Diversity and seasonal dynamics of airborne Archaea

Janine Fröhlich-Nowoisky[1], Cristina Ruzene Nespoli[1], Daniel A. Pickersgill[1, 2], Pierre E. Galand[3,4], Isabell Müller-Germann[1,5], Teresa Nunes[6], João Gomes Cardoso[6, 7], S. Marta Almeida[8] , Casimiro Pio[6], Meinrat O. Andreae[1], Ralf Conrad[9], Ulrich Pöschl[1], and Viviane R. Després[2]

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

[2]Institute of General Botany, Johannes Gutenberg University, Johannes-von-Müller-Weg 6, D-55128 Mainz, Germany

[3]UPMC, Univ Paris 06, Observatoire Océanologique, Banyuls-sur-Mer, France

[4]CNRS, UMR8222, Laboratoire d'écogéochimie des environnements benthiques, Observatoire Océanologique, Banyuls-sur-Mer, France

[5]Geosciences, Johannes Gutenberg University, Joh.-Joachim- Becher-Weg 21, D-55128 Mainz, Germany

[6]CESAM & Department of Environment, University of Aveiro, 3810-193 Aveiro, Portugal

[7]University of Cape Verde, 279 CPraia, Santiago, Cabo Verde

[8]ITN/MCTES - 2686-953 Sacavém, Portugal

[9]Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Straße 10, 35043 Marburg, Germany

Correspondence to: V. R. Després (despres@uni-mainz.de) & J. Fröhlich-Nowoisky (j.frohlich@mpic.de)

Abstract

Archaea are widespread and abundant in many terrestrial and aquatic environments, thus outside extreme environments, accounting for up to ~10% of the prokaryotes. Compared to Bacteria and other microorganisms, however, very little is known about the abundance, diversity, and dispersal of Archaea in the atmosphere. By DNA analysis and Sanger Sequencing targeting the 16S rRNA (435 sequences) and *amoA* genes in samples of air particulate matter collected over one year at a continental sampling site in Germany, we obtained first insights into the seasonal dynamics of airborne Archaea. The detected Archaea were identified as *Thaumarchaeota* or *Euryarchaeota*, with soil *Thaumarchaeota* (group *I.1b*) being present in all samples. The normalized species richness of *Thaumarchaeota* correlated positively with relative humidity and negatively with temperature. This together with an increase of bare agricultural soil surfaces may explain the diversity peaks observed in fall and winter. The detected *Euryarchaeota* were mainly predicted methanogens with a low relative frequency of occurrence. A slight increase in their frequency during spring may be linked to fertilization processes in the surrounding agricultural fields. Comparison with samples from the Cape Verde islands (72 sequences) and from other coastal and continental sites indicates that the proportions of *Euryarchaeota* are enhanced in coastal air, which is consistent with their suggested abundance in marine surface waters. We conclude that air transport may play an important role for the dispersal of Archaea, including assumed ammonia-oxidizing *Thaumarchaeota* and methanogens.

1 Introduction

Archaea have long been thought to occur only in restricted, extreme environments. However, since their formal recognition as an independent domain of life about two decades ago (Woese et al., 1990), they have been found in a wide variety of habitats from hydrothermal vents to aquatic and soil environments (Auguet et al., 2010; Bintrim et al., 1997; Boetius et al., 2000; Jurgens et al., 2000; Ochsenreiter et al., 2003; Schleper et al., 1997; Takai et al., 2001). Archaea are now known to be abundant, diverse, and widespread organisms that act as major players in the nitrogen and carbon cycle (Offre et al., 2013; Pester et al., 2012; Schleper et al., 2005).

Despite the recognition of the ubiquitous nature of Archaea, their possible presence in the atmosphere is poorly investigated. While the atmosphere may not be a key habitat of Archaea, it may act as the primary medium for the dispersal of microorganisms among the different ecosystems on the Earth's surface (Fröhlich-Nowoisky et al., 2012). For other bioaerosols such as Bacteria and Fungi, studies have been conducted that describe the composition, diversity, abundance, seasonal, and even daily variation of these organisms in the atmosphere, and explore their distribution and biogeography (e.g., Després et al., 2012; Fröhlich-Nowoisky et al., 2012 and references therein). These questions are still unanswered for Archaea.

The few published reports on airborne Archaea are from anthropogenic environments such as compost piles or biosolids (e.g., Baertsch et al., 2007; Moletta et al., 2007; Thummes et al., 2007) and a handful of natural and urban environments (Bowers et al., 2013; Brodie et al., 2007; Cao et al., 2014; Cho and Hwang, 2011; Radosevich et al., 2002; Robertson et al., 2013; Smith et al., 2013; Yooseph et al., 2013). The extent of these studies remains very limited and the number of Archaea sequences presented in the literature is small. Thus, the first insights into the presence of Archaea in the atmosphere could draw only very preliminary conclusions about the composition from single sites and specific sampling times. In particular, nothing is known about which metabolic groups of Archaea are present in the air.

Archaea have various types of metabolisms, some of which are involved in key biogeochemical processes like the production of methane, the assimilation of amino acids, or the oxidation of ammonia (Cavicchioli, 2011; Pester et al., 2011). Nitrification is a central process of the global nitrogen cycle that can be divided into two steps: the oxidation of ammonia to nitrite and the oxidation of nitrite to nitrate. Ammonia oxidizing Bacteria and Archaea (*AOB* and *AOA*, respectively) use the enzyme ammonia monooxygenase (encoded by the *amoA* gene,

Falkowski et al., 2008; Gruber and Galloway, 2008; Kowalchuk and Stephen, 2001) to perform the conversion to nitrite, which is the rate-limiting step. Since the discovery of the importance of the *amoA* gene, it has been widely used to explore the diversity and abundance of *AOA*, even though the presence of the gene does not necessarily prove the potential for ammonia oxidation (Pester et al., 2011). Thus, *amoA* gene based studies have allowed a better understanding of the ecological and evolutionary factors shaping the community assembly in *AOA* and the factors that drive community distribution (Auguet et al., 2010; Cao et al., 2011; Cavicchioli, 2011).

The main purpose of this study was first to determine whether airborne Archaea are present all year round, and if they display seasonal dynamics. Furthermore, we wanted to test the hypothesis of whether the composition of airborne Archaea depends on the type of ecosystem found below the air masses. For the first purpose we gathered samples for one year in continental boundary layer air at one sampling site and identified archaeal communities by sequencing the 16S rRNA and the *amoA* gene. For the second objective, we compared the results of the continental air site with discrete sampling from other geographical regions, including both coastal and continental air, to understand the composition of airborne Archaea under both a local and a more global perspective.

2 Material and Methods

2.1 Aerosol sampling

As it is well established that decontamination of sampling devices is a necessary prerequisite for reliable molecular genetic analysis (Després et al., 2007), the quartz and glass fiber filters used were baked at high temperatures to eliminate possible biological traces. To secure a continuous quality assurance, we regularly cleaned the samplers with ethanol to prevent colony growth in the interiors and checked the air flow rates weekly. For some sampling sites (e.g., Germany), we separated the particles into coarse and fine mode, while for other sites or campaigns (e.g., North America) total suspended particulate matter was analyzed. As within the sampling strategies small variations occur between the different sampling sites, the specific details are given per site. With the main analyses focusing on the samples from Germany and Cape Verde only the detailed sampling strategies for these sites are presented here, while the information for the other sites is available in the supplementary material. However, for an overall comparison the main sampling information is listed in Tables 1 and S1.

2.1.1 Germany

Over one year, 47 pairs of fine and coarse particle samples were collected in Mainz, Germany (March 2006 - April 2007). A self-built high-volume-dichotomous sampler (Solomon et al., 1983) was used, which was operated with a rotary vane pump (Becker VT 4.25) at a total flow rate of ~0.3 m^3 min^{-1}, corresponding to a nominal cut-off diameter of ~3 µm. Thus, coarse particles with an aerodynamic diameter larger than ~3 µm were collected on one glass fiber filter (~0.03 m^3 min^{-1}), while the fine particles from the same air sample were collected on a second glass fiber filter (~0.27 m^3 min^{-1}). The sampling period was generally ~7 days, corresponding to a sampled air volume of 3000 m^3. A few samples were collected over shorter periods (volumes of ~400-2000 m^3). The sampling station was positioned on a mast about 5 m above the flat roof of the three-story high Max Planck Institute for Chemistry building (MPIC) on the campus of the University of Mainz (49°59'31.36'' N, 8°14'15.22'' E). The sampled air masses represent a mix of urban and rural continental boundary layer air in central Europe. All glass fiber filters were baked overnight at 500 °C prior to sampling and the loaded filters sealed in decontaminated aluminum foil and stored at -80 °C until analysis.

To detect possible contaminants from the sampler and sample handling, blank samples were taken at 4-week intervals. Prebaked filters were mounted in the sampler like for regular sampling, but the pump was either not turned on at all ("mounting blanks") or only for 5 s ("start-up blank").

2.1.2 Cape Verde

Filter sampling was part of the Cape Verde Dust project (CV-Dust). The sampling station was positioned ~1500 m from the coast at 69 m elevation, latitude 14°55'33.96'' N, longitude 23°29'40.92'' W. For the 26 Cape Verde samples (Table 1 and Table S1), Whatman Quartz filters were used and sampling was done between February and March 2011 with a PM10 High-Volume sampler (Tisch PM10, TE-6001) placed 14 m above ground level and with a flow rate of 1 m^3 min^{-1}. Sampling periods ranged from less than 24 h to 3 days. Prior to sampling, filters were decontaminated by baking for 6 h at 500 °C and after sampling were stored in decontaminated bags and kept frozen at -80 °C until DNA extraction. The air at this site is primarily coastal with influence from continental air masses, which may carry particles from Saharan or Sahel dust events.

2.2 DNA extraction and amplification

Filter sample aliquots (about 1-25% of the filter) were extracted with a commercial soil DNA extraction kit (LysingMatrixE, FastDNASpin Kit for Soil, MP Biomedicals) according to the supplier's instructions with the following modifications (depending on the character of the filter): centrifugation (10 - 15 min) after the lysis, followed by addition of 900 µl buffer provided by the extraction kit and a second round of beating and centrifugation. The supernatants of both extraction steps were combined for the further extraction process. Finally, the DNA was dissolved in 100 µl elution buffer. Decontaminated filters were included during the extractions as extraction blanks. Extraction blanks without filters were prepared as well.

For the laboratory work, regular quality control and assurance measurements were taken. All PCRs were set up with sterile equipment in a DNA free environment. As it was difficult to amplify airborne Archaea (see also supplementary material) we used 3 primer pairs as well as nested primers to increase the amplification success. Thus, with the DNA extract from each of the filters up to four polymerase chain reactions (PCR) were performed using normal and nested primer pairs targeting the 16S rRNA gene. The number of PCR was dependent on the successful amplification. To detect even minimal amounts of archaeal DNA, an additional PCR and eventually nested PCRs were only performed in case the first PCR failed. The 25 - 50 µl reaction mixture contained the template DNA (0.5 - 1 µl sample extract), 1 x PCR buffer, 0.2 mM each dNTP (Roth), 0.33 µM each primer (Sigma Aldrich), 2 mM $MgCl_2$ and 2.5 units of JumpStartTM REDTaq DNA polymerase. A negative control was included in all PCR runs.

The thermal profile of the 16S PCR reactions was as follows: initial denaturing at 94 °C for 3 min; 34 cycles with 94 °C for 30 s, annealing at 55 °C for 1 min for primer pair A2F/958r (Baker and Cowan, 2004; Wani et al., 2006) or at 52 °C for 20 s for primer pair 109f/934r (Grosskopf et al., 1998), elongation at 72 °C for 1.5 min, and a final extension step at 72 °C for 5 min. Nested PCRs were done with primer pair 109f/934r and primer pair 109f/927r (Baker and Cowan, 2004; Grosskopf et al., 1998), respectively. The thermal profile for both pairs was: initial denaturing at 94 °C for 2 min, 29 cycles with 94 °C for 30 s, annealing at 52 °C for 1 min, elongation at 72 °C for 1 min and a final extension at 72 °C for 10 min. Within one extraction blank in the samples of Cape Verde a contamination occurred. The PCR product was cloned and sequenced and the sequences grouped into OTUs, thus treated as all other PCR products. The OTUs occurring in the extraction blank PCR were completely removed from the analysis thus also when they occurred in any other filer sample.

An *amoA* analysis was performed for 18 coarse filters from Mainz, Germany (Table S1). For the archaeal *amoA* gene the primer pair amo19F/amo643R with the nested primer pair amo111F/amo643R was used (Leininger et al., 2006; Treusch et al., 2005). The PCR mix was as described above. The PCR reaction consisted of an initial denaturation at 94 °C for 4 min, 9 cycles at 94 °C for 45 s, annealing at 55 - 65 °C for 1 min, elongation at 72 °C for 1 min and 19 cycles at 94 °C for 45 s, 55 °C for 1 min and 72 °C for 1 min, and a final extension step at 72 °C for 10 min.

Amplification products of each filter sample selected for sequencing were cloned using the TOPO TA Cloning Kit (Invitrogen) following the supplier's instructions. Colonies containing inserts were identified by blue-white selection and lysed in 20 µl water for 10 min at 95 °C. The cloning efficiency was between 5 and 200 colonies per cloned PCR product. The inserts of 12 - 24 randomly picked colonies were amplified for each PCR-product using 1.5 - 3 µl lysate in a 25 - 40 µl reaction. In case fewer colonies were available all colonies were amplified and sequenced.

2.3 Sequence and phylogenetic analysis

DNA sequences were determined with ABI Prism 377, 3100, and 3730 sequencers (Applied Biosystems) using BigDye-Terminator v3.1 chemistry at the Max Planck Genome Centre, Cologne. To assure quality we studied the individual chromatograms for background noise, and cut the vector sequences. Sequences were screened for possible chimera using the Bellerophon program (Huber et al., 2004). The alignment was conducted using the multiple sequence comparison by log expectation (MUSCLE) package (Edgar, 2004) and thereafter manually checked. The taxonomic identification was only based on sequence data. The sequences from each sampling site were clustered into operational taxonomic units (OTUs), which represent artificial species. For the 16S rRNA sequence the similarity of sequences within a species is thought to be at least 97% (Stackebrandt and Goebel, 1994). As during the cloning process the PCR products can be inserted into the vector sequence in both directions and the sequencing was conducted from one direction, the reads sometimes cover different sections of the analyzed sequence or they only have a few base pairs of overlap. Thus, for the identification of OTUs, sequences were grouped together according to their sequence coverage to provide the longest possible length for an automatic analysis with the Mothur software package (Schloss et al., 2009). Sequence blocks were cut to the longest possible length to retain as many sequences as possible. For Germany this was ~370 bp, for North America ~470 bp, for China ~600 bp and for the UK ~760 bp. Sequences that were too short for the chosen

208 blocks were compared manually to the other sequences to determine whether they fitted into
209 one of the automatically identified OTUs, or formed a new one. Operational taxonomic units
210 and their reference sequences were determined with the Mothur package. For the sequences
211 from Cape Verde, OTUs were created by using the USEARCH v5.2.3.2 package (Edgar,
212 2010) on the 97% discrimination level and manual BLAST procedures.

213 For the phylogenetic identification of the OTUs two approaches were used. Firstly, a broad
214 taxonomic identification by comparing the sequences of interest with the sequences in well-
215 known databases and, secondly, a more detailed analysis by calculating phylogenetic trees.
216 The results from the comparison against databases were used to identify potentially closely
217 related Archaea strains. From these strains we used then type species to build a basic tree for
218 the phylogenetic tree analysis. The basic tree sequences are listed in Table S2.

219 For these comparisons, with known sequences, both the Basic Local Alignment Search Tool
220 (BLAST), via the website of the National Center for Biotechnology Information (NCBI,
221 http://www.ncbi.nlm.nih.gov/), as well as the SINA aligner (Pruesse et al., 2012), which in-
222 cludes several microbial databases, were used.

223 For the phylogenetic analysis, a basic tree was constructed that mainly included sequences
224 from cultured, identified type species and a few from uncultured species (Table S2). For iden-
225 tification the representative sequences of each OTU were then included. For OTUs where the
226 representative sequences did not overlap the area used for the tree calculation, the best match
227 sequences were used as placeholders, but only when the similarity between the sequence and
228 its best match was higher than 97%. The model of evolution best suitable for the sequence
229 sets was chosen using the "find best DNA/Protein Model-option" in Mega5.2.2 (Tamura et
230 al., 2011). Phylogenetic trees were constructed by the maximum likelihood algorithm using
231 the substitution model Tamura-3-parameter model with gamma distributed rates for the analy-
232 sis of the 16S sequences and 1000 replicates for the boot strap values. Additionally, we also
233 tested Neighbor-Joining algorithm which agreed with maximum likelihood results.

234 The *amoA* gene DNA sequences were likewise aligned with MUSCLE. Operational taxonom-
235 ic units were formed when the sequence similarity scores were >85% according to Pester et
236 al. (2012). Phylogenetic trees were built on the nucleotide level based on Tamura-3-parameter
237 model as described above and on the amino acid level using the WAG+G+F model (Whelan
238 and Goldman, 2001).

The sequences of the present study have been deposited in the GenBank database under the following accession numbers: For 16S rRNA genes KF683446 – KF683516 (Cape Verde), KF558331 - KF558337 (China), JQ249390 - JQ249751 (Germany), JQ976019 - JQ976037 (North America), KF558321 - KF558330 (United Kingdom) and for the *amoA* gene sequences (KF824904-KF824908).

2.4 Quantitative PCR assays

All quantitative PCR (qPCR) reactions were performed on an iCycler thermocycler equipped with a MyiQ detection system (Bio-Rad, Munich, Germany) and the data were analyzed using the iQ3 Optical System software (Bio-Rad). As a standard for Archaea, a pure culture of *Methanosarcina thermophila* was used (Lueders et al., 2004) and for the Bacteria, chromosomal DNA of *Escherichia coli K12* (Stubner, 2002) was used. Both standards were serially diluted and used for the construction of calibration curves in each reaction. For the standards duplicates and for the filter extracts triplicates were analyzed.

For the actual quantification process the following primer pairs were chosen: for Archaea 364f (5'-CGG GG(CT) GCA (GC)CA GGC GCG AA-3') and 934r (5'-GTG CTC CCC CGC CAA TTC CT-3') and for Bacteria 519f (5´-CAG C(AC)G CCG CGG TAA (AGCT) (AT)C-3´) and 907f/926r (5´-CCG TCA ATT CMT TTR AGT TT-3´) as commonly used (Burggraf et al., 1997; Grosskopf et al., 1998; Lane, 1991). The qPCR assays were based on SYBR Green (Ambion). Each reaction had a volume of 25 µl and the following mixture for Archaea (conditions for Bacteria in brackets): 12.5 µl SYBR Green JumpStart Taq ReadyMix, 1.5 (4) mM $MgCl_2$, 0.66 (0.5) µM of each primer, and 5 µl of the template DNA and 1 µM FITC. For the assay, the program used was: 94 °C for one cycle for 6 (8) min, followed by 45 (50) cycles at 94 °C for 35 s (20 s), 66 °C (50 °C) for 30 s (20 s), 72 °C for 45 s (50 s) and 85.5 °C (75 °C) for 6 s for signal reading.

2.5 Statistical analysis

Although the number of the here analyzed Sanger sequences is low compared to modern NGS techniques, we performed statistical analyses to get a first indication of species richness and possible correlations with meteorological conditions. However, species richness measurements were only performed for the sites with most sequences (Germany and Cape Verde). Correlation analyses with meteorological conditions were only performed for Germany, where most sequences were analyzed.

To assess whether there were significant correlations between the diversity of sampled Archaea and the meteorological factors, the relative species richness per sample j ($S_{R,j}$), as well as the relative species richness per month k ($S_{R,k}$) were calculated (see Table S3). The statistical analysis was performed using the R-Package (R-Development-Core-Team, 2005). Pearson's linear regression coefficients were calculated between the relative species richness and the sampling period or monthly averages for temperature, relative humidity and wind speed and sum of precipitation per sample, to assess the influence of the single factors on the sampled diversity.

To examine if a combination of meteorological factors correlates best with the sampled diversity, the stepwise AIC Method (Akaike Information Criterion) for variable selection (Yamashita et al., 2007) was conducted. A more detailed linear regression analysis (Neter et al., 1996) was then performed on the best results of the Pearson and stepwise AIC test. A more detailed description of the analysis can be found in the supplement material.

3 Results and Discussion

3.1 Abundance and diversity of airborne Archaea in continental boundary layer air

From the 47 air filters sampled at Mainz, Germany (Table S1) we retrieved a total of 435 archaeal 16S rRNA gene sequences that grouped into 57 OTUs. The aerosols at this site were sampled using size selection with a cut-point at 3 µm. Archaea sequences could be amplified in all coarse (>3 µm), but only in 21% of the fine particle filter samples (<3 µm). To better understand this phenomenon we quantified copies (cp) of Archaea and Bacteria 16S rRNA genes in a subset of six air filter samples. We found the abundance of Archaea in air to be only between ~1 to ~10 gene copies per m^3 air, while that of Bacteria was found to be ~10^4 to ~10^6 in the same air samples (Table S4). These results are strikingly different from observations in soil and ocean surface water, where Archaea represent about 10% of the total prokaryotic abundance (Bates et al., 2011; Cao et al., 2012; DeLong, 1992; Karner et al., 2001; Kemnitz et al., 2007; Yin et al., 2013). One of the explanations might be that Archaea could be more susceptible to UV light or other damage caused by, e.g., components from the fine aerosol than Bacteria. However, the fact that Archaea were more numerous in coarse particle filters may indicate that Archaea are often attached to soil dust particles as it also has been shown for bacteria (Jones and Harrison, 2004). The observation that Archaea are more diffi-

cult to amplify than other airborne microbes has been stated in other publications as well (Bowers et al., 2009, 2013; Fierer et al., 2008; Woo et al., 2013; Woo, 2012).

This study gives first insights that airborne Archaea were present all year long in air from the continental boundary layer. Bioaerosols in air have a lifetime of about a week (Burrows et al., 2009), implying that the aerosols sampled may have traveled several thousand kilometers. Therefore, this study's data represent not only a snapshot of the actual diversity, but also include information about possible sources and distribution patterns.

All archaeal 16S rRNA gene OTUs at the German sampling site belonged either to *Euryarchaeota* (14%) or *Thaumarchaeota* (86%, see also Figs. 1a and S1, Table S5). *Thaumarchaeota* are widespread on Earth (Pester et al., 2011) and are detected in various environments such as freshwater, ocean, sediments, and hot springs, with an especially high abundance in soil (Bates et al., 2011; Brochier-Armanet et al., 2008; Brochier-Armanet et al., 2011; Leininger et al., 2006; Ochsenreiter et al., 2003; Schleper and Nicol, 2010). Based on phylogenetic analysis, all OTUs of the *Thaumarchaeota* were affiliated with the 16S rRNA sequences from the soil group *I.1b* (Fig. S1, Table S5). Group *I.1b* Archaea were composed of a high number of related phylotypes (Fig. S1) similar to what is observed in soils (Auguet et al., 2010).

In addition to the 16S rRNA gene, the phylogenetic relationships of the *amoA* gene, which is used as a reliable marker to study the diversity of archaeal ammonia oxidizers was also analyzed (Cao et al., 2011, 2013; Junier et al., 2010). The *amoA* gene was successfully detected on a subset of 18 coarse particulate matter samples, spanning all four seasons (Table S1). The 176 archaeal *amoA* sequences grouped into 5 OTUs. Thus, the diversity of archaeal *amoA* genes was lower than that of archaeal 16S rRNA genes. Strengthening the results from the analysis of archaeal 16S rRNA OTUs, all *amoA* OTUs belonged to the genus *Nitrososphaera* affiliated with the group *I.1b* of soil Archaea identified using the nucleotide (Fig. 2) as well as the amino acid sequences. Both findings support observations made for airborne Bacteria showing that airborne microbial diversity is closely linked to soil diversity, and that soil and soil dust might be a primary source of airborne microorganisms (Brodie et al., 2007; Després et al., 2007; Després et al., 2012; Fierer et al., 2008; Jones and Harrison, 2004; Lighthart and Shaffer, 1995; Prospero et al., 2005).

In comparison to *Thaumarchaeota* the number of OTUs of *Euryarchaeota* was low. Euryarchaeotal 16S rRNA gene sequences were present in 32% of the samples and attributed to

5 well-known orders. *Halobacteriales*, *Thermoplasmatales*, *Methanomicrobiales* and *Methanosarcinales* were represented by a single OTU each, while 50% of the OTUs belonged to the *Methanobacteriales* (Figs. 1a, S1 and Table S5). The presence of predicted methanogens may be surprising, as Ochsenreiter et al. (2003) suggested that oxic soils support only the *Thaumarchaeota* group *I.1b*. However, a recent study by Angel et al. (2012) proved that methanogenic *Euryarchaeota*, despite being obligate anaerobic, are present in oxic soils and regain their methanogenesis activity when deprived of oxygen. Other studies finding methanogens in soil attributed their presence to the effect of livestock introducing methanogens into the soil with their feces (Gattinger et al., 2007; Nicol et al., 2003; Radl et al., 2007). Similarly to *Thaumarchaeota*, *Euryarchaeota* in air are likely to originate from aerosolized soil or other emission sources.

Thaumarchaeota had a normalized species richness (S_n, number of detected OTUs divided by number of investigated samples, Table S3) of ca. 2 to 3 throughout the year (Fig. 3a) with OTU *Arch1* being the most frequent OTU present in 70% of the samples. In contrast, the S_n of *Euryarchaeota* was always below one, and the two most frequent OTUs (*Arch12* and *Arch13*) were detected in only 11% of the investigated air samples. Based on the Chao-1 index (Table S3), we obtained an estimate of ~137 OTUs for the actual species richness of Archaea in the investigated air samples (see also supplementary material for detailed discussion). This species richness is only about 50% of that reported by Brodie and colleagues (2007) for two urban sites in North America. However, in urban air, the composition and abundance of aerosol microbes seems to be unique. Bacteria, for instance, are often released from strong point sources (Després et al., 2012; Fang et al., 2007; Shaffer and Lighthart, 1997) and this may also be true for Archaea. While the number of investigated samples and DNA amplification products used in this study certainly underestimates the actual diversity of Archaea in air, Fröhlich-Nowoisky and colleagues (2009) detected 368 fungal species in a slightly smaller but otherwise identical data set..

3.2 Seasonal and meteorological derived dynamics of airborne Archaea

Seasonal variation within the species richness of *Thaumarchaeota* shows highest diversity in the fall/ winter period is illustrated in Fig. 3a. The sampling site in Germany is situated in an area with extensive agriculture. After the harvest of particular grains or rapeseed at the end of summer, fields either remain bare over fall and winter or are prepared, e.g., by plowing and tilling of the upper 10 - 30 cm for seeding. Soil turnover might influence the number of *Thaumarchaeota* available for aerosolization, as the surface soil is enriched with *Thaumar-*

366 *chaeota* while their frequency decreases with depth (Jia and Conrad, 2009; Leininger et al.,
367 2006).

368 As meteorological conditions are also known to influence airborne communities, we
369 compared the statistical variance of archaeal diversity in relation to average wind speed, tem-
370 perature, relative humidity and the sum of precipitation, since temperature and relative humid-
371 ity have already been found to correlate with airborne fungal composition (Fröhlich-
372 Nowoisky et al., 2009). We found that on a short term sample-wise basis the relative species
373 diversity of Archaea correlates significantly with wind speed (negative correlation, p-
374 value=0.01, Table S6, Fig. S2a). On a long-term monthly basis multiple linear correlation
375 analysis revealed that the normalized species richness (S_n; Fig. 4a) negatively correlated with
376 both temperature and wind speed, while the relative species richness ($S_{R,k}$ Fig. 4b) showed a
377 significant positive correlation with relative humidity and wind speed, albeit that wind speed
378 played a secondary role in both cases (p-value (S_n)=0.02 and ($S_{R,k}$)=0.01). These findings co-
379 incide with our results that Archaea diversity is highest in the fall/winter period, as in the
380 sampling regions temperature decreases during that time. In fall and winter also relative hu-
381 midity increases. At this point, however, it cannot be answered whether the meteorological
382 factors are causing the diversity change, or whether the correlations are merely coincidental
383 with the diversity and meteorology following similar annual cycle and other factors, e.g.,
384 increased aerosolization sources cause the phenomenon. Although, as illustrated in e.g. Fig.
385 4a between December 2006 and March 2007 the successive steep rise and fall of S_n is predict-
386 ed rather accurately by the linear model, indicating that the correlation might not be coinci-
387 dental.

388 Although S_n stayed constant for *Euryarchaeota* throughout the year, the relative frequency of
389 occurrence (*RFO*), that is, the proportion of samples in which these species were detected,
390 exhibited variations (Fig. 3b). In spring *Euryarchaeota* were present in almost 50% of the air
391 samples, while in the rest of the year they were found only in 20 - 30%. A constant source of
392 *Euryarchaeota* could be the nearby river Rhine, as *Methanomicrobiales* and *Thermoplasma-*
393 *tales* can be found in freshwater (Auguet et al., 2010; Cao et al., 2013; Galand et al., 2006)
394 and might thus be aerosolized. The increase of the *RFO* in spring may be linked to anthropo-
395 genic activities like fertilization of the soil. Although methanogens have been shown to be
396 present in low numbers in aerated soils (Angel et al., 2012), the increase of the here predicted
397 methanogens in spring might be due to fields being fertilized with manure and thus their
398 methanogen composition may be affected (Gattinger et al., 2007; Nicol et al., 2003; Radl et

al., 2007). All of the *Euryarchaeota* detected here are known to be common in livestock manure (St-Pierre and Wright, 2013).

3.3 Ammonia-Oxidizing-Archaea in air

Within the air samples analyzed over a whole year, a total of five OTUs were discovered among the *amoA* gene sequences. As visualized in Fig. 5, two OTUs were present throughout the year, while the others were detected only sporadically. Comparing the two most common OTUs, *AmoA2* showed a high relative frequency of occurrence, while the other was always below 60%. Within our air filter samples we found that the 16S rRNA gene OTU *Arch1* and *amoA* OTU *AmoA2* occur on 94% of the filter samples together, while no other OTUs showed such a high co-occurrence. This co-occurrence suggests that both genes may actually have been amplified from the same microorganisms. This observation confirms that the *Nitrosophaera amoA* cluster may be equivalent to the 16S rRNA cluster *I.1b* as already indicated by other evidence (Bartossek et al., 2012; Pester et al., 2012).

The composition of *amoA* OTUs in the sampled air seems to be a combination of OTUs that are present throughout the year and others that occur only sporadically. Operational taxonomic units that occur over the entire season might originate from a constant source like the Rhine river or sparsely vegetated soil, as it has been shown that *AOA* are present in rivers, freshwater environments, and soils (Cao et al., 2013; Liu et al., 2011; Pouliot et al., 2009). In contrast OTUs that occur only sporadically might originate from unique sources like localized spreading of compost.

The presence of different *amoA* OTUs in air could indicate that air might serve as a medium to distribute *AOA* globally. In this case, fertilization, which is known to decrease *AOA* frequency and diversity in comparison to *AOB* (He et al., 2012; Xu et al., 2012), might also decrease to some extent the spreading of *AOA*.

3.4 Comparison between locations

Initial analyses suggest that the composition of airborne microorganisms shows biogeographic patterns especially between continental, coastal, and marine sites but also across different land-use types, though only a few studies have considered, or attempted to study this phenomenon with molecular or cultural methods (e.g., Bovallius et al., 1978; Bowers et al., 2011; Després et al., 2007; Fröhlich-Nowoisky et al., 2012; Shaffer and Lighthart, 1997; Womack et al., 2010).The comparison between sampling sites can in general be challenging when differ-

ent sampling and analysis procedures are used. Although in a comparison between sites the local geographic and meteorological conditions need to be taken into account, the height in which the samples were taken for this study is negligible as the continental boundary layer air up to 1000 m is fairly well mixed and all sorts of atmospheric measurements can be compared. We thus compared Archaea diversity detected in the continental air of Mainz with the Archaea diversity we discovered in an additional subset of 26 air filter samples from coastal air in Cape Verde, which were collected in winter 2011. The approximately 6 m difference in sampling height is negligible. The air of Cape Verde represents typical coastal air and is influenced by a combination of marine and continental air masses. The 72 sequences detected in Cape Verde air were separated into 25 different OTUs. Both *Euryarchaeota* as well as *Thaumarchaeota* were present (Fig. 1b, Table S5).

While in Mainz 86% of the OTUs were *Thaumarchaeota*, in Cape Verde only 60% could be affiliated with this archaeal group. Out of the *Thaumarchaeota* OTUs detected in Cape Verde, 73% were identical to OTUs detected in Mainz and more than 50% were discovered in several air samples. This might imply that the airborne *Thaumarchaeota* of Cape Verde disperse on a larger scale. However, the few OTUs abundant in Germany were discovered only at the Mainz site. This might indicate that OTUs that are abundant and present throughout the year at this site are dispersed locally and do not originate from far away.

Euryarchaeota in Cape Verde were not very diverse and only contained *Halobacteriales* and *Thermoplasmatales* (Fig. 1b). In Germany, in addition to one OTU of *Halobacteriales*, different methanogens were found. The lack of methanogens in Cape Verde and the dominance of a limited set of *Euryarchaeaota* may be due to the relatively stronger influence of the marine environment compared to terrestrial ecosystems impacted by small-scale agriculture. Interestingly, we did not find *Euryarchaeota* from the marine Group II which are typical for ocean surface layers (Galand et al., 2009; Hugoni et al., 2013; Massana et al., 1997). Although the results for Cape Verde are based on a preliminary data set only, our results might point to general differences in the composition of airborne Archaea in continental versus coastal air.

To follow up on the hypothesis that the composition of airborne Archaea depends on the type of ecosystem found beneath the air masses, samples from other continental and coastal sites were investigated. We successfully amplified Archaea from one of twenty random filter samples taken in Colorado (Huffman et al., 2013). The 20 sequences obtained from this filter all fell into one OTU which was identical to the 16S rRNA gene OTU *Arch31*, a *Thaumarchaeo-*

462 *ta,* also found in Mainz. In China and the United Kingdom we tried to amplify Archaea 16S
463 rRNA gene sequences from 14 and 12 filter samples, respectively, taken from coastal sites
464 with influences from marine and continental air. We could only successfully amplify Archaea
465 16S rRNA gene sequences from one filter sample at each site. All seven sequences from the
466 chinese site belonged to one OTU, which was identical to *Arch16,* a predicted methanogen,
467 observed in Mainz air. In the UK sample set, we amplified 11 sequences from a filter on
468 which primarily air masses from the North Sea were sampled (Gysel et al., 2007). These se-
469 quences grouped into 4 OTUs, all of which belonged to the order *Thermoplasmatales* (*Eu-*
470 *ryarchaeota*). These results support the observations from Cape Verde that in coastal samples
471 the diversity of *Euryarchaeota* might be higher and the composition of them different com-
472 pared to continental samples.

473 We also compared our results with the scarce literature available for airborne Archaea (Table
474 2). Continental airborne Archaea from Salt Lake City (Radosevich et al., 2002) belonged to
475 *Thaumarchaeota* supporting the hypothesis that *Thaumarchaeota* prevail in continental air.
476 Using microarray techniques targeting 16S rRNA genes, Brodie et al (2007) counted, in two
477 cities in Texas, 30% *Crenarchaeota (*under the assumption that these would in a reanalysis be
478 assigned to *Thaumarchaeota)* and 70% *Euryarchaeota*. However, the influence of marine air
479 from the nearby Gulf of Mexico cannot be judged. To our knowledge there are no sequences
480 of airborne Archaea from coastal sites available in the literature. However, Cho and col-
481 leagues (2011) found one single Archaea OTU in marine air from the East Sea (Korea), which
482 was identified as belonging to the *Halobacteriales* (*Euryarchaeota*). Smith and colleagues
483 (2013) using microarrays found that within the very few Archaea detected in air influenced by
484 marine air masses 82% of the OTUs belonged to the *Euryarchaeota*. As the number of availa-
485 ble sequences and literature is limited, more analyses need to be done to confirm or reject the
486 hypothesis about biogeography in airborne Archaea.

487

488 **4 Conclusions**

489 With this study we gained first insights into the seasonal behavior of airborne Archaea diver-
490 sity, composition, and abundance in a continental, agriculture-dominated sampling site. The
491 measured abundance of airborne Archaea (one to ten 16S rRNA gene copies per m^3 air) was
492 much lower than the abundance of airborne Bacteria (10^4-10^6 cp m^{-3}) and the ratio of Archaea
493 to Bacteria was orders of magnitudes lower than in soil and marine environments. We show

that the diversity of common soil *Thaumarchaota* (*I.1b*), which are present throughout the year, correlates positively with relative humidity and negatively with temperature. This, together with the increased number of bare agricultural fields after harvest being available for aerosolization processes, may lead to the observed diversity peaks in fall and winter. In general, the diversity of the *amoA* gene was much lower than in the 16S rRNA gene. The OTU *AmoA1* correlates with *Arch1*, the most abundant OTU from the 16S region. This could point to associations of an *amoA* OTU with an archaeal 16S rRNA gene OTU. The few predicted methanogenic *Euryarchaeota* found show no seasonal trends in diversity but a slightly increased relative frequency of occurrence in spring, when fields are fertilized with manure. The comparison of this continental site with airborne Archaea at a coastal site in Cape Verde indicates differences in community composition and that the diversity of *Euryarchaeota* might increase in coastal air. These results were supported by data from additional sampling sites in the United Kingdom, China, and North America, and also concurred with the very limited information available in the literature.

We conclude that air as a transport medium might play an important role in habitat and niche formation for Archaea. Thus, the analysis of Archaea diversity, abundance, and composition in air is important in order to better understand their behavior and evolutionary history in their primary ecosystems.

Acknowledgement

We thank T. Andreae, R. Burgess, R. Garland, M. Grant, A. Huffman, M. Klose, C. Pöhlker, D. Rose, B. Schmer, B. Schmitt, and H. Yang for technical assistance, H. Paulsen and Z. Jia for helpful discussions. This work was funded by the Max Planck Society, the German Research Foundation (DE1161/2-1, PO1013/5-1, FOR 1525 INUIT), the LEC Geocycles in Mainz funded by the state Rheinland-Pfalz, and the Portuguese Science Foundation through the project PTDD/AAC-CLI/100331/2008 (CV-Dust). J. Cardoso acknowledges the PhD grant SFRH-BD-6105-2009 from FCT. P.E. Galand is supported by the Agence Nationale de la Recherche (ANR) project MICADO (ANR-11JSSV7-003-01).

References

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.

526 Auguet, J.-C., Barberan, A. and Casamayor, E. O.: Global ecological patterns in uncultured
527 Archaea, ISME J, 4(2), 182–190, doi:10.1038/ismej.2009.109, 2010.

528 Baertsch, C., Paez-Rubio, T., Viau, E. and Peccia, J.: Source tracking aerosols released from
529 land-applied class B biosolids during high-wind events, Appl. Environ. Microb., 73(14),
530 4522–4531, doi:10.1128/AEM.02387-06, 2007.

531 Baker, G. C. and Cowan, D.: 16 S rDNA primers and the unbiased assessment of thermophile
532 diversity, Biochem. Soc. T., 32, 218–221, doi:10.1042/BST0320218, 2004.

533 Bartossek, R., Spang, A., Weidler, G., Lanzen, A. and Schleper, C.: Metagenomic analysis of
534 ammonia-oxidizing archaea affiliated with the soil group, Front. Microbiol., 3, 208,
535 doi:10.3389/fmicb.2012.00208, 2012.

536 Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R. and Fierer, N.:
537 Examining the global distribution of dominant archaeal populations in soil, ISME J, 5, 908–
538 917, doi:10.1038/ismej.2010.171, 2011.

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

541 Boetius, A., Ravenschlag, K., Schubert, C. J., Rickert, D., Widdel, F., Gieseke, A., Amann,
542 R., Jørgensen, B. B., Witte, U. and Pfannkuche, O.: A marine microbial consortium
543 apparently mediating anaerobic oxidation of methane, Nature, 407(6804), 623–626,
544 doi:10.1038/35036572, 2000.

545 Bovallius, A., Bucht, B., Roffey, R. and Anäs, P.: Long-range air transmission of bacteria,
546 Appl. Environ. Microbiol., 35(6), 1231–1232, 1978.

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

550 Bowers, R. M., Lauber, C. L., Wiedinmyer, C., Hamady, M., Hallar, A. G., Fall, R., Knight,
551 R. and Fierer, N.: Characterization of airborne microbial communities at a high-elevation site
552 and their potential to act as atmospheric ice nuclei, Appl. Environ. Microb., 75(15), 5121–
553 5130, doi:10.1128/AEM.00447-09, 2009.

554 Bowers, R. M., McLetchie, S., Knight, R. and Fierer, N.: Spatial variability in airborne
555 bacterial communities across land-use types and their relationship to the bacterial
556 communities of potential source environments, ISME J., 5, 601–612, 2011.

557 Brochier-Armanet, C., Boussau, B., Gribaldo, S. and Forterre, P.: Mesophilic Crenarchaeota:
558 proposal for a third archaeal phylum, the Thaumarchaeota, Nat. Rev. Microbiol., 6(3), 245–
559 252, doi:10.1038/nrmicro1852, 2008.

560 Brochier-Armanet, C., Forterre, P. and Gribaldo, S.: Phylogeny and evolution of the Archaea:
561 one hundred genomes later, Curr. Opin. Microbiol., 14(3), 274–281,
562 doi:10.1016/j.mib.2011.04.015, 2011.

Brodie, E. L., DeSantis, T. Z., Parker, J. P. M., Zubietta, I. X., Piceno, Y. M. and Andersen, G. L.: Urban aerosols harbor diverse and dynamic bacterial populations, P. Natl. Acad. Sci. USA., 104(1), 299–304, doi:10.1073/pnas.0608255104, 2007.

Burggraf, S., Huber, H. and Stetter, K. O.: Reclassification of the Crenarchaeal Orders and Families in Accordance with 16S rRNA Sequence Data, Int. J. Syst. Bacteriol., 47(3), 657–660, 1997.

Burrows, S. M., Butler, T., Jöckel, P., Trost, H., Kerkweg, A., Pöschl, U. and Lawrence, M. G.: Bacteria in the global atmosphere – Part 2 : Modeling of emissions and transport between different ecosystems, Atmos. Chem. Phys., 9, 9281–9297, 2009.

Cao, C., Jiang, W., Wang, B., Fang, J., Land, J., Tian, G., Jingkun, J. and Zhu, T. F.: Inhalable microorganisms in Beijing's PM2.5 and PM10 pollutants during a severe smog event., Environ. Sci. Technol., 48, 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.

Cao, H., Dang, H. and Gu, J.-D.: Responses of aerobic and anaerobic ammonia/ammonium-oxidizing microorganisms to anthropogenic pollution in coastal marine environments, Methods Enzymol., 496, 35–62, 2011.

Cao, P., Zhang, L., Shen, J., Zheng, Y.-M., Hong, J. D. and He, J.-Z.: Distribution and diversity of archaeal communities in selected Chinese soils, FEMS Microbiol. Ecol., 80(1), 146–158, 2012.

Cavicchioli, R.: Archaea- timeline of the third domain, Nat. Rev. Microbiol., 9, 51–61, 2011.

Cho, B. C. and Hwang, C. Y.: Prokaryotic abundance and 16S rRNA gene sequences detected in marine aerosols on the East Sea (Korea), FEMS Microbiol. Ecol., 76(2), 327–341, doi:10.1111/j.1574-6941.2011.01053.x, 2011.

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

Després, V. R., Huffman, A. 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.

Després, V. R., Nowoisky, J., 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, 2007.

Edgar, R. C.: MUSCLE: a multiple sequence alignment method with reduced time and space complexity, BMC Bioinformatics, 5, 113, doi:10.1186/1471-2105-5-113, 2004.

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

Falkowski, P. G., Fenchel, T. and Delong, E. F.: The microbial engines that drive Earth's biogeochemical cycles., Science (80-.)., 320(5879), 1034–1039, doi:10.1126/science.1153213, 2008.

Fang, Z., Ouyang, Z., Zheng, H., Wang, X. and Hu, L.: Culturable airborne bacteria in outdoor environments in Beijing, China, Microb. Ecol., 54(3), 487–496, doi:10.1007/s00248-007-9216-3, 2007.

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.

Fröhlich-Nowoisky, J., Burrows, S. M., Xie, Z., Engling, G., Solomon, P. A., Fraser, M. P., Mayol-Bracero, O. L., Artaxo, P., Begerow, D., Conrad, R., Andreae, M. O., Després, V. R. and Pöschl, U.: Biogeography in the air: fungal diversity over land and oceans, Biogeosciences, 9(3), 1125–1136, doi:10.5194/bg-9-1125-2012, 2012.

Fröhlich-Nowoisky, J., Pickersgill, D. A., Després, V. R. and Pöschl, U.: High diversity of fungi in air particulate matter, P. Natl. Acad. Sci. USA., 106(31), 12814–12819, doi:10.1073/pnas.0811003106, 2009.

Galand, P. E., Casamayor, E. O., Kirchman, D. L., Potvin, M. and Lovejoy, C.: Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing, ISME J, 3(7), 860–869, doi:10.1038/ismej.2009.23, 2009.

Galand, P., Lovejoy, C. and Vincent, W.: Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean, Aquat. Microb. Ecol., 44, 115–126, doi:10.3354/ame044115, 2006.

Gattinger, A., Höfle, M. G., Schloter, M., Embacher, A., Böhme, F., Munch, J. C. and Labrenz, M.: Traditional cattle manure application determines abundance, diversity and activity of methanogenic Archaea in arable European soil, Environ. Microbiol., 9(3), 612–624, 2007.

Grosskopf, R., Janssen, P. H. and Liesack, W.: Diversity and structure of the methanogenic community in anoxic rice paddy soil microcosms as examined by cultivation and direct 16S rRNA gene sequence retrieval, Appl. Environ. Microb., 64(3), 960–969, 1998.

Gruber, N. and Galloway, J. N.: An Earth-system perspective of the global nitrogen cycle, Nature, 451(7176), 293–296, doi:10.1038/nature06592, 2008.

Gysel, M., Crosier, J., Topping, D. O., Whitehead, J. D., Bower, K. N., Cubison, M. J., Williams, P. I., Flynn, M. J., McFiggans, G. B. and Coe, H.: Closure study between chemical composition and hygroscopic growth of aerosol particles during TORCH2, Atmos. Chem. Phys., 7, 6131–6144, 2007.

He, J.-Z., Hu, H.-W. and Zhang, L.-M.: Current insights into the autotrophic thaumarchaeal ammonia oxidation in acidic soils, Soil Biol. Biochem., 55(June), 146–154, doi:10.1016/j.soilbio.2012.06.006, 2012.

Huber, T., Faulkner, G. and Hugenholtz, P.: Bellerophon: a program to detect chimeric sequences in multiple sequence alignments, Bioinformatics, 20(14), 2317–2319, doi:10.1093/bioinformatics/bth226, 2004.

Huffman, J. A., Prenni, A. J., DeMott, P. J., Pöhlker, C., Mason, R. H., Robinson, N. H., Fröhlich-Nowoisky, J., Tobo, Y., Després, V. R., Garcia, E., Gochis, D. J., Harris, E., Müller-Germann, I., Ruzene, C., Schmer, B., Sinha, B., Day, D. A., Andreae, M. O., Jimenez, J. L., Gallagher, M., Kreidenweis, S. M., Bertram, A. K. and Pöschl, U.: High concentrations of biological aerosol particles and ice nuclei during and after rain, Atmos. Chem. Phys., 13, 6151–6164, doi:10.5194/acp-13-6151-2013, 2013.

Hugoni, M., Taib, N., Deboras, D., Domaizon, I., Dufournel, I. J., Bronner, G., Salter, I., Agogúe, H., Mary, I. and Galand, P. E.: Sturcture of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters, P. Natl. Acad. Sci. USA., 110(15), 6004–6009, 2013.

Jia, Z. and Conrad, R.: Bacteria rather than Archaea dominate microbial ammonia oxidation in an agricultural soil, Environ. Microbiol., 11(7), 1658–1671, doi:10.1111/j.1462-2920.2009.01891.x, 2009.

Jones, A. M. and Harrison, R. M.: The effects of meteorological factors on atmospheric bioaerosol concentrations-a review, Sci.Total Environ., 326(1-3), 151–180, doi:10.1016/j.scitotenv.2003.11.021, 2004.

Junier, P., Molina, V., Dorador, C., Hadas, O., Kim, O.-S., Junier, T., Witzel, J.-P. and Imhoff, J. F.: Phylogenetic and functional marker genes to study ammonia-oxidizing microorganisms (AOM) in the environment, Appl. Microbiol. Biot., 85(3), 425–440, doi:10.1007/s00253-009-2228-9, 2010.

Jurgens, G., Glöckner, F.-O., Amann, R., Saano, A., Montonen, L., Likolammi, M. and Münster, U.: Identification of novel Archaea in bacterioplankton of a boreal forest lake by phylogenetic analysis and fluorescent in situ hxbridization, FEMS Microbiol. Ecol., 34, 45–56, 2000.

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, 2001.

Kemnitz, D., Kolb, S. and Conrad, R.: High abundance of Crenarchaeota in a temperate acidic forest soil., FEMS Microbiol. Ecol., 60(3), 442–448, doi:10.1111/j.1574-6941.2007.00310.x, 2007.

Kowalchuk, G. A. and Stephen, J. R.: Ammonia-oxidizing bacteria: a model for molecular microbial ecology, Annu. Rev. Microbiol., 55, 485–529, doi:10.1146/annurev.micro.55.1.485, 2001.

Lane, D. J.: 16S/23S rRNA sequencing., in Nucleic acid techniques in bacterial systematics, edited by E. Stackebrandt and M. Goodfello, pp. 205–248, Wiley, Chichester., 1991.

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, 806–809, 2006.

Lighthart, B. and Shaffer, B. T.: Airborne Bacteria in the Atmospheric Surface Layer: Temporal Distribution above a Grass Seed Field, Appl. Environ. Microb., 61(4), 1492–1496, 1995.

Liu, Z., Huang, S., Sun, G., Xu, Z. and Xu, M.: Diversity and abundance of ammonia-oxidizing archaea in the Dongjiang River, China, Microbiol. Res., 166(5), 337–345, doi:10.1016/j.micres.2010.08.002, 2011.

Lueders, T., Manefield, M. and Friedrich, M. W.: Enhanced sensitivity of DNA- and rDNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients, Environ. Microbiol., 6(1), 73–78, 2004.

Massana, R., Murray, a 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.

Moletta, M., Delgenes, J.-P. and Godon, J.-J.: Differences in the aerosolization behavior of microorganisms as revealed through their transport by biogas, Sci. Total Environ., 379(1), 75–88, doi:10.1016/j.scitotenv.2007.02.019, 2007.

Neter, J., Kutner, M. and Wasserman, W.: Applied linear statistical models, 4th ed., edited by Irwin, McGraw-Hill, Irwin, Chicago., 1996.

Nicol, G. W., Glover, L. A. and Prosser, J. I.: Molecular analysis of methanogenic archaeal communities in managed and natural upland pasture soils, Glob. Chang. Biol., 9, 1451–1457, 2003.

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.

Offre, P., Spang, A. and Schleper, C.: Archaea in biogeochemical cycles, Annu. Rev. Microbiol., 67, 437–457, 2013.

Pester, M., Rattei, T., Flechl, S., Gröngröft, A., Richter, A., Overmann, J., Reinhold-Hurek, B., Loy, A. and Wagner, M.: amoA-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic regions, Environ. Microbiol., 14(2), 525–539, doi:10.1111/j.1462-2920.2011.02666.x, 2012.

Pester, M., Schleper, C. and Wagner, M.: The Thaumarchaeota : an emerging view of their phylogeny and ecophysiology, Curr. Opin. Microbiol., 14, 300–306, doi:10.1016/j.mib.2011.04.007, 2011.

Pouliot, J., Galand, P. E., Lovejoy, C. and Vincent, W. F.: Vertical structure of archaeal communities and the distribution of ammonia monooxygenase A gene variants in two meromictic High Arctic lakes., Environ. Microbiol., 11(3), 687–699, 2009.

Prospero, J. M., Blades, E., Mathison, G. and Naidu, R.: Interhemispheric transport of viable fungi and bacteria from Africa to the Caribbean with soil dust, Aerobiologia (Bologna)., 21(1), 1–19, doi:10.1007/s10453-004-5872-7, 2005.

Pruesse, E., Peplies, J. and Glöckner, F. O.: SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes, Bioinformatics, 28(14), 1823–1829, 2012.

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(5), 443–452, doi:10.1038/ismej.2007.60, 2007.

Radosevich, J. L., Wilson, W. J., Shinn, J. H., DeSantis, T. Z. and Andersen, G. L.: Development of a high-volume aerosol collection system for the identification of air-borne micro-organisms, Lett. Appl. Microbiol., 34(3), 162–167, 2002.

R-Development-Core-Team: R: A language and environment for statistical computing, reference index version 2.2.1, R Foundation for statistical computing, Vienna., 2005.

Robertson, C. E., Baumgartner, L. K., Harris, J. K., Peterson, K. L., Stevens, M. J., Frank, D. N. and Pace, N. R.: Culture-independent analysis of aerosol microbiology in a metropolitan subway system, Appl. Environ. Microb., 79(11), 3485–3493, 2013.

Schleper, C., Holben, W. and Klenk, H. P.: Recovery of crenarchaeotal ribosomal DNA sequences from freshwater-lake sediments., Appl. Environ. Microb., 63(1), 321–323, 1997.

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

Schleper, C. and Nicol, G. W.: Ammonia-oxidising archaea-physiology, ecology and evolution, Adv. Microb. Ecol., 57, 1–41, doi:10.1016/B978-0-12-381045-8.00001-1, 2010.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. and Weber, C. F.: Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities, Appl. Environ. Microb., 75(23), 7437–7541, 2009.

Shaffer, B. and Lighthart, B.: Survey of Culturable Airborne Bacteria at Four Diverse Locations in Oregon: Urban, Rural, Forest, and Coastal, Microb. Ecol., 34(3), 167–177, 1997.

Smith, D. J., Timonen, H. J., Jaffe, D. A., Griffin, D. W., Birmele, M. N., Perry, K. D., Ward, P. D. and Roberts, M. S.: Intercontinental dispersal of bacteria and archaea by transpacific winds, Appl. Environ. Microb., 79(4), 1134–1139, 2013.

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. Tech., 2(4), 455–464, doi:10.1080/02786828308958649, 1983.

752 Stackebrandt, E. and Goebel, B. M.: Taxonomic Note: A Place for DNA-DNA Reassociation
753 and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology, Int. J.
754 Syst. Bacteriol., 44(4), 846–849, doi:10.1099/00207713-44-4-846, 1994.

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

757 Stubner, S.: Enumeration of 16S rDNA of Desulfotomaculum lineage 1 in rice field soil by
758 real-time PCR with SybrGreen detection, J. Microbiol. Meth., 50(2), 155–164, 2002.

759 Takai, K. E. N., Moser, D. P., Flaun, M. D. E., Onstott, T. C. and Fredrickson, J. K.: Archaeal
760 Diversity in Waters from Deep South African Gold Mines, Appl. Environ. Microb., 67(12),
761 5750–5760, doi:10.1128/AEM.67.21.5750, 2001.

762 Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S.: MEGA5:
763 molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance,
764 and maximum parsimony methods., Mol. Biol. Evol., 28(10), 2731–2739,
765 doi:10.1093/molbev/msr121, 2011.

766 Thummes, K., Kämpfer, P. and Jäckel, U.: Temporal change of composition and potential
767 activity of the thermophilic archaeal community during the composting of organic material.,
768 Syst. Appl. Microbiol., 30(5), 418–429, doi:10.1016/j.syapm.2007.01.006, 2007.

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

773 Wani, A. A., Surakasi, V. P., Siddharth, J., Raghavan, R. G., Patole, M. S., Ranade, D. and
774 Shouche, Y. S.: Molecular analyses of microbial diversity associated with the Lonar soda lake
775 in India: an impact crater in a basalt area, Res. Microbiol., 157(10), 928–937,
776 doi:10.1016/j.resmic.2006.08.005, 2006.

777 Whelan, S. and Goldman, N.: A general empirical model of preotein evolution derived from
778 multiple protein families using a maximum-likelihood approach., Mol. Biol. Evol., 18, 691–
779 699, 2001.

780 Woese, C. R., Kandler, O. and Wheelis, M. L.: Towards a natural system of organisms:
781 Proposal for the domains Archaea, Bacteria, and Eucarya, P. Natl. Acad. Sci. USA., 87,
782 4576–4579, 1990.

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

785 Woo, A. C., Brar, M. S., Chan, Y., Lau, M. C. Y., Leung, F. C. C., Scott, J. A., Vrijmoed, L.
786 L. P., Zawar-Reza, P. and Pointing, S. B.: Temporal variation in airborne microbial
787 populations and microbially-derived allergens in a tropical urban landscape, Atmos. Environ.,
788 74, 291–300, doi:10.1016/j.atmosenv.2013.03.047, 2013.

789 Woo, C. A.: Molecular ecology and public health risks of urban bioaerosols, 1–119 pp., The
790 University of Hong Kong., 2012.

Xu, Y., Yu, W., Ma, Q. and Zhou, H.: Responses of bacterial and archaeal ammonia oxidisers of an acidic luvisols soil to different nitrogen fertilization rates after 9 years, Biol. Fert. Soils, 48(7), 827–837, doi:10.1007/s00374-012-0677-2, 2012.

Yamashita, T., Yamashita, K. and Kamimura, R.: A Stepwise AIC Method for Variable Selection in Linear Regression, Commun. Stat. - Theory Methods, 36(13), 2395–2403, doi:10.1080/03610920701215639, 2007.

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 1–12, doi:10.1371/journal.pone.0055148, 2013.

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.

807 **Table 1. Overview and diversity parameters of aerosol filter samples and detected Archaea.** Sampling details of continental and coastal sites
808 are given: number of air samples, aerosol size range (total aerosol particle (TSP); particulate matter (PM)), and obtained number of DNA sequences
809 as well as the statistical parameters: species richness (*S* measured, *S** estimated), Shannon index (*H'*), Shannon Evenness (*E*), and Simpson's index
810 (*D*), not available (n.a.), Table S3). The measured species richness *S* in North America, China, and UK needs to be interpreted with caution as only
811 few sequences were available.

	Continental		**Coastal**		
Sampling region	**Germany**	**North America**	**Cape Verde**	**China**	**UK**
Elevation	130 m	2370 m	69 m	21 m	0 m
Above ground	20 m	1 m	14 m	1.5 m	40 m
Latitude	49°59'31.36'' N	39°6'0" N	14°55'33.96'' N	23°54'80.56'' N	52°57'02'' N
Longitude	8°14'15.22'' E	105°5'30" W	-23°29'40.92'' W	113°06'63.89'' E	1°07'19'' E
Samples	47, 47	20	26	14	12
Sampling duration	Mar2006 - April 2007	July 2011	Feb 2011 –Mar 2011	July 2006	Spring 2004
Filter type	glass	glass	quartz	quartz	glass
Average flow rate	0.3 m^3 min^{-1}	1 m^3 min^{-1}	1 m^3 min^{-1}	1 m^3 min^{-1}	1.120 m^3 min^{-1}
Size range	<3 µm, >3 µm	TSP	PM10	TSP	<4.49 µm
Sequences	435	20	72	7	11
S	57	1	25	1	4

S*	137	n.a.	67	n.a.	n.a.
H'	3.32	n.a.	3.07	n.a.	n.a.
E	0.82	n.a.	0.95	n.a.	n.a.
D	0.05	n.a.	0.05	n.a.	n.a.

812 **Table 2. Summary of published airborne Archaea sequence information.** Sampling location, the most likely source of the air masses, the num-
813 ber of reported OTUs, as well as the information as to which archaeal phyla they were attributed to and the performed sequencing technique.
814 *another study by Bowers et al (2013) analyzed bacteria and fungi in air collected at the Colorado Front Range over a 14 month period with the
815 Illumina HiSeq technique. Within this study Archaea were amplified as well, however, no detailed information is provided.

Location	**Air mass**	**OTUs**	***Euryarchaeota* (%)**	***Thaumarchaeota* (%)**	**Other (%)**	**Technique**	**Publication**
North America (Salt Lake City, Utah	Continental	3	0	100	0	Sanger sequencing	(Radosevich et al., 2002)
Middle America (San Antonio & Austin, Texas)	Continental/ urban	307	70	15	15	Microarray	(Brodie et al., 2007)
Asia (East Korea)	Coastal	1	100	0	0	Sanger sequencing	(Cho and Hwang, 2011)
North America (Mt.Bachelor, Oregon)	Transpacific plume	11	82	0	18	Microarray	(Smith et al., 2013)
Asia (Beijing)	Maritime Influence	-	0.8	-	-	Hiseq Illumina*	(Cao et al., 2014)
Europe (Mainz, Germany)	Continental	58	14	86	0	Sanger sequencing	This study
North America (Colorado)	Continental	1	0	100	0	Sanger sequencing	
Africa (Cape Verde)	Coastal	25	40	60	0	Sanger sequencing	
Europe (UK)	Coastal	4	100	0	0	Sanger sequencing	

	Asia (China)	Coastal	1	100	0	0	Sanger sequencing

816

Figure Captions

Fig. 1. Species richness of airborne Archaea. Relative proportions of different phyla (*Thaumarchaeota* and *Euryarchaeota*) and, within the *Euryarchaeota*, of different orders for continental air, Mainz (a) and coastal air, Cape Verde (b), respectively.

Fig. 2. Phylogeny of *Euryarchaeota* and predicted ammonia-oxidizing *Thaumarchaeota*. Schematic illustration of a collapsed Maximum Likelihood tree based on the OTU sequences found in Mainz, Germany, in the 16S rRNA gene (left) and archaeal *amoA* gene (right) as well as basic tree sequences (Table S2). Black dots represent boot strap values >80%. Red stars indicate where OTUs from airborne Archaea (Mainz, Germany) fall. The *amoA* cluster names are given according to Pester et al (2012), the figure has been layed out as in Pester et al (2011).

Fig. 3. Overview of the species richness in Mainz. Seasonal variations of the species richness of *Euryarchaeota* and *Thaumarchaeota* in total suspended particles (TSP) normalized by the number of investigated air samples (a) and in the relative frequency of occurrence (*RFO*) of the phyla *Euryarchaeota* and *Thaumarchaeota* in Mainz, Germany (b).

Fig. 4. Time series depicting the most significant results from a multiple linear regression analysis between the two diversity parameters, normalized species richness per month (S_n) and relative species diversity per month ($S_{R,k}$) and meteorological factors. The observed S_n and $S_{R,k}$ curves (blue) were calculated with the equations given in Table S3. The modeled S_n and $S_{R,k}$ curves (red) were calculated using the monthly averages of the corresponding meteorological factors in the equations depicted in the plots. (a): Time series of S_n. Best variable subset being *T* and *WS* (p-value = 0.011) with error bars being $SE_n(S_n)$.(b): Time series of $S_{R,k}$. Best variable subset: *RH* and *WS* (p-value = 0.022).) with error bars being $SE_n(S_{R,k})$. *RH* average relative humidity [%]; *WS*: average wind speed [m/s]; *T*: average temperature [°C].

Fig. 5. Seasonal variations in the relative frequency of occurrence (*RFO*) of the individual *amoA* OTUs detected in Mainz, Germany**.**

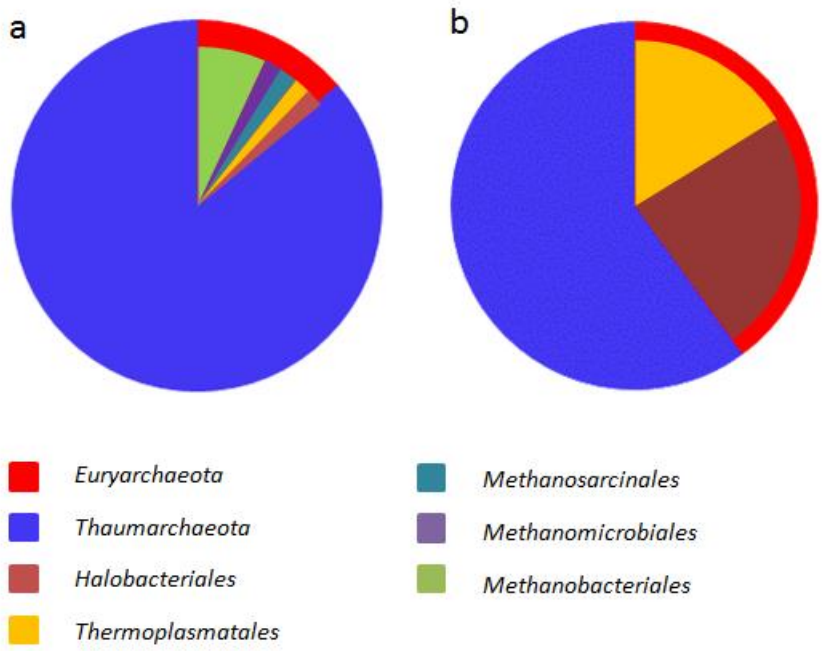
a
b
Euryarchaeota
Thaumarchaeota
Halobacteriales
Thermoplasmatales
Methanosarcinales
Methanomicrobiales
Methanobacteriales

Figure 1

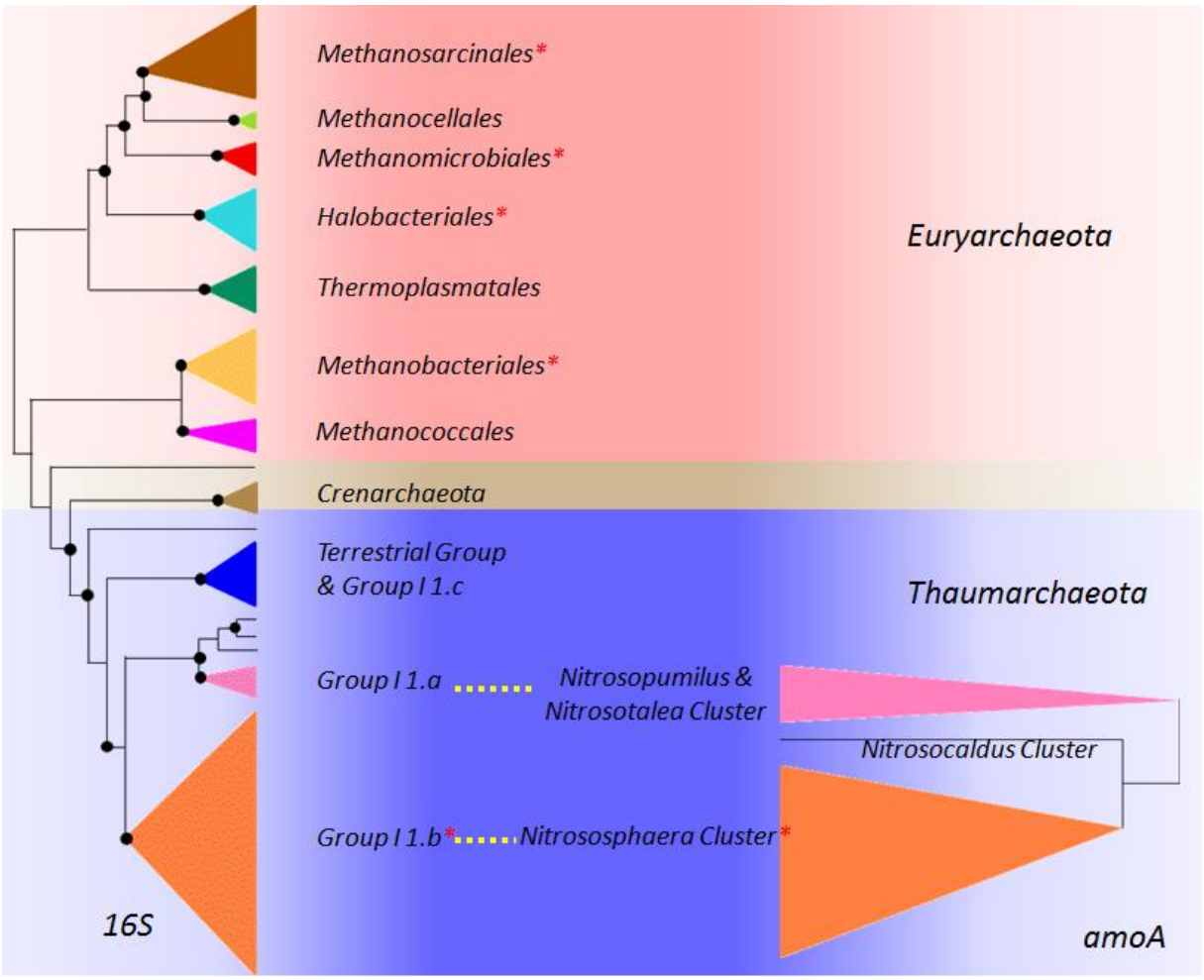
Methanosarcinales*
Methanocellales
Methanomicrobiales*
Halobacteriales*
Thermoplasmatales
Euryarchaeota
Methanobacteriales*
Methanococcales
Crenarchaeota
Terrestrial Group
& Group I 1.c
Thaumarchaeota
Group I 1.a
Nitrosopumilus &
Nitrosotalea Cluster
Nitrosocaldus Cluster
Group I 1.b*
Nitrososphaera Cluster*
16S
amoA

Figure 2

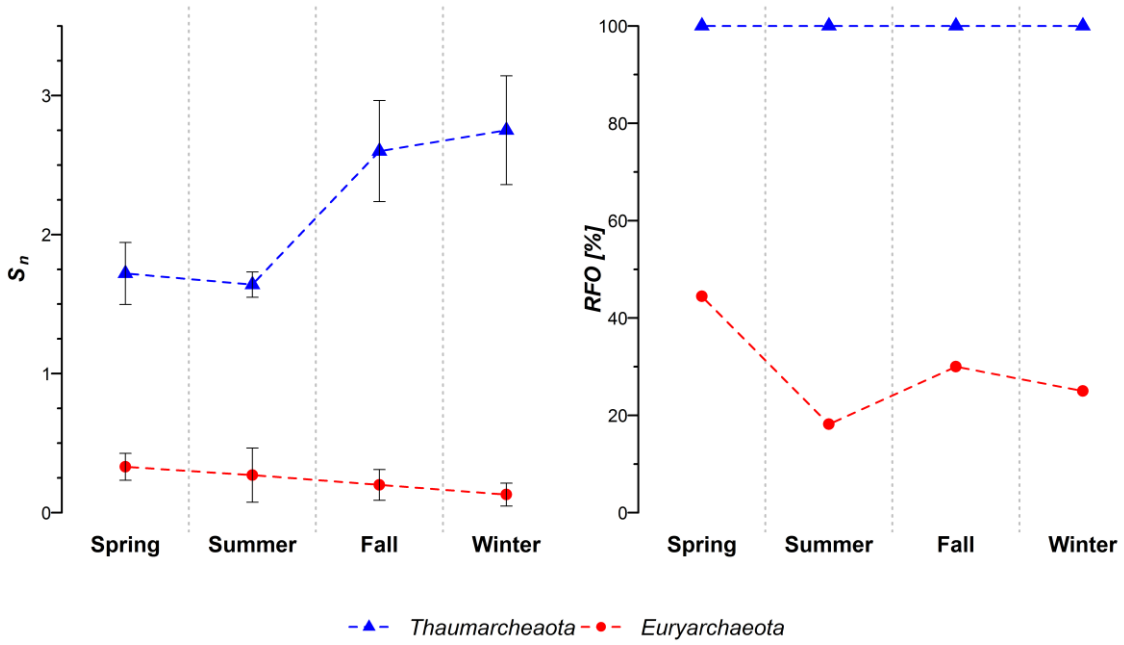
S_n
3
2
1
0
Spring
Summer
Fall
Winter
RFO [%]
100
80
60
40
20
0
Spring
Summer
Fall
Winter
Thaumarcheaota
Euryarchaeota

Figure 3

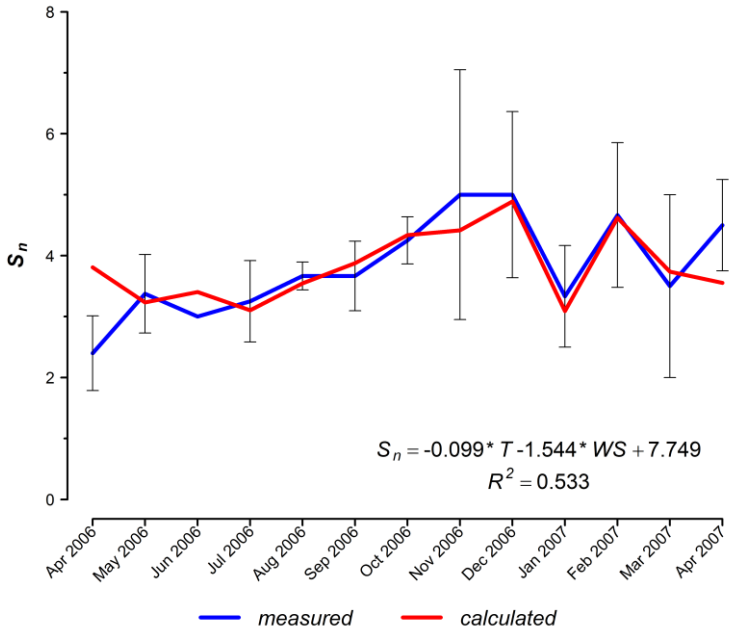
8
6
4
2
0
S_n
$S_n = -0.099 * T - 1.544 * WS + 7.749$
$R^2 = 0.533$
Apr 2006
May 2006
Jun 2006
Jul 2006
Aug 2006
Sep 2006
Oct 2006
Nov 2006
Dec 2006
Jan 2007
Feb 2007
Mar 2007
Apr 2007
measured
calculated

Figure 4a

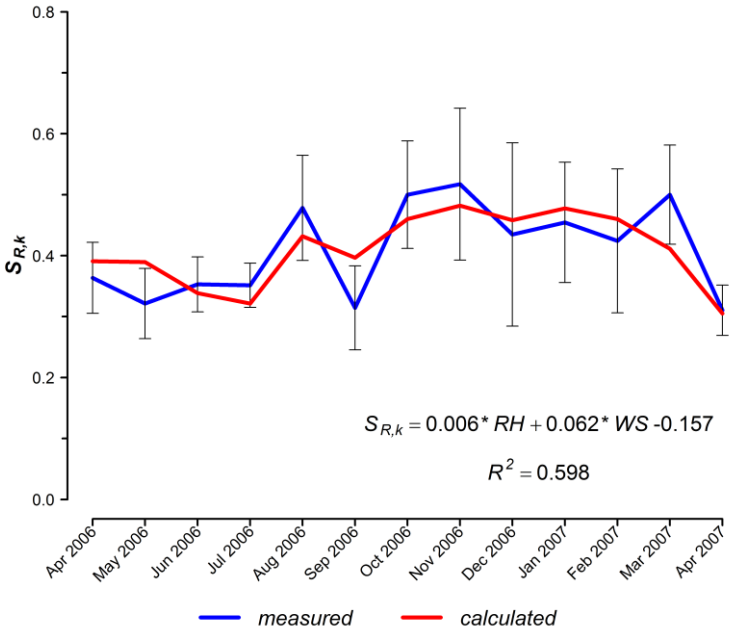
0.8
0.6
0.4
0.2
0.0
$S_{R,k}$
$S_{R,k} = 0.006 * RH + 0.062 * WS - 0.157$
$R^2 = 0.598$
Apr 2006
May 2006
Jun 2006
Jul 2006
Aug 2006
Sep 2006
Oct 2006
Nov 2006
Dec 2006
Jan 2007
Feb 2007
Mar 2007
Apr 2007
measured
calculated

Figure 4b

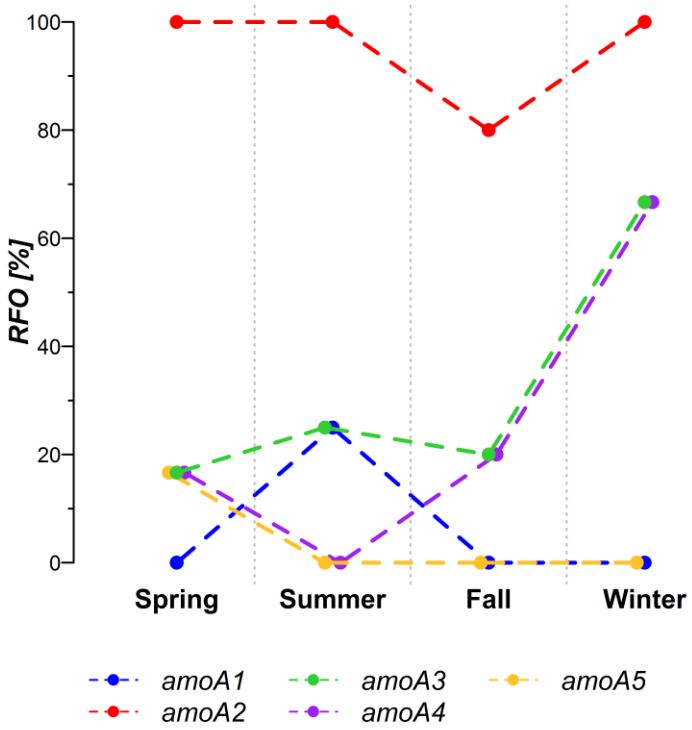
100
80
60
40
20
0
RFO [%]
Spring
Summer
Fall
Winter
amoA1
amoA2
amoA3
amoA4
amoA5

Figure 5