJ/mol CH_4, catabolisms were then considered to be equally favourable. This corresponds to central coloured areas in the diagram where either two or three metabolisms are shown as equally favourable. The dotted line indicates particular values of C(methanol) and $p(H_2)$ for which all three catabolisms have exactly the same ΔG_{cat}. Ranges of C(methanol) and $p(H_2)$ found in the literature for either gut or marine sediments environments were also mapped on the graph: dots correspond to mean values and bars indicate minimal and maximal values. See Material and Methods section for the references on the studies providing the concentrations.

It also reinforces the hypothesis that the gut environment is particularly propitious for this kind of methanogenesis, which could have led to the transition from methylotrophic and CO_2-reducing methanogenesis to methyl-reducing methanogenesis in *Methanimicrococcus* and *Methanosphaera*, respectively (Thomas *et al.*, 2021). Other methane metabolisms (based on dismutation of methyl-compounds or acetate) are almost absent from the animal gut and may occur in the few *Methanosarcina* members (0.6 % of the total reads) identified in our dataset (Figure 7a). The few data on methanol concentration and H_2 partial pressure in the gut highly

contrast with those observed in marine sediments where dismutation of methyl-compounds is prevalent (Figure 7b). Gut conditions with relatively high partial pressure of hydrogen are generally more favourable for CO_2-reducing hydrogenotrophic methanogenesis but can vary to be equally favourable for methyl-reducing hydrogenotrophic methanogenesis or even for dismutation of methyl-compounds (Figure 7b). Why methanogens having this latter metabolism are not more prevalent in the gut is unclear, it may be related to the conditions generally favouring other metabolisms but this could also be related to the lower affinity of their enzymes for methanol and H_2 than methyl-reducing methanogens (Sprenger *et al.*, 2007).

The influence of diet is further substantiated by the link between the archaea present in the gut microbiota of both predators and their preys. Indeed, within *Methanobrevibacter*, sequences of insectivorous mammals from distinct orders (Carnivora, Pilosa, Pholidota, Cingulata, Afrotheria) are clustered with sequences from insect clades (Figure 4) reported in the literature (Brune, 2018). Except for cane toad (Urodela), no sequences from non-mammal insectivores are present in these clades. Moreover, a similar phylogenetic clustering of insectivores and insect-derived archaeal sequences was not observed for other known insect clades outside of *Methanobrevibacter* (even if a few non-mammal insectivores are present in the *Methanimicrococcus* insect clade). This suggests that *Methanobrevibacter* may also develop in the gut of insectivorous mammals while other insect-associated methanogens are mostly transients.

Impact of digestive tract physiology

Both the coefficient of gut differentiation (Chivers *et al.*, 1980) (i.e., proportion of the gut dedicated to fermentation) and where the fermentation takes place (e.g. foregut, hindgut, caecum) explain part of the variance in the beta-diversity (Table 1). In addition, many ASVs are almost ubiquitous in the ruminant Cetartiodactyla (paraphyletic, Ruminantia and Tylopoda), but mostly absent from non-ruminant Cetartiodactyla or other animals, highlighting a possible cross influence of gut physiology and host-phylogeny (Figure S12). Whether these archaea found in faeces originate from the rumen compartment or can colonize more largely the gut of these animal is currently unknown. The total abundance of methanogens is positively correlated with gut differentiation coefficient in mammals (R^2 =

0.33, p = 0.0036, n = 25), while there was no correlation with abundance of Thaumarchaeota and Bacteria (Figure S13).

It was previously reported that digesta mean retention time (MRT) is positively correlated with methane emission in herbivorous mammals (Clauss *et al.*, 2020). Also, a positive relationships between methanogen abundance and MRT was reported in humans (Triantafyllou *et al.*, 2014). When considering only Primate species, we also highlight a positive correlation between MRT and methanogen (or total archaea) abundance (Figure 6h). However, we found only a weak positive correlation between MRT and methanogen (or total archaea) abundance in all animals (R^2 = 0.11; p = 0.02, n = 69). However, the distribution of the values suggests that the abundance of methanogens is mostly influenced by the lower range of MRT values. Indeed, there is a stronger positive correlation (R^2 = 0.33; p = 4.3e-06, n = 54) for MRT values ranging from 0.5 to 50 h and significantly less archaea in animals with an MRT < 20h than animals with an MRT > 20h (Kruskal-Wallis p = 0.0008, n = 69). Diet type and MRT are generally related since digestion of fibre-rich diet relies on microbial fermentation ("allo-enzymatic" digestion) which requires longer MRT than digestion of protein/soluble carbohydrate-rich diets that are processed at high rate by animal enzymes ("auto-enzymatic" digestion) (Karasov *et al.*, 1986; Clauss *et al.*, 2008). An exception is the extreme case of the giant panda, an herbivore with a short MRT (8 h) that is a clear outlier in the relationships between fibre content and both methanogen absolute and relative abundances (Figure S10). Conversely, carnivorous reptiles have a long MRT, which can exceed a week for some large snakes (Secor *et al.*, 1995). Among them, boa constrictor and reticulated python have a high abundance of methanogens compared to other carnivorous animals which suggests that a long MRT can allow a substantial development of methanogens on meat diet. An increase in the relative abundance of Firmicutes, potential partners of methanogens, has also previously been recorded in Burmese pythons during digestion (Costello *et al.*, 2010). However, while long MRT may facilitate methanogen development on meat-diet, it should be stressed that most carnivorous reptiles have a low abundance of methanogens. Outside of Mammals and Reptiles, most birds, fish, amphibian, and invertebrates have generally a low concentration of archaea and especially of methanogens (Figure 5). Many flying birds feeding on plant materials use only readily digestible components of their diet, and rapidly expel recalcitrant cell-wall constituent without significant microbial fermentation (Mackie, 2002). This may be an adaptation to improve flight power by decreasing the body mass (Mackie, 2002). It is thus

likely that the short transit time and the low level of plant fermentation have a negative impact on methanogen abundance in birds. In addition to low concentrations of methanogens in Carnivora, bird and fish, we found no clear archaeal clades associated with these animals (with a few exceptions, like a small Carnivora-associated clade in *Methanosphaera*; Figure 4), suggesting that no lineage of methanogens developed strong adaptations to these hosts.

It was proposed that some animals, including birds, rely relatively little on their gut microbiota (Hammer *et al.*, 2019). In addition, bacteria recovered from birds show little host specificity and do not display phylosymbiotic patterns with their host or correlation with diet, differently from what has been generally observed in most mammals gut microbiota (Song *et al.*, 2020). In our dataset, the low abundance of bacteria in the gut microbiota of bird supports the hypothesis of Hammer et al., (Hammer *et al.*, 2019) and extends the observations of Song et al., (Song *et al.*, 2020) on the particularity of the gut microbiota of these animals. However, we found that concentrations of faecal bacteria in other animals proposed to rely less on their gut microbiota, such as Carnivora species (Hammer *et al.*, 2019), are as high as in other mammals (Figure S18).

Conclusions

Our work provides first key insights into the lifestyle and role of intestinal archaea across a diverse range of animal hosts. Increased sampling efforts, time-series analyses, and metagenomic investigation will help to answer standing questions about the impact of geography, captivity, residency, and adaptations of intestinal archaea throughout the animal phylogeny.

Material and Methods

Sample collection and DNA extraction

A majority of animal faecal samples were donated from various zoological institutions in France (Table S1). Fresh faecal samples (n = 392) were stored at -20°C until DNA extraction. Total DNA was extracted using a modified QIAamp PowerFecal DNA Kit (Hilden, Germany) protocol. Cells were lysed using the Fastprep (MP Biomedicals) cell homogenizer ‘faecal sample’ default setting in the lysis buffer provided in the PowerFecal DNA kit. For subsequent analyses, genomic DNA was diluted ten times, to limit the effect of PCR inhibitors.

Quantitative PCR

Total bacteria, total archaea, and specific archaeal lineages (Methanobacteriales, Methanomassiliicoccales, Methanomicrobiales, *Methanimicrococcus*, Thaumarchaeota) were quantified using quantitative PCR with lineage specific primers (Table S5). qPCR was performed on a qTower3 Touch device (Analytik Jena GmbH) using SensiFAST SYBR® & Fluorescein Kit (Bioline, Paris, France). For each run, a standard curve was prepared using a 10-fold serial dilution (10^9 to 10^1 copies/µl) of a plasmid containing a 16S rRNA. Plasmids containing a partial archaeal or bacterial 16S rRNA gene were generated through cloning PCR amplified 16S rRNA gene of the groups into *E. coli*. Bacterial 16S rRNA genes were amplified from a faeces sample using the B-27F-YM/B-1492R primer set (Lane, 1991; Nercessian *et al.*, 2005). Archaeal 16S rRNA genes were amplified from *Methanimicrococcus blatticola*, *Methanocorpusculum aggregans*, *Methanomethylophilus alvus*, *Methanosphaera stadtmanae* and *Nitrososphaera viennensis* using the A-21F/A-1386R primer set (Delong, 1992; Skillman *et al.*, 2004). PCR products were cloned with a pGEM-T vector according to the manufacturer's instructions (Promega, Charbonnières-les-Bains, France). The accuracy of the plasmid construction was confirmed through sequencing and all plasmids were diluted to 10^9 copies/µl, aliquoted and stored at -20°C. The accuracy of the qPCR assay was confirmed through melting curve analysis. All quantifications were performed twice in independent runs. The final concentration of all the microbial was averaged between replicates and normalized as copies of 16S rRNA gene per gram of faeces.

16S rRNA gene amplicon sequencing

Archaeal 16S rRNA genes were amplified in two steps (Nested-PCR; Table S6) to allow the inclusion of a larger range of samples. Prokaryotic 16S rRNA genes were directly amplified with Illumina tagged primer pairs (Table S6). Sequencing was performed on an Illumina MiSeq platform (Biofidal, Vaulx-en-Velin, France) according to the Illumina protocols for PE 2x300 bp, and resulted in more than 21 million reads and more than 16.7 million reads for the prokaryotic and archaea specific sequencing, respectively.

Microbial Diversity Analyses

Reads were processed and assigned to amplicon sequence variants (ASVs) using the DADA2 software (v1.12.1) in R (v3.6.0). Briefly, reads were trimmed and quality-filtered using the standard parameters - maximum expected errors for forward and reverse reads = 2, quality score = 2, and trimming length = 273 and 170 base pairs for forward and reverse reads, respectively. Forward and reverse reads were merged with a 20 base pair overlap, ASVs were generated, and chimeras were discarded. ASV annotation was performed using the Silva 16S rRNA database (v132). Assignment of ASVs to a main type of methane metabolism (hydrogenotrophic CO_2-reducing, hydrogenotrophic CH_3-reducing, acetoclastic and methylotrophic (methyl-dismutation)), was done based on their taxonomic affiliation, since all members of almost all methanogen genera/families have the same dominant type of methane metabolism (Table S4). *Methanosarcina* is the main exception, as species from this group can have one or several types of methane metabolisms. All ASVs that were not annotated as archaea were removed from the archaeal-specific primer generated sequences, and ASVs annotated as archaea or bacteria were kept from the universal primer generated sequences. Samples from the same species were merged by summing ASV abundances. These approaches resulted in 1307 archaeal ASVs from the archaea specific primers, as well as 140 archaeal ASVs and 19,145 bacterial ASVs from the universal primers. To estimate the novelty of the archaeal ASVs (obtained with the archaea-specific primers), we compared them using BLAST to 16S rRNA genes of isolated archaea retrieved from the SILVA Living Tree Project LTP database (Quast *et al.*, 2013) plus additional sequences of candidate species belonging to *Methanomassiliicoccales* and Thaumarchaeota. For diversity analyses, rarefaction was performed to normalize sequencing depth to 3,000 reads, leading to 1,253 archaeal ASVs. Bacterial ASVs were normalized to a sequencing depth of 12,000 reads per sample. Observed richness (alpha diversity) was estimated and all beta diversity analyses were performed using the 'phyloseq' package in R (v1.30.0). Subsequent statistical analyses were performed using the base Rstudio 'stats' package (v3.6.0) as well as the R package 'vegan' (v2.5-6). To test for significant differences using the various beta diversity metrics (Table 1) a permutational multivariate analysis of variation (PERMANOVA) from the R package 'vegan' (function adonis) was used. A pairwise Wilcoxon rank sum test with continuity correction from the R package 'stats' (function pairwise.wilcox.test) was used to determine differences between the absolute abundance of archaea and bacteria in animal diet types, as well as between animal classes. Linear regressions from the R package 'stats' (function lm) were used to determine

the relationships between the abundance (log-transformed) of methanogens, Thaumarchaeota and bacteria, and mean retention time (MRT) and dietary fibre consumption. Significance cut-off was $p < 0.05$ for all analyses. Type I errors were corrected using the Benjamini–Hochberg (BH) approach in all pairwise comparisons.

Placement of ASVs within Reference 16S rRNA gene trees

All archaeal ASVs were filtered on a per sample basis, to keep only ASVs representing at least 1% of the total number of reads of the sample. Reference sequences > 1200 bp with a quality >95% were obtained from the Silva SSU 138 database (Yilmaz *et al.*, 2014), RDP database (Cole *et al.*, 2014), and an in-house dataset. Redundancy was removed from reference sequences with a 98% or 97% sequence identity threshold using the VSEARCH software (Rognes *et al.*, 2016). For each archaeal order, long reference sequences were combined with the ASV sequences and were aligned using the G-IN-SI algorithm in MAFFT (Katoh *et al.*, 2013). Phylogenetic trees were generated using the GTR+G4+I model in the IQTREE software (Nguyen *et al.*, 2015). The distribution of ASVs host orders were mapped using ITOL (Letunic *et al.*, 2019). Reference sequences were ultimately removed from the tree to only keep the ASVs sequences.

Co-occurrence of Archaea and Bacteria

To identify co-occurrence signal between archaea and bacteria across Mammalia, Reptilia, and Aves, we integrated the sequences from both the Universal and Archaea specific 16S rRNA gene amplicon sequencing. Only bacterial reads were selected from the Universal 16S rRNA gene amplicon sequencing for this analysis. We used VSEARCH (Rognes *et al.*, 2016) to cluster ASVs into OTUs at 97% in order to reduce the size of the dataset and to filter out truly low abundance lineages of microbes. Then, to merge these datasets in a way that accurately represented the microbial community in terms of relative abundance between archaea and bacteria, we normalized the two datasets both in terms of sequence depth and in terms of archaea-bacterial ratios -information which was gathered through qPCR data. OTUs that were present in less than 10% of the animal classes – Mammalia, Aves, and Reptilia independently- were removed. Following this, we implemented both the SPIEC-EASI (Spiec.Easi package v1.1.0, (Kurtz *et al.*, 2015)) and the SparCC algorithms (Friedman *et al.*, 2012) (part of the Spiec.Easi package (v1.1.0)) in Rstudio (v3.6.0) to determine co-occurrence trends between

archaea and bacteria. Networks were calculated with 1000 iterations. The output from these analyses were filtered using a 0.5 minimum threshold of edge stability (SPIEC-EASI) (Table S3) and a *p*-value $\leq$ 0.05 (SparCC), independently. Only the co-occurrence patterns identified by both algorithms were further analysed.

Investigation of archaea distribution in the gut and other environment

All archaeal 16S rRNA gene sequences from Silva database (Quast *et al.*, 2013) longer than 800 bp and with more than 80% sequence quality, alignment quality and pintail quality were downloaded. Sequences from metagenomes were removed because their environmental origin was not clearly indicated. The annotation of each sequence was retrieved from GenBank and used to classify them as "Gut", "Environmental" or "Human built" origin. Sequences from sponge, animal environments (e.g. nest) or polluted sites (e.g. dump) were not included. The relative abundance of each category was mapped on a tree of archaea built with genomic sequences used in Borrel et al., (Borrel *et al.*, 2019) as well as additional DPANN sequences not present in this study.

Gibbs free energies of methanogenic pathways

The following chemical reactions were considered for methanogenic catabolisms:

CH_3-R dismutation:

4/3 methanol → CH_4 + 1/3 CO_2 + 2/3 H_2O

CH_3-R + CO_2:

methanol + H_2 → CH_4 + H_2O

CO_2 + H_2:

CO_2 + 4 H_2 → CH_4 + 2 H_2O

For each catabolism Gibbs free energy (ΔG_{cat}) calculations were performed using the R package CHNOSZ(Dick, 2019) considering C(methanol) between 10^{-3} and 10^{-7} mol/l, p(H2) between 1 and 10^{-7} bar, T = 298 K, pH = 7 and $p(CO_2) = p(CH_4) = 10^{-1}$ bar.

Origin of the metadata

Animal metadata were collected from various literature sources and online databases. Diet information for mammals and birds were downloaded from the EltonTraits database (Wilman *et al.*, 2014), and information for other animal diets were annotated using the Animal Diversity

Web database (Museum of Zoology, University of Michigan, https://animaldiversity.org/). Information on body weight were also gathered on this website. Information about coefficients of gut differentiation, pH, diet fibre content, and intestinal tract structure and mean retention time were gather from (Chivers *et al.*, 1980; Stevens *et al.*, 1998; Steuer *et al.*, 2011; Beasley *et al.*, 2015; Lin *et al.*, 2015; Langer, 2017; Abraham *et al.*, 2020). Information on methanol, H_2 CO_2 and CH_4 in the gut and in marine sediments originate from (Steggerda, 1968; Czerkawski *et al.*, 1971; Lloyd *et al.*, 1989; Sprenger *et al.*, 2007; Ametaj *et al.*, 2010; Morgavi *et al.*, 2012; Zhuang *et al.*, 2019, 2016; Borrel *et al.*, 2017; Kalantar-Zadeh *et al.*, 2018; Xu *et al.*, 2021).

Data Availability

Data have been deposited in GenBank under the bioproject PRJNAXXXX.

Acknowledgments

C.T. is supported by a PhD fellowship from Paris Sorbonne Université Science and by funds from the doctoral school Bio Sorbonne Paris Cité “BioSPC”. G.B. acknowledges funding from the French National Agency for Research Grant Methevol (ANR-19-CE02-0005-01) and S.G. acknowledges Archaevol (ANR-16-CE02-0005-01). This study has received funding from the French Government's Investissement d'Avenir program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant n°ANR-10-LABX-62-IBEID). We thank the computational and storage services (TARS cluster) provided by the IT department at Institut Pasteur, Paris. We also thank all people who generously collected and provided samples (or made it possible to do so) for this study: Jérôme Fuchs from the Museum National d’Histoire Naturelle (MNHN), Frédéric Delsuc (Institut des Sciences de l'Evolution, UMR5554), Alexis Lécu and Olivier Marquis from Parc Zoologique de Paris, Baptiste Mulot and Hanae Pouillevet from Zoo de Beauval, Paola Dvihally and Dominique Gitton from Parc des Mammelles, Antoine Talarmin, Yann Reynaud, Stéphanie Guyomard, Matthieu Pot, Gaëlle Gruel, Séverine Ferdinand from Institut Pasteur de Guadeloupe, Thomas Godoc from Aquarium de Guadeloupe, Patrick Buisson and Sabine Collin at Palais de la Découverte, Jérôme Contignac and David Luis Garcias Warner from Nabau Projects/Grupo Atrox, Aude Bourgeois from Jardin des Plantes de Paris, Jérémy Sauvanet, Gaspard, Anaïs Tibi and Johnatan Aparicio. We thank

Christa Schleper for providing a plasmid with a Thaumarchaeota 16S rRNA gene for qPCR standard, Jérôme Mathieu for data on gastrointestinal tract types and Marcus Clauss for helping us collecting the MRT data from an abundant literature.

Author contributions

C.M.T., G.B. and S.G. conceived the study. C.M.T. did the experiments and E.D. the thermodynamic calculations. C.M.T. and G.B. analyzed the data and wrote the manuscript with the input of S.G.

Competing interests

The authors declare no competing interests.

Chapter 1.A: Supplementary Data

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,*]

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 faeces) 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 Thaumarchaeota 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 (≥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 ≥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" [Sup](Borrel, O'Toole, *et al.*, 2013; Lang *et al.*, 2015; Söllinger *et al.*, 2016), 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 faeces 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 [Sup](Raymann *et al.*, 2017) (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 [Sup](Hackstein *et al.*, 1994; Leadbetter *et al.*, 1996, 1998; Sprenger *et al.*, 2000, 2005, 2007; Brune *et al.*, 2015), but recently was shown to form two distinct clades -one of sequences obtained from insects and another from mammals [Sup](Thomas *et al.*, 2021). 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 [Sup](Thomas *et al.*, 2021) (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 [Sup](Ormerod *et al.*, 2016; Lagkouvardos *et al.*, 2019), 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 faeces [Sup](Kageyama *et al.*, 2000). 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 faeces), 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 [Sup](Ruaud *et al.*, 2020). 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 [Sup](Biddle *et al.*, 2013).

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

In Aves, no archaea-bacteria relationships shared between the SparCC and SPIEC-EASI approaches (Table S7). However, the SparCC approach, identified one significant archaea-bacteria relationship between a *Methanosphaera* OTU and a Clostridiales (Clostridium *sensu stricto* 1); and the SPIEC-EASI approach identified 57 relationships between archaea and bacteria. Interestingly this approach identified that *M. smithii* (OTUarc_1), is negatively correlated to several bacterial OTUs, including common gut bacterial groups such as Clostridia and *Arthrobacter* – a species previously characterized as being part of the goose core-gut microbiome [Sup](Wang *et al.*, 2016). 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 using both approaches. 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 [Sup](Manos *et al.*, 2006). 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.

Chapter 1.A: Supplementary Figures

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

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 donnor (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.

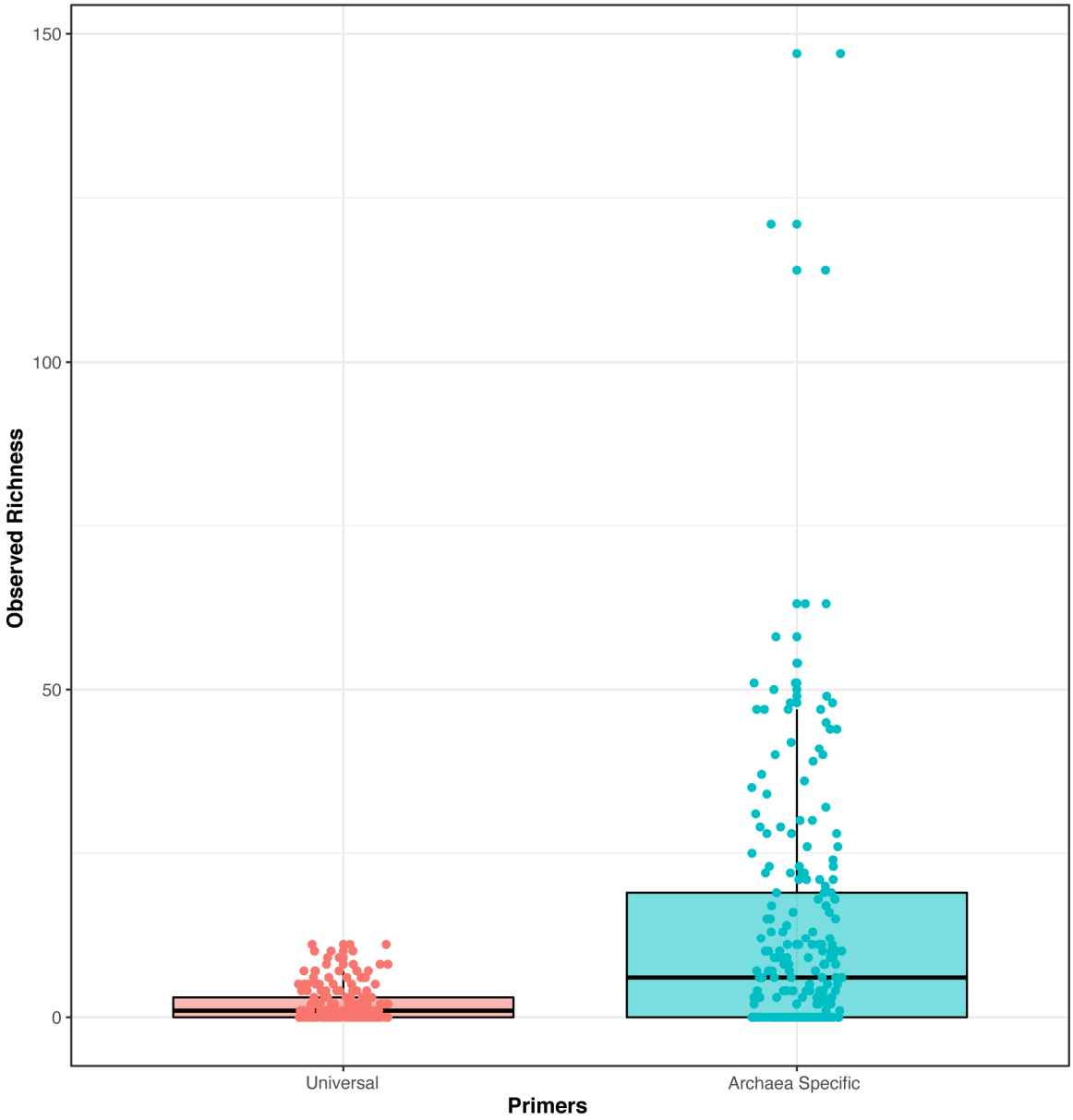


Figure S1: Comparison of the observed archaeal richness captured using universal and Archaea-specific 16S rRNA gene primer.

Figure S2.

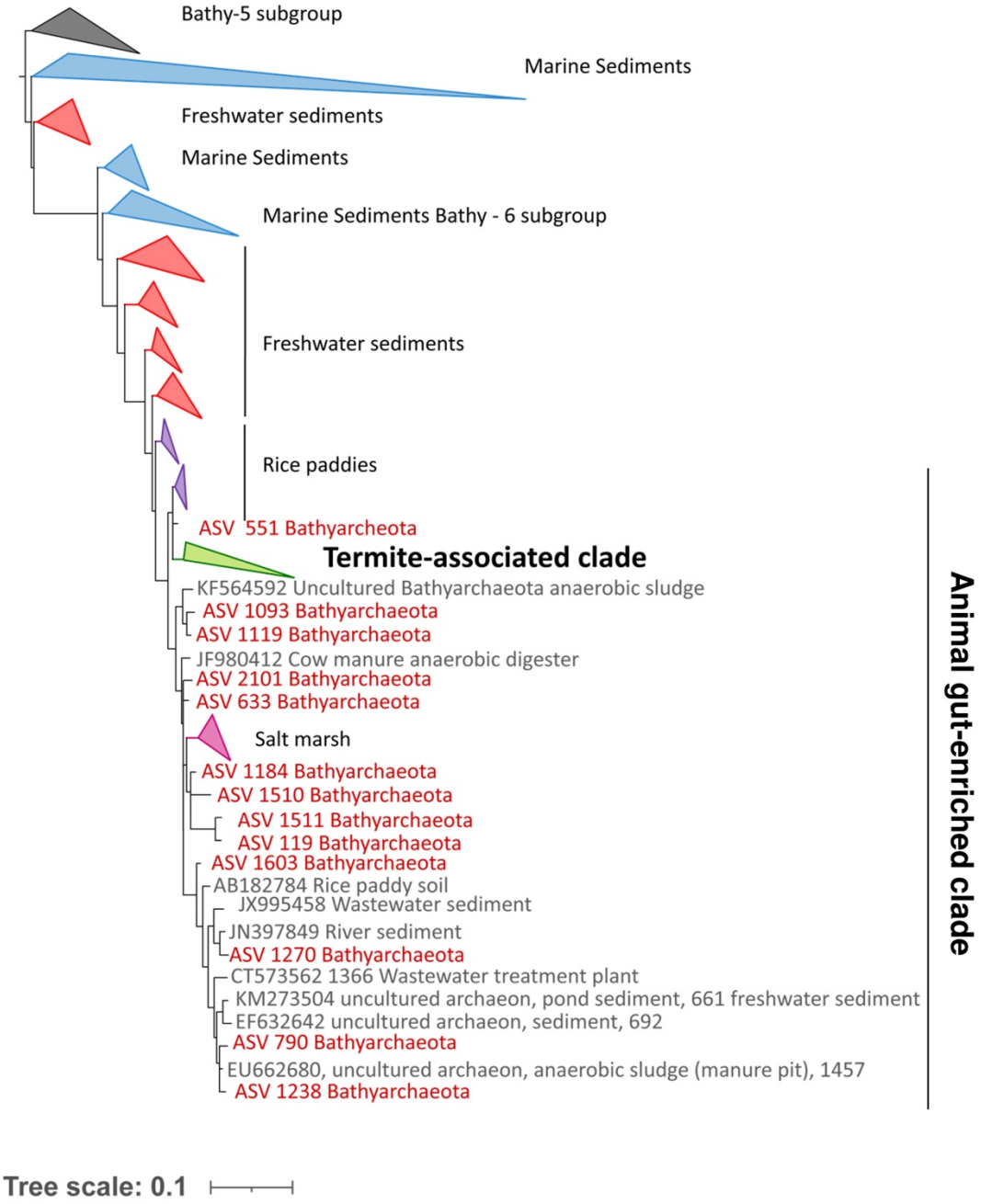


Figure S2: Close relationship between Bathyarchaeota sequences recovered from the animal gut. Sequences from this study are in red. They correspond to ASVs representing at least 1% of the gut archaeome of one animal species. Maximum-likelihood tree based on 16S rRNA gene sequences and built with GTR+G4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

Figure S3.

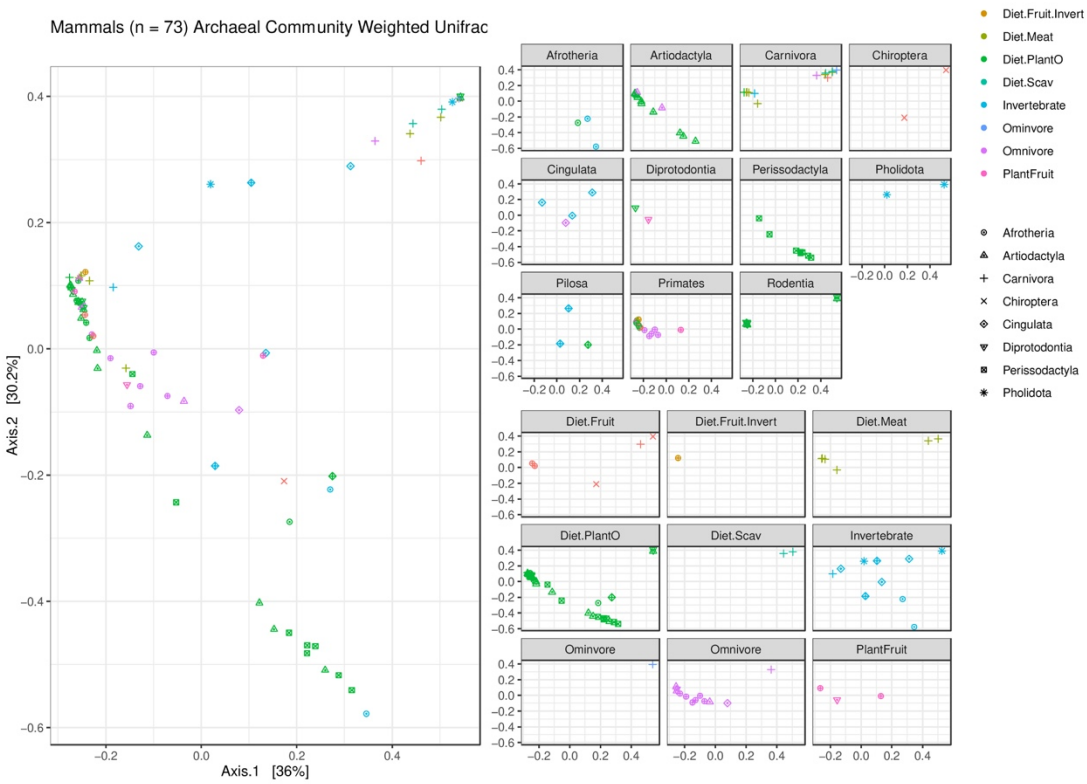


Figure S3: Weighted UniFrac analysis of the gut archaeome of Mammals, based on 73 rarefied samples with ≥3000 reads.

Figure S4

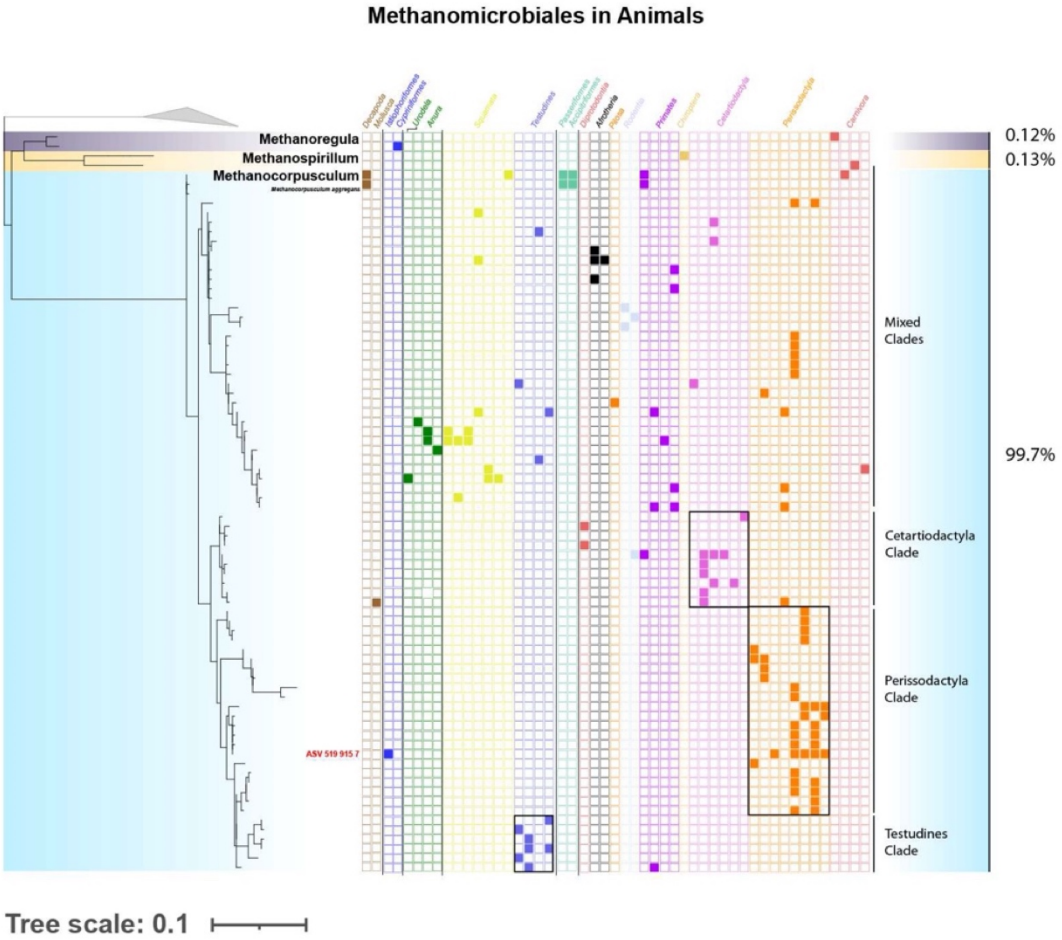


Figure S4: Distribution of Methanomicrobiales ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was built with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were then removed 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.

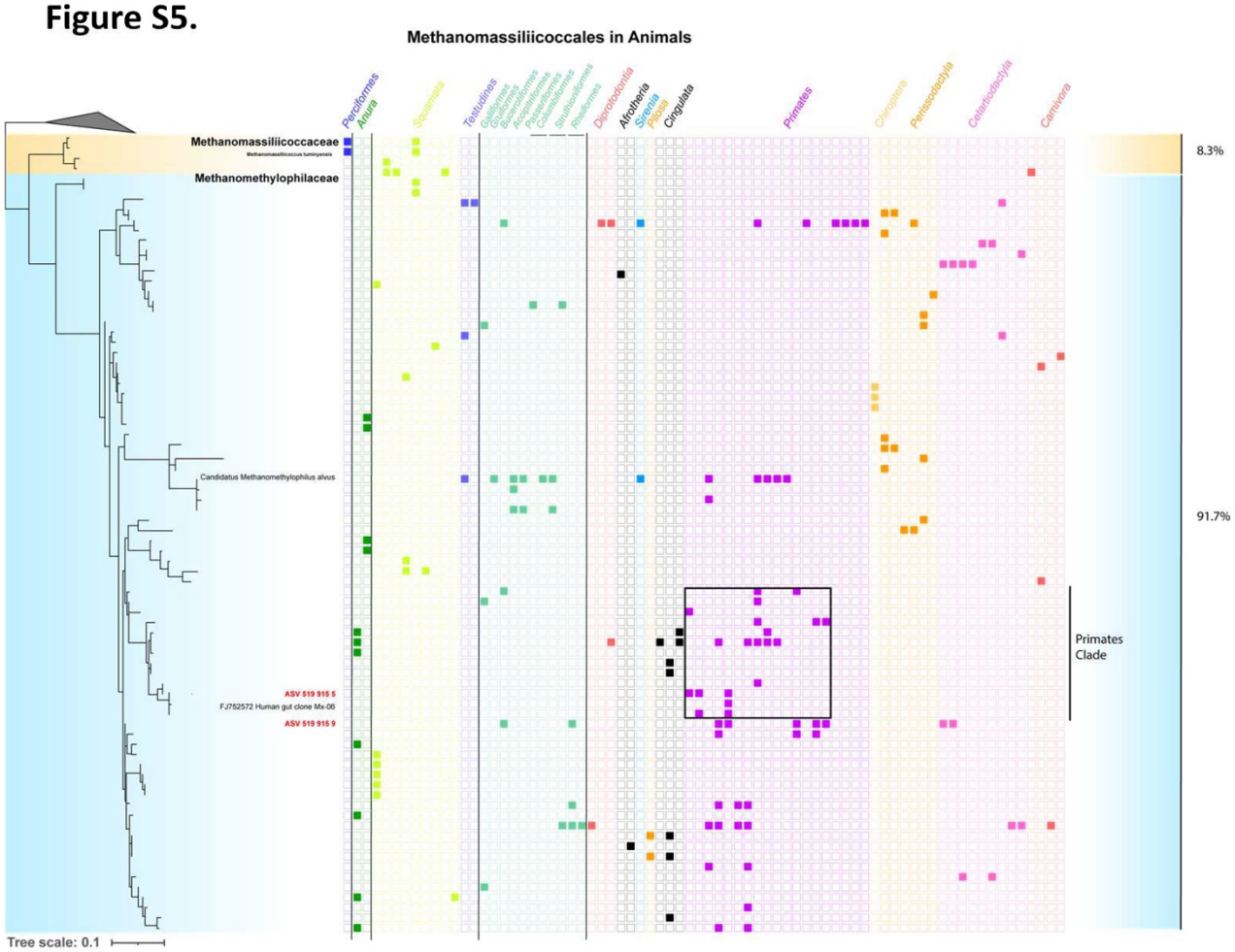


Figure S5: Distribution of Methanomassiliicoccales ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was constructed with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were 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.

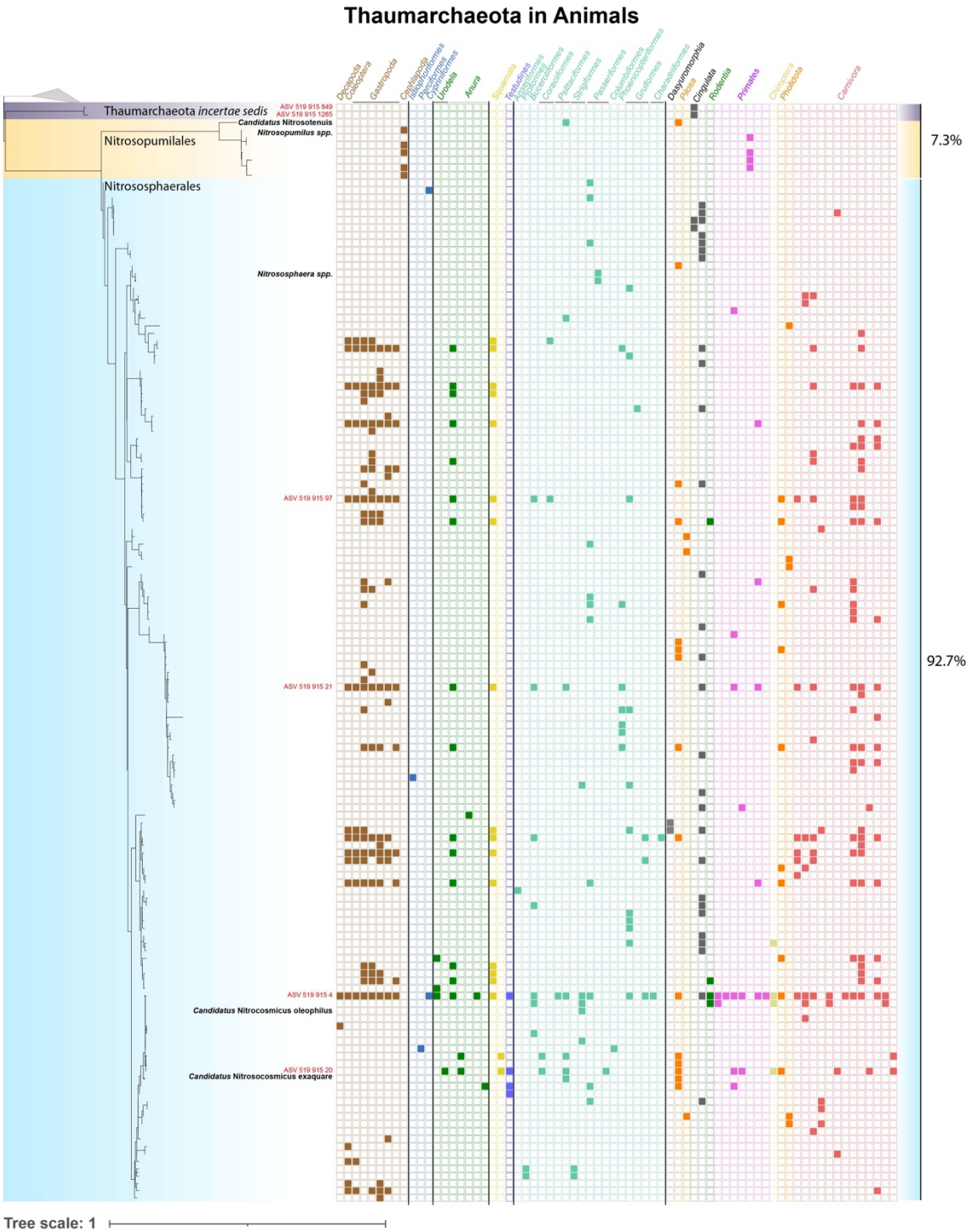


Figure S6: Distribution of Thaumarchaeota ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was constructed with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were then removed 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.

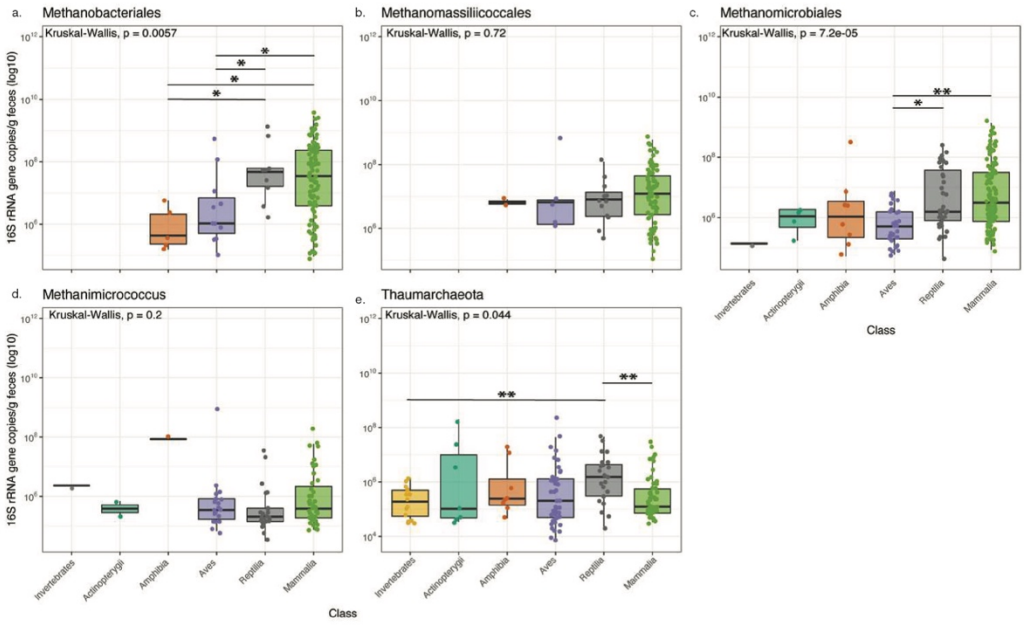


Figure S7: Absolute abundance of five archaeal lineages determined via qPCR. Significant differences across all groups were determined via the Kruskal-Wallis test, $p < 0.05$ is significant. a) Methanobacteriales (n = 121); b) Methanomassiliicoccales (n = 99); c) Methanomicrobiales (n = 204); d) Methanimicrococcus (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.

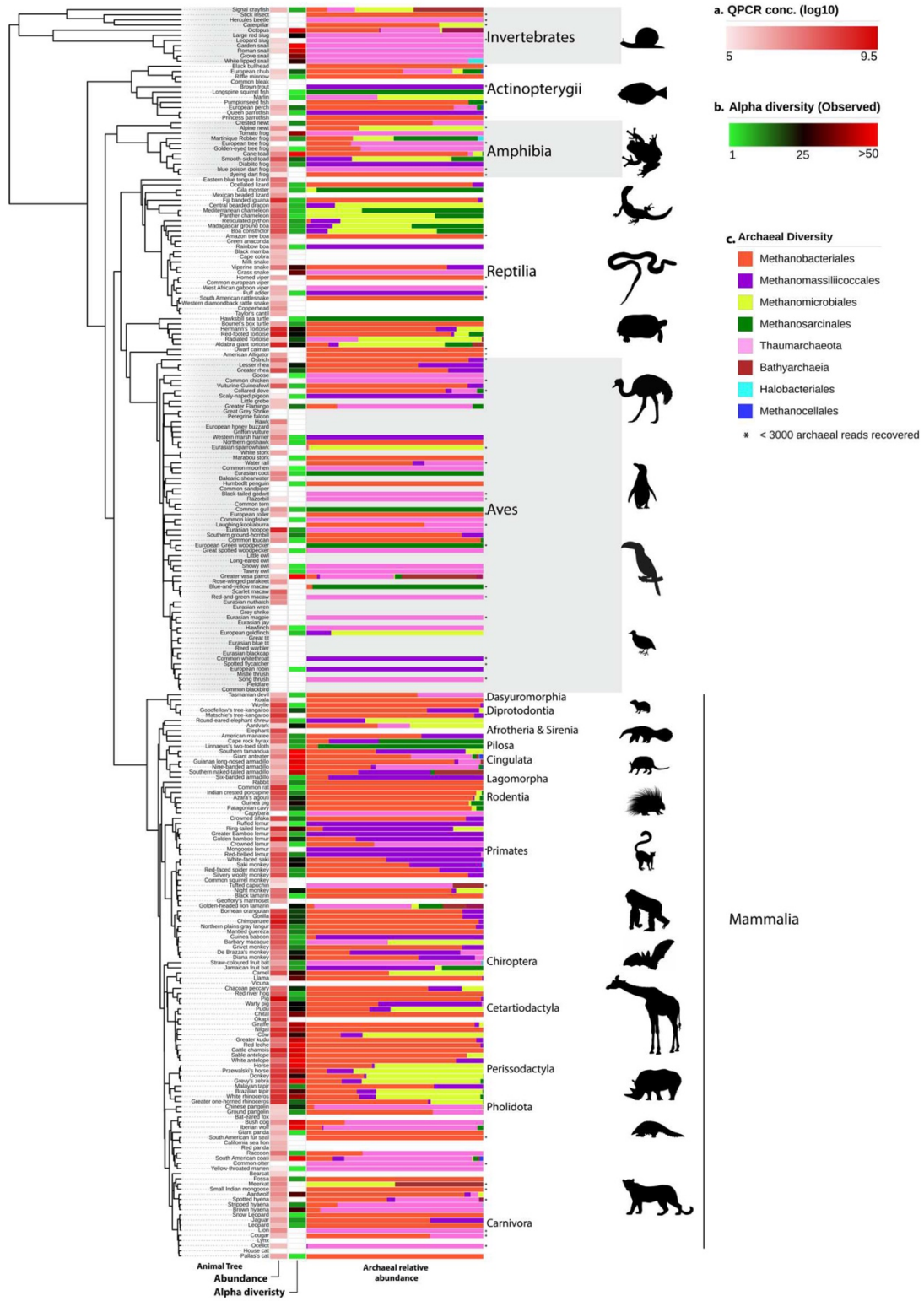


Figure S8: Archaeal diversity across animals including species for which a low number (< 3000 reads) of reads or no reads were obtained. Animal species with less 5 to 3000 reads are indicatde by a *.

Figure S9.

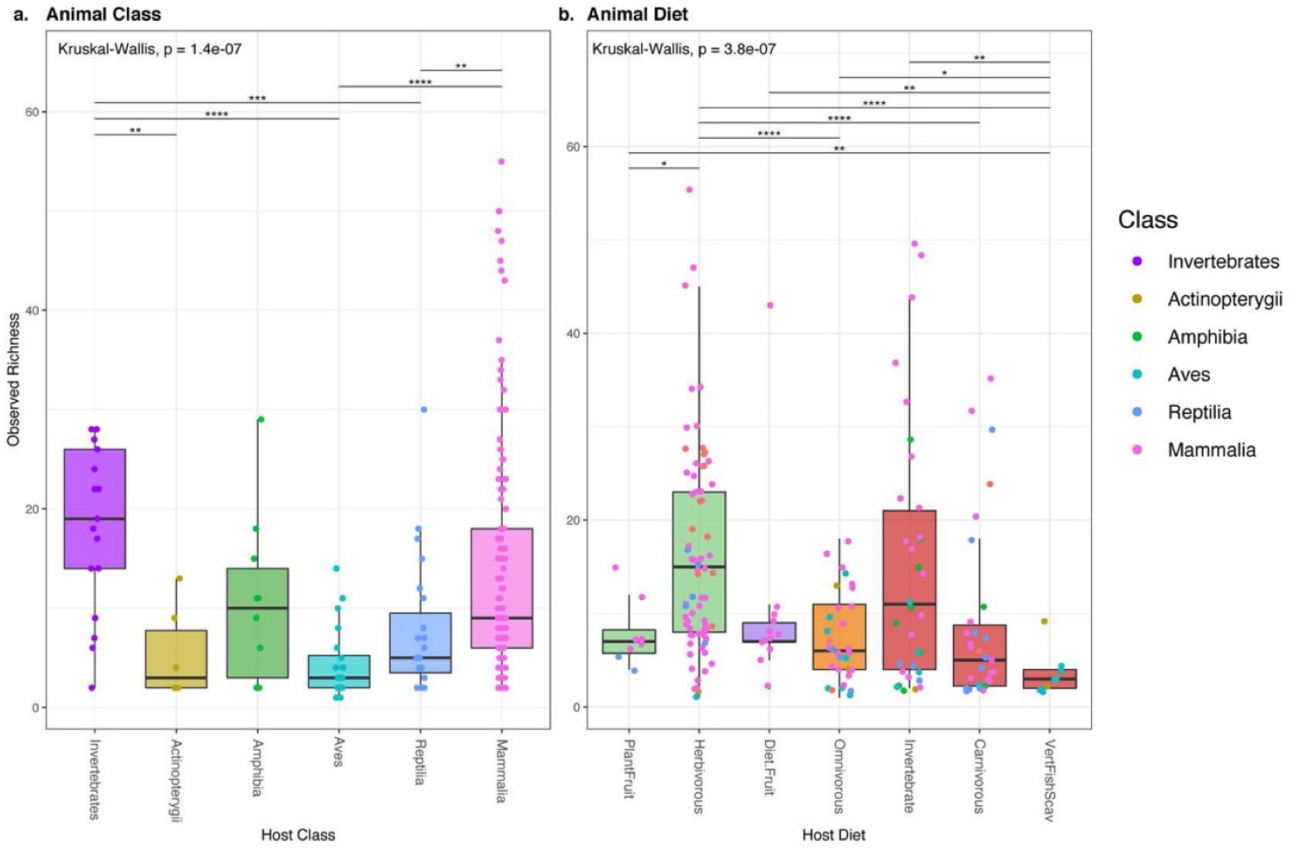


Figure S9: a) Observed richness of Archaea between animal classes (n = 218). b) Observed richness of Archaea between animal diets (n = 197), diet types with fewer than 3 representatives were removed. Wilcoxon rank sum test *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$.

Figure S10.

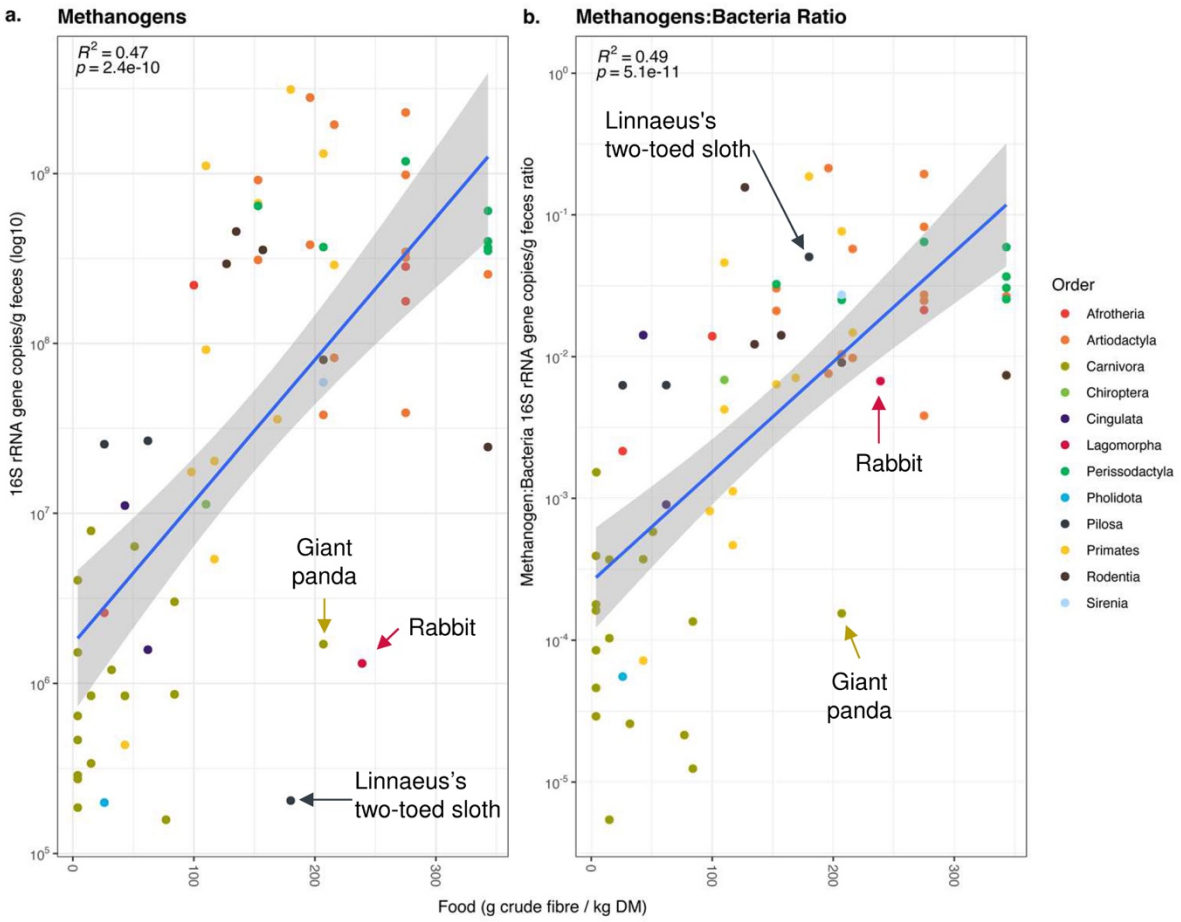


Figure S10: Correlation between diet-fibre content and a) absolute abundance and b) relative abundance of methanogens in mammals faeces (n = 110). The abundance methanogen is determined by the sum of the abundance of abundance of Methanobacteriales, Methanomassiliicoccales, Methanomicrobiales, and *Methanimicrococcus*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

Figure S11.

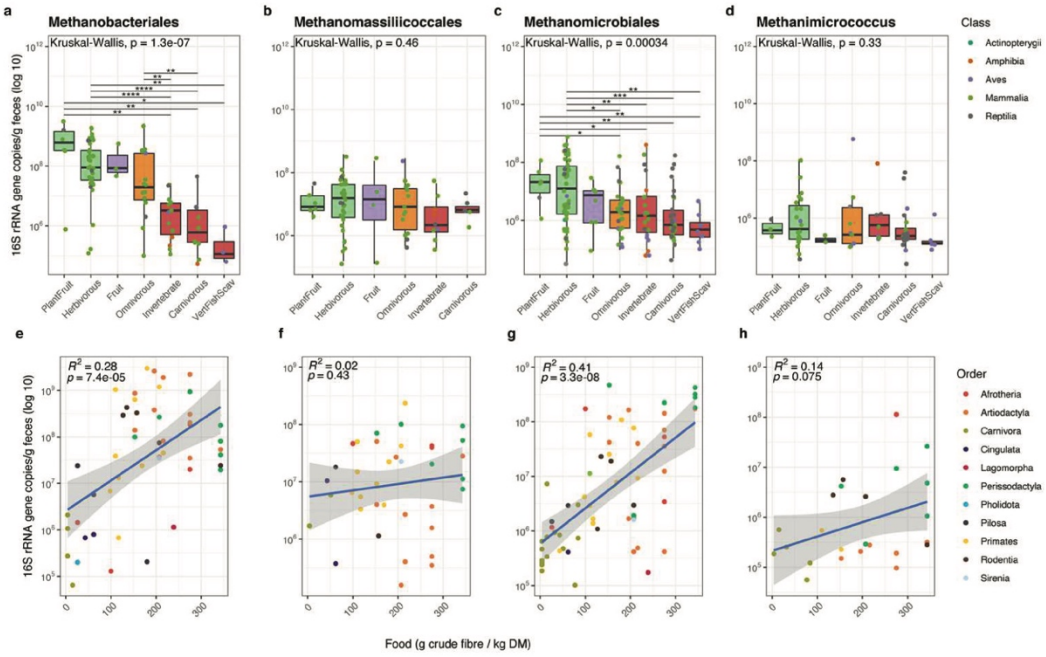


Figure S11: Influence of host diet-type and diet-fibre content on the absolute abundance of four methanogen lineages. Abundance of **a**) Methanobacteriales (n = 91), **b**) Methanomassiliicoccales (n = 73), **c**) Methanomicrobiales (n = 148) and **d**) Methanimicrococcus (n = 69) according to host 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.

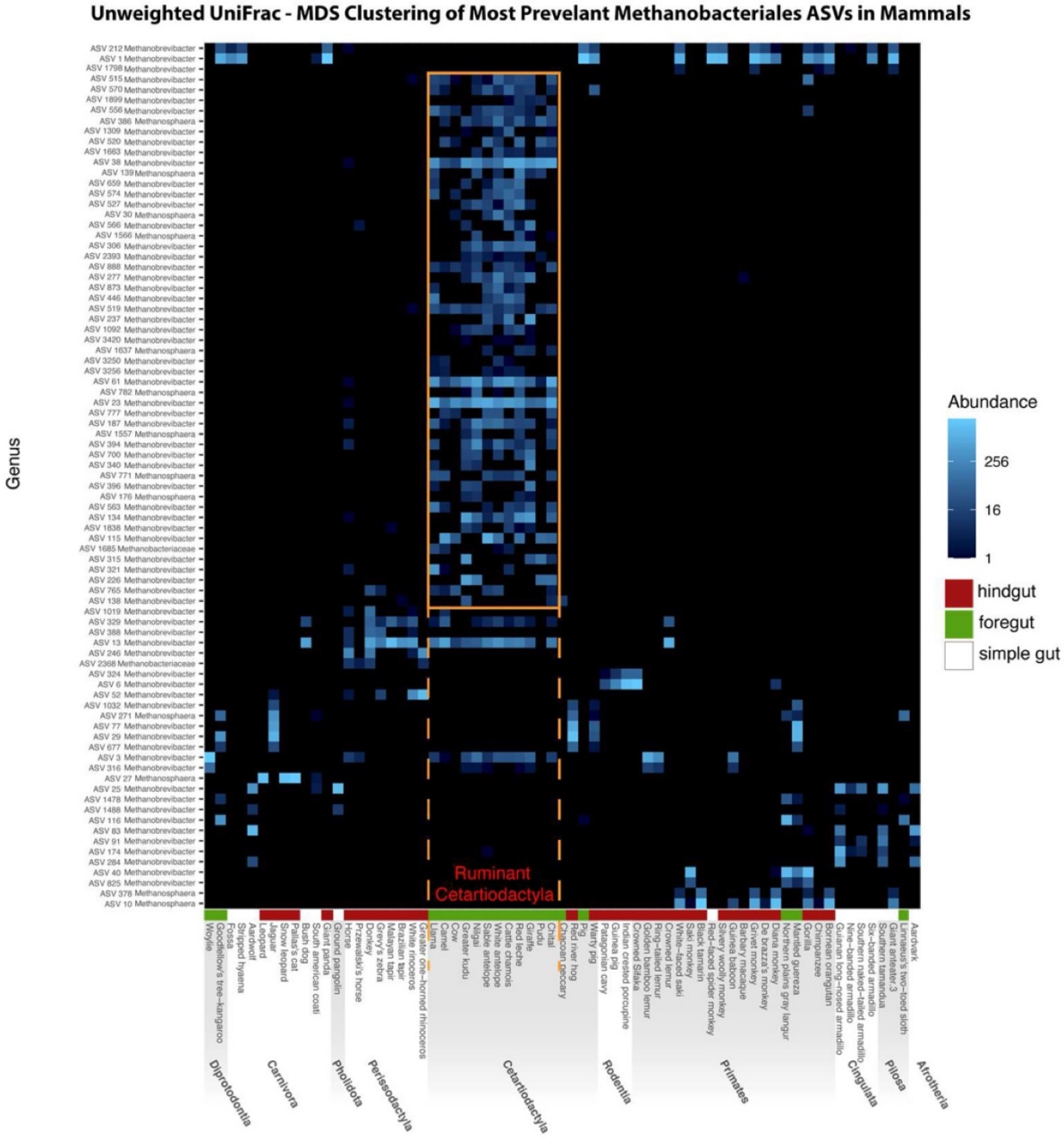


Figure S12: Unweighted UniFrac – MDS clustering of most prevelant Methanobacteriales ASVs in mammals. Methanobacteriales ASVs that are found in ≥ 5% of mammals (n = 67; rarefied 3000 reads/sample). Animals are organized according to mammal phylogeny, and ASVs on y-axis are clustered by the unweighted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

Figure S13.

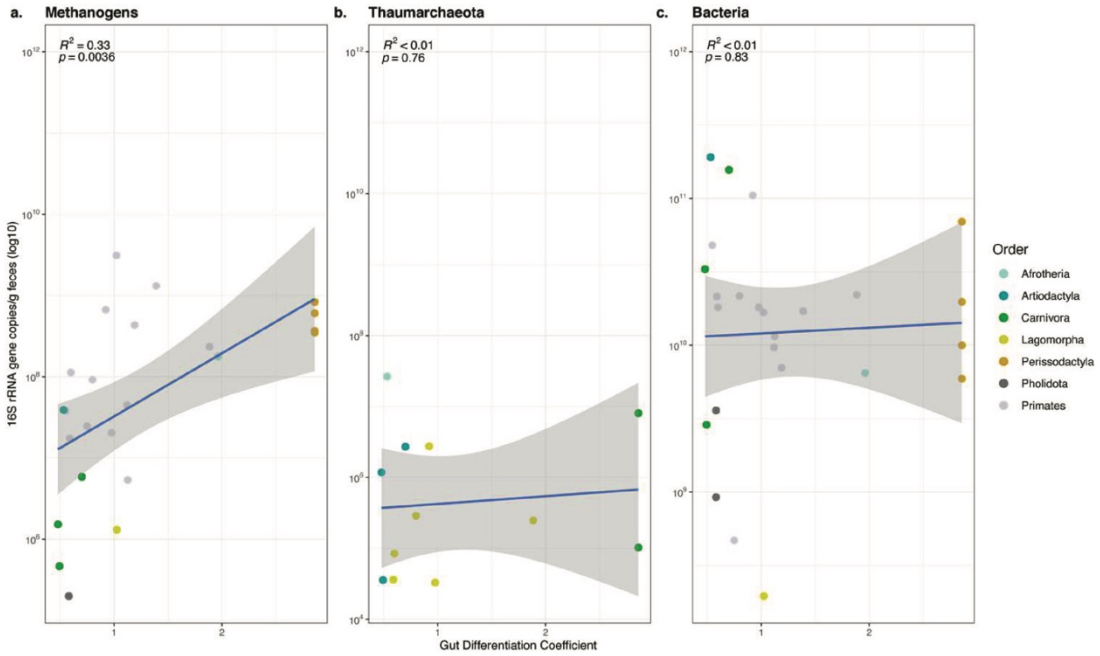


Figure S13: Correlation between gut differentiation coefficient and the absolute abundance of a) methanogens (n = 25), b) Thaumarchaeota (n = 12) and c) bacteria (n = 25) (c) in mammals. The coefficient of gut differentiation corresponds to the surface area of the stomach, caecum, and colon, relative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

Figure S14.

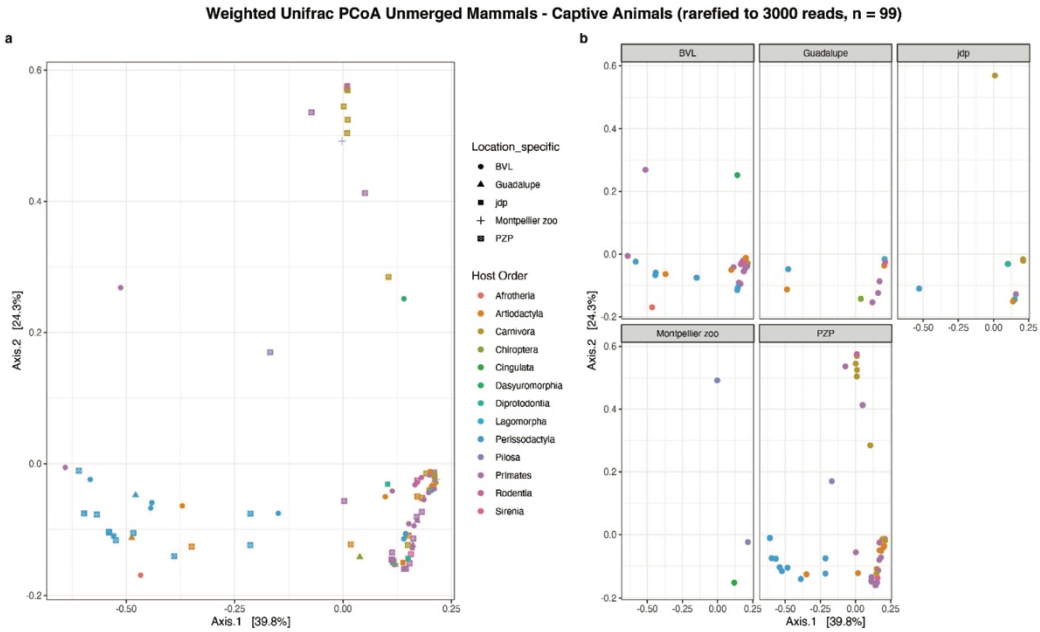


Figure S14: Weighted Unifrac analysis of individual Captive Mammals. Mammals with ≥3000 reads per sample (n = 99) were analyzed using a Weighted Unifrac analysis, location did not significantly explain the variance between the archaeal community in captive mammals ($p > 0.05$).

Figure S15.

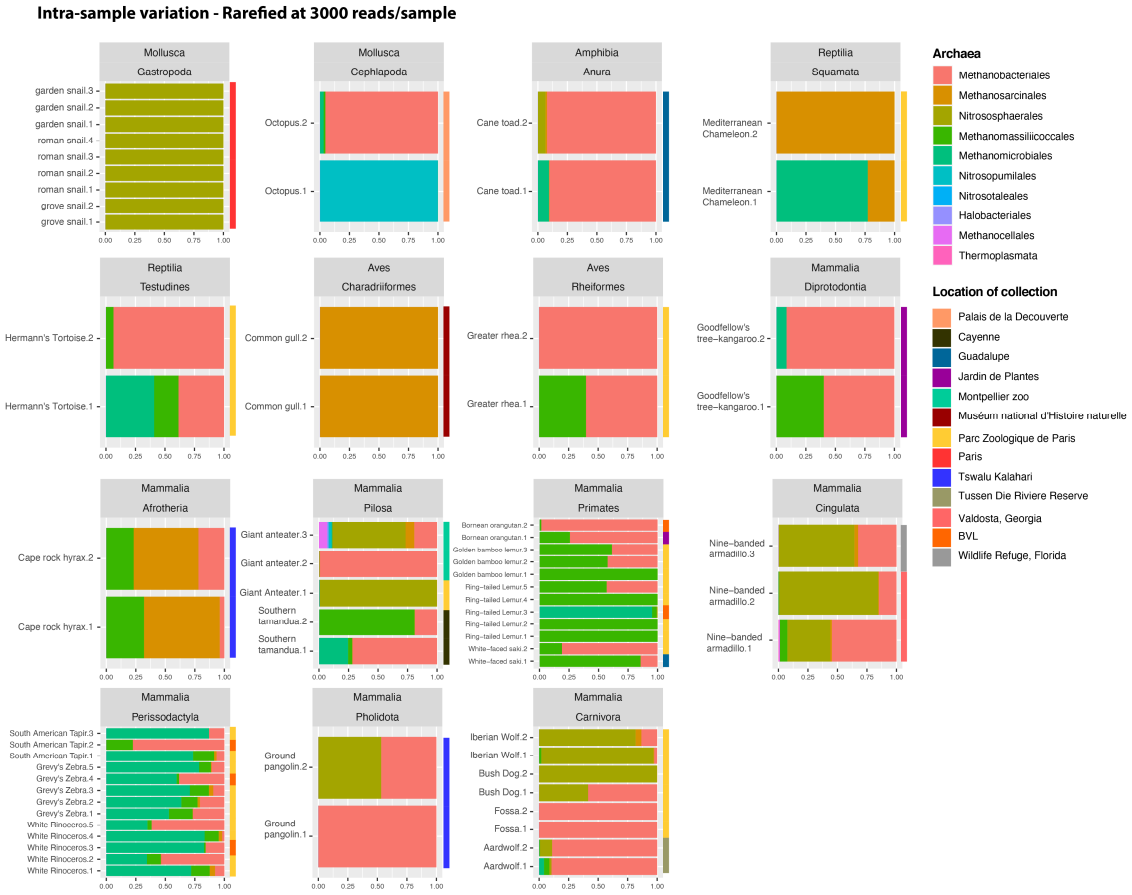


Figure S15: Intra-species variation rarefied at 3000 reads per sample. Animals for which multiple fecal samples were used the variation in relative abundance of archaeal ASVs between samples is low, even between samples from varying locations.

Figure S16.

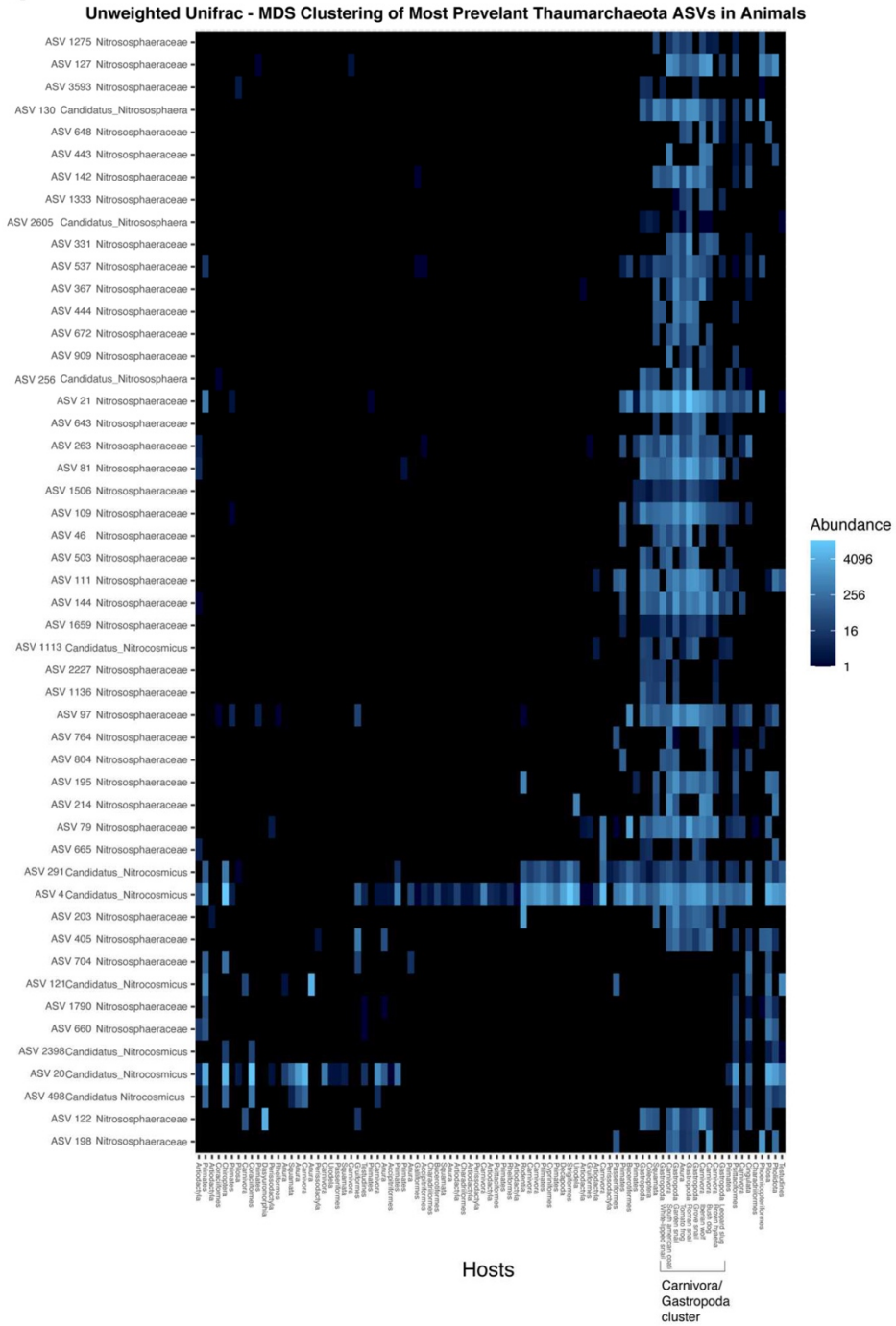


Figure S16: Unweighted UniFrac – MDS Thaumarchaeota of most prevelant Thaumarchaeota ASVs in all animals. Thaumarchaeota ASVs that are found in ≥ 5% of samples (n = 89). Animals and ASVs are clustered by the unweighted UnifFrac dissimilarity measure. Carnivora and Gastropoda host similar Thaumarchaeota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

Figure S17.

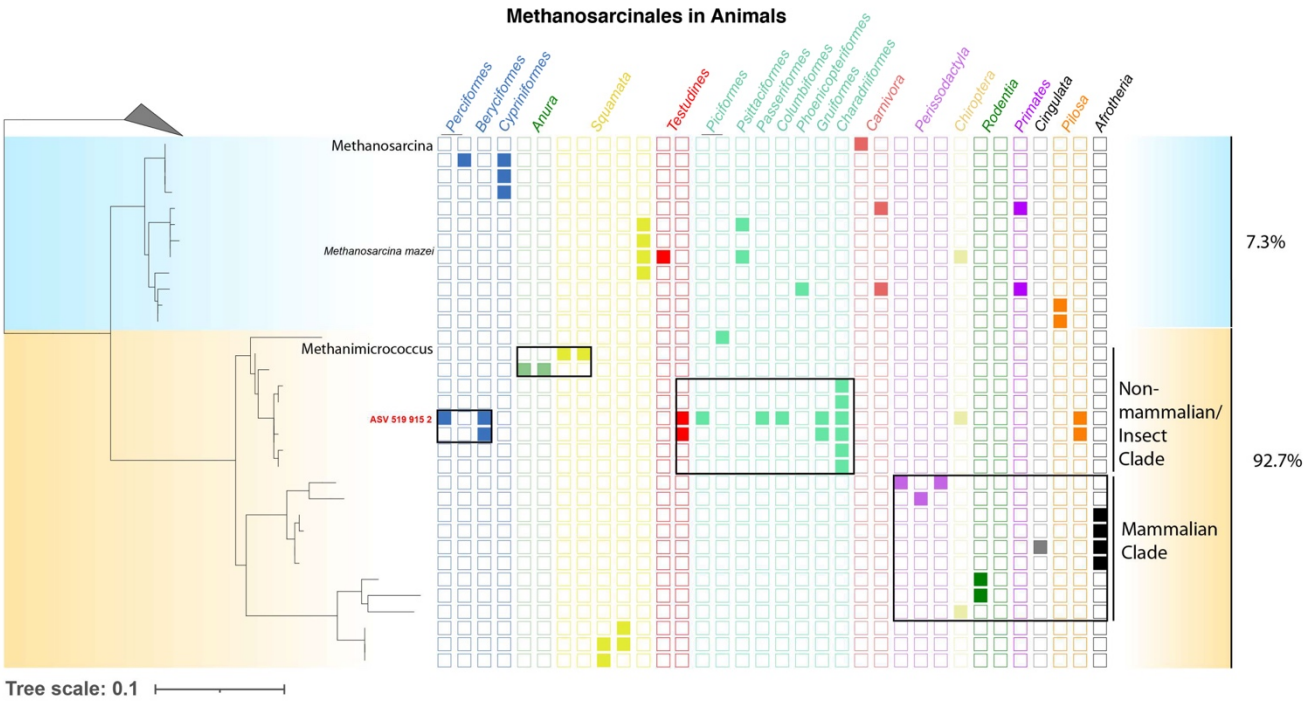


Figure S17: Distribution of Methanosarcinales ASVs among animals. The phylogenetic tree (maximum-likelihood, GTR+G4) was built with nearly full length 16S rRNA genes sequences from literature and the ASVs sequences from this study. For clarity, the full 16S rRNA genes from literature were then removed 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.

Figure S18.

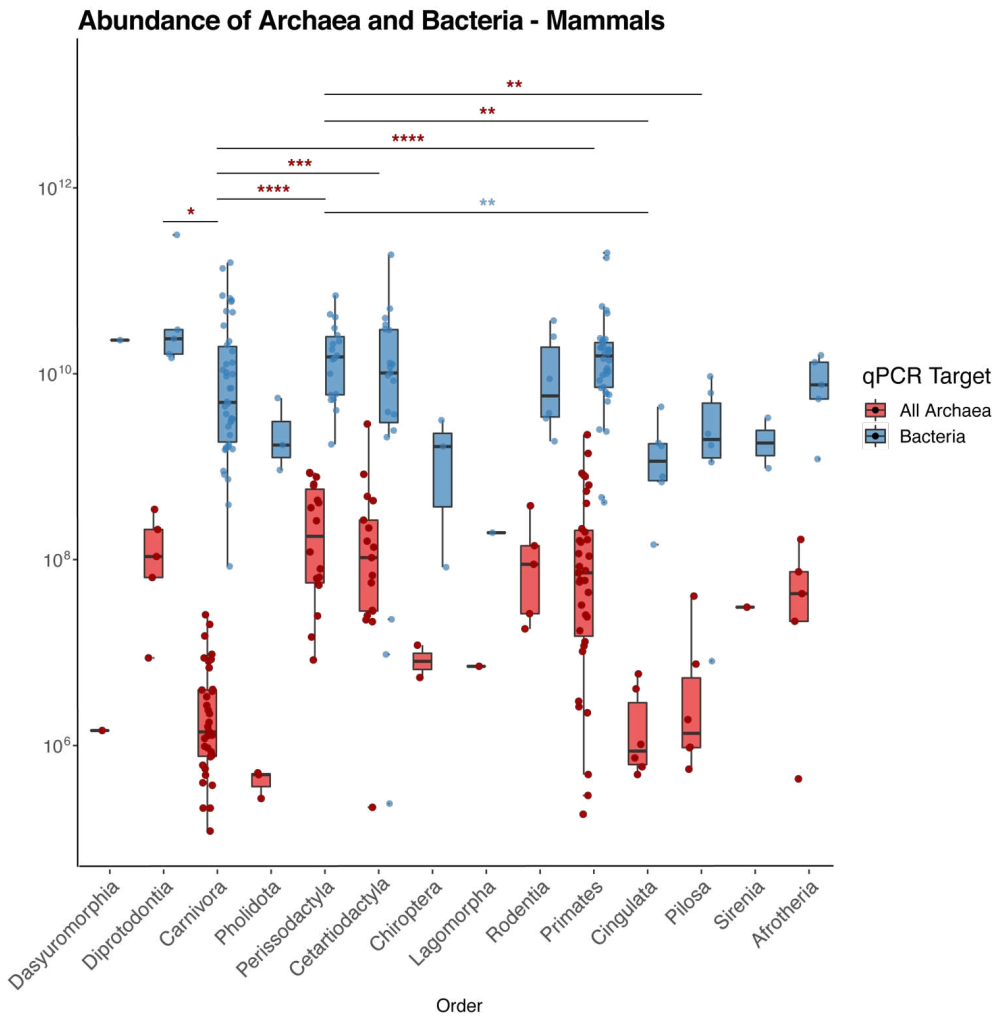


Figure S18: Absolute Abundance of Archaea (red) and Bacteria (blue) determined via qPCR mammalian orders (n = 156). Animal lineages with significantly different archaeal/bacterial abundances are labeled. Wilcoxon rank sum *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$.

Chapter 1.B: Analysis of publicly available intestinal microbiome data and quantification of Sulfate-Reducing Bacteria

The analyses presented in Chapter 1.A. resulted in one of the first insights into the intestinal archaeome across a wide range of animal hosts. This work was largely conducted on fecal samples from captive animals. It was previously determined that captivity impacts the bacterial community of some animals (Ley *et al.*, 2008; McKenzie *et al.*, 2017; Gibson *et al.*, 2019). In contrast, the impacts of captivity on the intestinal archaeal community have yet to be thoroughly examined. In the following section, I discuss a preliminary analysis of publicly available microbiome datasets, which includes samples from various wild species of animals and yielded several interesting results.

Additionally, the methanogens are one of the major groups of hydrogenotrophs in the intestinal tract, and are believed to compete with sulfate reducing bacteria and homoacetogens for this substrate (Dore *et al.*, 2006; Nakamura *et al.*, 2011; Kushkevych *et al.*, 2017). However, the intensity of this competition has yet to be systematically examined in most animals. To begin to understand the nature of the relationship between intestinal hydrogenotrophs, an intern that I mentored in the lab, Tomás Rodríguez, collected quantitative data on the sulfate reducing bacteria (SRB) in 161 of the same fecal samples analyzed in Chapter 1.A. The results of analyses are discussed further below.

I. Mining of publicly available datasets for host-associated Archaea

The results presented in Chapter 1.A clearly illustrate that archaea-specific primers are useful in capturing a higher diversity of intestinal archaea, especially when archaea are present at low concentrations. However, universal primers were capable of identifying archaea for

samples with the highest concentrations. Indeed, I managed to detect a variety of archaeal 16S rRNA gene sequences from publicly available datasets that did not use archaea specific primers (see **Chapter 1.B Figure 1** and the accompanying legend). Out of more than 100 previously published studies, 57 contained archaeal sequences. These 57 studies were associated to 51 species belonging to four major groups of animals (wild and captive: Mammalia, Aves, Reptilia, Actinopterygii, and Invertebrates) and the trends identified in these works largely support what we found in our analyses. For example, the Methanomicrobiales dominate the archaeome of the Perissodactyla, the Methanomassiliicoccales are particularly abundant in the lemurs, and the Nitrososphaerales are the dominant archaea in the Chiroptera (**Chapter 1.B Figure 1**). The similarity in taxa identified through this analysis suggests that captivity may not have as large of an impact on the intestinal archaeome as it does on the bacterial community, however a large-scale analysis is needed to determine if this is indeed the case.

This meta-analysis also highlighted unique trends such as the prevalence of the Nitrosopumilales in marine invertebrates and the high relative abundance of Methanomassiliicoccales in several whales (**Chapter 1.B Figure 1**). Thus, a thorough investigation of marine animals would greatly increase the overall view of intestinal archaea throughout the animal kingdom, possibly revealing novel evolutionary trends of intestinal/host colonization of archaea in the marine environment.

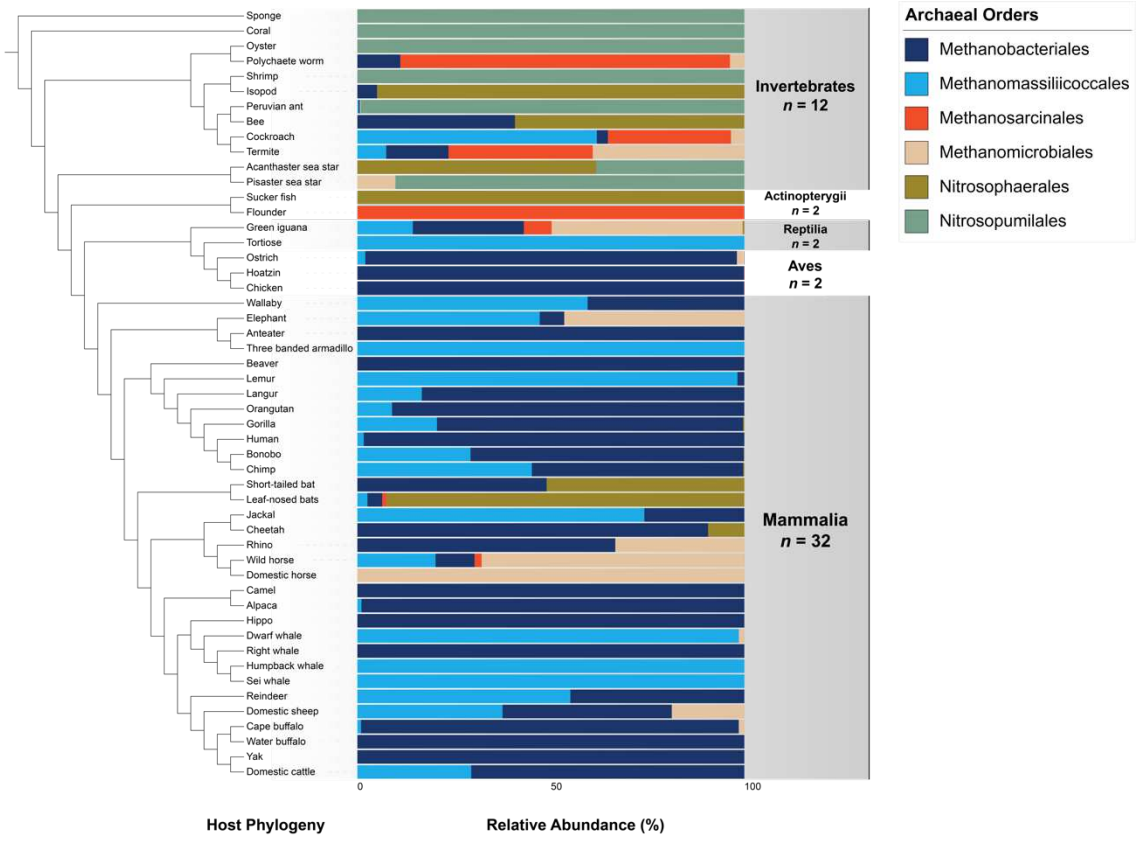


Figure 1. Composition of the archaeal community in animal intestinal microbiomes available in public databanks gathered in 2020. 57 studies containing archaeal sequences that targeted invertebrates, fish, birds, and mammals were kept for analysis. Sequences of the microbiome of a few animals without a digestive tract (e.g., sponge) were also included. All sequences were run through an in-house bioinformatic pipeline to identify and annotate any potential archaeal reads (Volant *et al.*, 2020). Then, all archaeal sequences were aligned to a reference archaeal 16S rRNA gene dataset and placed in a phylogenetic tree to ensure that they were all of archaeal origin. To further confirm the annotations of these reads, we used the classification function available on the Silva ribosomal RNA database (Cole *et al.*, 2014), and reads that represented less than 1% of the total community were removed.

II. Quantification of other hydrogen-consuming microbes in the intestine

Consumption of hydrogen in the intestinal tract is extremely important for the functioning of this ecosystem (Bui *et al.*, 2019). As already mentioned in the Introduction section, the role is largely fulfilled by three functional groups within the microbiota: the sulfate reducing bacteria

(SRB), the methanogens, and the homoacetogens. In non-intestinal environments such as anoxic sediments containing sulfate, the exclusion of CO_2-reducing hydrogenotrophic methanogens and homoacetogenic bacteria by sulfate-reducing bacteria has been well characterized (Lovley *et al.*, 1982, 1988). In the intestine, such relationships has long been debated but it is generally considered that these groups directly compete for hydrogen (Ozuolmez *et al.*, 2015; Smith *et al.*, 2019). In primates (including humans), it was suggested that SRB and methanogens are mutually exclusive (Christl *et al.*, 1992) but more recent studies have refuted this hypothesis (Gibson *et al.*, 1993; Dore *et al.*, 2006; Nakamura *et al.*, 2009).

The quantitative data presented in Chapter 1.A show that high fibre consumption is related to an increase of methanogenic archaea in the gut. We also found that animals with a diet consisting of plants and fruits host a higher diversity of intestinal archaea. This is consistent with what has been previously found in humans, pigs, and ruminants in terms of gene expression levels by methanogens and of methane production (Jensen *et al.*, 1994; Shibata *et al.*, 2010; Tap *et al.*, 2015). In humans, an increase in plant-based fibre results in an increased microbial diversity (Filippo *et al.*, 2010; Cotillard *et al.*, 2013; Sonnenburg *et al.*, 2016) as well as a shift in the overall functional profile of the intestinal microbiota - e.g. higher levels of glycan-degrading carbohydrate active enzymes (CAZymes) predicted to degrade plant cell walls and to increase SCFAs production (Wastyk *et al.*, 2021). This supports the hypothesis that an increased amount of fibre results in an increased concentration of metabolites that methanogens have access to, including hydrogen (Tap *et al.*, 2015; Wastyk *et al.*, 2021). However, the relationship between fibre consumption and SRB within most animals is still unclear.

To better understand the dynamics of the two most abundant hydrogenotrophic groups of microbes in the intestinal tract, a qPCR analysis targeting the functional marker gene for sulfate-reduction (*dsr)* was performed on the same samples used to quantify archaea in the study presented in Chapter 1.A. If there is a significant level of competition for hydrogen in the intestinal tract between these groups, a negative correlation between their abundances would be expected. Further, if the SRB and methanogens are mutually exclusive, they should rarely occur within the same host. Finally, if the main driver of SRB abundance is the availability of H_2, a direct correlation of their abundance with fibre content would also be

observed, as our analyses suggest that elevated fibre consumption is related to increased H_2 production in the intestine.

Across animal classes, the quantification of SRB reveals that they are always less abundant than methanogens, except in the amphibians, where both groups were present in almost equal amounts **(Chapter 1.B Figure 2a)**. Within animal orders, SRB and methanogens are present at similar amounts in the Carnivora, Crocodilia, and Squamata **(Chapter 1.B Figure 2b**). Interestingly, the abundance of SRB is more uniform between different groups of animals than the abundance of archaea. SRB were previously identified to be less abundant than methanogens in the intestinal tract of primates (Amato *et al.*, 2013), but to our knowledge this is the first time this observation has been made in such a large number of animals. These trends are interesting to observe, as in mammals there was a clear increase in methanogens in species that consume large amounts of fibre (e.g., Cetartiodactyla/Perissodactyla) (Chapter 1.A). However, the same patterns do not emerge with the SRB. Indeed, fibre consumption was not significantly related to the abundance of methanogens (**Chapter 1.B Figure 2**). There was also no correlation between the CO_2-reducing hydrogenotrophic methanogens and SRB (data not shown). Taken together, these results suggest that there is no direct competition for hydrogen in the animal intestine between SRB and CO_2-reducing hydrogenotrophic methanogens.

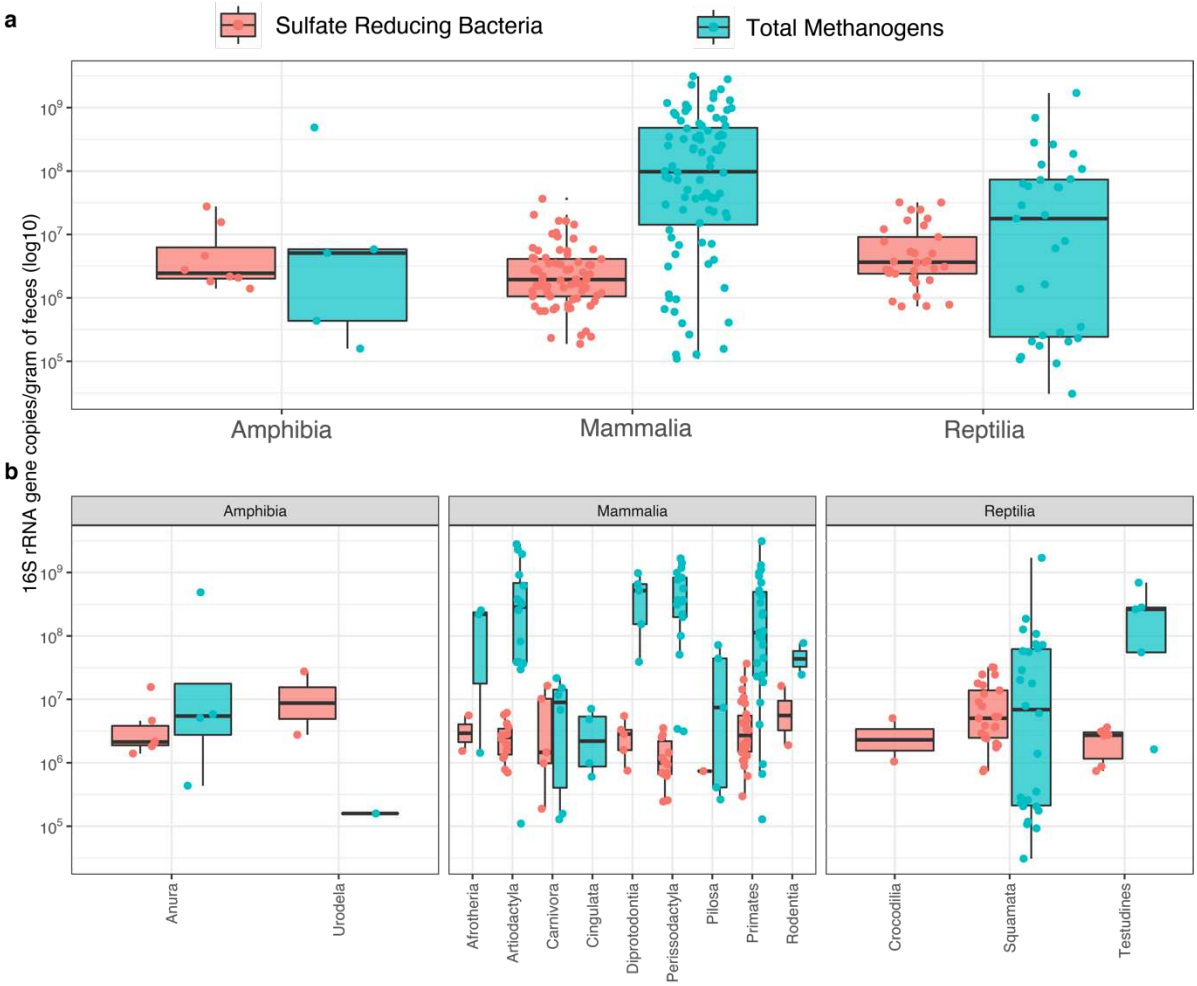


Figure 2. Absolute Abundance of Sulfate Reducing Bacteria and Methanogens within **a)** amphibians, mammals, and reptiles, and **b)** across orders within these classes.

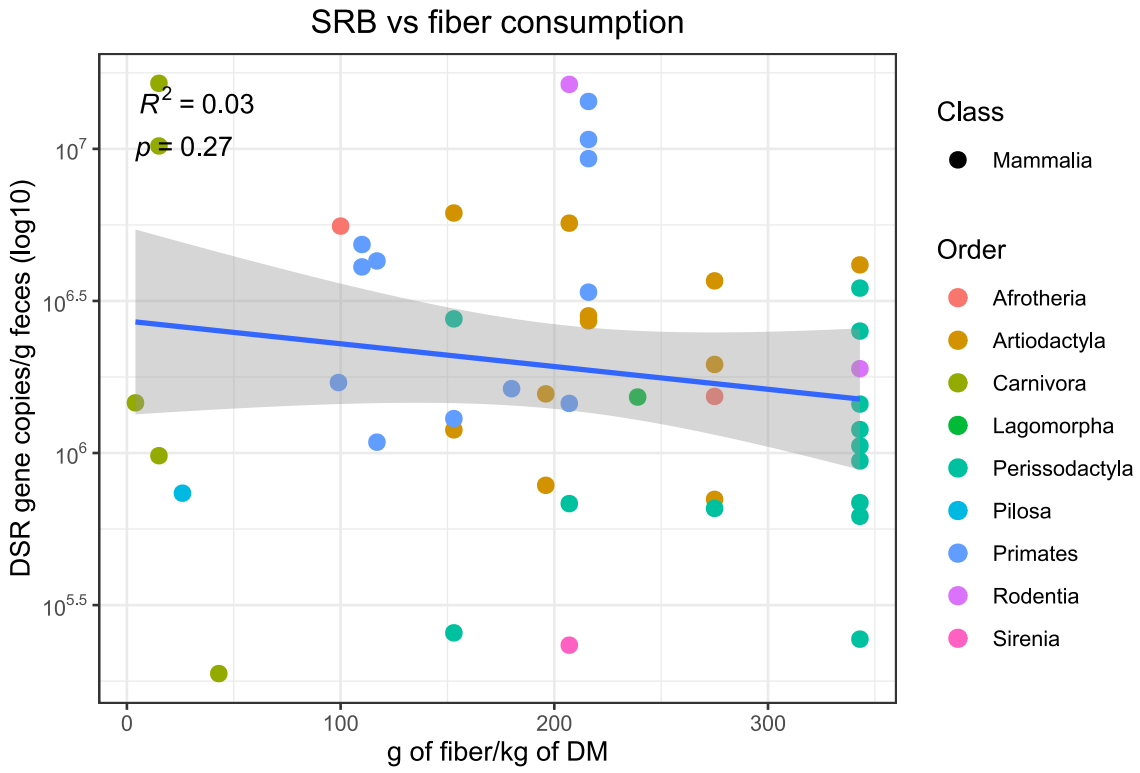


Figure 3. Sulfate reducing bacteria are not significantly correlated to fibre consumption in mammals.

However, it was surprising to find a weak positive correlation between the most abundant methyl-reducing hydrogenotrophic lineage of methanogens, the *Methanomassiliicoccales,* and the SRB. This signal was particularly prevalent in the mammals **(Chapter 1.B Figure 4)**. The positive association between *Methanomassiliicoccales* and SRBs could be due to the fact that these groups both have a high affinity for hydrogen (Morvan *et al.*, 1996; Feldewert *et al.*, 2021) and so they may be more competitive for H_2 than CO_2-reducing methanogens when H_2 concentrations are low. The dominant species of SRB in the intestine of these samples is *Desulfovibrio* which has been shown to produce trimethylamines (TMA) in the intestinal tract (Fennema *et al.*, 2016), a product specifically consumed by *Methanomassiliicoccales*. These analyses also determined that reptiles (mostly carnivorous) host the highest abundance of SRB, which could be due to their high protein diets **(Chapter 1.B Figure 4c**). Indeed, *Desulfovibrio* has genes annotated to be involved in the metabolism of both inorganic & organic sulfur compounds (Carbonero *et al.*, 2012).

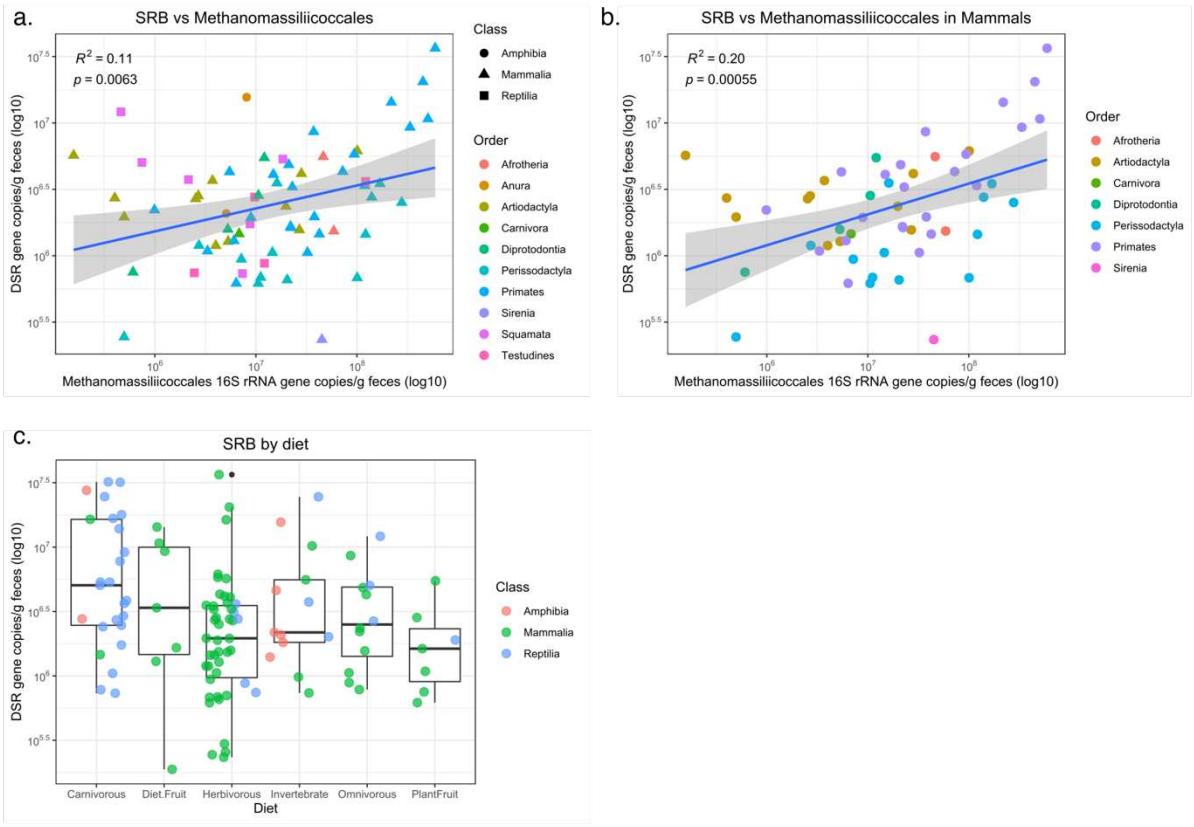


Figure 4. Sulfate reducing bacteria in Animals. **a)** Absolute abundance of SRB vs. Methanomassiliicoccales in amphibians, mammals, and reptiles. **b)** Absolute abundance of SRB vs Methanomassiliicoccales in mammals. **c)** Absolute abundance of SRB across different diets.

Conclusions

The use of qPCR and 16S rRNA gene amplicon sequencing has revealed that intestinal archaea are widespread in the animal kingdom, and that most of the intestinal archaea belong to five methanogenic genera/families. These analyses indicated that host taxonomy and diet (specifically fibre consumption) are drivers of archaeal abundance and diversity in the intestinal tract. Further investigation on H_2 competition in the gut of a diverse range of samples would provide important additional information concerning other factors that influence the abundance of intestinal methanogens. Quantification of other groups of intestinal hydrogenotrophs, such as the homoacetogens and the fumarate reducers, would be useful to better understand how the methanogens compete for H_2 in a diverse range of animal intestinal tracts. Also, despite the limitations of a meta-analysis from different studies, this preliminary work shows that the intestinal archaea in wild animals are largely similar to the

taxa found in captive individuals. It also highlights that further investigation is needed concerning intestinal archaea of marine animals.

Finally, the results of the Chapter 1.A suggest that there were at least five independent adaptation events within the domain Archaea. Thus, clarification of how these archaea have evolved to colonize the intestinal tract would reveal unique information concerning the colonization of this unique type of environment. The following chapter discusses how one of these intestinal lineages, the *Methanimicrococcus* has adapted to the host intestinal tract.

Chapter 2.A: Comparative genomic analysis of Methanimicrococcus blatticola provides insights into host-adaptation in archaea and the evolution of methanogenesis

---

In the previous chapter, I presented the abundance and diversity of intestinal archaea in a wide range of animal hosts. These results suggest host-related adaptation events in the lineages of archaea that have colonized the intestinal tract. More broadly, the dominance of five distantly related genera/families of methanogens across different animal classes suggests that several independent events of adaptation related to host colonization have occurred in Archaea. The work presented in this chapter aims at further investigating and better understand these adaptations.

In the first part of the chapter, a comparative genomic analysis of the dominant intestinal Methanosarcinales, *Methanimicrococcus blatticola*, was performed. *M. blatticola* was selected for this study for several reasons. First, it is one of the few intestinal methanogens isolated that does not belong to the Methanobacteriales. It is consistently identified in the intestinal tract and our results in Chapter 1 show that it is indeed the most commonly detected Methanosarcinales in fecal samples. Finally, *M. blatticola* is an obligate methyl-reducing hydrogenotrophic methanogen, which is unique among currently known the Methanosarcinales. Thus, it was of interest to understand how the genome of this archaea has changed. Results show that *M. blatticola* has undergone a significant shuffling of genomic content as a consequence of the colonization of the intestinal tract. These host-associated changes include the acquisition of a distinct repertoire of cell surface modification genes as well as the loss of other genes that are likely redundant in a nutrient rich environment like the intestinal tract. All of which suggests that there has been some level of convergent adaptations in the intestinal methanogens, as similar genomic characteristics have been identified in the host associated Methanobacteriales and Methanomassiliicoccales. Additionally, the extensive

loss of genes involved in H_4MPT methyl branch of the Wood–Ljungdahl pathway provides unique insights into the evolution of the methyl-reducing hydrogenotrophic methanogenesis pathway.

This work was published in ISME Communications in September of 2021.

In the second part of the chapter, I will present first results on novel host-associated archaeal genomes. For this analysis, 64 of the fecal samples from Chapter 1 were subjected to shotgun metagenomic sequencing, and 191 high-quality archaeal metagenome-assembled-genomes (MAGs) were reconstructed. These MAGs plus those assembled from publicly available datasets greatly expand the diversity of intestinal archaeal genomes currently available, both in terms of archaeal taxa and hosts. This dataset will be used in the future to perform a comparative metagenomic analysis similar to what was done in *Methanimicrococcus* to further elucidate host-associated adaptations in the Archaea.

ISME

www.nature.com/ismecomms

ARTICLE OPEN

Check for updates

Comparative genomic analysis of *Methanimicrococcus blatticola* provides insights into host adaptation in archaea and the evolution of methanogenesis

Courtney M. Thomas[1,2], Najwa Taib[1], Simonetta Gribaldo[1] and Guillaume Borrel[1]✉

Other than the Methanobacteriales and Methanomassiliicoccales, the characteristics of archaea that inhabit the animal microbiome are largely unknown. *Methanimicrococcus blatticola*, a member of the Methanosarcinales, currently reunites two unique features within this order: it is a colonizer of the animal digestive tract and can only reduce methyl compounds with H_2 for methanogenesis, a increasingly recognized metabolism in the archaea and whose origin remains debated. To understand the origin of these characteristics, we have carried out a large-scale comparative genomic analysis. We infer the loss of more than a thousand genes in *M. blatticola*, by far the largest genome reduction across all Methanosarcinales. These include numerous elements for sensing the environment and adapting to more stable gut conditions, as well as a significant remodeling of the cell surface components likely involved in host and gut microbiota interactions. Several of these modifications parallel those previously observed in phylogenetically distant archaea and bacteria from the animal microbiome, suggesting large-scale convergent mechanisms of adaptation to the gut. Strikingly, *M. blatticola* has lost almost all genes coding for the H_4MPT methyl branch of the Wood–Ljungdahl pathway (to the exception of *mer*), a phenomenon never reported before in any member of Class I or Class II methanogens. The loss of this pathway illustrates one of the evolutionary processes that may have led to the emergence of methyl-reducing hydrogenotrophic methanogens, possibly linked to the colonization of organic-rich environments (including the animal gut) where both methyl compounds and hydrogen are abundant.

ISME Communications; https://doi.org/10.1038/s43705-021-00050-y

INTRODUCTION

Methanogenic archaea are common components of the intestinal microbiota of animals ranging from insects to humans [1–4]. However, archaea are generally overlooked in intestinal microbiome studies, leaving their ecology and diversity largely undescribed in this context. Additionally, only a small number of gut-associated archaeal isolates exists outside of the *Methanobrevibacter* and *Methanosphaera* genera (both belonging to the order Methanobacteriales), resulting in little knowledge concerning the adaptations that allowed archaea to colonize the animal gut. The physical/chemical conditions of a microbe's niche exert selective pressures that can influence its genome content [5]. The gut of animals is characterized by a high and almost constant intake of fresh organic matter that distinguishes it from most aquatic and terrestrial environments. At the same time, gut microbes have to adapt to stressors such as potential removal from the host by the peristaltic movement of the digestive tract [5]. Host-associated archaea lineages have not only distinct genetic background due to their distant evolutionary relationships but also different histories of adaptation to the gut, and they can cover a wide variety of hosts. This poses the question of whether these distantly related archaea may display convergent adaptations or not. Genomic adaptations to the intestinal microbiome have been proposed in some methanogens: these include the gain of genes coding for cell surface proteins facilitating interactions or adhesion and possibly genomic streamlining [3, 6–10].

Within the order Methanosarcinales, two host-associated methanogens have been isolated, *Methanimicrococcus blatticola* PA from a cockroach (*Periplaneta americana*) [11] and *Methanosarcina barkeri* CM1 from the rumen of a cow [12]. However, the genome of *M. barkeri* CM1 is very similar to that of its freshwater close relative *M. barkeri* fusaro [13] and *Methanosarcina* members do not regularly occur in the rumen microbiota, suggesting a generalist lifestyle or a transient presence in the gut. In contrast, *Methanimicrococcus* has been reported multiple times in the gut of animals, and in some cases, it was found to represent a large majority of the methanogenic community in termites and some ruminants [14, 15]. Based on this apparent niche specificity, *M. blatticola* should display specific genomic adaptations to a host-associated lifestyle.

In addition to being an interesting model to study adaptations of archaea to the gut environment, *M. blatticola* is the only known member of Methanosarcinales that obligately uses H_2 to reduce methyl compounds [11, 16, 17] (i.e., methyl-reducing hydrogenotrophic methanogenesis), a metabolism that has been recently

[1]Department of Microbiology, UMR 2001, Unit Evolutionary Biology of the Microbial Cell, Institut Pasteur, Paris, France. [2]Université Paris Diderot, Sorbonne Paris Cité, Paris, France. ✉email: guillaume.borrel@pasteur.fr

Received: 18 March 2021 Revised: 8 July 2021 Accepted: 26 August 2021
Published online: 09 September 2021

reported in a growingly large number of newly discovered methanogen lineages [18–23]. Here we have analyzed the genome of *M. blatticola* PA, currently the only available isolated Methanimicrococcus strain, and carried out a large-scale comparative genomic analysis to understand the emergence of its unique characteristics. Our results provide new insights into the processes leading to the adaptation of archaea to the animal digestive tract and highlight one of the possible paths that led to the emergence of methyl-reducing hydrogenotrophic methanogenesis.

MATERIALS AND METHODS

Genome sequencing and annotation

The *M. blatticola* PA genome was sequenced from DNA ordered at the DSMZ German culture collection (DSM 13328), using Illumina MiSeq Nano V2 (2 ×250 PE). Reads were assembled using Spades 3.11 [24]. A total of 12 contigs >1 kb were obtained, with an average coverage of 157×. Almost all reads were assembled, as 99.97% of them aligned on these contigs. Genes were predicted using prodigal [25]. The genome and protein sequences are available in GenBank, PRJNA731512, and in Supplementary Dataset 1 and 2. Gene functions were annotated with Kyoto Encyclopedia of Genes and Genomes (KEGG) [26] and Eggnog [27]. Genes associated with the methanogenesis pathways were specifically targeted with HMM searches using PFAM [28], TIGRFAMs [29], and custom HMM profiles. Presence of transmembrane domains was determined using TMHMM 2.0 [30]. In the course of our analysis, this strain was also sequenced by JGI (2756170388) and released in GenBank (GCA_004363215.1). Unsurprisingly, the sequences and statistics of the two genomes are practically identical (size difference of <1 kb, average nucleotide identity (ANI) of 99.99%, 13 contigs). The analyses presented here were conducted on the genome we sequenced.

Dataset selection

Phylogenetic and comparative genomic analyses were carried out using 21 Methanosarcinales genomes, including 15 *Methanosarcinaceae* (Table S1). We prioritized the selection of high-quality genomes, a majority of which came from cultured and well-characterized species. Genomes from metagenomes were only included when no genome from an isolated strain was available and only if the estimated completeness level was >90% and the contamination <5%, as estimated with CheckM [31]. Four additional genomes were used as outgroups, three Methanocellales and one "*Ca*. Methanoflorentaceae" genomes, corresponding to close relatives to the Methanosarcinales.

Glycosyltransferase (GT) and PGF-CTERM annotation

GTs were annotated using dbCAN2 [32]. dbCAN2 was run on 507 archaeal genomes from a local database (Table S2), covering all major lineages of archaea and having >75% of completeness and <5% contamination. GTs in Methanosarcinales of the selected genome dataset described above was determined using the annotation found in the CAZy database [33].

Proteins containing a PGF-CTERM domain, which are associated with the PGF-CTERM/archaeosortase A system and possibly N-glycosylated, were first identified in the 21 Methanosarcinales plus the *M. blatticola* PA genomes using TIGR04126 HMM profile. Because this domain is short, new HMM profiles were generated from the PGF-CTERM domains of Methanosarcinales and *Methanimicrococcus* proteins (Supplementary Dataset 3). These new HMM profiles allowed the identification of extra proteins with a PGF-CTERM domain.

Phylogenetic analyses

The phylogenetic position of *M. blatticola* PA was established using three different datasets: (i) the concatenation of 40 universal markers as described in [18] that correspond to 36 proteins of the Phylosift dataset [34], plus the alpha and beta subunits of the RNA polymerase and two universal ribosomal proteins (L30, S4), (ii) the concatenation of the three MCR subunits (McrABG), and (iii) the 16S rRNA gene. The protein sequences for the phylogenetic analysis were retrieved from *M. blatticola* PA and the other 25 selected genomes using HMMer [35]. Single-protein datasets were aligned using MAFFT [36] (mafft-linsi), trimmed with BMGE [37] (BLOSUM30 substitution matrix) and concatenated. The 16S rRNA genes of *M. blatticola* PA and the 25 selected genomes were also aligned with MAFFT. A Bayesian phylogeny was built for the concatenation of the 40 universal markers using PhyloBayes [38] under the CAT + GTR + Γ4 model with four independent MMC chains, until convergence (maxdiff < 0.05). For the concatenation of the McrABG and the 16S rRNA gene sequences, Maximum Likelihood phylogenies were calculated with IQ-Tree [39] with the NEWTEST option for best model selection (McrABG: LG + F + Γ4; 16S: GTR + R3) and 1000 ultrafast bootstrap iterations. To investigate the environmental distribution of *Methanimicrococcus*, 16S rRNA gene sequences were searched in GenBank (Nucleotide collection nr/nt) and in all assembled metagenomes of the IMG databases using BLAST. Sequences >750 bp and with >91% identity with *M. blatticola* PA were downloaded. This identity cutoff was based on the minimal distance between *M. blatticola* PA and Methanosarcina spp. 16S rRNA gene sequences. A phylogeny of these sequences was built as described above for the 16S rRNA genes. Fasta files of the sequences used for phylogeny are available in Supplementary Dataset 4.

Gene flow reconstruction

This analysis was based on the phylogeny of the 40 concatenated markers described above and on the distribution of protein families among the 26 taxa used in this phylogeny. The protein families were determined using Silix with a 40% identity and 80% coverage cut-off [40], generating 21,377 protein families whose distribution was mapped on the reference phylogeny. The events of gene gains and losses that occurred during the evolution of the Methanosarcinales, and in particular on the branch leading to *M. blatticola* PA, were assessed using a death and birth model implemented in Count [41]. The rate of variation across families were optimized iteratively from uniform rates of gain, loss, and duplication to three discrete categories for the gamma distribution, using default parameters. Family history was calculated using posterior probabilities. *M. blatticola* PA protein sequences and representative of *Methanosarcinaceae* protein families that have been lost in *M. blatticola* PA were compared with three recently released Methanimicrococcus metagenome-assembled genomes (MAGs; i.e., GCA_009784005.1, GCA_009783635.1, and GCA_009776675.1, having 92.1/0%, 82.5/2.3%, 75.1/4.6% of completeness/contamination, respectively, according to CheckM [31], Table S3). A 40% identity and 50% coverage cut-off was used to determine whether proteins present or predicted to be lost in *M. blatticola* PA were coded in these MAGs. A lower coverage cut-off than in the main analysis was chosen because two of these MAGs are sheared (138 and 239 contigs), and as a consequence, a number of their proteins can be cut and split between different contigs.

RESULTS AND DISCUSSION

Widespread association of the *Methanimicrococcus* genus with gut microbiomes and extensive genome reduction in *M. blatticola*

We sequenced the genome of *M. blatticola* PA from the DSMZ German culture collection (DSM 13328). The draft genome (12 contigs) is 1.78 Mb long with a GC content of 42% and is predicted to encode 1569 proteins (Table S1). It is practically identical to the one sequenced by JGI (2756170388) in the course of our analysis and released in GenBank (GCA_004363215.1) (size difference of <1 kb, ANI of 99.99%, 13 contigs). The analyses presented hereafter were conducted on the genome we sequenced.

We first established the phylogenetic position of *M. blatticola* PA using three different datasets: the concatenation of 40 universal markers as described in [18] (Fig. 1a), the 16S rRNA gene (Fig. S1), and a concatenation of the three MCR subunits (McrABG; Fig. S2). In all three resulting phylogenies, *M. blatticola* PA consistently and robustly branches within *Methanosarcinaceae*, at the base of *Methanosarcina*. This placement is different from the one previously proposed, at the base of the *Methanosarcinaceae*, but the phylogeny was poorly resolved [11].

Investigation of the environmental distribution of *Methanimicrococcus* members based on 16S rRNA gene sequences retrieved from GenBank and IMG databases revealed that they commonly occur in the animal digestive tract (Fig. 1b), in contrast to other members of the Methanosarcinales that are generally found in wetland soils and sediments [42]. Three distinct clades can be

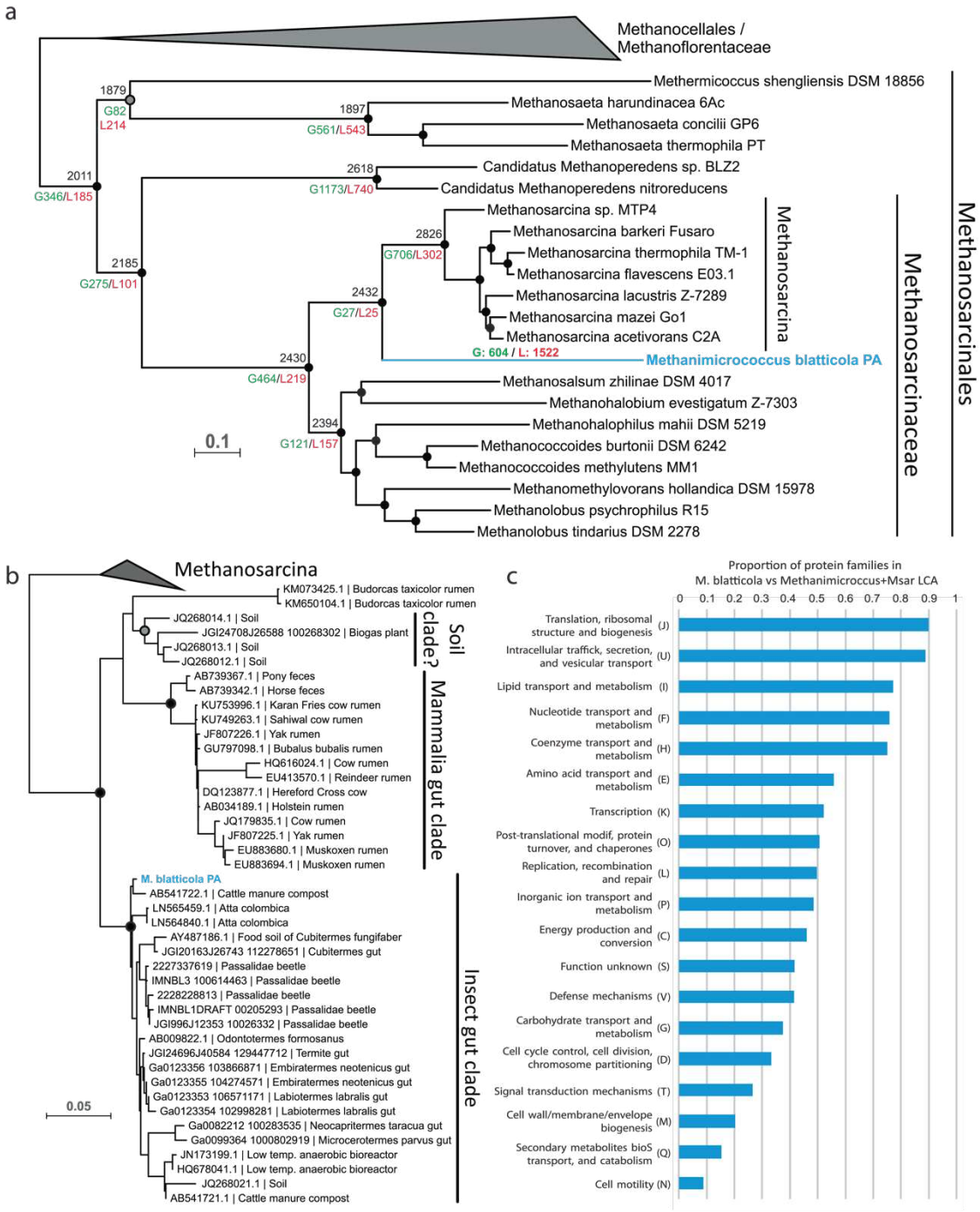


Fig. 1 Phylogenetic position, gene flow, and environmental distribution of *Methanimicrococcus*. **a** Bayesian phylogeny (PhyloBayes, CAT + GTR + Γ4) of the Methanosarcinales based on a concatenation of 40 markers universal markers (10,126 positions). Numbers on the nodes indicate the genes present (black), gained (green, preceded by "G"), and lost (red preceded by "L") along the evolution of the Methanosarcinales, as predicted by Count program [41]. The color of the circles at the nodes indicates the posterior probability values (black, >0.95; gray, between 0.70 and 0.95). **b** Maximum likelihood phylogeny (GTR + I + Γ4) based on the 16S rRNA gene sequences (1497 positions), displaying the environments where *Methanimicrococcus* occurs, mainly insect and mammalian digestive tract. The soil clade is followed by a question mark as all sequences of this poorly supported clade come from a single unpublished study, and their origin is therefore not totally reliable. The color of the circles at the nodes indicates ultrafast bootstrap values (black, >95; gray, between 70 and 95). The supports of the branches are only displayed for the main clades. **c** For each COG category, proportion of protein families in *Methanimicrococcus blatticola* PA relative to those present in the last common ancestor (LCA) of *Methanimicrococcus* + *Methanosarcina* (Msar).

observed within the *Methanimicrococcus* genus: one is formed by sequences mainly from the digestive tract of insects (including *M. blatticola* PA), the second corresponds to sequences from mammalian digestive tracts, and the third consists of a few sequences from a single unpublished study on soil (Fig. 1b). These three clusters may indicate either one ancestral or two independent specializations to the animal digestive tract (mammals and insects) in *Methanimicrococcus*.

To investigate how genome content was affected during the transition from an open environment to the animal gut, we inferred the gene gains and losses that occurred between the last common ancestor (LCA) of *Methanosarcina/Methanimicrococcus* and *M. blatticola* PA. This analysis revealed an important shift in the gene content (Fig. 1a and Table S4). Indeed, one-third of *M. blatticola* PA genes (604 genes) were not present in *Methanosarcina/Methanimicrococcus* LCA (Fig. 1a) whereas two-thirds of the genes present in *Methanosarcina/Methanimicrococcus* LCA were lost (1522 genes), corresponding by far to the largest genome reduction across all Methanosarcinales (Fig. 1a). The small number of proteins encoded in the *M. blatticola* genome (1569) compared to *Methanosarcina* (3400 in average) or most other *Methanosarcinaceae* (2300 in average) further supports this prediction (Table S1). All Clusters of Orthologous Genes (COG) categories display a net loss in *M. blatticola* PA compared to *Methanosarcina/Methanimicrococcus* LCA (Fig. 1c) but some are more impacted than others. Several core cellular processes, such as "Translation, ribosomal structure and biogenesis (J)," "Lipid transport and metabolism (I)," or "Nucleotide transport and metabolism (F)" are only weakly affected, while others, such as cell envelope biogenesis, display a massive reduction in the number of protein families involved (Fig. 1c). A majority of the acquired genes have only a general or no predicted function (60% have no COG category and 80% no KEGG annotation), in contrast with those that were lost (4% have no COG category and 57% no KEGG annotation; Table S4).

Extensive modification of the cell–host interaction surface

An expected trait of adaptation to a host-associated lifestyle is modification of the repertoire of membrane-bound proteins, situated at the interface between the cell and its environment. Indeed, we observed that two-thirds of the proteins inferred to be membrane bound in *M. blatticola* PA were not present in *Methanosarcina/Methanimicrococcus* LCA and were therefore specifically acquired in the lineage leading to this archaeon. Similar to the cytoplasmic proteins acquired by *M. blatticola* PA, most of the membrane-associated proteins have an unknown function (Table S4). However, among them, several correspond to putative adhesin-like proteins (ALPs) (Tables S4 and 1). They are annotated as cell wall-binding repeat-containing protein or collagen-binding protein, and two contain collagen or cellulose (CBM44) binding domains. Sixteen acquired proteins contain one or several Listeria-Bacteroidetes repeats/Flg_New (PF09479) domains that display structural similarities with β-grasp folds, having various binding functions [43]. In Archaea, proteins with these domains were previously only known from Methanomassiliicoccales of the human gut (up to 38 in "*Ca*. Methanomassiliicoccus intestinalis") where they have been suggested to be involved in the attachment to specific sites in the digestive tract [3]. Bacteria with a high number of genes coding for Listeria-Bacteroidetes repeats/Flg_New domains are also associated with an animal host [3]. The presence of these putative ALPs in two phylogenetically distant gut-associated archaeal lineages (belonging to Methanomassiliicoccales and Methanosarcinales) and in host-associated bacteria further supports the hypothesis that they are specific adaptations to the digestive tract [3] and indicate convergent adaptation. Among the acquired genes coding for membrane-bound proteins, 15 contain a GLUG motif of unknown function (2 of them also have Flg_new domains) and may also have a role in adhesion. In the absence of flagella, these proteins may prevent the washout of *M. blatticola* PA cells from the insect gut and help them to remain at specific sites where they are most competitive for substrates [14]. It was shown in fact that many methanogens associated with cockroaches are loosely associated with chitin bristles present in the hindgut [1].

While *M. blatticola* PA has kept the potential to express major S-layer proteins present in other Methanosarcinales, its cell surface was likely highly modified through the loss of many enzymes involved in the synthesis of nucleotide-activated sugars that are precursors of cell-surface glycoconjugates/extracellular polysaccharides (Figs. 1c and 2). Moreover, *M. blatticola* PA has lost up to 35 glycosyl transferases (GT), including 12 with transmembrane domains. Within Archaea, this is the genome that currently encodes the lowest number of GTs (2) (Table S2) and it also contrasts with other Methanosarcinales that have 33 GTs genes on average (Table S5). Even if not to the same extent, other host-associated methanogens (Methanobacteriales and Methanomassiliicoccales) also tend to have a lower number of GTs than their closest free-living relatives (Fig. 3), suggesting that the loss of GTs is another convergent characteristic related to the adaptation to gut conditions in archaea. The two GTs conserved in *M. blatticola* PA are involved in N-glycosylation of cell surface proteins by (i) transferring the first nucleotide-activated sugar on the lipid carrier (AglJ) and (ii) transferring the (oligo)saccharide from the lipid carrier to the nascent protein (AglB). Considering the almost complete loss of GTs and nucleotide-activated sugars synthesis genes, the nature of the glycans on the N-glycosylated proteins remains unclear (Fig. 2). Four membrane-bound proteins of unknown function and one with Flg_new domains have a PGF-CTERM domain, indicating that they are recognized by the archeosortase A present in *M. blatticola* PA and possibly N-glycosylated [44]. The number of such proteins is much lower in *M. blatticola* PA compared to other Methanosarcinales, which on average encode 20 proteins with this signal, suggesting a lower number of N-glycosylated proteins in *M. blatticola* PA (Table S5). Moreover, *M. blatticola* PA has lost two archaeosortases systems (PEF-CTERM and VPXXXP-CTERM) that may be involved in export and posttranslational modification of additional membrane-bound proteins [44].

Glycans originating from commensal microorganisms cover many roles in the digestive tract, including persistence in this environment, microbe–microbe interactions, protection from phages, and host–microbe interactions, such as host-immune system stimulation [45]. However, there are no data on the role of archaeal glycans in the host microbiome and very few on the cell wall of the *Methanosarcinaceae* members outside the *Methanosarcina* genus. When they are aggregated, *Methanosarcina* members are surrounded by methanochondroitin [46], a fibrillar polymer that is analogous to the chondroitin present in tissues of animals, including cockroaches [47]. The ability to synthesize two of the sugars that constitute methanochondroitin (*N*-acetylgalactosamine and glucuronic acid) was lost in *M. blatticola* PA (Fig. 2). To our knowledge, such a large loss of genes associated with cell-surface glycoconjugates or exopolysaccharide synthesis has never been reported in the archaea, and how this event is linked to adaptation of *M. blatticola* to the digestive tract remains to be elucidated.

Adaptations to the gut as a nutrient/substrate rich environment

Several other gene losses reflect an adaptation to the gut as a nutrient/substrate-rich environment (Table 1), relative to non-gut environments that are essentially oligotrophic [48]. *M. blatticola* PA has lost a high-affinity phosphate transporter (*pstABC*) and an alkaline phosphatase (*phoA*) for organo-phosphate utilization, both of which are known to be upregulated under low phosphate concentrations in *Methanosarcina mazei* [49]. In contrast, the PitA low-affinity

Table 1. Summary of the genes gained and lost in *M. blatticola* PA, discussed in the text as adaptations to a host-associated lifestyle, and comparison with the presence/absence patterns in other host-associated archaea and bacteria reported in the literature.

Genes	State	Function	Similar pres/abs patterns in 3 *Methanimicrococcus* MAGs[a]	Similar pres/abs patterns in other host-associated archaea[b]	Similar to gut-associated bacteria[c]
fwdABCDEFG, fmdBD, ftr, mch, mtd, mfnBDEF, mooABC/mobAB/moeA, modABCDE, wtpAB, *tupAB, Rnf, Fpo-like*	Lost	Methylotrophic and CO_2-reducing methanogenesis	Yes	*mooABC/moeA* in *Methanosphaera*	
comDE	Lost	Coenzyme-M biosynthesis	Yes	"*Ca.* Methanoplasma termitum", ISO4-H4, ISO4-G1, *Methanobrevibacter ruminantium*	
pstABC	Lost	High-affinity phosphate transporter	Absent in GCA009776675		
phoA	Lost	Organo-phosphate utilization	Yes		
nifKDU	Lost	N_2 fixation	Yes	*Methanobrevibacter*	
cdhAB	Lost	CO_2 fixation	Yes	*Methanobrevibacter*	
asnB, trpABDFGE	Lost	Amino acid synthesis	Most *trp* genes present in GCA009776675		
cstA	Gained	Uptake of pyruvate and peptides	No	"*Ca.* Methanomassiliicoccus intestinalis"	
m19, m21, m24	Lost	MCR-associated markers; regulatory processes of methane metabolism related to changes in substrate/nutrient availability[d]	Yes	Host-associated Methanomassiliicoccales, *Methanobrevibacter*, *Methanosphaera*	
gys, gyp, PK_C, pfkC, galU	Lost	Glycogen synthesis and degradation	Yes		
cheABCFDRWY	Lost	Sensing environment	Yes		Yes
19 histidine kinase/PAS sensor	Lost	Sensing environment	Yes (18/19 also absent in the three MAGs)		Yes
flaBCEFGH	Lost	Motility	Yes		Yes
43 transcriptional regulators	Lost	Modification of gene expression	Yes (37/43 also absent in the three MAGs)	*Methanobrevibacter smithii, Methanosphaera stadtmanae*	Yes
16 genes encoding a Flg_new domain	Gained	Adhesion[d]	Yes (3 to 13 present)	Host-associated Methanomassiliicoccales	Yes
MBLAb0210	Gained	Binding to cellulose (CBM44)	Yes		
MBLAb0400	Gained	Binding to pectins[d]	Yes		
MBLAb1410	Gained	Binding to cell wall[d] (CBM37)	No		
MBLAb1461	Gained	Binding to collagen[d]	No		
MBLAb1469	Gained	Cell wall-binding repeat	No		
glk, rfbA, rfbB, rfbC, rfbD, galU, galE, ugd, glmS, glmU, wecB, wecC, wbpP, aglA/ manAC, gmd, flc	Lost	Synthesis of nucleotide-activated sugars, precursors of extracellular polysaccharides	Yes		
35 glycosyltransferases	Lost	Cell surface modification	Yes	Host-associated Methanomassiliicoccales, *Methanobrevibacter*	
artB	Lost	VPXXXP-CTERM-specific archaeosortases	Yes		
artC	Lost	PEF-CTERM-specific archaeosortases	Yes		

pres/abs presence/absence.
[a]Similar pattern means that gene lost in *M. blatticola* PA are absent in the three *Methanimicrococcus* MAGs or that genes gained in *M. blatticola* PA are present in these MAGs.
[b]Lineages indicated in this column have patterns of gene presence/absence (from literature cited in the main text) that match the gene gains and losses reported in *M. blatticola* PA.
[c]See literature cited in the main text.
[d]Putative function.

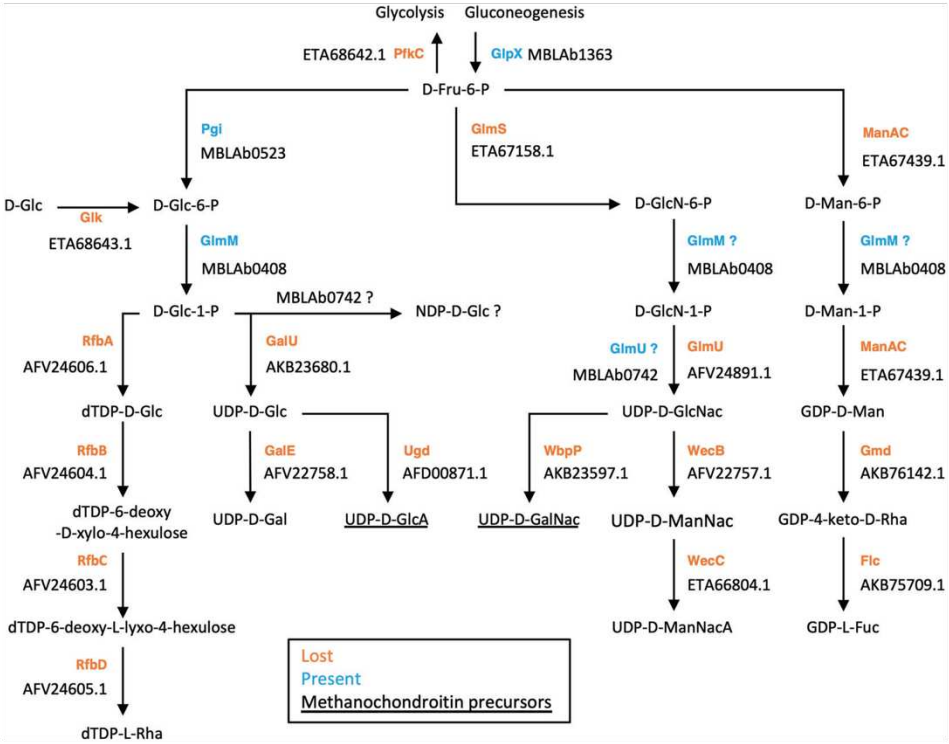


Fig. 2 Loss of multiple enzymes involved in the synthesis of nucleotide-activated sugars that are precursors of cell-surface glycoconjugates/extracellular polysaccharides. The enzyme names in orange indicate that the corresponding genes were lost in *M. blatticola* PA, while those in blue correspond to genes that have been kept in *M. blatticola* PA. The genes lost in *M. blatticola* PA are also absent in the three *Methanimicrococcus* MAGs and those present in *M. blatticola* PA are also present in the three *Methanimicrococcus* MAGs. In other organisms, GlmM could perform three different reactions, albeit with different level of activity. Methanosarcinales genomes have two copies of *glmM*, but *M. blatticola* PA has only one. This enzyme is likely involved in the conversion of glucose-6-P into glucose-1-P. The accession numbers below the protein names provide an example of sequences in the *Methanosarcinaceae*.

phosphate transporter [50] was kept in *M. blatticola* PA. Moreover, *M. blatticola* PA has lost the capacity to fix atmospheric nitrogen (nitrogenase *nifKDU* and regulatory genes P-II) and carbon dioxide (*cdhAB*) for autotrophic growth. The loss of these genes has also been previously reported in other gut methanogens belonging to *Methanosphaera* and *Methanobrevibacter* [51], and a loss of the nitrogenase was also suggested in host-associated Methanomassiliicoccales [52]. In contrast, termite-associated *Methanobrevibacter* spp. have not lost these genes [51], which may reflect adaptation to a different gut compartment or a lesser degree of specialization to the gut environment. The dependence of *M. blatticola* PA to a nutrient-rich environment is also reflected by the loss of several genes involved in amino acid synthesis (e.g., *asnB* for asparagine, *trpABDFGE* for tryptophan). The gene *cstA* was acquired in *M. blatticola* PA, possibly allowing the uptake of organic carbon and nitrogen in the form of pyruvate and peptides, as reported in gut-associated bacteria [53, 54]. This gene was independently acquired in a gut-associated Methanomassiliicoccales (*Ca.* Methanomassiliicoccus intestinalis Mx-01) suggesting convergent adaptation (Table 1 and Fig. S3). All these characteristics fit well with the high growth requirements of *M. blatticola* PA (acetate, yeast extract, tryptic soy, and vitamins [11]), as compared to most other *Methanosarcinaceae* members [42]. With the loss of *comDE* genes coding for coenzyme-M, the dependency of *M. blatticola* PA on other members of the gut microbiota includes this key cofactor of methanogenesis, consistently with the previously reported need of *M. blatticola* PA for an external source of this coenzyme [11]. The absence of *comDE* genes has been previously reported in several other gut methanogens such as *Methanobrevibacter ruminantium* [8] and several Methanomassiliicoccales (ISO4-H4 [55], ISO4-G1 [56]), including the termite-associated *Methanoplasma termitum* [9].

Seven markers recently identified as co-occurring with MCR and that are mostly conserved in Class I/Class II methanogens [18] were lost in *M. blatticola* PA. Three of them (m19, m21, and m24) are also missing in methanogens from nutrient-rich environments, including the host-associated Methanomassiliicoccales (m19 and m21) and host-associated Methanobacteriales (m21 and m24) (Tables 1 and S6; [18]). The function of these genes is currently unknown, but they have been suggested to be involved in regulatory processes of methanogenesis related to changes in substrate/nutrient availability [18]. Various methanogens, including *Methanosarcinaceae* [57, 58], are able to store carbon/energy as glycogen when substrates are available and to use it in periods of starvation. Genes involved in glycogen synthesis and degradation and those involved in glycolysis were lost in *M. blatticola* PA (Tables 1 and S4), indicating an adaptation to more stable conditions in terms of substrate availability.

The lower variability in the conditions faced by *M. blatticola* PA are also reflected in a drastically reduced capacity to sense its environment and move to more favorable conditions (Fig. 1C), with the loss of the *cheABCFDRWY* chemotaxis genes, 19 histidine kinase/PAS sensor genes, and the whole motility machinery (*flaBCEFGH*) (Tables 1 and S4). In addition, 43 genes coding for transcriptional regulators were lost. This matches previous observations reported for gut-associated *Methanobrevibacter*/*Methanosphaera* and Methanomassiliicoccales, which have a significantly smaller number of genes involved in chemotaxis, signal transduction, and transcriptional

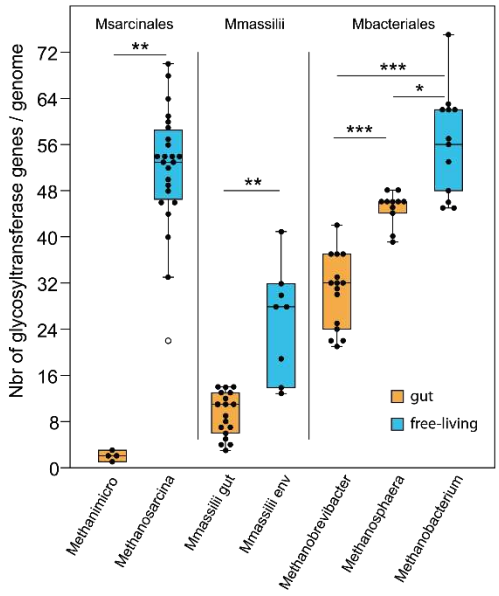


Fig. 3 Number of glycosyltransferase genes in host-associated and environmental methanogens belonging to Methanosarcinales ($n = 28$), Methanomassiliicoccales ($n = 26$) and Methanobacteriales ($n = 38$). For Methanosarcinales and Methanobacteriales, number of glycosyltransferases genes was compared between host-associated genera and the closest related genus (i.e., *Methanosarcina* for *Methanimicrococcus*; *Methanobacterium* for *Methanobrevibacter* and *Methanosphaera*). The three *Methanimicrococcus* MAGs were added to this analysis. For Methanomassiliicoccales, "*Ca.* M. intestinalis" and *Methanomethylophilaceae* MAGs were pooled in "Mmassilii gut" and all other MAGs in "Mmassilii env". Stars represent the significant differences between groups, determined using a Wilcoxon test (continuity correction: ***p.adj < 1e−4, **p.adj < 0.001, *p.adj < 0.05). Methanimicro *Methanimicrococcus*, Msarcinales Methanosarcinales, Mmassilii Methanomassiliicoccales, Mbacteriales Methanobacteriales.

regulation, as compared to non-gut methanogens [6, 52] (Table 1). These characteristics are also consistent with what has been reported from gut bacteria, which have generally a smaller repertoire of genes involved in motility, chemotaxis [59], and transcriptional regulation [60], relative to non-gut bacteria (Table 1). These results reveal the existence of largely shared traits of adaptation to a host-associated lifestyle in both archaea and bacteria.

A unique case of loss of the H_4MPT methyl-branch of the Wood–Ljungdahl pathway among Class I/II methanogens and evolution toward methyl-reducing hydrogenotrophy

Among the enzymes needed for energy conservation, the most notable losses correspond to several genes involved in methanogenesis, particularly those coding for the H_4MPT methyl branch of the Wood–Ljungdahl (H_4MPT mWL) pathway (*fwdABCDEFG*, *fmdBD*, *ftr*, *mch*, and *mtd*; Fig. 4 and Table 1). Moreover, these losses also include the genes coding for the biosynthesis of two cofactors of Fwd/Fmd, (methanofuran (*mfnBDEF*) and molybdopterin (*moaABC/mobAB/moeA*)) as well as those used to import the ions associated with these complexes (molybdate (*modABCDE*, *wtpAB*) and tungstate (*tupAB*, *wtpAB*)). Altogether, this explains the inability of *M. blatticola* PA to grow by reducing CO_2 with H_2 or by the disproportionation of methyl compounds (methylotrophic methanogenesis) as reported by Sprenger et al. [11]. This pathway was previously suggested to be absent in *M. blatticola* PA [16], but it was only inferred indirectly by the low enzymatic activity of the F_{420}-reducing hydrogenase, which did not provide information on the presence/absence of the genes of the H_4MPT mWL pathway. The almost complete absence of the genes of the H_4MPT mWL pathway is remarkable, as it has never been previously observed in a member of Class I (Methanobacteriales, Methanopyrales, Methanococcales) and Class II (Methanosarcinales, Methanomicrobiales, Methanocellales) methanogens, even in those that do not use it for methanogenesis, such as *Methanosphaera stadtmanae* [61].

Interestingly, one of the genes of the H_4MPT mWL pathway (*mer*) and those coding for the MTR complex were kept (Fig. 4). In the absence of the remaining components of the H_4MPT mWL pathway, their implication in methanogenesis is unlikely. However, these enzymes are presumably functional, as *M. blatticola* PA has also retained the capacity to synthesize their associated cofactors, H_4MPT (Table S4; [62]) and F_{420} (Fig. 4). The presence of the F_{420} cofactor is supported by the previously reported observation of fluorescence at 420-nm excitation in *M. blatticola* PA [11]. Mer and MTR may have been kept in *M. blatticola* PA for their anabolic role, as suggested by the presence of MetE and ThyA homologs, possibly involved in the synthesis of methionine using H_4MPT-CH_3 [63] and thymidylate (dTMP) using H_4MPT-CH_2 [64] (Fig. 4). Consistently with this hypothesis, it was shown that *M. barkeri* growing on methanol and H_2 is dependent on Na^+ for growth, but not for ATP and methane formation [65]. This further supports the potential anabolic role of MTR (that is Na^+ dependent) during the growth of *Methanosarcinaceae* spp. on methyl compounds with H_2. In *M. blatticola* PA, the methyl and methylene groups needed for these anabolic reactions would thus be mainly provided by the methyl compounds used for methanogenesis (Fig. 4). The methylene groups may additionally be derived from formaldehyde produced by the pentose phosphate pathway [66], because a bifunctional Fae-Hps enzyme (a fused formaldehyde-activating/3-hexulose-6-phosphate synthase) is encoded by *M. blatticola* PA.

The specialization of *M. blatticola* PA on methyl-reducing hydrogenotrophic methanogenesis was also likely associated with loss of the membrane-bound Fpo and Rnf complexes (Fig. 4), which are involved in energy conservation in other Methanosarcinales [67]. The absence of the Fpo complex and of an Ech complex in *M. blatticola* PA was previously reported [68]. The Ech complex, which was likely acquired in *Methanosarcina*, is notably used to reduce ferredoxin for anabolic purpose in species growing by reduction of methanol with H_2 [69]. One can therefore wonder how *M. blatticola* PA reduces ferredoxin for anabolic purposes (e.g., pyruvate synthesis; Fig. 4). An HdrABC/MvhD/Fd(r)hB complex may play this role by reducing ferredoxin and CoM-S-S-CoB using $F_{420}H_2$ (generated during the oxidation of H_4MPT-CH_3 by Mer, or by FrhABG using H_2 as electron donor) through electron bifurcation, in a similar way as the HdrABC/MvhADG complex [70] but using $F_{420}H_2$ instead of H_2 (Fig. 4). This potential complex was first reported in a methanotrophic member of the Methanosarcinales (ANME-2d) where it was referred as HdrABC/MvhD/FrhB and proposed to perform electron confurcation (the reverse reaction of what we propose to occur in *M. blatticola*) for ferredoxin and CoM-S-S-CoB recycling [71]. This putative complex is encoded by a gene cluster that was reported in many methane and short-chain alkane-oxidizing archaea [18].

Except for *M. blatticola* PA, all other *Methanosarcinaceae* representatives have the capacity to disproportionate methyl compounds. Thus, the specialization of *M. blatticola* PA on methyl-reducing hydrogenotrophic methanogenesis likely emerged from this type of metabolism. For methyl-compound disproportionation, *Methanosarcinaceae* spp. use the H_4MPT mWL pathway to oxidize one methyl group into CO_2, producing three reducing equivalents for the reduction of three additional methyl groups into CH_4 [72]. However, when *M. barkeri* grows on methanol in the presence of H_2, methanol is totally reduced by H_2 and no longer oxidized into CO_2 [65], suggesting that the H_4MPT mWL pathway is no longer used for methanogenesis under such conditions. It is thus possible that an ancestor of *M. blatticola* PA lost the genes encoding the H_4MPT mWL pathway after stable colonization of the animal gut environment

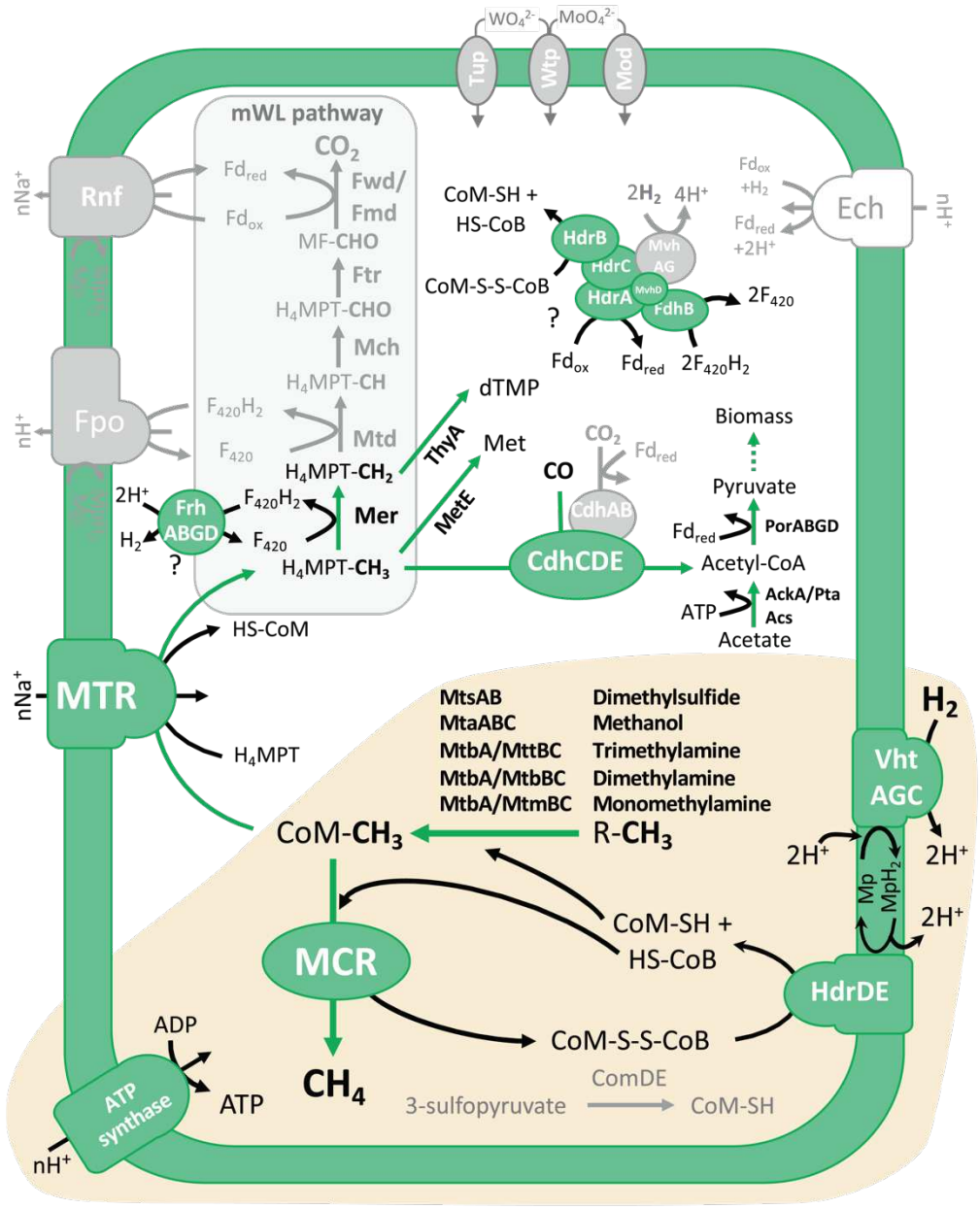


Fig. 4 Methanogenesis-associated enzymes in *M. blatticola* PA: losses and role in catabolism and anabolism. Enzymes in gray were lost in *M. blatticola* PA. The Ech complex (in white) was not present in the last common ancestor of *Methanimicrococcus* and *Methanosarcina* but likely acquired in *Methanosarcina*. Enzymes/enzymatic complexes within the beige frame (MCR, VhtAGC, HdrDE, MtsAB, MtaABC, MtbA, MttBC, MtbBC, MtmBC, and ATP synthase) are involved in energy conservation. Other enzymes/enzymatic complexes that can be involved in methane metabolisms in other archaea but predicted to be involved in anabolism in *M. blatticola* PA: MTR complex, Frh, Mer, AckA/Pta, Acs, CdhCDE, and the putative HdrABC/MvhD/FdhAB-like complex. CdhCDE may also have a role in CO detoxification. The full names of the proteins are displayed in Table S4.

where both H_2 and methyl compounds are available. Indeed, the concentrations of H_2 in the gut (4 µM in cockroaches [17]; 100 nM–50 µM in ruminants [73]; 168 µM in mice [74]; 5–156 µM in humans [75, 76]) are generally 10–1000-folds higher than those in anoxic soils or sediments (10–30 nM [77, 78]). This is likely due to the constant load of fresh organic matter making H_2 production rates much higher in the animal gut (5 µM h^{-1} and 2 mM h^{-1} [79]) than in the environment (1 µM h^{-1} in eutrophic lake sediments [77]). For similar reasons, methyl-compound concentrations in the gut (e.g., for methanol, 10 µM in cockroaches [17]; 23–72 µM in the rumen [80]; 70 µM in humans [3]) are also higher than those in sediments, where they are generally around or below the micromolar level [81, 82]. The availability of methanol and hydrogen also potentially drove the specialization on methyl-reducing hydrogenotrophic methanogenesis

in the members of the genus *Methanosphaera* (Class I methanogen, Methanobacteriales). Several Methanobacteriales have the *mtaABC* genes for methanol utilization, but except *Methanosphaera* spp. only a few have been shown to grow on methanol + H_2 [83, 84]. It is therefore possible that the ancestors of *Methanosphaera* spp. evolved stepwise from (i) CO_2-reducing methanogenesis (which is shared by all Methanobacteriales) to (ii) facultative methyl-reducing methanogenesis and then to (iii) obligate methyl-reducing methanogenesis. The transition from facultative to obligate methyl-reducing methanogenesis likely occurred when molybdopterin biosynthesis genes (absent in *Methanosphaera spp*. [51, 61] were lost. All the main enzymes of the H_4MPT mWL pathway are present in *Methanosphaera* and involved in anabolic reactions [61].

With *M. blatticola* PA, the only other known methanogens missing part or all of the H_4MPT mWL pathway belong to recently discovered lineages branching all over the tree of archaea: Methanomassiliicoccales [22] Methanofastidiosa [19], Verstraetearchaeota [21], Methanonatronarchaeia [20], and NM3 Acherontia [18]. All these methanogens have been experimentally characterized or predicted to be methyl-reducing hydrogenotrophic methanogens. The origin of this metabolism in these evolutionarily distant lineages has been discussed and investigated previously [18, 85]. Briefly, methyl-reducing hydrogenotrophic methanogenesis could have originated early in the evolution of archaea and been vertically inherited in these lineages or rather have been acquired by horizontal gene transfers and/or emerged through the loss of the H_4MPT mWL pathway and the MTR complex. These hypotheses are not mutually exclusive, and according to phylogenetic analyses of the MCR complex, the acquisition of this metabolism via horizontal gene transfer is very likely at least for some of these lineages [18]. The hypothesis of the emergence of methyl-reducing hydrogenotrophic methanogenesis from methanogenesis involving the H_4MPT mWL pathway is difficult to test in these deep-branching lineages. The case of *M. blatticola* therefore represents the first concrete proof of a partial loss of the H_4MPT mWL pathway in a methanogen. When methanogens stop using the H_4MPT mWL pathway for methanogenesis, they tend to keep these enzymes for anabolic purposes, but whether those enzymes are replaced by other enzymes over long evolutionary periods remains unknown.

Three novel MAGs support the patterns of gene gains/losses in *Methanimicrococcus*

During the course of this analysis, three *Methanimicrococcus* MAGs from a termite gut sample were released in GenBank [86] (Table S3 and Fig. S4). By comparing the protein sequences of these MAGs with those of *M. blatticola* PA, we observe that most of the genes (93.3%) that we predicted to have been lost in *M. blatticola* PA are also absent in these three MAGs (Table S4). This shows that most of these losses occurred before the LCA of these four *Methanimicrococcus* species. The proteins absent in the three *Methanimicrococcus* MAGs cover those discussed above and are involved in methanogenesis (mWL pathway and MCR-associated markers), glycogen synthesis/utilization, chemotaxis, nucleotide-activated sugar synthesis, GTs (1–3 only are present, Fig. 3), and 86% (37) of the transcriptional regulators lost in *M. blatticola* PA (Table 1). Only 0.01% of genes predicted to be lost in *M. blatticola* PA are present in all 3 MAGs and could therefore represent losses specific to this species. A smaller fraction of the genes that were gained in *M. blatticola* PA are also present in the 3 *Methanimicrococcus* MAGs: 50% are present in at least 1 of them and only 19.4% are present in all 3 MAGs, suggesting that there is a weaker pattern of gene acquisition than gene loss. However, some of the gained proteins—in particular ALPs—may have evolved faster than average following acquisition and could have been missed by our identity threshold. For example, no genes coding for homologs of the *M. blatticola* PA proteins with a Flg_new domain were found in the 3 MAGs using a 40% identity threshold, but we detected 3 (in the least complete MAG) to 13 proteins with a Flg_new domain in the proteome of these MAGs using a specific HMM search.

CONCLUDING REMARKS

Our results uncover how the transition from the open environment to the gut can deeply modify the physiology of an archaeon, including its central energy metabolism. Moreover, it reveals one of the possible paths toward the emergence of methyl-reducing hydrogenotrophic methanogenesis. In contrast to Methanomassiliicoccales, *M. blatticola* PA and *Methanosphaera* spp. have specialized relatively recently (at the genus level) on methyl-reducing hydrogenotrophic methanogenesis. This probably occurred after gut colonization, being triggered by the conditions encountered in this environment. After they became obligate methyl-reducing hydrogenotrophic methanogens, *M. blatticola* PA and *Methanosphaera* spp. kept one or all of the H_4MPT mWL enzymes, respectively, for the synthesis of amino acids and purines. This may represent an intermediate state toward the complete loss of H_4MPT mWL pathway and the MTR complex, as observed in the deep-branching lineages of obligate methyl-reducing hydrogenotrophic methanogens that were recently discovered [18–22]. There are multiple other (non-exclusive) explanations for the presence of this metabolism in free-living methanogens [18]: (i) it could have arisen from horizontal gene transfer to previously non-methanogenic archaea, (ii) the mWL pathway could have been lost in CO_2-reducing methanogens thriving in some specific non-gut environment where the conditions are favorable to this type of methanogenesis, and (iii) the methyl-reducing hydrogenotrophic methanogenic pathway may be the ancestral type of methanogenesis. In this latter case, the specialization of *Methanimicrococcus* and *Methanosphaera* on this type of metabolism because of gut colonization would correspond to a step back to the original type of methanogenesis.

Our analysis also highlights multiple adaptations to the digestive tract of animals, several of which have also been previously proposed in phylogenetically distant host-associated archaea belonging to the Methanobacteriales and Methanomassiliicoccales. These convergent adaptations include the loss of the ability to grow autotrophically, to fix nitrogen, and of specific methanogenesis/MCR-associated markers. They also indicate a reduced capacity to sense and respond to variations in the environment, and an important shift in cell surface elements, notably by reduction in the number of GTs and gain of ALPs such as the Listeria-Bacteroidetes repeats/Flg_New domain proteins. The gain and loss of some of these genes also fit patterns of gene depletion or enrichment observed in gut bacteria, revealing common adaptation mechanisms across the two prokaryotic domains. Several gene losses observed in *M. blatticola* PA, and suggested in other gut methanogens, are also similar to those reported for two methanogens endosymbionts of ciliates, *Methanobrevibacter* sp. NOE and *Methanocorpusculum* sp. MCE, affiliated to Methanobacteriales and Methanomicrobiales, respectively [87]. Indeed, these two endosymbionts have also lost transcriptional regulators, aromatic amino acid biosynthesis, and cell surface modification enzymes [87]. Beyond the gut environment, some of these losses could thus be more generally indicative of host adaptation.

Currently, the availability of archaeal genomes associated with a host is essentially restricted to humans, ruminants, and, to a lesser extent, termites and sponges. Sequencing of new archaeal genomes from a wide variety of animal hosts will therefore provide key information to draw a comprehensive picture of host-adaptation mechanisms in the archaea.

REFERENCES

1. Hackstein JH, Stumm CK. Methane production in terrestrial arthropods. Proc Natl Acad Sci USA. 1994;91:5441–5.
2. Hackstein JHP, van Alen TA. Fecal methanogens and vertebrate evolution. Evolution. 1996;50:559–72.
3. Borrel G, McCann A, Deane J, Neto MC, Lynch DB, Brugère JF, et al. Genomics and metagenomics of trimethylamine-utilizing archaea in the human gut microbiome. ISME J. 2017;11:2059–74.
4. Raymann K, Moeller AH, Goodman AL, Ochman H. Unexplored archaeal diversity in the great ape gut microbiome. mSphere. 2017;2:e00026-17.

5. Douglas AE. Multiorganismal insects: diversity and function of resident microorganisms. Annu Rev Entomol. 2015;60:17–34.
6. Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henrissat B, Fulton R, et al. Genomic and metabolic adaptations of *Methanobrevibacter smithii* to the human gut. Proc Natl Acad Sci USA. 2007;104:10643–8.
7. Gaci N, Borrel G, Tottey W, O'Toole PW, Brugère JF. Archaea and the human gut: new beginning of an old story. World J Gastroenterol. 2014;20:16062–78.
8. Leahy SC, Kelly WJ, Altermann E, Ronimus RS, Yeoman CJ, Pacheco DM, et al. The genome sequence of the rumen methanogen *Methanobrevibacter ruminantium* reveals new possibilities for controlling ruminant methane emissions. PLoS ONE. 2010;5:e8926.
9. Lang K, Schuldes J, Klingl A, Poehlein A, Daniel R, Brunea A. New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of '*Candidatus* Methanoplasma termitum'. Appl Environ Microbiol. 2015;81:1338–52.
10. Borrel G, Brugère JF, Gribaldo S, Schmitz RA, Moissl-Eichinger C. The host-associated archaeome. Nat Rev Microbiol. 2020;18:622–36.
11. Sprenger WW, van Belzen MC, Rosenberg J, Hackstein JH, Keltjens JT. 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. 2000;50:1989–99.
12. Jarvis GN, Strömpl C, Burgess DM, Skillman LC, Moore ER, Joblin KN. Isolation and identification of ruminal methanogens from grazing cattle. Curr Microbiol. 2000;40:327–32.
13. Lambie SC, Kelly WJ, Leahy SC, Li D, Reilly K, McAllister TA, et al. The complete genome sequence of the rumen methanogen *Methanosarcina barkeri* CM1. Stand Genomic Sci. 2015;10:57.
14. Brune, A. Methanogens in the digestive tract of termites. In: Hackstein JHP, editor. (Endo)symbiotic methanogenic archaea. Berlin: Springer; 2018. p. 81–101.
15. Li Z, Wang X, Alberdi A, Deng J, Zhong Z, Si H, et al. Comparative microbiome analysis reveals the ecological relationships between rumen methanogens, acetogens, and their hosts. Front Microbiol. 2020;11:1311.
16. Sprenger WW, Hackstein JHP, Keltjens JT. The energy metabolism of *Methanomicrococcus blatticola*: physiological and biochemical aspects. Antonie van Leeuwenhoek. 2005;87:289–99.
17. Sprenger WW, Hackstein JHP, Keltjens JT. 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. 2007;60:266–75.
18. Borrel G, Adam PS, McKay LJ, Chen LX, Sierra-García IN, Sieber C, et al. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. Nat Microbiol. 2019;4:603–13.
19. Nobu MK, Narihiro T, Kuroda K, Mei R, Liu WT. Chasing the elusive Euryarchaeota class WSA2: genomes reveal a uniquely fastidious methyl-reducing methanogen. ISME J. 2016;10:2478–87.
20. Sorokin DY, Makarova KS, Abbas B, Ferrer M, Golyshin PN, Galinski EA, et al. Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. Nat Microbiol. 2017;2:17081.
21. Vanwonterghem I, Evans PN, Parks DH, Jensen PD, Woodcroft BJ, Hugenholtz P, et al. Methylotrophic methanogenesis discovered in the novel archaeal phylum Verstraetearchaeota. Nat Microbiol. 2016;1:16170.
22. Borrel G, O'Toole PW, Harris HM, Peyret P, Brugère JF, Gribaldo S. Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis. Genome Biol Evol. 2013;5:1769–80.
23. Söllinger A, Urich T. Methylotrophic methanogens everywhere—physiology and ecology of novel players in global methane cycling. Biochem Soc Trans. 2019;47:1895–907.
24. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77.
25. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.
26. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol. 2016;428:726–31.
27. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019;47:D309–14.
28. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. Nucleic Acids Res. 2019;47:D427–32.
29. Haft DH, Selengut JD, White O. The TIGRFAMs database of protein families. Nucleic Acids Res. 2003;31:371–3.
30. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305:567–80.
31. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.
32. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. DbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2018;46:W95–101.
33. Coutinho PM, Deleury E, Davies GJ, Henrissat B. An evolving hierarchical family classification for glycosyltransferases. J Mol Biol. 2003;328:307–17.
34. Darling AE, Jospin G, Lowe E, Matsen FA, Bik HM, Eisen JA. PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ. 2014;2:e243.
35. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics. 2010;11:431.
36. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
37. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10:210.
38. Lartillot N, Lepage T, Blanquart S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics. 2009;25:2286–8.
39. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.
40. Miele V, Penel S, Duret L. Ultra-fast sequence clustering from similarity networks with SiLiX. BMC Bioinformatics. 2011;12:116.
41. Csurös, M. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. Bioinformatics. 2010;26:1910–2.
42. Oren, A. The family methanosarcinaceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes: other major lineages of bacteria and the archaea. Berlin: Springer; 2014. p. 259–81.
43. Ebbes M, Bleymüller WM, Cernescu M, Nölker R, Brutschy B, Niemann HH. Fold and function of the InlB B-repeat. J Biol Chem. 2011;286:15496–506.
44. Haft DH, Payne SH, Selengut JD. Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. J Bacteriol. 2012;194:36–48.
45. Porter NT, Martens EC. The critical roles of polysaccharides in gut microbial ecology and physiology. Annu Rev Microbiol. 2017;71:349–69.
46. Albers SV, Meyer BH. The archaeal cell envelope. Nat Rev Microbiol. 2011;9:414–26.
47. Ashhurst DE, Costin NM. Insect mucosubstances. III. Some mucosubstances of the nervous systems of the wax-moth (*Galleria mellonella*) and the stick insect (*Carausius morosus*). Histochem J. 1971;3:379–87.
48. Morita, RY. Bacteria in oligotrophic environments. New York, NY: Chapman & Hall; 1997.
49. Paula FS, Chin JP, Schnürer A, Müller B, Manesiotis P, Waters N, et al. The potential for polyphosphate metabolism in archaea and anaerobic polyphosphate formation in *Methanosarcina mazei*. Sci Rep. 2019;9:17101.
50. Harris RM, Webb DC, Howitt SM, Cox GB. Characterization of PitA and PitB from *Escherichia coli*. J Bacteriol. 2001;183:5008–14.
51. Poehlein A, Schneider D, Soh M, Daniel R, Seedorf H. Comparative genomic analysis of members of the genera methanosphaera and methanobrevibacter reveals distinct clades with specific potential metabolic functions. Archaea. 2018;2018:609847.
52. Borrel G, Parisot N, Harris HM, Peyretaillade E, Gaci N, Tottey W, et al. Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. BMC Genomics. 2014;15:679.
53. Hwang S, Choe D, Yoo M, Cho S, Kim SC, Cho S, et al. Peptide transporter CstA imports pyruvate in *Escherichia coli* K-12. J Bacteriol. 2018;200:e00771-17.
54. Rasmussen JJ, Vegge CS, Frøkiær H, Howlett RM, Krogfelt KA, Kelly DJ, et al. *Campylobacter jejuni* carbon starvation protein A (CstA) is involved in peptide utilization, motility and agglutination, and has a role in stimulation of dendritic cells. J Med Microbiol. 2013;62:1135–43.
55. Li Y, Leahy SC, Jeyanathan J, Henderson G, Cox F, Altermann E, et al. The complete genome sequence of the methanogenic archaeon ISO4-H5 provides insights into the methylotrophic lifestyle of a ruminal representative of the Methanomassiliicoccales. Stand Genomic Sci. 2016;11:59.
56. Kelly WJ, Li D, Lambie SC, Jeyanathan J, Cox F, Li Y, et al. Complete genome sequence of methanogenic Archaeon ISO4-G1, a member of the *Methanomassiliicoccales*, isolated from a sheep rumen. Genome Announc. 2016;4:e00221-16.
57. Maitra PK, Bhosale SB, Kshirsagar DC, Yeole TY, Shanbhag AN. Metabolite and enzyme profiles of glycogen metabolism in *Methanococcoides methylutens*. FEMS Microbiol Lett. 2001;198:23–9.

58. Santiago-Martínez MG, Encalada R, Lira-Silva E, Pineda E, Gallardo-Pérez JC, Reyes-García MA, et al. The nutritional status of *Methanosarcina acetivorans* regulates glycogen metabolism and gluconeogenesis and glycolysis fluxes. FEBS J. 2016;283:1979–99.
59. Dobrijevic D, Abraham AL, Jamet A, Maguin E, van de Guchte M. Functional comparison of bacteria from the human gut and closely related non-gut bacteria reveals the importance of conjugation and a paucity of motility and chemotaxis functions in the gut environment. PLoS ONE. 2016;11:e0159030.
60. Merhej V, Royer-Carenzi M, Pontarotti P, Raoult D. Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. Biol Direct. 2009;4:13.
61. Fricke WF, Seedorf H, Henne A, Krüer M, Liesegang H, Hedderich R, et al. The genome sequence of Methanosphaera stadtmanae reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis. J Bacteriol. 2006;188:642–58.
62. Adam PS, Borrel G, Gribaldo S. An archaeal origin of the Wood–Ljungdahl H4MPT branch and the emergence of bacterial methylotrophy. Nat Microbiol. 2019;4:2155–63.
63. Schröder I, Thauer RK. Methylcobalamin:homocysteine methyltransferase from *Methanobacterium thermoautotrophicum*. Identification as the metE gene product. Eur J Biochem. 1999;263:789–96.
64. Krone UE, McFarlan SC, Hogenkamp HPC. Purification and partial characterization of a putative thymidylate synthase from *Methanobacterium thermoautotrophicum*. Eur J Biochem. 1994;220:789–94.
65. Muller V, Blaut M, Gottschalk G. Utilization of methanol plus hydrogen by *Methanosarcina barkeri* for methanogenesis and growth. Appl Environ Microbiol. 1986;52:269–74.
66. Kato N, Yurimoto H, Thauer RK. The physiological role of the ribulose monophosphate pathway in bacteria and archaea. Biosci Biotechnol Biochem. 2006;70:10–21.
67. Welte C, Deppenmeier U. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. Biochim Biophys Acta Bioenerg. 2014;1837:1130–47.
68. Kurth JM, den Camp HJMO, Welte CU. Several ways one goal—methanogenesis from unconventional substrates. Appl Microbiol Biotechnol. 2020;104:6839–54.
69. Meuer J, Kuettner HC, Zhang JK, Hedderich R, Metcalf WW. Genetic analysis of the archaeon *Methanosarcina barkeri* Fusaro reveals a central role for Ech hydrogenase and ferredoxin in methanogenesis and carbon fixation. Proc Natl Acad Sci USA. 2002;99:5632–7.
70. Wagner T, Koch J, Ermler U, Shima S. Methanogenic heterodisulfide reductase (HdrABC-MvhAGD) uses two noncubane [4Fe-4S] clusters for reduction. Science. 2017;357:689–703.
71. Arshad A, Speth DR, de Graaf RM, Op den Camp HJ, Jetten MS, Welte CU. A metagenomics-based metabolic model of nitrate-dependent anaerobic oxidation of methane by Methanoperedens-like archaea. Front Microbiol. 2015;6:6.
72. Hedderich R, Whitman WB. Physiology and biochemistry of the methane-producing archaea. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes - prokaryotic biology and symbiotic associations. Berlin: Springer; 2013. p. 1050–79.
73. Morgavi DP, Martin C, Jouany JP, Ranilla MJ. Rumen protozoa and methanogenesis: not a simple cause-effect relationship. Br J Nutr. 2012;107:388–97.
74. Maier RJ, Olczak A, Maier S, Soni S, Gunn J. Respiratory hydrogen use by *Salmonella enterica serovar* Typhimurium is essential for virulence. Infect Immun. 2004;72:6294–9.
75. Carbonero F, Benefiel AC, Gaskins HR. Contributions of the microbial hydrogen economy to colonic homeostasis. Nat Rev Gastroenterol Hepatol. 2012;9:504–18.
76. Kalantar-Zadeh K, Berean KJ, Ha N, Chrimes AF, Xu K, Grando D, et al. A human pilot trial of ingestible electronic capsules capable of sensing different gases in the gut. Nat Electron. 2018;1:79–87.
77. Conrad R, Phelps TJ, Zeikus JG. Gas metabolism evidence in support of the juxtaposition of hydrogen-producing and methanogenic bacteria in sewage sludge and lake sediments. Appl Environ Microbiol. 1985;50:595–601.
78. Lovley DR, Goodwin S. Hydrogen concentrations as an indicator of the predominant terminal electron-accepting reactions in aquatic sediments. Geochim Cosmochim Acta. 1988;52:2993–3003.
79. Hackstein JHP. Genetic and evolutionary constraints for the symbiosis between animals and methanogenic bacteria. Environ Monit Assess. 1996;42:39–56.
80. Ametaj BN, Zebeli Q, Saleem F, Psychogios N, Lewis MJ, Dunn SM, et al. Metabolomics reveals unhealthy alterations in rumen metabolism with increased proportion of cereal grain in the diet of dairy cows. Metabolomics. 2010;6:583–94.
81. Mausz MA, Chen Y. Microbiology and ecology of methylated amine metabolism in marine ecosystems. Curr Issues Mol Biol. 2019;33:133–48.
82. King GM, Klug MJ, Lovley DR. Metabolism of acetate, methanol, and methylated amines in intertidal sediments of Lowes Cove, Maine. Appl Environ Microbiol. 1983;45:1848–53.
83. Borrel G, Joblin K, Guedon A, Colombet J, Tardy V, Lehours AC, et al. *Methanobacterium lacus* sp. nov., isolated from the profundal sediment of a freshwater meromictic lake. Int J Syst Evol Microbiol. 2012;62:1625–9.
84. Krivushin KV, Shcherbakova VA, Petrovskaya LE, Rivkina EM. *Methanobacterium veterum* sp. nov., from ancient Siberian permafrost. Int J Syst Evol Microbiol. 2010;60:455–9.
85. Wang Y, Wegener G, Williams TA, Xie R, Hou J, Tian C, et al. A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism. Sci Adv. 2021;7:eabj1453.
86. Hervé V, Liu P, Dietrich C, Sillam-Dussès D, Stiblik P, Šobotník J, et al. Phylogenomic analysis of 589 metagenome-assembled genomes encompassing all major prokaryotic lineages from the gut of higher termites. PeerJ. 2020;8:e8614.
87. Lind AE, Lewis WH, Spang A, Guy L, Embley TM, Ettema T. Genomes of two archaeal endosymbionts show convergent adaptations to an intracellular lifestyle. ISME J. 2018;12:2655–67.

ACKNOWLEDGEMENTS

We thank the computational and storage services (TARS cluster) provided by the IT department at Institut Pasteur, Paris.

AUTHOR CONTRIBUTIONS

G.B. conceived the study. G.B., C.M.T. and N.T. performed the analyses. G.B., C.M.T. and S.G. wrote the manuscript.

FUNDING INFORMATION

CMT is supported by a PhD fellowship from Paris Sorbonne Université Science and by funds from the doctoral school Bio Sorbonne Paris Cité "BioSPC". GB acknowledges funding from the French National Agency for Research Grant Methevol (ANR-19-CE02-0005-01) and SG acknowledges Archaevol (ANR-16-CE02-0005-01). This study has received funding from the French Government's Investissement d'Avenir program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant no. ANR-10-LABX-62-IBEID).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43705-021-00050-y.

Correspondence and requests for materials should be addressed to Guillaume Borrel.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Chapter 2.A: Supplementary Figures

Comparative genomic analysis of Methanimicrococcus blatticola provides insights into host-adaptation in archaea and the evolution of methanogenesis

Courtney M. Thomas[1,2], Najwa Tiab[1], Simonetta Gribaldo[1] and Guillaume Borrel[1,*]

* Corresponding author: guillaume.borrel@pasteur.fr

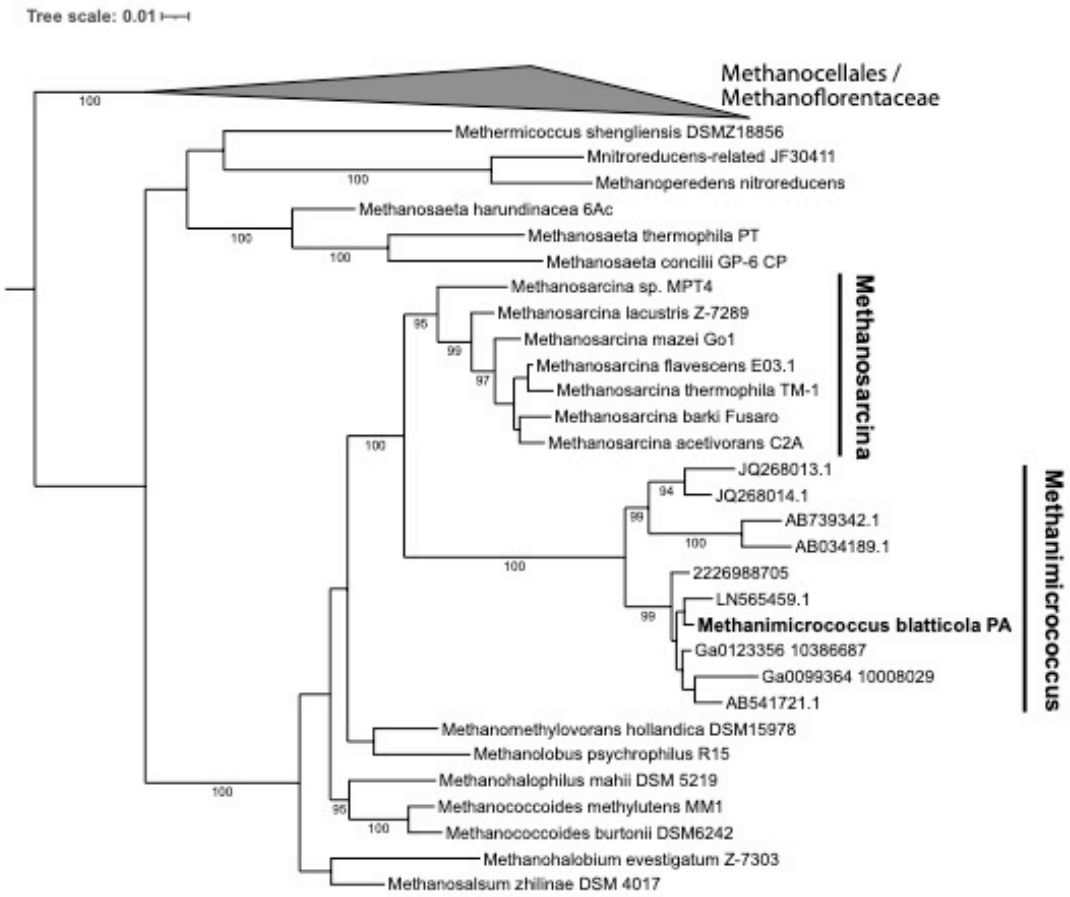


Figure S1: Phylogenetic position of Methanimicrococcus blatticola PA using 16S rRNA gene sequences (1,431 positions; Maximum likelihood inference with GTR+R3). Branch support (ultrafast bootstraps; 1000 iterations) above 90 are shown.

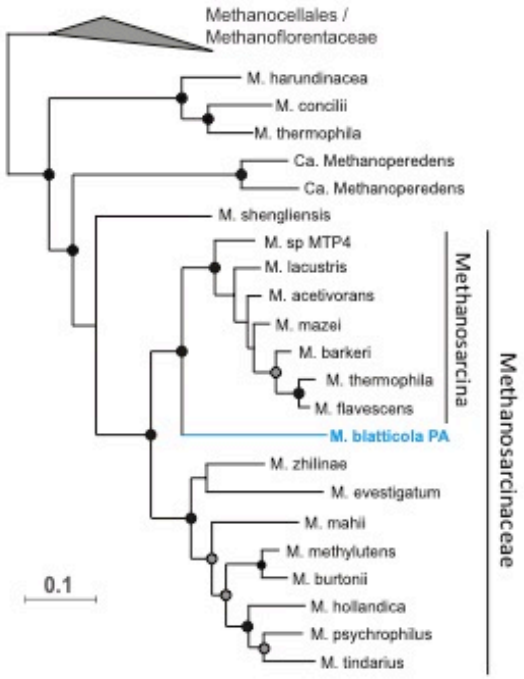


Figure S2: Maximum likelihood phylogeny (IQ-TREE, LG+F+Γ4) of the Methanosarcinales based on a concatenation of McrABG (1,176 positions). The color of the circles at the nodes indicates ultrafastbootstrap values (black, >95; grey, between 70 and 95).

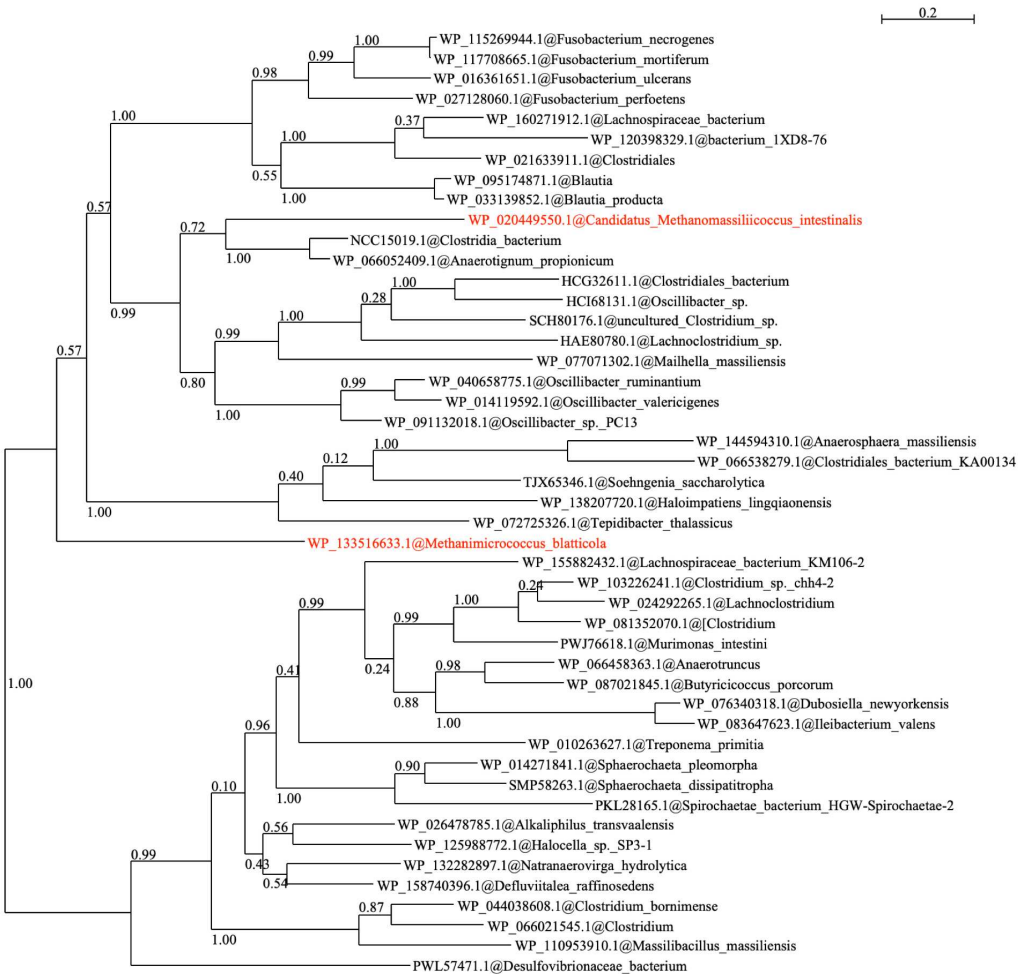


Figure S3: Maximum likelihood (LG+G4) tree of CstA (668 positions). Sequences of methanogens are in red, those of bacteria are in black.

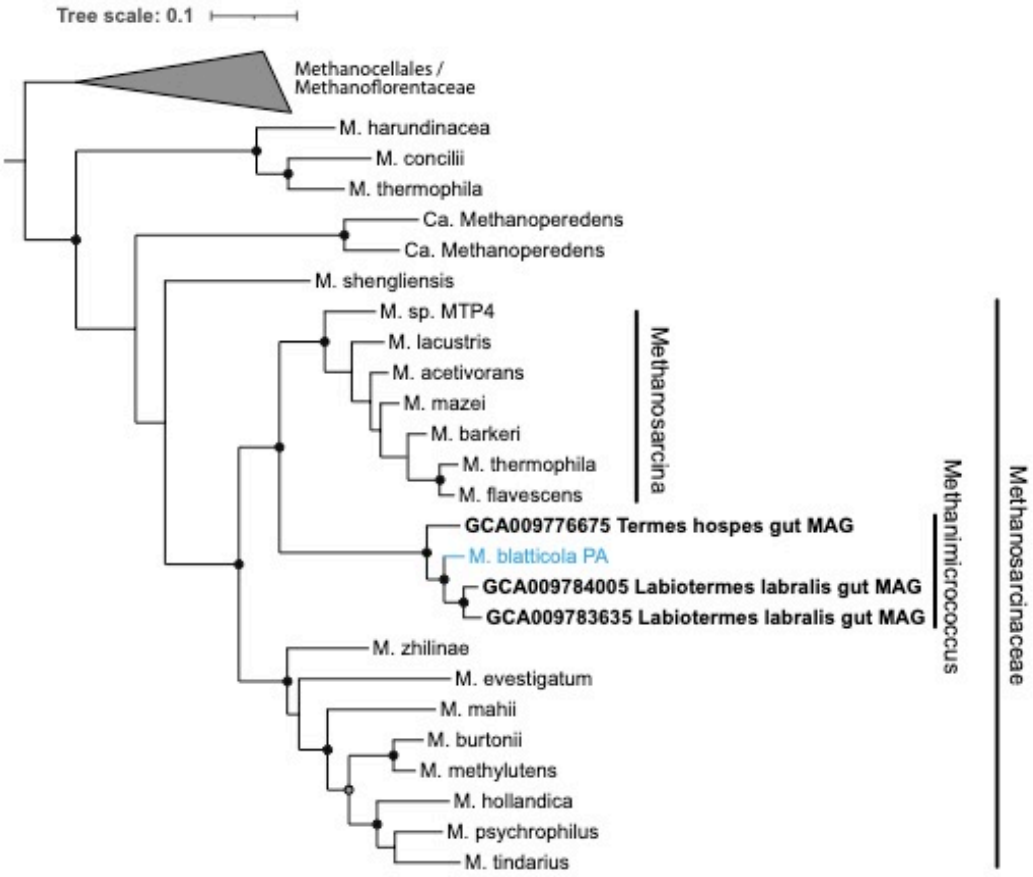


Figure S4: Maximum likelihood phylogeny (IQ-TREE, LG+F+Γ4) of the Methanosarcinales based on a concatenation of McrABG (1,233 positions), including three *Methanimicrococcus* MAGs sequenced from termite gut (bold). The color of the circles at the nodes indicates ultrafastbootstrap values (black, >95; grey, between 70 and 95).

Table S1: Genome characteristics of *M. blatticola* PA and comparison with genomes used for the analysis of gene gains and losses. Colors range from blue for the lowest values to orange for the highest ones.

NCBI TaxID	Species name	Order	Family	Contig number	Size (bp)	Protein nbr	GC%	Genome/ MAG	Isolation source[a]
91560	**Methanimicrococcus blatticola PA**	**Methanosarcinales**	**Methanosarcinaceae**	**12**	**1783575**	**1569**	**42**	**Genome**	**Periplaneta americana hindgut**
188937	Methanosarcina acetivorans C2A	Methanosarcinales	Methanosarcinaceae	1[b]	5751492	4567	42,7	Genome	Marine sediment
269797	Methanosarcina barkeri Fusaro	Methanosarcinales	Methanosarcinaceae	1[b]	4873766	3739	39,3	Genome	Lake sediment
1715806	Methanosarcina flavescens	Methanosarcinales	Methanosarcinaceae	74	3283688	2846	41,31	Genome	Biogas plant
1434111	Methanosarcina lacustris Z-7289	Methanosarcinales	Methanosarcinaceae	1[b]	4139808	3264	41,8	Genome	Lake sediment
192952	Methanosarcina mazei Go1	Methanosarcinales	Methanosarcinaceae	1[b]	4096345	3347	41,5	Genome	Sewage digester
1434100	Methanosarcina sp. MTP4	Methanosarcinales	Methanosarcinaceae	1[b]	4211468	3422	45,9	Genome	Marine sediment
523844	Methanosarcina thermophila TM-1	Methanosarcinales	Methanosarcinaceae	1[b]	3127379	2597	41,1	Genome	Sludge digestor
259564	Methanococcoides burtonii DSM 6242	Methanosarcinales	Methanosarcinaceae	1[b]	2575032	2406	40,8	Genome	Lake hypolimnion
1434104	Methanococcoides methylutens MM1	Methanosarcinales	Methanosarcinaceae	1[b]	2394636	2245	44	Genome	Marine sediment
644295	Methanohalobium evestigatum Z-7303	Methanosarcinales	Methanosarcinaceae	1[b]	2406232	2157	36,6	Genome	Saline lagoon
547558	Methanohalophilus mahii DSM 5219	Methanosarcinales	Methanosarcinaceae	1[b]	2012424	1955	42,6	Genome	Lake sediment
1094980	Methanolobus psychrophilus R15	Methanosarcinales	Methanosarcinaceae	1[b]	3072769	2841	44,6	Genome	Wetland
1090322	Methanolobus tindarius DSM 2278	Methanosarcinales	Methanosarcinaceae	1[b]	3151883	2917	39,8	Genome	Marine sediment
867904	Methanomethylovorans hollandica DSM 15978	Methanosarcinales	Methanosarcinaceae	1[b]	2714013	2281	42,6	Genome	Freswater pond sediment
679901	Methanosalsum zhilinae DSM 4017	Methanosarcinales	Methanosarcinaceae	1[b]	2138444	1955	39,2	Genome	Alkaline lake sediment
1392998	Candidatus Methanoperedens nitroreducens	Methanosarcinales	Ca. Methanoperedenaceae	72 (10 scaffolds)	3203386	3435	43,1	MAG	Bioeactor
2035255	Candidatus Methanoperedens sp. BLZ2	Methanosarcinales	Ca. Methanoperedenaceae	85	3735583	3790	40,3	MAG	Bioeactor
990316	Methanosaeta concilii GP6	Methanosarcinales	Methanosaetaceae	1[b]	3026645	2878	51	Genome	Sewage sludge
1110509	Methanosaeta harundinacea 6Ac	Methanosarcinales	Methanosaetaceae	1[b]	2571034	2385	60	Genome	Sludge blanket reactor
349307	Methanosaeta thermophila PT	Methanosarcinales	Methanosaetaceae	1[b]	1879471	1696	53,5	Genome	Thermophilic digester
1122233	Methermicoccus shengliensis DSM 18856	Methanosarcinales	Methermicoccaceae	15	1513583	1611	54,7	Genome	Offshore oil field
351160	Methanocella arvoryzae MRE50	Methanocellales	Methanocellaceae	1[b]	3179916	3058	54,6	Genome	Rice field
1041930	Methanocella conradii HZ254	Methanocellales	Methanocellaceae	1[b]	2378438	2436	52,7	Genome	Rice field
304371	Methanocella paludicola SANAE	Methanocellales	Methanocellaceae	1[b]	2957635	3004	54,9	Genome	Rice field
Na[c]	Methanoflorens stordalenmirensis	Methanoflorentales	Methanoflorentaceae	117 (10 scaffolds)	2064259	2373	51,58	MAG	Thawing permafrost

Information about *Methanimicrococcus blatticola* PA sequenced in this study are in bold. [a]Origin of the strain/enrichment culture/sample from which the genome/MAG was sequenced, [b]closed genome, [c]this MAG is available in IMG database (ID 2518645542)

Table S3: Characteristics of three recently released *Methanimicrococcus* MAGs.

NCBI TaxID	Assemby	Contig number	Size (bp)	Protein number	GC%	Comp (%)	Conta (%)	Isolation source[a]
2022468	GCA009784005	62	1688578	1532	45.5	92.14	0.00	Labiotermes labralis P3 gut compartment
2022468	GCA009783635	138	1333240	1316	45.9	84.52	2.29	Labiotermes labralis P3 gut compartment
2022468	GCA009776675	239	1397194	1511	39.6	75.09	4.58	Termes hospes gut

Comp, Completion; Conta, Contamination

Table S5: Number of proteins with a PGF-CTERM signal (likely N-glycosylated) and number of glycosyltransferase families in the Methanosarcinales members. The PGF-CTERM signal was first identified using TIGR04126 hmm profile, and manually refined profile targeting more specifically the Methanosarcinales sequences. Glycosyltransferases families were retrieved from CAZy database (http://www.cazy.org/), with the exception of M. *blatticola* PA, *Methanolobus tindarius* DSM 2278, *Methermicoccus shengliensis* DSM 18856 and the two '*Ca*. Methanoperedens' MAGs, that were annotated using dbCAN2. Colors range from blue for the lowest values to orange for the highest ones.

Species	Taxon family	PGF-CTERM proteins	Glycosyltransferases families												Total GTs
			1	2	3	4	20	35	66	81	83	84	94	NC	
Methanimicrococcus blatticola PA	**Methanosarcinaceae**	**5**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**0**	**2**
Methanosarcina acetivorans C2A	Methanosarcinaceae	44	3	13	1	20	0	1	4	0	0	0	0	0	42
Methanosarcina barkeri Fusaro	Methanosarcinaceae	34	2	21	1	23	0	2	3	0	0	0	0	0	52
Methanosarcina flavescens	Methanosarcinaceae	30	3	17	1	20	0	1	3	0	0	0	0	0	45
Methanosarcina lacustris Z-7289	Methanosarcinaceae	19	2	20	1	31	0	1	2	0	0	0	0	0	57
Methanosarcina mazei Go1	Methanosarcinaceae	29	2	23	1	26	0	1	3	0	0	0	0	0	56
Methanosarcina sp. MTP4	Methanosarcinaceae	27	3	10	1	14	0	1	2	0	0	0	0	0	31
Methanosarcina thermophila TM-1	Methanosarcinaceae	13	1	18	1	21	0	1	3	0	0	0	0	0	45
Methanococcoides burtonii DSM 6242	Methanosarcinaceae	24	1	10	1	8	0	0	1	1	0	0	0	0	22
Methanococcoides methylutens MM1	Methanosarcinaceae	21	1	5	1	8	0	0	1	1	0	0	0	0	17
Methanohalobium evestigatum Z-7303	Methanosarcinaceae	9	1	5	0	8	0	1	1	1	0	0	0	0	17
Methanohalophilus mahii DSM 5219	Methanosarcinaceae	16	1	5	1	3	0	0	1	1	0	0	0	0	12
Methanolobus psychrophilus R15	Methanosarcinaceae	19	4	13	1	14	0	1	0	2	0	0	0	0	35
Methanolobus tindarius DSM 2278	Methanosarcinaceae	32	2	5	1	8	0	1	1	1	1	0	0	1	21
Methanomethylovorans hollandica DSM 15978	Methanosarcinaceae	6	0	7	1	18	0	1	2	0	0	0	0	0	29
Methanosalsum zhilinae DSM 4017	Methanosarcinaceae	18	0	2	1	3	0	1	1	2	0	0	0	1	11
Candidatus Methanoperedens nitroreducens	Methanoperedentaceae	26	0	23	0	17	1	0	2	0	6	0	0	1	50
Candidatus Methanoperedens sp. BLZ2	Methanoperedentaceae	25	0	27	0	12	1	0	2	0	2	0	1	2	47
Methanosaeta concilii GP6	Methanosaetaceae	2	1	16	1	15	0	2	2	0	0	1	0	1	39
Methanosaeta harundinacea 6Ac	Methanosaetaceae	1	0	8	1	9	0	4	4	0	0	0	0	0	26
Methanosaeta thermophila PT	Methanosaetaceae	2	0	7	1	10	1	1	3	0	0	0	0	1	24
Methermicoccus shengliensis DSM 18856	Methermicoccaceae	7	0	4	0	14	1	0	1	1	1	0	0	0	22

Glycosyltransferases families were retrieved from CAZy database (http://www.cazy.org/), with the exception of Methanimicrococcus blatticola, Methanolobus tindarius DSM 2278, Methermicoccus shengliensis DSM 18856 and the two Methanoperedens MAGs, that were annotated using dbCAN2.

Table S6: Losses of "MCR-associated markers" in *Methanimicrococcus blatticola* PA and comparison with the prevalence of these genes in other gut-associated lineages and their phylogenetically closest non-gut relatives. "*Ca.* Methanomassiliicoccus intestinalis" (M. int.) is a Methanomassiliicoccaceae that is only present in the human gut. Mbac: Methanobacteriales; Mmass: Methanomassiliicoccales.

		Gut-associated	Nbr of genomes / lineage	m1_mcrA	m2_mcrB	m3_mcrG	m4_Predicted rotamase	m5_DUF2102	m6_DUF2112	m7_YjiL-like	m8_DUF2113	m9_arCOG03226	m10_atwA	m11_mcrC	m12_mcrD	m13_cfbD	m14_cfbE	m15_mtxX	m16_AIR-synthase-related	m17_Zn-ribbon protein	m18_DUF2099	**m19_DUF2117**	m20_Thioglicine modif (McrA PTM)	**m21_DUF2124**	m22_DUF2098	m23_Soluble P-type ATPase	**m24_DUF2111**	m25_arCOG04853	m26_DUF2114	m27_mtrA	m28_mtrB	m29_mtrC	m30_mtrD	m31_mtrE	m32_arCOG04885	m33_Arg-methyltransf (McrA PTM)	m34_DUF1894	m35_DUF1890	m36_DUF2115	m37_DUF2119	m38_DUF2121
Methanosarcinaceae	**Methanimicrococcus blatticola**	Y	1	MBLAb1417	MBLAb1421	MBLAb1418	MBLAb0808	MBLAb0807	MBLAb0806	MBLAb0805	MBLAb0804	MBLAb0803	MBLAb0809	MBLAb1419	MBLAb1420	MBLAb0389	MBLAb0388	MBLAb1464	MBLAb1101	MBLAb0230	MBLAb0116	MBLAb1068	-	MBLAb0762	-	MBLAb1428	MBLAb0810	-	MBLAb1151	MBLAb1427	MBLAb1241	MBLAb1240	MBLAb1239	MBLAb1238	MBLAb1237	MBLAb0561	MBLAb1414	MBLAb1513	-	-	-	-
	Methanosarcina	N	23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**	1	**1**	1	1	**1**	1	1	1	1	1	1	1	1	1	1	0	1	1
	Other Methanosarcinaceae	N	38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.97	1	1	1	0.92	1	0.87	1	1	**1**	1	1	1	1	1	1	1	1	1	1	0.67	0.21	0.9	0.67
	Methanosaetaceae	N	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**	1	1	1	1	1	1	1	0.66	1	1	1	0	0	1
Mbac	**Methanobrevibacter**	Y	17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.94	1	1	**0.12**	1	1	**0**	1	1	1	1	1	1	1	0.94	1	1	1	1	0.94	1
	Methanosphaera	Y	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**0**	1	1	**0**	1	1	1	1	1	1	1	1	1	0	0	1	0	0
	Methanobacterium	N	9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**	1	1	**1**	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Methanothermobacter	N	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**	1	1	**1**	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Mmass	**Methanomethylophilaceae**	Y	21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**0**	1	**0**	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Methanomassiliicoccus intestinalis	Y	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**0**	0.93	**0**	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Methanomassiliicoccaceae (no M. int.)	N	6	1	1	1	1	1	1	1	1	1	1	1	1	0.5	0.5	1	1	1	0.83	**1**	1	**0.3**	1	0	1	0	0	0.67	0	0	0	0	0	0	0	0	0	0	0

Chapter 2.B: Large-Scale Comparative Genomic analysis of Host-associated Archaeal Genomes

Current understanding of how archaea have adapted to the intestinal tract of animals is largely based on a handful of studies, most of which have focused on a single lineage of archaea and a small number of archaeal genomes. While these efforts have identified several host-associated genomic adaptations in the archaea, as discussed in Chapter 2.A, a large-scale study that includes all the lineages of intestinal archaea has been missing. Indeed, until recently the number of available intestinal archaeal genomes came from a narrow range of hosts and were available from only a limited diversity of intestinal methanogens. To address this, I started performing a comparative genomic analysis based on archaeal genomes associated with a broader host range. Quantitative PCR analyses (presented in Chapter 1.A) allowed to us identify 64 animal fecal samples with a high portion of methanogens. These samples were then subjected to shotgun sequencing. The obtained reads, as well as those from previously published metagenomic studies, were processed using the Let-It-Bin pipeline (Borrel *et al.*, 2019) to obtain host-associated archaeal genomes.

I. Large-scale comparative genomic analysis of host-associated archaeal genomes

Up to now, most of the gut-associated archaeal genomes/MAGs were derived from humans, ruminants (mainly cattle) and to lesser extent, termites (**Chapter 2.B Figure 1A**). Indeed, public databanks currently contain only nine MAGs/genomes of intestinal archaeal associated with animals that are not humans, ruminants, or termites (**Chapter 2.B Figure 1B**). The shotgun metagenomic sequencing from this work added 218 archaeal MAGs from 12 orders of mammals, as well as reptiles, birds, and amphibians (**Chapter 2.B Figure 1**). These MAGs were estimated to have at least 50% of completeness and less than 5% of contamination using CheckM (Parks *et al.*, 2015). Among them, 191 were placed in a reference phylogeny as they contained at least 50% of the necessary marker genes necessary for building a robust phylogeny **(Chapter 2.B Figure 2)**.

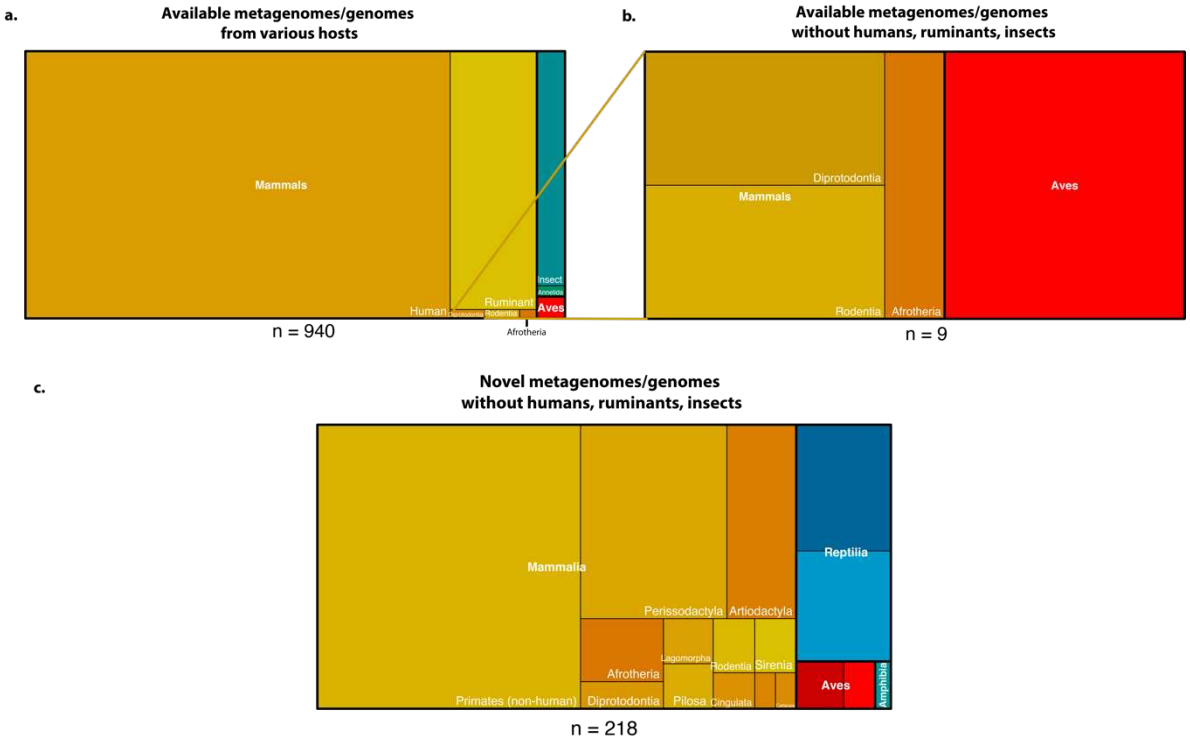


Figure 1. Tree maps showing the distribution of hosts for which archaeal genomes/MAGs are available as of the summer of 2021. **a.** Available MAGs/genomes from hosts before this work, including the most sampled groups of animals: humans, ruminants, and insects. **b.** Same as panel "a" but without humans, ruminants, and insects. **c**. Novel metagenome/genomes made available by this work.

These MAGs were included in a database with a diverse range of reference genomes from non-gut environments and host. The 191 novel high-quality archaeal MAGs were placed in a phylogeny including the reference genomes. For this purpose, HMM profiles were used to search for 41 marker genes that are universally present in archaeal genomes as described in (Adam *et al.*, 2017). These markers were aligned, trimmed, and concatenated to build a character supermatrix that was used to build a Maximum Likelihood phylogenetic tree using the LG+C60+F+I model **(Chapter 2.B Figure 2)**. In addition to expanding the diversity of animals from which these archaeal MAGs were obtained, this analysis also resulted in the first MAG from a member of Bathyarchaeota collected from a reptile **(Chapter 2.B Figure 2)**. Interestingly, this MAG is the closest relative to the recently sequenced Bathyarchaeota MAGs from the termite gut (Hervé *et al.*, 2020; Loh *et al.*, 2021), supporting the existence of a host-associated clade of Bathyarchaeota having specific adaptations to this niche.

Figure 2. Placement of our 191 newly reconstructed host-associated MAGs in the reference phylogeny of Archaea. Phylogeny (IQTREE, LG+C60+F+I) based on concatenation of 41 conserved phylogenetic markers (6,838 amino acid positions) from 658 genomes/MAGs. Colored genome names indicate the most abundant host-associated clades of archaea. The first layer on the tree indicates the location from which the genomes/MAGs were collected. The second and the third layers indicate the animal classes and the mammal orders from which genomes/MAGs were collected, respectively. Stars on the third layer indicate the novel MAGs that were assembled in this work.

The other MAGs belong to Methanobacteriales (52), Methanomassiliicoccales (98), Methanomicrobiales (34) and Methanosarcinales (6) **(Chapter 2.B Figure 2)**. In the last two orders, there was previously only one and four host-associated archaeal genome/MAGs available, respectively.

A comparative genomic analysis of the genomes/MAGs in the host-associated clades relative to the genomes/MAGs of archaea only found in the environment is expected to reveal new

information concerning how these groups have adapted to the host intestinal tract. More specifically, it will be possible to determine if the genomic adaptations previously identified in the Methanobacteriales, Methanomassiliicoccales, and now *Methanimicrococcus* (Samuel, 2007; Hansen, 2011; Lurie-Weinberger, 2012; Lurie-Weinberger, 2012; Borrel, 2014; Thomas, 2021) are present in other host-associated archaea like the Methanomicrobiales and the Bathyarchaeota. This would mean that there are even more widespread and convergent adaptations to the intestinal microbiome in the Archaea.

A large-scale comparative genomic analysis would also reveal unique adaptations of certain lineages of archaea to their specific hosts - i.e., archaea that live in the mammalian versus the reptile intestinal tracts. In Chapter 1.A, ASVs placed in a reference 16S rRNA gene tree for each methanogenic order formed host-specific clades, suggesting that genetically, host-associated methanogens may have preferences for specific hosts. For example, within the Methanomicrobiales there are two distinct clades – one with Methanomicrobiales from perissodactyls and tortoises and the other is a mixed clade with ASVs coming from various mammals, birds, and reptiles **(Chapter 1.A Supplemental Figure 4)**. This was also seen when placing the novel Methanomicrobiales MAGs in a reference phylogeny **(Chapter 2.B Figure 2)**. Thus, it would be particularly interesting not only to compare the genomic content of environmental and intestinal Methanomicrobiales but also to compare the genomes of the Methanomicrobiales from each host-associated clade (perissodactyls and tortoises and the mixed animal clade), to highlight possible unique adaptations between the two clades.

Discussion & Perspectives

This work has provided key insights into which archaea are present in the animal intestinal microbiome, including what decides how many and which archaeal taxa are present, and some of the genomic adaptations related to the colonization of this unique environment (e.g., adhesin-like proteins). The following sections expand upon some of the major results presented in Chapter 1 and Chapter 2 and discuss perspective lines of research within the field. Major questions discussed in this section include: (i) how do methanogens interact with bacteria and their host; (ii) what is the impact of captivity on the intestinal archaeome; (iii) how does the archaeome change across various time scales (i.e. seasonally, within lifetimes, and across generations); (iv) is there a shared evolutionary path of adaptation to the intestinal environment in the Archaea; and (v) can the hypothesized interactions between microbes be validated in the laboratory.

I. The dominant intestinal archaea

I.a Methanogens are the most abundant intestinal archaea

We found that most abundant archaea across animals belong to five genera/families of methanogenic archaea – *Methanobrevibacter*, *Methanosphaera*, "*Ca*. Methanomethylophilaceae", *Methanimicrococcus*, and *Methanocorpusculum*. This is a surprising result as there are a number of mesophilic archaea that are capable of diverse carbohydrate metabolisms which are never detected in the intestinal tract of animals (Borrel *et al.*, 2020). What has led to the dominance of these specific groups of methanogens in the intestinal tract of animals while other lineages of archaea are never detected in this environment?

Firstly, the archaea that are not detected in the intestine are not commonly in contact with animals due to the unique environments that they have colonized such as extreme salt lakes, deep ocean sediments, etc. Evolutionarily, it has been proposed that a common trait of

archaea is an adaptation to chronic energy stress which has allowed them to out-compete bacteria in these environments (Valentine, 2007). However, these adaptations may have also resulted in most archaea being poorly suited for inhabiting host-associated niches.

This said, the Archaea that are found in the intestine have acquired several distinct genomic adaptations that cumulatively allow them to thrive in the gut. As previously discussed, intestinal methanogens have reduced genomes (Borrel *et al.*, 2014; Lang *et al.*, 2015; Söllinger *et al.*, 2016; de la Cuesta-Zuluaga *et al.*, 2021; Thomas *et al.*, 2021) as well as unique repertories of cell surface variation genes like adhesin-like proteins (ALP) and glycosyltransferases (GT) (Samuel *et al.*, 2007; Lurie-Weinberger, Peeri and Gophna, 2012; Lurie-Weinberger, Peeri, Tuller, *et al.*, 2012; Kelly, Li, *et al.*, 2016; Thomas *et al.*, 2021). Additionally, the removal of H_2 by hydrogenotrophic methanogens allows hydrogen-producing fermentative bacteria to efficiently continue their metabolisms (Borrel *et al.*, 2020); making these archaea important members of the microbial food web in the intestinal tract.

I.b Host-adapted non-methanogenic archaea?

The detection of aerobic ammonia-oxidizing Thaumarchaeota in all classes of animal examined in Chapter 1 was surprising. It is currently unclear from these data whether the Thaumarchaeota are resident members of the intestinal microbiome. Moreover, no genome of an intestinal Thaumarchaeota has been sequenced so far.

However, a genome of "*Ca.* Nitrosocosmicus" (Thaumarchaeota) collected from a wastewater treatment plant is available (Sauder *et al.*, 2017). This genus is closely related to the most prevalent Thaumarchaeota ASV from Chapter 1.A. A recent study of the genome of "*Ca*. Nitrosocosmicus hydrocola[1]", revealed that it has genes coding for oxidative stress (Sauder *et al.*, 2017), like the insect-associated *Methanobrevibacter* (Hackstein *et al.*, 1994; Leadbetter *et al.*, 1996, 1998). "*Ca*. Nitrosocosmicus hydrocola" and the closely related "*Ca*. Nitrosocosmicus franklandus" (from soil) both lack the genes encoding for chemotaxis and motility (Sauder *et al.*, 2017; Nicol *et al.*, 2019), similar to *Methanimicrococcus blatticola*

[1] Previously published as **"*Candidatus* Nitrosocosmicus exaquare"**, Sauder et al. 2017

(Thomas *et al.*, 2021). A more distant relative to the Thaumarchaeota ASVs detected in Chapter 1.A, "*Ca.* Nitrosotenuis cloacae" (also collected from a wastewater treatment plant), lacks the genes necessary for producing osmolyte transporters, which are usually detected in the genomes of free-living ammonia oxidizing archaea (Spang *et al.*, 2012; Yamamoto *et al.*, 2012; Santoro *et al.*, 2015). Taken together, these distinct genomic characteristics could suggest that these Thaumarchaeota have started to adapt to nutrient rich environments, like the gut. It is also interesting to note that the Thaumarchaeota were more abundant in carnivorous animals, which could be linked to a high protein diet that results in an increase of available nitrogen in the intestinal tract (Silvester *et al.*, 1995; Russell *et al.*, 2003; Reese *et al.*, 2018). The Thaumarchaeota could also inhabit other locations in the intestinal tract, besides the lower sections, however, here a majority of the samples examined were faeces. Therefore, it is possible that there are aerobic niches elsewhere along the gastrointestinal tract that maybe better suited for these microbes. Altogether, this leads to the hypothesis that Thaumarchaeota representatives found in the gut are indeed resident members of this environment.

An alternative hypothesis may explain the nearly ubiquitous presence of the Thaumarchaeota in fecal samples. These sequences which are closely related to dominant soil phylotypes of the Nitrososphaerales and may therefore have been picked up by animals from grooming habits, diet, or other soil interactions. This could be a mode of dispersal for these microbes in the environment. A similar process occurs with seeds that are ingested by animals (Schupp, 1993). Obtaining data on the viability of Thaumarchaeota detected in the intestinal tract may allow to test this hypothesis.

The Bathyarchaeota have recently been identified in soil-feeding termites (Hervé *et al.*, 2020; Loh *et al.*, 2021), a tortoise, a parrot, and an armadillo (Chapter 1.A). These archaea form a distinct termite-specific cluster (Hervé *et al.*, 2020), suggesting that they have specifically adapted to this unique ecosystem. In termites, they are predicted to be homoacetogenic using H_2 and either CO_2 or methyl-compounds to perform their metabolism (Loh *et al.*, 2021). However, further genomic adaptations to the intestinal tract have not yet been identified in the Bathyarchaeota, and would be interesting to investigate. More specifically, it would be of

interest to determine if the Bathyarchaeota have acquired similar adaptations to the methanogens while colonizing the intestine.

II. Widespread co-occurrence of methanogens and Firmicutes

Most of the co-occurrences identified in Chapter 1 are between methanogens and bacteria belonging to the Firmicutes phylum. Methanogen-Firmicutes relationships were previously identified in humans and ruminants (Chassard *et al.*, 2010; Arumugam *et al.*, 2011; Kittelmann *et al.*, 2013; Ruaud *et al.*, 2020), but it was remarkable to find that they also occur across the entire mammalian class and even in some reptiles. Those associations are not restricted to a particular group of archaea but involve members of *Methanobrevibacter*, "*Ca.* Methanomethylophilaceae", *Methanocorpusculum* and *Nitrososphaeraceae*. What are the drivers of these associations?

II.a Positive and negative associations between Methanobrevibacter and Christensenellaceae members

In the individual orders of mammals that were targeted for co-occurrence analyses (Carnivora, Perissodactyla/Cetartiodactyla, Primates), most of the relationships we identified occurred between the Methanobacteriales and Firmicutes (14/24). Among these fourteen cases was the previously identified *Methanobrevibacter – Christensenellaceae* co-occurrence (Ruaud *et al.*, 2020). *Christensenellaceae* can support the growth of *M. smithii* through a hydrogen-based syntrophic association (Ruaud *et al.*, 2020). Interestingly, among data gathered from Perissodactyla/Cetartiodactyla, there are both positive and negative co-occurrences between *Methanobrevibacter* and *Christensenellaceae* OTUs. An OTU that is closely related to *Methanobrevibacter ruminantium* was negatively correlated to the presence of a *Christensenellaceae* OTU (OTU_2376) while a *Methanobrevibacter wolinii* OTU was positively correlated to a different *Christensenellaceae* OTU (OTU_503). This suggests that the association between members of these groups is highly specialized. It would be interesting to co-culture these different species of *Methanobrevibacter* with various members of the *Christensenellaceae* to further understand the nature of these relationships. This will first

require the cultivation and characterization of a larger range of *Methanobrevibacter* and *Christensenellaceae* species.

II.b "Ca. Methanomethylophilaceae" are related to methyl-compound producing bacteria

"Ca. Methanomethylophilaceae*"* were found to be positively related to several Firmicutes shown to produce useful metabolites for their mode of methanogenesis, like *L. pectinoschiza* and *Sarcina spp.* in mammals, and in reptiles *Methanomassiliicoccaceae* positively co-occurs with *Clostridium spp*. Both *Sarcina spp.* and some *Clostridium spp*. have been linked to trimethylamine (TMA) production in humans (Fennema *et al.*, 2016). TMA is made through the bacterial degradation of components found in some fruit, eggs, soy, and red meat (Zhang *et al.*, 1999; Zeisel *et al.*, 2003). Interestingly, *"Ca.* Methanomethylophilaceae*"* also co-occur with a member of the *Muribaculaceae* family (Bacteroidetes phylum), previously known as S24-7 or *Homeothermaceae*. These bacteria are almost exclusively found in the intestine of animals (Ormerod *et al.*, 2016; Lagkouvardos *et al.*, 2019), and some have been predicted to degrade plant glycans like hemicellulose and pectin (Ormerod *et al.*, 2016). The identification of methyl-reducing methanogens with bacteria able to provide them with methyl-compounds supports the hypothesis that methyl-reducing methanogens may not be limited by hydrogen availability in the intestine but by the availability of methyl groups (Feldewert *et al.*, 2021). If hydrogen is the only basis of these archaea-bacteria relationships, then one would expect to find co-occurrence relationships similar to those observed with CO_2-reducing hydrogenotrophic methanogens, but this is not the case.

II.c Methanocorpusculum co-occurs with hydrogen-producing bacteria

In Perissodactyla/Cetartiodactyla, a *Methanocorpusculum* (Methanomicrobiales) OTU positively co-occurs with a *Roseburia* OTU (phylum Firmicutes, class Clostridia). In the human intestine, *Roseburia* is viewed as beneficial because it produces butyrate through the fermentation of plant cell wall polysaccharides (Duncan *et al.*, 2006; Louis *et al.*, 2017; La Rosa *et al.*, 2019). This process releases large amounts of hydrogen, which likely plays a role in the interaction between CO_2-reducing hydrogenotrophic *Methanocorpusculum* and *Roseburia*.

In primates, a *Methanocorpusculum* OTU significantly co-occurs with an OTU belonging to the *Ruminococcaceae* family, a common constituent of the intestinal microbiota of humans and primates also commonly associated to fibre fermentation (Barelli *et al.*, 2015; Katherine R. Amato *et al.*, 2019). The fact that *Methanocorpusculum* is significantly correlated to the presence of beneficial fibre-fermenting intestinal Firmicutes highlights the potential importance of *Methanocorpusculum* in the intestinal tract of these animals, which play an important role in keeping the hydrogen levels low.

It is also interesting to observe these co-occurrence trends between *Methanocorpusculum* and fibre degrading bacteria because the abundance of *Methanocorpusculum* is positively correlated to the host diet fibre content (Chapter 1.A). It is likely that increased fibre consumption results in an increase in H_2 production by fibre fermenting bacteria, like the *Roseburia* and *Ruminococcaceae*, which is beneficial for CO_2-reducing hydrogenotrophic *Methanocorpusculum*.

Taken together, the co-occurrences identified in Chapter 1.A suggest that CO_2-reducing hydrogenotrophic methanogens are establishing bacteria-archaea relationships based on H_2 production; while methyl-reducing hydrogenotrophic methanogens are forming relationships based on the generation of methyl-compounds.

This is further supported by the positive correlation between the sulfate reducing bacteria (SRB) and Methanomassiliicoccales discussed in Chapter 1.B. The SRB are major consumers of H_2 in the intestinal tract, and the dominant SRB, *Desulfovibrio spp*. have been suggested to produce TMA in the intestine (Fennema *et al.*, 2016). Moreover, both Methanomassiliicoccales and SRB can use H_2 at a much lower concentration than CO_2-reducing hydrogenotrophic methanogens (Lovley *et al.*, 1988; Feldewert *et al.*, 2021). These two aspects could explain why there is a positive correlation between the Methanomassiliicoccales and the SRB, and conversely, why there is no (positive/negative) correlation between the CO_2-reducing hydrogenotrophic methanogens and the abundance of SRB. Interestingly, we found that there is no positive correlation between the *Methanimicrococcus* and SRB abundance. Thus, the apparent relationship between SRB and

the Methanomassiliicoccales is likely complex, or is only the reflect of independent factors having similar influence on their abundance, and warrants further investigation.

III. Interaction with hosts, importance of cell surface variations

Archaeal-host interactions are overall poorly understood, especially in the intestinal microbiota. It has been shown that *Methanosphaera stadtmanae* have a stronger immunogenic effect than *M. smithii* (Bang *et al.*, 2014). Moreover, *M. stadtmanae* is rapidly phagocytosed by dendritic cell while *M. smithii* is not (Bang *et al.*, 2014). Beyond this, a specific analysis of a direct interaction between animal cells and intestinal archaea has not been carried out. However, it is likely that a part of the interaction between animal cells and archaeal cells occurs through cell surface glycoproteins. This is why, genes involved in cell surface variations, including the synthesis of glycoconjugates, are often used as predictors of putative host-microbe interactions (Tuson *et al.*, 2013; Varki, 2017).

III.a Host and Diet derived glycans exert a selective pressure on the intestinal microbiota

Animal genes involved in the production of glycans (sugar structures that coat the outside of cells) exert selective pressure on intestinal microbes and thus are thought to be constantly adapting to select for the most beneficial intestinal microbes (Hooper *et al.*, 2001; Dethlefsen *et al.*, 2006). This selection is possible because bacteria produce adhesins that bind to host glycans. Indeed, some epithelial glycans produced by hosts are even microbe dependent, so they are not expressed without the right microbes being present (Hooper *et al.*, 2001). For example, in mice, the common intestinal bacterium *Bacteroides thetaiotaomicron* induces the production of fructose bearing epithelial glycans that are used by the bacterium as a carbon source (Hooper *et al.*, 2001). When mice start to wean, these fructose-containing epithelial glycans are produced at low levels in the intestine, allowing the early colonization of *B. thetaiotaomicron*. Once the bacterial population grows, the production of epithelial glycans increases, which helps the *B. thetaiotaomicron* population outcompete other bacteria. Glycan production is thought to be an evolutionary characteristic of animals aimed at helping the selection of positive host-microbe interactions, i.e., commensal/beneficial microbes get to

bind and interact with host cells while 'bad' microbes are not able to recognize these glycans and cannot bind/interact with host cells (Hooper *et al.*, 2001). In general, bacteria have adapted either to use a wide range of host-glycans or a very narrow/specific range (Varki, 2017). Currently, it is not clear where archaea fit into these interactions, as the most commonly encountered archaea in the intestine, the methanogens, do not seem to have the genomic potential to degrade glycans (Coutinho *et al.*, 1999; Nkamga *et al.*, 2017). However, genomic adaptations suggest that the intestinal archaea interact with host in a different way than bacteria. This is discussed in the following section.

III.b Glycans as cell surface decorations of methanogens

Methanogens have genes coding for distinct glycosyltransferases (GTs) that belong to GT families responsible for transferring glycans to other sugars and to cell surface proteins (Samuel *et al.*, 2007; Guerry, 2011; Lurie-Weinberger, Peeri and Gophna, 2012; Jarrell *et al.*, 2013; Kandiba *et al.*, 2013). Thus, while intestinal archaea cannot degrade glycans in the gut they can utilize modified sugars to modify their cell surfaces. Interestingly, *M. smithii* is able to decorate its cell with a number of sugars, one of which is pseudaminic acid (Lewis *et al.*, 2009), a host-generated glycan commonly found in the intestine (Samuel *et al.*, 2007; Kandiba *et al.*, 2013). Even though the ability to produce pseudaminic acid is present in a number of environmental archaea, most intestinal archaea do not have the genomic potential to produce it (Kandiba *et al.*, 2013). It is possible that this ability was retained by the Methanobacteriales and has aided their colonization of the intestinal tract by allowing them to mimicry host-glycans and evade immune system detection. However, to-date this hypothesis has not been verified and will have to be thoroughly investigated.

Other GTs in intestinal archaeal genomes originated from horizontal gene transfers from bacteria (Lurie-Weinberger, Peeri and Gophna, 2012). Paradoxically, as discussed in Chapter 2 (Thomas *et al.*, 2021), many intestinal methanogens have a reduced number of GTs relative to environmental lineages. However, this further highlights the importance of the GTs that intestinal archaea have conserved or acquired, because it suggests that they have played an important role in the colonization of the intestinal tract. Moreover, a unique repertoire of GTs

could imply a specialization to specific hosts' intestinal microbiota, a hypothesis that will be interesting to test experimentally.

III.c Adhesin-like proteins

Additionally, adhesin-like genes are relatively enriched in the intestinal methanogens' genomes (Samuel *et al.*, 2007; Leahy *et al.*, 2010; Lurie-Weinberger *et al.*, 2012; Kelly, Li, *et al.*, 2016; Kelly, Pacheco, *et al.*, 2016; Ng *et al.*, 2016). In the *Methanimicrococcus blatticola* genome, there is an enrichment of putative ALPs containing the Flg_new domain. These proteins are absent from environmental relatives (Thomas et al., 2021), but have been detected in other intestinal methanogens (Borrel *et al.*, 2017). This suggests that these genes are involved in the process of intestine colonization. Some of these ALPs have been proposed to play a role in attachment to host cells by the Methanomassiliicoccales and intestinal bacteria (Borrel *et al.*, 2020). In *Methanobrevibacter*, many of the ALPs were horizontally acquired from bacteria (Hansen *et al.*, 2011; Lurie-Weinberger, Peeri and Gophna, 2012). In human twins, it was shown that different strains of *M. smithii* have acquired different types of adhesin-like proteins from intestinal bacteria, and can regulate their expression in a way that would allow *M. smithii* to adapt to an increased number of niches by forming unique bacterial-archaeal and host-archaeal relationships (Hansen *et al.*, 2011). Further, rumen methanogens like *Methanobrevibacter ruminantium* have been shown to use their adhesins to interact with rumen protozoa (Ng *et al.*, 2016). This species of *Methanobrevibacter* is predicted to have the largest number of ALPs (105) among the intestinal Methanobacteriales (Leahy *et al.*, 2010). In co-culture experiments, several of these ALPs were significantly upregulated and helped *M. ruminantium* attach to the bacterium with which it was co-cultured (Leahy *et al.*, 2010). Clearly, these results show that ALPs can aide in microbe-microbe interactions within the intestinal microbiota.

IV. Impact of captivity on the intestinal microbiome

This thesis work was primarily conducted using captive animal fecal samples, but does the captive animal's archaeal community reflect what is found in wild animals? And more generally, what is the current understanding of the impact of captivity on the intestinal microbiota?

A systematic analysis comparing the intestinal archaeome between captive and wild animals is currently not available. However, a recent study on bioRxiv did not find captivity to significantly impact the beta diversity of the intestinal archaeome (Youngblut *et al.*, 2020) – even though, this study was not specifically designed with this comparison in mind. It was also interesting to see that some of the results from this study support what we present in Chapter 1, such as the influence of host phylogeny and diet on the overall structure of the intestinal archaeal community, as well as a high relative abundance of Methanomicrobiales in the Perissodactyla.

This study also observed differences in the archaeal community between captive and wild animals. More specifically in black rhinos (Perissodactyla), wild individuals were shown to host relatively higher concentrations of *Methanocorpusculum*, while captive individuals hosted more *Methanobrevibacter* (Gibson *et al.*, 2019). This is of interest as in our results discussed in Chapter 1, the Perissodactyla consistently hosted high concentrations of *Methanocorpusculum*, despite being in captivity. Thus, it would be useful to use archaea-specific 16S rRNA gene primers in both captive and wild Perissodactyla to determine how captivity impacts the overall community of archaea in this order of animals.

In another study of rock ptarmigans, a species of artic bird, the abundance of archaea has been shown to be higher in wild individuals than in captive ones (Salgado-Flores *et al.*, 2019). Here, Methanomassiliicoccales from the family *Methanomassiliicoccaceae* was the dominant methanogenic lineage, possibly due to the increased consumption of pectin in wild rock ptarmigans (Salgado-Flores *et al.*, 2019). Additionally, a study in howler monkeys found *Methanosphaera stadtmanae* to be the most abundant methanogen in captive species, but it was largely absent from wild monkeys (Nakamura *et al.*, 2011). Another study in howler monkeys found that the absolute abundance of methanogens (determined via qPCR targeting the *mcrA* gene) decreased signficantly in captive animals (Amato *et al.*, 2013). Thus, in these primates, both the diversity and abundance of methanogens appear to decrease as a result of captivity. It is worth noting that, with the exception of the Youngblut *et al.* 2020 study, these works did not use archaea-targeting 16S rRNA gene sequencing approaches and may have

thus missed some of the finer scale diversity differences between captive and wild animals, as pointed out in (Raymann *et al.*, 2017).

More generally speaking, captivity does not appear to have a uniform impact on the overall microbiome of animals. Indeed, captivity decreases the bacterial richness of some mammals like canids, primates, equids and rhinoceros, but in others like the bovids, giraffes, anteaters, and aardvarks, no significant differences between wild and captive individuals have been observed (Amato *et al.*, 2013; McKenzie *et al.*, 2017). However, even when richness does not change in these mammals, captivity does impact the bacterial community composition and/or relative abundance of some lineages (Nelson *et al.*, 2013; Clayton *et al.*, 2016; McKenzie *et al.*, 2017; Gibson *et al.*, 2019). These community level changes, either in alpha or beta diversity, also result in distinct shifts in the functional profile of the intestinal bacterial community (Zhou *et al.*, 2014; Clayton *et al.*, 2016; Gibson *et al.*, 2019).

There are several hypotheses for why we observe these changes. One thought is that captive animals have a diet that is also more uniform over time, providing less diverse niches for gut bacteria. For example, captive primates consume less fibre from a less diverse range of plants, so they won't need as many different types of bacteria to help them break down the plant material (McKenzie *et al.*, 2017). This is also seen in captive black rhinos, where there is an increase in genes associated to glycolysis and amino acid synthesis pathways likely indicating that the intestinal microbiota has adapted to a less varied range of dietary material and is more prepared for digesting food that is high in glucose (Gibson *et al.*, 2019). In other groups of animals, less is known concerning the impact of captivity on the intestinal microbiome. However, in single species studies from reptiles and birds, the same trends can be seen as with mammals - a significant shift in the relative abundances of major bacterial lineages (changes in beta diversity) and as a result, changes in the overall functional profile of the intestinal microbiome (Xenoulis *et al.*, 2010; Wienemann *et al.*, 2011; Zhou *et al.*, 2014; Kohl *et al.*, 2017). As in mammals, these changes are possibly related to the more restricted diet of captive reptiles and birds.

Another thought as to why there are clear differences between captive and wild animals' intestinal microbial communities is that captive individuals interact with humans on a regular

basis. For example, in primates (McKenzie *et al.*, 2017), captivity has resulted in a convergence of the intestinal bacterial community that is more similar to humans than to wild primates. Similarly in black rhinos (order Perissodactyla), groups of microbes found in domesticated livestock are often detected in captive individuals but never in wild rhinos (Gibson *et al.*, 2019). In captive leopard seals, which are carnivorous marine mammals, there is even an increase in alpha diversity likely because of their interaction with humans/human handled food (Nelson *et al.*, 2013). Captive cheetahs, birds, and reptiles host higher concentrations in potentially pathogenic lineages of intestinal bacteria possibly because of an increased interactions with humans (Xenoulis *et al.*, 2010; Wienemann *et al.*, 2011; Zhou *et al.*, 2014; Jiang *et al.*, 2017; Wasimuddin *et al.*, 2017).

Interestingly, these captivity-associated trends can be seen in the data presented in Chapter 1, as *M. smithii* was particularly prevalent. However, biases associated with captivity do not mask important variations in the abundance and composition of the archaeal community related to different factors. We observed associations of methanogen clades with specific animal lineages, such as the high absolute and relative abundance of *Methanocorpusculum* in the perissodactyls and *Methanomethylophilaceae* in frugivorous animals. Further, other human associated methanogens such as *Methanosphaera* and Methanomassiliicoccales were not as widespread as *M. smithii*. Taken together, this suggests that *M. smithii* has a higher capacity to colonize a wide range of animal hosts, relative to other intestinal methanogens.

However, it is possible that a systematic comparison between the intestinal archaeome of captive vs wild animals would reveal a shift in which archaea are the most abundant in the intestinal tract in Primates, Perissodactyla, and Carnivora, as well as a change in the overall richness of archaea in various animals.

Finally, some wild animals change diets seasonally based on plant availability (Amato *et al.*, 2014; Hicks *et al.*, 2018). A decrease in fibre consumption may result in a decrease in H_2 production and consequently of methanogenic archaea in the gut. In contrast, captive animals do not experience the same seasonal dietary shifts (Amato *et al.*, 2013), and may thus not experience the same seasonal dietary/microbiota changes. This question leads us to the following section.

V. The Intestinal Archaeome across Time

Generally speaking, the stability of the intestinal archaeome over various time scales is poorly understood. At short time-scales (six months to one year) humans have been shown to host relatively stable populations of Methanomassiliicoccales and *M. smithii* (Miller *et al.*, 1983; Borrel *et al.*, 2017). Across longer periods of time, sporadic, single-species studies in primates, ruminants, and humans have observed shifts in the intestinal archaeal community between seasons (due to dietary changes) (Amato *et al.*, 2014; Hicks *et al.*, 2018), or with age (Gaci *et al.*, 2014; Polag *et al.*, 2014). These studies have even identified specific lineages of archaea as being 'heritable' (Goodrich *et al.*, 2016; Grieneisen *et al.*, 2021), meaning they are passed from parent to offspring across generations. However, the stability of the intestinal archaeome in most other animals at various timescales has yet to be examined.

V.a Seasonal impacts on the archaeome

Many animals drastically shift their dietary niches with different seasons and this also changes the intestinal bacterial community (Maurice *et al.*, 2015; Hicks *et al.*, 2018; Guo *et al.*, 2021). In primates (Amato *et al.*, 2014; Hicks *et al.*, 2018) and wild mice (Maurice *et al.*, 2015), seasonal changes in bacteria as a result of changing diets have been observed. In human populations that switch seasonally between hunting and gathering, seasonal changes have been observed in the bacterial community both taxonomically and functionally (Smits *et al.*, 2017). More specifically, a larger number of microbial enzymes degrading animal-derived carbohydrate were detected in seasons where the population relies heavily on hunting (Smits *et al.*, 2017). Because the archaeal community is highly dependent on the substrates generated by the bacterial community and because the requirements for H_2 and the carbon substrates for methanogenesis are different between different methanogens, it is possible that the archaeal community also changes as a result of seasonal dietary variations.

These shifts may occur between the two types of methanogens that dominate the intestinal microbial community across animals - CO_2-reducing and methyl-reducing hydrogenotrophic methanogens. CO_2-reducing methanogens require four H_2 molecules while methyl-reducing methanogens require only one per molecule of methane produced; and methyl-reducers have a higher affinity for H_2 than CO_2-reducing methanogens (Lovley, 1985; Feldewert *et al.*, 2021).

Even though we encountered methyl-reducing methanogens more frequently in our data than what has been observed in open environments (Cozannet *et al.*, 2021), CO_2-reducing methanogens were still the overwhelming majority of methanogenic sequences we detected (~60%), especially in herbivores. This supports what has been suggested previously (Feldewert *et al.*, 2021), that methyl-reducing methanogens are limited by the availability of methylated compounds in the intestine and not by the availability of H_2. Since both H_2 and methylated compounds that are used by methanogens are the by-products of bacterial fermentation of dietary fibre, changes in the type of fibre would change the availability these substrates for the methanogens. For example, in species that alternate between leaves and fruit based on availability -as seen in Howler monkeys (Amato *et al.*, 2014)-, would we observe a shift from CO_2-reducing to methyl-reducing methanogens in the intestinal community? Interestingly, Hicks *et al.* did observe a change in relative abundance of intestinal methanogens between seasons in gorillas, with relatively more methanogens being detected when fruit availability was high (Hicks *et al.*, 2018). However, the most abundant archaea in their samples are typical free-living or even extremophilic archaea (e.g. *Methanopyrales*, *Methanococcaceae*, *Methanoregulaceae*), this suggests a bias in their annotation methodology relying on metagenomic reads, which warrants a reanalysis.

Our current study, and others (Turnbaugh *et al.*, 2006; Ley *et al.*, 2008; Delsuc *et al.*, 2014; Hird *et al.*, 2015; Vital *et al.*, 2015; Song *et al.*, 2019), draw conclusions about the adaptations of intestinal microbes to their hosts based on their presence or absence and relative abundance from a single time point. A large-scale archaea-centric study on how seasonal dietary changes impact the intestinal archaeal community would certainly provide unique insights into how the intestinal archaeome assembles across varying time scales. A more robust understanding of how the entire intestinal microbial population changes seasonally would also provide important information to shape our understanding of the intestinal microbiome evolutionarily.

V.b Changes in the archaeome with age

With age, the gut microbiome changes both in terms of taxonomic composition and metabolic function (Bosco *et al.*, 2021). Human aging is accompanied by a decrease in bacterial diversity,

with groups like *Prevotella* (Bacteroidetes) and *Bifidobacterium* (Actinobacteria) becoming less abundant (Jeffery *et al.*, 2016). The diversity of the metabolic capacity of intestinal microbes also decreases with age, resulting in reduced concentrations of short-chain fatty acids (SCFAs) (Nagpal *et al.*, 2018). Interestingly, humans aging also results in an increase in the proportion of intestinal methanogens (Mihajlovski *et al.*, 2010; Dridi, Henry, *et al.*, 2012; Gaci *et al.*, 2014; Vanderhaeghen *et al.*, 2015).

There are various hypotheses as to why we see a decrease in bacterial diversity while the abundance of methanogens seems to increase with age in humans. One thought has to do with how long it takes for material to travel through the intestinal tract. Indeed, longer intestinal transit times are associated with an increase in the abundance of methanogens in the intestinal tract (Nagpal *et al.*, 2018). More specifically, in the elderly and people with chronic constipation (groups of humans with significantly longer intestinal transit times) a higher abundance of intestinal methanogens or methane gas production has been observed (Kunkel *et al.*, 2011; Triantafyllou *et al.*, 2014). We also support this as we found that in mammals the mean retention time is positively correlated to the absolute abundance of methanogens. This fits with what is observed in humans concerning the overall richness of the bacterial community, which decreases as transit time increases (Tottey *et al.*, 2017). Interestingly, this study also found a decrease in sulfate-reducing bacteria in correlation with higher transit times (Tottey *et al.*, 2017). As sulfate-reducing bacteria are thought to be in competition with methanogens for hydrogen in the intestine, this surely warrants further investigation.

Similarly in rats and ruminants, there is an increase the abundance of intestinal methanogens with age (Maczulak *et al.*, 2014; Liu *et al.*, 2017). In young ruminants like calves and lambs, methanogens are present very early in their life cycle - even at birth (Zhou *et al.*, 2014; Guzman *et al.*, 2015). Information concerning how methanogens change with age is largely lacking in other animals and would be of great interest to study both on long- and short-time scales. More specifically, an archaea-centric time series study would build the understanding on how stable the intestinal archaeome is between seasons but also over lifetimes of animals. A recent metagenomic study followed the intestinal microbiome of baboons over their lifetime, across multiple generations (Grieneisen *et al.*, 2021). It did not consider how archaea change with

age, and so these data might be re-used to investigate how this factor influence archaeal communities in non-human primate. However, this study found that some methanogens are highly heritable, an aspect that will be discussed in the following section.

V.c Inheritance of Archaea between Generations

Previously, Hackstein et al. (1996) suggested that host evolution is tightly linked to the presence of methanogenic archaea, and our findings largely support this. Certain archaea preferentially associate with certain animal lineages, e.g., Perissodactyla (*Methanocorpusculum*), and Primates ("*Ca*. Methanomethylophilaceae"). Further, animals that are closely related to each other, like the perissodactyls and cetartiodactyls, host methanogens that are genetically more similar to each other than they are to those detected in more distant animal lineages, like rodents. The existence of distinct clades suggests that over evolutionary time these methanogens have co-speciated with their host, being propagated through these lineages of mammals.

Host-genetics has also been determined to play a major role in shaping the intestinal archaeome at shorter time scale. Certain groups of intestinal microbes including methanogens like *Methanobrevibacter* and "*Ca*. Methanomethylophilaceae" have been shown to be highly heritable in humans and baboons, respectively, as well as bacterial families like *Christensenellaceae* in both animal species (Hansen *et al.*, 2011; Goodrich *et al.*, 2014; Grieneisen *et al.*, 2021). Genomically, *M. smithii* strains isolated from homozygotic twins are more similar to each other than to strains isolated from the mothers of the twins and other people, suggesting that the genetics of the human hosts heavily influences the genomic composition of their intestinal microbes (Hansen *et al.*, 2011). In ruminants, *Methanobrevibacter ruminantium* (Li *et al.*, 2019) and other *Methanobrevibacter* were found to be heritable members of the rumen intestinal microbiota (Difford *et al.*, 2018; Mizrahi *et al.*, 2021). It has also been suggested that methane production (high vs low) is a heritable trait in ruminants, and that this trait is largely a result of the inherited bacterial community which either competes with or produces substrates for different methanogenic lineages (Mizrahi *et al.*, 2021).

The mode in which microbes are introduced between the mother to the offspring varies between animal classes. For example in Mammalia, offspring is exposed to their mother's microbiota as part of the birthing process (Blaser, 2006), and possibly even before in the placenta (Satokari *et al.*, 2009). Other factors such as breast-feeding have been shown to also play an important role in the establishment of the intestinal microbiota (Dominguez-Bello *et al.*, 2010). *M. smithii* has been detected in human breast milk (Togo *et al.*, 2019), and this may be one way in which maternal microbes are introduced into the offspring's intestinal tract, as methanogens have been detected as early as one week after birth (Hudson *et al.*, 1993). However, whether and how the intestinal microbiota is vertically inherited in other groups of animals is not well understood. In viviparous lizards, up to 30% of the intestinal bacterial community is shared between mother and offspring, even when the pair were separated after birth (Kohl *et al.*, 2017). In chicken embryos host microbes are inherited from maternal hens and development of the intestinal bacterial community is heavily influenced by host genetics (Jiang *et al.*, 2017). Even in termites, vertical inheritance is a major driving force in the assembly of the intestinal bacterial community (Rahman *et al.*, 2015). However, a fuller understanding of inheritably of methanogens across a wide range of animals is largely lacking. Indeed, determining the heritability of microbes requires large cohorts of individual from the same species, making these types of analyses difficult to conduct on a large range of species. Nevertheless, a study focused on a few distantly related animal species, would undoubtedly provide important insights into how/if other intestinal methanogens are inherited in animals.

VI. Future comparative metagenomics

As discussed in Chapter 2, the use of comparative metagenomics can be particularly useful to investigate the dynamics of gene loss and gains linked to colonization of the intestinal tract. Some adaptations to the host intestinal tract have been already identified in three orders of intestinal methanogens (Methanobacteriales, Methanomassiliicoccales, and Methanosarcinales (Borrel *et al.*, 2014; Poehlein *et al.*, 2018; Thomas *et al.*, 2021)). However, because of the limited number of genomes available until recently, a large-scale comparative genomics analysis of the intestinal archaea has not been carried out. Additionally, there are various lineages -such as the Methanomicrobiales- for which there is no data concerning how these microbes have successfully colonized their hosts.

It will be important to investigate both the transition from the environment to the intestinal tract as well as the genomic differences that can be observed in archaea from various animal hosts, as different animal digestive systems pose very different physical stressors. For example, the insect-associated *Methanobrevibacter* spp. have commonly been identified to have a number of genes allowing them to deal with oxidative stress, as well as adhesin/filaments which allow them to live in close association to the hindgut wall (Hackstein *et al.*, 1994; Leadbetter *et al.*, 1996, 1998; Tholen *et al.*, 2007). Similar unique adaptations may be observed in the Methanomicrobiales, as the 16S rRNA gene analysis and metagenomic analyses from Chapters 1.A and 2.B clearly show that there are two distinct clades of host-associated *Methanocorpusculum* clades (mammalian only and a mixed host clade) and as well as a rumen-adapted genus (*Methanomicrobium*). These clusters of sequences from certain host lineages have also been observed in the *Methanimicrococcus*.

We have now assembled ~200 novel MAGs from gut archaea. This represents a unique dataset which will undoubtedly aide in these much-needed investigations. Further, as archaeal MAGs are starting to become available from a more diverse range of animal hosts, a re-evaluation of previously analyzed lineages like the Methanobacteriales and Methanomassiliicoccales will likely reveal previously unidentified adaptations.

VII. Validation of hypotheses in the laboratory

Most of the work presented here is the result of culture-independent approaches – e.g., qPCR, amplicon sequencing, and comparative genomics, and many of the hypotheses suggested by our results can now be further investigated at the bench. For example, the numerous relationships suggested by co-occurrence analyses between methyl-reducing hydrogenotrophic methanogens and methyl-producing bacteria could be validated through co-cultures, microscopy, and transcriptomic analyses to establish how these microbes are interacting. A co-culture between a hydrogen producing bacterium and the two different hydrogenotrophic methanogens under different physical conditions would be particularly useful for better understanding the potential competition between methanogens in the intestine. Similar approaches were used recently to elucidate the syntrophic relationship

between *Methanobrevibacter smithii* and *Christensenella spp.* (Ruaud *et al.*, 2020). Indeed, *M. smithii* goes better with *Christensenella* as a hydrogen donor than it does with *Bacteroides thetaiotaomicron* (Ruaud *et al.*, 2020).

Host-methanogen interactions would also be particularly interesting to investigate in a laboratory setting as a number of the proteins identified in methanogens as adaptations to the host intestinal tract could be involved in these interactions (Samuel *et al.*, 2007; Hansen *et al.*, 2011; Lurie-Weinberger, Peeri and Gophna, 2012; Lurie-Weinberger, Peeri, Tuller, *et al.*, 2012; Thomas *et al.*, 2021). Thus, validation of the annotation of these proteins is needed. More specifically, binding assays of the Flg_new proteins and other adhesin-like proteins (ALPs) would allow for functional characterization and likely provide key insights into the mechanism that *Methanimicrococcus* uses to maintain its presence in the intestinal tract of some animals. For example, in the insect hindgut, *Methanobrevibacter spp*. have been found to be closely associated to the chitinous bristles lining the organ (Leadbetter *et al.*, 1996), but it is currently unclear if *Methanimicrococcus* uses a similar mechanism to maintain its position in the intestinal tract of insects. Additionally, host-archaea interactions could be monitored similarly to what was done with *Methanosphaera* and human cells (Bang *et al.*, 2017) to determine if and how different methanogens like *Methanimicrococcus* are recognized by the host immune system.

Finally, previous efforts have managed to isolate aerobic archaea from the human intestinal tract (Nam *et al.*, 2008; Oxley *et al.*, 2010). Further efforts should therefore be made to isolate the Nitrososphaerales from different fecal samples where they are shown to be most abundant. Isolation and genome sequencing would be especially useful in determining the residency of these aerobic ammonia-oxidizing archaea in the intestinal tract.

Conclusions

The intestinal microbiome has been studied since the early days of microbiology (Metchnikoff *et al.*, 1909). Even though the methanogens have been known to play an active part of this complex microbial consortia for decades (Wasburn *et al.*, 1937), they have been consistently overlooked in large-scale studies. Thus, in most animals, basic information such as the diversity and abundance of intestinal archaea had never been examined. As a consequence, the overall understanding of how these microorganisms successfully colonized the intestinal tract remained unanswered. Before the work presented in this thesis, it was largely assumed that there were two major host associated methanogens, the *Methanobrevibacter* and Methanomassiliicoccales, and two rarer orders, the Methanomicrobiales and Methanosarcinales. These assumptions were the result of sporadic single species studies largely in humans, primates, and ruminants and methane gas detection approaches (Hackstein *et al.*, 1996). With this work, I have significantly contributed to the expansion of our current understanding of the intestinal archaeome by using a combination of approaches such as qPCR, 16S rRNA gene amplicon sequencing, and comparative genomics in a diverse set of animal fecal samples.

The significance of these results is multifold. First, this is among the first large-scale, archaea-focused, sequenced-based analyses conducted in a large range of animal hosts. The information gathered on the diversity and abundance of intestinal archaea will serve as a foundation for future work aimed at investigating the intestinal archaeome. This is also one of the first analyses showing a strong association between mean retention time, diet fibre content, and intestinal methanogens across such a diverse range of mammals. Further, we established that there are two major methanogenesis pathways that are particularly well suited for the intestinal tract - CO_2-reducing and methyl-reducing hydrogenotrophic methanogenesis. The later type of methanogenesis could have emerged twice independently (in *Methanimicrococcus* and *Methanosphaera*) in link with the colonization of the methyl-compounds/H_2 rich gut, revealing one of the possible evolutionary paths that led to the large diversity of the methyl-reducing hydrogenotrophic methanogens. More generally, our

analysis of the gains and losses of genes and functions characterized in *Methanimicrococcus* mirrors what has previously been found both in distantly related intestinal methanogens and intestinal bacteria, suggesting that there has been some level of convergent evolution across the intestinal microbial life.

To conclude, this PhD thesis work opens multiple avenues of research on the archaeome, a young subject among the most promising in the field of microbiome studies.

Résumé en français

Le microbiome intestinal a été étudié depuis les débuts de la microbiologie (Metchnikoff et al., 1909). Bien que l'on sache depuis des décennies que les méthanogènes jouent un rôle important dans ce consortium microbien complexe (Wasburn et Brody, 1937), ils ont toujours été négligés dans les études à grande échelle. Ainsi, chez la plupart des animaux, des informations fondamentales telles que la diversité et l'abondance des archées intestinales n'ont jamais été examinées. Par conséquent, le processus qui a permis à ces microorganismes de réussir à coloniser le tractus intestinal est resté pendant longtemps inconnu. Antérieurement aux travaux présentés dans le chapitre 1, il était largement admis qu'il existe deux groupes majeurs de méthanogènes associés à l'hôte, les Methanobrevibacter et les Methanomassiliicoccales, et deux ordres moins abondants, les Methanomicrobiales et les Methanosarcinales. Ces hypothèses reposaient sur des études sporadiques portant sur une seule espèce d'archée principalement chez les humains, les primates et les ruminants ; et des méthodes de détection du méthane (Hackstein, 1996 ; Hackstein et van Alen, 1996). Afin d'approfondir les connaissances actuelles sur l'archéome intestinal, j'ai combiné différentes approches allant de la qPCR et le séquençage des amplicons du gène de l'ARNr 16S jusqu'à la génomique comparative, sur un ensemble d'échantillons fécaux de différents animaux.

La méthodologie appliquée dans cette thèse souligne l'importance d'utiliser des amorces du gène de l'ARNr 16S adaptées pour les études du microbiome, ainsi que l'importance de l'approche quantitative - qui fait largement défaut dans ce domaine. En effet, la qPCR s'est avérée la plus efficace pour détecter les archées dans les échantillons fécaux (figure 1, chapitre 1). Les données quantitatives ont également permis d'obtenir une vision plus précise des lignées d'archées abondantes et rares. Par exemple, certaines archées ont une faible abondance absolue alors qu'elles dominent l'ensemble de la communauté archéenne. Une étude basée uniquement sur les données de séquençage (abondance relative) aurait pu conclure que ces lignées sont des archées importantes associées à l'hôte, alors qu'elles sont présentes en faibles abondances chez la plupart des animaux et seulement détectées lorsque les archées dominantes et mieux adaptées sont absentes. De plus, nous avons déterminé la

cause de la quasi-absence des méthanogènes dans le microbiome intestinal des carnivores (Hackstein et van Alen, 1996). En effet, les méthanogènes sont systématiquement moins abondants chez ces animaux et ne produisent peut-être pas assez de méthane pour être détectés par les méthodes de détection des gaz.

L'utilisation d'amorces spécifiques des archées nous a permis de mieux caractériser leur diversité dans un plus grand nombre d'échantillons par rapport aux amorces procaryotes universelles (figure 1, chapitre 1.A). Autrement dit, si les archées sont présentes en concentrations significativement faibles, elles ne seront pas détectées par l'utilisation d'amorces universelles. Ainsi, si l'objectif d'une étude est de cibler l'archéome intestinal, il est recommandé d'utiliser des amorces dessinées spécifiquement pour les archées. Cependant, il est important de noter que les principaux groupes d'archées comme les Methanobacteriales, Methanomassiliicoccales, Methanomicrobiales et Methanosarcinales ont été systématiquement détectés dans les échantillons fécaux de différents animaux en utilisant les deux types d'amorces. Cela signifie que les études qui ont déjà été réalisées en utilisant des amorces universelles (sans mésappariement avec la plupart des séquences d'archées) et qui ont occulté l'archéome intestinal, méritent d'être réexaminées afin d'étendre le travail effectué ici (chapitre 1.B).

Grâce à ces approches ciblées, j'ai caractérisé l'archéome intestinal d'un large spectre d'animaux hôtes. Ce travail a mis en évidence l'existence de cinq lignées principales de méthanogènes dans l'intestin : *Methanobrevibacter*, *Methanosphaera*, "*Ca*. Methanomethylophilaceae" (Methanomassiliicoccales), *Methanimicrococcus*, et *Methanocorpusculum*, et a révélé la présence potentielle d'une lignée supplémentaire d'archées aérobies oxydant l'ammoniac (Nitrososphaerales). Les approches de qPCR montrent que tandis que Methanobrevibacter et "*Ca*. Methanomethylophilaceae " sont abondants, les deux autres genres de méthanogènes ne font pas partie de la fraction rare puisqu'ils ont été détectés dans toutes les classes d'animaux. De plus, ce travail souligne l'omniprésence de ces microorganismes dans le tractus intestinal, quel que soit le régime alimentaire. Cependant, les animaux carnivores hébergent un archéome intestinal nettement moins abondant que leurs homologues herbivores. En effet, le régime alimentaire et la phylogénie de l'hôte déterminent l'abondance et la diversité des archées qui coloniseront le tractus intestinal des

animaux. L'étude des relations de parenté entre les archées associées à l'hôte et des archées environnementaux sur une phylogénie d'espèces a montré que ces microorganismes intestinaux ont divergé de leurs parents environnementaux et se sont adaptés, par des changements affectant leurs génomes, pour coloniser les animaux. Ces travaux indiquent l'existence d'au moins cinq événements majeurs et indépendants d'adaptation des archées au tractus intestinal animal.

Une étude comparative des génomes appartenant à l'une de ces lignées, *Methanimicrococcus*, confirme cette hypothèse. En effet, les membres de *Methanimicrococcus* ont subi un remaniement génomique important suite à la colonisation du tractus intestinal. Nos analyses ont révélé que les Methanimicrococcus ont subi la plus grande réduction de génome observée jusqu'à présent chez les Methanosarcinales. Ces pertes concernent la plupart des gènes impliqués dans la branche méthyle de H4MPT de Wood-Ljungdahl (H4MPT mWL), observées pour la première fois chez des méthanogènes de classe I/II et elle fournit des informations uniques sur l'évolution de la voie de la méthanogenèse hydrogénotrophique réductrice de méthyle. D'autres gènes perdus sont impliqués dans la chimiotaxie et la motilité (plus généralement dans la transduction des signaux), la synthèse de certains acides aminés, la fixation du N_2 et du CO_2, ainsi que le stockage du carbone sous forme de glycogène. Ces fonctions sont probablement devenues redondantes/non nécessaires dans un environnement riche en nutriments comme l'intestin. Methanimicrococcus possède un répertoire distinct de gènes impliqués dans la modification de la surface cellulaire, dont certains ont été acquis par des transferts horizontaux de gènes. Il est important de noter que les pertes et acquisitions de gènes chez Methanimicrococcus suivent la même tendance précédemment observée chez les Methanomassiliicoccales et les Methanobacteriales libres et associées à l'hôte (Lurie-Weinberger, Peeri et Gophna, 2012 ; Borrel et al., 2017 ; de la Cuesta-Zuluaga et al., 2021), ainsi que les bactéries intestinales. De manière générale, ces résultats indiquent plusieurs adaptations convergentes dans des lignées distantes d'archées intestinales. Les ~200 nouveaux MAG d'archées obtenus dans ce travail à partir d'une grande variété d'animaux permettront d'aborder ces questions d'adaptations convergentes à la fois par rapport au microbiome intestinal en général et à des groupes spécifiques d'animaux.

En résumé, ce travail met en évidence plusieurs points importants. Tout d'abord, il s'agit de la première analyse à grande échelle des données de séquençage, axée sur la diversité des archées chez un large spectre d'animaux hôtes. Les informations recueillies sur la diversité et l'abondance des archées intestinales serviront de base aux futurs travaux sur l'archéome intestinal. Il s'agit également de l'une des premières analyses mettant en évidence la forte association entre le temps de rétention moyen dans le transit digestif, la teneur en fibres du régime alimentaire et les méthanogènes intestinaux dans une large et diverse gamme de mammifères. En effet, ce travail a permis d'établir la présence de deux grandes voies de méthanogénèse particulièrement bien adaptées au tractus intestinal : la méthanogenèse hydrogénotrophique réductrice de CO_2 et réductrice de méthyle. Ce dernier type de méthanogenèse pourrait avoir émergé deux fois indépendamment (chez Methanimicrococcus et Methanosphaera) suite à la colonisation du microbiome intestinal riche en composés méthyliques et en H_2, révélant ainsi l'une des voies évolutives possibles qui aurait conduit à la grande diversité des méthanogènes hydrogénotrophes réducteurs de méthyle. Plus généralement, les gains et les pertes de gènes et des fonctions caractérisés chez Methanimicrococcus correspondent à ce qui a été précédemment observé à la fois chez les méthanogènes intestinaux et les bactéries intestinales apparentées, ce qui suggère une convergence évolutive de la flore microbienne intestinale.

Bibliography

Abraham, A.J. *et al.* (2020) 'Improved estimation of gut passage time considerably affects trait-based dispersal models', *Functional Ecology*, (November 2020), pp. 1–10. doi:10.1111/1365-2435.13726.

Adam, P.S., Borrel, G., Brochier-Armanet, C. and Gribaldo, S. (2017) 'The growing tree of Archaea: New perspectives on their diversity, evolution and ecology', *ISME Journal*, 11(11), pp. 2407–2425. doi:10.1038/ismej.2017.122.

Alibardi, L. and Cossu, R. (2016) 'Effects of carbohydrate, protein and lipid content of organic waste on hydrogen production and fermentation products', *Waste Management*, 47, pp. 69–77.

Amato, K.R. *et al.* (2013) 'Habitat degradation impacts black howler monkey (Alouatta pigra) gastrointestinal microbiomes', *The ISME Journal 2013 7:7*, 7(7), pp. 1344–1353. doi:10.1038/ismej.2013.16.

Amato, K.R. *et al.* (2014) 'The Gut Microbiota Appears to Compensate for Seasonal Diet Variation in the Wild Black Howler Monkey (Alouatta pigra)', *Microbial Ecology 2014 69:2*, 69(2), pp. 434–443. doi:10.1007/S00248-014-0554-7.

Amato, Katherine R. *et al.* (2019) 'Convergence of human and Old World monkey gut microbiomes demonstrates the importance of human ecology over phylogeny', *Genome Biology 2019 20:1*, 20(1), pp. 1–12. doi:10.1186/S13059-019-1807-Z.

Amato, Katherine R *et al.* (2019) 'Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes', *The ISME Journal*, 13, pp. 576–587. doi:10.1038/s41396-018-0175-0.

Ametaj, B.N. *et al.* (2010) 'Metabolomics reveals unhealthy alterations in rumen metabolism with increased proportion of cereal grain in the diet of dairy cows', *Metabolomics*, 6(4), pp. 583–594. doi:10.1007/s11306-010-0227-6.

Arrieta, M.-C. *et al.* (2014) 'The Intestinal Microbiome in Early Life: Health and Disease', *Frontiers in Immunology*, 5(September), pp. 1–18. doi:10.3389/fimmu.2014.00427.

Arumugam, M. *et al.* (2011) 'Enterotypes of the human gut microbiome', *Nature*, 473(7346), pp. 174–180. doi:10.1038/nature09944.

Bang, C. *et al.* (2014) 'The intestinal archaea Methanosphaera stadtmanae and Methanobrevibacter smithii activate human dendritic cells', *PLoS ONE*, 9(6), pp. 1–9. doi:10.1371/journal.pone.0099411.

Bang, C. *et al.* (2017) 'Immunogenic properties of the human gut-associated archaeon Methanomassiliicoccus luminyensis and its susceptibility to antimicrobial peptides', *PLoS ONE*, 12(10), pp. 1–11. doi:10.1371/journal.pone.0185919.

Bang, C. and Schmitz, R.A. (2015) 'Archaea associated with human surfaces: Not to be underestimated', *FEMS Microbiology Reviews*, 39(5), pp. 631–648. doi:10.1093/femsre/fuv010.

Barelli, C. *et al.* (2015) 'Habitat fragmentation is associated to gut microbiota diversity of an endangered primate: implications for conservation', *Scientific Reports 2015 5:1*, 5(1), pp. 1–12. doi:10.1038/srep14862.

Barko, P.C., McMichael, M.A., Swanson, K.S. and Williams, D.A. (2018) 'The Gastrointestinal Microbiome: A Review', *Journal of Veterinary Internal Medicine*, 32(1), pp. 9–25. doi:10.1111/jvim.14875.

Bates, S.T. *et al.* (2011) 'Examining the global distribution of dominant archaeal populations in soil', *The ISME journal*, 5(5), pp. 908–917.

Beasley, D.E. *et al.* (2015) 'The evolution of stomach acidity and its relevance to the human microbiome', *PloS one*, 10(7), p. e0134116.

Beijer, W.H. (1952) 'Methane Fermentation in the Rumen of Cattle', *Nature*, 170, p. 367. doi:https://doi.org/10.1038/170576a0.

Biavati, B., Vasta, M. and Ferry, J.G. (1988) 'Isolation and characterization of "Methanosphaera cuniculi" sp. nov.', *Applied and Environmental Microbiology*, 54(3), pp. 768–771.

Biddle, A., Stewart, L., Blanchard, J. and Leschine, S. (2013) 'Untangling the genetic basis of fibrolytic specialization by Lachnospiraceae and Ruminococcaceae in diverse gut communities', *Diversity*, 5(3), pp. 627–640. doi:10.3390/d5030627.

Blaser, M.J. (2006) 'Who are we?', *EMBO reports*, 7(10), pp. 956–960. doi:10.1038/SJ.EMBOR.7400812.

Blaut, M. and Clavel, T. (2007) 'Metabolic diversity of the intestinal microbiota: implications for health and disease', *Journal of Nutrition*, 137(3). doi:10.1093/jn/137.3.751s.

Bond, J.H., Engel, R.R. and Levitt, M.D. (1971) 'Factors influencing pulmonary methane

excretion in man. An indirect method of studying the in situ metabolism of the methane-producing colonic bacteria', *J Exp Med*, 1(133), pp. 572–588. doi:10.1084/jem.133.3.572.

Borrel, G. *et al.* (2012) 'Genome sequence of "Candidatus Methanomethylophilus alvus" Mx1201, a methanogenic archaeon from the human gut belonging to a seventh order of methanogens', *Journal of Bacteriology*, 194(24), pp. 6940–6941. doi:10.1128/JB.01867-12.

Borrel, G., Harris, H., *et al.* (2013) 'Genome sequence of "Candidatus Methanomassiliicoccus intestinalis" Issoire-Mx1, a third Thermoplasmatales-related methanogenic archaeon from human feces', *Genome Announcements*, 1(4), pp. 1–2. doi:10.1186/1471-21.

Borrel, G., O'Toole, P.W., *et al.* (2013) 'Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis', *Genome Biology and Evolution*, 5(10), pp. 1769–1780. doi:10.1093/gbe/evt128.

Borrel, G. *et al.* (2014) 'Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine', *BMC Genomics*, 15(1), p. 679. doi:10.1186/1471-2164-15-679.

Borrel, G. *et al.* (2017) 'Genomics and metagenomics of trimethylamine-utilizing Archaea in the human gut microbiome', *ISME Journal*, 11(9), pp. 2059–2074. doi:10.1038/ismej.2017.72.

Borrel, G. *et al.* (2019) 'Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea', *Nature Microbiology*, 4(4), pp. 603–613. doi:10.1038/s41564-019-0363-3.

Borrel, G. *et al.* (2020) 'The host-associated archaeome', *Nature Reviews Microbiology* [Preprint]. doi:10.1038/s41579-020-0407-y.

Bosco, N. and Noti, M. (2021) 'The aging gut microbiome and its impact on host immunity', *Genes & Immunity 2021*, pp. 1–15. doi:10.1038/s41435-021-00126-8.

Brugère, J.F. *et al.* (2013) 'Archaebiotics: Proposed therapeutic use of archaea to prevent trimethylaminuria and cardiovascular disease', *Gut Microbes*, 5(1). doi:10.4161/gmic.26749.

Brune, A. (2018) 'Methanogens in the Digestive Tract of Termites', in Hackstein, J.H.P. (ed.)

(Endo)symbiotic Methanogenic Archaea)symbiotic Methanogenic Archaea. Springer N. Springer Nature, pp. 81–101. doi:10.1007/978-3-642-13615-3_6.

Brune, A. and Dietrich, C. (2015) 'The Gut Microbiota of Termites: Digesting the Diversity in the Light of Ecology and Evolution', *Annual Review of Microbiology*, 69(1), pp. 145–166. doi:10.1146/annurev-micro-092412-155715.

Bui, T.P.N. *et al.* (2019) 'Mutual Metabolic Interactions in Co-cultures of the Intestinal Anaerostipes rhamnosivorans With an Acetogen, Methanogen, or Pectin-Degrader Affecting Butyrate Production', *Frontiers in Microbiology*, 10. doi:10.3389/fmicb.2019.02449.

Calloway, D.H., Colastio, D.J. and Mathews, R.D. (1966) 'Gases produced by Human Intestinal Microflora', *Nature 1966 212:5067*, 212(5067), pp. 1238–1239. doi:10.1038/2121238a0.

Carbonero, F., Benefiel, A.C., Alizadeh-Ghamsari, A.H. and Gaskins, H.R. (2012) 'Microbial pathways in colonic sulfur metabolism and links with health and disease', *Frontiers in Physiology*, 0, p. 448. doi:10.3389/FPHYS.2012.00448.

Chassard, C., Delmas, E., Robert, C. and Bernalier-Donadille, A. (2010) 'The cellulose-degrading microbial community of the human gut varies according to the presence or absence of methanogens', *FEMS Microbiology Ecology*, 74(1), pp. 205–213. doi:10.1111/j.1574-6941.2010.00941.x.

Chivers, D.J. and Hladik, C.M. (1980) 'Morphology of the gastrointestinal tract in primates: Comparisons with other mammals in relation to diet', *Journal of Morphology*, 166(3), pp. 337–386. doi:10.1002/jmor.1051660306.

Christl, S.U., Gibson, G.R. and Cummings, J.H. (1992) 'Role of dietary sulphate in the regulation of methanogenesis in the human large intestine', *Gut*, 33(9), pp. 1234–1238. doi:10.1136/gut.33.9.1234.

Chung, Y.W. *et al.* (2020) 'Functional dynamics of bacterial species in the mouse gut microbiome revealed by metagenomic and metatranscriptomic analyses', *PLOS ONE*, 15(1), p. e0227886. doi:10.1371/JOURNAL.PONE.0227886.

Clauss, M. *et al.* (2008) 'The influence of natural diet composition, food intake level, and body size on ingesta passage in primates', *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 150(3), pp. 274–281.

Clauss, M. *et al.* (2020) 'Review: Comparative methane production in mammalian

herbivores', *Animal*, 14(S1), pp. S113–S123. doi:10.1017/S1751731119003161.

Clavel, T., Lagkouvardos, I., Blaut, M. and Stecher, B. (2016) 'The mouse gut microbiome revisited: From complex diversity to model ecosystems', *International Journal of Medical Microbiology*, 306(5), pp. 316–327. doi:10.1016/j.ijmm.2016.03.002.

Clayton, J.B. *et al.* (2016) 'Captivity humanizes the primate microbiome', *PNAS*, 113(37), pp. 10376–10381. doi:10.1073/PNAS.1521835113.

Clemente, J.C. *et al.* (2015) 'The microbiome of uncontacted Amerindians', *Science Advances*, 1, p. e1500183. doi:10.1126/sciadv.1500183.

Cole, J.R. *et al.* (2014) 'Ribosomal Database Project: Data and tools for high throughput rRNA analysis', *Nucleic Acids Research*, 42(D1), pp. 633–642. doi:10.1093/nar/gkt1244.

Conrad, R. (2020) 'Importance of hydrogenotrophic, aceticlastic and methylotrophic methanogenesis for methane production in terrestrial, aquatic and other anoxic environments: A mini review', *Pedosphere*, 30(1), pp. 25–39. doi:10.1016/S1002-0160(18)60052-9.

Cornick, N.A. *et al.* (1994) 'Lachnospira pectinoschiza sp. nov., an anaerobic pectinophile from the pig intestine', *International Journal of Systematic Bacteriology*, 44(1), pp. 87–93. doi:10.1099/00207713-44-1-87.

Costello, E.K., Gordon, J.I., Secor, S.M. and Knight, R. (2010) 'Postprandial remodeling of the gut microbiota in Burmese pythons', *The ISME journal*, 4(11), pp. 1375–1385.

Cotillard, A. *et al.* (2013) 'Dietary intervention impact on gut microbial gene richness', *Nature 2013 500:7464*, 500(7464), pp. 585–588. doi:10.1038/nature12480.

Coutinho, P.M. and Henrissat, B. (1999) 'Life with No Sugars?', *J. Mol. Microbiol. Biotechnol*, 1(2), pp. 307–308. Available at: www.caister.com/bacteria-plant (Accessed: 27 August 2021).

Cozannet, M. *et al.* (2021) 'New Insights into the Ecology and Physiology of Methanomassiliicoccales from Terrestrial and Aquatic Environments', *Microorganisms*, 9(30), p. 30. doi:10.3390/MICROORGANISMS9010030.

Cruzen, P.J., Aselmann, I. and Seiler, W. (1986) 'Methane production by domestic animals, wild ruminants, other herbivorous fauna, and humans', *Tellus B*, 38 B(3–4), pp. 271–284. doi:10.1111/j.1600-0889.1986.tb00193.x.

Cryan, J.F. and Dinan, T.G. (2012) 'Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour', *Nature Reviews Neuroscience 2012 13:10*, 13(10),

pp. 701–712. doi:10.1038/nrn3346.

Czerkawski, J.W. and Breckenridge, G. (1971) 'Determination of concentration of hydrogen and some other gases dissolved in biological fluids.', *Laboratory practice*, 20(5), pp. 403–413.

Danielsson, R. *et al.* (2017) 'Methane production in dairy cows correlates with rumen methanogenic and bacterial community structure', *Frontiers in Microbiology*, 8(FEB), pp. 1–15. doi:10.3389/fmicb.2017.00226.

Deevong, P. *et al.* (2004) 'Isolation and Detection of Methanogens from the Gut of Higher Termites', *Microbes and Environments*, 19(3), pp. 221–226. doi:10.1264/jsme2.19.221.

Dehority, B.A. (1969) 'Pectin-fermenting bacteria isolated from the bovine rumen', *Journal of Bacteriology*, 99(1), pp. 189–196.

Delong, E.F. (1992) 'Archaea in coastal marine environments', *Proceedings of the National Academy of Sciences of the United States of America*, 89(12), pp. 5685–5689. doi:10.1073/pnas.89.12.5685.

Delong, E.F. (1998) 'Everything in moderation: Archaea as "non-extremophiles"', *Current Opinion in Genetics and Development*, 8(6), pp. 649–654. doi:10.1016/S0959-437X(98)80032-4.

Delsuc, F. *et al.* (2014) 'Convergence of gut microbiomes in myrmecophagous mammals', *Molecular Ecology*, 23(6), pp. 1301–1317. doi:10.1111/mec.12501.

Dethlefsen, L., Eckburg, P.B., Bik, E.M. and Relman, D.A. (2006) 'Assembly of the human intestinal microbiota', *Trends in Ecology & Evolution*, 21(9), pp. 517–523. doi:10.1016/J.TREE.2006.06.013.

Dick, J.M. (2019) 'CHNOSZ: Thermodynamic calculations and diagrams for geochemistry', *Frontiers in Earth Science*, 7, p. 180.

Difford, G.F. *et al.* (2018) 'Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows', *PLOS Genetics*, 14(10), p. e1007580. doi:10.1371/JOURNAL.PGEN.1007580.

Dominguez-Bello, M.G. *et al.* (2010) 'Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns', *PNAS* [Preprint]. doi:10.1073/pnas.1002601107.

Dordević, D., Jančíková, S., Vítězová, M. and Kushkevych, I. (2021) 'Hydrogen sulfide toxicity in the gut environment: Meta-analysis of sulfate-reducing and lactic acid bacteria in

inflammatory processes', *Journal of Advanced Research*, 27, pp. 55–69. doi:10.1016/j.jare.2020.03.003.

Dore, J. *et al.* (2006) 'Enumeration of H2-utilizing methanogenic archaea, acetogenic and sulfate-reducing bacteria from human feces', *FEMS Microbiology Ecology*, 17(4), pp. 279–284. doi:10.1111/j.1574-6941.1995.tb00152.x.

Douglas, A.E. (2015) 'Multiorganismal Insects: Diversity and function of resident microorganisms', *Annual Review of Microbiology*, 60, pp. 17–34. doi:10.1016/j.physbeh.2017.03.040.

Dridi, B., Henry, M., *et al.* (2012) 'Age-related prevalence of Methanomassiliicoccus luminyensis in the human gut microbiome', *Apmis*, 120(10), pp. 773–777. doi:10.1111/j.1600-0463.2012.02899.x.

Dridi, B., Fardeau, M.L., *et al.* (2012) 'Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces', *International Journal of Systematic and Evolutionary Microbiology*, 62(8), pp. 1902–1907. doi:10.1099/ijs.0.033712-0.

Dridi, B., Raoult, D. and Drancourt, M. (2011) 'Archaea as emerging organisms in complex human microbiomes', *Anaerobe*, 17(2), pp. 56–63. doi:10.1016/j.anaerobe.2011.03.001.

Drochner, W., Kerler, A. and Zacharias, B. (2004) 'Pectin in pig nutrition, a comparative review', *Journal of animal physiology and animal nutrition*, 88(11-12), pp. 367–380.

Duncan, S.H. *et al.* (2006) 'Proposal of Roseburia faecis sp. nov., Roseburia hominis sp. nov. and Roseburia inulinivorans sp. nov., based on isolates from human faeces', *International Journal of Systematic and Evolutionary Microbiology*, 56(10), pp. 2437–2441. doi:10.1099/IJS.0.64098-0.

Edwards, J.E. *et al.* (2020) 'Multi-kingdom characterization of the core equine fecal microbiota based on multiple equine (sub)species', *Animal Microbiome*, 2(1), pp. 1–16. doi:10.1186/s42523-020-0023-1.

Enzmann, F., Mayer, F., Rother, M. and Holtmann, D. (2018) 'Methanogens: biochemical background and biotechnological applications', *AMB Express*, 8(1), pp. 1–22. doi:10.1186/s13568-017-0531-x.

Evans, P.N. *et al.* (2009) 'Community composition and density of Methanogens in the foregut of the Tammar Wallaby (Macropus eugenii)', *Applied and Environmental Microbiology*,

75(8), pp. 2598–2602. doi:10.1128/AEM.02436-08.

Facey, H. V., Northwood, K.S. and Wright, A.D.G. (2012) 'Molecular Diversity of Methanogens in Fecal Samples From Captive Sumatran Orangutans (Pongo abelii)', *American Journal of Primatology*, 74(5), pp. 408–413. doi:10.1002/ajp.21992.

Feldewert, C., Lang, K. and Brune, A. (2021) 'The hydrogen threshold of obligately methyl-reducing methanogens', *FEMS Microbiology Letters*, 367(17), pp. 1–7. doi:10.1093/femsle/fnaa137.

Fennema, D., Phillips, I.R. and Shephard, E.A. (2016) *Trimethylamine and trimethylamine N-oxide, a Flavin-Containing Monooxygenase 3 (FMO3)-mediated host-microbiome metabolic axis implicated in health and disease*, *Drug Metabolism and Disposition*. American Society for Pharmacology and Experimental Therapy. doi:10.1124/dmd.116.070615.

Fernandes, K.A. *et al.* (2014) 'Faecal microbiota of forage-fed horses in new zealand and the population dynamics of microbial communities following dietary change', *PLoS ONE*, 9(11). doi:10.1371/journal.pone.0112846.

Filippo, C. De *et al.* (2010) 'Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa', *Proceedings of the National Academy of Sciences*, 107(33), pp. 14691–14696. doi:10.1073/PNAS.1005963107.

Fisher, R.M. *et al.* (2017) 'The evolution of host-symbiont dependence', *Nature Communications*, 8(1), pp. 1–8. doi:10.1038/ncomms15973.

Flint, H.J. *et al.* (2012) 'Microbial degradation of complex carbohydrates in the gut', *Gut Microbes*, 3(4), p. 289. doi:10.4161/GMIC.19897.

Frey, J.C. *et al.* (2010) 'Comparative studies of microbial populations in the rumen, duodenum, ileum and faeces of lactating dairy cows', *Journal of Applied Microbiology*, 108(6), pp. 1982–1993. doi:10.1111/j.1365-2672.2009.04602.x.

Friedman, J. and Alm, E.J. (2012) 'Inferring Correlation Networks from Genomic Survey Data', *PLoS Computational Biology*, 8(9), p. 1002687. doi:10.1371/journal.pcbi.1002687.

Friedrich, M.W., Schmitt-Wagner, D., Lueders, T. and Brune, A. (2001) 'Axial Differences in Community Structure of Crenarchaeota and Euryarchaeota in the Highly Compartmentalized Gut of the Soil-Feeding Termite Cubitermes orthognathus', *Applied and Environmental Microbiology*, 67(10), pp. 4880–4890. doi:10.1128/AEM.67.10.4880-

4890.2001.

Gaci, N. *et al.* (2014) 'Archaea and the human gut: New beginning of an old story', *World Journal of Gastroenterology*, 20(43), pp. 16062–16078. doi:10.3748/wjg.v20.i43.16062.

Gibson, G.R., MacFarlane, G.T. and Cummings, J.H. (1993) 'Sulphate reducing bacteria and hydrogen metabolism in the human large intestine', *Gut*, 34(4), pp. 437–439. doi:10.1136/gut.34.4.437.

Gibson, K.M. *et al.* (2019) 'Gut microbiome differences between wild and captive black rhinoceros – implications for rhino health', *Scientific Reports*, 9(1), pp. 1–11. doi:10.1038/s41598-019-43875-3.

Glad, T. *et al.* (2010) 'Ecological Characterisation of the Colonic Microbiota in Arctic and Sub-Arctic Seals', *Microbial Ecology*, 60(2), pp. 320–330. doi:10.1007/s00248-010-9690-x.

Goodrich, J.K. *et al.* (2014) 'Human Genetics Shape the Gut Microbiome', *Cell*, 159(4), pp. 789–799. doi:10.1016/J.CELL.2014.09.053.

Goodrich, J.K. *et al.* (2016) 'Genetic Determinants of the Gut Microbiome in UK Twins', *Cell Host & Microbe*, 19(5), pp. 731–743. doi:10.1016/J.CHOM.2016.04.017.

Gordon, J.I., Dewey, K.G., Mills, D.A. and Medzhitov, R.M. (2012) 'The human gut microbiota and undernutrition', *Sci Transl Med*, 4(137), p. 137ps12. doi:10.1126/scitranslmed.3004347.

Greening, C. *et al.* (2019) 'Diverse hydrogen production and consumption pathways influence methane production in ruminants', *ISME Journal*, 13(10), pp. 2617–2632. doi:10.1038/s41396-019-0464-2.

Gribaldo, S. and Brochier-Armanet, C. (2006) 'The origin and evolution of Archaea: a state of the art', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1470), p. 1007. doi:10.1098/RSTB.2006.1841.

Grieneisen, L. *et al.* (2021) 'Gut microbiome heritability is nearly universal but environmentally contingent', *Science*, 373(6551), pp. 181–186. doi:10.1126/SCIENCE.ABA5483.

Groussin, M. *et al.* (2017) 'Unraveling the processes shaping mammalian gut microbiomes over evolutionary time', *Nature Communications*, 8. doi:10.1038/ncomms14319.

Groussin, M., Mazel, F. and Alm, E.J. (2020) 'Co-evolution and Co-speciation of Host-Gut Bacteria Systems', *Cell Host and Microbe*, 28(1), pp. 12–22. doi:10.1016/j.chom.2020.06.013.

Guerry, P. (2011) 'N-linked glycosylation in Archaea: two paths to the same glycan', *Molecular Microbiology*, 81(5), pp. 1133–1135. doi:10.1111/J.1365-2958.2011.07782.X.

Guo, N. *et al.* (2021) 'Seasonal dynamics of diet–gut microbiota interaction in adaptation of yaks to life at high altitude', *Biofilms and Microbiomes*, 7(1), pp. 1–11. doi:10.1038/s41522-021-00207-6.

Gutiérrez, G. (2012) 'Draft genome sequence of methanobacterium formicicum DSM 3637, an archaebacterium isolated from the methane producer amoeba pelomyxa palustris', *Journal of Bacteriology*, 194(24), pp. 6967–6968. doi:10.1128/JB.01829-12.

Guzman, C.E., Bereza-Malcolm, L.T., Groef, B. De and Franks, A.E. (2015) 'Presence of selected methanogens, fibrolytic bacteria, and Proteobacteria in the gastrointestinal tract of neonatal dairy calves from birth to 72 hours', *PLOS ONE*, 10(7), p. e0133048. doi:10.1371/JOURNAL.PONE.0133048.

Hackstein, J. and Stumm, C.K. (1994) 'Methane production in terrestrial arthropods.', *Proceedings of the National Academy of Sciences of the United States of America*, 91(12), pp. 5441–5445. doi:10.1073/pnas.91.12.5441.

Hackstein, J.H.P. (1997) 'Genetic and evolutionary aspects of methanogenesis', *EDP Sciences*, 37, pp. 5–8.

Hackstein, J.H.P. and van Alen, T.A. (1996) 'Fecal Methanogens and Vertebrate Evolution', *Evolution*, 50(2), pp. 559–572. doi:10.2307/2410831.

Hammer, T.J., Sanders, J.G. and Fierer, N. (2019) 'Not all animals need a microbiome', *FEMS microbiology letters*, 366(10), pp. 1–11. doi:10.1093/femsle/fnz117.

Hansen, E.E. *et al.* (2011) 'Pan-genome of the dominant human gut-associated archaeon, Methanobrevibacter smithii, studied in twins', *Proceedings of the National Academy of Sciences*, 108(Supplement_1), pp. 4599–4606. doi:10.1073/pnas.1000071108.

Henderson, G. *et al.* (2015) 'Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range', *Scientific Reports*, 5(October), p. 14567. doi:10.1038/srep14567.

Hervé, V. *et al.* (2020) 'Phylogenomic analysis of 589 metagenome-assembled genomes encompassing all major prokaryotic lineages from the gut of higher termites', *PeerJ*, 2020(2), p. e8614. doi:10.7717/peerj.8614.

Hicks, A.L. *et al.* (2018) 'Gut microbiomes of wild great apes fluctuate seasonally in response to diet', *Nature Communication*, 9(1), pp. 1–18. doi:10.1038/s41467-018-04204-w.

Hird, S.M., Sánchez, C., Carstens, B.C. and Brumfield, R.T. (2015) 'Comparative Gut Microbiota of 59 Neotropical Bird Species', *Frontiers in Microbiology*, 0(DEC), p. 1403. doi:10.3389/FMICB.2015.01403.

Hoedt, E.C. *et al.* (2016) 'Differences down-under: Alcohol-fueled methanogenesis by archaea present in Australian macropodids', *ISME Journal*, 10(10), pp. 2376–2388. doi:10.1038/ismej.2016.41.

Hoffmann, C. *et al.* (2013) 'Archaea and Fungi of the Human Gut Microbiome: Correlations with Diet and Bacterial Residents', *PLoS ONE*, 8(6), p. 66019. doi:10.1371/journal.pone.0066019.

Hong, P.Y., Wheeler, E., Cann, I.K.O. and Mackie, R.I. (2011) 'Phylogenetic analysis of the fecal microbial community in herbivorous land and marine iguanas of the Galápagos Islands using 16S rRNA-based pyrosequencing', *ISME Journal*, 5(9), pp. 1461–1470. doi:10.1038/ismej.2011.33.

Hooper, L. V. and Gordon, J.I. (2001) 'Glycans as legislators of host–microbial interactions: spanning the spectrum from symbiosis to pathogenicity', *Glycobiology*, 11(2), pp. 1R-10R. doi:10.1093/GLYCOB/11.2.1R.

Hooper, L. V., Littman, D.R. and Macpherson, A.J. (2012) 'Interactions between the microbiota and the immune system', *Science*. American Association for the Advancement of Science, pp. 1268–1273. doi:10.1126/science.1223490.

Hristov, A. *et al.* (2015) 'An inhibitor persistently decrease enteric methane emission from dairy cows with no negative effect on milk production', *PNAS*, 112(34), pp. 10663–10668.

Hudson, M.J. and Roberts, A.K. (1993) 'Establishment of Methanogens in the Infant Intestine', *MICROBIAL ECOLOGY IN HEALTH AND DISEASE*, 6, pp. 30–31.

Hungate, R.E. (1966) *The rumen and its microbes*. Elsevier.

Huse, S.M. *et al.* (2007) 'Accuracy and quality of massively parallel DNA pyrosequencing', *Genome biology*, 8(7), pp. 1–9.

Jackson, S.A. *et al.* (2013) 'Archaea appear to dominate the microbiome of Inflatella pellicula deep sea sponges', *PLoS ONE*, 8(12), pp. 1–8. doi:10.1371/journal.pone.0084438.

Jarrell, K.F., Ding, Y., Nair, D.B. and Siu, S. (2013) 'Surface appendages of archaea: Structure, function, genetics and assembly', *Life*. Multidisciplinary Digital Publishing Institute, pp. 86–117. doi:10.3390/life3010086.

Jarvis, G.N. *et al.* (2000) 'Isolation and identification of ruminal methanogens from grazing cattle', *Current Microbiology*, 40(5), pp. 327–332. doi:10.1007/s002849910065.

Jeffery, I.B., Lynch, D.B. and O'Toole, P.W. (2016) 'Gut microbiota and aging'. doi:10.1038/ismej.2015.88.

Jensen, B.B. and Jørgensen, H. (1994) 'Effect of dietary fiber on microbial activity and microbial gas production in various regions of the gastrointestinal tract of pigs.', *Applied and Environmental Microbiology*, 60(6), p. 1897. Available at: /pmc/articles/PMC201578/?report=abstract (Accessed: 1 September 2021).

Jiang, H.-Y. *et al.* (2017) 'Diets Alter the Gut Microbiome of Crocodile Lizards', *Frontiers in Microbiology*, 0(OCT), p. 2073. doi:10.3389/FMICB.2017.02073.

Jones, B. V. *et al.* (2008) 'Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome', *Proceedings of the National Academy of Sciences*, 105(36), pp. 13580–13585. doi:10.1073/PNAS.0804437105.

Jung, M. *et al.* (2016) 'A hydrophobic ammonia-oxidizing archaeon of the Nitrosocosmicus clade isolated from coal tar-contaminated sediment', *Environmental microbiology reports*, 8(6), pp. 983–992.

Kageyama, A. and Benno, Y. (2000) 'Phylogenic and phenotypic characterization of some Eubacterium-like isolates from human feces: Description of Solobacterium moorei gen. nov., sp. nov.', *Microbiology and Immunology*, 44(4), pp. 223–227. doi:10.1111/J.1348-0421.2000.TB02487.X.

Kalantar-Zadeh, K. *et al.* (2018) 'A human pilot trial of ingestible electronic capsules capable of sensing different gases in the gut', *Nature Electronics*, 1, pp. 79–87. doi:10.1038/s41928-017-0004-x.

Kamada, N., Chen, G.Y., Inohara, N. and Núñez, G. (2013) 'Control of pathogens and pathobionts by the gut microbiota', *Nature Immunology*. Nature Publishing Group, pp. 685–690. doi:10.1038/ni.2608.

Kandiba, L. and Eichler, J. (2013) 'Analysis of putative nonulosonic acid biosynthesis pathways in Archaea reveals a complex evolutionary history', *FEMS Microbiology Letters*, 345(2), pp. 110–120. doi:10.1111/1574-6968.12193.

Kapitan, M. *et al.* (2018) 'Fungi as Part of the Microbiota and Interactions with Intestinal Bacteria.', in *Fungal Physiology and Immunopathogenesis*. Springer, Cham. doi:10.1007/82_2018_117.

Karasov, W.H., Petrossian, E., Rosenberg, L. and Diamond, J.M. (1986) 'How do food passage rate and assimilation differ between herbivorous lizards and nonruminant mammals?', *Journal of Comparative Physiology B*, 156(4), pp. 599–609.

Katoh, K. and Standley, D.M. (2013) 'MAFFT multiple sequence alignment software version 7: Improvements in performance and usability', *Molecular Biology and Evolution*, 30(4), pp. 772–780. doi:10.1093/molbev/mst010.

Kelly, W.J. *et al.* (2014) 'The complete genome sequence of the rumen methanogen Methanobacterium formicicum BRM9', *Standards in Genomic Sciences*, 9(1), pp. 1–8. doi:10.1186/1944-3277-9-15.

Kelly, W.J., Li, D., *et al.* (2016) 'Draft genome sequence of the rumen methanogen Methanobrevibacter olleyae YLM1', *Genome Announcements*, 4(2), pp. 4–5. doi:10.1128/genomeA.00232-16.

Kelly, W.J., Pacheco, D.M., *et al.* (2016) 'The complete genome sequence of the rumen methanogen Methanobrevibacter millerae SM9', *Standards in Genomic Sciences*, 11(1), pp. 1–9. doi:10.1186/s40793-016-0171-9.

Kelly, W.J. *et al.* (2019) 'Occurrence and expression of genes encoding methyl-compound production in rumen bacteria', 1(1), pp. 1–13. doi:10.1186/S42523-019-0016-0.

Kerou, M. and Schleper, C. (2015) 'Nitrososphaera', *Bergey's manual of systematics of archaea and bacteria*, pp. 1–10.

Khelaifia, S. *et al.* (2018) 'Genome sequence and description of Haloferax massiliense sp. nov., a new halophilic archaeon isolated from the human gut', *Extremophiles*, 22(3), pp. 485–498. doi:10.1007/s00792-018-1011-1.

Khelaifia, S., Caputo, A., Djossou, F. and Raoult, D. (2017) 'Draft genome sequence of a human-associated isolate of Haloferax alexandrinus strain Arc-hr, an extremely halophilic archaea', *New Microbes and New Infections*, 15(Umr 7278), pp. 44–45. doi:10.1016/j.nmni.2016.11.012.

Khelaifia, S. and Raoult, D. (2016) 'Haloferax massiliensis sp. nov., the first human-associated halophilic archaea', *New Microbes and New Infections*, 12(Umr 7278), pp. 96–98. doi:10.1016/j.nmni.2016.05.007.

Kim, J.J.Y.J.-H. *et al.* (2020) 'The human gut archaeome: identification of diverse haloarchaea in Korean subjects', *Microbiome 2020 8:1*, 8(1), pp. 1–17. doi:10.1186/S40168-020-00894-X.

Kittelmann, S. *et al.* (2013) 'Simultaneous Amplicon Sequencing to Explore Co-Occurrence Patterns of Bacterial, Archaeal and Eukaryotic Microorganisms in Rumen Microbial Communities', *PLoS ONE*, 8(2). doi:10.1371/journal.pone.0047879.

Klemm, P. and Schembri, M.A. (2000) 'Bacterial adhesins: Function and structure', *International Journal of Medical Microbiology*, 290(1), pp. 27–35. doi:10.1016/S1438-4221(00)80102-2.

Koeth, R.A. *et al.* (2013) 'Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis', *Nature medicine*, 19(5), pp. 576–585.

Koh, A. and Bäckhed, F. (2020) 'From Association to Causality: the Role of the Gut Microbiota and Its Functional Products on Host Metabolism', *Molecular Cell*. Cell Press, pp. 584–596. doi:10.1016/j.molcel.2020.03.005.

Kohl, K.D. *et al.* (2017) 'Gut microbial ecology of lizards: insights into diversity in the wild, effects of captivity, variation across gut regions and transmission', *Molecular Ecology*, 26(4), pp. 1175–1189. doi:10.1111/MEC.13921.

Kopeýný, J., Hajer, J. and Mrázek, J. (2004) 'Detection of Cellulolytic Bacteria from the Human Colon', *Folia Microbiol*, 49(2), pp. 175–177. Available at: http://www.biomed.cas.cz/mbu/folia/ (Accessed: 21 September 2021).

Koskinen, K. *et al.* (2017) 'First Insights into the Diverse Human Archaeome : Specific Detection of Archaea in the Gastrointestinal Tract ... crossm First Insights into the Diverse Human Archaeome : Specific Detection of Archaea in the Gastrointestinal Tract ', *mBio*, 8(6), pp. 1–17. doi:10.1128/mBio.00824-17.

Kumar, S., Stecher, G., Suleski, M. and Hedges, S.B. (2017) 'TimeTree: A Resource for Timelines, Timetrees, and Divergence Times', *Molecular biology and evolution*, 34(7), pp. 1812–1819. doi:10.1093/molbev/msx116.

Kunkel, D. *et al.* (2011) 'Methane on breath testing is associated with constipation: a systematic review and meta-analysis', *Digestive diseases and sciences*, 56(6), pp. 1612–1618. doi:10.1007/S10620-011-1590-5.

Kurtz, Z.D. *et al.* (2015) 'Sparse and Compositionally Robust Inference of Microbial Ecological Networks', *PLoS Computational Biology*, 11(5), pp. 1–25. doi:10.1371/journal.pcbi.1004226.

Kushkevych, I., Vítězová, M., Vítěz, T. and Bartoš, M. (2017) 'Production of biogas: Relationship between methanogenic and sulfate-reducing microorganisms', *Open Life*

Sciences, 12(1), pp. 82–91. doi:10.1515/biol-2017-0009.

de la Cuesta-Zuluaga, J., Spector, T.D., Youngblut, N.D. and Ley, R.E. (2021) 'Genomic Insights into Adaptations of Trimethylamine-Utilizing Methanogens to Diverse Habitats, Including the Human Gut', *mSystems*, 6(1). doi:10.1128/MSYSTEMS.00939-20.

Lagkouvardos, I. *et al.* (2019) 'Sequence and cultivation study of Muribaculaceae reveals novel species, host preference, and functional potential of this yet undescribed family', *Microbiome*, 7(1), pp. 1–15. doi:10.1186/s40168-019-0637-2.

Lambie, S.C. *et al.* (2015) 'The complete genome sequence of the rumen methanogen Methanosarcina barkeri CM1', *Standards in Genomic Sciences*, 10, p. 57. doi:10.1186/s40793-015-0038-5.

Lane, D.J. (1991) '16S/23S rRNA sequencing', *Nucleic acid techniques in bacterial systematics*, pp. 115–175.

Lang, K. *et al.* (2015) 'New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of "Candidatus Methanoplasma termitum"', *Applied and Environmental Microbiology*, 81(4), pp. 1338–1352. doi:10.1128/AEM.03389-14.

Langer, P. (2017) *Comparative anatomy of the gastrointestinal tract in Eutheria I: Taxonomy, Biogeography and Food: Afrotheria, Xenarthra and Euarchontoglires*. Walter de Gruyter GmbH & Co KG.

Leadbetter, J.R. and Breznak, J.A. (1996) 'Physiological ecology of Methanobrevibacter cuticularis sp. nov. and Methanobrevibacter curvatus sp. nov., isolated from the hindgut of the termite Reticulitermes flavipes', *Applied and Environmental Microbiology*, 62(10), pp. 3620–3631.

Leadbetter, J.R., Crosby, L.D. and Breznak, J.A. (1998) 'Methanobrevibacter filiformis sp. nov., a filamentous methanogen from termite hindguts', *Archives of Microbiology*, 169(4), pp. 287–292. doi:10.1007/s002030050574.

Leadbetter, J.R., Schmidt, T.M., Graber, J.R. and Breznak, J.A. (1999) 'Acetogenesis from H2 plus CO2 by spirochetes from termite guts', *Science*, 283(5402), pp. 686–689. doi:10.1126/science.283.5402.686.

Leahy, S.C. *et al.* (2010) 'The Genome Sequence of the Rumen Methanogen Methanobrevibacter ruminantium Reveals New Possibilities for Controlling Ruminant Methane Emissions', *PLoS ONE*, 5(1), p. e8926. doi:10.1371/journal.pone.0008926.

Leahy, S.C. *et al.* (2013) 'The Complete genome sequence of Methanobrevibacter sp. ABM4', *Standards in Genomic Sciences*, 8(2), pp. 215–227. doi:10.4056/sigs.3977691.

Lehtovirta-Morley, L.E. *et al.* (2016) 'Isolation of "Candidatus Nitrosocosmicus franklandus", a novel ureolytic soil archaeal ammonia oxidiser with tolerance to high ammonia concentration', *FEMS Microbiology Ecology*, 92(5).

Letunic, I. and Bork, P. (2019) 'Interactive Tree of Life (iTOL) v4: Recent updates and new developments', *Nucleic Acids Research*, 47(W1), pp. 256–259. doi:10.1093/nar/gkz239.

Levitt, M.D., Furne, J.K., Kuskowski, M. and Ruddy, J. (2006) 'Stability of Human Methanogenic Flora Over 35 Years and a Review of Insights Obtained From Breath Methane Measurements', *Clinical Gastroenterology and Hepatology*, 4(2), pp. 123–129. doi:10.1016/J.CGH.2005.11.006.

Lewis, A.L. *et al.* (2009) 'Innovations in host and microbial sialic acid biosynthesis revealed by phylogenomic prediction of nonulosonic acid structure', *Proceedings of the National Academy of Sciences of the United States of America*, 106(32), pp. 13552–13557. doi:10.1073/PNAS.0902431106.

Lewis, W.H., Sendra, K.M., Embley, T.M. and Esteban, G.F. (2018) 'Morphology and Phylogeny of a New Species of Anaerobic Ciliate, Trimyema finlayi n. sp., with Endosymbiotic Methanogens', *Frontiers in Microbiology*, 9(February), p. 140. doi:10.3389/fmicb.2018.00140.

Ley, R.E. *et al.* (2008) 'Evolution of mammals and their gut microbes', *Science*, 320(5883), pp. 1647–1651. doi:10.1126/science.1155725.Evolution.

Ley, R.E., Peterson, D.A. and Gordon, J.I. (2006) 'Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine', *Cell*, 124(4), pp. 837–848. doi:10.1016/J.CELL.2006.02.017.

Li, Y. *et al.* (2016) 'The complete genome sequence of the methanogenic archaeon ISO4-H5 provides insights into the methylotrophic lifestyle of a ruminal representative of the Methanomassiliicoccales', *Standards in Genomic Sciences*, 11(1), pp. 1–12. doi:10.1186/s40793-016-0183-5.

Li, Z. *et al.* (2019) 'Heterogeneous development of methanogens and the correlation with bacteria in the rumen and cecum of sika deer (Cervus nippon) during early life suggest different ecology relevance', *BMC Microbiology*, 19(1), pp. 1–13. doi:10.1186/s12866-019-1504-9.

Lim, S.J. and Bordenstein, S.R. (2020) 'An introduction to phylosymbiosis', *Proceedings of the Royal Society B: Biological Sciences: Biological Sciences*, 287(1922), p. 287. doi:10.1098/rspb.2019.2900.

Lin, M.F., Chang, C., Yang, C.W. and Dierenfeld, E.S. (2015) 'Aspects of digestive anatomy, feed intake and digestion in the Chinese pangolin (Manis pentadactyla) at Taipei zoo', *Zoo biology*, 34(3), pp. 262–270.

Lind, A.E. *et al.* (2018) 'Genomes of two archaeal endosymbionts show convergent adaptations to an intracellular lifestyle', *The ISME Journal*, p. 1. doi:10.1038/s41396-018-0207-9.

Lindinger, W. *et al.* (1997) 'Endogenous production of methanol after the consumption of fruit', *Alcoholism: Clinical and Experimental Research*, 21(5), pp. 939–943.

Liu, C. *et al.* (2017) 'Role of Age-Related Shifts in Rumen Bacteria and Methanogens in Methane Production in Cattle', *Frontiers in Microbiology*, 0(AUG), p. 1563. doi:10.3389/FMICB.2017.01563.

Liu, Y. and Whitman, W.B. (2008) 'Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea', *Annals of the New York Academy of Sciences*, 1125, pp. 171–189. doi:10.1196/annals.1419.019.

Lloyd, D., Hillman, N., Yarlett, N. and Williams, A.G. (1989) 'Hydrogen production by rumen holotrich protozoa: effects of oxygen and implications for metabolic control by in situ conditions', *The Journal of protozoology*, 36(2), pp. 205–213.

Loh, H.Q., Hervé, V. and Brune, A. (2021) *Metabolic Potential for Reductive Acetogenesis and a Novel Energy-Converting [NiFe] Hydrogenase in Bathyarchaeia From Termite Guts – A Genome-Centric Analysis*, *Frontiers in Microbiology*. doi:10.3389/FMICB.2020.635786.

López-Legentil, S., Erwin, P.M., Pawlik, J.R. and Song, B. (2010) 'Effects of Sponge Bleaching on Ammonia-Oxidizing Archaea: Distribution and Relative Expression of Ammonia Monooxygenase Genes Associated with the Barrel Sponge Xestospongia muta', *Microbial Ecology*, 60(3), pp. 561–571. doi:10.1007/s00248-010-9662-1.

Louis, P. and Flint, H.J. (2017) 'Formation of propionate and butyrate by the human colonic microbiota', *Environmental Microbiology*, 19(1), pp. 29–41. doi:10.1111/1462-2920.13589.

Lovley, D.R. (1985) 'Minimum Threshold for Hydrogen Metabolism in Methanogenic Bacteria', *Applied and environmental microbiology*, 49(6), pp. 1530–1531.

Lovley, D.R., Dwyer, D.F. and Klug, M.J. (1982) 'Kinetic analysiis of competition between sulfate reducers and methanogens for hydrogen in sediments', *Applied and Environmental Microbiology*, 43(6), pp. 1373–1379. doi:10.1128/AEM.43.6.1373-1379.1982.

Lovley, D.R. and Goodwin, S. (1988) 'Hydrogen concentrations as an indicator of the predominant terminal electron-accepting reactions in aquatic sediments', *Geochimica et Cosmochimica Acta*, 52(12), pp. 2993–3003. doi:10.1016/0016-7037(88)90163-9.

Lozupone, C.A. *et al.* (2012) 'Diversity, stability and resilience of the human gut microbiota', *Nature*, 489(7415), pp. 220–230. doi:10.1038/nature11550.

Lurie-Weinberger, M.N., Peeri, M. and Gophna, U. (2012) 'Contribution of lateral gene transfer to the gene repertoire of a gut-adapted methanogen', *Genomics*, 99(1), pp. 52–58. doi:10.1016/j.ygeno.2011.10.005.

Lurie-Weinberger, M.N., Peeri, M., Tuller, T. and Gophna, U. (2012) 'Extensive inter-domain lateral gene transfer in the evolution of the human commensal Methanosphaera stadtmanae', *Frontiers in Genetics*, 3, pp. 1–11. doi:10.3389/fgene.2012.00182.

Van Der Maarel, M.J.E.C., Artz, R.R.E., Haanstra, R. and Forney, L.J. (1998) 'Association of marine archaea with the digestive tracts of two marine fish species', *Applied and Environmental Microbiology*, 64(8), pp. 2894–2898.

Mackie, R.I. (2002) 'Mutualistic fermentative digestion in the gastrointestinal tract: diversity and evolution', *Integrative and Comparative Biology*, 42(2), pp. 319–326.

Maczulak, A.E., Wolin, M.J. and Miller, T.L. (2014) 'Increase in Colonic Methanogens and Total Anaerobes in Aging Rats', *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, 55, pp. 2468–2473.

Maeder, D.L. *et al.* (2006) 'The Methanosarcina barkeri genome: Comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes', *Journal of Bacteriology*, 188(22), pp. 7922–7931. doi:10.1128/JB.00810-06.

Manos, J.I.M. and Belas, R. (2006) 'The Genera *Proteus*, *Providencia*, and *Morganella*', pp. 245–269.

Maurice, C.F. *et al.* (2015) 'Marked seasonal variation in the wild mouse gut microbiota', *The ISME Journal 2015 9:11*, 9(11), pp. 2423–2434. doi:10.1038/ismej.2015.53.

McKenney, E.A., O'Connell, T.M., Rodrigo, A. and Yoder, A.D. (2018) 'Feeding strategy shapes

gut metagenomic enrichment and functional specialization in captive lemurs', *Gut Microbes*, 9(3), pp. 202–217. doi:10.1080/19490976.2017.1408762.

McKenzie, V.J. *et al.* (2017) 'The effects of captivity on the mammalian gut microbiome', *Integrative and Comparative Biology*, 57(4), pp. 690–704. doi:10.1093/icb/icx090.

Metchnikoff *et al.* (1909) 'Bacteriotherapie, vaccination, serotherapie', *Medicaments microbiens* [Preprint].

Mihajlovski, A. *et al.* (2010) 'Molecular evaluation of the human gut methanogenic archaeal microbiota reveals an age-associated increase of the diversity', 2(2), pp. 272–280. doi:10.1111/J.1758-2229.2009.00116.X.

Mihajlovski, A., Alric, M. and Brugère, J.F. (2008) 'A putative new order of methanogenic Archaea inhabiting the human gut, as revealed by molecular analyses of the mcrA gene', *Research in Microbiology*, 159(7–8), pp. 516–521. doi:10.1016/j.resmic.2008.06.007.

Miller, T.L. and Wolin, M.J. (1983) 'Stability of Methanobrevibacter smithii populations in the microbial flora excreted from the human large bowel', *Applied and Environmental Microbiology*, 45(1), pp. 317–318. doi:10.1128/AEM.45.1.317-318.1983.

Miller, T.L. and Wolin, M.J. (1985) 'Methanosphaera stadtmaniae gen. nov., sp. nov.: a species that forms methane by reducing methanol with hydrogen', 141(2), pp. 116–122. doi:10.1007/BF00423270.

Miller, T.L. and Wolin, M.J. (1986) 'Methanogens in human and animal intestinal Tracts', *Systematic and Applied Microbiology*, 7(2–3), pp. 223–229. doi:10.1016/S0723-2020(86)80010-8.

Miller, T.L., Wolin, M.J., Demacario, E.C. and Macario, A.J.L. (1982) 'Isolation of Methanobrevibacter-smithii from Human Feces', *Applied and Environmental Microbiology*, 43(1), pp. 227–232. doi:10.1128/AEM.43.1.227-232.1982.

Milliken, W., Albert, B. and Gomez, G.G. (1999) *Yanomami: a forest people*. Royal Botanic Gardens, Kew.

Mittl, P.R.E. and Schneider-Brachert, W. (2007) 'Sel1-like repeat proteins in signal transduction', *Cellular Signalling*. Pergamon, pp. 20–31. doi:10.1016/j.cellsig.2006.05.034.

Mizrahi, I., Wallace, R.J. and Moraïs, S. (2021) 'The rumen microbiome: balancing food security and environmental impacts', *Nature Reviews Microbiology*, pp. 1–14.

doi:10.1038/s41579-021-00543-6.

Moeller, A.H. *et al.* (2016) 'Cospeciation of gut microbiota with hominids', *Science*, 353(6297), pp. 380–382. doi:10.1126/SCIENCE.AAF3951.

Moitinho-Silva, L. *et al.* (2017) 'Integrated metabolism in sponge-microbe symbiosis revealed by genome-centered metatranscriptomics', *ISME Journal*, 11(7), pp. 1651–1666. doi:10.1038/ismej.2017.25.

Moraïs, S. and Mizrahi, I. (2019) 'Islands in the stream: From individual to communal fiber degradation in the rumen ecosystem', *FEMS Microbiology Reviews*. Oxford University Press, pp. 362–369. doi:10.1093/femsre/fuz007.

Moran, N.A., McLaughlin, H.J. and Sorek, R. (2009) 'The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria', *Science*, 323(5912), pp. 379–382. doi:10.1126/science.1167140.

Morgavi, D.P., Martin, C., Jouany, J.P. and Ranilla, M.J. (2012) 'Rumen protozoa and methanogenesis: Not a simple cause-effect relationship', *British Journal of Nutrition*, 107, pp. 388–397. doi:10.1017/S0007114511002935.

Morvan, B., Bonnemoy, F., Fonty, G. and Gouet, P. (1996) 'Quantitative determination of H2-utilizing acetogenic and sulfate-reducing bacteria and methanogenic archaea from digestive tract of different mammals', *Current Microbiology*, 32(3), pp. 129–133. doi:10.1007/s002849900023.

Mosoni, P., Chaucheyras-Durand, F., Béra-Maillet, C. and Forano, E. (2007) 'Quantification by real-time PCR of cellulolytic bacteria in the rumen of sheep after supplementation of a forage diet with readily fermentable carbohydrates: Effect of a yeast additive', *Journal of Applied Microbiology*, 103(6), pp. 2676–2685. doi:10.1111/j.1365-2672.2007.03517.x.

Mowat, A.M.I. (2021) 'Historical Perspective: Metchnikoff and the intestinal microbiome', *Journal of Leukocyte Biology*, 109(3), pp. 513–517. doi:10.1002/JLB.4RI0920-599.

Moya, A., Peretó, J., Gil, R. and Latorre, A. (2008) 'Learning how to live together: Genomic insights into prokaryote-animal symbioses', *Nature Reviews Genetics*. Nat Rev Genet, pp. 218–229. doi:10.1038/nrg2319.

Nagpal, R. *et al.* (2018) 'Gut microbiome and aging: Physiological and mechanistic insights', *Nutrition and healthy aging*, 4(4), pp. 267–285. doi:10.3233/NHA-170030.

Nakamura, N. *et al.* (2011) 'Analysis of the hydrogenotrophic microbiota of wild and captive

black howler monkeys (Alouatta pigra) in palenque national park, Mexico', *American Journal of Primatology*, 73(9), pp. 909–919. doi:10.1002/AJP.20961.

Nakamura, N., Leigh, S.R., Mackie, R.I. and Gaskins, H.R. (2009) 'Microbial community analysis of rectal methanogens and sulfate reducing bacteria in two non-human primate species', *Journal of Medical Primatology*, 38(5), pp. 360–370. doi:10.1111/j.1600-0684.2009.00361.x.

Nam, Y. Do *et al.* (2008) 'Bacterial, archaeal, and eukaryal diversity in the intestines of Korean people', *Journal of Microbiology*, 46(5), pp. 491–501. doi:10.1007/s12275-008-0199-7.

Nelson, T.M., Rogers, T.L., Carlini, A.R. and Brown, M. V. (2013) 'No Title', *Environmental Microbiology*, 15(4), pp. 1132–1145. doi:10.1111/1462-2920.12022.

Nercessian, O. *et al.* (2005) 'Diversity of Bacteria and Archaea associated with a carbonate-rich metalliferous sediment sample from the Rainbow vent field on the Mid-Atlantic Ridge', *Environmental Microbiology*, 7(5), pp. 698–714.

Ng, F. *et al.* (2016) 'An adhesin from hydrogen-utilizing rumen methanogen Methanobrevibacter ruminantium M1 binds a broad range of hydrogen-producing microorganisms', *Environmental microbiology*, 18(9), pp. 3010–3021. doi:10.1111/1462-2920.13155.

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A. and Minh, B.Q. (2015) 'IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies', *Molecular Biology and Evolution*, 32(1), pp. 268–274. doi:10.1093/molbev/msu300.

Nicol, G.W. *et al.* (2019) 'Genome Sequence of " Candidatus Nitrosocosmicus franklandus" C13, a Terrestrial Ammonia-Oxidizing Archaeon', *Microbiology Resource Announcements*, 8(40). doi:10.1128/MRA.00435-19.

Nishida, A.H. and Ochman, H. (2017) 'Rates of Gut Microbiome Divergence in Mammals', *Molecular Ecology*, 27(December), pp. 1884–1897. doi:10.1111/mec.14473.

Nkamga, V.D. *et al.* (2016) 'Methanobrevibacter smithii and Bacteroides thetaiotaomicron cultivated from a chronic paravertebral muscle abscess', *Clinical Microbiology and Infection*, 22(12), pp. 1008–1009. doi:10.1016/j.cmi.2016.09.007.

Nkamga, V.D., Henrissat, B. and Drancourt, M. (2017) 'Archaea: Essential inhabitants of the human digestive microbiota', *Human Microbiome Journal*, 3, pp. 1–8. doi:10.1016/j.humic.2016.11.005.

Nkrumah, J.D. *et al.* (2006) 'Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle', *Journal of Animal Science*, 84(1), pp. 145–153. doi:10.2527/2006.841145X.

Nonoh, J.O. (2013) *Archaeal diversity and community structure in the compartmented gut of higher termites*.

Nottingham, P.M. and Hungate, R.E. (1968) 'Isolation of methanogenic bacteria from feces of man.', *Journal of Bacteriology*, 96(6), pp. 2178–2179.

Odamaki, T. *et al.* (2016) 'Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study', *BMC microbiology*, 16(1), pp. 1–12.

Oremland, R.S. (1979) 'Methanogenic activity in plankton samples and fish intestines A mechanism for in situ methanogenesis in oceanic surface waters', *Limnology and Oceanography*, 24(6), pp. 1136–1141. doi:10.4319/lo.1979.24.6.1136.

Oren, A. (2014) 'The Family Methanocorpusculaceae', *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*, pp. 225–230. doi:10.1007/978-3-642-38954-2_314.

Ormerod, K.L. *et al.* (2016) 'Genomic characterization of the uncultured Bacteroidales family S24-7 inhabiting the guts of homeothermic animals', *Microbiome*, 4(1), pp. 1–17. doi:10.1186/s40168-016-0181-2.

Ouboter, H.T. *et al.* (2020) ' Draft Genome Sequence of a New Methanobacterium Strain Potentially Resistant to Bile Salts, Isolated from Deer Feces ', *Microbiology Resource Announcements*, 9(32), pp. 6–7. doi:10.1128/mra.00548-20.

Oxley, A.P.A. *et al.* (2010) 'Halophilic archaea in the human intestinal mucosa', *Environmental Microbiology*, 12(9), pp. 2398–2410. doi:10.1111/j.1462-2920.2010.02212.x.

Ozuolmez, D. *et al.* (2015) 'Methanogenic archaea and sulfate reducing bacteria co-cultured on acetate: Teamwork or coexistence?', *Frontiers in Microbiology*, 6(MAY), pp. 1–12. doi:10.3389/fmicb.2015.00492.

Parks, D.H. *et al.* (2015) 'CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes', *Genome Research*, 25(7), pp. 1043–1055. doi:10.1101/gr.186072.114.

Paul, K., Nonoh, J.O., Mikulski, L. and Brune, A. (2012) '"Methanoplasmatales," thermoplasmatales-related archaea in termite guts and other environments, are the

seventh order of methanogens', *Applied and Environmental Microbiology*, 78(23), pp. 8245–8253. doi:10.1128/AEM.02193-12.

Paynter, M.J. and Hungate, R.E. (1968) 'Characterization of Methanobacterium mobilis, sp. n., isolated from the bovine rumen.', *Journal of Bacteriology*, 95(5), pp. 1943–1951.

Pester, M., Schleper, C. and Wagner, M. (2011) 'The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology', *Current opinion in microbiology*, 14(3), pp. 300–306.

Pimentel, M. *et al.* (2006) 'Methane, a gas produced by enteric bacteria, slows intestinal transit and augments small intestinal contractile activity', *https://doi.org/10.1152/ajpgi.00574.2004*, 290(6), pp. 1089–1095. doi:10.1152/AJPGI.00574.2004.

Poehlein, A. *et al.* (2018) 'Comparative Genomic Analysis of Members of the Genera Methanosphaera and Methanobrevibacter Reveals Distinct Clades with Specific Potential Metabolic Functions', 2018.

Polag, D. *et al.* (2014) 'Age dependent breath methane in the German population', *Science of the Total Environment*, 481(1), pp. 582–587. doi:10.1016/j.scitotenv.2014.02.086.

Poulsen, M. *et al.* (2013) 'Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine rumen', *Nature Communications*, 4. doi:10.1038/ncomms2432.

Preston, C.M., Wu, K.Y., Molinski, T.F. and Delong, E.F. (1996) 'A psychrophilic crenarchaeon inhabits a marine sponge: Cenarchaeum symbiosum gen. nov., sp. nov.', *Proceedings of the National Academy of Sciences of the United States of America*, 93(13), pp. 6241–6246. doi:10.1073/pnas.93.13.6241.

Pryde, S.E. *et al.* (2002) *The microbiology of butyrate formation in the human colon*. doi:10.1111/j.1574-6968.2002.tb11467.x.

Qin, J. *et al.* (2010) 'A human gut microbial gene catalogue established by metagenomic sequencing', *Nature*, 464(7285), pp. 59–65. doi:10.1038/nature08821.

Quast, C. *et al.* (2013) 'The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools', *Nucleic Acids Research*, 41(D1), pp. 590–596. doi:10.1093/nar/gks1219.

Quigley, E.M.M. (2011) 'Microflora Modulation of Motility', *Journal of Neurogastroenterology and Motility*, 17(2), p. 140. doi:10.5056/JNM.2011.17.2.140.

Rahman, N.A. *et al.* (2015) 'A molecular survey of Australian and North American termite genera indicates that vertical inheritance is the primary force shaping termite gut microbiomes', *Microbiome 2015 3:1*, 3(1), pp. 1–16. doi:10.1186/S40168-015-0067-8.

Raymann, K., Brochier-Armanet, C. and Gribaldo, S. (2015) 'The two-domain tree of life is linked to a new root for the Archaea', *Proceedings of the National Academy of Sciences*, 112(21), pp. 6670–6675. doi:10.1073/PNAS.1420858112.

Raymann, K., Moeller, A.H., Goodman, A.L. and Ochman, H. (2017) 'Unexplored Archaeal Diversity in the Great Ape Gut Microbiome', *mSphere*, 2(1), pp. e00026-17. doi:10.1128/mSphere.00026-17.

Reese, A.T. *et al.* (2018) 'Microbial nitrogen limitation in the mammalian large intestine', *Nature Microbiology*, 3(12), pp. 1441–1450. doi:10.1038/s41564-018-0267-7.

Di Rienzi, S.C. *et al.* (2013) 'The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria', *eLife* [Preprint], (2). doi:10.7554/eLife.01102.001.

Rieu-Lesme, F., Delbès, C. and Sollelis, L. (2005) 'Recovery of Partial 16S rDNA Sequences Suggests the Presence of Crenarchaeota in the Human Digestive Ecosystem', *Current Microbiology*, 51(5), pp. 317–321. doi:10.1007/s00284-005-0036-8.

Robert, C. and Bernalier-Donadille, A. (2003) 'The cellulolytic microflora of the human colon: Evidence of microcrystalline cellulose-degrading bacteria in methane-excreting subjects', *FEMS Microbiology Ecology*, 46(1), pp. 81–89. doi:10.1016/S0168-6496(03)00207-1.

Rognes, T. *et al.* (2016) 'VSEARCH: A versatile open source tool for metagenomics', *PeerJ*, 2016(10), pp. 1–22. doi:10.7717/peerj.2584.

La Rosa, S.L. *et al.* (2019) 'The human gut Firmicute Roseburia intestinalis is a primary degrader of dietary β-mannans', *Nature Communications 2019 10:1*, 10(1), pp. 1–14. doi:10.1038/s41467-019-08812-y.

Rothschild, L. and Mancinelli, R. (2001) 'Life in extreme environments', *Nature*, 409(6823), pp. 1092–1101. doi:10.1038/35059215.

Ruaud, A. *et al.* (2020) 'Syntrophy via Interspecies H2 Transfer between Christensenella and Methanobrevibacter Underlies Their Global Cooccurrence in the Human Gut', *mBio*, 11(1). doi:10.1128/mBio.03235-19.

Russell, K., Lobley, G.E. and Millward, D.J. (2003) 'Whole-body protein turnover of a

carnivore, Felis silvestris catus', *British Journal of Nutrition*, 89(1), pp. 29–37. doi:10.1079/BJN2002735.

Saengkerdsub, S. and Ricke, S.C. (2014) 'Ecology and characteristics of methanogenic archaea in animals and humans', *Critical Reviews in Microbiology*, 40(2), pp. 97–116. doi:10.3109/1040841X.2013.763220.

Salgado-Flores, A. *et al.* (2019) 'Characterization of the cecum microbiome from wild and captive rock ptarmigans indigenous to Arctic Norway', *PLOS ONE*, 14(3), p. e0213503. doi:10.1371/JOURNAL.PONE.0213503.

Samuel, B.S. *et al.* (2007) 'Genomic and metabolic adaptations of Methanobrevibacter smithii to the human gut', *Proceedings of the National Academy of Sciences*, 104(25), pp. 10643–10648. doi:10.1073/pnas.0704189104.

Santoro, A.E. *et al.* (2015) 'Genomic and proteomic characterization of "Candidatus Nitrosopelagicus brevis": An ammonia-oxidizing archaeon from the open ocean', *PNAS*, 112(4), pp. 1173–1178. doi:10.1073/pnas.1416223112.

Satokari, R. *et al.* (2009) 'Bifidobacterium and Lactobacillus DNA in the human placenta', *Letters in Applied Microbiology*, 48(1), pp. 8–12. doi:10.1111/J.1472-765X.2008.02475.X.

Sauder, L.A. *et al.* (2017) 'Cultivation and characterization of Candidatus Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system', *ISME Journal*, 11(5), pp. 1142–1157. doi:10.1038/ismej.2016.192.

Schink, B. and Zeikus, J.G. (1980) 'Microbial methanol formation: a major end product of pectin metabolism', *Current Microbiology*, 4(6), pp. 387–389.

Schleper, C., Jurgens, G. and Jonuscheit, M. (2005) 'Genomic studies of uncultivated archaea', *Nature Reviews Microbiology 2005 3:6*, 3(6), pp. 479–488. doi:10.1038/nrmicro1159.

Schupp, E.W. (1993) 'Quantity, quality and the effectiveness of seed dispersal by animals', *Vegetatio 1993 107:1*, 107(1), pp. 15–29. doi:10.1007/BF00052209.

Scott, K.P. *et al.* (2006) 'Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium "Roseburia inulinivorans"', *Journal of Bacteriology*, 188(12), pp. 4340–4349. doi:10.1128/JB.00137-06.

Secor, S.M. and Diamond, J. (1995) 'Adaptive responses to feeding in Burmese pythons: pay before pumping.', *The Journal of experimental biology*, 198(6), pp. 1313–1325.

Sherrill-Mix, S. *et al.* (2018) 'Allometry and Ecology of the Bilaterian Gut Microbiome', *mBio*, 9(2), pp. e00319-18. doi:10.1128/mBio.00319-18.

Shi, Y. *et al.* (2015) 'Phylogenetic diversity of Archaea in the intestinal tract of termites from different lineages', *Journal of Basic Microbiology*, 55(8), pp. 1021–1028. doi:10.1002/jobm.201400678.

Shibata, M., Terada, F., M, S. and F, T. (2010) 'Factors affecting methane production and mitigation in ruminants', 81(1), pp. 2–10. doi:10.1111/J.1740-0929.2009.00687.X.

Shiffman, M.E. *et al.* (2017) 'Gene and genome-centric analyses of koala and wombat fecal microbiomes point to metabolic specialization for Eucalyptus digestion', *PeerJ*, 2017(11), pp. 1–32. doi:10.7717/peerj.4075.

Sieber, J.R., McInerney, M.J. and Gunsalus, R.P. (2012) 'Genomic Insights into Syntrophy: The Paradigm for Anaerobic Metabolic Cooperation', *Annual Review of Microbiology*, 66, pp. 429–452. doi:10.1146/ANNUREV-MICRO-090110-102844.

Silvester, K. and Cummings, J. (1995) 'Does digestibility of meat protein help explain large bowel cancer risk?', *Nutrition and cancer*, 24(3), pp. 279–288. doi:10.1080/01635589509514417.

Simister, R.L. *et al.* (2011) 'Sponge-specific clusters revisited: A comprehensive phylogeny of sponge-associated microorganisms', *Environmental Microbiology*, 14(2), pp. 517–524. doi:10.1111/j.1462-2920.2011.02664.x.

Skillman, L.C. *et al.* (2004) '16S ribosomal DNA-directed PCR primers for ruminal methanogens and identification of methanogens colonising young lambs', *Anaerobe*, 10(5), pp. 277–285. doi:10.1016/j.anaerobe.2004.05.003.

Smith, N.W. *et al.* (2019) 'Hydrogen cross-feeders of the human gastrointestinal tract', *Gut Microbes*. Taylor and Francis Inc., pp. 270–288. doi:10.1080/19490976.2018.1546522.

Smith, P.H. and Hungate, R.E. (1958) 'Isolation and characterization of Methanobacterium ruminantium', pp. 713–718.

Smits, S.A. *et al.* (2017) 'Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania', *Science*, 357(6353), pp. 802–806. doi:10.1126/SCIENCE.AAN4834.

Sogodogo, E., Drancourt, M. and Grine, G. (2019) 'Methanogens as emerging pathogens in anaerobic abscesses', *European Journal of Clinical Microbiology & Infectious Diseases 2019 38:5*, 38(5), pp. 811–818. doi:10.1007/S10096-019-03510-5.

Söllinger, A. *et al.* (2016) 'Phylogenetic and genomic analysis of Methanomassiliicoccales in wetlands and animal intestinal tracts reveals clade-specific habitat', *FEMS Microbiology Ecology*, 92(1), pp. 1–12. doi:10.1093/femsec/fiv149.

Song, J. *et al.* (2019) 'Is there convergence of gut microbes in blood-feeding vertebrates?' doi:10.1098/rstb.2018.0249.

Song, S.J. *et al.* (2020) 'Comparative Analyses of Vertebrate Gut Microbiomes Reveal Convergence between Birds and Bats', *mBio*, 11(1), pp. 1–14. doi:10.1128/mBio.02901-19.

Sonnenburg, J.L. and Bäckhed, F. (2016) 'Diet–microbiota interactions as moderators of human metabolism', *Nature 2016 535:7610*, 535(7610), pp. 56–64. doi:10.1038/nature18846.

Spang, A. *et al.* (2012) 'The genome of the ammonia-oxidizing Candidatus Nitrososphaera gargensis: insights into metabolic versatility and environmental adaptations', 14(12), pp. 3122–3145. doi:10.1111/J.1462-2920.2012.02893.X.

Sprenger, W. *et al.* (2000) 'Methanimicrococcus blatticola gen. nov., sp. nov., a methanol- and methylamine-reducing methanogen from the hindgut of the cockroach Periplaneta americana.', *International Journal of Systematic and Evolutionary Microbiology*, 50(6), pp. 1989–1999. doi:10.1099/00207713-50-6-1989.

Sprenger, W., Hackstein, J. and Keltjens, J. (2007) 'The competitive success of Methanomicrococcus blatticola, a dominant methylotrophic methanogen in the cockroach hindgut, is supported by high substrate affinities and favorable thermodynamics', *FEMS Microbiology Ecology*, 60(2), pp. 266–275. doi:10.1111/j.1574-6941.2007.00287.x.

Sprenger, W., Hackstein, J.H.P. and Keltjens, J.T. (2005) 'The energy metabolism of Methanomicrococcus blatticola: Physiological and biochemical aspects', *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 87(4), pp. 289–299. doi:10.1007/s10482-004-5941-5.

Spring, S. *et al.* (2016) 'Characterization of the first cultured representative of Verrucomicrobia subdivision 5 indicates the proposal of a novel phylum', *ISME Journal*, 10(12), pp. 2801–2816. doi:10.1038/ismej.2016.84.

St-Pierre, B. and Wright, A.D.G. (2013) 'Diversity of gut methanogens in herbivorous animals', *Animal*, 7(SUPPL.1), pp. 49–56. doi:10.1017/S1751731112000912.

Steggerda, F.R. (1968) 'Gastrointestinal gas following food consumption', *Annals of the New York Academy of Sciences*, 150(1), pp. 57–66.

Steiner, K. *et al.* (2021) 'The microbiome of Chinook salmon (Oncorhynchus tshawytscha) in a recirculation aquaculture system', *Aquaculture*, 534, p. 736227. doi:10.1016/J.AQUACULTURE.2020.736227.

Steuer, P. *et al.* (2011) 'Is there an influence of body mass on digesta mean retention time in herbivores? A comparative study on ungulates', *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 160(3), pp. 355–364. doi:10.1016/j.cbpa.2011.07.005.

Stevens, C.E. and Hume, I.D. (1998) 'Contributions of microbes in vertebrate gastrointestinal tract to production and conservation of nutrients', *Physiological Reviews*, 78(2), pp. 393–427. doi:10.1152/physrev.1998.78.2.393.

Sugimoto, A. *et al.* (1998) 'Methane and hydrogen production in a termite-symbiont system', *Ecological Research 1998 13:2*, 13(2), pp. 241–257. doi:10.1046/J.1440-1703.1998.00262.X.

Tang, W.H.H.W.W. *et al.* (2014) 'Gut microbiota-dependent trimethylamine N-oxide (TMAO) pathway contributes to both development of renal insufficiency and mortality risk in chronic kidney disease', *Circulation Research*, 116(3), pp. 448–455. doi:10.1161/CIRCRESAHA.116.305360.

Tap, J. *et al.* (2015) 'Gut microbiota richness promotes its stability upon increased dietary fibre intake in healthy adults', *Environmental microbiology*, 17(12), pp. 4954–4964. doi:10.1111/1462-2920.13006.

Tapio, I., Snelling, T.J., Strozzi, F. and Wallace, R.J. (2017) 'The ruminal microbiome associated with methane emissions from ruminant livestock', *Journal of Animal Science and Biotechnology*, 8(1), pp. 1–11. doi:10.1186/s40104-017-0141-0.

Tennekes, M. (2017) 'treemap: Treemap Visualization'. Available at: https://cran.r-project.org/package=treemap.

Tholen, A., Pester, M. and Brune, A. (2007) 'Simultaneous methanogenesis and oxygen reduction by Methanobrevibacter cuticularis at low oxygen fluxes', *FEMS Microbiology Ecology*, 62(3), pp. 303–312. doi:10.1111/j.1574-6941.2007.00390.x.

Thomas, C.M., Taib, N., Gribaldo, S. and Borrel, G. (2021) 'Comparative genomic analysis of Methanimicrococcus blatticola provides insights into host adaptation in archaea and

the evolution of methanogenesis', *ISME Communications 2021 1:1*, 1(1), pp. 1–11. doi:10.1038/s43705-021-00050-y.

Thomas, T. *et al.* (2016) 'Diversity, structure and convergent evolution of the global sponge microbiome', *Nature Communications*, 7(May). doi:10.1038/ncomms11870.

Togo, A.H. *et al.* (2019) 'Culture of Methanogenic Archaea from Human Colostrum and Milk', *Scientific Reports 2019 9:1*, 9(1), pp. 1–10. doi:10.1038/s41598-019-54759-x.

Tottey, W. *et al.* (2017) 'Colonic Transit Time Is a Driven Force of the Gut Microbiota Composition and Metabolism: In Vitro Evidence', *Journal of Neurogastroenterology and Motility*, 23(1), p. 124. doi:10.5056/JNM16042.

Traore, S.I. *et al.* (2019) 'Isolation and culture of Methanobrevibacter smithii by co-culture with hydrogen-producing bacteria on agar plates', *Clinical Microbiology and Infection*, 25(12), pp. 1561.e1-1561.e5. doi:10.1016/j.cmi.2019.04.008.

Triantafyllou, K., Chang, C. and Pimentel, M. (2014) 'Methanogens, methane and gastrointestinal motility', 20(1), p. 31. doi:10.5056/JNM.2014.20.1.31.

Tripathi, B.M. *et al.* (2015) 'Soil pH and biome are both key determinants of soil archaeal community structure', *Soil Biology and Biochemistry*, 88, pp. 1–8.

Turnbaugh, P.J. *et al.* (2006) 'An obesity-associated gut microbiome with increased capacity for energy harvest', *Nature*, 444(7122), pp. 1027–1031. doi:10.1038/nature05414.

Tuson, H.H. and Weibel, D.B. (2013) 'Bacteria-surface interactions', *Soft matter*, 9(18), p. 4368. doi:10.1039/C3SM27705D.

Ueno, Y. *et al.* (2006) 'Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era', *Nature*, 440(7083), pp. 516–519. doi:10.1038/nature04584.

Ungerfeld, E.M. (2020) 'Metabolic Hydrogen Flows in Rumen Fermentation: Principles and Possibilities of Interventions', *Frontiers in Microbiology*, 11(April). doi:10.3389/fmicb.2020.00589.

Valentine, D.L. (2007) 'Adaptations ot energy stress dictate the ecology and evolution of the Archaea', *Nature Reviews*, 5, pp. 316–323. Available at: www.nature.com/reviews/micro (Accessed: 4 October 2021).

Vanderhaeghen, S., Lacroix, C. and Schwab, C. (2015) 'Methanogen communities in stools of humans of different age and health status and co-occurrence with bacteria', *FEMS Microbiology Letters*, 362(13), pp. 1–8. doi:10.1093/femsle/fnv092.

Varki, A. (2017) 'Biological roles of glycans', *Glycobiology*, 27(1), pp. 3–49.

doi:10.1093/GLYCOB/CWW086.

Vester, B.M. *et al.* (2008) 'Nutrient digestibility and fecal characteristics are different among captive exotic felids fed a beef-based raw diet', *Zoo Biology: Published in affiliation with the American Zoo and Aquarium Association*, 27(2), pp. 126–136.

Vital, M. *et al.* (2015) 'Diet is a major factor governing the fecal butyrate-producing community structure across Mammalia, Aves and Reptilia', *ISME Journal*, 9, pp. 832–843. doi:10.1038/ismej.2014.179.

Volant, S. *et al.* (2020) 'SHAMAN: a user-friendly website for metataxonomic analysis from raw reads to statistical analysis', *BMC Bioinformatics 2020 21:1*, 21(1), pp. 1–15. doi:10.1186/S12859-020-03666-4.

Voragen, A.G.J.J., Coenen, G.-J.J., Verhoef, R.P.R.P. and Schols, H.A. (2009) 'Pectin, a versatile polysaccharide present in plant cell walls', *Structural Chemistry*, 20(2), pp. 263–275. doi:10.1007/s11224-009-9442-z.

Wang, L. *et al.* (2012) 'Elevated fecal short chain fatty acid and ammonia concentrations in children with autism spectrum disorder', *Digestive diseases and sciences*, 57(8), pp. 2096–2102.

Wang, W. *et al.* (2016) 'High-throughput sequencing reveals the core gut microbiome of Bar-headed goose (Anser indicus) in different wintering areas in Tibet', *MicrobiologyOpen*, 5(2), pp. 287–295. doi:10.1002/mbo3.327.

Wang, Y. *et al.* (2021) 'A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism', *Science Advances*, 7(27), p. eabj1453. doi:10.1126/SCIADV.ABJ1453.

Wang, Z. *et al.* (2011) 'Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease.', *Nature*, 472(7341), pp. 57–63. doi:10.1038/nature09922.

Wasburn, L.E. and Brody, S. (1937) 'Methane, hydrogen, and carbon doxine production in the digestive tract of ruminants in relation to the respiratory exchange', *Grown and development with Special Reference to Domestic Animals* [Preprint].

Wasimuddin *et al.* (2017) 'Gut microbiomes of free-ranging and captive Namibian cheetahs: Diversity, putative functions and occurrence of potential pathogens', *Molecular Ecology*, 26(20), pp. 5515–5527. doi:10.1111/MEC.14278.

Wastyk, H.C. *et al.* (2021) 'Gut-microbiota-targeted diets modulate human immune status', *Cell*, 184(16), pp. 4137-4153.e14. doi:10.1016/J.CELL.2021.06.019.

Wedlock, D. *et al.* (2013) 'Progress in the development of vaccines against rumen methanogens', *Animal : an international journal of animal bioscience*, 7 Suppl 2, pp. 244–252. doi:10.1017/S1751731113000682.

Wienemann, T. *et al.* (2011) 'The bacterial microbiota in the ceca of Capercaillie (Tetrao urogallus) differs between wild and captive birds', *Systematic and Applied Microbiology*, 34(7), pp. 542–551. doi:10.1016/J.SYAPM.2011.06.003.

Williams, A.G., Withers, S.E. and Joblin, K.N. (1994) 'The effect of cocultivation with hydrogen-consuming bacteria on xylanolysis by Ruminococcus flavefaciens', *Current Microbiology*, 29(3), pp. 133–138.

Wilman, H. *et al.* (2014) 'EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals: Ecological Archives E095-178', *Ecology*, 95(7), p. 2027.

Woese, C.R. and Fox, G.E. (1977) 'Phylogenetic structure of the prokaryotic domain: The primary kingdoms', *Proceedings of the National Academy of Sciences*, 74(11), pp. 5088–5090. doi:10.1073/pnas.74.11.5088.

Worm, P. *et al.* (2010) 'Syntrophy in Methanogenic Degradation', in *(Endo)symbiotic Methanogenic Archaea*. Springer, Berlin, Heidelberg, pp. 143–173. doi:10.1007/978-3-642-13615-3_9.

Wright, A.D.G., Northwood, K.S. and Obispo, N.E. (2009) 'Rumen-like methanogens identified from the crop of the folivorous South American bird, the hoatzin (Opisthocomus hoazin)', *ISME Journal*, 3(10), pp. 1120–1126. doi:10.1038/ismej.2009.41.

Wu, Z. *et al.* (2019) 'High-throughput sequencing reveals the gut and lung prokaryotic community profiles of the Chinese giant salamander (Andrias davidianus)', *Molecular Biology Reports*, 46(5), pp. 5143–5154. doi:10.1007/s11033-019-04972-8.

Xenoulis, P.G. *et al.* (2010) 'Molecular characterization of the cloacal microbiota of wild and captive parrots', *Veterinary microbiology*, 146(3–4), pp. 320–325. doi:10.1016/J.VETMIC.2010.05.024.

Xu, L. *et al.* (2021) 'Methyl-compounds driven benthic carbon cycling in the sulfate-reducing sediments of South China Sea', *Environmental Microbiology*, 23(2), pp. 641–651.

Yamamoto, N. *et al.* (2012) 'Ammonia-Oxidizing Bacteria Rather than Ammonia-Oxidizing Archaea were Widely Distributed in Animal Manure Composts from Field-Scale Facilities', *Microbes and Environments*, 27(4), p. 519. doi:10.1264/JSME2.ME12053.

Yang, J. *et al.* (2021) 'Gut region induces gastrointestinal microbiota community shift in Ujimqin sheep (Ovis aries): from a multi-domain perspective', *Environmental Microbiology* [Preprint]. doi:10.1111/1462-2920.15782.

Yilmaz, P. *et al.* (2014) 'The SILVA and "all-species Living Tree Project (LTP)" taxonomic frameworks', *Nucleic Acids Research*, 42(D1), pp. 643–648.

Youngblut, N.D. *et al.* (2019) 'Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades', *Nature Communications*, 10(1), pp. 1–15. doi:10.1038/s41467-019-10191-3.

Youngblut, N.D. *et al.* (2020) 'Strong influence of vertebrate host phylogeny on gut archaeal diversity', *bioRxiv*, 5000, p. 147. doi:10.1101/2020.11.10.376293.

Zeisel, S.H., Mar, M.-H., Howe, J.C. and Holden, J.M. (2003) 'Concentrations of Choline-Containing Compounds and Betaine in Common Foods', *The Journal of Nutrition*, 133(5), pp. 1302–1307. doi:10.1093/JN/133.5.1302.

Zhang, A.Q., Mitchell, S.C. and Smith, R.L. (1999) 'Dietary Precursors of Trimethylamine in Man: A Pilot Study', *Food and Chemical Toxicology*, 37(5), pp. 515–520. doi:10.1016/S0278-6915(99)00028-9.

Zhang, L. *et al.* (2017) 'Characteristics of rumen microorganisms involved in anaerobic degradation of cellulose at various pH values', *RSC Advances*, 7(64), pp. 40303–40310. doi:10.1039/c7ra06588d.

Zhang, S. *et al.* (2019) 'Comparative Genomics Reveals Ecological and Evolutionary Insights into Sponge-Associated Thaumarchaeota', *mSystems*, 4(4), pp. 1–16. doi:10.1128/msystems.00288-19.

Zheng, L., Kelly, C. and Colgan, S. (2015) 'Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A Review in the Theme: Cellular Responses to Hypoxia', *American journal of physiology. Cell physiology*, 309(6), pp. C350–C360. doi:10.1152/AJPCELL.00191.2015.

Zhou, M., Chen, Y., Griebel, P.J. and Guan, L.L. (2014) 'Methanogen prevalence throughout the gastrointestinal tract of pre-weaned dairy calves', *Gut Microbes*, 5(5), p. 628. doi:10.4161/19490976.2014.969649.

Zhuang, G.-C. *et al.* (2016) 'Multiple evidence for methylotrophic methanogenesis as the dominant methanogenic pathway in hypersaline sediments from the Orca Basin, Gulf of Mexico', *Geochimica et Cosmochimica Acta*, 187, pp. 1–20.

Zhuang, G. *et al.* (2019) 'Biogeochemistry, microbial activity, and diversity in surface and subsurface deep-sea sediments of South China Sea', *Limnology and Oceanography*, 64(5), pp. 2252–2270.

De Zwart, F.J. *et al.* (2003) 'Glycine betaine and glycine betaine analogues in common foods', *Food chemistry*, 83(2), pp. 197–204.