Exploring archaeal and bacterial diversity and co-occurrence in Lake Geneva

Jade Ezzedine, Yves Desdevises, Stéphan Jacquet

Exploring archaeal and bacterial diversity and co-occurrence in Lake Geneva

Jade A. Ezzedine,[1] Yves Desdevises,[2] Stéphan Jacquet[1*]

[1]Université Savoie Mont-Blanc, INRAE, UMR CARRTEL, Thonon-les-Bains; [2]CNRS, Biologie Intégrative des Organismes Marins, Observatoire Océanologique, 2Sorbonne Université, F-66650 Banyuls-sur-Mer, France

ABSTRACT

The diversity and relationships between Archaea and Bacteria remain poorly examined in lakes. Using universal primers targeting 16S rRNA gene *via* Hiseq sequencing, we explored archaeal and bacterial diversity, structure and relationships in the largest natural deep lake in Western Europe, *i.e.*, Lake Geneva. Despite being less diverse than bacteria, archaeal dominant OTUs assigned to the phylum Thaumarchaeota and Nanoarchaeota displayed significant links with a variety of nitrifying bacteria and other bacteria as suggested by co-occurrence networks and function profile predictions. We propose that archaeal OTUs are most likely involved in nitrogen and methanogen cycles, and formed a consortium with other bacteria that also operate these cycles in the deep layers of the lake. These probable syntrophic or mutualistic associations suggest that dominant archaeal OTUs share with some bacteria a similar niche for mutual benefits.

INTRODUCTION

In aquatic ecosystems, microorganisms play a variety of important roles at the foundation of the microbial food webs as prey, predators or parasites, but also as actors in decomposing organic matter and (re)cycling nutrients. Among such microorganisms, archaeal contribution to the microbial loop and the interactions involved between archaea and other prokaryotes have been partially examined so that many pieces of information still remain to be explored and discovered (Braga *et al.*, 2016; Parada and Fuhrman, 2017; Seyler *et al.*, 2019). For a long time, archaea were believed to only inhabit extreme environments. High temperature, salinity and extreme anaerobiosis conditions were thought to be obligate prerequisite for archaeal growth. Therefore, possible links between archaea and other prokaryotes, typically bacteria, were not examined because not expected in such habitats where bacterial growth is improbable or even not possible (Seyler *et al.*, 2019). For the last two decades, however, it has been shown that archaea can be found everywhere (Fuhrman and Davis, 1997; Robertson *et al.*, 2005) including pelagic and surface freshwater ecosystems (Hugoni *et al.*, 2013; Casamayor, 2017). Along the last two decades, archaea were even reported to be as abundant as bacteria in marine sediments (Wurzbacher *et al.*, 2017) or dominant in deep ocean prokaryotic communities (Varela *et al.*, 2008). Archaea, like bacteria, are important players in biogeochemical cycles, *i.e.*, the carbon and nitrogen cycles. Moreover, they are the only life forms to operate some fundamental processes such as methanogenesis (Cavicchioli, 2011) which is a very important function of anaerobic lake waters and sediments (Koizumi *et al.*, 2004; Bomberg *et al.*, 2008). Archaea are also known to oxidize ammonia (Könneke *et al.*, 2005), degrade ammonium urea and, like bacteria, to operate the nitrification process (Alonso-Sáez *et al.*, 2012).

Operating comparable or same functions may logically generate a competition between archaea and bacteria. Therefore, a variety of factors including nutrient availability is likely to structure differently these communities (Berdjeb *et al.*, 2011, 2013). This does not exclude that they can coexist in a non-limiting environment for growth with different responses to resource availability and/or predation pressure by viruses or planktonic grazers (Seyler *et al.*, 2019). It is noteworthy, however, that archaeal growth seems to be favored over bacteria in oligotrophic environments (Vuillemin *et al.*, 2019). Still today, the relationships between archaea and bacteria remain largely unknown. In the microbial realm, we can find relationships such as mutualism, amensalism, competition or elimination for example through antibiotics secretion. Therefore, the outcomes of such relations can be described as positive, neutral or negative (Pacheco and Segrè, 2019). For instance, we can cite the parasitic life style for Pacearchaeota and Woesearchaeota with bacteria (Ortiz-Alvarez and Casamayor, 2016). Another contact dependence has also been reported for two archaea, where *Nanoarchaeum equitans* directly depends on *Igniococcus hospitalis* for its survival (Hu *et al.*, 2018). This could be explained by the fact that some archaea have a compact genome (Koonin and Wolf, 2008) and limited metabolic capabilities (Ortiz-Alvarez and Casamayor, 2016) so they could need a partnership to compensate. Another example is provided by some nitrogen-fixing archaea, designated as a single entity which is known to engage an interaction with the bacterial genus *Desulfosarcina* (Dekas *et al.*, 2009).

The aim of this study was to assess, for the first time using high throughput sequencing, the diversity of both bacterial and archaeal assemblages in Lake Geneva and to investigate the potential relationships between these two communities using co-occurrence network analysis and functional profile prediction. We tested the hypothesis that archaea and bacteria may share similar niches and display significant links for a variety of ecological or biogeochemical functions.

METHODS

Study site

Lake Geneva is a deep and large warm monomictic lake (surface area: 580 km^2; volume: 89 km^3; maximum depth: 309.7 m; mean depth: 152.7 m), located in the western part of the Alps at an altitude of 372 m. Lake Geneva has been monitored since 1974 as a part of a long-term water quality and biological monitoring program. Sampling has been continuously undertaken in the middle of the lake at the deepest point, referred to as SHL2, once or twice a month. This scientific survey revealed that the lake has switched from an oligotrophic to eutrophic state with annual phosphorus concentrations reaching 90 µg L^{-1} in 1979 (Anneville *et al.*, 2002). Thanks to effective management measures, Lake Geneva turned back to a mesotrophic state in the early 2000s with total phosphorus concentrations about 20 µg L^{-1} in 2010 (Jacquet *et al.*, 2014).

Sampling strategy

During the TRANSLEM project, we collected water samples at three sites, including the reference station SHL2. These sites referred to as pt2, pt4 (SHL2) and pt6 (Supplementary Fig. S1), separated from each other by approximately 14 km, were sampled at four different dates and seasons: February 20 (TL1), June 4 (TL2), August 7 (TL3) and November 20 (TL4) 2014, and at three different depths, *i.e.*, 2, 15 and 200 m. It is noteworthy that pt2 could not be sampled on June 4 due to bad weather conditions. A filtration system was set on the boat and a volume of 300 mL of each water sample was filtered through 5 (to select all prokaryotes and avoid clogging of the 2-µm filter), 2 and 0.2 µm polycarbonate 47 mm filters (Merck Millipore, Burlington, MA, USA). Samples were then kept at -20°C until DNA extraction.

Descriptors such as temperature, pH, conductivity, chlorophyll *a* and dissolved oxygen concentrations of the water column were measured using a multiparametric probe (Sea & Sun Technology GmbH, Trappenkamp, Germany). Transparency was measured using a normalized 25 cm diameter Secchi disk. Total organic carbon (TOC), total nitrogen (TN), dissolved ammonium (NH_4-N), dissolved nitrates (NO3-N), total phosphorus (TP) and orthophosphates (PO_4-P) were measured only at pt4, at the different depths and dates, according to the standard French protocols AFNOR.

DNA extraction

The filters were subjected to DNA extraction using a homemade protocol with GenEluteTM-LPA (Sigma-Aldrich, St. Louis, MO, USA) solution. The protocol started with a lysis step in Eppendorf tubes by adding 300 µL of TE buffer (TRIS 1M – pH 8, EDTA 0.5M – pH 8) and 200 µL of a lysis solution (TRIS 1M – pH 8, EDTA 0.5 M – pH 8 and sucrose 0.7 M). Next, a thermic shock was carried out by first placing the tubes at -80°C for 15 min and then thawed into a block heater at 55°C for 2 min. After, 50 µL of a 10% sodium dodecyl sulfate (SDS) and 10 µL of proteinase K (20 mg mL^{-1}) were added to the solution. The solution was incubated at 37°C for 1 h with gentle stirring and placed again in the block heater at 55°C for 20 min. After a quick centrifugation step (13,000 rpm at 4°C for 3 min), the supernatant was collected. Then, 50 µL of sodium acetate (3 M – pH 5.2) and 1 µL of GenEluteTM-LPA (Sigma-Aldrich, 25 µg $µL^{-1}$) were added. One volume of isopropanol was then added and the tubes were centrifuged for 10 min at 12,000 g and 4°C. The supernatant was discarded and 2 washing rounds using ethanol (80%) were carried to purify the DNA. The remaining ethanol was evaporated using a Speed-Vac for 20 min. Finally, 30 µL of TE were added and tubes were left for 1 h at 37°C. DNA concentration was measured using a NanoDrop 1000 spectrophotometer. The amount of DNA extracted from 5 µm size filters ranged between 4.25 and 138.09 ng $µL^{-1}$ with an average of 45.37 ng $µL^{-1}$. For the 2 µm filters, the minimum, maximum and mean concentrations were 6.71, 166.77, and 59.75 ng $µL^{-1}$, respectively. For the 0.2 µm filters, the minimum, maximum and mean concentrations were 2.93, 215.18 and 66.89 ng $µL^{-1}$ respectively. Afterwards, all tubes were stored at -20°C until analysis.

PCR and sequencing

Total DNA extracts were set at 25 ng $µL^{-1}$. Total DNA extracted from the 5 and 2 µm filters were pooled since we had low amplification when considering only 2 µm filters. 0.2 µm filter was considered as it is. The PCR amplification of 16S rRNA gene fragments was performed using universal set of primers. The primers, namely forward primer 515F (GTGYCAGCMGCCGCGGTA) (Wang and Qian, 2009) and reverse primer 909R (CCCCGYCAATTCMTTTRAGT) (Wang *et al.*, 2018) had tags attached to them. It should be noted here that we verified, using TestPrime (*not shown*), that this set of primers had good prokaryotic coverage. The final number of samples was 66 and a PCR replicate was made for each sample giving a total of 132 samples. Each sample was identified with a different tag. PCR mixture volume was 30 µL and consisted of (final concentration): 1x buffer, 0.5 mM dNTP, 1.5 mM $MgCl_2$, 0.5 mg mL^{-1} bovine serum albumin (BSA) and 0.75 U Biotaq DNA polymerase (Bioline). In a second step, a unique combination of tagged primers (forward and reverse) was added to each sample. Finally, 1 µL of template DNA (25 ng $µL^{-1}$) was added. A negative control was included and the PCR program was as follows: 95°C – 2 min, 30 x (94°C

– 30 sec, 58°C – 30 sec, 72°C – 30 sec), with a final extension step at 72°C for 5 min. Agarose gel analysis was performed to check the PCR products. When samples showed a non-specific band (approximately 550 pb) close to the target band (approximately 450 pb), the target bands were then captured using Pippin prep system (sage science) following the manufacturer's instructions. These captured DNA were checked with TapeStation (Agilent 2200) system for size and quality assessment following the manufacturer's instructions. All amplified DNA were quantified using the Quant-iT PicoGreen ds DNA Reagent kit (Invitrogen) and fluorescence was read using the plate reader Fluoroskan Ascent FL. All DNA samples were then pooled as one equimolar tube. DNA was then purified using the Clean PCR kit (CleanNA) according to the manufacturer's instructions to remove dNTP and dimers. Again, the pool was quantified using PicoGreen. Then, the pool tube containing the 132 different tagged DNA was sent to GATC-Eurofins platform for DNA sequencing using Illumina HiSeq paired end technology (2 x 250 bp).

Bioinformatics pipeline

Two paired fastq files, referred to as R1 (forward sequences) and R2 (reverse sequences), were received from the sequencing platform. Files were processed using the pipeline developed by Frederic Mahé which combines Vsearch (Rognes *et al.*, 2016), Cutadapt (Martin, 2011) and Swarm (Mahé *et al.*, 2015) and reported at https://github.com/frederic-mahe/swarm/wiki/Fred's-metabarcoding-pipeline. All default parameters of the pipeline were left unchanged except when mentioned. Briefly, the pipeline started with merging reads of R1 and R2 files using Vsearch. Next, Cutadapt was used to demultiplex the sequences according to a list of tags. Here, each sequence was assigned to its sample in an individual file by its tag. Primers were also removed and sequences containing ambiguous bases were discarded. In each file, sequences were dereplicated using Vsearch. All files were then assembled as one file and dereplicated in the process. The Swarm algorithm was then used to cluster the sequences. The "d" parameter (*i.e.*, the number of different nucleotides between sequences) was changed from 1 to 13 in order to be close to the identity threshold of 97% which describes same species. For the calculation, we took into account the median length of sequences, which is 377 bp. Next, *de novo* chimera detection was applied to representative sequences of each cluster using Vsearch. Using Stampa (Sequence Taxonomic Assignment by Massive Pairwise Alignments), the representative sequences were taxonomically assigned to a reference database downloaded from arb-SILVA (release number 132; (Quast *et al.*, 2013)) and prepared as required. The final OTU table was built using the python script provided in the pipeline. Then, multiple filtering was carried out on the OTUs table and these steps were named "default", "stringent", and "shared". In the "default" step, the OTU table was filtered by removing chimera, low quality sequences (<0.0002) and OTUs with less than 3 reads in a sample, unless they were present in 2 or more samples. After examining the data, we applied the "stringent" filter. It consisted of keeping OTUs with 10 or more reads present in 2 or more samples, removing OTUs with an identity score lower than 90% to arb-SILVA and discarding unassigned OTUs. Finally, we applied the last filter, "shared", where we kept OTUs that are common between the PCR replicates. On the other hand, we first intended to study microorganisms attached to particles, *i.e.*, those captured on the 2 µm filter *vs* the free-living part collected between 2 and 0.2 µm. However, due to technical issues, we chose to associate reads of all filters (5+2 µm and 0.2 µm), providing prokaryotes of all size. At the end, after applying all the aforementioned filters and association, only 33 out of 132 samples were analyzed.

Statistical analysis

Statistical analyses and plots were performed using R, version 3.5.0 (R Core Team, 2019) with the ggplot2 package (Wickham *et al.*, 2018). The OTU table was transformed to relative abundance using the "decostand" function from the vegan package (Oksanen *et al.*, 2019). Thus, all statistical analysis was performed using relative abundance. Alpha diversity indices (*i.e.*, Shannon, Pielou, Chao1) were calculated with the OTUtable package (Linz, 2018). Simpson and inverse Simpson were computed using the vegan package (Oksanen *et al.*, 2019). The indexes comparison for each condition (Site, Month, and Depth) was performed using ANOVA or Kruskal-Wallis test according to the data distribution after initial testing with Shapiro and Bartlett test. When the p-value of ANOVA or Kruskal-Wallis test was inferior to 0.05 (alpha), a Tukey HSD or a Dunn test with Bonferroni correction (Dinno, 2017) were performed. An NMDS to illustrate beta diversity was computed using the "metaMDS" function from vegan and also the goeveg package (Goral and Schellenberg, 2017). Also, Adonis and Anosim (Oksanen *et al.*, 2019) tests were performed to analyse whether the groups observed were significantly different. Moreover, a Simper test (Oksanen *et al.*, 2019) was performed to assess the contribution of each OTU to the observed dissimilarity between samples. Regarding environmental variables, they were only available for pt4. OTUs of pt4 were extracted and if an OTU had no read in that site, the OTU was removed from the final OTUs table. The raw data was transformed as $\log(1 + x)$ in order to stabilize variances, and converted to relative abundance in order to perform a CCA following the online tutorial of Umer Zeeshan Ijaz (Torondel *et al.*, 2016) found at

http://userweb.eng.gla.ac.uk/umer.ijaz/bioinformatics/ecological.html. Redundant environmental variables were removed after calculating a variance inflation factor (VIF). Any environmental variable with VIF superior to 10 were deleted. Co-occurrence network analysis was performed on all sites for OTUs and phylum-class clusters following Ju Feng's R and python scripts (Ju *et al.*, 2014; Ju and Zhang, 2015; Hu *et al.*, 2017) found at https://github.com/RichieJu520/Co-occurrence_Network_Analysis. Only positive interactions between community members were considered. The Spearman correlation and p-value cutoffs were set to 0.9 and 0.01 in the script. As for the C-score of the network, it was performed with the R package EcoSimR (Gotelli *et al.*, 2015) based on a presence/absence OTU matrix. The visualization and customization of the network was done with Gephi software (Bastian *et al.*, 2009). Finally, we used the computational method called pangenome-based functional profiles (PanFP) (Jun *et al.*, 2015) to infer functional profiles for bacterial and archaeal classes. The software can be found at (https://github.com/srjun/PanFP). Generated KOs where then analyzed using KEGG website (https://www.genome.jp/kegg/ko.html) (Kanehisa *et al.*, 2016). Pairwise intersection of KOs between archaea and bacteria were obtained from the website Molbiotools (http://www.molbiotools.com/listcompare.html). At last, the total shared KOs were calculated using the dedicated Bioinformatics and Evolutionary Genomics website tool, available at http://bioinformatics.psb.ugent.be/webtools/Venn/.

Sequence data

Paired end fastq files, tags list, unfiltered bacterial and archaeal OTU tables can be found at Zenodo repository website following this link: http://doi.org/10.5281/zenodo.3678385

RESULTS

The sequencing results along with the details regarding archaeal and bacterial reads and OTUs can be found in the Supplementary Material 1 and Supplementary Figures S2-S8.

Archaeal diversity and distribution

The archaeal reads and OTUs represented 9% and 4% of the dataset, respectively. OTU 2, assigned to Nitrososphaeria (Thaumarchaeota), constituted 97.4% of archaeal reads and was the second OTU to hold the highest number of reads among bacterial and archaeal OTUs. Overall, 5 phyla and 7 classes were detected after taxonomic assignment. Furthermore, Thaumarchaeota was dominant in all samples (Supplementary Fig. S4) followed by Nanoarchaeaeota, mostly present in all samples and characterized by a lower number of reads. Archaeal richness illustrated with the Chao1 index (Supplementary Fig. S5) was similar to bacteria, and the highest values for this index were obtained at 200 m depth ($p=1.859^{-05}$). Evenness (Pielou index) was significantly different between seasons (p=0.01758), especially between winter (February) and summer (August), but not between depths (Supplementary Table S1). Archaeal diversity calculated with Shannon, Simpson and InvSimpson indices (Fig. 1A and Supplementary Table S1) also revealed significant differences between depths (p=0.006761), especially between near surface and deep waters (p=0.0028). Similarly, depth was a factor explaining the difference in OTUs composition among samples as showed by the NMDS (Fig. 2A), Anosim (p=0.001) and Adonis (p=0.003) (Supplementary Table S2). Undoubtedly, one OTU contributed significantly to the beta diversity between treatments, *i.e.*, OTU 2 (Thaumarchaeota, Nitrososphaeria), in particular regarding depth (p=0.006761) since it drove respectively more than 45%, 47% and 40% of the difference between 2 and 15 m, 2 and 200 m and 15 and 200 m for the entire community (Supplementary Table S3). Finally, the CCA (Fig. 3) performed on Pt4 OTUs revealed that these 77 OTUs were positively correlated to total phosphorus, especially at 200 m, and also to conductivity, total nitrogen and ammoniacal nitrogen. In contrast, archaeal OTUs were negatively correlated to chlorophyll *a*, dissolved oxygen and total organic carbon.

Bacterial diversity and distribution

Forty-one bacterial phyla and 110 classes were assigned according to the pipeline. Alphaproteobacteria (Proteobacteria) yielded the highest number of OTUs (*i.e.*, 411) (Supplementary Fig. S6) and Actinobacteria the highest number of reads, *i.e.*, 8,948,606 (Supplementary Fig. S7). Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Planctomycetes and Proteobacteria were highly present in all samples (Fig. 4). Samples at 200 m contained a higher phylum diversity than any other sample. Indeed, when looking at the richness and diversity indices performed on the OTUs, the results showed a significant difference according to depth, while no significant diversity shifts were observed for sites or seasons (Supplementary Table S4). Samples at 200 m displayed higher and significant values for all indices (Chao1 p-value=7.710^{-10}, Pielou p-value=4.706^{-6}, Shannon p-value=4.322^{-08}, Simpson p-value=4.757^{-08} and InvSimpson p-value=2.335^{-05}) compared to the two other depths (2 and 15 m) (Fig. 1-B and Supplementary Fig. S8). All investigated sites were not significantly different in OTUs composition (Fig. 2-B) which was affected mainly by depth (Anosim and Adonis p-value were respectively 0.001 and 0.001) and by season (Anosim and Adonis p-value were 0.006 and 0.011, respectively) (Sup-

plementary Table S5). The Simper test (Supplementary Table S6) indicated that OTUs 4 and 5, assigned to the class Anaerolineae (Chloroflexi) and Oxyphotobacteria (Cyanobacteria) respectively, drove significantly the beta diversity. Indeed, OTU 4 contributed ($p=1.714^{-05}$) to the disparity between surface and depth (up to 11%), as well as between 15 and 200 m depth (up to 12%). In addition, the average dissimilarity between 2 and 200 m was 65% and it was about 62% between 15 m and 200 m. OTU 5 contributed significantly (p=0.0003058) to the disparity between February and the 3 other months (June 14%, August 14% and November 15%). The CCA (Fig. 5) that was performed on 1964 OTUs found at pt4, revealed that the variance is drawn in all directions. On one hand, total phosphorus revealed to be well linked to the OTUs present at 200 m. On the other hand, total nitrogen, chlorophyll *a*, conductivity and dissolved oxygen were linked to bacterial OTUs present at 2 and 15 m.

Relations between archaea and bacteria

To assess possible interactions between bacteria and archaea we used two different and complementary approaches performed on the OTUs tables: a co-occurrence network analysis to determine significant connections between microbes and a functional profile prediction analysis.

The co-occurrence network (made of 260 and 24 nodes and 1230 and 33 undirected edges, for bacteria and archaea respectively) revealed connections among and between bacterial and archaeal OTUs. We only reported here strong (correlations ≥0.9) and significant ($p<0.01$) links. The C-score test with sim9 algorithm confirmed that the network was non-random. Indeed, the observed C-score (10.523) was higher than the mean value (c-score mean=10.321, $p<0.001$) expected under the null model. On one hand, the archaeal OTU 2 assigned to the genus *Nitrosopumilus* (Thaumarchaeota, Nitrososphaeria, Nitrosopumilales, *Nitrosopumilaceae*) was linked to 8 bacterial OTUs. These OTUs were identified as *Anaerolineaceae* (Chloroflexi, Anaerolineae, Anaerolineales, OTU 4), BSV26 (Bacteroidetes, Ignavibacteria, Kryptoniales, OTU 10), CL500-3 (Planctomycetes, Phycisphaerae, Phycisphaerales, *Phycisphaeraceae*, OTU 19), *Nitrospira* (Nitrospirae, Nitrospira, Nitrospirales, *Nitrospiraceae,* OTU 25), IMCC26256 (Actinobacteria, Acidimicrobiia, OTU 46), SAR11 clade (Proteobacteria, Alphaproteobacteria, OTU 124), Elsterales (Proteobacteria, Alphaproteobacteria, OTU 135), and *Methylomonaceae* (Proteobacteria, Gammaproteobacteria, Methylococcales, OTU 149). On the other hand, OTU 67 identified as the archaeal class Woesearchaeia (Nanoarchaeaeota) was linked to 4 bacterial OTUs (Fig. 6). These

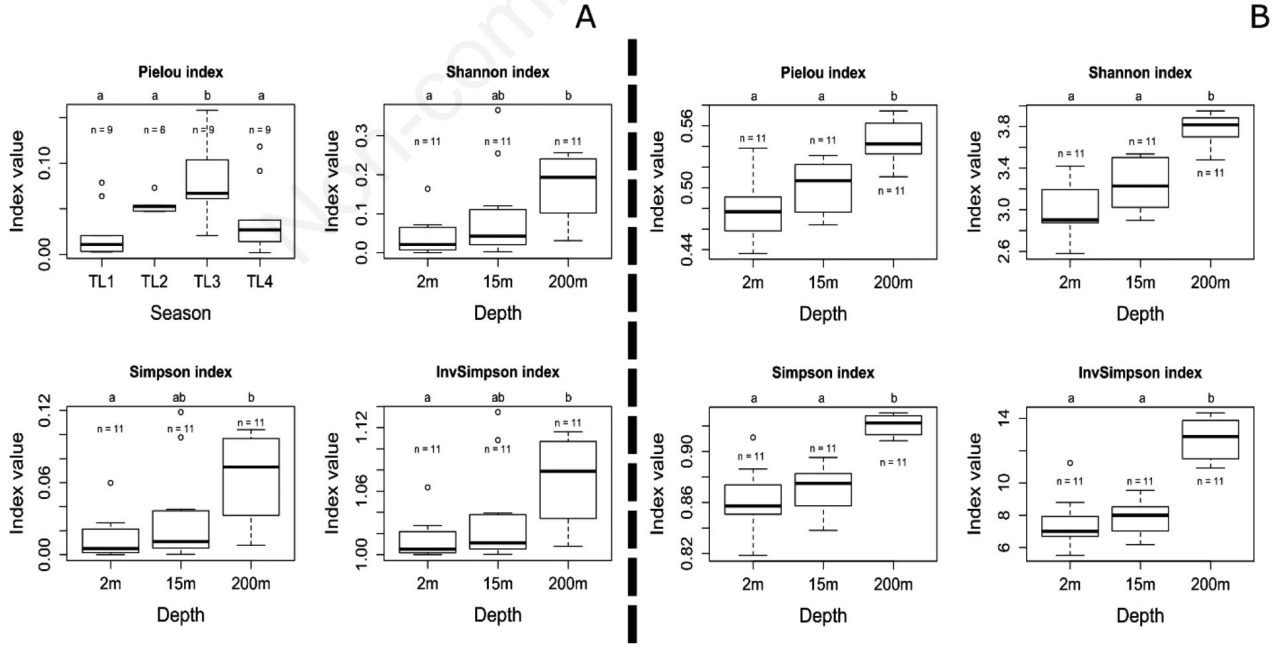

Fig. 1. Boxplot diagram comparing alpha diversity indices for archaea (A) and bacteria (B) (only significant results for indices are shown). For archaea, significant differences ($p<0.01$) are detected among the depth and season (TL1: February; TL2: June; TL3: August; TL4: November) variables. For bacteria, significant differences ($p<0.01$) are detected among depths. N, number of samples; a, b, significant differences.

OTUs are identified as BSV26 (Bacteroidetes, Ignavibacteria, Kryptoniales, OTU 10), Peribacteria (Patescibacteria, Gracilibacteria, OTU 188), P2-11E (Chloroflexi, OTU 417) and KD4-96 (Chloroflexi, OTU 587). At the class level (Supplementary Fig. S9), Nitrososphaeria (Thaumarchaeota) was only linked to Anaerolinae (Chloroflexi). The Woesearchaeia (Nanoarchaeaeota) class maintained 3 links which were P2-11E (Chloroflexi), KD4-96 (Chloroflexi) and Gracilibacteria (Patescibacteria) and a new link with JG30.KF.CM66 (Chloroflexi) appeared. The C-score test also confirmed that this network is non-random with observed C-score (5.1653) being higher than the mean value (c-score mean=4.4757, $p<0.001$) expected under the null model.

The prediction of functional profiles from lineages (phylum – class) that are involved in the co-occurrence network were only possible for Thaumarchaeota - Nitrososphaeria and its associated bacteria *i.e.*, Chloroflexi – Anaerolineae; Bacteroidetes – Ignavibacteria; Planctomycetes – Phycisphaerae; Nitrospirae – Nitrospira; Actinobacteria – Acidimicrobiia; Proteobacteria – Alphaproteobacteria; Proteobacteria – Gammaproteobacteria (Supplementary Fig. S10). No functional profiles were found by the software for Nanoarchaeaeota – Woesearchaeia. Nitrososphaeria had the minimum number of functional orthologs (KO=736), and Gammaproteobacteria the highest (*i.e.*, 5217). Overall, the archaeal class Nitrososphaeria had less KOs involved in metabolic pathways, enzymes, carbon, sulfur and pyruvate metabolism, but methane and nitrogen metabolism. Moreover, we calculated the shared number of KOs among all these class and found a value of 186. When looking at the pairwise intersections (Supplementary Fig. S11) Nitrososphaeria had the less shared KOs with Alpha and Gammaproteobacteria (Jaccard index=0.1231 and 0.0989, respectively), while most KOs were shared among Alpha and Gammaproteobacteria (Jaccard index=0.6833). This result showed that some pathways may be lacking in Nitrososphaeria.

DISCUSSION

The choices we made shaped our observations

We investigated the genetic diversity of bacteria and archaea at four months representing the different seasons of the year 2014. We are aware that the time scale of this study was relatively short (*i.e.*, four dates) compared to other studies (Bomberg *et al.*, 2008; Hugoni *et al.*, 2013; Parada and Fuhrman, 2017) and the relatively low number of sampled sites, depths and months might not be representative to illustrate all the complex diversity of both archaea and bacteria in Lake Geneva.

Using Illumina HiSeq, although being very useful to assess deeply the 16S rRNA-based genetic diversity of these prokaryotes, an important issue could be the primer specificity since we noted that the pair of primers targeting the V4-V5 region used in our study preferably amplified Actinobacteria and Thaumarchaeota. The *in silico*

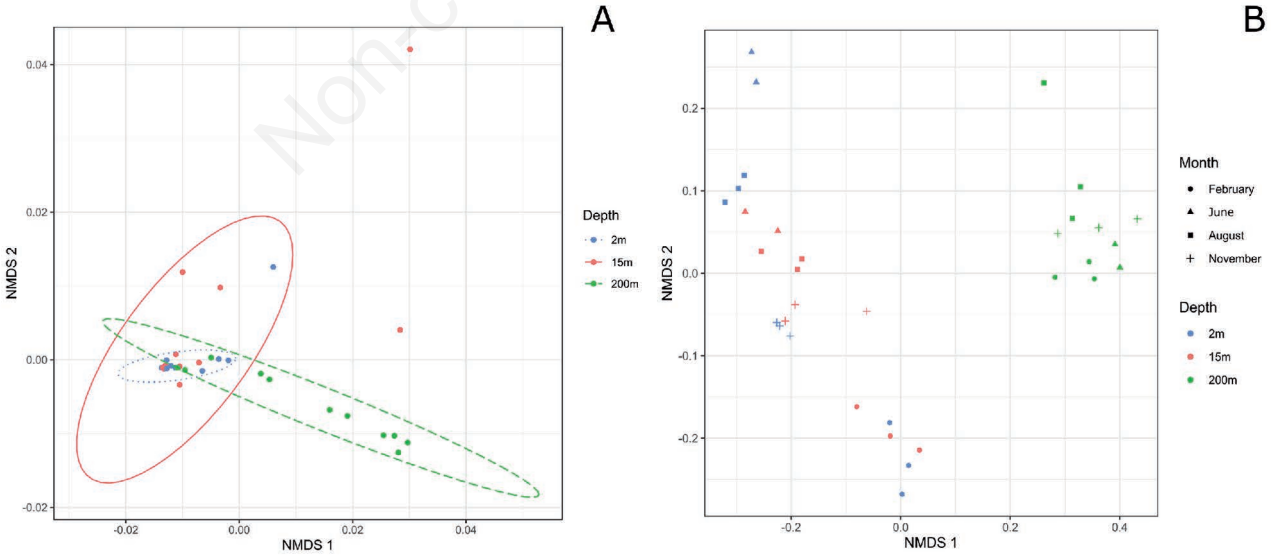

Fig. 2. NMDS plot illustrating separation of samples based upon differences in archaeal (A; stress value 0.03) and bacterial (B; stress value 0.05) OTUs structure. Shape and color correspond respectively to different seasons February (TL1), June (TL2), August (TL3) and November (TL4), and depths (2, 15 and 200 m). For archaea, ellipses indicate 95% confidence intervals of OTUs grouped by different depths. For bacteria, the ellipses could not be generated.

results showed that these primers had a good taxonomic resolution, covering 80.2% and 98.8% of the taxa at 0 and 3 mismatches respectively. It is noteworthy, however, that the behavior of primer pairs *in vitro vs in silico* may greatly differ as demonstrated in a previous study (Ezzedine *et al.*, 2020). Ideally, different universal primers should be tested with the same sequencing technology to compare results.

It is also noteworthy that the taxonomic assignment is the result of bioinformatics algorithms, a procedure not perfect that can result in a variety of bias. Also, the sequence identity threshold for 16S rRNA OTUs can induce misidentification. As stated in the literature, sequences with 95% identity are likely to represent a same genus, whereas sequences with 97% identity may represent a same species (Schloss and Handelsman, 2005; Schloss *et al.*, 2009). In this study, we used OTUs, set the threshold at 96.6% and chose to analyze the taxonomic assignment to the genus level, whenever possible. Indeed, the Swarm algorithm was then used to cluster the sequences. The "d" parameter (*i.e.*, the number of different nucleotides between sequences) was changed from 1 to 13 in order to be close to the identity threshold of 97% which describes same species. For the calculation, we took into account the median length of the sequences that was 377 bp. In addition, we chose stringent cutoff to filter the OTUs tables and considered only shared OTUs among PCR replicates.

As we explored relationships between dominant archaea OTUs and other bacterial OTUs *via* a co-occurrence network analysis, we have to remind here again that it was constructed using a stringent cutoff (>0.9) in order to only reveal significant and strong correlations. In addition, we made the choice by focusing on positive correlation. The co-occurrence network along with the detected functional profiles from 16S rRNA (and not from metagenome) are presented as a suggestion based on strong statistical analysis.

Regardless, this study sheds light on both bacterial and archaeal diversity in Lake Geneva using a high throughput sequencing approach and highlights, for this large and deep lake, possible relationships between prokaryotic communities, known to play a variety of key roles in aquatic ecosystems.

Archaeal and bacterial diversity

Our data set was mainly represented by two archaeal phyla, Thaumarchaeota and Nanoarchaeaeota, while ar-

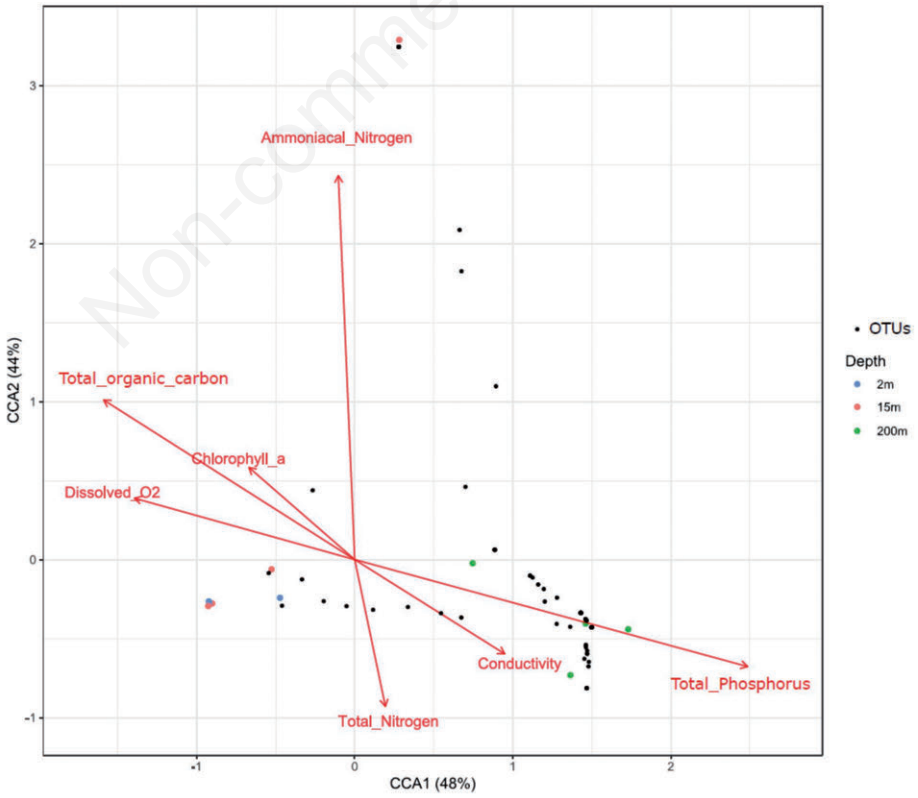

Fig. 3. CCA illustrating the separation of archaeal OTUs and samples based upon physicochemical descriptors. The CCA was only performed on pt4 site were physicochemical variables were obtained. 12 samples and 77 archaeal OTUs were used in this analysis.

chaea are generally more diversified in aquatic ecosystems (Kozubal *et al.*, 2013; Evans *et al.*, 2015; Vanwonterghem *et al.*, 2016; Eme *et al.*, 2017). Other phyla were in fact detected, such as Euryarchaeota, Diapherotrites and Crenarchaeota but they were clearly not dominant and only found with a low number of reads and OTUs. Also, the diversity was found to be higher in deep-water (*i.e.*, 200 m). Globally, it has been reported that archaea may represent less than 10% of the microbial population in freshwater habitats (Bomberg *et al.*, 2008; Bahram *et al.*, 2019). Our results are in accordance with these previous studies since archaea reads represented only 9% of the whole sequencing data set. This result could be related to the archaeal genome which is compact with a low number of ribosomal copies (Koonin and Wolf, 2008; Angly *et al.*, 2014), that may lead to underestimation of archaeal abundance (Wurzbacher *et al.*, 2017). In addition, the quite generic primers we used could have played a part in overlooking the archaeal diversity. As a matter of fact, (Bahram *et al.*, 2019) proved that employing an adequate pair of specific primers enabled to get accurate estimate of archaeal diversity by detecting readily, the Asgard, the TACK and the DPANN superphyla in a variety of ecosystems. That being said, (Haller *et al.*, 2011) in their study of bacterial and archaeal communities at different contamination levels, using a cloning-sequencing approach also in Lake Geneva, reported that Euryarchaeota were mainly found in contaminated sediments, rich in organic matter. (Eme *et al.*, 2017) found that ammonia oxidizing archaea (AOA) Thaumarchaea living in the neighboring Lake Bourget, another (French) peri-alpine lake, were favored when ammonia concentrations were the lowest and

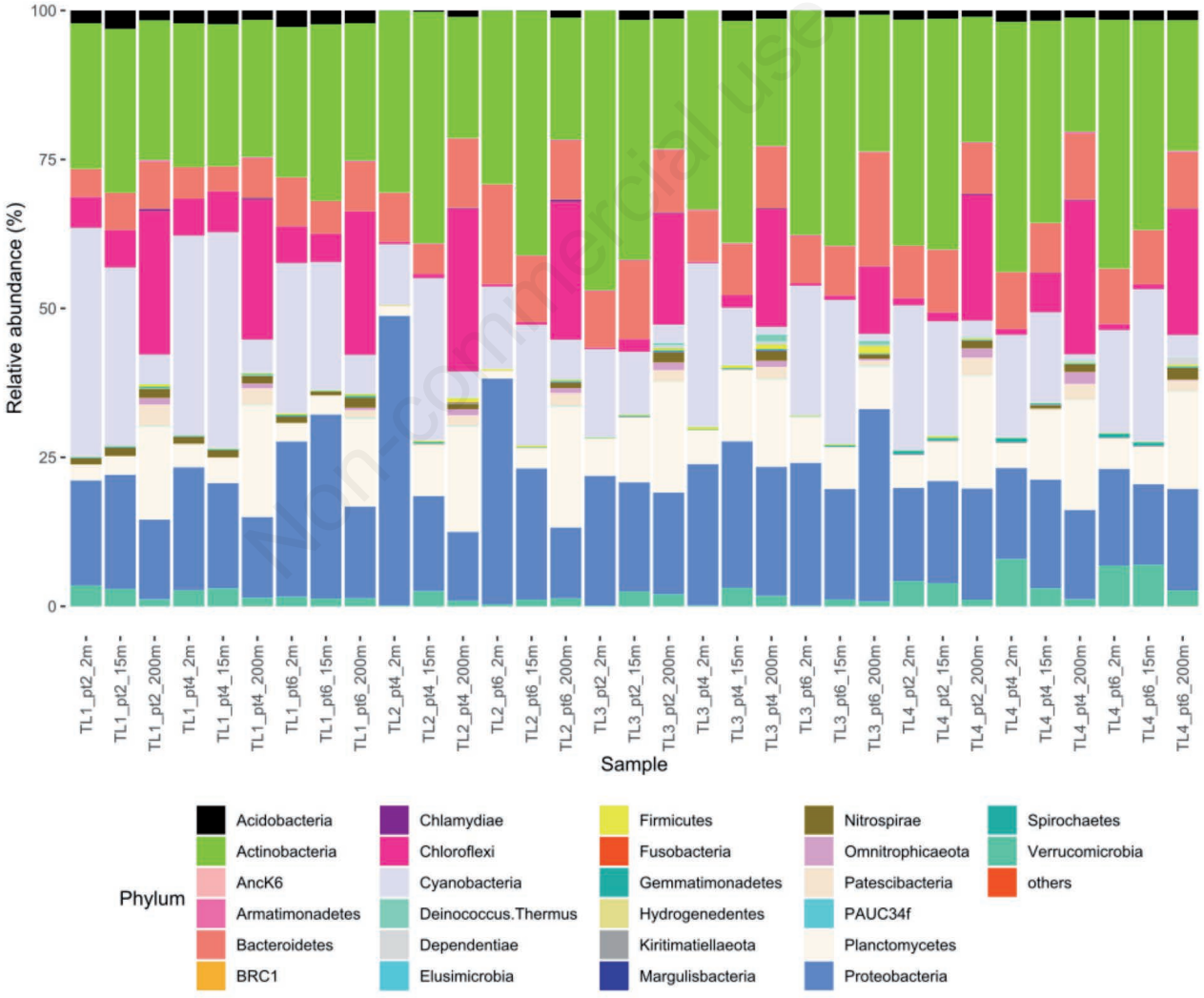

Fig. 4. Stacked histogram of bacterial phylum. Phylum abundances are represented by read relative abundance per sample. Phyla with less than 5% of read relative abundance in a sample are merged and identified as “others”.

during winter. For the same lake, other studies reported that archaea can occupy different ecological niches due to the influence of different abiotic and biotic factors. Indeed, archaeal OTUs were shown to be structured by factors such as ammoniacal nitrogen, total nitrogen and total phosphorus (Berdjeb *et al.*, 2013; Pollet *et al.*, 2018). At last, (Parada and Fuhrman, 2017) reported that depth and seasonality influenced archaeal community composition, and that archaeal group selected different ecotypes depending on how they fit to environmental conditions at each depth. Overall, this suggests that archaea have preferential niches and that Thaumarchaea of Lake Geneva are most probably adapted to thrive at the three sampled depths and the studied months, compared to other groups. For instance, OTU 2 that contributed to sample dissimilarity also held the highest number of reads for archaea and the second in the data set. This OTU was assigned to Candidatus *Nitrosopumilus* from the family of the *Nitrosopumilaceae* (Thaumarchaeota), a genus reported to be ubiquitous in marine environments and an important ammonia oxidizing archaea.

Bacteria were much more diverse than archaea. Actinobacteria, Cyanobacteria, Proteobacteria, Chloroflexi and Bacteroidetes dominated this community. This result agrees with previous studies which identified most of these groups as predominant phyla of freshwater systems (Glöckner *et al.*, 1999; Debroas *et al.*, 2009; Newton *et al.*, 2011; Zwirglmaier *et al.*, 2015). As for the archaea, depth mainly explained the differences found in the bacterial diversity and the highest values in bacteria richness and diversity were also found at 200 m depth (with a dominance of Actinobacteria). No difference in richness, diversity and evenness was found among the different sampling sites. However, the season also played a role since bacterial OTUs found at 2 and 15 m (but not those at 200 m) were segregated by months. OTU 5 (Oxyphotobacteria) was the most influential OTU to drive the beta diversity within seasons. OTU 4 (Anaerolineae) was more present at 200 m and it drove the beta diversity between 2 and 200 m, and 15 and 200 m. Physicochemical descriptors explained 41% of the bacterial OTUs distribution present at pt4. Total phosphorus explained well the presence of some OTUs at 200 m, a result reminding the study of (Berdjeb *et al.*, 2011) who reported that the bacterial community structure present in the hypo- and epilimion in two other peri-alpine lakes, Bourget and Annecy, was affected mainly by bottom-up factors regardless of the local environmental conditions of these lakes.

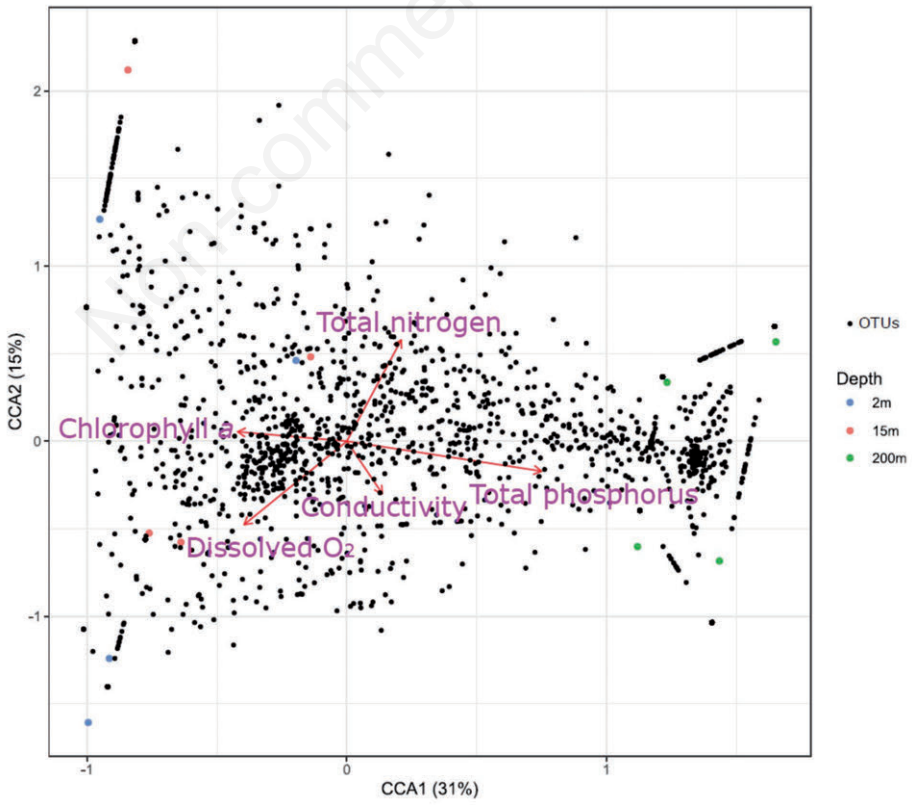

Fig. 5. CCA illustrating the separation of bacterial OTUs and samples based upon physicochemical variables. The CCA was only performed on pt4 site were physicochemical variables were obtained. 12 samples and 1,964 bacterial OTUs were used in this analysis.

Potential relationships between prokaryotes

A variety of relationships was found between the prokaryotes, suggesting potential ecological and/or biogeochemical interactions.

The archaeal OTU 2 assigned to the phylum – class; Thaumarchaeota – Nitrososphaeria co-occurred with 8 bacterial OTUs. Recently, in the release of SILVA SSU database 138 (December 16, 2019), Nitrososphaeria was transferred to the Crenarchaeota phylum. However, the analysis was performed with database 132 and prior to the release of the new database version, explaining why we maintained the former classification. The detection of functional profiles revealed that Nitrososphaeria is less equipped in term of pathway in comparison to associate bacteria. This is expected since archaea possesses a compact genome. In addition, only few KOs were common between Nitrososphaeria and the other bacteria, suggesting that some functions are missing in the archaeal genome. These results explained probably why Nitrososphaeria co-occurred with other bacteria. The archaeal OTU 2 (*i.e.*, *Nitrosopumilus*), mentioned above, co-occurred with the bacterial OTU 25 assigned to *Nitrospira* (Nitrospirae). This relationship is coherent with the finding of (Parada and Fuhrman, 2017) who also reported a correlation between marine Thaumarchaea and ammonia oxidizing bacteria (AOB) such as *Nitrospina*, both responsible for nitrification. Archaeal OTU 2 displayed also a link with the bacterial OTU 19, identified as CL500-3 (*i.e.*, a Planctomycetes). These bacteria were already reported to be associated to AOA such as Thaumarchaeota in oxygenated hypolimnion of deep freshwater lakes (Okazaki *et al.*, 2017). *Nitrosopumilus* is known to operate the methane cycle and produce methylphosphonic acid (Carini *et al.*, 2014). Methane-oxidizing archaea (MOA) and bacteria (MOB) such as α- and γ-proteobacteria (Hanson and Hanson, 1996; Lüke *et al.*, 2010) can either produce or consume methane. These microorganisms are essential for the methane release and rate in the atmos-

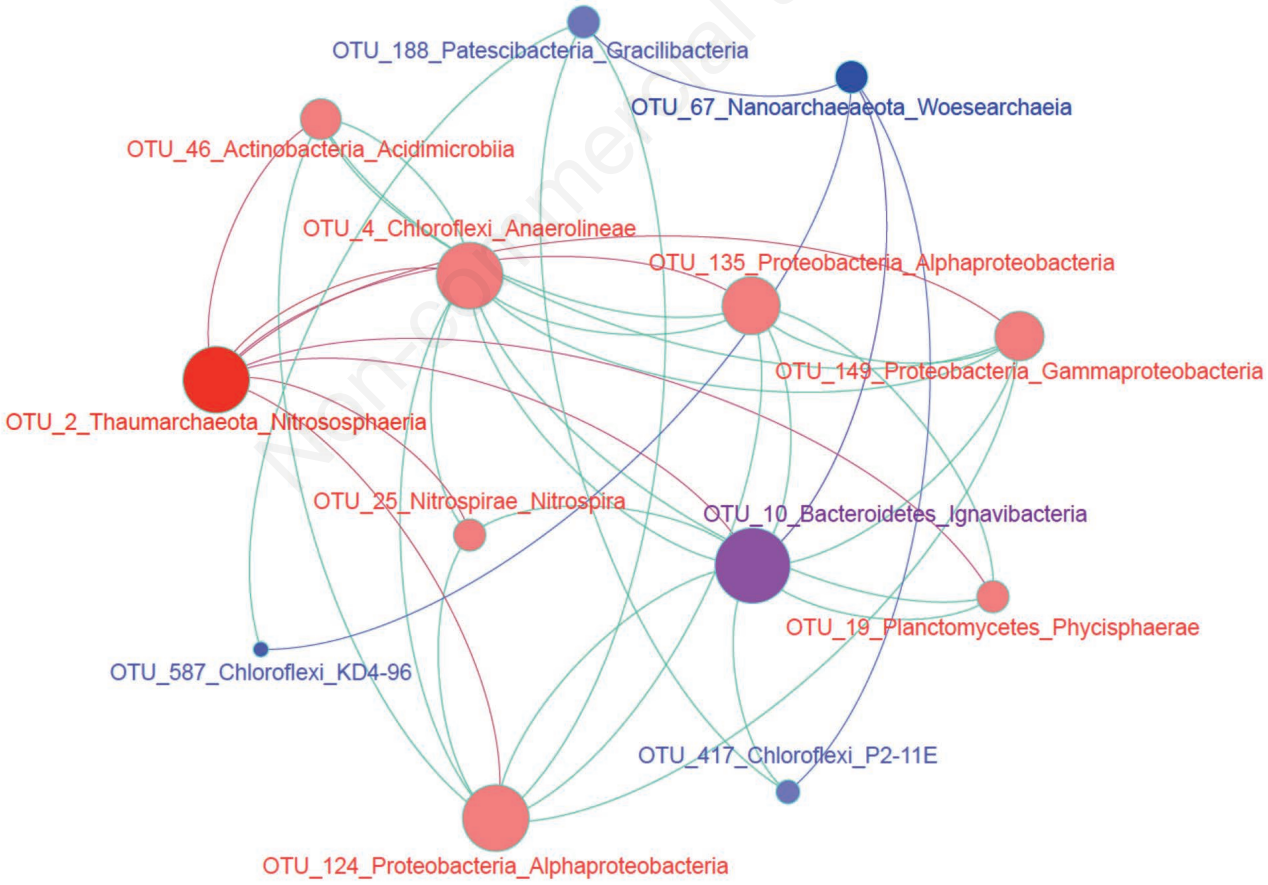

Fig. 6. Co-occurrence network between bacteria and archaea OTUs. Only OTUs that passed the cutoff (correlation coefficient ≥0.9 and $p<0.01$) are illustrated. Only positive relationships are shown. Red and salmon color are the connections from the archaeal OTU 2 to its associate bacteria. While blue color represents the connections between archaeal OTU 67 and its associate bacteria. The purple color indicates that both archaeal OTU shared a connection with a bacterial OTU.

phere. Therefore, the occurrence between the archaeal OTU 2 and the bacterial OTU 4 assigned to *Anaerolineaceae* (Chloroflexi) makes sense, with the possibility that *Anaerolineaceae* provide organic acid such as acetate to other microorganisms like acetoclastic methanogens. Moreover, *Anaerolineaceae* could form a syntrophic cooperation with some archaea involved in methanogenic degradation of alkanes (Lüke *et al.*, 2010; Liang *et al.*, 2015). Similarly, we observed MOB co-occurrence through OTU 149 assigned to the *Methylomonaceae* (Gammaproteobacteria). Furthermore, a mutualistic interaction might be occurring between archaea OTU 2 and bacteria OTU 124 assigned to SAR11 clade (Alphaproteobacteria) (Parada and Fuhrman, 2017). A link between some Thaumarchaea and SAR groups such as SAR406 and SAR86 was also found in the ocean. In fact, *Nitrosopumilus* produce methylphosphonic acid (MPn), which is decomposed by phosphate-starved microorganisms such as SAR11 clade (chemoheterotrophic bacteria) that produce CH_4 from MPn when P_i (inorganic phosphate) starved (Carini *et al.*, 2014). This was also suggested by the functional profile where Nitrososphaeria, Alpha and Gammaproteobacteria shared 526 to 536 from 736 KOs pointing out that the bacterial class lack some pathways that can be compensated with an interaction with Nitrososphaeria. Finally, the co-occurrence with OTU 10 assigned to BSV26 (Kryptoniales) and OTU 46 identified as IMCC26256 (Acidimicrobiia) could be ecologically relevant since these two OTUs were originally isolated together from extremely acidic environments or geothermal sites were archaea blossom (Ludwig *et al.*, 2012; Hu *et al.*, 2017).

The second most important archaeal OTU (*i.e.*, OTU 67) showed important relationships with bacterial OTUs 10, 188, 417 and 587. OTU 67 was assigned to the class Woesearchaeia of the phylum Nanoarchaeaeota. Woesearchaeota are widely spread in diverse environments and (Liu *et al.*, 2018) reported a syntrophic relationship between Woesearchaeota and other methanogenic archaea. This syntrophic and/or mutualistic partnership emerged from deficiencies in metabolic pathways within the Woesearchaeota class. (Liu *et al.*, 2018) also suspected that Woesearchaeota develop a partnership with some bacterial methanogens, proposing a Woesearchaeota-methanogens consortium, likely to influence methane formation. However, these authors lacked reliable bacterial data to analyze collectively these microorganisms and prove this assumption. Our results support their finding. A relationship also existed with Chloroflexi OTU 417 (P2-11E), 587 (KD4-96) and to JG30.KF.CM66 (Supplementary Fig. S9). In fact, Chloroflexi possess a wide diversity of metabolisms (*e.g.*, aerobic respiration, denitrification and phototrophy), and ecological roles, but are best known as photoheterotrophs (Ward *et al.*, 2018). They are reported to be dominant and abundant in methanogenic reactors (Bovio *et al.*, 2019). Recently, (Ward *et al.*, 2019) reported that within the Chloroflexi phylum some of them such as OHK40 are capable of methane oxidation (photomethanotrophy). We cannot pinpoint the exact need or relationship between OTU 67 and the other bacterial OTUs but a syntrophic or/and a mutualistic link is very probable between them. This could explain the abundance of Chloroflexi observed at 200 m (Fig. 1) where oxygen concertation was found to be lower than at the surface (*i.e.*, 7 *vs* 9-10 mg L^{-1}). Indeed, if some Chloroflexi OTUs can produce methane they are likely to occur in less oxic layers since methanogenesis is not favored where oxygen is abundant (Carini *et al.*, 2014), and, as pointed out by (Mayr *et al.*, 2019), MOB can occupy different niches according to oxygen and methane concentration. At last, the co-occurrence observed with OTU 188 assigned to *Peribacteria* (Gracilibacteria) also make sense since the latter lack some pathways and is predicted to be either a symbiont or closely dependent on other community members for key building blocks (Sieber *et al.*, 2019).

CONCLUSIONS

In parallel to studies dealing with physiology, biodiversity, biological indication and assessment, assessing interactions between microorganisms in pelagic communities are of great ecological interest but still remain poorly documented in lakes. Using Illumina HiSeq combined with 16S metabarcoding *via* universal primer, co-occurrence networks and functional profile predictions, we showed that Lake Geneva is booming with connections, and our results suggest that some OTUs co-occur for scavenging or sharing similar ecological niches for mutual benefits. Sequencing different functional genes and conducting experiments with isolated organisms could help to better understand the relationship between key "species" and biotic interactions sustaining the microbial ecosystem functioning.

ACKNOWLEDGMENTS

We are grateful to the editor and the reviewer for helping us to improve the article. We would like to thank Pascal Perney, Philippe Quetin, and Jean-Christophe Hustache for sampling. We are also thankful to Louis Jacas, Cécile Chardon, Eric Capo and Gaétan Maechler for technical assistance during the project. We thank Julie Gueguen, Valentin Vasselon, Viet Tran-Khac and François Kek for bioinformatics and statistical help. Finally, special thanks are attributed to Dr. Se-ran Jun for providing some help to generate KOs from the OTUs tables.

Corresponding author: stephan.jacquet@inrae.fr

Keywords: Bacteria; Archaea; Lake Geneva; co-occurrence; diversity; functional profiles.

Authors contributions: SJ, designed the study and collected the samples; JE, performed bioinformatics and statistical analysis; JE, SJ, analyzed and interpreted the results; JE, SJ wrote the manuscript with revisions by YD.

Conflict of interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 May 2020.
Accepted: 18 August 2020.

Licensee PAGEPress, Italy
Advances in Oceanography and Limnology, 2020; 11:9099
DOI: 10.4081/aiol.2020.9099

REFERENCES

Alonso-Sáez L, Waller AS, Mende DR, Bakker K, Farnelid H, Yager PL, Lovejoy C, Tremblay JÉ, Potvin M, Heinrich F, Estrada M, Riemnann L, Bork P, Pedros-Alio C, Bertilsson S, 2012. Role for urea in nitrification by polar marine Archaea. Proc. Natl. Acad. Sci. U. S. A. 109: 17989–17994.

Angly FE, Dennis PG, Skarshewski A, Vanwonterghem I, Hugenholtz P, Tyson GW, 2014. CopyRighter: A rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. Microbiome 2:1–13.

Anneville O, Souissi S, Ibanez F, Ginot V, Druart JC, Angeli N, 2002. Temporal mapping of phytoplankton assemblages in Lake Geneva: Annual and interannual changes in their patterns of succession. Limnol. Oceanogr. 47: 1355–1366.

Bahram M, Anslan S, Hildebrand F, Bork P, Tedersoo L, 2019. Newly designed 16S rRNA metabarcoding primers amplify diverse and novel archaeal taxa from the environment. Environ. Microbiol. Rep. 11: 487–494.

Bastian M, Heymann S, Jacomy M, 2009. Gephi: An open source software for exploring and manipulating networks. BT - International AAAI Conference on Weblogs and Social. Int. AAAI Conf. Weblogs Soc. Media: 361–362.

Berdjeb L, Ghiglione JF, Jacquet S, 2011. Bottom-up versus top-down control of hypo-and epilimnion free-living bacterial community structures in two neighboring freshwater lakes. Appl. Environ. Microbiol. 77: 3591–3599.

Berdjeb L, Pollet T, Chardon C, Jacquet S, 2013. Spatio-temporal changes in the structure of archaeal communities in two deep freshwater lakes. FEMS Microbiol. Ecol. 86: 215–230.

Bomberg M, Montonen L, Münster U, Jurgens G, 2008. Diversity and function of archaea in freshwater habitats. Curr. Trends Microbiol. 4: 61–89.

Bovio P, Cabezas A, Etchebehere C, 2019. Preliminary analysis of Chloroflexi populations in full-scale UASB methanogenic reactors. J. Appl. Microbiol. 126: 667–683.

Braga RM, Dourado MN, Araújo WL, 2016. Microbial interactions: ecology in a molecular perspective. Brazilian J. Microbiol. 47: 86–98.

Carini P, White AE, Campbell EO, Giovannoni SJ, 2014. Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. Nat. Commun. 5: 1–7.

Casamayor EO, 2017. Towards a microbial conservation perspective in high mountain lakes, p. 157–180. In: Catalan J, Ninot JM and Aniz MM (eds). High mountain conservation in a changing world. Springer, Cham.

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

Debroas D, Humbert JF, Enault F, Bronner G, Faubladier M, Cornillot E, 2009. Metagenomic approach studying the taxonomic and functional diversity of the bacterial community in a mesotrophic lake (Lac du Bourget - France). Environ. Microbiol. 11: 2412–2424.

Dekas AE, Poretsky RS, Orphan VJ, 2009. Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. Sci. Mag. 326: 422–427.

Dinno A, 2017. Dunn's Test of Multiple Comparisons Using Rank Sums.

Eme L, Spang A, Lombard J, Stairs CW, Ettema TJG, 2017. Archaea and the origin of eukaryotes. Nat. Rev. Microbiol. 15: 711–723.

Evans PN, Parks DH, Chadwick GL, Robbins SJ, Orphan VJ, Golding SD, Tyson GW, 2015. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. Science (80-.). 350: 434–438.

Ezzedine JA, Chardon C, Jacquet S, 2020. New 16S rRNA primers to uncover Bdellovibrio and like organisms diversity and abundance. J. Microbiol. Methods 175: 105996.

Fuhrman JA, Davis AA, 1997. Widespread Archaea and novel Bacteria from the deep sea as shown by 16S rRNA gene sequences. Mar. Ecol. Prog. Ser. 150: 275–285.

Glöckner FO, Fuchs BM, Amann R, 1999. Bacterioplankton compositions of lakes and oceans: A first comparison based on fluorescence in situ hybridization. Appl. Environ. Microbiol. 65: 3721–3726.

Goral F, Schellenberg J, 2017. goeveg: functions for community data and ordinations.

Gotelli N, Hart E, Ellison A, 2015. EcoSimR: Null Model Analysis for Ecological Data.

Haller L, Tonolla M, Zopfi J, Peduzzi R, Wildi W, Poté J, 2011. Composition of bacterial and archaeal communities in freshwater sediments with different contamination levels (Lake Geneva, Switzerland). Water Res. 45: 1213–1228.

Hanson RS, Hanson TE, 1996. Methanotrophic bacteria. Microbiol. Rev. 60: 439–471.

Hu A, Ju F, Hou L, Li J, Yang X, Wang H, Mulla SI, Sun Q, Burgmann H, Yu C-P, 2017. Strong impact of anthropogenic contamination on the co-occurrence patterns of a riverine microbial community. Environ. Microbiol. 19: 4993–5009.

Hu D, Cha G, Gao B, 2018. A phylogenomic and molecular markers based analysis of the class Acidimicrobiia. Front. Microbiol. 9: 1–12.

Hugoni M, Etien S, Bourges A, Lepère C, Domaizon I, Mallet

C, Bronner G, Debroas D, Mary I, 2013. Dynamics of ammonia-oxidizing Archaea and Bacteria in contrasted freshwater ecosystems. Res. Microbiol. 164: 360–370.

Jacquet S, Domaizon I, Anneville O, 2014. The need for ecological monitoring of freshwaters in a changing world: A case study of Lakes Annecy, Bourget, and Geneva. Environ. Monit. Assess. 186: 3455–3476.

Ju F, Zhang T, 2015. Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant. ISME J. 9: 683–695.

Ju F, Xia Y, Guo F, Wang Z, Zhang T, 2014. Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. Environ. Microbiol. 16: 2421–2432.

Jun SR, Robeson MS, Hauser LJ, Schadt CW, Gorin AA, 2015. PanFP: Pangenome-based functional profiles for microbial communities. BMC Res. Notes 8: 1–7.

Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M, 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44: D457–D462.

Koizumi Y, Takii S, Fukui M, 2004. Depth-related change in archaeal community structure in a freshwater lake sediment as determined with denaturing gradient gel electrophoresis of amplified 16S rRNA genes and reversely transcribed rRNA fragments. FEMS Microbiol. Ecol. 48: 285–292.

Könneke M, Bernhard AE, De La Torre JR, Walker CB, Waterbury JB, Stahl DA, 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature 437: 543–546.

Koonin E V., Wolf YI, 2008. Genomics of bacteria and archaea: The emerging dynamic view of the prokaryotic world. Nucleic Acids Res. 36: 6688–6719.

Kozubal MA, Romine M, Jennings RD, Jay ZJ, Tringe SG, Rusch DB, Beam JP, McCue LA, Inskeep WP, 2013. Geoarchaeota: A new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone National Park. ISME J. 7: 622–634.

Liang B, Wang LY, Mbadinga SM, Liu JF, Yang SZ, Gu JD, Mu BZ, 2015. Anaerolineaceae and Methanosaeta turned to be the dominant microorganisms in alkanes-dependent methanogenic culture after long-term of incubation. AMB Express 5:37.

Linz A, 2018. OTUtable: North Temperate Lakes - Microbial Observatory 16S Time Series Data and Functions. Available from: https://rdrr.io/cran/OTUtable/

Liu X, Li M, Castelle CJ, Probst AJ, Zhou Z, Pan J, Liu Y, Banfield JF, Gu JD, 2018. Insights into the ecology, evolution, and metabolism of the widespread Woesearchaeotal lineages. Microbiome 6: 1–16.

Ludwig W, Euzéby J, Schumann P, Busse H-J, Trujillo ME, Kämpfer P, Whitman WB, 2012. Road map of the phylum Actinobacteria. Bergey's Manual® Syst. Bacteriol.: 1–28.

Lüke C, Krause S, Cavigiolo S, Greppi D, Lupotto E, Frenzel P, 2010. Biogeography of wetland rice methanotrophs. Environ. Microbiol. 12: 862–872.

Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M, 2015. Swarmv2: Highly-scalable and high-resolution amplicon clustering. PeerJ 2015: 1–12.

Martin M, 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 1–3.

Mayr MJ, Zimmermann M, Guggenheim C, Brand A, Bürgmann H, 2019. Niche partitioning of methane-oxidizing bacteria along the oxygen–methane counter gradient of stratified lakes. ISME J.: 274–287.

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S, 2011. A guide to the natural history of freshwater lake bacteria. Microbiol. Mol. Biol. Rev. 75:14–49.

Okazaki Y, Fujinaga S, Tanaka A, Kohzu A, Oyagi H, Nakano SI, 2017. Ubiquity and quantitative significance of bacterioplankton lineages inhabiting the oxygenated hypolimnion of deep freshwater lakes. ISME J. 11: 2279–2293.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H, 2019. Package 'vegan' Title Community Ecology Package. Community Ecol. Packag. 2: 1–297.

Ortiz-Alvarez R, Casamayor EO, 2016. High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. Environ. Microbiol. Rep. 8: 210–217.

Pacheco AR, Segrè D, 2019. A multidimensional perspective on microbial interactions. FEMS Microbiol. Lett. 366: 1–11.

Parada AE, Fuhrman JA, 2017. Marine archaeal dynamics and interactions with the microbial community over 5 years from surface to seafloor. ISME J. 11: 2510–2525.

Pollet T, Berdjeb L, Chardon C, Jacquet S, 2018. Contrasting temporal patterns in ammonia-oxidizing archaeal community dynamics in two peri-alpine lakes with different trophic status. Aquat. Microb. Ecol. 81: 95–108.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Glo FO, Yarza P, 2013. The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. Nucleic Acids Res. 41: 590–596.

R Core Team, 2019. R: A language and environment for statistical computing (v. 3.5.0). R Foundation for Statistical Computing, Wien, Austria.

Robertson CE, Harris JK, Spear JR, Pace NR, 2005. Phylogenetic diversity and ecology of environmental Archaea. Curr. Opin. Microbiol. 8: 638–642.

Rognes T, Flouri T, Nichols B, Quince C, Mahé F, 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4: 1–22.

Schloss PD, Handelsman J, 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. 71:1501–1506.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF, 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75: 7537–7541.

Seyler LM, Tuorto S, McGuinness LR, Gong D, Kerkhof LJ, 2019. Bacterial and Archaeal Specific-Predation in the North Atlantic Basin. Front. Mar. Sci. 6: 1–10.

Sieber CMK, Paul BG, Castelle CJ, Hu P, Tringe SG, Valentine DL, Anderson GL, Banfield JF, 2019. Unusual metabolism and hypervariation in the genome of a Gracilibacteria (BD1-5) from an oil degrading community. bioRxiv. doi: 10.1101/595074.

Torondel B, Ensink JHJ, Gundogdu O, Ijaz UZ, Parkhill J, Abdelahi F, Nguyen VA, Sudgen S, Gibson W, Walker AW, Quince C, 2016. Assessment of the influence of intrinsic environmental and geographical factors on the bacterial ecology of pit latrines. Microb. Biotechnol. 9:209–223.

Vanwonterghem I, Evans PN, Parks DH, Jensen PD, Woodcroft BJ, Hugenholtz P, Tyson GW, 2016. Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. Nat. Microbiol. 1: 1–9.

Varela MM, Van Aken HM, Sintes E, Herndl GJ, 2008. Latitudinal trends of Crenarchaeota and Bacteria in the meso- and bathypelagic water masses of the Eastern North Atlantic. Environ. Microbiol. 10: 110–124.

Vuillemin A, Wankel SD, Coskun ÖK, Magritsch T, Vargas S, Estes ER, Spivack AJ, Smith DC, Pockalny R, Murray RW, D'Hondt S, Orsi WD, 2019. Archaea dominate oxic subseafloor communities over multimillion-year time scales. Sci. Adv. 5: 1–12.

Wang F, Men X, Zhang G, Liang K, Xin Y, Wang J, Li A, Zhang H, Liu H, Wu L, 2018. Assessment of 16S rRNA gene primers for studying bacterial community structure and function of aging flue-cured tobaccos. AMB Express 8: 1–9.

Wang Y, Qian PY, 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One 4:e7401.

Ward L, Shih PM, Hemp J, Kakegawa T, Fischer WW, McGlynn SE, 2019. Phototrophic Methane Oxidation in a Member of the Chloroflexi Phylum. bioRxiv: 531582. doi: 10.1101/531582.

Ward LM, Hemp J, Shih PM, McGlynn SE, Fischer WW, 2018. Evolution of phototrophy in the Chloroflexi phylum driven by horizontal gene transfer. Front. Microbiol. 9: 1–16.

Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, Woo K, 2018. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. Available from: https://rdrr.io/cran/ggplot2/

Wurzbacher C, Fuchs A, Attermeyer K, Frindte K, Grossart HP, Hupfer M, Casper P, Monaghan MT, 2017. Shifts among eukaryota, bacteria, and archaea define the vertical organization of a lake sediment. Microbiome 5: 1–16.

Zwirglmaier K, Keiz K, Engel M, Geist J, Raeder U, 2015. Seasonal and spatial patterns of microbial diversity along a trophic gradient in the interconnected lakes of the Osterseen Lake District, Bavaria. Front. Microbiol. 6: 1–18.

ADVANCES IN OCEANOGRAPHY AND LIMNOLOGY

DOI: 10.4081/aiol.2020.9099

SUPPLEMENTARY MATERIAL

Exploring archaeal and bacterial diversity and co-occurrence in Lake Geneva

Jade A. Ezzedine,[1] Yves Desdevises,[2] Stéphan Jacquet[1*]

[1]Université Savoie Mont-Blanc, INRAE, UMR CARRTEL, Thonon-les-Bains
[2]CNRS, Biologie Intégrative des Organismes Marins, Observatoire Océanologique, [2]Sorbonne Université, F-66650 Banyuls-sur-Mer, France

***Corresponding author**: stephan.jacquet@inrae.fr

Sequencing results

Following sequencing, each raw data file (R1 forward and R2 reverse sequence) contained 55,679,272 reads. After merging, we obtained 49,545,789 paired end reads. At this step, 11% of reads were lost, and the median read length was 437 bp. After demultiplexing, and after a first dereplication, the number of reads reached 11,079,438, with a median length of 377 pb. The second dereplication of reads after reuniting them in one file gave 6,861,908 reads. The clustering process yielded 127,052 representative sequences, including 84.3% chimeric sequences and 15.3% non-chimeric sequences. However, when considering abundance into account, this corresponded to 1.3% chimeras and 98.7% non-chimeras. The median read length after the clustering was 376 pb.

For the archaeal OTU table, we obtained with the default filter, 631 OTUs and 2,868,729 reads. After applying the stringent filter, this number fell down to 194 OTUs and 2,862,662 reads. In addition, only 112 OTUs (2,858,335 reads) were common to both replicates. After applying all these filters, we only lost 0.4% of archaeal reads. Overall, 5 phyla and 7 classes were detected after taxonomic assignment. Woesearchaeia (Nanoarchaeaeota) had the highest number of OTUs (*i.e.,* 91), but was classed second regarding the number of reads that was 67,646. Nitrososphaeria

(Thaumarchaeota) followed with 11 OTUs, corresponding however to 2,790,583 reads. For Methanomicrobia (Euryarchaeota), Iainarchaeia (Diapherotrites), Bathyarchaeia (Crenarchaeota), Methanobacteria (Euryarchaeota), and Micrarchaeia (Diapherotrites) the number of OTUs varied between 1 and 3, with 3 to 43 reads (Supplementary Fig. S2 and S3).

For the bacterial OTU table, the "default" step gave 7,987 bacterial OTUs and 29,846,037 reads. After applying the "stringent" step we obtained 3,712 OTUs and 29,502,677 reads. This step resulted in a loss of only 1.15% in read number. The "shared" filter yielded 2,674 OTUs and 29,383,517 reads. Only 0.4% of reads were lost between these two steps, for a total loss of 1.55%. In general, 41 bacterial phyla and 110 classes were assigned according to the pipeline. Alphaproteobacteria (Proteobacteria) yield the highest number of OTUs (*i.e.,* 411) (Supplementary Fig. S6), followed by Bacteroidetes (Bacteroidia) (*i.e.,* 400). The Gammaproteobacteria (Proteobacteria) class followed with 381 OTUs, the Deltaproteobacteria (Proteobacteria) with 223 OTUs, the Planctomycetacia (Planctomycetes) with 122 OTUs, and the Oxyphotobacteria (Cyanobacteria) with 110 OTUs. Actinobacteria was characterized by the highest number of reads, *i.e.*, 8,948,606. It was followed by the Oxyphotobacteria (Cyanobacteria) with 4,699,878 reads, the Gammaproteobacteria (Proteobacteria) with 2,220,792 reads, the Anaerolineae (Chloroflexi) with 2,209,038 reads, the Bacteroidetes (Bacteroidia) with 2,076,034 reads and the Alphaproteobacteria (Proteobacteria) with 1,955,836 reads. 48 bacterial classes had less than 1,000 reads (Supplementary Fig. S7).

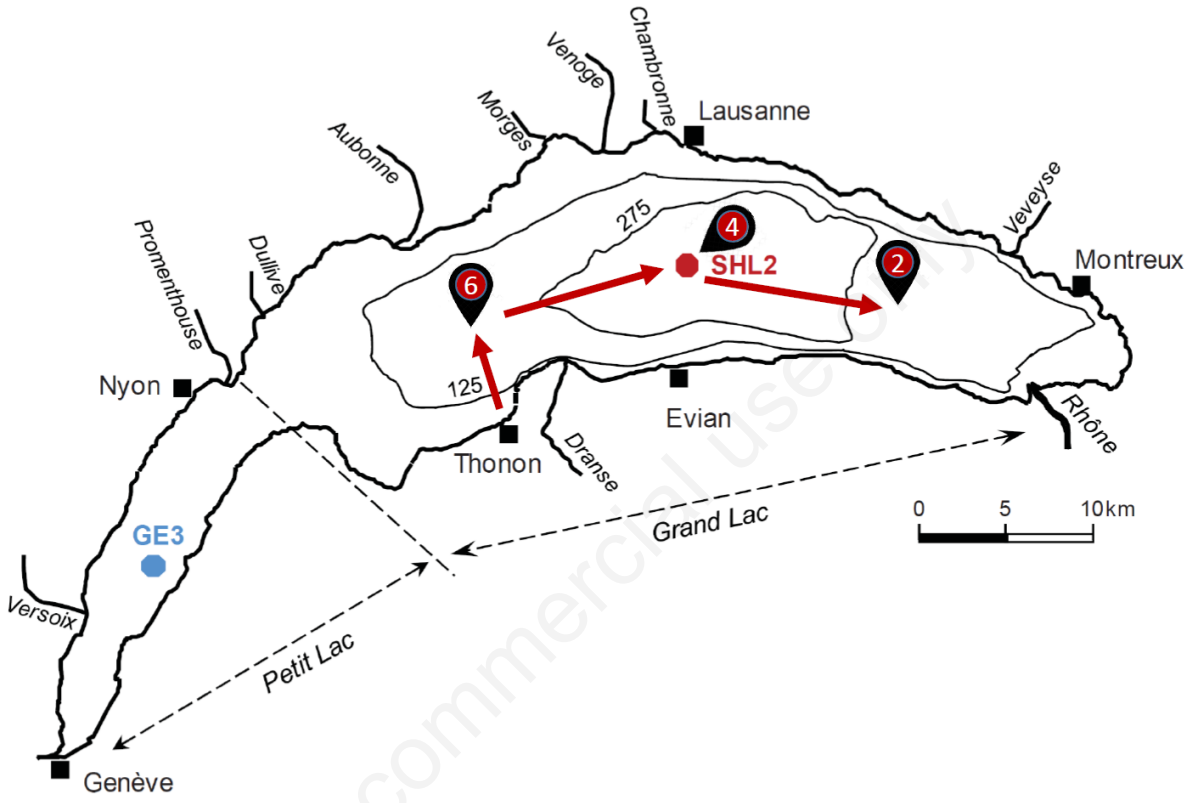

Fig. S1. Map of Lake Geneva and coordinates of the different sampling points selected during TRANSLEM: Site 2 (N 46° 26.206 / E 006° 46.848), Site 4-SHL2 (N 46° 27.207 / E 006° 35.654) and Site 6 (N 46° 25.061 / E 006° 24.957).

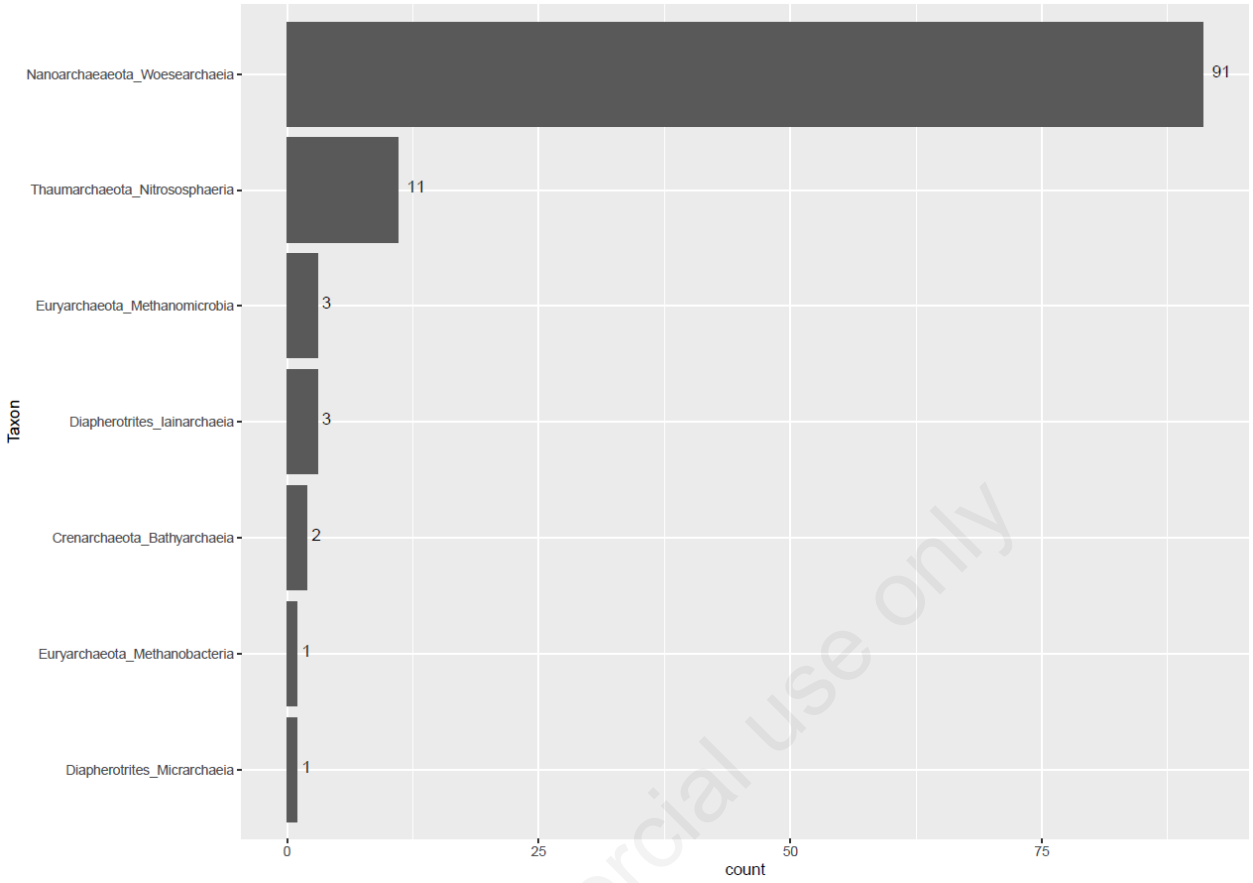

Fig. S2. Number of OTUs per archaeal class.

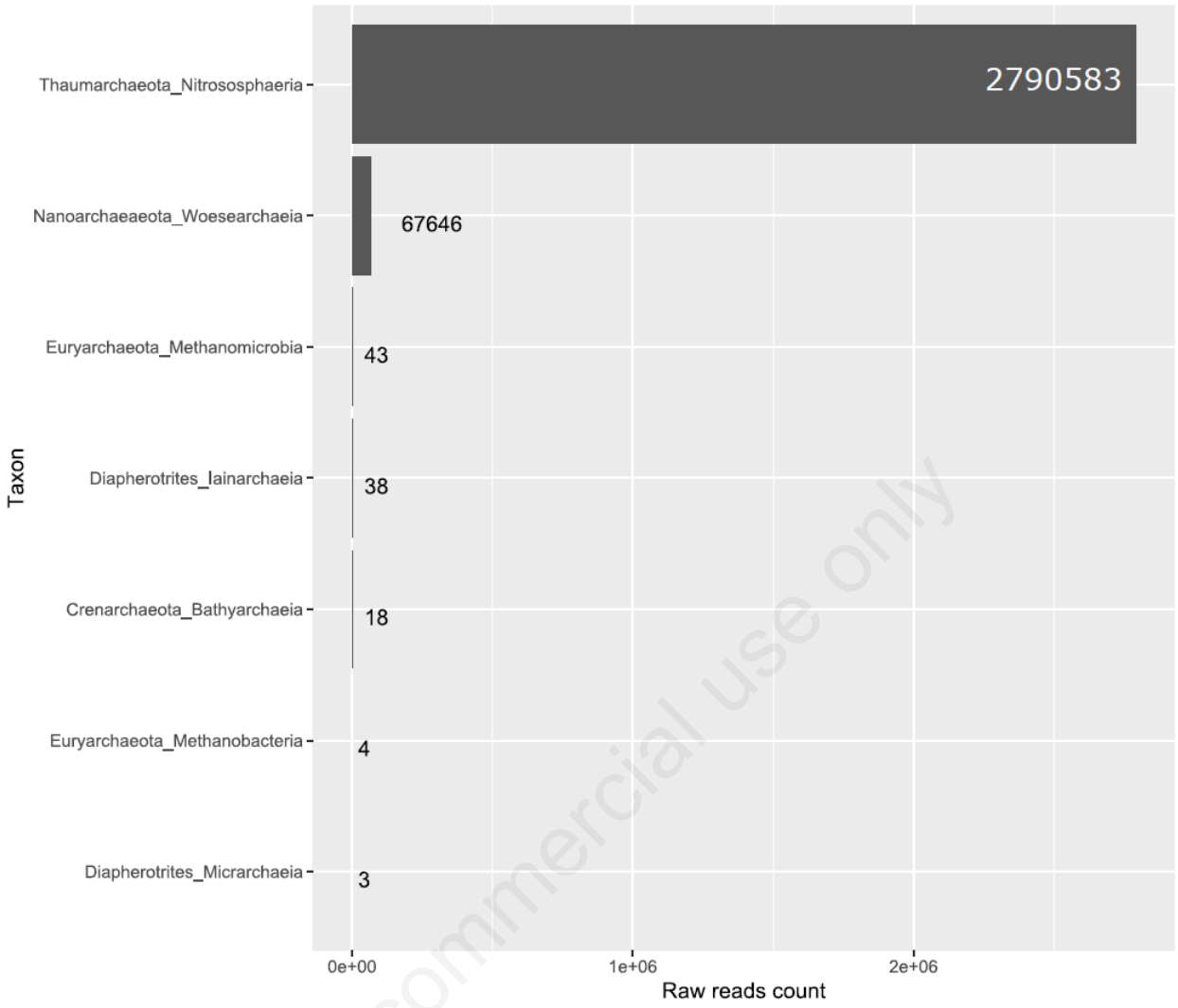

Fig. S3. Number of raw reads per archaeal class.

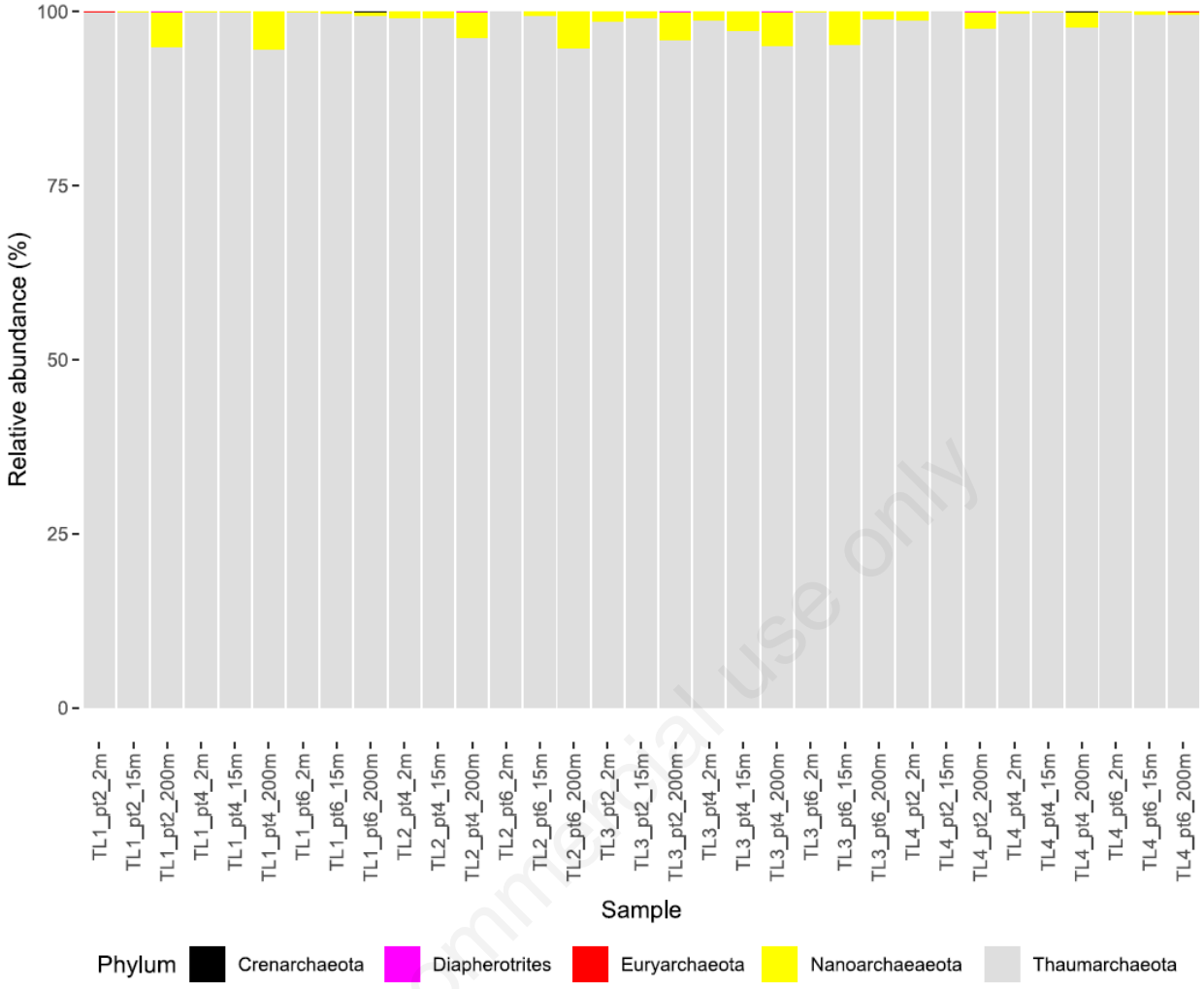

Fig. S4. Stacked histogram of archaeal phylum. Phylum abundances are represented by reads relative abundance per sample. Thaumarchaeota is dominant in all samples.

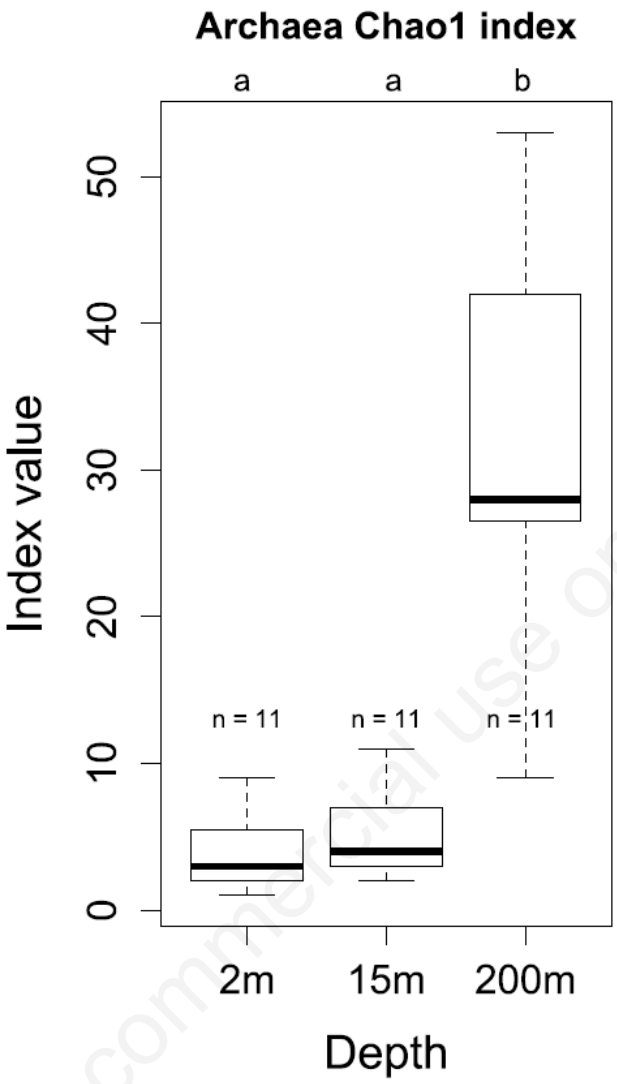

Fig. S5. Chao1 richness index for archaea showing significant differences between depths. N, number of samples; a, b, significant differences.

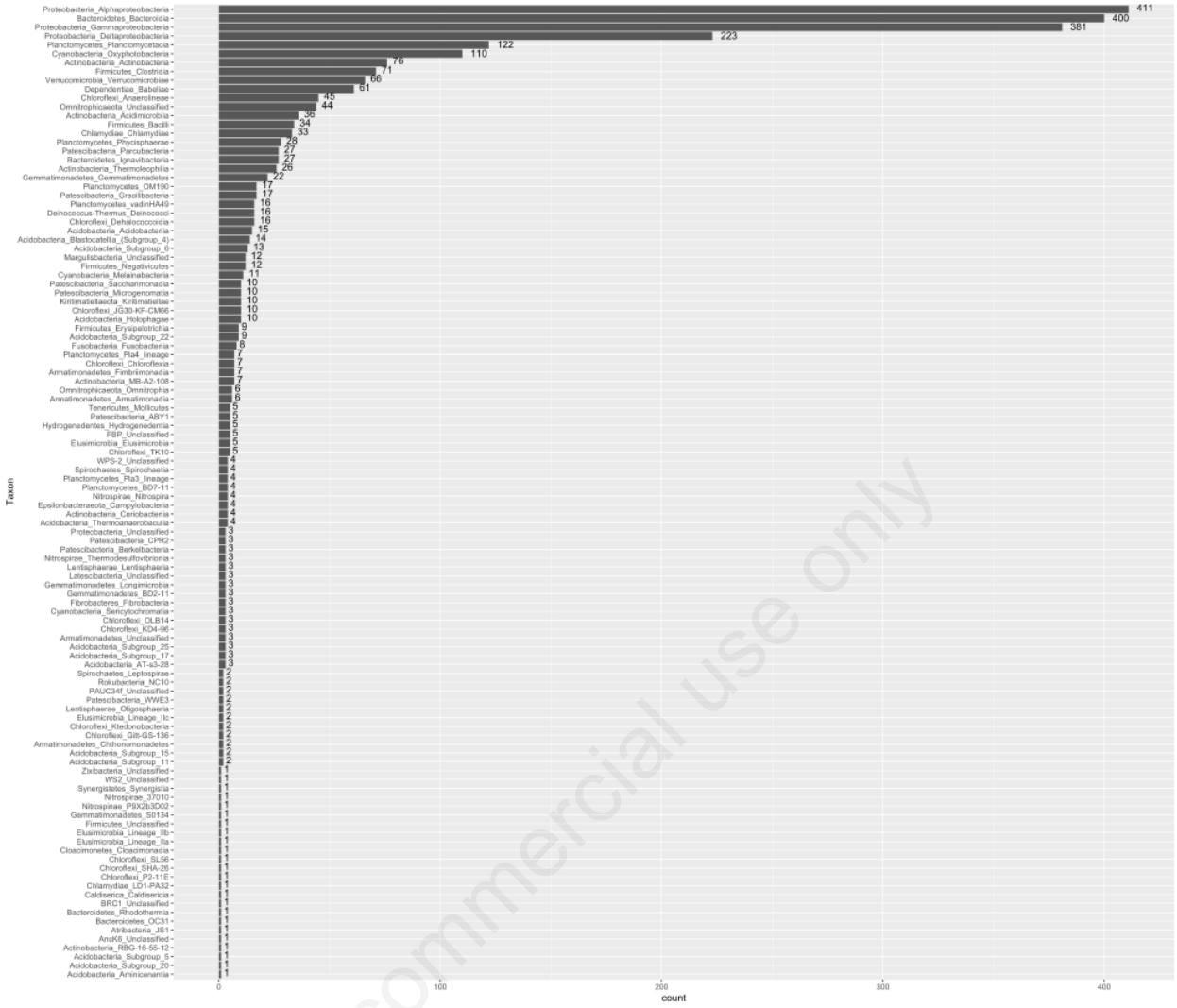

Fig. S6. Number of OTUs per bacterial class.

Fig. S7. Number of raw reads per bacterial class.

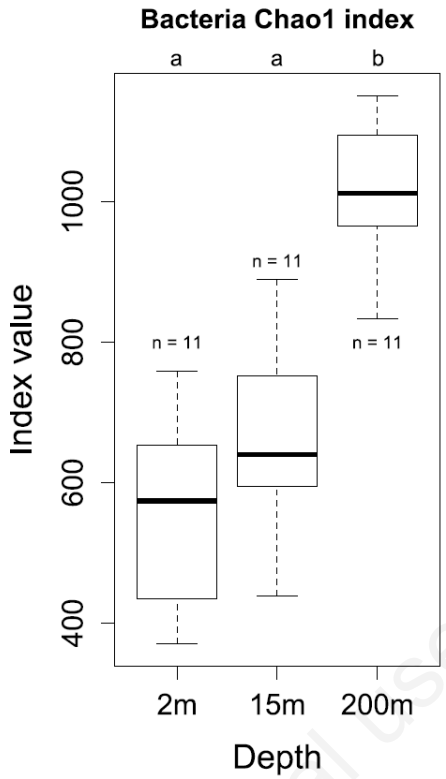

Fig. S8. Chao1 richness index for bacteria showing significant differences between depths. N, number of samples; a and b show where the differences are.

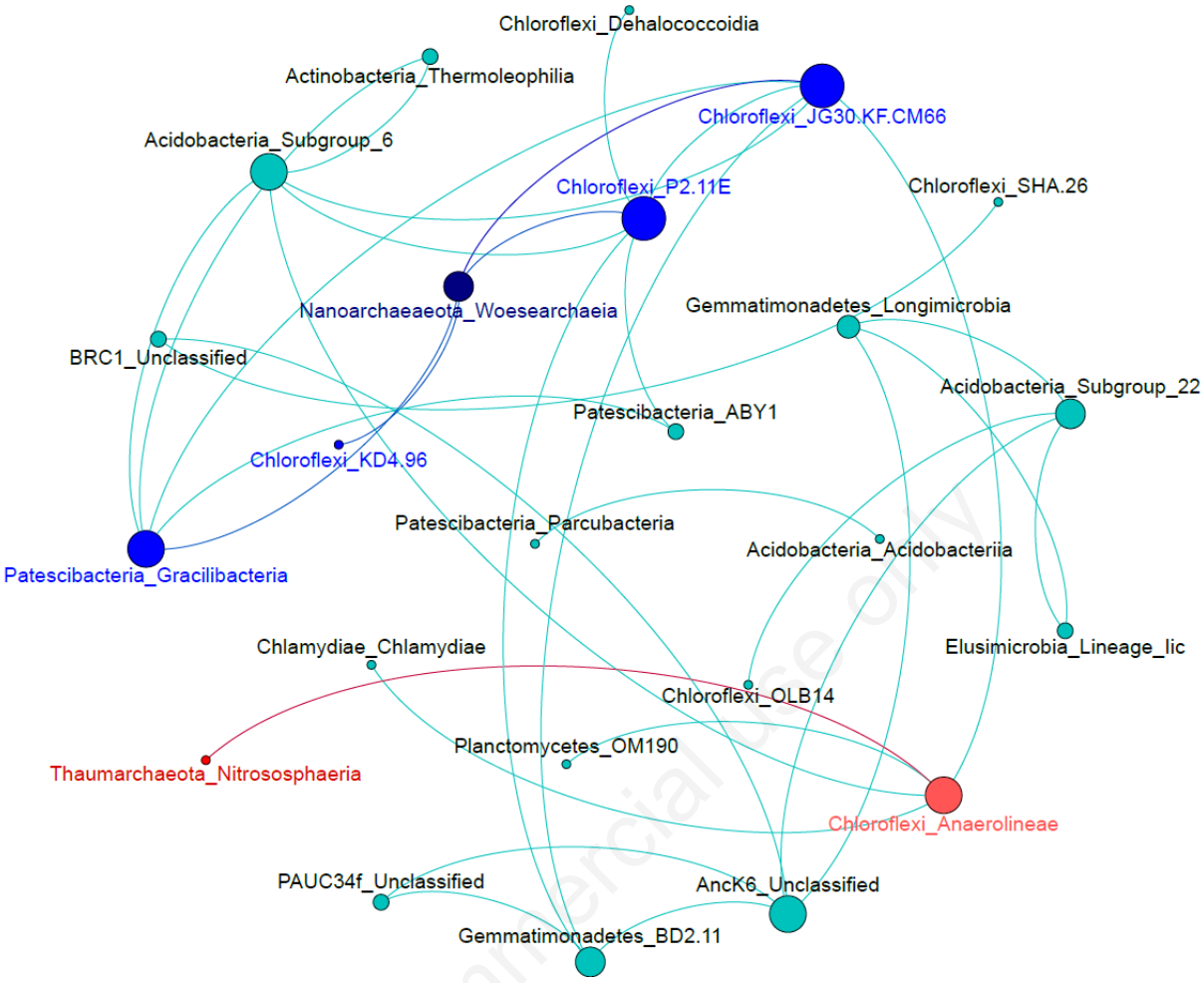

Fig. S9. Co-occurrence network between archaeal and bacterial phyla. Only phylum that passed the cutoff (correlation coefficient ≥0.9 and p<0.01) are shown.

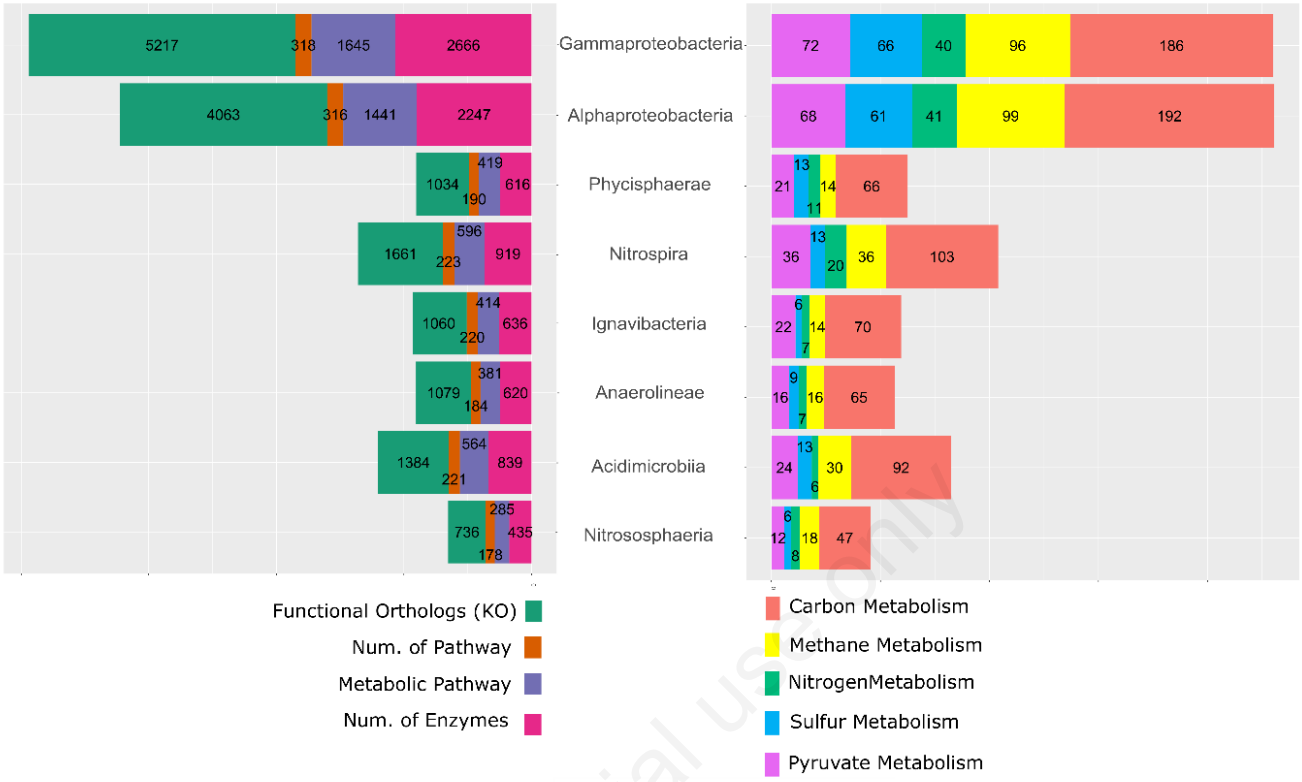

Fig. S10. Functional prediction based on PanFP software of the archaeal OTU 2 Nitrososphaeria and its bacteria associates showed in the co-occurrence network (Fig. 6). The archaea Nitrososphaeria is shown to possess less functional orthologs and pathway than its bacterial associates except for the nitrogen and methane metabolism.

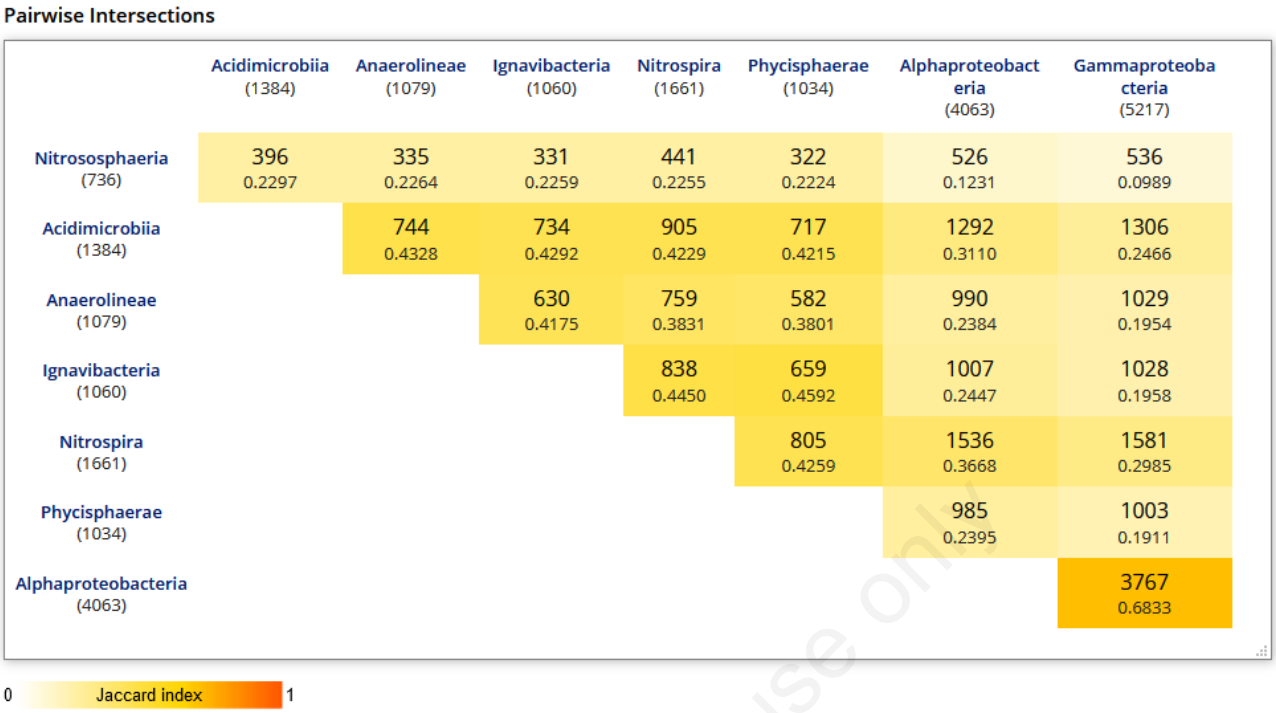

Fig. S11. Pairwise intersection comparison between the numbers of KOs generated *via* PanFP software. The presented class are from the co-occurrence network of Fig. 6.

Tab. S1. Results of the statistical analyses on richness and diversity indices for archaeal OTUs table.

Chao1	Site	Month	Depth
Kruskal p-value	0.5268	0.3121	1.859^{-05} ***
Dunn test			2 m ≠ 200 m (0.0000) 15 m ≠ 200 m (0.0006) 2 m = 15 m (0.5)
Pielou	Site	Month	Depth
Kruskal p-value	0.4466	0.01758*	0.6221
Dunn test		TL1 ≠ TL3 (0.0084*)	
Shannon	Site	Month	Depth
Kruskal p-value	0.3802	0.1672	0.006761
Dunn test			2 m ≠ 200 m (0.0028) 15 m = 200 m (0.0605) 2 m = 15 m (0.4)
Simpson/ InvSimpson	Site	Month	Depth
Kruskal p-value	0.3579	0.1792	0.006761***
Dunn			2 m ≠ 200 m (0.0028) 15 m = 200 m (0.0605) 2 m = 15 m (0.4)

Tab. S2- Anosim and Adonis results performed on the Bray-Curtis dissimilarity matrix generated from the archaeal OTUs table in order to validate the NMDS graph.

	Site	Month	Depth
Anosim p-value	0.227	0.137	0.001***
Adonis p-value	0.592	0.26	0.003***

Tab. S3- Simper test results and significance test analysis for archaeal OTUs table.

Group		Most influential OTU	Kruskal-Wallis (p-value)	Simper cumulative contributions	Average abundances in each compared treatment	Average dissimilarity between the two treatment
Month	February – June	2_Thaumarchaeota_Nitrososphaeria	0.1792	48%	February > June	2%
	February – August	2_Thaumarchaeota_Nitrososphaeria	0.1792	45%	February > August	3%
	February - November	2_Thaumarchaeota_Nitrososphaeria	0.1792	46%	February < November	2%
	June - August	2_Thaumarchaeota_Nitrososphaeria	0.1792	43%	June > August	3%
	June - November	2_Thaumarchaeota_Nitrososphaeria	0.1792	45%	June < November	2%
	August - November	2_Thaumarchaeota_Nitrososphaeria	0.1792	43%	August < November	3%
Site	pt2 - pt4	2_Thaumarchaeota_Nitrososphaeria	0.3579	43%	pt2 = pt4	3%
	pt2 – pt6	2_Thaumarchaeota_Nitrososphaeria	0.3579	45%	pt2 > pt6	2%
	pt4 – pt6	2_Thaumarchaeota_Nitrososphaeria	0.3579	46%	pt4 > pt6	2%
Depth	2m – 15m	2_Thaumarchaeota_Nitrososphaeria	0.006761***	45%	2m > 15m	2%
	2m – 200m	2_Thaumarchaeota_Nitrososphaeria	0.006761***	47%	2m > 200m	3%
	15m – 200m	2_Thaumarchaeota_Nitrososphaeria	0.006761***	40%	15m > 200m	3%

Tab. S4. Results of the statistical analyses on richness and diversity indices for bacterial OTUs table.

Chao1	Site	Month	Depth
Anova p-value	0.619	0.5945	7.71^{-10}***
Tukey multiple comparisons test			2 m ≠ 200 m (0.000000) 15 m ≠ 200 m (0.0000003) 2 m = 15 m (0.07)
Pielou	Site	Month	Depth
Anova p-value	0.8347	0.4889	4.706^{-06} ***
Tukey multiple comparisons test			2 m ≠ 200 m (0.000004) 15 m ≠ 200 m (0.0006579) 2 m = 15 m
Shannon	Site	Month	Depth
Anova p-value	0.8597	0.4149	4.322^{-08} ***
Tukey multiple comparisons test			2 m ≠ 200 m (0.000000) 15 m ≠ 200 m (0.0000210) 2 m = 15 m (0.07)
Simpson	Site	Month	Depth
Anova p-value	0.9225	0.6342	4.757^{-08} ***
Tukey multiple comparisons test			2 m ≠ 200 m (0.000016) 15 m ≠ 200 m (0.000001) 2 m = 15 m (0.6)
InvSimpson	Site	Month	Depth
Kruskal p-value	0.916	0.6285	2.335^{-05} ***
Dunn test			2 m ≠ 200 m (0.0000) 15 m ≠ 200 m (0.0003) 2 m = 15 m (0.8)

Tab. S5. Anosim and Adonis results performed on the Bray-Curtis dissimilarity matrix generated from the bacterial OTUs table.

	Site	Month	Depth
Anosim p-value	0.982	0.006 ***	0.001 ***
Adonis p-value	0.975	0.011 **	0.001 ***

Tab. S6. Simper test results with differences tested by statistical analysis for bacterial OTUs table.

Group		Most influential OTUs	Kruskal-Wallis (p-value)	Simper cumulative contributions	Average abundances in each compared treatment	Average dissimilarity between the two treatment
Month	February – June	5_Cyanobacteria_Oxyphotobacteria	0.0003058 ***	14%	February > June	55%
	February – August	5_Cyanobacteria_Oxyphotobacteria	0.0003058 ***	14%	February < August	54%
	February - November	5_Cyanobacteria_Oxyphotobacteria	0.0003058 ***	15%	February < November	49%
	June - August	3_Cyanobacteria_Oxyphotobacteria	0.3754	9%	June < August	50%
	June - November	4_Chloroflexi_Anaerolineae	0.1434	8%	June < November	51%
	August - November	4_Chloroflexi_Anaerolineae	0.1434	8%	August < November	48%
Site	pt2 - pt4	5_Cyanobacteria_Oxyphotobacteria	0.4996	9%	pt2 > pt4	47%
	pt2 – pt6	5_Cyanobacteria_Oxyphotobacteria	0.4996	8%	pt2 > pt6	48%
	pt4 – pt6	4_Chloroflexi_Anaerolineae	0.6618	8%	pt4 > pt6	49%
Depth	2m – 15m	5_Cyanobacteria_Oxyphotobacteria	0.8098	13%	2m > 15m	39%
	2m – 200m	4_Chloroflexi_Anaerolineae	1.714^{-05} ***	11%	2m < 200m	65%
	15m – 200m	4_Chloroflexi_Anaerolineae	1.714^{-05} ***	12%	15m > 200m	62%