



First Insights into the Diverse Human Archaeome: Specific Detection of Archaea in the Gastrointestinal Tract, Lung, and Nose and on Skin

Kaisa Koskinen,^{a,b} Manuela R. Pausan,^a Alexandra K. Perras,^{a,c} Michael Beck,^{a*} Corinna Bang,^{d*} Maximilian Mora,^a Anke Schilhabel,^{d*} Ruth Schmitz,^d Christine Moissl-Eichinger^{a,b}

Medical University of Graz, Graz, Austria^a; BioTechMed-Graz, Graz, Austria^b; University of Regensburg, Regensburg, Germany^c; Christian-Albrechts-Universität zu Kiel, Kiel, Germany^d

ABSTRACT Human-associated archaea remain understudied in the field of microbiome research, although in particular methanogenic archaea were found to be regular commensals of the human gut, where they represent keystone species in metabolic processes. Knowledge on the abundance and diversity of human-associated archaea is extremely limited, and little is known about their function(s), their overall role in human health, or their association with parts of the human body other than the gastrointestinal tract and oral cavity. Currently, methodological issues impede the full assessment of the human archaeome, as bacteria-targeting protocols are unsuitable for characterization of the full spectrum of *Archaea*. The goal of this study was to establish conservative protocols based on specifically archaea-targeting, PCR-based methods to retrieve first insights into the archaeomes of the human gastrointestinal tract, lung, nose, and skin. Detection of *Archaea* was highly dependent on primer selection and the sequence processing pipeline used. Our results enabled us to retrieve a novel picture of the human archaeome, as we found for the first time *Methanobacterium* and *Woesearchaeota* (DPANN superphylum) to be associated with the human gastrointestinal tract and the human lung, respectively. Similar to bacteria, human-associated archaeal communities were found to group biogeographically, forming (i) the thaumarchaeal skin landscape, (ii) the (methano)eurarchaeal gastrointestinal tract, (iii) a mixed skin-gastrointestinal tract landscape for the nose, and (iv) a woesearchaeal lung landscape. On the basis of the protocols we used, we were able to detect unexpectedly high diversity of archaea associated with different body parts.

IMPORTANCE In summary, our study highlights the importance of the primers and NGS data processing pipeline used to study the human archaeome. We were able to establish protocols that revealed the presence of previously undetected *Archaea* in all of the tissue samples investigated and to detect biogeographic patterns of the human archaeome in the gastrointestinal tract, on the skin, and for the first time in the respiratory tract, i.e., the nose and lungs. Our results are a solid basis for further investigation of the human archaeome and, in the long term, discovery of the potential role of archaea in human health and disease.

KEYWORDS archaeome, methanogens, microbiome, Archaea

Our invisible organ, the human microbiome, is composed of *Bacteria*, eukaryotes, *Archaea*, and viruses. As *Bacteria* are considered to be the predominant component of the human microbiome, the overwhelming majority of studies focus on their contribution (1). Recently, research has been directed to the eukaryome, mycobiome

Received 16 May 2017 Accepted 27 September 2017 Published 14 November 2017

Citation Koskinen K, Pausan MR, Perras AK, Beck M, Bang C, Mora M, Schilhabel A, Schmitz R, Moissl-Eichinger C. 2017. First insights into the diverse human archaeome: specific detection of archaea in the gastrointestinal tract, lung, and nose and on skin. *mBio* 8:e00824-17. <https://doi.org/10.1128/mBio.00824-17>.

Invited Editor Christa M. Schleper, University of Vienna

Editor Janet K. Jansson, Pacific Northwest National Laboratory

Copyright © 2017 Koskinen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Christine Moissl-Eichinger, christine.moissl-eichinger@medunigraz.at.

* Present address: Michael Beck, Roche Diagnostics GmbH, Penzberg, Germany; Corinna Bang, University Hospital Schleswig-Holstein, Institute of Clinical Molecular Biology, Kiel, Germany; Anke Schilhabel, University Hospital Schleswig-Holstein, Department of Hematology, Kiel, Germany.

K.K., M.R.P., and A.K.P. contributed equally to this work.

(2–4), and virome (5). However, the human archaeome remains largely uninvestigated, although *Archaea* are known to strongly impact human health and well-being (6, 7). In particular, methanoarchaea (methane-producing archaea, methanogens) are believed to be keystone species that may have a greater influence on the composition and function of the whole gastrointestinal microbial community than we currently know (8–10).

To date, four species of methanogenic archaea have been cultivated and isolated from human body samples: the first isolated and described human-associated archaeon *Methanobrevibacter smithii* (11), *Methanosphaera stadtmanae* (12), and most recently *Methanomassiliicoccus luminyensis* (13) were all isolated from human feces. A fourth isolate, *Methanobrevibacter oralis* (14) was cultivated and described from oral mucosa. In recent years, with the help of molecular methods, several other archaeal groups have been found in the human body. For example, two candidate species, “*Candidatus* *Methanomassiliicoccus intestinalis*” and “*Candidatus* *Methanomethylophilus alvus*,” as well as several unknown members of the orders *Methanosarcinales*, *Methanobacteriales*, *Methanococcales*, *Methanomicrobiales*, and *Methanopyrales*, have been found to inhabit the human gut (15). It is estimated that up to 96% of all people carry *M. smithii* in their guts (16, 17). The abundance of this archaeon in the human gut differs in individual human subjects, but it can account up to 10% of the anaerobic microbes in the colon (18). In addition, *M. stadtmanae* was reported to be highly abundant and was detected in about 30% of the subjects tested (16). Together with the recently discovered species *M. luminyensis*, with a prevalence of 4% in the population studied (17), *M. smithii* and *M. stadtmanae* are currently known to predominantly inhabit the human gut.

In natural environments, most methane-producing archaea use C₁ and C₂ carbon sources and hydrogen for methane production in the last step of the anaerobic food chain. The constant removal of molecular hydrogen drives preceding bacterial metabolic processes (primary and secondary fermenters), as higher hydrogen partial pressure turns these activities thermodynamically unfavorable (19). On the other hand, methanogenic archaea have been reported to be associated with different disease patterns, such as the development of inflammatory bowel disease (20), dental disease (21, 22), or potentially even brain abscesses (23).

However, not all human gut-associated archaea are methanogens; unknown members of the orders *Desulfurococcales*, *Sulfolobales*, *Thermoproteales*, *Nitrososphaerales*, and *Halobacteriales* have also been detected in the human intestine (15). Recently, a new species of halophilic archaea, *Haloferax massiliensis* (24), has been isolated from the human gut, again raising the question of whether halophilic archaea could potentially be permanent residents (25). Besides the gut, *Archaea* have also been reported to be widely distributed in other human body sites, such as the vaginal cavity (*M. smithii*) and skin (*Thaumarchaeota*) (26–28). Although their role is largely unknown, skin-associated *Thaumarchaeota* can make up as much as 10% of the microbiome, in particular in skin samples from elderly persons and children (29).

Research on microbiomes and on the human microbiome in particular was pushed forward years ago by the application of next-generation sequencing (NGS) in amplicon-based studies. Since then, the search for optimal methods in preparatory laboratory work, subsequent primer selection, and PCR amplification has begun. The ability to create more data in less time also increased the need to optimize data processing workflows and displaying methods after sequencing. Although the results of human microbiome research are exciting, a certain incomparability between microbiome studies that have used various methods on different levels has recently been observed. The influence of the chosen methodology was recognized as the major driving force behind the dissimilarity of results, as shown in one example of indoor microbiome studies (30). On the basis of these observations, standardized procedures have been proposed by world-leading projects such as the Human Microbiome Project (HMP) (1) and the Earth Microbiome Project (EMP) (31) to simplify comparisons between different studies. In the HMP, the focus is on the bacteriome (the diversity of bacteria), with

standard operating procedures (SOPs) proposing the use of so-called “universal,” but mainly bacteria-targeting, primers for the 16S rRNA gene amplicon sequencing approach. Consequently, the diversity and abundance of archaea are mostly inferred from “side products” of studies targeting bacteria (27, 32). As a consequence, archaeal 16S rRNA gene information is frequently filtered from the data set during sequence processing (33). Such problems are also reflected in metagenomic analyses, in which assignable archaeal reads typically represent a small minority as well (34). The specific detection of archaeal signatures suffers from methodological problems on various levels, including DNA extraction, selection of suitable primers, a low abundance of archaeal DNA, or incompleteness of 16S rRNA gene reference databases (for more detailed information on these issues, see Text S1 in the supplemental material; for a summary of the targeted archaeal groups and genes, sample types, molecular methods, and primers that have been used to detect *Archaea* in human-associated samples, see Table S1).

Besides those methodological problems, archaeome research faces another challenge; with no known archaeal pathogen, these species are mostly ignored and the need for a detailed analysis is not recognized within the big and well-funded field of clinical microbiology (35, 36).

In this study, we are taking the first step in a long row of necessary improvements for the specific detection of archaea in large studies. Here, we focused on optimal primer selection and analyzed the ability of available NGS sequence processing pipelines, namely mothur (37), QIIME (Quantitative Insights Into Microbial Ecology) (38), and DADA2 (Divisive Amplicon Denoising Algorithm) (39), to expand the picture of the human archaeome.

More specifically, we present two main approaches. We first used a PCR-based detection method to determine the presence or absence of the well-known human-associated methanoarchaea in a variety of biopsy samples from the gut, providing a sound basis for determining the involvement of these archaea in inflammatory processes in the human gut.

The second approach addresses the need for archaea-specific NGS-based diversity analyses. We critically assessed the capacity of archaeal and “universal” primers to detect archaea-specific signals via NGS in different human tissue samples and processed data with different NGS pipelines. As a proof of principle, we applied our most specific and conservative protocol to a variety of human tissue samples, including skin samples, lung samples, biopsy specimens from the gastrointestinal tract (GIT), stool samples, and nasal samples, uncovering a previously unknown diversity of archaea associated with specific sites of the human body.

RESULTS AND DISCUSSION

Human-associated archaea remain, similar to fungi or viruses, understudied in the field of microbiome research. Numerous studies have indicated the potential of human-associated archaea (in particular, methanogens) to represent keystone species, as they can be major drivers of metabolic processes in the gut. Recent studies have shown that the human archaeome still holds many surprises, including the discovery of the seventh order of methanogens in the human gut (13), the cultivation of a halophilic archaeon from human samples (24), and the detection of *Thaumarchaeota* on human skin (27). Unfortunately, studies on the human archaeome are rare, mainly due to methodological issues and a lack of motivation from a medical point of view because no firm evidence of an archaeal pathogen has been discovered. In this study, we focused on primer choice and sequence processing protocols to optimize specific detection of (methano)archaea in human samples. The most conservative and most specific protocol was then used to determine the diversity of archaea associated with human skin, lung, GIT, stool, and nasal samples.

A PCR-based methanoarchaea-targeting approach allows the specific detection of human-associated methanogens (*M. smithii*, *M. stadtmanae*, and *M. luminyensis*) in healthy and inflamed gut biopsy samples (approach 1). In the first approach tested in

TABLE 1 Results obtained for the species-specific detection of *Methanoarchaea* in gut biopsy samples of healthy controls and UC and CD patients^a

Target or parameter	Healthy controls					Inflamed UC					Uninflamed UC	
	1 ^b	2	3	4	5	1	2	3	4	5	1	2
<i>M. smithii</i>	-/+/+	-	+/+/+	-	-	-/+/-	+/+/+	+/+/-	+/-/-	+/+/-	-/+/-	+/-/-
<i>M. stadtmanae</i>	-/+/-	-/-	-/+/-	-/-	-/+	-/-/-	-/-/+	-/-/-	-/-/-	-/-/-	+/-/-	-/+/+
<i>M. luminyensis</i>	-/-/-	-/-/-	+/-/-	-/-/-	+/-/-	-/-/-	-/-/-	-/-/-	-/-/-	-/-/-	-/-/-	-/-/-
Archaea ^c	-/-/-	-/+/-	+/+/+	-/-/-	+/-/+	+/+/+	-/+/-	-/-/-	-/-/+	-/-/-	-/-/-	-/+/+
Identity ^d		<i>M. smithii</i>	<i>M. smithii</i> , <i>M. stadtmanae</i>		Euryarchaeote SCGC AAA003-G23	<i>M. smithii</i>	<i>M. smithii</i> , <i>M. stadtmanae</i>		<i>M. stadtmanae</i>			<i>M. smithii</i>

^aNested PCR of the *rpoB* gene (and of the 16S rRNA gene as a control) was performed as described in Materials and Methods. Positive (+) and negative (-) results out of three replicates are shown (each symbol reflects one replicate). Where only one or two symbols are shown, three replicates could not be analyzed because of a low DNA yield.

^bBiopsy sample number.

^c16S rRNA gene.

^dObtained by Sanger sequencing.

TABLE 1 (Continued)

Uninflamed UC			Inflamed CD					Uninflamed CD				
3	4	5	1	2	3	4	5	1	2	3	4	5
-/+/-	-/-/-	-/+/-	+/+/-	+/-/+	+/+/-	+/-/-	-/+/+	+/-/-	+/+/-	-/+/-	-/+/-	-/+/-
-/-/-	-/-/-	-/-/-	-/-/-	+/-/-	-/-/-	-/-/-	-/+/-	-/-/-	-/+/-	-/-/-	-/-/-	-/-/-
-/-/-	-/-/-	-/-/-	-/-/-	+/+/-	+/-/-	+/-/-	-/+/-	-/-/-	-/-/-	-/-/-	-/-/-	-/-/-
-/+/+	-/-/-	-/+/+	-/+/-	-/+/-	-/+/-	-/-/-	-/+/+	-/+/+	+/+/+	+/+/-	-/+/-	-/-/-
<i>M. smithii</i> , <i>M. stadtmanae</i>		<i>M. smithii</i>	<i>M. smithii</i>	<i>M. smithii</i> , <i>M. stadtmanae</i>	<i>M. smithii</i>		<i>M. stadtmanae</i>	<i>M. smithii</i> , <i>M. stadtmanae</i>	<i>M. smithii</i>	<i>M. stadtmanae</i>	<i>M. smithii</i>	

this study, we concentrated on facilitated, NGS-independent detection of the three most abundant (identified) methanoarchaea of the human gut microbiome. With this setup, we addressed the question of whether there is a pattern of co-occurrence of various mucosa-associated methanoarchaeal strains with the manifestation of inflammatory processes involving the human GIT.

Biopsy samples retrieved from healthy controls and diseased persons (Crohn’s disease [CD], ulcerative colitis [UC]) at both inflamed and uninflamed locations were analyzed for the presence of *M. smithii*, *M. stadtmanae*, and *M. luminyensis*. Primers were selected and designed on the basis of a previously described quantitative reverse transcription-PCR setup for quantifying *M. smithii* and *M. stadtmanae* *rpoB* genes (16). Primers for the detection of *M. luminyensis* targeted the 16S rRNA gene and were specifically designed in this study (see Materials and Methods). In all cases, the evaluation demonstrated that a nested PCR was required, since a single PCR with strain-specific primers never provided immediate positive results for *M. luminyensis* and *M. stadtmanae* and only in rare cases for *M. smithii*. The results of the second PCR are shown in Table 1. All nested PCRs were performed in triplicates if not mentioned otherwise.

Overall, signatures of *M. smithii* were found to be the most abundant, as 10 of 25 samples (at least two positive results out of three replicates) or 21 of 25 samples (at least one positive result) revealed positive amplification. *M. stadtmanae* was found in 1 of 25 samples (at least two positive results out of three replicates) or in 8 of 25 samples (at least one positive result). These findings were in congruence with the archaea-targeting results (see below). *M. luminyensis* signatures were detected in 1 of 25 samples (at least two positive results out of three replicates) or 7 of 25 samples (at least one positive result out of three replicates; Table 1). For verification, all samples were also subjected to 16S rRNA gene amplification (also by a nested PCR approach) targeting archaea in general (Table 1; Text S1; Fig. S1). With this experimental setup, the presence of archaeal signatures in 9 of 25 samples (at least two positive results out of three replicates) or 17 of 25 samples (at least one positive result out of three replicates) could be shown. Sanger sequencing of these PCR products identified mainly gene signatures of *M. smithii* (12 samples), as well as *M. stadtmanae* (8 samples), whereas

signatures of *M. luminyensis* were not confirmed. Although the results basically overlapped, several samples that were negative in the archaea-targeting PCR were positive when a species-specific nested PCR assay was performed (Table 1).

In general, the nested PCR approach proposed was found to be suitable for the detection of specific strains in human biopsy specimens. An advantage of the strain-specific detection method described is proof of the presence of known strains despite their comparatively low abundance. However, the overall results of this method were not always clear, probably because of the comparatively small proportion of archaeal DNA (40, 41), and no obvious pattern was obtained regarding the presence of methanoarchaeal strains in patients with inflamed or uninflamed UC or CD. Information needs to be obtained from quantitative studies and larger cohorts to resolve this question. One reason for the inconsistency of PCR-based results might be the presence of a large amount of eukaryotic DNA in the purified DNA fractions. Thus, determination of archaeal strains in samples containing large amounts of eukaryotic DNA, such as biopsy specimens, might be made more consistent by using, e.g., NGS-based analyses. This will be addressed later. It is noteworthy that the proven presence of methanoarchaea in biopsy samples suggests that these microorganisms most likely are associated with the mucosa and are not only present in the lumen of the GIT. Previous studies have hypothesized the attachment of these strains to the intestinal epithelia because of their ability to form biofilms on nonliving surfaces (42). This hypothesis is supported by genomic studies of *M. smithii* and *M. stadtmanae* that showed that both species express adhesion-enabling proteins similar to bacterial adhesins (18).

Primer pairs and sequence data processing pipelines determine the detection of archaeal operational taxonomic units (OTUs) and ribosomal sequence variants (RSVs) in natural mock communities (approach 2). In a preliminary step, we tested the efficiency of three different primer pairs suitable for NGS (short read length), namely, 515f/806r targeting *Bacteria* and *Archaea* (43), 349af/519ar (344f/915r) targeting *Archaea*, and 519af/785ur (344f/915r) targeting *Bacteria* and *Archaea* (44), with respect to their (specific) ability to detect the diverse archaea in human stool samples, serving as natural mock communities. We used two human stool samples from healthy persons, as this type of sample is often used for microbiome analyses and as archaea have been successfully detected in human stool samples in the past (16).

The first primer pair we tested was the commonly used “universal” primer pair 515f/806r, targeting both *Bacteria* and *Archaea* (43) (proposed by the HMP and the EMP). The other two primer pairs were used in a nested PCR strategy with an archaea-specific first PCR. For the first PCR, we used the primer pair 344f/915r (27, 45, 46). The nested PCR method was used to (i) preselect archaeal fragments, (ii) increase the amount of archaeal signatures, and (iii) avoid the formation of primer dimers, which were observed when primers tagged with Illumina adapter and barcode were used without previous amplification.

After the initial PCR with the well-established primer pair 344f/915r, which is very selective for the archaeal domain, either primer combination 349af/519ar (coverage without a mismatch, 78.7% of *Archaea* and 0% of *Bacteria* on the basis of *in silico* analysis), targeting *Archaea*, or primer combination 519af/785ur (coverage without a mismatch, 88.7% of *Archaea* and 88.5% of *Bacteria*), targeting both *Bacteria* and *Archaea* (44), were used in a subsequent PCR approach (see Materials and Methods for details). All three protocols allowed the successful amplification of 16S rRNA gene fragments, which were sequenced by Illumina MiSeq. The raw sequence data set retrieved was then subjected to three different analysis pipelines based on QIIME, mothur, and DADA2 (37–39) to assess the impact of the data processing pipeline used on the results. All three pipelines were used in accordance with the SOPs proposed by their developers.

The different primer combinations revealed diverse levels of specificity with respect to *Archaea* detection (Table 2) (detailed results [OTU and RSV tables] are available upon request). The primer pair 349af/519ar (following the first PCR with 344f/915r) allowed the retrieval of solely archaeal sequences and was, as expected, found to be the most

TABLE 2 Summary of the bioinformatics pipelines and primer combinations used and archaeal and bacterial OTUs/RSVs detected in the stool samples studied^a

Primers	QIIME						Mothur						DADA2					
	Archaeal OTUs		Bacterial OTUs		% of archaeal OTUs		Archaeal OTUs		Bacterial OTUs		% of archaeal OTUs		Archaeal OTUs		Bacterial OTUs		% of archaeal OTUs	
	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
349af/519ar (344f/915r)	484 ± 80 ^b	622 ± 27	0 ± 0.5	0 ± 0	100	0 ± 0	320 ± 70	411 ± 31	0 ± 0	0 ± 0	100	0 ± 0	8 ± 2	15 ± 5	0 ± 0	0 ± 0	0 ± 0	100
519af/785ur (344f/915r)	36 ± 2.5	37 ± 2	216 ± 90	84 ± 44	16–44	8,990 ± 2,030	7,869 ± 1,182	146 ± 45	83 ± 40	98–99	63 ± 22	52 ± 13	69 ± 27	21 ± 19	48–71			
515f/806r	0 ± 0	1 ± 0.5	4,873 ± 1,146	6,618 ± 874	0–0.02	1 ± 0	1 ± 0	20,648 ± 6,208	25,063 ± 7,033	0.004–0.005	0 ± 1	1 ± 1	166 ± 24	274 ± 43	0–0.36			

^aThe primer pair in parentheses (344f/915r) was used in the first round of a nested PCR. The samples were amplified in three replicates. The same sequence data sets were used for the different pipelines.

^bMedian ± interquartile range.

specific approach. With all of the data processing pipelines used, the percentage of archaeal reads was 100%. However, the number of archaeal OTUs/RSVs observed was significantly less than that obtained with the 519af/785ur primer combination when the sequence data were processed with mothur (t test, $P = 0.00007$) or DADA2 (t test, $P = 0.002$). Conversely, a different picture was obtained from the QIIME processing pipeline, as the same data sets resulted in a higher number of archaeal OTUs from the more specific primer pair 349af/519ar (344f/915r) than from the combination 519af/785ur (344f/915r) (t test, $P = 0.00004$).

Primer pair 519af/785ur (344f/915r) was capable of retrieving the broadest diversity of archaea, but because of the different primer specificities, a vast variety of bacterial signatures were also obtained. The percentage of OTUs classified as archaea and the number of archaeal OTUs/RSVs detected depended largely on the data processing pipeline used (one-way analysis of variance [ANOVA] $P = 1.09E-10$) (Table 2). Additionally, our results show that the universal and widely used primer pair 515f/806r failed to detect the diverse archaea in stool samples. By this approach, all of the data processing pipelines tested were able to capture only one archaeal OTU per sample, corresponding to 0 to 0.36% of the OTUs and RSVs retrieved (Table 2, last column).

On the basis of our analyses, primer pair 349af/519ar proved superior to the other primer pairs tested with respect to archaeal specificity and thus was used later in this study to analyze human tissue samples. All of the different pipelines reflected a similar picture of the results, although the mothur pipeline tended to overestimate the number of OTUs. DADA2 allowed the retrieval of the most conservative data set, with a lower number of RSVs than mothur and QIIME. In principle, and as also shown by *in silico* analysis, primer pair 349af/519ar could even be used for NGS amplicon studies without a nested approach with a first PCR with 344f/915r, as the amplicon has an appropriate length and the overall coverage for archaea would allow the identification of a wide taxonomic diversity (44). However, our methodology required the addition of Illumina adapter sequences to the primers, which resulted in severe primer dimer formation during a single-step PCR, making this approach ineffective in this experimental study (for further details, see Text S1).

The combination of primer pair 349af/519ar (344f/915r) and the DADA2 pipeline allows for the most conservative detection of archaeal signatures in human samples. Our NGS approach was based on the use of the Illumina MiSeq platform to study the community composition and diversity of archaea in human tissue samples. In proof-of-principle experiments, we assessed the efficiency of one primer pair, namely, 349af/519ar (344f/915r) (selected on the basis of its clear archaeal specificity and promising results in the natural mock community experiment), to analyze the diversity of archaea in various human microbiome samples, including gut biopsy, lung (bronchoalveolar lavage [BAL] fluid), olfactory mucosa (nasal) samples, and skin wipe samples from the forearm, chest, and back. As the Illumina data error profiles have not been well understood, and the widely used publicly available sequence data analysis pipelines were originally not designed for the characteristics of Illumina data (47), we were particularly interested in the impact of the sequence processing platforms chosen, namely, QIIME, mothur, and DADA2, on the results acquired (for further details on QIIME and mothur, see Text S1). DADA2, published in 2016, is an open-source software package designed to model and correct Illumina-sequenced amplicon errors in particular. It has been demonstrated to identify more variants and produce fewer spurious sequence reads than other sequence processing tools. DADA2 does not have a clustering step, and it is a full amplicon workflow with steps including filtering, dereplication, sample inference, chimera identification, and merging of paired-end reads (39). The different principle of DADA2 was clearly reflected in our results, as DADA2 retrieved approximately 200 bacterial RSVs from the unspecific approach with primer pair 515f/806r, whereas mothur detected more than 20,000 OTUs in the same sequence data sets. The same trend was observed with archaeal OTUs and RSVs, respectively, as we obtained 911 archaeal OTUs with QIIME, 999 OTUs with mothur, and 50 RSVs with DADA2 (for full details, see Table S2).

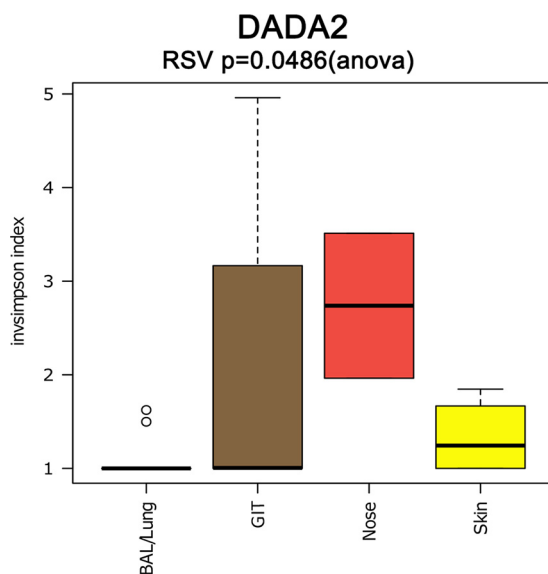


FIG 1 Alpha diversity based on the inverse Simpson diversity index in BAL fluid/lung (blue), GIT (brown), nasal (red), and skin (yellow) samples.

When analyzing the stool samples that served as natural mock communities, the most conservative, and thus most stringent, approach was found to be the combination of primer pair 349af/519ar (344f/915r) amplicons with data processing through DADA2. Here, DADA2 also retrieved the most conservative estimation of archaeal diversity as estimates based on the number of RSVs but also reflected the broadest diversity based on the number of taxa identified. Consequently, in the following, we will focus on the results retrieved from the DADA2 pipeline. However, additional results of pipeline comparisons are fully described in Text S1.

First insights into the archaeome of the human GIT, skin, nose, and lung reveal unexpected diversity and novel human archaeome signatures. On the basis of the above-described conservative NGS protocol, the nasal samples carried the highest diversity of archaea, followed by skin and gut samples (Fig. 1; Text S1; Fig. S2A), whereas the lowest diversity was retrieved from lung samples. The difference in alpha diversity between the body sites was found to be significant ($P = 0.0486$; DADA2 output; Fig. 1).

In total, signatures from five archaeal phyla (*Euryarchaeota*, *Thaumarchaeota*, *Diapherotrites*, *Woesearchaeota*, *Aenigmarchaeota*; the last three phyla are associated with the DPANN [*Diapherotrites*, *Parvarchaeota*, *Aenigmarchaeota*, *Nanoarchaeota*, *Nanohaloarchaea*] superphylum) (48) were identified in the human samples, along with numerous unclassified sequences (Table S2; more information is shown below).

Besides the finding that not only the skin (27) and gut, but also the nose and lung harbor a unique archaeal community, we were most surprised by the retrieval of *Methanobacterium* signatures from the human gut (see below) and the detection of a broad variety of DPANN superphylum-associated sequences from diverse human samples. In particular, members of *Woesearchaeota* were identified as the dominating archaeal taxon in the human lung. Signatures from the DPANN superphylum have never been associated with the human microbiome before. However, the presence of *Woesearchaeota* signatures has been reported in human-associated areas such as door handles (49) and in dust on the International Space Station (50), as well as in other environmental samples (49, 51, 52). Information on this clade of archaea is sparse, except for speculations on a potential parasitic/symbiotic lifestyle based on the observation of a small, reduced genome (53). Very recent studies suggest a close relationship to *Eukarya* (54), but their role in the environment is still unclear.

Besides methanogens, haloarchaea have also been reported to be coinhabitants of the colon and potentially also of the skin (24, 27, 55). Notably, haloarchaeal taxa were

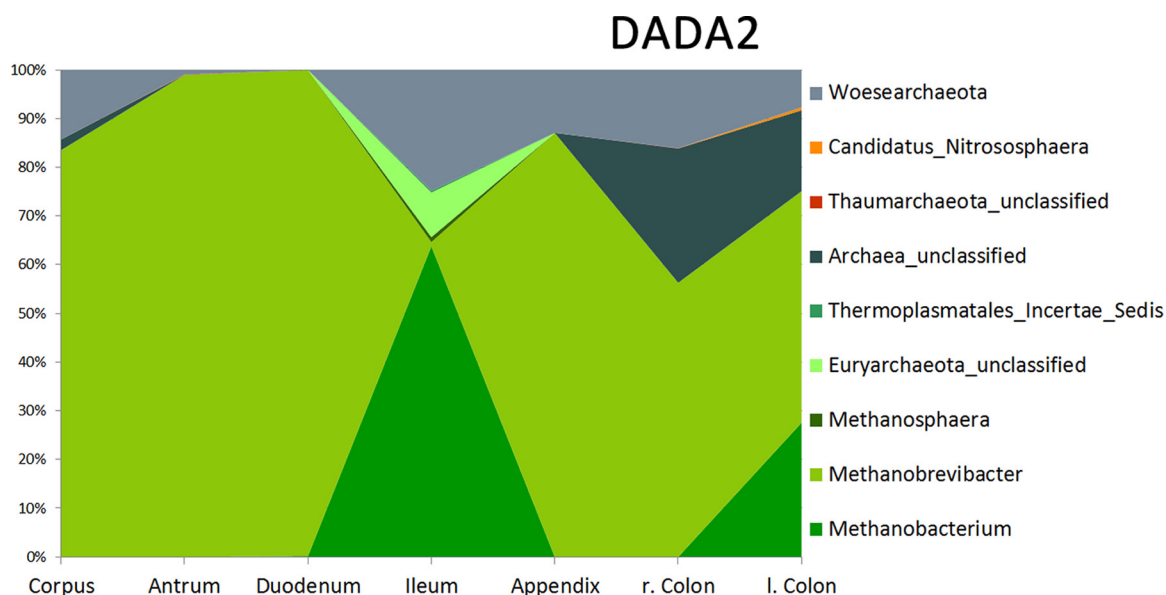


FIG 2 Area plot of the relative abundance (percent) of archaeal signatures (at the genus level) detected in GIT biopsy samples. *Methanoarchaea* are displayed in shades of green. The arrangement on the x axis reflects the sequential of the locations in the human GIT (from the stomach to the anus).

not detected in any of the tissue samples in our study. Since the primers used were found to be able to also target haloarchaea *in silico*, our results indicate an absence of this archaeal clade in our samples.

Archaea-specific amplification, NGS, and sequence processing procedures allow novel insights into the composition and distribution of the human GIT archaeome. On the basis of the data sets retrieved, we further explored the archaeal communities associated with different parts of the human GIT. As stool samples do not allow a deeper insight into the mucosa-associated microbiota, we selected biopsy specimens from seven different GIT locations, the corpus, antrum, duodenum, ileum, appendix, right colon, and left colon, for this study.

To explore possible differences between the archaeal GIT communities, alpha diversity (inverse Simpson index) was assessed and multivariate analyses were performed (Text S1; Fig. S3A). Statistical analyses did not show a significant ($P > 0.05$) influence of the sample location within the GIT on the diversity of archaea and community composition. Redundancy analysis (RDA), however, indicated a potential separate clustering of left colon and ileum communities (Text S1; Fig. S3B).

This was confirmed by area plots on the genus level, which also revealed a dissimilar picture in different GIT areas (Fig. 2; Text S1; Fig. S4).

Methanobrevibacterial signatures were detected in each location within the GIT, clearly representing the most abundant archaeal taxa therein. However, many additional sequences fell into the *Methanobrevibacter* clade but could not be classified to species level on the basis of the outcomes of the sequence processing pipelines. Taxonomic analysis (tree construction; see below), however, indicated clustering with known *Methanobrevibacter* signatures, including *M. smithii*, *M. oralis*, *M. arboriphilus*, *M. filiformis*, and *M. woesei*.

Overall, 70 to 90% of the reads from GIT samples (depending on the pipeline used) belonged to *Methanobrevibacter*, followed by *Methanobacterium* signatures (5 to 15%). More specifically, DADA2 retrieved 0.8% *Methanobrevibacter* signatures in ileum samples and 99.7% in duodenum biopsy specimens; *Methanobacterium* signatures made up 0.18% of the duodenum samples and 63.7% of the ileum biopsy specimens, indicating a prevalence of *Methanobacterium* particularly in the ileum. Signatures of the *Methanosphaera* clade were also detected in GIT samples, as well as *Methanomassiliicoccales* ("*Methanoplasmatales*"/*Thermoplasmatales*) signatures (56).

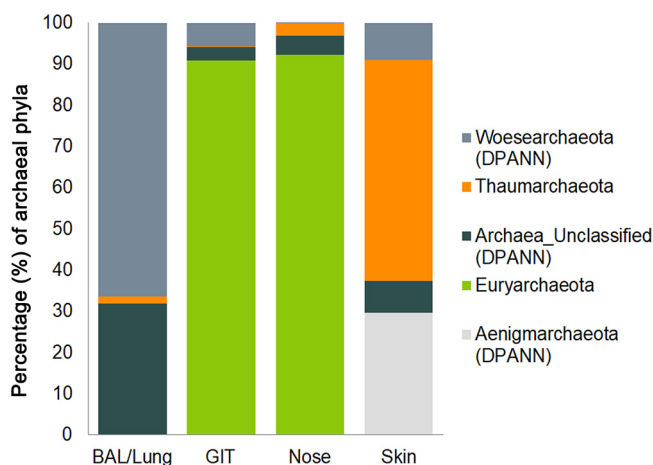


FIG 3 Archaeal composition at the phylum level in BAL fluid/lung, gut, nasal, and skin samples. The bar chart shows the relative abundance (percent) of the archaeal phyla in samples taken from different body sites. The different phyla are indicated by different colors, as shown on the right. DPANN archaeal signatures are shown in different shades of gray.

As mentioned above, the detection of *Methanobacterium* signatures was unexpected, as the presence of this genus in the human GIT has not been reported before. *Methanobacterium* is usually found in anaerobic digesters, freshwater sediments, marshy soils, and the rumen of cattle or sheep (57). Similar to *Methanobrevibacter*, it can use H₂ and CO₂, as well as formate, as a substrate for methanogenesis. However, as no other study has so far reported the presence of *Methanobacterium* in human gut samples and our observations were retrieved from NGS data only, we cannot infer physiological consequences from our findings. Further, as we were able to detect *Methanobacterium* solely in biopsy samples from the ileum and left colon, one might speculate about a potential association with the gut mucosa, hindering proper detection in stool samples.

The archaeome of the human body is site specific. To explore possible differences between the archaeal communities of the different body sites, multivariate analyses were performed. RDA was used to visualize how the archaeal communities of the different sites (lung, gut, nose, skin) relate to the other body sites and to test if the variation can be explained by the differences between sampling sites (Text S1; Fig. S2B).

On the basis of our results, we were able to assign a specific archaeal community to each of the body parts analyzed; independently of the data processing pipeline used, the body location was found to be the determining factor of the microbial community's composition ($P < 0.05$) (Fig. 3; DADA2 data set; Text S1; Fig. S5). This finding is in congruence with previous findings on bacterial communities in general (33). Costello et al. reported a predictable biogeographic bacterial pattern for body sites and even within different parts of body sites (58), which evidently is also true for archaeal communities.

Although RDA plots revealed a grouping of nose- and skin-associated archaeal communities, the lung and gut communities did not separate clearly, indicating a certain similarity of these communities on the RSV level (Text S1; Fig. S2B). These observations were partially confirmed by principal-coordinate analysis (PCoA) (Fig. 4; Text S1; Fig. S6A).

Overall, the information retrieved indicated a niche differentiation of human-associated archaeal communities; independently of the pipeline used, (anaerobic) methanogenic *Euryarchaeota* signatures were frequently detected throughout the anoxic human GIT (15) and nasal samples but were hardly found in skin or lung samples. Nasal samples harbored both methanogenic *Euryarchaeota* and skin-associated *Thaumarchaeota* signatures, indicating an overlap of the niche properties of

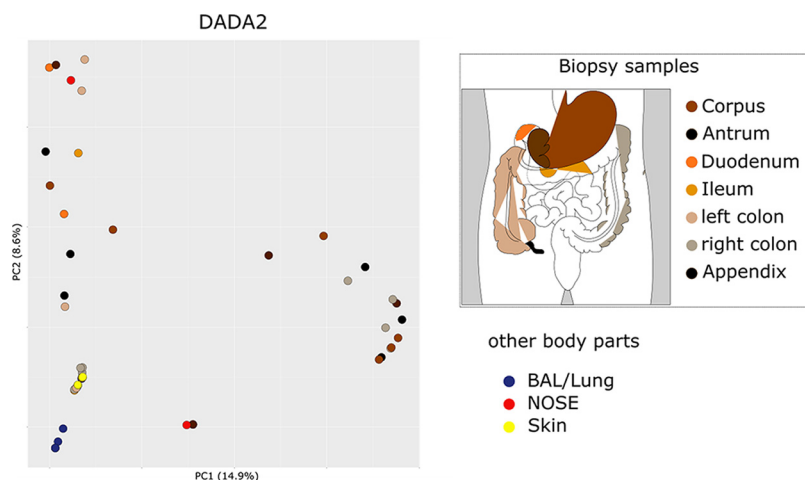


FIG 4 PCoA plot of the archaeal communities retrieved from human GIT, BAL fluid (lung), nasal, and skin samples. The color key explains the coloring of the samples according to their origins.

the skin and digestive tract. The predominance of thaumarchaeal signatures on oxygenated, ammonia/urea-rich human skin (27, 29) was confirmed, and DPANN-associated archaea (*Woesearchaeota* and unassigned archaea) were detected specifically in samples from the lung (Fig. 5).

On the basis of these findings, the archaeal biogeographic pattern can be divided into (i) a thaumarchaeal skin landscape, (ii) a (methano)eurarchaeal GIT landscape, (iii) a mixed skin-GIT landscape for the nose, and (iv) a woesearchaeal lung landscape.

Network analyses reveal the connection of the archaeal communities on the OTU/RSV level. OTUs/RSVs of all three processing pipelines were visualized as networks (Text S1; Fig. S6B). For a better overview, the network generated on the basis of the DADA2 output was arranged according to the different body sites as shown in Fig. 6.

The most abundant RSVs (assigned to the genus *Methanobrevibacter*) were connected with archaeal communities from all parts of the GIT. However, some GIT archaeal communities (i.e., from the left colon, right colon, ileum, and corpus) remained unconnected on the RSV level and contained only one RSV per sample. These RSVs were assigned as unclassified *Archaea*, *Woesearchaeota*, or *Euryarchaeota*.

Notably, only one GIT sample (right colon) shared one RSV node (thaumarchaeal signature) with two BAL fluid/lung samples; however, a remarkably high percentage of biopsy samples (56%) were connected with RSVs also abundant in the nose (mainly assigned to *Methanobrevibacter*). Only one skin sample was connected with the antrum sample through a thaumarchaeal node. Overall, archaeal signatures deriving from the GIT were connected to all other body parts (in particular, nasal samples), but skin and lung samples were not well connected with those from other body parts on the basis of RSVs.

Outlook. Although our study is one of the first steps in the elucidation of the human archaeome, much more investigation and development is required to assess the abundance (i.e., the ratio of *Bacteria* to *Archaea*), distribution, and function of the human-associated archaea. Other methods and technologies, including visualization, cultivation, and functional *in vivo* studies, are necessary to answer the final question of whether archaea are really “never bad” (59).

In summary, our study highlights the importance of the choice of primers and data processing pipeline for studying the human archaeome. Despite these pitfalls, we were able to establish protocols that allow the specific detection of a broad diversity of human-associated archaea. Moreover, this technique allowed us to identify potential novel archaeal members of the human microbial communities, detect archaeal signatures in the nose and lung, and finally reveal body site specificity for the human archaeome.

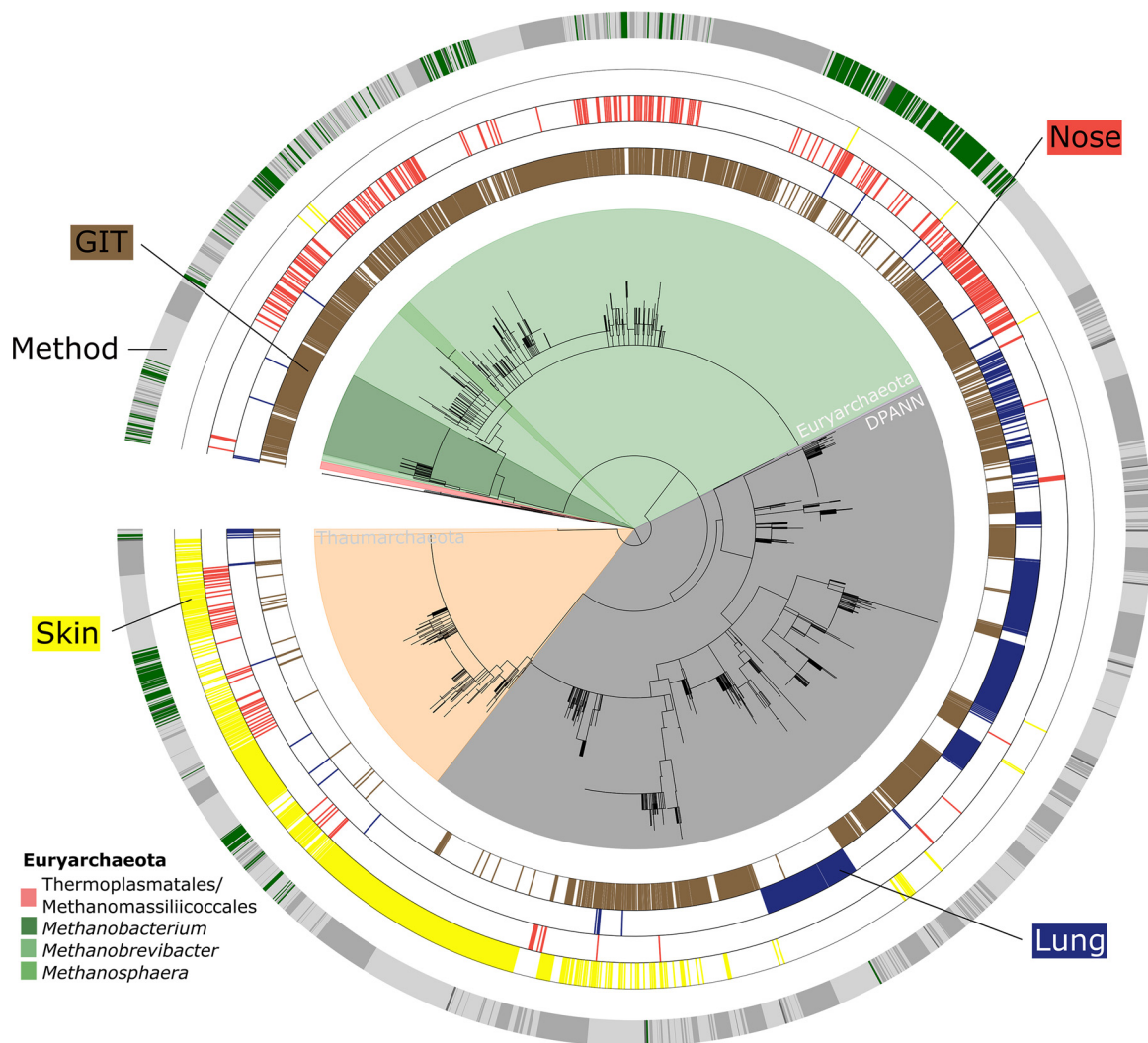


FIG 5 Phylogenetic tree of retrieved archaeal signatures affiliated with *Euryarchaeota*, *Thaumarchaeota*, and the DPANN superphylum. The “method” circle displays the origins of the respective OTUs/RSVs. Green represents database sequences of next neighbors as identified through SILVA SINA aligner. Dark gray represents RSV sequences (DADA2), medium gray represents OTUs retrieved through the mothur approach, and light gray represents the QIIME procedure. The four inner circles display the origins of the respective OTUs, i.e., the skin (yellow), nose (red), lung (blue), and gut (brown).

MATERIALS AND METHODS

General remarks. Research involving human material was performed in accordance with the Declaration of Helsinki and was approved by the local ethics committees (the Ethics Committee at the Medical University of Graz, Graz, Austria, and the University of Kiel, Kiel, Germany). (Bacterial) microbiome studies of these samples have already been published elsewhere (60–63), and details of the ethics approvals obtained are also shown there. DNA extracted in those previous experiments was used to analyze the archaeome in this study.

Primer sequences and target genes for the specific detection of archaea in human samples. The primer sequences used, target genes, and references are shown in Text S1.

Samples, sample processing, and amplification procedure for the detection of specific methanoarchaeal gene signatures in gut biopsy samples (approach 1). For this approach, 25 biopsy samples (3 to 20 mg) taken from the lower part of the colon were processed, including 5 samples from healthy volunteers, 10 samples from patients with UC (5 samples from inflamed tissue and 5 from uninfamed tissue), and 10 samples from patients with CD (5 samples from inflamed tissue and 5 from uninfamed tissues). Samples were immediately frozen at -80°C . DNA was isolated by a combination of mechanical lysis (glass beads) and chemical and enzymatic digestion. For details, see Text S1.

DNA extracted from gut biopsy specimens (5 healthy, 5 UC inflamed, 5 UC uninfamed, 5 CD inflamed, 5 CD uninfamed) was subjected to a nested PCR approach specifically targeting *M. smithii*, *M. stadtmanae*, and *M. luminyensis*, the most abundant methanogenic archaea in the human gut. On the basis

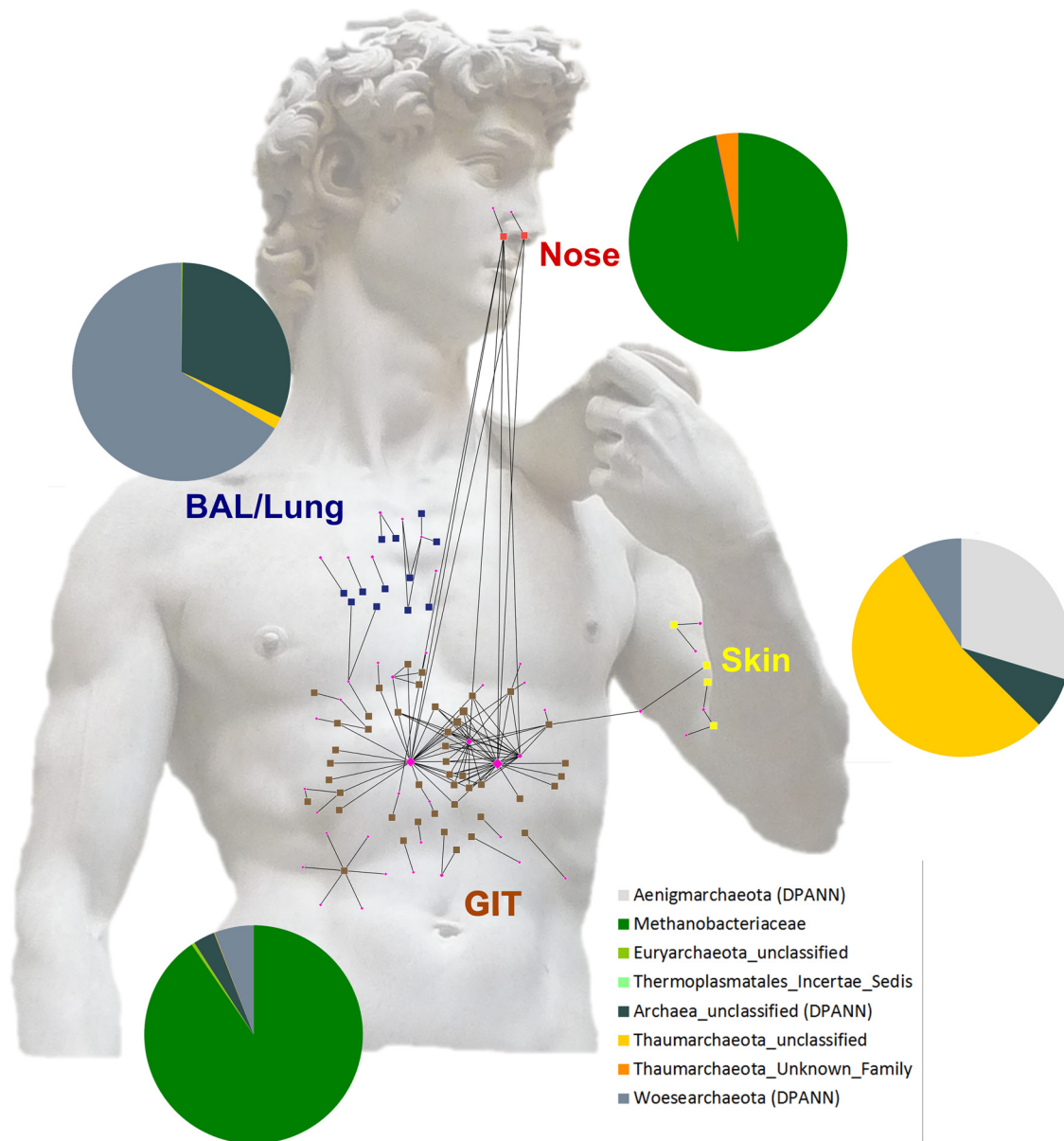


FIG 6 The archaeal taxonomic landscape of the human body visualized as a network. RSVs (DADA2 output) obtained from four sampling sites (BAL fluid/lung, blue; skin, yellow; nose, red; GIT, brown) are plotted and arranged according to their sampling origins.

of the study of Dridi et al. (16), we decided to target the *rpoB* gene of *M. smithii* and *M. stadtmanae* because of its higher specificity than the 16S rRNA gene system. This gene encodes the β subunit of RNA polymerase and has been previously used to unravel phylogenetic relationships within the archaeal domain (16, 64) (Text S1).

DNA extracted from *M. smithii* DSM861, *M. stadtmanae* DSM3091, and *M. luminyensis* B10 DSM25720, obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ), Braunschweig, Germany, served as positive controls, whereas sterile double-distilled water and DNA extracted from *Escherichia coli* K-12 served as negative controls.

Template DNA (100 ng of biopsy sample DNA, 100 pg of control DNA) was added to a 25- μ l PCR mixture including GoTaq DNA polymerase (Promega, Mannheim, Germany), 10 μ M primer, and 0.4 mM deoxynucleoside triphosphates. For the cycling conditions used, see Text S1.

Following the PCR or nested PCRs, amplified fragments were analyzed by agarose gel electrophoresis. To confirm the identities of the signatures retrieved, selected nested PCR products were subjected to Sanger sequencing at the sequencing facility of the Institute of Clinical Molecular Biology, University of Kiel, Kiel, Germany. For details, see Text S1. As the PCR products obtained were too long for Illumina-based NGS, they were TOPO-TA cloned (see Text S1). For the corresponding sequences, see the appendix in Text S1.

Stool samples, sample processing, and amplicon generation for NGS-based specific assessment of the diversity of archaea in a natural mock community (method establishment and verification, approach 2). In our study, stool samples from two healthy adults served as a natural mock community. For sample processing details, see Text S1. All samples were immediately stored at -80°C until DNA isolation. The DNA was extracted by a combination of mechanical lysis and chemical and enzymatic digestion. For details, see Text S1. DNA extracted from two different stool samples, serving as natural mock communities, was subjected to PCR with three primer combinations to evaluate the quality of the sequencing results retrieved by NGS with regard to the diversity of archaea obtained and the specificity of the approach chosen. First, we chose primers 515f and 806r, targeting both *Bacteria* and *Archaea*, which currently represents the primer pair used the most for microbiome studies, as proposed by the HMP and the EMP (43). For the other approaches, we selected more archaea-specific procedures, based on a nested PCR approach using primers 344f and 915r in the first PCR and primer pair S-D-Arch-0349-a-S-17/S-D-Arch-0519-a-A-16 (here 349af/519ar), targeting *Archaea*, or primer pair S-D-Arch-0519-a-S-15/S-D-Bact-0785-b-A-18 (here 519af/785ur), targeting both *Bacteria* and *Archaea*, in the subsequent PCR (44). For the primer sequences and amplification protocols used, see Text S1. Briefly, amplified fragments from the first PCR were gel purified and $1\ \mu\text{l}$ of the eluate obtained was subjected to the subsequent PCR for 25 additional cycles. We used DNA from *Pyrococcus furiosus* as a positive control and PCR grade water as a negative control. All samples were amplified and sequenced in triplicate. All of the data sets retrieved were subjected to data processing as described below.

Samples, sample processing, and amplicon generation for the NGS-based, specific assessment of the diversity of archaea in human tissue samples (method application, approach 2). For this part of the study, we processed a variety of human samples, including biopsy specimens from the GIT, BAL fluid, and nasal and skin swab samples (for details, see Text S1). The biopsy samples were collected at the Medical University of Graz from seven different sites in the GIT ([gastric] corpus [$n = 11$], [gastric] antrum [$n = 11$], duodenum [$n = 11$], [terminal] ileum [$n = 11$], appendiceal orifice [appendix; $n = 11$], ascending [right] colon [$n = 11$], and sigmoid [left] colon [$n = 11$]). For biopsy specimen collection, 11 volunteers underwent gastroduodenoscopy, followed by colonoscopy on the next day (60, 61). Patients (>18 years old) were recruited for BAL fluid sampling ($n = 36$). The study subjects were nonneutropenic intubated and mechanically ventilated patients in an intensive care unit (62). Lung microbiome samples were obtained by bronchoscopy of the right lung through endotracheal tubes. Nasal microbiome samples from healthy volunteers were taken by an ear-nose-throat physician from the olfactory mucosa located at the ceiling of the nasal cavity with ultra minitip nylon flocked swabs (Copán, Brescia, Italy; $n = 2$). Skin samples were taken from two volunteers at three different body locations, i.e., the exterior side of the left forearm ($n = 2$), chest ($n = 2$), and back ($n = 2$), with the BD Culture Swabs EZ Collection and Transport system (63).

For the screening of different human samples (biopsy specimens and lung, nose, and skin samples), we chose one amplification approach, namely, the first PCR with primer pair 344f/915r and a subsequent nested PCR with primer pair 349af/519ar as described in Text S1. In this case, amplicons from the first PCR were purified with the MinElute PCR Purification kit (Qiagen, Hilden, Germany). For an overview of the samples processed (type and number) and the approach chosen, see Text S1.

NGS, bioinformatics, and statistical analyses. All of the amplicons produced were sequenced at ZMF Core Facility Molecular Biology in Graz, Austria, with the available MiSeq platform (63).

Sequence data processing was performed with three different sequence data analysis programs, namely, QIIME (v1.9.1), mothur (v1.36.1), and DADA2 (v1.1). Data processing with QIIME and mothur was conducted with an in-house Galaxy server setup (65, 66). The parameters for the pipelines were created in accordance with the SOPs provided by their developers (67) (Illumina overview tutorial [http://nbviewer.jupyter.org/github/biocore/qiime/blob/1.9.1/examples/ipynb/illumina_overview_tutorial.ipynb]). Briefly, for QIIME and mothur, paired-end reads were joined, trimmed, and checked for chimeric sequences. OTUs were picked in accordance with the open reference strategy in QIIME and with the average neighbor algorithm in mothur with a cutoff of 97% identity. Taxonomic affiliations were determined with SILVA v123 (68) as the reference database.

The third type of amplicon processing was performed with the open source package DADA2 (39) in R (version 3.2.2) as described previously (50). In general, the DADA2 approach is based on the modeling and correction of amplicon errors without constructing OTUs, and in contrast to mothur and QIIME, DADA2 performs merging of paired-end reads after denoising. In brief, the DADA2 approach turns paired-end fastq files into merged, denoised, chimera-free, and inferred sample sequences, so-called RSVs with fewer incorrect sequences than any OTU method (39). In our analysis, sequences were quality checked, filtered, and trimmed to a consistent length (140 bp). Passed sequences were dereplicated and subjected to the DADA2 core algorithm. After chimera filtering and merging of paired reads, sequences were assigned a taxonomic classification by using the RDP classifier and the SILVA database v123 (68). In the resulting RSV table, each row corresponds to a nonchimeric inferred sample sequence with a separate taxonomic classification.

During data processing, samples with no archaeal OTUs/RSVs were removed and only samples with archaeal signals were processed further.

Every step of the analysis included negative controls consisting of blank DNA extractions for each group of samples and negative controls during PCR amplifications. OTUs/RSVs that overlapped negative controls and samples were removed from the data sets.

After data processing, alpha diversity and beta diversity (RDA) were calculated and visualized with Calypso (69). PCoA (based on Bray-Curtis distance) for OTU/RSV clustering was performed with the package phyloseq (70). The OTU/RSV tables obtained were used to summarize taxon abundance at

different taxonomic levels (domain, phylum, class, order, family, and genus). The taxonomic profiles obtained were used to generate bar graphs at the phylum level for all of the body locations investigated, area plots at the genus level for the biopsy specimens from the GIT, and pie charts at the genus level on the DADA2 results for all samples. In the pretest to evaluate the performance of primer pairs in the detection of archaeal OTUs/RSVs in human stool samples, we assessed the differences in the richness of OTUs/RSVs by one-way ANOVA and a *t* test.

To analyze and visualize the diversity of archaea detected in human samples, a phylogenetic tree was constructed on the basis of one representative sequence from each OTU retrieved, as obtained through the QIIME and mothur pipelines, and on the basis of RSV sequences taken from the DADA2 output. To include closely related reference sequences from a quality-checked and regularly updated database, SILVA SINA (71) was used to identify the 10 most closely related available sequences (neighbors). The neighbor sequences were downloaded and included in the retrieved sequence data set. All sequences were aligned and cropped to the same length before a tree based on the maximum-likelihood algorithm (MEGA6) (72) was constructed. The Newick output was processed further with the iTOL interactive online platform (73). Data sets for, e.g., the method or the origin of the OTU were added in accordance with the instructions on the iTOL help pages.

To visualize the relationship between archaeal OTUs/RSVs within the human archaeome, a network analysis of the data sets retrieved from the processing pipelines (QIIME, mothur, and DADA2) was performed. Respective OTUs and RSVs were clustered, and E weights were calculated with a stochastic spring-embedded algorithm. The resulting edge and node tables were visualized with cytoscape 2.8.3 (74). OTUs/RSVs were colored according to their sample origins, and their abundance was correlated with node size.

Accession number(s). The MiSeq amplicon sequence data obtained in this study were deposited in the European Nucleotide Archive under study accession number PRJEB19529.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.00824-17>.

TEXT S1, DOCX file, 0.2 MB.

FIG S1, TIF file, 0.3 MB.

FIG S2, TIF file, 0.1 MB.

FIG S3, TIF file, 0.2 MB.

FIG S4, TIF file, 0.5 MB.

FIG S5, TIF file, 0.8 MB.

FIG S6, TIF file, 0.4 MB.

TABLE S1, XLSX file, 0.03 MB.

TABLE S2, XLSX file, 1.5 MB.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of BioTechMed-Graz, Medical University of Graz (Ph.D. program in molecular medicine) and DFG (Schm1052/11-2).

We thank Christoph Högenauer (Division of Gastroenterology and Hepatology, Medical University of Graz, Graz, Austria), Thomas Pieber and Mina Bashir (Department of Internal Medicine, Medical University of Graz, Graz, Austria), and Stephan Ott (Institute of Clinical Molecular Biology, University of Kiel, Kiel, Germany) for providing biopsy samples, as well as Robert Krause (Department of Internal Medicine, Medical University of Graz, Graz, Austria) for providing the BAL fluid samples. We are grateful for nasal swab samples provided by Veronika Schöpf (Institute of Psychology, University of Graz, Graz, Austria) and skin wipe samples provided by Peter Wolf (Department of Dermatology and Venereology, Medical University of Graz, Graz, Austria) and Ingeborg Klymiuk (Center for Medical Research, Medical University of Graz, Graz, Austria). M. Mora and M. R. Pausan were trained within the frame of the Ph.D. program in Molecular Medicine of the Medical University of Graz.

REFERENCES

- Human Microbiome Project Consortium. 2012. A framework for human microbiome research. *Nature* 486:215–221. <https://doi.org/10.1038/nature11209>.
- Parfrey LW, Walters WA, Knight R. 2011. Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. *Front Microbiol* 2:153. <https://doi.org/10.3389/fmicb.2011.00153>.
- Huffnagle GB, Noverr MC. 2013. The emerging world of the fungal microbiome. *Trends Microbiol* 21:334–341. <https://doi.org/10.1016/j.tim.2013.04.002>.
- Halwachs B, Madhusudhan N, Krause R, Nilsson RH, Moissl-Eichinger C, Högenauer C, Thallinger GG, Gorkiewicz G. 2017. Critical issues in Mycobacteria analysis. *Front Microbiol* 8:180. <https://doi.org/10.3389/fmicb.2017.00180>.
- Zou S, Caler L, Colombini-Hatch S, Glynn S, Srinivas P. 2016. Research on

- the human virome: where are we and what is next. *Microbiome* 4:32. <https://doi.org/10.1186/s40168-016-0177-y>.
6. Lurie-Weinberger MN, Gophna U. 2015. Archaea in and on the human body: health implications and future directions. *PLoS Pathog* 11: e1004833. <https://doi.org/10.1371/journal.ppat.1004833>.
 7. Nkamga VD, Henrissat B, Drancourt M. 2017. Archaea: essential inhabitants of the human digestive microbiota. *Hum Microbiome J* 3:1–8. <https://doi.org/10.1016/j.humic.2016.11.005>.
 8. Horz HP, Conrads G. 2010. The discussion goes on: what is the role of Euryarchaeota in humans? *Archaea* 2010:967271. <https://doi.org/10.1155/2010/967271>.
 9. Horz HP, Conrads G. 2011. Methanogenic *Archaea* and oral infections—ways to unravel the black box. *J Oral Microbiol* 3:5940. <https://doi.org/10.3402/jom.v3i0.5940>.
 10. Hajishengallis G, Darveau RP, Curtis MA. 2012. The keystone-pathogen hypothesis. *Nat Rev Microbiol* 10:717–725. <https://doi.org/10.1038/nrmicro2873>.
 11. Miller TL, Wolin MJ, Conway de Macario EC, Macario AJ. 1982. Isolation of *Methanobrevibacter smithii* from human feces. *Appl Environ Microbiol* 43:227–232.
 12. Miller TL, Wolin MJ. 1985. *Methanosphaera stadtmaniae* gen. nov., sp. nov.: a species that forms methane by reducing methanol with hydrogen. *Arch Microbiol* 141:116–122. <https://doi.org/10.1007/BF00423270>.
 13. Dridi B, Fardeau ML, Ollivier B, Raoult D, Drancourt M. 2012. *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int J Syst Evol Microbiol* 62:1902–1907. <https://doi.org/10.1099/ijso.0.033712-0>.
 14. Ferrari A, Brusa T, Rutili A, Canzi E, Biavati B. 1994. Isolation and characterization of *Methanobrevibacter oralis* sp. nov. *Curr Microbiol* 29:7–12. <https://doi.org/10.1007/BF01570184>.
 15. Gaci N, Borrel G, Tottey W, O'Toole PW, Brugère JF. 2014. Archaea and the human gut: new beginning of an old story. *World J Gastroenterol* 20:16062–16078. <https://doi.org/10.3748/wjg.v20.i43.16062>.
 16. Dridi B, Henry M, El Khéchine A, Raoult D, Drancourt M. 2009. High prevalence of *Methanobrevibacter smithii* and *Methanosphaera stadtmanae* detected in the human gut using an improved DNA detection protocol. *PLoS One* 4:e7063. <https://doi.org/10.1371/journal.pone.0007063>.
 17. Dridi B, Henry M, Richet H, Raoult D, Drancourt M. 2012. Age-related prevalence of *Methanomassiliicoccus luminyensis* in the human gut microbiome. *APMIS* 120:773–777. <https://doi.org/10.1111/j.1600-0463.2012.02899.x>.
 18. Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henrissat B, Fulton R, Latreille P, Kim K, Wilson RK, Gordon JL. 2007. Genomic and metabolic adaptations of *Methanobrevibacter smithii* to the human gut. *Proc Natl Acad Sci U S A* 104:10643–10648. <https://doi.org/10.1073/pnas.0704189104>.
 19. McInerney MJ, Sieber JR, Gunsalus RP. 2009. Syntrophy in anaerobic global carbon cycles. *Curr Opin Biotechnol* 20:623–632. <https://doi.org/10.1016/j.copbio.2009.10.001>.
 20. Blais Lecours PB, Marsolais D, Cormier Y, Berberli M, Haché C, Bourdages R, Duchaine C. 2014. Increased prevalence of *Methanosphaera stadtmanae* in inflammatory bowel diseases. *PLoS One* 9:e87734. <https://doi.org/10.1371/journal.pone.0087734>.
 21. Vianna ME, Conrads G, Gomes BP, Horz HP. 2006. Identification and quantification of archaea involved in primary endodontic infections. *J Clin Microbiol* 44:1274–1282. <https://doi.org/10.1128/JCM.44.4.1274-1282.2006>.
 22. Nguyen-Hieu T, Khelaifia S, Aboudharam G, Drancourt M. 2013. Methanogenic archaea in subgingival sites: a review. *APMIS* 121:467–477. <https://doi.org/10.1111/apm.12015>.
 23. Drancourt M, Nkamga VD, Lakhe NA, Régis JM, Dufour H, Fournier PE, Bechah Y, Scheld WM, Raoult D. 2017. Evidence of archaeal methanogens in brain abscess. *Clin Infect Dis* 65:1–5. <https://doi.org/10.1093/cid/cix286>.
 24. Khelaifia S, Raoult D. 2016. *Haloferax massiliensis* sp. nov., the first human-associated halophilic archaea. *New Microbes New Infect* 12: 96–98. <https://doi.org/10.1016/j.nmni.2016.05.007>.
 25. Nam YD, Chang HW, Kim KH, Roh SW, Kim MS, Jung MJ, Lee SW, Kim JY, Yoon JH, Bae JW. 2008. Bacterial, archaeal, and eukaryal diversity in the intestines of Korean people. *J Microbiol* 46:491–501. <https://doi.org/10.1007/s12275-008-0199-7>.
 26. Belay N, Mukhopadhyay B, Conway de Macario E, Galask R, Daniels L. 1990. Methanogenic bacteria in human vaginal samples. *J Clin Microbiol* 28:1666–1668.
 27. Probst AJ, Auerbach AK, Moissl-Eichinger C. 2013. Archaea on human skin. *PLoS One* 8:e65388. <https://doi.org/10.1371/journal.pone.0065388>.
 28. Moissl-Eichinger C, Pausan M, Taffner J, Berg G, Bang C, Schmitz RA. 18 August 2017. Archaea are interactive components of complex microbiomes. *Trends Microbiol* <https://doi.org/10.1016/j.tim.2017.07.004>.
 29. Moissl-Eichinger C, Probst AJ, Birarda G, Auerbach A, Koskinen K, Wolf P, Holman H-YN. 2017. Human age and skin physiology shape diversity and abundance of Archaea on skin. *Sci Rep* 7:4039. <https://doi.org/10.1038/s41598-017-04197-4>.
 30. Adams RI, Bateman AC, Bik HM, Meadow JF. 2015. Microbiota of the indoor environment: a meta-analysis. *Microbiome* 3:49. <https://doi.org/10.1186/s40168-015-0108-3>.
 31. Gilbert JA, Jansson JK, Knight R. 2014. The Earth Microbiome Project: successes and aspirations. *BMC Biol* 12:69. <https://doi.org/10.1186/s12915-014-0069-1>.
 32. Fischer MA, Güllert S, Neulinger SC, Streit WR, Schmitz RA. 2016. Evaluation of 16S rRNA gene primer pairs for monitoring microbial community structures showed high reproducibility within and low comparability between datasets generated with multiple archaeal and bacterial primer pairs. *Front Microbiol* 7:1297. <https://doi.org/10.3389/fmicb.2016.01297>.
 33. Ding T, Schloss PD. 2014. Dynamics and associations of microbial community types across the human body. *Nature* 509:357–360. <https://doi.org/10.1038/nature13178>.
 34. Oh J, Byrd AL, Deming C, Conlan S, NISC Comparative Sequencing Program, Kong HH, Segre JA. 2014. Biogeography and individuality shape function in the human skin metagenome. *Nature* 514:59–64. <https://doi.org/10.1038/nature13786>.
 35. Cavicchioli R, Curmi PMG, Saunders N, Thomas T. 2003. Pathogenic archaea: do they exist? *Bioessays* 25:1119–1128. <https://doi.org/10.1002/bies.10354>.
 36. Gill EE, Brinkman FSL. 2011. The proportional lack of archaeal pathogens: do viruses/phages hold the key? *Bioessays* 33:248–254. <https://doi.org/10.1002/bies.201000091>.
 37. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541. <https://doi.org/10.1128/AEM.01541-09>.
 38. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JL, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>.
 39. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583. <https://doi.org/10.1038/nmeth.3869>.
 40. Gaskins HR, Croix JA, Nakamura N, Nava GM. 2008. Impact of the intestinal microbiota on the development of mucosal defense. *Clin Infect Dis* 46:S80–S86; discussion S144–S151. <https://doi.org/10.1086/523336>.
 41. Pinto AJ, Raskin L. 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS One* 7:e43093. <https://doi.org/10.1371/journal.pone.0043093>.
 42. Bang C, Ehlers C, Orell A, Prasse D, Spinner M, Gorb SN, Albers SV, Schmitz RA. 2014. Biofilm formation of mucosa-associated methanarchaeal strains. *Front Microbiol* 5:353. <https://doi.org/10.3389/fmicb.2014.00353>.
 43. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621–1624. <https://doi.org/10.1038/ismej.2012.8>.
 44. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41:e1. <https://doi.org/10.1093/nar/gks808>.
 45. Stahl DA, Amann R. 1991. Development and application of nucleic acid probes in bacterial systematics, p 205–248. *In* Stackebrandt E, Goodfellow M (ed), *Nucleic acid techniques in bacterial systematics*. John Wiley & Sons, Chichester, United Kingdom.

46. Raskin L, Stromley JM, Rittmann BE, Stahl DA. 1994. Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Appl Environ Microbiol* 60:1232–1240.
47. Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C. 2015. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res* 43:e37. <https://doi.org/10.1093/nar/gku1341>.
48. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stephanoukaskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499:431–437. <https://doi.org/10.1038/nature12352>.
49. Ross AA, Neufeld JD. 2015. Microbial biogeography of a university campus. *Microbiome* 3:66. <https://doi.org/10.1186/s40168-015-0135-0>.
50. Mora M, Perras AK, Alekhova TA, Wink L, Krause R, Aleksandrova A, Novozhilova T, Moissl-Eichinger C. 2016. Resilient microorganisms in dust samples of the International Space Station—survival of the adaptation specialists. *Microbiome* 4:65. <https://doi.org/10.1186/s40168-016-0217-7>.
51. Kinsman-Costello LE, Sheik CS, Sheldon ND, Allen Burton G, Costello DM, Marcus D, Den Uyl PA, Dick GJ. 2017. Groundwater shapes sediment biogeochemistry and microbial diversity in a submerged Great Lake sinkhole. *Geobiology* 15:225–239. <https://doi.org/10.1111/gbi.12215>.
52. Long Y, Yi H, Chen S, Zhang Z, Cui K, Bing Y, Zhuo Q, Li B, Xie S, Guo Q. 2016. Influences of plant type on bacterial and archaeal communities in constructed wetland treating polluted river water. *Environ Sci Pollut Res Int* 23:19570–19579. <https://doi.org/10.1007/s11356-016-7166-3>.
53. Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, Frischkorn KR, Tringe SG, Singh A, Markillie LM, Taylor RC, Williams KH, Banfield JF. 2015. Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* 25:690–701. <https://doi.org/10.1016/j.cub.2015.01.014>.
54. Furukawa R, Nakagawa M, Kuroyanagi T, Yokobori SI, Yamagishi A. 2017. Quest for ancestors of eukaryal cells based on phylogenetic analyses of aminoacyl-tRNA synthetases. *J Mol Evol* 84:51–66. <https://doi.org/10.1007/s00239-016-9768-2>.
55. Oxley AP, Lanfrancioni MP, Würdemann D, Ott S, Schreiber S, McGenity TJ, Timmis KN, Nogales B. 2010. Halophilic archaea in the human intestinal mucosa. *Environ Microbiol* 12:2398–2410. <https://doi.org/10.1111/j.1462-2920.2010.02212.x>.
56. Paul K, Nonoh JO, Mikulski L, Brune A. 2012. “*Methanoplasmatales*,” *Thermoplasmatales*-related archaea in termite guts and other environments, are the seventh order of methanogens. *Appl Environ Microbiol* 78:8245–8253. <https://doi.org/10.1128/AEM.02193-12>.
57. Boone DR, Boone RD. 2015. *Methanobacterium*, p 1–8. In Whitman WB (ed), *Bergey's manual of systematics of Archaea and Bacteria*. John Wiley & Sons, Chichester, United Kingdom.
58. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. 2009. Bacterial community variation in human body habitats across space and time. *Science* 326:1694–1697.
59. Reeve JN, Schleper C. 2011. Archaea: very diverse, often different but never bad? *Curr Opin Microbiol* 14:271–273. <https://doi.org/10.1016/j.mib.2011.04.011>.
60. Bashir M, Prietl B, Tauschmann M, Mautner SI, Kump PK, Treiber G, Wurm P, Gorkiewicz G, Högenauer C, Pieber TR. 2016. Effects of high doses of vitamin D₃ on mucosa-associated gut microbiome vary between regions of the human gastrointestinal tract. *Eur J Nutr* 55:1479–1489. <https://doi.org/10.1007/s00394-015-0966-2>.
61. Kump PK, Gröchenig HP, Lackner S, Trajanoski S, Reicht G, Hoffmann KM, Deutschmann A, Wenzl HH, Petritsch W, Krejs GJ, Gorkiewicz G, Högenauer C. 2013. Alteration of intestinal dysbiosis by fecal microbiota transplantation does not induce remission in patients with chronic active ulcerative colitis. *Inflamm Bowel Dis* 19:2155–2165. <https://doi.org/10.1097/MIB.0b013e31829ea325>.
62. Krause R, Halwachs B, Thallinger GG, Klymiuk I, Gorkiewicz G, Hoenigl M, Prattes J, Valentin T, Heidrich K, Buzina W, Salzer HJF, Rabensteiner J, Prüller F, Raggam RB, Meinitzer A, Moissl-Eichinger C, Högenauer C, Quehenberger F, Kashofer K, Zollner-Schwetz I. 2016. Characterisation of candida within the mycobiome/microbiome of the lower respiratory tract of ICU patients. *PLoS One* 11:e0155033. <https://doi.org/10.1371/journal.pone.0155033>.
63. Klymiuk I, Bambach I, Patra V, Trajanoski S, Wolf P. 2016. 16S based microbiome analysis from healthy subjects' skin swabs stored for different storage periods reveal phylum to genus level changes. *Front Microbiol* 7:2012. <https://doi.org/10.3389/fmicb.2016.02012>.
64. Klenk HP, Zillig W. 1994. DNA-dependent RNA polymerase subunit B as a tool for phylogenetic reconstructions: branching topology of the archaeal domain. *J Mol Evol* 38:420–432. <https://doi.org/10.1007/BF00163158>.
65. Blankenberg D, Taylor J, Schenck I, He J, Zhang Y, Ghent M, Veeraraghavan N, Albert I, Miller W, Makova KD, Hardison RC, Nekrutenko A. 2007. A framework for collaborative analysis of ENCODE data: making large-scale analyses biologist-friendly. *Genome Res* 17:960–964. <https://doi.org/10.1101/gr.5578007>.
66. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster J, Von Kuster G, Rasche W, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44:W3–W10. <https://doi.org/10.1093/nar/gkw343>.
67. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120. <https://doi.org/10.1128/AEM.01043-13>.
68. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>.
69. Zakrzewski M, Proietti C, Ellis JJ, Hasan S, Brion M-J, Berger B, Krause L. 2017. Calypso: a user-friendly web-server for mining and visualizing microbiome-environment interactions. *Bioinformatics* 33:782–783. <https://doi.org/10.1093/bioinformatics/btw725>.
70. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <https://doi.org/10.1371/journal.pone.0061217>.
71. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829. <https://doi.org/10.1093/bioinformatics/bts252>.
72. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>.
73. Letunic I, Bork P. 2007. Interactive Tree of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128. <https://doi.org/10.1093/bioinformatics/btl529>.
74. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>.