1 Dear Dr. Akob,

2 We are very grateful for all the comments, which were very helpful and significantly improve the
3 manuscript.

4 We revised the manuscript according to your comments. We submit (1) the point-by-point
5 response and (2) the revised version of our manuscript with all changes marked with track
6 changes.

7

8 Yours sincerely,

9 Jörn Wehking

10

11 1. I think that Archaea should be capitalized throughout the paper since you are referring to a
12 Domain of life. Same with Bacteria if you are referring to the Domain and not to all single-celled
13 organisms.
14 We capitalized "Archaea" as well as "Bacteria" throughout the whole manuscript when referring
15 to the Domain.
16
17 2. The gene studied should always be written as "16S rRNA gene" and not abbreviated to "16S"
18 or "16S rRNA". "16S rRNA" would imply that your analysis was performed on a RNA level.
19 We changed to "16S rRNA gene" in the entire manuscript.
20
21 3. In SI and throughout the paper use "." instead of "," for decimal places.
22 We replaced the „ , " with „ . " in the whole manuscript.
23
24 4. Pg. 8, L. 22: change to 16S rRNA gene.
25 Was changed to "16S rRNA gene".
26
27 5. Pg. 9, l. 13: you definitely need to capitalize the domain names here
28 We capitalized the domain names as proposed throughout the manuscript.
29
30 6. Pg. 10, l. 17: the units are not correct. It should be "gene cp kg-1" not DNA copies. In
31 addition, you need to clarify if you mean Bacteria and Archaea gene copies. I also would not
32 bother abbreviating copies.
33 We changed the units as proposed and abstain from using "cp" as abbreviation for copies.
34
35 7. Pg. 12, l. 14: add a "." To the end of the sentence.
36 The full stop was already set.
37
38 8. Pg. 12, l. 6-19: the sampling time frame is not clearly described in this section. "over a 7 day
39 period isn't clear" and 24 pairs doesn't match a full year. The only way to know when samples
40 were collected you have to go to the SI. One-2 sentences stating that samples were collected
41 weekly over a 1 year time period by filtering for 7 days would be helpful. State that in some

1 cases samples were not collected because …I presume weather?
2 We accept that the sampling procedure is not clearly described and added the information that
3 the presented samples were chosen out of the whole year filter set in such a way that each season
4 is represented by 5 samples represent. We changed the original text of the section:
5 "Except for filter pairs MZ 11 (24 h) and MZ 15 (5 d), all filter pairs were collecting particles
6 over a 7 day period (Table S1)."
7 to
8 "To get a representative dataset for the whole year, five random samples, consisting of a coarse
9 and fine filter, were analysed for each of the four seasons of the sampling campaign. The
10 sampling period of a single filter pair was generally 7 days except for filter pairs MZ 11 (24 h),
11 MZ 15 (5 d) and MZ 31 (5 d; Table S1)."
12 We hope that the sampling time frames are clarified now. Especially in combination with the
13 changes of the next point dealing with Table S1.
14
15 9. Table S1 also has >24 samples but no information is provided about which filter each sample
16 is. Add a column of what filter is presented. Also explain whether extracts from different filter
17 types were combined. Provide details on which samples were which controls/blanks.
18 Table S1 gives all filter samples analyzed in the current study as well as in Fröhlich-Nowoisky et
19 al (2014). We marked all filter samples which were sequenced with NGS technique additional in
20 the updated version of table S1. Furthermore, we added the blank filters to table S1 to show in
21 which time of the campaign the blanks were taken. The Blank filters have originally been listed
22 in table S2, to clarify this we added "and S2" in the main manuscript. As no extracts were
23 combined we did not mention that in the manuscript.
24
25 10. Pg. 12, l. 25: can you provide some more information on how DNA was extracted? A kit
26 name would be sufficient
27 We added the missing information as proposed. The "MoBio PowerMag Soil DNA Isolation kit"
28 was used.
29
30 11. Pg. 12: l. 26: I don't think this is correct. Illumina did not develop or have instruments that
31 ran pyrosequencing chemistry. Illumina uses "sequencing by synthesis (SBS)" and according to
32 this fact sheet the GAIIx instrument used SBS chemistry: next In addition, only Illumina
33 protocols are listed on the EMP website. Please confirm the type of sequencing chemistry used
34 and use correct language throughout the paper. I usually use "Illumina sequencing" or "next-
35 generation sequencing".
36 We changed it to the term "next-generation-sequencing" and named it to "sequencing by
37 synthesis" as proposed.
38
39 12. Pg. 13, l. 15: change to QIIME
40 Was changed as proposed.
41
42 13. Pg. 16, l. 7: change to "," for 2,342 sequences
43 Was changed as proposed.
44
45 14. Make it clear that the sequences recovered from TSP are just a sum of the coarse and fine
46 samples. State this in the figure legends.

1 As the TSP flow is split into both fractions we added this information into the figure caption of
2 Fig.1 and Fig.2
3
4 15. Pg. 18, l. 14: remove "of"
5 Was changed as proposed.
6
7 16. Pg. 18, l. 26: were the exact same primers used? If not, that is another important factor that
8 could have contributed to differences in the results.
9 We added "and the usage of different primer pairs" to clarify that this might be another factor.
10
11 17. Figure 3: consider putting your data first (upper left most location) and putting "this study"
12 under the location. This way you can have the reader focus more on your findings than previous
13 work.
14 We changed the order like proposed.
15
16 18. Pg. 19, l. 15: if Yooseph et al. 2013 analyzed peptides then you can't present the data as
17 sequences in Figure 3. Revise as needed to present the data accurately.
18 We corrected this with "metagenomic reads" as they used DNA sequences but used a
19 combination of nucleotide and amino acid searches to assign taxonomy to their metagenomic
20 reads.
21
22 19. Pg. 20, l. 20: is euryarchaeotic the correct term?
23 We rephrased to "groups of Euryarchaeota".
24
25 20. Pg. 20, l. 24 & 25: change to livestock
26 We changed "live-stock" to "livestock".
27
28 21. Figure 1: be mindful of whether this figure is legible by those that are color blind. Add
29 "TSP" to the legend.
30 We optimized the figure for red-green color blindness and added (TSP) to the figure caption.
31
32 22. Figure 4: add RFO to the legend.
33 We added "RFO" to the figure caption.
34
35 23. Figure 5: specify in the legend that the data are from this study.
36 We added "within this study" to the figure caption.
37
38 24. Table 1: define the * after "S"
39 We defined it to "S_{Chao1}" to clarify.

Community composition and seasonal changes of ~~archaea~~Archaea in coarse and fine air particulate matter

Jörn Wehking[1,2], Daniel A. Pickersgill[1,2], Robert M. Bowers[3,4], David Teschner[1,2], Ulrich Pöschl[2], Janine Fröhlich-Nowoisky[2], Viviane R. Després[1,2]

[1]Institute of molecular Physiology, Johannes Gutenberg University, Johannes-von-Müller-Weg 6, D-55128 Mainz, Germany

[2]Max Planck Institute for Chemistry, P.O. Box 3060, D-55020 Mainz, Germany

[3]DOE Joint Genome Institute, Walnut Creek, CA, USA

[4]University of Colorado, Boulder, CO, USA.

Correspondence to:

Viviane R. Després (despres@uni-mainz.de)

Jörn Wehking (wehking@uni-mainz.de)

Abstract. ~~Archaea~~Archaea are ubiquitous in terrestrial and marine environments and play an important role in biogeochemical cycles. Although air acts as the primary medium for their dispersal among different habitats, their diversity and abundance is not well characterized. The main reasons for this lack of insight is that ~~archaea~~Archaea are difficult to culture, seem to be low in number in the atmosphere, and have so far been difficult to detect even with molecular genetic approaches. However, to better understand the transport, residence time, and living conditions of microorganisms in the atmosphere as well as their effects on the atmosphere and vice versa, it is essential to study all groups of bioaerosols. Here we present an in-depth analysis of airborne ~~archaea~~Archaea based on Illumina sequencing of 16S ~~-~~rRNA genes from atmospheric coarse and fine particulate matter samples and show seasonal dynamics and discuss anthropogenic influences on the diversity, composition, and abundance of airborne ~~archaea~~Archaea.

The relative proportions of ~~archaea~~Archaea to ~~bacteria~~Bacteria, the differences of the community composition in fine and coarse particulate matter, as well as the high abundance in coarse matter

of one typical soil related family, the Nitrososphaeraceae, points to local phyllosphere and soil habitats as primary emission sources of airborne ~~archaea~~Archaea.

We found comparable seasonal dynamics for the dominating Euryarchaeota classes and Crenarchaeota orders peaking in summer and fall. In contrast, the omnipresent Cenarchaeales and the Thermoplasmata occur only throughout summer and fall. We also gained novel insights into archaeal composition in fine particulate matter (<3 µm), with Cenarchaeaceae, Nitrososphaeraceae, Methanosarcinales, Thermoplasmata and the genus *Nitrosopumilus* as the dominating taxa.

The seasonal dynamics of methanogenic Euryarchaeota points to anthropogenic activities, like fertilization of agricultural fields with biogas substrates or manure, as sources of airborne ~~archaea~~Archaea. This study gains a deeper insight into the abundance and composition of ~~archaea~~Archaea in the atmosphere, especially within the fine particle mode, which adds to a better understanding of the overall atmospheric microbiome.

1 Introduction

Besides ~~bacteria~~Bacteria and E~~e~~ukaryotes, ~~archaea~~Archaea are regarded as a third independent domain of life (Woese et al., 1990). In the beginning of archaeal research in the 1880s primarily methanogenic ~~archaea~~Archaea were discovered and cultivated, so the belief arose that ~~archaea~~Archaea are exclusively extremophiles (Cavicchioli, 2011; Farlow, 1880; Schleper et al., 2005). However, during the last decades, cultivation and culture independent methods, like DNA sequencing, have substantially improved the understanding of ~~archaea~~Archaea and proved that they are also abundant in various environments such as marine, or soil habitats where they can represent more than 10 % of the microbial community (Buckley et al., 1998; Cao et al., 2012; Cavicchioli, 2011; Delong, 1998; Robertson et al., 2005; Yilmaz et al., 2016).

So far, diversity studies for ~~archaea~~Archaea have mainly concentrated on the major habitats also known for ~~bacteria~~Bacteria such as marine and soil environments (Bintrim et al., 1997; Buckley et al., 1998; DeLong, 1992; Ochsenreiter et al., 2003). In the global marine environment the abundance of ~~archaea~~Archaea is approximately 1 x 10^{28} archaeal compared to 3 x 10^{28} bacterial cells (Karner et al., 2001) with ~~archaea~~Archaea accounting for 2-10 % in surface waters and for 20-40 % in deep ocean water (Massana et al., 1997).

The abundance and composition of ~~archaea~~Archaea in soil varies between different soils types (Bates et al., 2011). All cultivated methanogens belong to the kingdom Euryarchaeota and are strictly dependent on anaerobic conditions with low redox potentials (Le Mer and Roger, 2010), thus they are only present in small numbers in many soils. The fertilization with life-stock manure adds anaerobic adapted organisms to the surface of agriculturally used soils. Thus, even in aerated soils, core anaerobic populations seem to survive albeit in low number (Angel et al., 2012). Another issue influencing the abundance and composition of ~~archaea~~Archaea in soil is - as also observed in water columns- the depth (Karner et al., 2001). Analyses of soil depth profiles revealed changing diversity patterns with depth (Bundt et al., 2001; Pesaro and Widmer, 2002) in composition and number.

Next to the well-established major habitats, the atmosphere is another environment in which microorganisms can be detected, however it remains unclear whether the atmosphere can be considered a natural habitat or if it only represents a medium of dispersal for terrestrial and marine microorganisms and their spores (Bowers et al., 2009, 2011, 2012, 2013; Smith et al., 2013; Womack et al., 2010; Yooseph et al., 2013). For airborne ~~bacteria~~Bacteria and ~~archaea~~Archaea the main known emission sources are surface waters and the surface layer of soils (Womack et al., 2010). Therefore, the different abundances and composition of ~~archaea~~Archaea within water and soil columns are of special interest to understand possible emission sources for airborne ~~archaea~~Archaea. For ~~bacteria~~Bacteria, which are abundant in air, the concentration of 16S rRNA gene copies ~~(cp)~~ quantified using qPCR in soil was 10^{11} to 10^{12} ~~DNA~~ gene copies ~~cp~~ kg^{-1} and for Archaea 10^{9} to 10^{11} gene copies~~cp~~ kg^{-1} ~~for archaea~~ (Cao et al., 2012; Kemnitz et al., 2007). In ocean surface waters the concentration is lower but estimated to be 10^{8} to 10^{9} gene copies~~cp~~ L^{-1} for ~~bacteria~~Bacteria and 10^{6} to 10^{7} gene copies~~cp~~ L^{-1} for ~~archaea~~Archaea (Kemnitz et al., 2007; Yin et al., 2013) whereas only 10^{4} to 10^{6} bacterial gene copies~~cp~~ m^{-3} air have been detected (Cao et al., 2012; Fröhlich-Nowoisky et al., 2014; Kemnitz et al., 2007; Yin et al., 2013). Interestingly in contrast to ~~bacteria~~Bacteria, it seems challenging to detect, amplify, and analyze ~~archaea~~Archaea in air, as their concentration of 100 ppm is much lower than the abundance of ~~bacteria~~Bacteria (Cao et al., 2012; Fröhlich-Nowoisky et al., 2014). Until now, it remains unclear whether these observations are biased by technical obstacles or reflect the true abundances. The largest study on airborne ~~archaea~~Archaea is to our knowledge by Fröhlich-Nowoisky et al., (2014) and is based on Sanger sequencing. However, in Fröhlich-

1 Nowoisky et al., (2014) the number of sequences were low, the observations are with little
2 statistical support and the analysis of the microbiome of aerosolized ~~archaea~~Archaea is difficult.
3 Therefore, we present an in-depth ~~pyrosequencing~~ next-generation sequencing study of airborne
4 ~~archaea~~Archaea collected on coarse and fine particulate matter filters over one year in Mainz,
5 Germany. We attempt to compare the composition, diversity, and abundance to the same
6 characteristics as in other habitats, which also allows an inference about the primary emission
7 sources of airborne ~~archaea~~Archaea.

2 Material and Methods

2.1 Aerosol sampling

As described in (Fröhlich-Nowoisky et al., 2009), in total 24 pairs of air filter samples (i.e., 20 filter pairs of one fine and one coarse particle filter sample each, two pairs of start-up air filter blanks and two pairs of mounting filter blanks) were analyzed within this data set. The air filters were installed on a self-built high-volume-dichotomous sampler (Solomon et al., 1983). The whole sampling campaign lasted one year in Mainz, Germany (March 2006 - April 2007). The rotary vane pump (Becker VT 4.25) worked with a flow rate of ~0.3 $m^3 min^{-1}$. The particles were split according to their aerodynamic diameter by a virtual impactor. Particles with an aerodynamic diameter larger than the nominal cut-off of ~3 µm and, due to the sampling device, an additional 10 % of the fine particles were sampled in line with the inlet on one glass fiber filter (flowrate: ~0.03 $m^3 min^{-1}$) representing the coarse fraction. The fine particles were collected on a second glass fiber filter perpendicular to the inlet (~0.27 $m^3 min^{-1}$) which was essentially free from coarse particles (Solomon et al., 1983). ~~Except for filter pairs MZ 11 (24 h) and MZ 15 (5 d), all filter pairs were collecting particles over a 7 day period.~~ To get a representative dataset for the whole year, five random samples, consisting of a coarse and fine filter, were analysed for each of the four seasons of the sampling campaign. The sampling period of a single filter pair was generally 7 days except for filter pairs MZ 11 (24 h), MZ 15 (5 d) and MZ 31 (5 d; Table S1).

The sampled air masses represent a mixture of urban and rural continental air, as the sampler was positioned on the roof of the Max Planck Institute for Chemistry on the campus of the University of Mainz (49°59’31.36’’N, 8°14’15.22’’E). To reduce the sampling of particles emitted from the ground, the sampling device was on a mast about 5 m above the flat roof of the three-story building.

2.2 Extraction, amplification, and sequencing

The DNA extraction and sequencing was part of the Earth Microbiome Project (EMP - http://www.earthmicrobiome.org/) using the MoBio PowerMag Soil DNA Isolation kit and the Illumina GAIIx sequencer with the ~~pyrosequencing~~ sequencing by synthesis technology. As shown before, this

technology is suitable for analyzing microbial communities in soil, water, and human skin (Caporaso et al., 2011).

For the PCR amplifications the 515f/806r primer set (Fwd:GTGCCAGCMGCCGCGGTAA; Rev:GGACTACHVGGGTWTCTAA) described in Caporaso et al., (2011) proved to be suitable, as shown by Bates et al.,(2011). It covers the conserved flanking regions ideal for amplifying ~~bacteria~~Bacteria and ~~archaea~~Archaea over the V4 region of the 16S rRNA gene (Bowers et al., 2013; Huse et al., 2008; Muyzer et al., 1993). In addition, the primer pair is preferred for this amplification as it exhibits only few biases against individual bacterial taxa. As suggested in Caporaso et al., (2011) each DNA extract was amplified in triplicate. These triplicates were combined and purified using a 96 well PCR clean-up kit from MO BIO. The utilized PCR reaction was performed; amplicons purified and sequenced using the GAIIx.

2.3 Grouping of sequences into OTUs and taxonomic identification

The sequences were analyzed using the Quantitative Insight Into Microbial Ecology (QIIME) toolkit (Caporaso et al., 2010). To assign sequences to OTUs, we used ~~Qiime's~~ QIIME's closed reference OTU picking script which uses Uclust (Edgar, 2010) and the greengenes reference database (gg_13_8_otus/rep_set/97_otus.fasta, last update 08/15/2013 ; McDonald et al., 2012) with 97 % similarity. For the actual identification process a corresponding taxonomy map provided by the greengenes database was used. Sequences, which did not match to any greengenes reference set OTU, were discarded for the downstream analysis.

2.4 Controls

Prior to the sampling procedure all filters were baked in sealed aluminum foil bags overnight at 500° C. To best conserve the DNA of the collected bioaerosols, after the sampling procedure all filter samples were stored at -80° C until analysis. To detect possible contaminants from the sampling device and the filter handling, blank filters were taken at 4-week intervals. Contamination free, prebaked filter pairs were mounted in the sampler as for regular sampling, but the pump was not turned on at all ("mounting blanks"). In addition, small environmental samples were taken to collect air exclusively around and

from the interior of the sampling device by turning the pump on for 5 s only (start-up filter blanks). A detailed list of all analyzed air and blank filter samples with their individual sampling details can be found in the supplemental material in Tab. S1 and S2.

The DNA of the blank filters was extracted and quantified in parallel to the actual filter samples. Often, the detected DNA concentrations on blanks can be too small to be quantifed or to build usable sequencing libraries (Cao et al., 2014). However, as the start-up blanks were briefly exposed to environmental air, they also could contain DNA. Within this study we controlled the actual filter changing process by sequencing two mounting blanks, i.e., MZ 23 und MZ 73. Two sequences were obtained from the fine particle filter of MZ 23 and 408 archaeal sequences (371 on the coarse and 37 sequences on the fine particle blank filter) were detected on MZ 73. On the coarse particle filter of MZ 23 no archaeal sequences were detected. Minimal DNA amounts are to be expected, as the mounting blanks were briefly exposed to the air during the mounting process. The sequences on the mounting blanks were assigned to five archaeal families (Cenarchaeaceae, Methanobacteriaceae, Methanoregulaceae, Methanosaetaceae, Methanomassiliicoccaceae). The handling of the sequences obtained with next generation sequencing techniques, e.g., for amplicon sequencing of environmental air samples controls is neither well established nor standardized. To ensure that all contaminants were removed comprehensively from the data set, we decided to omit all identified families from the data if present in more than 1% of all detected archaeal sequences of the mounting blanks.

The subsequently deleted families (from 404 sequences) Methanoregulaceae (8.5%, 3 OTUs), Methanomassiliicoccaceae (17.6%, 3 OTUs) and the largest family of the Methanobacteriaceae (72.4%, 4 OTUs) all belonged to the Euryarchaeota (see also Table S2). In total 2341 sequences remained for the downstream analysis.

Likewise, two pairs of start-up air filter blanks were sequenced. But as they sampled the air for five seconds the obtained sequences were not treated like the mounting blanks. On these four filter samples 709 archaeal sequences were found, distributed with 328 sequences on MZ 22 (326 sequences on coarse, 2 sequences on fine) and 381 sequences on MZ 72 (3 sequences on coarse, 378 sequences on fine).

2.5 Statistical analysis

All data management and most of the analyses were performed using a MySQL database and R-Statistics if not stated otherwise (R-Team, 2011).

To characterize the biodiversity of the ~~archaea~~Archaea community and thus to approximate the likely
5 diversity several statistical parameters were calculated: species richness estimators, rarefaction curves, and community diversity indices using the software tool EstimateS (Colwell et al., 2012)

2.6 Meteorological analysis

As a possible correlation between the abundance of taxonomic ranks in an air mass and meteorological parameters can be either following a monotone or specifically a linear relationship, in this study the
10 Pearson product-moment correlation coefficient (r_K) testing for a linear regression and the Spearman's Rank (r_R) for fine, coarse and total suspended particles. The meteorological parameters tested were: wind speed in m s^{-1} (average and maximum), temperature in °C (range and maximum), relative humidity in %, and the sum of precipitation in mm. The meteorological data were provided in hourly data for wind speed and half hourly values for all other meteorological parameters by the ZIMEN Luft
15 Messnetz of the Landesamt für Umwelt Wasserwirtschaft und Gewerbeaufsicht of Rhineland Palatine. All averages were calculated for the exact sampling periods (Tab. S1). The correlation analysis using the Pearson product-moment correlation coefficient (r_K) and Spearmans Rank (r_R) were calculated for different taxonomic levels, i.e., on kingdom, phylum, and class level. Only results with r_K or r_R over 0.5 or under -0.5 were interpreted. However, no significant correlations between the relative abundance and
20 the meteorological factors were found.

3 Results and Discussion

3.1 Overall Diversity

To determine the archaeal diversity in air, 20 air filter pairs were sampled and analyzed for one year in Mainz, Germany. Each filter pair consists of one filter collecting particles with aerodynamic diameters
25 smaller than 3 µm (fine particulate matter) and one collecting primarily coarse particles, which are

larger than 3 µm. On 39 (97.5 %) of the 40 analyzed filters (20 air filter pairs) archaeal DNA could be detected. In total 2,341 sequences could be assigned to Archaea (Tab. 1). More archaeal sequences were detected on coarse particle filters (109 sequences on average per sample) than on fine particle filters (8 sequences on average per sample) for which the number of sequences ranged from 0 to 42. On all but one fine particle filter, MZ 81 sampled in December 2006, archaeal sequences were discovered. The highest number of sequences, i.e., 601, were detected on the coarse particle filter MZ 74 from November 2006. The 2,341 archaeal sequences were assigned to 52 OTUs. Out of these OTUs, 17 OTUs were found in coarse as well as in fine particulate matter. As listed in Table 1 the coarse particle filters comprised 2180 sequences distributed among 41 OTUs whereas only 161 sequences assigned to 28 OTUs were identified on the fine particle filters.

In total only 7 % of all archaeal sequences stem from fine, whereas 93 % stem from coarse particle filters. Specifically, on 75 % of the coarse particle filters 20 or more archaeal sequences were found, while on 70 % of the fine particle filters less than six archaeal sequences could be detected.

The community structures of both size fractions differ remarkably in composition (Fig. 1). In the fine fraction the genus *Nitrosopumilus* is the dominant taxon. This Thaumarchaeota genus shows a relative abundance of 33.5 % over all Archaea sequences found on all samples in fine particulate matter. The cultivable *Nitrosopumilus maritimus* is a well-known representative of the genus *Nitrosopumilus*. These chemolithoautotrophic nitrifying Archaea have primarily been sampled from marine sources. They form straight rods with a diameter of 0.17–0.22 µm and a length of 0.5–0.9 µm (Könneke et al., 2005) and are thus one of the smallest organism known today. With this size even long distance transport from marine sources might be conceivable. The same can be said for species of Marine group II. However, *N. maritimus* and species of Marine group II have been found in soil samples (Leininger et al., 2006; Treusch et al., 2005). In contrast to the coarse fraction, where only the genera *Methanocella* and the *Candidatus Nitrososphaera* were found with relative proportions of more than 3 %. Due to the much higher number of sequences isolated from the coarse particle fraction in comparison to the fine fraction the TSP composition resembles that of the coarse particle fraction (Fig. 1).

Taking the relative distribution over the entire course of the year into account, on class level the Thaumarchaeota also dominate the fine particle fraction. Except for two fall fine filters where the Euryarchaeota even have a higher relative abundance than the Thaumarchaeota (93 % and 92 %).

The Crenarchaeota, primarily determined by Thaumarchaeota (99 %), are the dominating phylum in the coarse particle mode. Next to Thaumarchaeota a single OTU of the Miscellaneous Crenarchaeotal Group (MCG; Kubo et al., 2012) representing seven sequences was found on a single coarse spring filter sample. No Euryarchaeota were observed on 65 % of the fine and 50 % of the coarse particle filters. A closer look at taxonomic assignments and the contribution of sequences to individual families reveals that most sequences within the coarse particle fraction belong to the Nitrososphaeraceae family. While this family is only present in 10 % of the fine particle filters it was identified on 75 % of the coarse particle filters. In soil surveys the I.1.b group of Crenarchaeota has constantly been found (Ochsenreiter et al., 2003) with the Nitrososphaeraceae being one of the most abundant ~~archaea~~Archaea families therein. Thus, the aerosolization of soil and soil dust as a primary source can be hypothesized for this family. Within this family the genus *Nitrososphaera* is an abundant taxon specifically in agricultural soils (Zhalnina et al., 2013). The landscape of the surrounding area of the sampling location is dominated by agricultural fields and the emitted soil particles are thus, likely to contain the genus *Nitrososphaera*. Soil and soil dust are classically discussed as primary emission sources for airborne ~~bacteria~~Bacteria (Després et al., 2007, 2012; Fierer et al., 2008). Therefore, when attached to large soil particles these organisms should be mainly collected in the coarse particle fraction. To our knowledge, the only cultivated *Nitrososphaera* species, *Nitrososphaera viennensis,* has a much smaller diameter (irregular cocci with a diameter of 0.6–0.9 µm; Stieglmeier et al., 2014), which should be, if in single cell status, collected in the fine particle mode. The hypothesis that soil particles identified through Nitrososphaeraceae are mainly collected on coarse particles is also strengthened by the results of community analysis of the fine particle filters. The observed increase of the relative abundance of the Euryarchaeota could also be interpreted as the decline of Nitrososphaeraceae as soil particles are less frequent in the fine mode. On phylum level the Nitrososphaeraceae family forms the main difference between the two size fractions.

The diversity estimator Chao1 (Tab. 1) and the rarefaction curves (Fig. 2) predict a relatively low diversity for ~~archaea~~Archaea in Mainz air (S_{Chao1}; 64 and 41 for coarse and fine, respectively). On the other hand, the relative abundances of the OTUs and the diversity calculated by Shannon (H) or Simpson (D) (Tab. 1) is slightly higher for the fine particle fraction. This might be because of the small sequence number, but is surely driven by the relative dominance of Nitrososphaeraceae sequences in the coarse particulate matter (Fig. 1).

Most results of this study are in agreement with the previous Sanger-sequencing based study ~~of~~ by Fröhlich-Nowoisky et al. (2014) which analyzed 47 air filter pairs including the 20 filter pairs we focussed on in this study. However, in Fröhlich-Nowoisky et al., (2014), only a limited number of clones were sequenced resulting in a total of 435 sequences, as compared to 2,341 sequences obtained from the current study (Tab.1). Fröhlich-Nowoisky et al., (2014) concluded that ~~archaea~~Archaea occur far more often in coarse than in fine particulate matter as archaeal DNA could only be detected on 21% of the fine particle filters which is consistent with the results of this study. Further consistency is the high abundances of Group I.1.b on coarse particle samples monitored by Fröhlich-Nowoisky et al., (2014), which now can be explained by the higher relative abundance of Nitrososphaeraceae in the coarse fraction.

The main difference between the Sanger and the Illumina approach is the estimated species richness, with 137 species from Sanger estimating almost the double amount than the Illumina approach, which estimates 63 species. This can be caused by several issues. First, a possible lack of taxonomical depth caused by the shorter sequences compared to the Sanger approach and the usage of different primer pairs. Second, by the closed-reference based taxonomic assignment and a possible lack in the used reference dataset. And third, and most likely, by the smaller number of sequences from more filter samples used in Fröhlich-Nowoisky et al., (2014).

As the used primers also amplified bacterial sequences, the following observation could be made: The ratio of ~~archaea~~Archaea and ~~bacteria~~Bacteria suggests a very low proportion of airborne ~~archaea~~Archaea in comparison to airborne ~~bacteria~~Bacteria (Fig. 3). In total, 0.07 % of the total reads could be assigned to ~~archaea~~Archaea, while the rest ($5.7 * 10^6$ reads) consists of bacterial, mitochondrial, and plasmid DNA. After the sequences of mitochondria and plastids are eliminated, still

the ratio of ~~archaea~~Archaea to ~~bacteria~~Bacteria increases only to 0.1 %, which is widely different to the ratios discovered in soil and marine environment.

This extremely low ratio is an interesting phenomenon as in most possible emission sources the proportion of ~~archaea~~Archaea is higher than in air.

Several studies, focusing on airborne ~~bacteria~~Bacteria and ~~archaea~~Archaea found that archaeal DNA in air is extremely low (Cao et al., 2014; Yooseph et al., 2013; Smith et al., 2013). Cao et al. found a proportion of 0.8 % of ~~archaea~~Archaea when compared to ~~bacteria~~Bacteria in PM10 and PM2.5 using Illumina HiSeq data (2014). Yooseph et al., (2013), who analysed the urban prokaryotic metagenome of New York, on a multistep approach based on taxonomic classifications for their ~~peptides~~ metagenomic reads and assigned to the different organism groups, found that 0.48 % of their sequences were archaeal, with roughly 80 % Euryarchaeota and 20 % Crenarchaeota/Thaumarchaeota. Both studies therefore agree with the 0.1 % archaeal sequences found in Mainz air.

Next to comparisons of species diversity and composition, the ratio of ~~bacteria~~Bacteria to ~~archaea~~Archaea might be an indicator of the possible emission sources, as the aerosolization process is likely to equally affect all microorganisms from an emission source. We therefore compared the detected ratios with ratios of possible emission sources like soils, surface water and the phyllosphere reported in literature (Fig. 3).

We found that compared to soil, the microbial habitat, that is often discussed as the primarily emission source, differs strongly from our and other air studies. Although, archaeal abundance in aerated soils increases with depth (Kemnitz et al., 2007), the proportion known for surface soil is still much higher than the proportions in air. Thus, soil alone seems an unlikely emission source. Also in sea water their abundance increases with depth reaching up to 39 % (Karner et al., 2001). As Mainz is not close to oceans emission from sea water seems unlikely as a primary source. The only larger emission surface from water might be the river Rhine which is very likely one of the primary sources in the study area.

In a review by Vorholt (2012) it is convincingly shown that the abundance of ~~archaea~~Archaea in the phyllosphere is less than 1 % of the total microorganism load (Fig. 4), which is similar to the 0.1 % we found. With a total area of 10^9 km^2 of upper and smaller leaf surface, the phyllosphere surface habitat is approximately twice the size of the land surface and is supposed to comprise up to 10^{26} cells worldwide

(Vorholt, 2012), therefore it could present a significant emission source (Woodward and Lomas, 2004) in the studied area. Thus, the phyllosphere might be the local primary emission source.

The situation might, however, differ for individual groups found in the air filter samples, such as the Nitrososphaera family. This family includes typical soil microorganism, which would point to soil as primary emission source. The presence of this family in the air might be on the one hand caused by the diversity of the phyllosphere. Especially for annual plants the microorganism diversity of the phyllosphere is primarily driven by soil and the soil microbiome surrounding the sampling site (Knief et al., 2010). On the other hand, the explanation especially for the findings in the coarse fraction is that larger soil particles carry many typical soil ~~archaea~~Archaea. Thus, based on the proportions of ~~bacteria~~Bacteria and ~~archaea~~Archaea, the most likely interpretation is, that the microbiome detected in Mainz air is primarily originating from the phyllosphere and complemented by small soil particles, which add a large amount of typical soil ~~archaea~~Archaea. Unfortunately, there is a lack of literature on ~~archaea~~Archaea of the phyllosphere, thus the identification of the emission source based on the composition cannot be answered for certain.

Based on the identified genera, however, the phyllosphere and the soil can both be the primary emission source. But as the microbiome of the soil drives the composition of the phyllosphere comparing taxonomy alone will anyway not lead to a final answer.

3.2 Seasonal dynamics

To better understand the seasonal dynamics of ~~archaea~~Archaea in the atmosphere the availability of emission sources over different seasons per year can be analysed. As mentioned, from the 2,341 archaeal sequences 168 could be assigned to Euryarchaeota. By their relative frequencies of occurrence (RFO) Thaumarchaeota are present all year, whereas Euryarchaeota are less abundant and their RFO values show seasonal peaks in spring and fall (Fig. 4).

Although the seasonal increasing or decreasing trends of the RFO values over the year are similar to Fröhlich-Nowoisky et al., (2014) overall, they are higher.

Fröhlich-Nowoisky et al., (2014) suggested the nearby river Rhine as a potential permanent source for Methanomicrobiales and Thermoplasmatales as they are known to be present in river water throughout

the year (Auguet et al., 2009; Cao et al., 2013). The RFO values of the orders shown in Fig. 5 present a slightly different picture: Methanomicrobia were observed in every season with RFO values around 40 %, thus the Rhine could contribute continuously to the aerosolized Methanomicrobia. However, the Thermoplasmata group was exclusively found in summer and fall samples, arguing against an emission from an omnipresent source like the Rhine.

Alternatively to the Rhine, potential emission sources for several ~~euryarchaeotic~~ groups of Euryarchaeota -especially in agricultural areas as around Mainz - are biogas substrates and livestock ~~life stock~~ fertilization methods (Fröhlich-Nowoisky et al., 2014). Figure 5 shows that Methanomicrobia and Methanobacteria both have their highest relative RFO during fall and another increase during the springs in 2006 and 2007. This supports the hypothesis, of livestock ~~life-stock~~ manure being a possible emission source, as both classes are commonly known to be present in the microbiome of livestock ~~live-stock~~ and the typical times for fertilization of fields with manure is in spring and fall (Nicol et al., 2003; Radl et al., 2007). Like all methanogen groups, they have been reported in biogas reactors, too (Jaenicke et al., 2011). For the Thermoplasmata the peaks in summer and fall might be linked to the usage of biogas reactor substrates, which are also applied to agricultural fields as fertilizer. The differing RFO values of Thermoplasmata and other Euryarchaeota might be caused by their sensitivity to temperature and especially to pH, which only allows their survival in moderate to high temperatures and low pH environments.

The hypothesis that aerosolized ~~archaea~~Archaea are linked to agricultural activities is also supported by the seasonal variation of RFO of the order of the Nitrososphaerales within the Thaumarchaeota that is also present in the Euryarchaeota classes as discussed. Nitrososphaerales were found in agricultural soil samples close to the sampling area of our study (Ochsenreiter et al., 2003; Zhalnina et al., 2013), and thus can be considered a typical agricultural soil microorganism.

4 Conclusion

This study gains a deeper insight into the diversity of airborne ~~archaea~~Archaea. The overall abundance of ~~archaea~~Archaea in the atmosphere compared to ~~bacteria~~Bacteria is very low, comparable to the ratio found for the phyllosphere. We found the Nitrososphaeraceae family out of the I.1.b group of

Crenarchaeota to be the major archaeal family in course particulate matter. The groups Cenarchaeaceae, Nitrososphaeraceae, Methanosarcinales, Thermoplasmata and the genus *Nitrosopumilus* could be observed within the fine particulate matter.

The observed seasonal dynamics for the dominating Euryarchaeota classes and Crenarchaeota orders, which peak in summer and fall, might be a result of agriculture in the surrounding area. So anthropogenic activities like fertilization with livestock manure or substrates of biogas reactors might influence the diversity of airborne ~~archaea~~Archaea as their occurrence is increased during the main fertilization seasons.

This combination of findings provides support for the conceptual premise that the occurrence of ~~archaea~~Archaea in air might be driven by the microbiota of the phyllosphere but the influence of livestock manure gains an edge over the phyllosphere through the fertilization seasons. Additionally, groups emitted with soil as carrier particles seem to have a major influence on the community composition. For a further understanding of the dependencies of airborne microorganisms on their sources, future studies should additionally explore possible source habitats to gain as complete pictures as possible.

We conclude that the understanding of the seasonality, diversity, and composition of airborne ~~archaea~~Archaea as one very small fraction within the bioaerosols is an important contribution to understand the patterns driving the whole atmospheric microbiome.

5 Data availability

The post-library-split sequence dataset will be made available from the edmond digital repository http://edmond.mpdl.mpg.de/imeji/.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors are grateful to M. O. Andreae, T. Andreae, J. Cimbal, R. Conrad, P. E. Galand, M. Grant, M. Klose, I. Müller-Germann, C. Osterhof, C. Pio, C. Ruzene Nespoli, B. Schmitt, and H. Yang for technical assistance, as well as H. Paulsen, R. Stantscheff and Z. Jia for helpful discussions. We also thank the Mainz Bioaerosol Laboratory (MBAL) for support. This work was supported by the Max Planck Society.

J. Wehking acknowledges the PhD grant of the Konrad-Adenauer-Stiftung.

References

Angel, R., Soares, M. I. M., Ungar, E. D. and Gillor, O.: Biogeography of soil archaea and bacteria along a steep precipitation gradient, ISME J., 4(4), 553–563, doi:10.1038/ismej.2009.136, 2010.

Angel, R., Claus, P. and Conrad, R.: Methanogenic archaea are globally ubiquitous in aerated soils and become active under wet anoxic conditions, ISME J, 6(4), 847–862, doi:10.1038/ismej.2011.141, 2012.

Auguet, J.-C., Barberan, A. and Casamayor, E. O.: Global ecological patterns in uncultured Archaea, Isme J, 4(2), 182–190, doi:http://www.nature.com/ismej/journal/v4/n2/suppinfo/ismej2009109s1.html, 2009.

Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R. and Fierer, N.: Examining the global distribution of dominant archaeal populations in soil, ISME J, 5, 908–917, doi:10.1038/ismej.2010.171, 2011.

Bintrim, S. B., Donohue, T. J., Handelsman, J., Roberts, G. P. and Goodman, R. M.: Molecular phylogeny of Archaea from soil, P. Natl Acad. Sci. USA., 94, 277–282, 1997.

Bowers, R. M., Lauber, C. L., Wiedinmyer, C., Hamady, M., Hallar, A. G., Fall, R., Knight, R. and Fierer, N.: Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei., Appl. Environ. Microbiol., 75(15), 5121–30, doi:10.1128/AEM.00447-09, 2009.

Bowers, R. M., McLetchie, S., Knight, R. and Fierer, N.: Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments., ISME J., 5(4), 601–12, doi:10.1038/ismej.2010.167, 2011.

Bowers, R. M., McCubbin, I. B., Hallar, A. G. and Fierer, N.: Seasonal variability in airborne bacterial
5 communities at a high-elevation site, Atmos. Environ., 50, 41–49, doi:10.1016/j.atmosenv.2012.01.005, 2012.

Bowers, R. M., Clements, N., Emerson, J. B., Wiedinmyer, C., Hannigan, M. P. and Fierer, N.: Seasonal variability in bacterial and fungal diversity of the near-surface atmosphere, Environ. Sci. Technol., 47, 12097–12106, 2013.

10 Buckley, D. H., Graber, J. R. and Schmidt, T. M.: Phylogenetic analysis of nonthermophilic members of the kingdom Crenarchaeota and their diversity and abundance in soils, Appl. Environ. Microbiol., 64(11), 4333–4339, 1998.

Bundt, M., Widmer, F., Pesaro, M., Zeyer, J. and Blaser, P.: Preferential flow paths: Biological “hot spots” in soils, Soil Biol. Biochem., 33(6), 729–738, doi:10.1016/S0038-0717(00)00218-2, 2001.

15 Cao, C., Jiang, W., Wang, B., Fang, J., Lang, J., Tian, G., Jiang, J. and Zhu, T. F.: Inhalable microorganisms in Beijing’s PM2.5 and PM10 pollutants during a severe smog event, Environ. Sci. Technol., 48(3), 1499–1507, 2014.

Cao, H., Auguet, J.-C. and Gu, J.-D.: Global Ecological Pattern of Ammonia-Oxidizing Archaea, edited by S. Bertilsson, PLoS One, 8(2), 1–8, doi:10.1371/journal.pone.0052853, 2013.

20 Cao, P., Zhang, L. M., Shen, J. P., Zheng, Y. M., Di, H. J. and He, J. Z.: Distribution and diversity of archaeal communities in selected Chinese soils, FEMS Microbiol. Ecol., 80(1), 146–158, doi:10.1111/j.1574-6941.2011.01280.x, 2012.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley,
25 R. E., Lozupone, C. A., Mcdonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R.: QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing, Nat. Publ. Gr., 7(5), 335–336, doi:10.1038/nmeth0510-335, 2010.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N. and Knight, R.: Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample., Proc. Natl. Acad. Sci. U. S. A., 108 Suppl, 4516–22, doi:10.1073/pnas.1000080107, 2011.

Cavicchioli, R.: Archaea — timeline of the third domain, Nat. Rev. Microbiol., 9(1), 51–61, doi:10.1038/nrmicro2482, 2011.

Chen Cao, Wenjun Jiang, Buying Wang, Jianhuo Fang, Jidong Lang, Geng Tian, et al: Inhalable Microorganisms in Beijing 's PM 2.5 and PM 10 Pollutants during a Severe Smog Event, Environ. Sci. Technol., 48(3), 1499–1507 [online] Available from: http://pubs.acs.org/doi/pdf/10.1021/es4048472 (Accessed 16 April 2015), 2014.

Colwell, R. K., Chao, A., Gotelli, N. J., Lin, S. Y., Mao, C. X., Chazdon, R. L. and Longino, J. T.: Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages, J. Plant Ecol., 5(1), 3–21, doi:10.1093/jpe/rtr044, 2012.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C. and Vorholt, J. : Community proteogenomics reveals insights into the physiology of phyllosphere bacteria., Proc. Natl. Acad. Sci. U. S. A., 106(38), 16428–16433, doi:10.1073/pnas.0905240106, 2009.

Delong, E. F.: Everything in moderation: Archaea as "non-extremophiles," Curr. Opin. Genet. Dev., 8(6), 649–654, doi:10.1016/S0959-437X(98)80032-4, 1998.

DeLong, E. F.: Archaea in coastal marine environments., P. Natl. Acad. Sci. USA., 89(12), 5685–5689, 1992.

Després, V. R., Nowoisky, J. F., Klose, M., Conrad, R., Andreae, M. O. and Pöschl, U.: Characterization of primary biogenic aerosol particles in urban, rural, and high-alpine air by DNA sequence and restriction fragment analysis of ribosomal RNA genes, Biogeosciences, 4, 1127–1141 [online] Available from: http://hal.archives-ouvertes.fr/hal-00297660/ (Accessed 12 June 2013), 2007.

Després, V. R., Alex Huffman, J., Burrows, S. M., Hoose, C., Safatov, A. S., Buryak, G., Fröhlich-Nowoisky, J., Elbert, W., Andreae, M. O., Pöschl, U. and Jaenicke, R.: Primary biological aerosol particles in the atmosphere: a review, Tellus B, 64, doi:10.3402/tellusb.v64i0.15598, 2012.

Edgar, R. C.: Search and clustering orders of magnitude faster than BLAST, Bioninformatics, 26(19), 2460–2461, 2010.

Farlow, W.: On the Nature of the Peculiar Reddening of Salted Codfish during the Summer Season, United States Comm. Fish Fish. Part IV Rep. Comm. 1878., (XLIV), 969–974, 1880.

Fierer, N., Liu, Z., Rodríguez-Hernández, M., Knight, R., Henn, M. and Hernandez, M. T.: Short-term temporal variability in airborne bacterial and fungal populations, Appl. Environ. Microb., 74(1), 200–207, doi:10.1128/AEM.01467-07, 2008.

Fisher, R., Corbet, S. and Williams, C. B.: The relation between the number of species and the number of individuals in a random sample of an animal population, J. Anim. Ecol., 12(1), 42–58, doi:10.2307/1411, 1943.

Fröhlich-Nowoisky, J., Pickersgill, D. A, Després, V. R. and Pöschl, U.: High diversity of fungi in air particulate matter., Proc. Natl. Acad. Sci. U. S. A., 106(31), 12814–9, doi:10.1073/pnas.0811003106, 2009.

Fröhlich-Nowoisky, J., Ruzene Nespoli, C., Pickersgill, D. A., Galand, P. E., Müller-Germann, I., Nunes, T., Gomes Cardoso, J., Almeida, S. M., Pio, C., Andreae, M. O., Conrad, R., Pöschl, U. and Després, V. R.: Diversity and seasonal dynamics of airborne archaea, Biogeosciences, 11(21), 6067–6079, doi:10.5194/bg-11-6067-2014, 2014.

Huse, S. M., Dethlefsen, L., Huber, J. A., Welch, D. M., Relman, D. and Sogin, M. L.: Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing, PLoS Genet., 4(11), doi:10.1371/journal.pgen.1000255, 2008.

Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Dröge, M., Gartemann, K. H., Jünemann, S., Kaiser, O., Krause, L., Tille, F., Zakrzewski, M., Pühler, A., Schlüter, A. and Goesmann, A.: Comparative and joint analysis of two metagenomic datasets from a biogas fermenter obtained by 454-pyrosequencing, PLoS One, 6(1), doi:10.1371/journal.pone.0014519, 2011.

Karner, M. B., DeLong, E. F. and Karl, D. M.: Archaeal dominance in the mesopelagic zone of the Pacific Ocean., Nature, 409(6819), 507–510, doi:10.1038/35054051, 2001a.

Kemnitz, D., Kolb, S. and Conrad, R.: High abundance of Crenarchaeota in a temperate acidic forest soil., FEMS Microbiol. Ecol., 60(3), 442–8, doi:10.1111/j.1574-6941.2007.00310.x, 2007.

Knief, C., Ramette, A., Frances, L., Alonso-Blanco, C. and Vorholt, J. A.: Site and plant species are important determinants of the Methylobacterium community composition in the plant phyllosphere., ISME J., 4(6), 719–728, doi:10.1038/ismej.2010.9, 2010.

Knief, C., Delmotte, N., Chaffron, S., Stark, M., Innerebner, G., Wassmann, R., von Mering, C. and
5 Vorholt, J. : Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice, ISME J., 6(7), 1378–1390, doi:10.1038/ismej.2011.192, 2012.

Könneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B. and Stahl, D. a: Isolation of an autotrophic ammonia-oxidizing marine archaeon., Nature, 437(7058), 543–6, doi:10.1038/nature03911, 2005.

10 Kubo, K., Lloyd, K. G., F Biddle, J., Amann, R., Teske, A. and Knittel, K.: Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments, ISME J., 6(10), 1949–1965, doi:10.1038/ismej.2012.37, 2012.

Leininger, S., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., Prosser, J. I., Schuster, S. C. and Schleper, C.: Archaea predominate among ammonia-oxidizing prokaryotes in soils, Nature, 442, 806–
15 809, 2006.

Lueders, T. and Friedrich, M.: Archaeal Population Dynamics during Sequential Reduction Processes in Rice Field Soil Archaeal Population Dynamics during Sequential Reduction Processes in Rice Field Soil, Appl. Environ. Microbiol., 66(7), 2732–2742, doi:10.1128/AEM.66.7.2732-2742.2000.Updated, 2000.

20 Massana, R., Murray, E., Preston, C. M. and DeLong, E. F.: Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel, Appl. Environ. Microb., 63(1), 50–56, 1997.

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., Andersen, G. L., Knight, R. and Hugenholtz, P.: An improved Greengenes taxonomy with explicit ranks for ecological
25 and evolutionary analyses of bacteria and archaea, ISME J., 6(3), 610–618, doi:10.1038/ismej.2011.139, 2012.

Le Mer, J. and Roger, P.: Production, oxidation, emission and consumption of methane by soils: A review, Eur. J. Soil Biol., 37(1), 25–50, 2010.

Muyzer, G., De Waal, E. C. and Uitterlinden, A. G.: Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA, Appl. Environ. Microbiol., 59(3), 695–700, 1993.

Nicol, G. W., Glover, L. A. and Prosser, J. I.: Molecular analysis of methanogenic archaeal
5 communities in managed and natural upland pasture soils, Glob. Chang. Biol., 9(10), 1451–1457, doi:10.1046/j.1365-2486.2003.00673.x, 2003.

Nicol, G. W., Leininger, S., Schleper, C. and Prosser, J. I.: The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria, Environ. Microbiol., 10(11), 2966–2978, doi:10.1111/j.1462-2920.2008.01701.x, 2008.

10 Ochsenreiter, T., Selezi, D., Quaiser, A., Bonch-Osmolovskaya, L. and Schleper, C.: Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR, Environ. Microbiol., 5(9), 787–797, doi:10.1046/j.1462-2920.2003.00476.x, 2003.

Pesaro, M. and Widmer, F.: Identification of novel Crenarchaeota and Euryarchaeota clusters associated with different depth layers of a forest soil, FEMS Microbiol. Ecol., 42, 89–98, doi:10.1016/S0168-
15 6496(02)00302-1, 2002.

Pickersgill, D.: Design and Implementation of a Primary Bioaerosol DNA Database., 2013.

Radl, V., Gattinger, A., Chronáková, A., Nemcová, A., Cuhel, J., Simek, M., Munch, J. C., Schloter, M. and Elhottová, D.: Effects of cattle husbandry on abundance and activity of methanogenic archaea in upland soils., ISME J., 1, 443–452, doi:10.1038/ismej.2007.60, 2007.

20 Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y. and Fierer, N.: The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves, Environ. Microbiol., 12(11), 2885–2893, doi:10.1111/j.1462-2920.2010.02258.x, 2010.

Robertson, C. E., Harris, J. K., Spear, J. R. and Pace, N. R.: Phylogenetic diversity and ecology of environmental Archaea, Curr. Opin. Microbiol., 8(6), 638–642, doi:10.1016/j.mib.2005.10.003, 2005.

25 Schleper, C., Jurgens, G. and Jonuscheit, M.: Genomic studies of uncultivated archaea, Nat. Rev. Microbiol., 3(6), 479–488, doi:10.1038/nrmicro1159, 2005.

Smith, D., J. Timonen, D. Jaffe, D. Griffin, M. Birmele and M. Roberts. 2013. Intercontinental dispersal of bacteria and archaea by transpacific winds. Applied and Environmental Microbiology. 79(4):1134-1139

Solomon, P. A., Moyers, J. L. and Fletcher, R. A.: High-Volume dichotomous virtual impactor for the fractionation and collection of particles according to aerodynamic size, Aerosol Sci. Technol., 2(4), 455–464, 1983.

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

Stieglmeier, M., Klingl, A., Alves, R. J. E., Rittmann, S. K. M. R., Melcher, M., Leisch, N. and Schleper, C.: Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota, Int. J. Syst. Evol. Microbiol., 64(PART 8), 2738–2752, doi:10.1099/ijs.0.063172-0, 2014.

Suda, W., Nagasaki, A. and Shishido, M.: Powdery Mildew-Infection Changes Bacterial Community Composition in the Phyllosphere, Microbes Environ., 24(3), 217–223, doi:10.1264/jsme2.ME09114, 2009.

Team, R. D. C.: R: a language and environment for statistical computing, R Foundation for Statistical Computing. Available from: http://www.r-project.org/. 2011.

Treusch, A. H., Leininger, S., Kletzin, A., Schuster, S. C., Klenk, H.-P. and Schleper, C.: Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling, Environ. Microbiol., 7(12), 1985–1995, doi:10.1111/j.1462-2920.2005.00906.x, 2005.

UEDA, T. SUGA, Y. T. M.: Molecular phylogenetic analysis of a soil microbial community in a soybean field, Eur. J. Soil Sci., 46(3), 415–421, 1995.

Vorholt, J. A: Microbial life in the phyllosphere., Nat. Rev. Microbiol., 10(12), 828–40, doi:10.1038/nrmicro2910, 2012.

Woese, C. R., Kandler, O. and Wheelis, M. L.: Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya, P. Natl. Acad. Sci. USA. 87, 4576–4579, 1990.

Womack, A. M., Bohannan, B. J. M. and Green, J. L.: Biodiversity and biogeography of the atmosphere., Phil.Trans.R.Soc.B, 365(1558), 3645–3653, doi:10.1098/rstb.2010.0283, 2010.

Woodward, F. I. and Lomas, M. R.: Vegetation dynamics--simulating responses to climatic change., Biol. Rev. Camb. Philos. Soc., 79, 643–670, doi:10.1017/s1464793103006419, 2004.

5 Yilmaz, P., Yarza, P., Rapp, J. Z. and Glöckner, F. O.: Expanding the World of Marine Bacterial and Archaeal Clades, Front. Microbiol., 6 (January), 1524, doi:10.3389/fmicb.2015.01524, 2016.

Yin, Q., Fu, B., Li, B., Shi, X., Inagaki, F. and Zhang, X.-H.: Spatial Variations in Microbial Community Composition in Surface Seawater from the Ultra-Oligotrophic Center to Rim of the South Pacific Gyre, PLoS One, 8(2), e55148, doi:10.1371/journal.pone.0055148, 2013.

10 Yooseph, S., Andrews-Pfannkoch, C., Tenney, A., McQuaid, J., Williamson, S., Thiagarajan, M., Brami, D., Zeigler-Allen, L., Hoffman, J., Goll, J. B., Fadrosh, D., Glass, J., Adams, M. D., Friedman, R. and Venter, J. C.: A metagenomic framework for the study of airborne microbial communities., PLoS One, 8(12), e81862, doi:10.1371/journal.pone.0081862, 2013.

Zhalnina, K., De Quadros, P. D., Gano, K. A., Davis-Richardson, A., Fagen, J. R., Brown, C. T.,
15 Giongo, A., Drew, J. C., Sayavedra-Soto, L. A., Arp, D. J., Camargo, F. A. O., Daroub, S. H., Clark, I. M., McGrath, S. P., Hirsch, P. R. and Triplett, E. W.: Ca. nitrososphaera and bradyrhizobium are inversely correlated and related to agricultural practices in long-term field experiments, Front. Microbiol., 4(MAY), 1–13, doi:10.3389/fmicb.2013.00104, 2013.

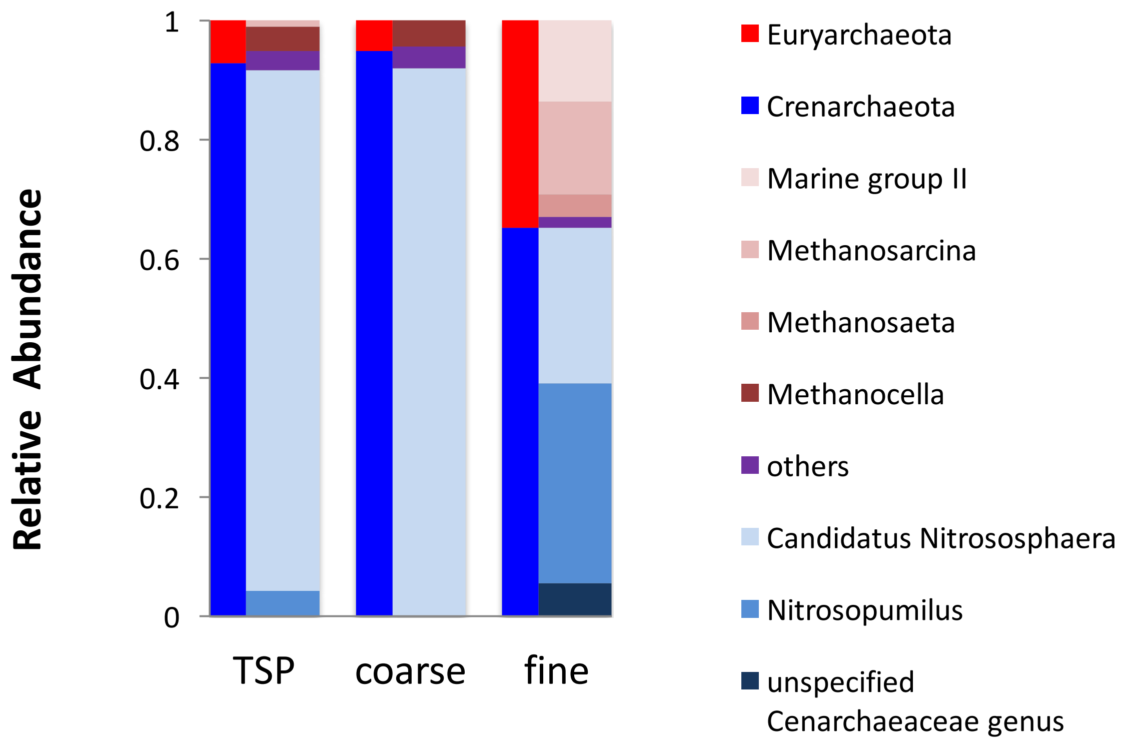


Figure 1: Archaeal community composition for total suspended (TSP), the same airstream split into coarse, and fine particulate matter on the level of phyla (red/blue) and genera (pastel colors).

5

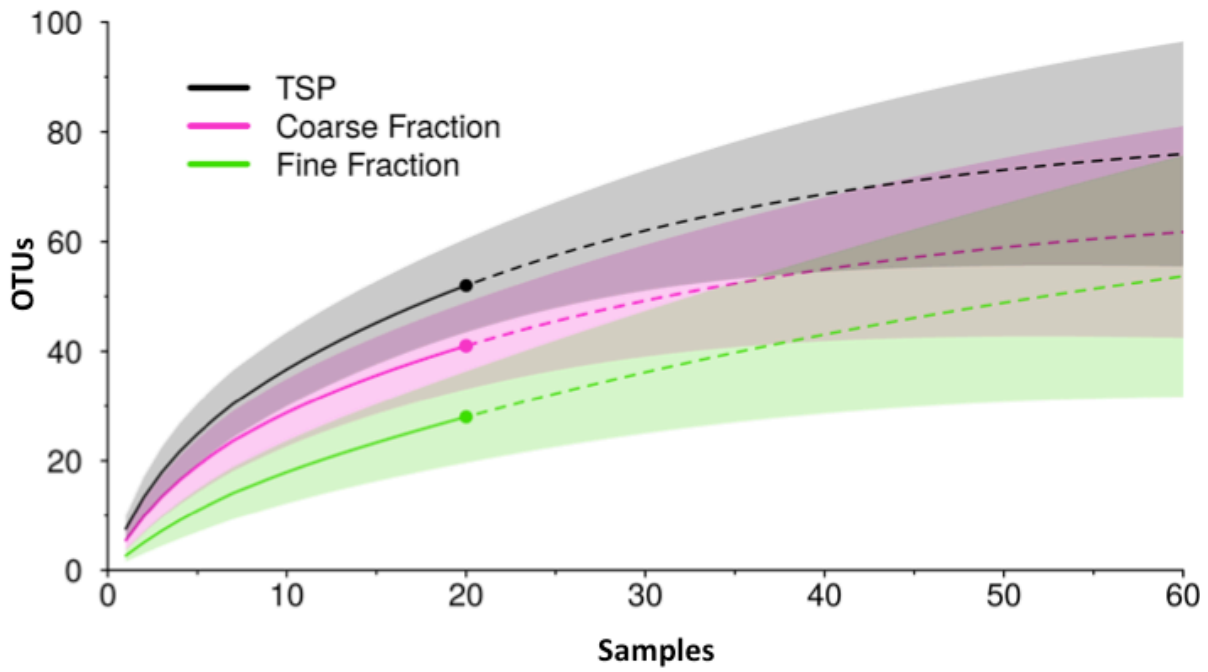
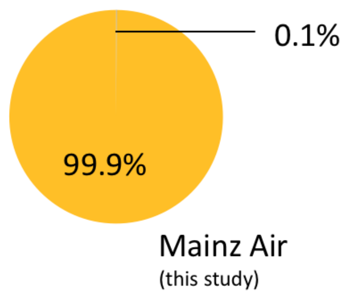
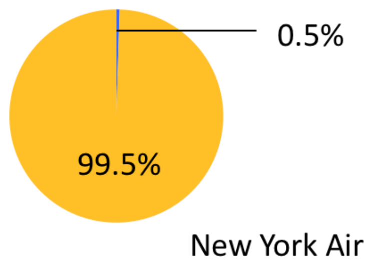
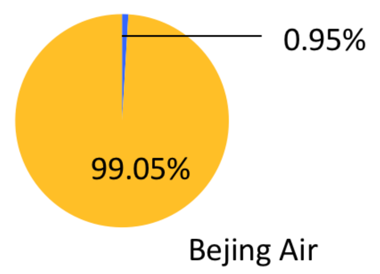
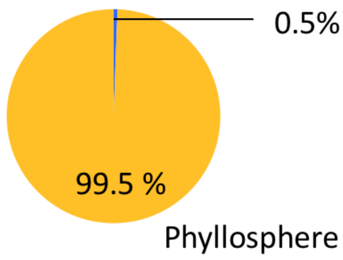


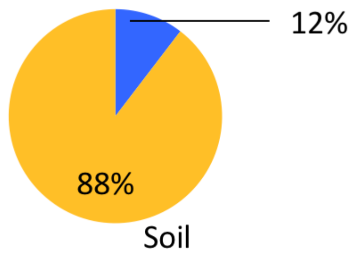
Figure 2: Rarefaction curve of species richness for total suspended, the same airstream split into coarse, and fine particulate matter. The solid curves represent the interpolated number of OTUs as a function against the number of samples. The dashed lines are according extrapolations and the dot marks the sample size of this study. The colored areas represent the 95 confidence intervals.

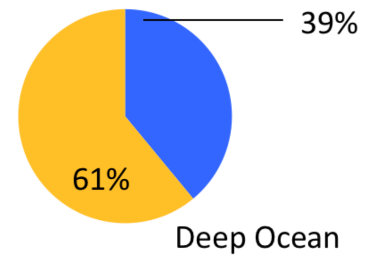












Bacteria Sequences Archaea Sequences

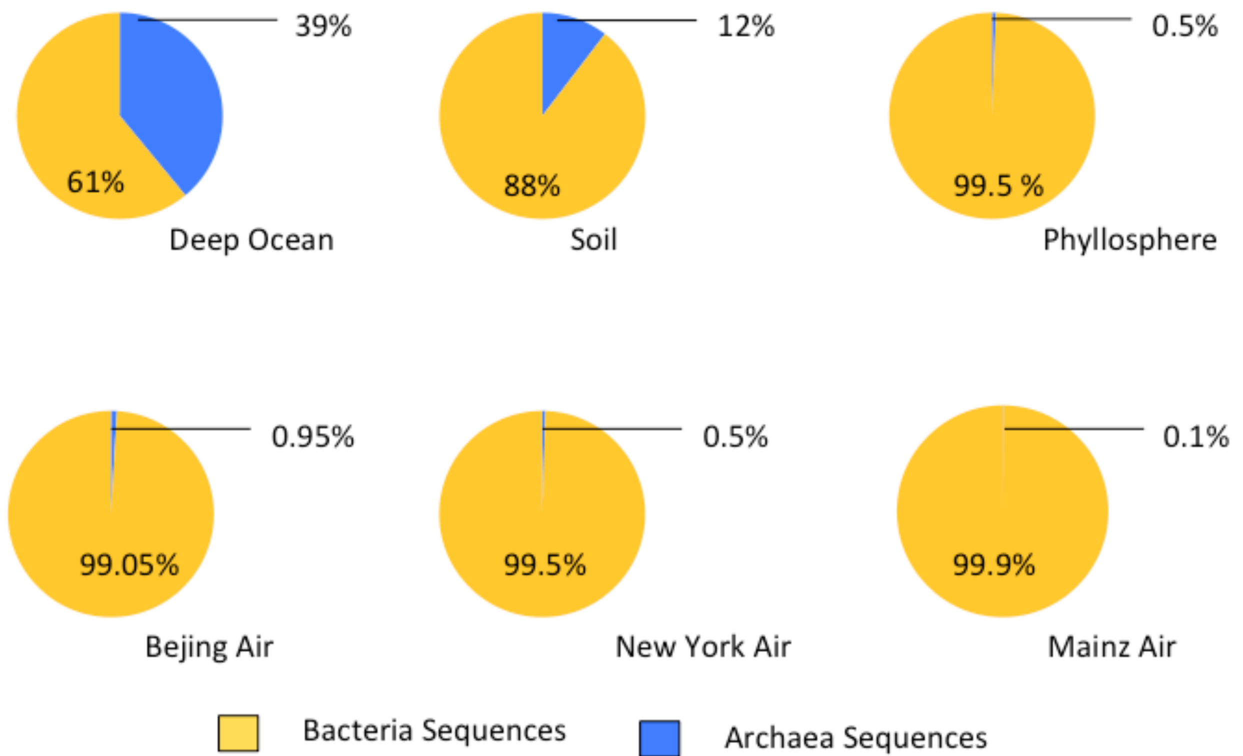


Figure 3: Relative proportions of archaeal (blue) and bacterial (yellow) sequences detected in environmental samples. Proportions for soil are based on Kemnitz et al., (2007), for the deep ocean on Karner et al., (2001), and for the phyllosphere on Delmotte et al., (2009) and Knief et al., (2012). The proportions of the Mainz air are based on this study. The data for the New York air are published in Yooseph et al., (2013) and the data of Bejing are based on Cao et al., (2014).

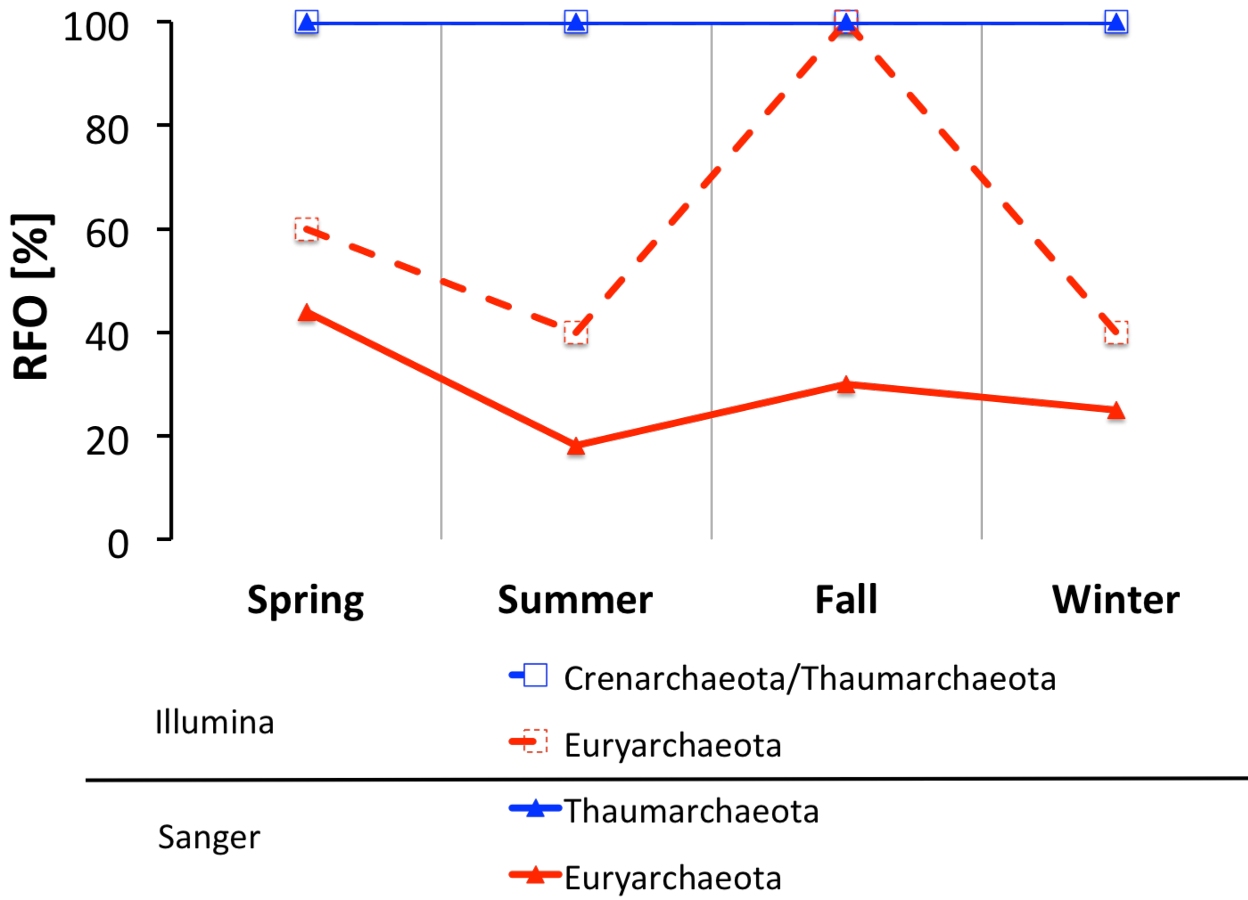


Figure 4: Seasonal variation in the relative frequency of occurrence (RFO) of airborne ~~archaea~~Archaea on phylum level. The relative frequency of occurrence – the proportion of samples in which these taxa were detected - is given for both phyla, i.e., Thaumarchaeota and Euryarchaeota. The graph based on Sanger sequencing represents the data published in Fröhlich-Nowoisky
5 **et al., (2014), whereas the remaining data comprises the results of this study.**

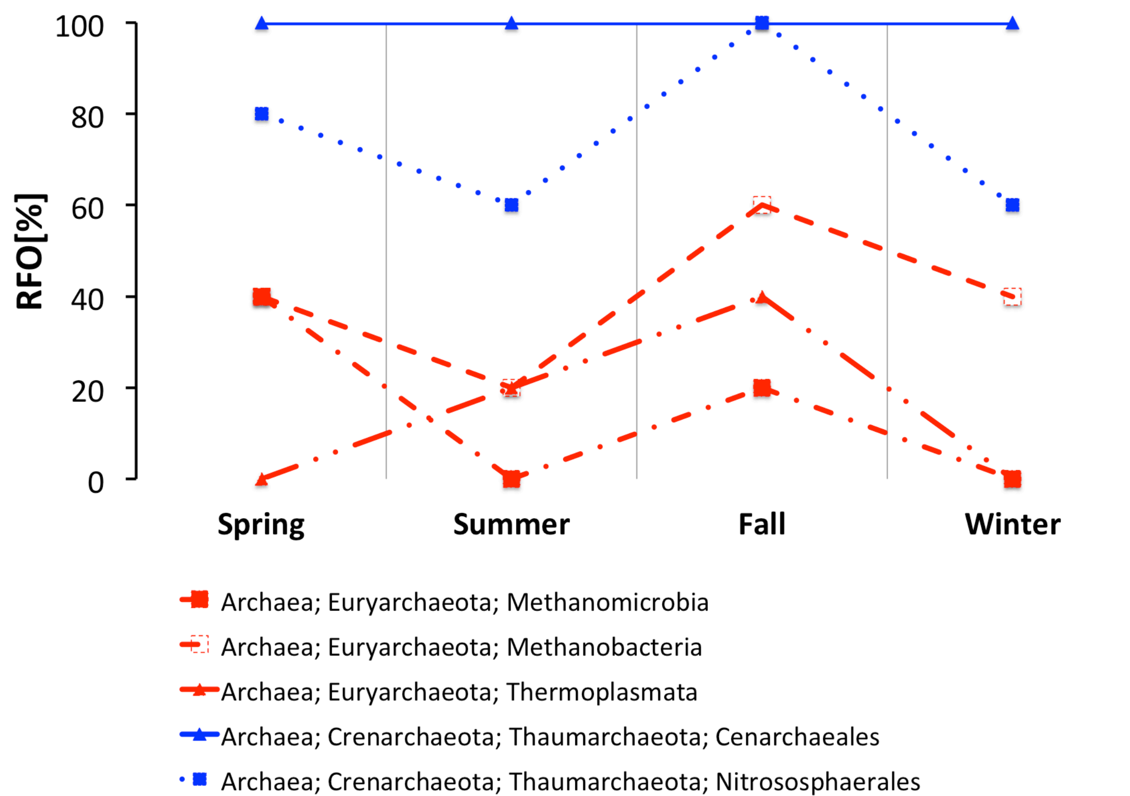


Figure 5: Seasonal variation in the relative frequency of occurrence of dominating Euryarchaeota classes and Crenarchaeota orders within this study.

5 **Table 1. Number of sequences and indices estimating the archaeal diversity in Mainz for coarse and fine particle filter samples and total suspended particles (TSP).**

Size Fraction	n (Samples)	Sq (Sequences)	Sq/n	OTU (Operational taxonomic unit)	S_{Chao1}* (Chao1)	H (Shannon)	D (Simpson)
Coarse	20	2180	109	41	64	3.09	0.83
Fine	20	161	8.1	28	41	3.65	0.88
TSP	20	2341	117.1	52	63	3.36	0.84
Fröhlich-Nowoisky et al., (2014)	47	435	9.3	57	137	3.32	0.82

Supplementary material.

Table S1: List of air filter samples analysed for A~~a~~rchaea presence. While the air masses collected on MZ 11, MZ 15, and MZ 31 were sampled over 1, 5, and 5 days, respectively, all other air filter samples were taken over a 7-day period corresponding to ~ 3000 m^3 of sampled air. Air masses analysed and sequenced with Illumina within this study are bold. For comparison the samples analysed and published in Fröhlich-Nowoisky et al (2014) were included, n.a. = not available as these air masses were not sequenced with Illumina but exclusively via Sanger sequencing.~~The sequences gained with the Sanger sequencing method are analysed and published in Fröhlich-Nowoisky et al (2014), n.a. = not available as these air masses were not sequenced with Illumina but exclusively via Sanger sequencing.~~

Sample ID	Sampling period	Number of sequences 16S Sanger (coarse, fine)	Number of sequences 16S Illumina (coarse, fine)
MZ 1	24.03.2006 - 31.03.2006	5	n.a.
MZ 2	31.03.2006 - 07.04.2006	8	n.a.
MZ 4	07.04.2006 - 12.04.2006	11	n.a.
MZ 6	15.04.2006 - 18.04.2006	4	n.a.
MZ 9	20.04.2006 - 27.04.2006	5	n.a.
MZ 10	27.04.2006 - 02.05.2006	5	n.a.
MZ 11	**02.05.2006 - 03.05.2006**	**7**	**3 - ~~,~~80**
MZ 15	**04.05.2006 - 09.05.2006**	**9 - ~~,~~ 6**	**195 - ~~,~~81**
MZ 18	12.05.2006 - 15.05.2006	7	n.a.
MZ 19	15.05.2006 - 16.05.2006	7	n.a.
MZ 21	17.05.2006 - 18.05.2006	4	n.a.
MZ23 blank	**18.05.2006**		**0 - 2**

MZ 24	18.05.2006 - 22.05.2006	13	n.a.
MZ 25	22.05.2006 - 23.05.2006	12	n.a.
MZ 26	**23.05.2006 - 30.05.2006**	**12 -; 6**	**301 -;103**
MZ 31	**01.06.2006 - 06.06.2006**	**8**	**208 -;5**
MZ 33	08.06.2006 - 13.06.2006	10	n.a.
MZ 35	14.06.2006 - 21.06.2006	5 -; 1	n.a.
MZ 36	21.06.2006 - 22.06.2006	10	n.a.
MZ 40	27.06.2006 - 04.07.2006	9	n.a.
MZ 41	**04.07.2006 - 11.07.2006**	**5**	**19 -;284**
MZ 45	19.07.2006 - 21.07.2006	14, 4	n.a.
MZ 47	**26.07.2006 - 02.08.2006**	**5**	**32 -;23**
MZ 50	**02.08.2006 - 09.08.2006**	**9**	**196 -;132**
MZ 51	09.08.2006 - 16.08.2006	5	n.a.
MZ 52	**16.08.2006 - 23.08.2006**	**9**	**234 -;2**
MZ 54	**30.08.2006 - 06.09.2006**	**4 -; 15**	**183 -;574**
MZ 59	**11.09.2006 - 18.09.2006**	**11**	**597 -;3**
MZ 60	18.09.2006 - 25.09.2006	5	n.a.
MZ 62	**02.10.2006 - 09.10.2006**	**6**	**318 -;351**
MZ 63	09.10.2006 - 16.10.2006	12	n.a.

MZ 66	16.10.2006 - 23.10.2006	4	n.a.
MZ 67	**23.10.2006 - 30.10.2006**	**12**	**29 - 321**
MZ 69	02.11.2006 - 09.11.2006	20	n.a.
MZ 71	16.11.2006 - 23.11.2006	1	n.a.
MZ73 blank	**23.11.2006**		**371 - 37**
MZ 74	**23.11.2006 - 30.11.2006**	**8**	**797 - 2**
MZ 75	30.11.2006 - 07.12.2006	11 - 7	n.a.
MZ 77	14.12.2006 - 21.12.2006	5	n.a.
MZ 81	**28.12.2006 - 04.01.2007**	**5**	**5 - 0**
MZ 82	**04.01.2007 - 11.01.2007**	**8**	**202 - 21**
MZ 84	18.01.2007 - 25.01.2007	5 - 4	n.a.
MZ 88	**01.02.2007 - 08.02.2007**	**3**	**93 - 2**
MZ 90	**15.02.2007 - 22.02.2007**	**9**	**122 - 485**
MZ 93	**22.02.2007 - 01.03.2007**	**15 - 6**	**597 - 2**
MZ 95	08.03.2007 - 15.03.2007	9	n.a.
MZ 97	22.03.2007 - 29.03.2007	5	n.a.
MZ 101	**05.04.2007 - 12.04.2007**	**11 - 4**	**36 - 4**
MZ 103	**19.04.2007 – 26.04.2007**	**10 - 4**	**109 - 5**
	∑ (+contaminants)		**4285 - 2507**

∑ (-contaminants)	**2180 - 161**

Table S2: Air filter blank samples analyzed for archaeal contamination. All families found on mounting blank filters comprise together 410 sequences: families were discarded from the data if present in more than 1% of all detected archaeal sequences on the mounting blanks, i.e., the Methanoregulaceae (8.54%), Methanomassiliicoccaceae (17.56%), and the Methanobacteriaceae (72.44%).

	MZ 23		**MZ 73**			
	coarse	**fine**	**coarse**	**fine**	**∑**	**%**
Cenarchaeaceae	0	0	1	1	2	0.49
Methanobacteriaceae	0	0	297	0	297	72.44
Methanoregulaceae	0	0	0	35	35	8.54
Methanosaetaceae	0	2	1	1	4	0.98
Methanomassiliicoccaceae	0	0	72	0	72	17.56