Ultrarare marine microbes contribute to key sulphur-related ecosystem functions

Dimitri Kalenitchenko, Nadine Le Bris, Erwan Peru, Pierre Galand

Ultra-rare marine microbes contribute to key sulfur related ecosystem functions

Dimitri Kalenitchenko[a], Nadine Le Bris[a], Erwan Peru[a], Pierre E. Galand[a1]

[a]Sorbonne Universités, UPMC Univ Paris 06, CNRS, Laboratoire d'Ecogéochimie des Environnements Benthiques (LECOB), Observatoire Océanologique, Banyuls sur Mer, France

Keywords: rare biosphere, bacteria, functional redundancy, sulfide, diversity, colonization, succession, sulfur cycle

[1]To whom correspondence should be addressed. E-mail: pierre.galand@obs-banyuls.fr.

Abstract

The description of a rare biosphere within microbial communities has created great interest because microbes play a fundamental role in the functioning of all ecosystems on earth. Despite recent progress in understanding the ecology of the rare biosphere, the concept itself is still discussed and fundamental questions remain. Here, we target the seed bank compartment of the rare biosphere, assess the level of rarity at which microorganisms are still able to colonize an ecosystem and investigate whether rare species are functionally redundant. Using an original experimental design where wood in aquaria was inoculated with increasingly diluted coastal seawater, we show that bacteria that represented as few as 0.00000002 % of the cells in the environment (or 1 cell in 10 L of seawater) were still able to grow and play key roles within the ecosystem. Our experiment further showed that some bacteria can be replaced by others that have the potential to fulfill the same metabolic tasks. This finding suggests some functional redundancy within bacterial species. However, when ultra-rare bacteria were progressively removed, productivity was reduced, and below a certain threshold some processes were lost, and the function of the ecosystem was altered. Overall the study shows that bacteria that are not detected by high-throughput sequencing approaches are nevertheless viable and able to colonize new ecosystems, suggesting the need to consider ultra-rare microbes in the marine environment.

Introduction

Microbial communities play a fundamental role in the functioning of all ecosystems on earth (Falkowski *et al.* 2008; McFall-Ngai *et al.* 2013). In the sea, they mediate key biogeochemical cycles and are involved in nutrient cycling, primary production, organic matter mineralization and ultimately climate regulation (Kirchman 2010). These processes are often conducted by a few dominant and active species that are present in the environment at a given time (Pedrós-Alió 2006). Such a feature is reflected by the representation of the distribution of microorganisms along a skewed rank-abundance curve that shows the presence within a community of a few dominant species and many less abundant organisms (Curtis *et al.* 2002; Pedrós-Alió 2012). This long tail of rare microbes contains most of the tremendous diversity of a microbial community (Crespo *et al.* 2016).

This large number of rare species was coined the "rare biosphere" following the first application of high-throughput sequencing to describe marine microbial communities (Sogin *et al.* 2006). The rare marine biosphere was first hypothesized to reflect the stochastic and potentially unlimited dispersal capacity of microbes in the ocean. Traces of these dispersed microorganisms would eventually be found everywhere. Furthermore, following Baas Becking's statement that "everything is everywhere, but the environment selects", these microbes would represent a seed bank containing organisms that can grow when the conditions become favorable (Jones & Lennon 2010; Lennon & Jones 2011; Pedrós-Alió 2006). However, it was shown that the rare biosphere represented a more complex picture. It has patterns of biogeography that suggests the occurrence of selection processes (Galand *et al.* 2009). Furthermore, some rare microbes are active and at times are more active than the abundant members of the community (Campbell *et al.* 2011; Jones & Lennon 2010), which indicates that the rare biosphere is not only composed of a seed bank of dormant microorganisms. It thus appears that there are large ecological differences between different

compartments of the rare biosphere (Alonso‐Sáez *et al.* 2015; Hugoni *et al.* 2013; Pedrós-Alió 2012). Despite recent progress in understanding the ecology of the rare biosphere, the concept itself is still discussed, and fundamental questions remain. For instance, some data suggest that some of the observed microbial diversity results from methodological artifacts (Kunin *et al.* 2010; Lee *et al.* 2012; Patin *et al.* 2013) or that the observed diversity does not represents viable microorganism but instead reflects free or dead DNA (Carini *et al.* 2016; Lynch & Neufeld 2015). In addition, although sequencing now allows the in-depth probing of the richness of the rare biosphere (Crespo *et al.* 2016), it was recently debated that there may not be as many rare microbes as originally thought (Amann & Rossello-Mora 2016; Amann & Rosselló-Móra 2016; Lennon & Locey 2016). However, there have been few if any convincing experiments designed to explore the limits of rarity in natural samples.

The existence of the rare biosphere also begs the question of why there are so many microbial species in the sea. High levels of diversity have been early explained by the presence in the environment of different taxa that are able to conduct the same metabolic processes. The paradigm of functional redundancy has long been assumed in microbial ecology (Bell *et al.* 2005; Finlay *et al.* 1997), and the assumption is that a change in microbial community composition will not have consequences for microbial-mediated processes (Allison & Martiny 2008). The recent use of metagenomic profiles revealed high functional redundancy for marine microbes at global scale (Sunagawa *et al.* 2015). Inversely, the fact that some bacterial communities that are distinct in composition show different functions (Leff *et al.* 2015) indicates that the paradigm of redundancy does not always hold in natural microbial communities, as shown recently in soil ecosystems (Fierer *et al.* 2013). In aquatic environments, recent findings showing intrapopulation complementary (Garcia *et al.* 2015) support the theory of a high level of functional complementary in marine microorganisms

illustrated by reductive genomic evolution or genome streamlining (Giovannoni *et al.* 2014; Morris *et al.* 2012).

Here, we focus on the seed bank compartment of the rare biosphere, which was conceptualized by Pedrós-Alió (2006, 2012) as rare taxa that are not growing but that become abundant when conditions become favorable, and investigate the extent of its diversity and viability. More precisely, the goal of this study was to determine a threshold of abundance at which rare organisms can still successfully colonize a complex substrate, to test whether the rare species are functionally redundant and to verify whether they can be detected in the environment. We earlier showed that wood incubated in seawater is a useful tool for testing hypothesis about the functional ecology of communities (Kalenitchenko *et al.* 2016). Here, wood incubated in aquaria were inoculated with different dilutions of coastal seawater, and the growth of the bacterial communities was followed both within the wood and on the wood surface, with a special focus on the bacteria involved in the sulfur cycle. These keystone species are responsible for key ecosystem processes in the ocean. They are active during the degradation of large organic falls on the sea floor (Bienhold *et al.* 2013; Kalenitchenko *et al.* 2015; Kalenitchenko *et al.* 2017), the mineralization of organic matter in marine sediments (Jørgensen 1982) and biomass production at hydrothermal vents or cold seeps (Campbell *et al.* 2006; Martin *et al.* 2008). We chose to target these organisms because of their global importance and because their activity can be conveniently monitored.

Materials and methods

Experimental set up

Wood logs originating from a same pine tree (*Pinus pinea*) in the Banyuls sur Mer ‘ Biodiversarium’ botanical garden were cut the day before the experiment started. The experimental design consisted of four separate 40-L aquaria that each contained four 15 cm in

diameter and 10 cm in length, fresh (non sterile), wood logs (Kalenitchenko *et al.* 2016). We did not want to take the risk of changing the integrity of the chemical composition of the wood by autoclave (possible impact on sugars). The aquaria were initially filled with increasing dilutions of coastal seawater from the SOLA station in Banyuls Bay sampled at 3-m depth on 27 January 2015. Seawater filtered through a UF100LL filtration module (Polymem, Castanet-Tolosan, France) with a 0.01 µm pore size membrane was used for the dilutions. The first aquaria was started with 100% (1/1 treatment) off coastal unfiltered seawater, the second with 4 L of coastal water and 36 L of filtered water (1/10 treatment), the third with 0.4 L of coastal water and 39.6 L of filtered water (1/100 treatment) and the last with 100% filtered water (filtered treatment). Wood logs were immersed without water renewal for the first 7 days of the experiment after which, the aquaria water was renewed with piped in coastal seawater filtered through the 0.01 µm pore size filtration module (Polymem) at a flow rate that allowed at least 2 renewals per day. The experiment lasted for 28 days.

One log was removed and sampled from each aquarium after 7 days (7d), 14 days (14d), 21 days (21d) and 28 days (28d). At each sampling time, four replicate 1-cm long cores were extracted from each log using a 4.35-mm-wide increment core borer. Mats growing on the wood surface were sampled when present (21d and 28d) by scratching the entire wood surface with a scalpel blade.

DNA extraction and sequencing

The microbial mat samples were lysed and wood core samples were powdered as described earlier (Kalenitchenko *et al.* 2016; Kalenitchenko *et al.* 2015). The cell lysate was then transferred to the Maxwell® 16 Blood DNA Purification kit cartridge and processed automatically with the Maxwell® 16 automated extractor (Promega, Fitchburg, MA, USA).

DNA was also extracted from the original coastal seawater as described previously (Galand *et al.* 2015). See the supplementary material and methods for a full description of the procedure.

A portion of the 16S rRNA gene was amplified using bacteria specific primers, 28F (5'-TTTGATCNTGGCTCAG-3') and 519R (5'-GTNTTACNGCGGCKGCTG-3'), and then sequenced by a commercial laboratory (Research and Testing Laboratory, Lubbock, TX) on an Illumina Miseq sequencer to produce 2x300bp long, paired-end sequences. The raw sequence data have been deposited in the NCBI Sequence Read Archive (accession no. SRP099419 and BioProject accession no. PRJNA374511).

Sequence analysis

The 16S rRNA sequences were processed by following (Galand *et al.* 2015). In brief, sequences were paired, producing ca. 500 bp fragments, quality trimmed, and chimeras were removed. Sequences were grouped in operational taxonomic units (OTU) at 97% similarity using the Uclust algorithm (Edgar 2010). The taxonomy of the most abundant sequence of each OTU was assigned using the SILVA SSU 123 database (Pruesse *et al.* 2007). All samples were resampled down to 5094 sequences per sample.

To identify the possible presence of rare species in seawater we used a dataset of sequences from the Banyuls Bay microbial observatory time series. Samples have been gathered twice a month from 2007 to 2014. DNA has been extracted and samples sequenced as described earlier (Galand *et al.* 2015). A total of 7,418,629 sequences were used to search for rare species in the time series. The occurrence of sequences that were not detected in the time series was calculated to be $< 1 / 7{,}418{,}629$, which is < 0.00001 %.

Bacterial diversity was estimated by calculating the Shannon diversity index (H') and the standardized effect size (SES) (Kembel *et al.* 2010), which is a standardized measure of phylogenetic diversity. Phylogenetic diversity was computed using the Picante package (Kembel *et al.* 2010) in R. To calculate SES, representative sequences from each OTU were

aligned using MUSCLE (Edgar 2004), and the alignment was then cleaned to remove non-overlapping sequence regions. A phylogenetic tree was constructed using FASTTREE (Price *et al.* 2010). The observed phylogenetic diversity was compared to the average phylogenetic diversity in a randomly generated community (null model) and divided by the standard deviation of the phylogenetic distances in the null model (Webb *et al.* 2008). The null model randomizes community data matrixes with the independent swap algorithm to maintain the species occurrence frequency and sample species richness (Kembel 2009). Positive SES values indicate greater phylogenetic distance among co-occurring species than expected by chance, whereas negative values indicate small phylogenetic distance.

The Bray Curtis similarity index was computed to compare the community composition between samples with an MDS analysis. The effect of dilution and time factors on composition differences was tested with PERMANOVA in the program PAST (Hammer *et al.* 2001). To identify OTUs or bacterial orders that had a significant difference in abundance between groups, we used a nested ANOVA with the tools described in (McDonald 2014). Differences in diversity (Shannon) were tested with ANOVA. Two outlier samples (1/1: 14d-Replicate1, 1/10/ 21d-Replicate1) were removed from the analysis.

Flow cytometry and most probable number

Heterotrophic prokaryotic abundance in the aquaria water was measured by flow cytometry (Supplementary methods). We used a method that is commonly used in microbiology, called the most probable number method (MPN) (Oblinger & Koburger 1975), to estimate the number of cells belonging to the OTUs that we monitored. This method relies on dilutions and the observation of microbial growth to estimate the number of cells present in a starting solution. To do so, we counted the number of replicate wood samples in which we detected the OTUs of interest at the end of the experiment (28d). We then used an MPN table to

estimate the number of cells in the original seawater sampled from Banyuls Bay. Their frequency was calculated by dividing the estimated number of cells by the total number of cell counted by flow cytometry.

Sulfide detection

Continuous monitoring of sulfide was performed using autonomous potentiometric sensors (NKE, Hennebont, France) equipped with lab made Ag/Ag_2S electrodes (Le Bris *et al.* 2012) to detect when sulfide started to diffuse at the surface of the wood logs. To measure the sulfide through the microbial mats, microsensor profiling was performed using amperometric microsensors (H2S100) and a picoammeter (PA2000) (UNISENSE, Aarhus, Denmark). Sulfide concentrations were profiled vertically, perpendicular to the surface of the wood, with a resolution of 250 µm at each sampling time using a micromanipulator. Calibrations were performed by standard additions of Na_2S stock solution in seawater as described in Laurent et al. (2009) (Laurent *et al.* 2009) for the potentiometric electrodes. The amperometric electrode was calibrated at pH of 6, 6.5, 7 and 7.5 using a Titrando pH controller (Metrohm, Herisau, Switzerland). The slope of the electrode response to H_2S was thus confirmed to be pH independent and was calculated combining the whole calibration series. The total free sulfide concentration (i.e. both H_2S and HS) is determined from the pH variability at the wood-water interface. To calculate the HS^- fraction, we used a pKa_1* of 6.7 at 13°C and salinity 37 (Rickard & Luther 2007).

The pH at the wood surface and wood water interface was measured with 1 mm resolution with a mini electrode (1.5 mm diameter) (Diamond General Co., Ann Arbor, USA) coupled with the sulfide microsensor. To prevent sulfide profiles from being disturbed by the larger pH electrode, the tips of both electrodes were 1 cm apart. Triplicates of the combined

pH and H_2S profiles were acquired to determine mean concentration of total sulfide and the corresponding standard deviation, which reflects the heterogeneity of the mat.

Results

Sulfide concentrations

We continuously measured sulfide concentrations inside the wood during the course of the experiment. In the undiluted treatment, sulfide was first detected inside the wood after 13 days of incubation. In the 1/10 dilution, sulfide was detected after 21 days, and in the 1/100 dilution, the electrode potential only showed a small shift from baseline (< 7 mV), which indicated that no significant amount of sulfide built up at the wood surface. In the aquaria with filtered water only, the sulfide electrode potential in the wood did not change over the experiment (Fig. 1a).

Surface mat sulfide profiles obtained with amperometric microsensors also reveal marked differences between the treatments (Fig. 1b). The mats were first visible at 21d, and the maximum sulfide concentration was higher in the undiluted experimental mats than in the 1/10 dilution mats. No mats were visible at the other dilutions. At 28d, the undiluted mat had thicker sulfide layer and higher maximum sulfide concentrations at the wood surface compared to those in the 1/10 dilution treatment (Fig. 1b). No sulfide was detected on the 1/100 and filtered treatments at 28d.

Community diversity

Community diversity measured as the Shannon index showed that at 7d, the communities from the undiluted treatment had significantly higher diversity (ANOVA, $p<0.05$), followed by the 1/10 dilution, the 1/100 and the filtered treatments (Fig. S1). At 14d, the undiluted, 1/10 and 1/100 samples had similar diversity, whereas the diversity of the filtered treatment

community remained low. At 21d, the pattern of diversity did not change markedly, and at 28d, the diversity in the undiluted treatment decreased significantly.

For the phylogenetic diversity, the largest difference between samples was observed at 7d among diluted treatments. The 1/10 communities were the more phylogenetically clustered (lowest SES values), followed by the 1/100 and filtered communities (ANOVA, $p<0.05$). As time passed, communities became less phylogenetically clustered (increasing SES values), and there were no significant differences between treatments (Fig. S2).

Bacterial community composition inside the wood

Community composition between samples was compared by computing the Bray Curtis dissimilarity index. For undiluted treatment (1/1) at 7d, the bacterial communities were very different from those of the 7d communities in the diluted treatments. Dilution had a stronger effect on community composition than incubation time had (Table S1 and S2). The community composition changed among 7d, 14d and 21d. However, the undiluted 21d and 28d communities were very similar to each other (Fig. 2).

For the 1/10, 1/100 and filtered treatments, the 7d communities were more similar to each other than to the communities from the undiluted treatment, as shown by their close proximity on the MDS plot (Fig. 2). After 7d, the 1/10 communities became more similar to the 1/1 communities, whereas the 1/100 and filtered communities became more similar to each other with time (Fig. 2). In all dilution treatments, the communities changed with time and continued changing after 21d. The diluted communities changed more than the undiluted ones did, as shown by larger distances observed between samples on the plot.

Composition at order level

At the order level, in the diluted treatments, after 7 days (7d), there were significantly fewer *Campylobacterales* and *Flavobacteriales* compared to the undiluted bacterial communities (1/1). In the 1/10 dilution, the proportion of *Oceanospirillales* and *Alteromonadales* was higher than that in the undiluted treatment. In the 1/100 and filtered treatments, the proportion of *Vibrionales* was higher (Fig. 3).

After 14 days (14d), the samples from the undiluted treatment (1/1) maintained the same groups as those present at 7 days but the proportion of *Aeromonadales* and *Alteromonadales* decreased. The proportion of *Campylobacterales* increased significantly in the 1/10 dilution. *Rhodospirillales* and *Flavobacteriales* sequences strongly increased in the 1/100 and filtered treatments (Fig. 3).

After 21 days (21d), the proportion of *Aeromonadales* and *Alteromonadales* continued to decrease in the undiluted treatment (1/1), and 2 new groups increased (*Desulfovibrionales* and *Bacteroidales*). *Desulfovibrionales* and *Bacteroidales* were also detected in the 1/10 dilution but at lower abundance. There were no significant changes otherwise for the 1/100 and filtered treatments (Fig. 3).

After 28 days (28d), the proportion of *Aeromonadales* and *Alteromonadales* kept decreasing in the undiluted treatment. In the 1/10 dilution, *Desulfovibrionales* and *Bacteroidales* sequences increased. In the 1/100 dilution, *Bacteroidales* became more abundant (Fig. 3) but more variable between replicates in the samples from the filtered treatment (Fig. S3).

Dynamics at the OTU level

We identified OTUs representing organisms known to be involved in sulfur cycling in chemosynthetic ecosystems. For sulfate reducers, we identified two OTUs that grew well in the undiluted treatment but not, or poorly, in the diluted treatments (Fig. 4). These two OTUs

were affiliated with *Desulfovibrio piezophilus* (OTU10 and OTU58, with 100% and 96 % identity, respectively, over 442 bases). OTU10 was first detected at 21d in the undiluted treatment and at 28d in the 1/10 dilution. It was only present at trace amounts in the 1/100 and filtered water treatments (Fig. 4a). OTU58 was also detected in the samples at 21d in the undiluted treatment but was not present in the 1/10 diluted treatment nor in the 1/100 or filtered treatments (Fig. 4b).

We also identified potential sulfide oxidizers OTUs that grew better in the undiluted treatment than in the diluted waters (Fig. S4). These OTUs were similar to *Arcobacter bivalviorum* (OTU6, 100% identity) and *Arcobacter* sp. (OTU15, 94% identity). OTU6 sequences were detected at 7d in the undiluted treatment. In the 1/10 dilution, sequences appeared mostly at 28d and were less abundant (Fig. S4a). OTU15 was only present in the undiluted treatment and was detected at 21d (Fig. S4b).

Some of the OTUs were more abundant in samples from the diluted compared to the undiluted treatment. OTU43 (100% identity over 452 bases to *Marinomonas foliarum*) was more abundant at 7d in the undiluted and 1/10 dilution treatments but was overall more abundant in the 1/100 dilution, especially from 14d (Fig. S5a). OTU19 (100% identity over 452 bases to *Vibrio oceanisediminis*) was rare in the undiluted treatment, more abundant and present at all times tested in the 1/10 dilution and most abundant in the 1/100 dilution, especially at 7d (Fig. S5b).

Proportion of sequences and cell counts

OTU10, OTU58, OTU6, OTU15, OTU19 and OTU43 sequences were not detected in the sequences obtained from the original coastal seawater sample from Banyuls Bay. Because we analyzed 23,743 sequences of bacteria from this inoculum water, the frequency of these OTUs represented <0.004 % of the sequences. We further verified whether the sequences

were present in a database of 7,418,629 sequences sampled at Banyuls Bay over a period of 9 years. OTU10, OTU58, OTU6, OTU15, OTU19 and OTU43 were never detected, and their occurrence was thus calculated to be < 0.00001 %.

Our flow cytometry counts validated the serial dilution of the Banyuls Bay seawater for the experiment (Fig. S6). We used the most probable number method (MPN) to estimate the number of cells belonging to the OTUs that we monitored (Table S3). Our MPN estimates showed that some OTUs had occurrences as low as 1 cell in 10 L of seawater (OTU15). That small proportion represents 0.00000002 % of the cells present in the aquaria.

OTUs in the microbial mats

We identified OTUs affiliated with *Arcobacter* sp. in the mats growing on the wood logs (Fig. S7). Mats were visible starting at 21d. In the undiluted treatment, among the sequences that were identified as *Arcobacter* sp., the OTU Arcobacter1 dominated, followed by the OTU Arcobacter6 and a few sequences belonging to the OTU Arcobacter3 (Fig. 5). In the 1/10 dilution, the *Arcobacter* sequences belonged mainly to Arcobacter3. In the 1/100 dilution, very few *Arcobacter* sequences were present. They were represented by Arcobacter31 and Arcobacter3. In the samples from the filtered treatment, there were also few *Arcobacter* sequences at 21d and 28d; half of those identified belonged to OTU1, and half belonged to OTU3. Arcobacter1 was similar to the OTU6 found inside the wood. The other 3 main *Arcobacter* OTUs were not detected inside the wood (Fig. 5).

Discussion

Our experiment demonstrated that seawater contains bacteria that are extremely rare but still have the potential to grow and multiply rapidly when they encounter a suitable niche. Our work suggests the existence of ultra-rare bacteria, microbes that are never detected by high

348 throughput sequencing approaches but that are nevertheless viable and able to colonize new
349 ecosystems with the potential to fulfill an important function in the marine ecosystem.
350 The presence of these ultra-rare microbes can, to date, only be revealed through
351 culturing or enrichment approaches. The organisms that grew in our experimental wood were
352 not detected by sequencing the inoculum water originating from Banyuls Bay. Their
353 occurrence in the inoculum in terms of sequence abundance was thus below 0.004%. These
354 ultra-rare microbes were not found either among sequences from Banyuls Bay collected over
355 8 years. By taking into account the total number of sequences from this long time series, these
356 ultra-rare microbes would represent < 0.00001 % of the naturally occurring community.
357 In terms of cell numbers, the flow cytometry counts and the most probable number
358 approach showed that, theoretically, the frequency of ultra-rare microbes was as low as 1 cell
359 in 10 L of seawater or as little as 0.00000002 % of the cells in the aquaria. They were not able
360 to grow below that threshold. Together, our count and sequencing data show a level of rarity
361 that is several orders of magnitude lower than thresholds used earlier to define the rare
362 biosphere in sequencing data: 0.1% (Pedrós-Alió 2012) or 0.01% (Galand *et al.* 2009). For the
363 first time, we also provide an estimation of rarity in terms of cell numbers. Theoretically, if
364 PCR or sequencing bias were not considered, one would thus need at least 8,038,282,890
365 sequences (the number of bacteria present in average in 10 L of Banyuls Bay seawater) to be
366 able to detect the ultra-rare *Arcobacter* cell (Fig. 6). Thus, current sequencing techniques may
367 not provide exhaustive maps of the microbial diversity even with one million final reads, a
368 number suggested recently to be required to yield quasi-exhaustive mapping of marine
369 bacterial diversity (Crespo *et al.* 2016).
370 These ultra-rare microbes could have been missed by our sequencing approach if they
371 had been present in the water as resting spores. Spore forming microbes are common (Lennon
372 & Jones 2011) and their DNA sometimes difficult to extract. Among the sulfate-reducing

bacteria, the *Firmicutes* form endospores that can be dispersed through the global ocean (Hubert *et al.* 2009). However, among the bacteria that we identified, *Arcobacter* and *Desulfovibrio* are considered non-spore forming bacteria (Vandamme *et al.* 2015) and should be detected by PCR based diversity surveys. The growth of these organisms in our experiments suggests that the ultra-rare biosphere contains living non-dormant bacteria, and such viability shows that they do not represent taphonomic DNA originating from decaying organisms, or extracellular DNA (Carini *et al.* 2016; Lynch & Neufeld 2015). In the context of the ongoing debate around the presence of a "inflated biosphere" originating from PCR or sequencing artifacts (Kunin *et al.* 2010; Lee *et al.* 2012; Patin *et al.* 2013), our findings argue for the true existence of the rare biosphere.

In our study, we focused on microorganisms involved in the sulfur cycle. For instance, sulfate reducing bacteria use the sulfate present in seawater as an electron acceptor to obtain energy and release hydrogen sulfide. By measuring the concentration of sulfide, the accumulation rates could be used to indirectly assess microbial activity and thus monitor this key ecosystem process occurring in the wood and wood surface. Our data demonstrate that after diluting the inoculum 10 times, the ecosystem processes were maintained. At a 100 times dilution, however, sulfide was no longer detected. Thus, when ultra-rare microbes were progressively removed, ecosystem processes were maintained but only until a certain threshold of rarity. Below that threshold some processes were lost and ecosystem functioning was impaired. Interestingly, even though sulfide was still detected after the first dilution, it appeared later and the concentration plateau was not reached, with a slower buildup of sulfide at the interface. It might be that at 1/10 dilution, the activity started later because the community had to restore a critical biomass before ecosystem processes could be detected. Our data suggest, despite the fact that ecosystem functioning was maintained, that activity was reduced when the original community was diluted.

Dilution had the effect of reducing the community diversity and the phylogenetic diversity at the beginning of the experiment. Dilution certainly removed some rare species and probably decreased the abundance of individuals within more abundant species. It affected the composition of the community used as inoculum, which in turn affected the composition of the communities colonizing the wood. As a result, the composition of the wood communities in the diluted treatments differed from those observed in the wood from the undiluted treatment. For instance, one sulfate reducing bacteria (*Desulfovibrio*, OTU 58) and a potential sulfide oxidizer, *Arcobacter* (OTU 15), were never detected in the diluted treatments. Interestingly, community diversity later increased with time and the functioning of the community was maintained, as revealed by the production of sulfide. It clearly shows that some microorganisms from the original community were replaced by others with the potential to fulfill the same metabolic tasks. Functional redundancy can be difficult to demonstrate because it requires knowledge about the microbial populations that perform a specific process (Allison & Martiny 2008). Here, a change in microbial community composition did not have any consequences for the microbial-mediated production of sulfide. It is an indication of the functional redundancy within the bacterial species that are involved in the degradation of large organic falls in the sea. Metagenomic data from planktonic marine bacteria recently suggested high functional redundancy at global scale (Sunagawa *et al.* 2015). However, the recent report that microorganisms in soils did not show a high degree of functional redundancy (Fierer *et al.* 2013) could suggest that redundancy may vary depending on ecosystems or, more likely, that the scale of the study is important. It has been emphasized that redundancy has to be tested under common environmental conditions (Bradford & Fierer 2012).

In our study, a change in community composition, concomitant with the maintenance of a function, was also clearly visible in the mats growing on the wood. Our results show that

423 at each level of dilution different *Arcobacter* bacterium were found in the mat. Again, it
424 points toward a certain level of functional redundancy within the genera. In addition, these
425 differences at the species level indicate that when some rare bacteria are lost or reduced in
426 abundance by dilution, other rare microbes can colonize the substrate and use the available
427 niche. Interestingly, some microorganisms that grew in the wood in the diluted treatments
428 were not detected in the woods from the undiluted treatment. For example, *Marinomonas*
429 (OTU43) and *Vibrio oceanisediminis* (OTU19) were less abundant or not detected in the
430 undiluted treatment. Dilution probably liberated a niche for these microorganisms that
431 otherwise would have been outcompeted. It should also be noted that a few *Arcobacter* were
432 present in the lowest dilutions after 21d, even though sulfide was not detected, which possibly
433 reflects the complete consumption of the sulfide produced and that wood could have produced
434 sulfide if the experiment had been longer.
435 In conclusion, we demonstrate the presence in the sea of very rare but viable bacteria
436 that are key to important ecosystem processes, such as organic matter degradation. These
437 functioning bacteria occurred at concentrations as low as 1 cell per 10 L of water and thus
438 would not be detected by current amplicon based environmental diversity surveys. In the
439 debate on the extent of microbial diversity, we found that the bacterial community richness
440 remains under estimated (Crespo *et al.* 2016; Locey & Lennon 2016). The rare biosphere
441 described to date only represents the emerged part of an iceberg floating in a sea of ultra-rare
442 microbes, which represent an underappreciated source of hidden diversity.
443
444 **Acknowledgment**. The work of PEG was supported by the Agence Nationale de la
445 Recherche (ANR) through the projects EUREKA (ANR-14-CE02-0004-01). We
446 acknowledge the support of Fondation Total to the UPMC ‘chair biodiversity, extreme marine

environment and global change'. We are thankful to Connie Lovejoy for helpful advices and comments on the manuscript.

Author contributions. P.E.G. and D.K. designed research, D.K, E.P., N.L.B. and P.E.G. performed research, P.E.G contributed new reagents/analytic tools, D.K, N.L.B. and P.E.G. analyzed data, D.K, E.P., N.L.B. and P.E.G. wrote the paper.

Data Accessibility. The raw sequence data have been deposited in the NCBI Sequence Read Archive (accession no. SRP099419 and BioProject accession no. PRJNA374511).

References

Allison SD, Martiny JBH (2008) Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences*, **105**, 11512-11519.

Alonso‐Sáez L, Díaz‐Pérez L, Morán XAG (2015) The hidden seasonality of the rare biosphere in coastal marine bacterioplankton. *Environmental Microbiology*, **17**, 3766-3780.

Amann R, Rossello-Mora R (2016) Reply to "The Underestimation of Global Microbial Diversity". *mBio*, **7**, e01623-01616.

Amann R, Rosselló-Móra R (2016) After all, only millions? *mBio*, **7**, e00999-00916.

Bell T, Newman JA, Silverman BW, Turner SL, Lilley AK (2005) The contribution of species richness and composition to bacterial services. *Nature*, **436**, 1157-1160.

Bienhold C, Pop Ristova P, Wenzhöfer F, Dittmar T, Boetius A (2013) How Deep-Sea Wood Falls Sustain Chemosynthetic Life. *PLoS ONE*, **8**, e53590.

Bradford MA, Fierer N (2012) The biogeography of microbial communities and ecosystem processes: implications for soil and ecosystem models. In: *Soil Ecology and Ecosystem Services* (eds. Wall DH, Bardgett RD), p. 424. Oxford University Press, Oxford.

Campbell BJ, Engel AS, Porter ML, Takai K (2006) The versatile ε-proteobacteria: key players in sulphidic habitats. *Nature Reviews Microbiology*, **4**, 458-468.

Campbell BJ, Yu L, Heidelberg JF, Kirchman DL (2011) Activity of abundant and rare bacteria in a coastal ocean. *Proceedings of the National Academy of Sciences*, **108**, 12776-12781.

Carini P, Marsden PJ, Leff JW*, et al.* (2016) Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *bioRxiv*, 043372.

Crespo BG, Wallhead PJ, Logares R, Pedrós-Alió C (2016) Probing the Rare Biosphere of the North-West Mediterranean Sea: An Experiment with High Sequencing Effort. *PLoS ONE*, **11**, e0159195.

Curtis TP, Sloan WT, Scannell JW (2002) Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences*, **99**, 10494-10499.
Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.
Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460-2461.
Falkowski PG, Fenchel T, Delong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science*, **320**, 1034-1039.
Fierer N, Ladau J, Clemente JC*, et al.* (2013) Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. *Science*, **342**, 621-624.
Finlay BJ, Maberly SC, Cooper JI (1997) Microbial diversity and ecosystem function. *Oikos*, 209-213.
Galand P, Salter I, Kalenitchenko D (2015) Microbial productivity is associated with phylogenetic distance in surface marine waters *Molecular Ecology*, **24**, 5785-5795.
Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 22427-22432.
Garcia SL, Buck M, McMahon KD*, et al.* (2015) Auxotrophy and intrapopulation complementary in the 'interactome'of a cultivated freshwater model community. *Molecular Ecology*, **24**, 4449-4459.
Giovannoni SJ, Thrash JC, Temperton B (2014) Implications of streamlining theory for microbial ecology. *The ISME journal*, **8**, 1553-1565.
Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, **4**, 9.
Hubert C, Loy A, Nickel M*, et al.* (2009) A constant flux of diverse thermophilic bacteria into the cold Arctic seabed. *Science*, **325**, 1541-1544.
Hugoni M, Taib N, Debroas D*, et al.* (2013) Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 6004-6009.
Jones SE, Lennon JT (2010) Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences*, **107**, 5881-5886.
Jørgensen BB (1982) Mineralization of organic matter in the sea bed-the role of sulphate reduction. *Nature*, **296**, 643-645.
Kalenitchenko D, Dupraz M, Le Bris N*, et al.* (2016) Ecological succession leads to chemosynthesis in mats colonizing wood in sea water. *The ISME journal*, **10**, 2246-2258.
Kalenitchenko D, Fagervold SK, Pruski AM*, et al.* (2015) Temporal and spatial constraints on community assembly during microbial colonization of wood in seawater. *The ISME journal*, **9**, 2657-2670.
Kalenitchenko D, Le Bris N, Dadaglio L*, et al.* (2017) Bacteria alone establish the chemical basis of the wood-fall chemosynthetic ecosystem in the deep-sea. *The ISME journal*, **in press**.
Kembel SW (2009) Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. *Ecology Letters*, **12**, 949-960.
Kembel SW, Cowan PD, Helmus MR*, et al.* (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463-1464.

533 Kirchman DL (2010) *Microbial ecology of the oceans* John Wiley & Sons.
534 Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare
535 biosphere: pyrosequencing errors can lead to artificial inflation of diversity
536 estimates. *Environmental Microbiology*, **12**, 118-123.
537 Laurent MC, Gros O, Brulport J-P, Gaill F, Le Bris N (2009) Sunken wood habitat for
538 thiotrophic symbiosis in mangrove swamps. *Marine Environmental Research*, **67**,
539 83-88.
540 Le Bris N, Contreira-Pereira L, Yücel M (2012) In situ sensors for benthic ecosystem
541 studies. In: *Sensors for Ecology: Towards Integrated Knowledge of Ecosystems* (eds.
542 Le Galliard JF, Guarini JM, Gaill F), pp. 185-208. CNRS, Paris.
543 Lee CK, Herbold CW, Polson SW*, et al.* (2012) Groundtruthing next-gen sequencing for
544 microbial ecology–biases and errors in community structure estimates from PCR
545 amplicon pyrosequencing. *PLoS ONE*, **7**, e44224.
546 Leff JW, Jones SE, Prober SM*, et al.* (2015) Consistent responses of soil microbial
547 communities to elevated nutrient inputs in grasslands across the globe.
548 *Proceedings of the National Academy of Sciences*, **112**, 10967-10972.
549 Lennon JT, Jones SE (2011) Microbial seed banks: the ecological and evolutionary
550 implications of dormancy. *Nature Reviews Microbiology*, **9**, 119-130.
551 Lennon JT, Locey KJ (2016) The underestimation of global microbial diversity. *mBio*, **7**,
552 e01298-01216.
553 Locey KJ, Lennon JT (2016) Scaling laws predict global microbial diversity. *Proceedings*
554 *of the National Academy of Sciences*, **113**, 5970–5975.
555 Lynch MD, Neufeld JD (2015) Ecology and exploration of the rare biosphere. *Nature*
556 *Reviews Microbiology*, **13**, 217-229.
557 Martin W, Baross J, Kelley D, Russell MJ (2008) Hydrothermal vents and the origin of life.
558 *Nature Reviews Microbiology*, **6**, 805-814.
559 McDonald JH (2014) *Handbook of Biological Statistics* Sparky House Publishing,
560 Baltimore.
561 McFall-Ngai M, Hadfield MG, Bosch TC*, et al.* (2013) Animals in a bacterial world, a new
562 imperative for the life sciences. *Proceedings of the National Academy of Sciences*,
563 **110**, 3229-3236.
564 Morris JJ, Lenski RE, Zinser ER (2012) The Black Queen Hypothesis: evolution of
565 dependencies through adaptive gene loss. *mBio*, **3**, e00036-00012.
566 Oblinger J, Koburger J (1975) Understanding and teaching the most probable number
567 technique. *Journal of Milk and Food Technology (JMFT)*, **38**, 540-545.
568 Patin NV, Kunin V, Lidström U, Ashby MN (2013) Effects of OTU clustering and PCR
569 artifacts on microbial diversity estimates. *Microbial Ecology*, **65**, 709-719.
570 Pedrós-Alió C (2006) Marine microbial diversity: can it be determined? *Trends in*
571 *Microbiology*, **14**, 257.
572 Pedrós-Alió C (2012) The rare bacterial biosphere. *Annual Review of Marine Science*, **4**,
573 449-466.
574 Price MN, Dehal PS, Arkin AP (2010) FastTree 2–approximately maximum-likelihood
575 trees for large alignments. *PLoS ONE*, **5**, e9490.
576 Pruesse E, Quast C, Knittel K*, et al.* (2007) SILVA: a comprehensive online resource for
577 quality checked and aligned ribosomal RNA sequence data compatible with ARB.
578 *Nucleic Acids Research*, **35**, 7188-7196.
579 Rickard D, Luther GW (2007) Chemistry of iron sulfides. *Chemical reviews*, **107**, 514-562.

Sogin ML, Morrison HG, Huber JA, *et al.* (2006) Microbial diversity in the deep sea and the under explored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 12115-12120.
Sunagawa S, Coelho LP, Chaffron S, *et al.* (2015) Structure and function of the global ocean microbiome. *Science*, **348**, 1261359.
Vandamme P, Dewhirst FE, Paster BJ, On SLW (2015) Arcobacter. In: *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Ltd.
Webb CO, Ackerly DD, Kembel SW (2008) Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, **24**, 2098-2100.

Figure legends

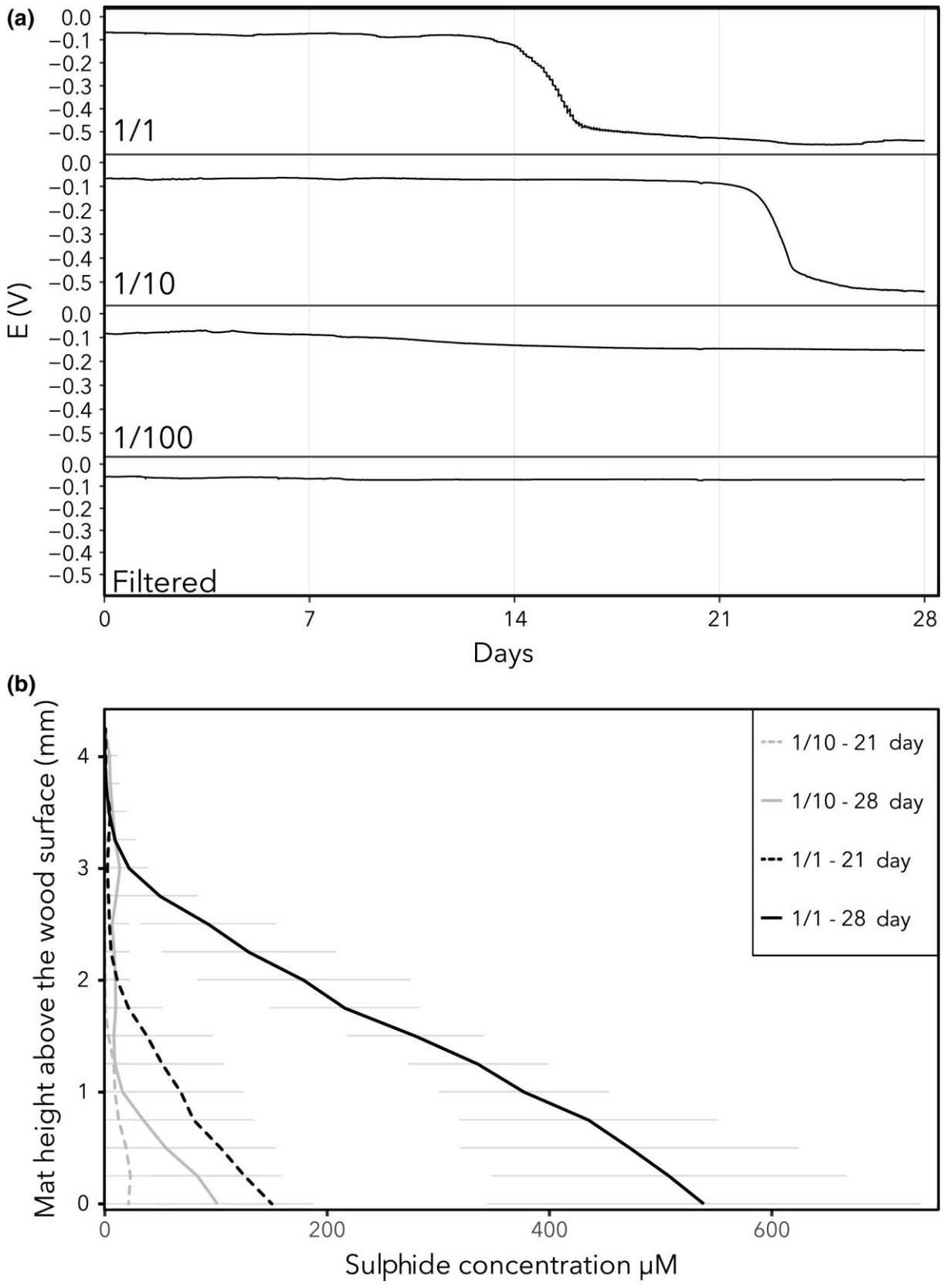
Figure 1. (a) Electrode potential as function of time indicating the presence of sulfide on the surface of wood inoculated with different dilutions of seawater (1/1, 1/10, 1/100 and filtered). A decrease in electrode potential indicates an accumulation of sulfide. 7d, 14d, 21d and 28d indicate sampling times. (b) Microprofiles of sulfide concentration through the mats colonizing the wood surface in the undiluted (1/1) (black lines) and 1/10 dilution experiments (grey lines) measured after 21 days (dashed lines) and 28 days (full lines). Error bars represent the standard deviation obtained from triplicate measurements.

Figure 2. Trajectories of the bacterial communities. The similarity between bacterial communities growing inside the wood is represented by an MDS ordination based on the Bray-Curtis index. Colors group communities sampled from a same dilution treatment. The size of the points is proportional to the time elapsed after the incubation started. The percent value for each axis represents the proportion of total variation explained.

Figure 3. Proportions of sequences at the order level detected inside woods incubated in aquaria inoculated with seawater with a range of dilutions (1/1, 1/10, 1/100 and only filtered) and sampled at four different time points (7 to 28 days). Each bar shows the average values of 4 replicate samples.

Figure 4. Abundance of sequences belonging to two sulfate reducing bacteria identified as *Desulfovibrio piezophilus* OTU10 (a) and *Desulfovibrio* sp. OTU58 (b) during the experiment in the 4 different inoculation conditions. The number of sequences represents averages over four replicates.

Figure 5. Proportion of the 4 most abundant *Arcobacter* OTUs in the mats colonizing the wood surface in the 4 different treatment conditions (1/1, 1/10, 1/100 and filtered) at 21d and 28d. The proportion is relative to the total number of sequences found in each sample. The whole wood surface was sampled for sequencing.



635
636
637 Fig.1

638

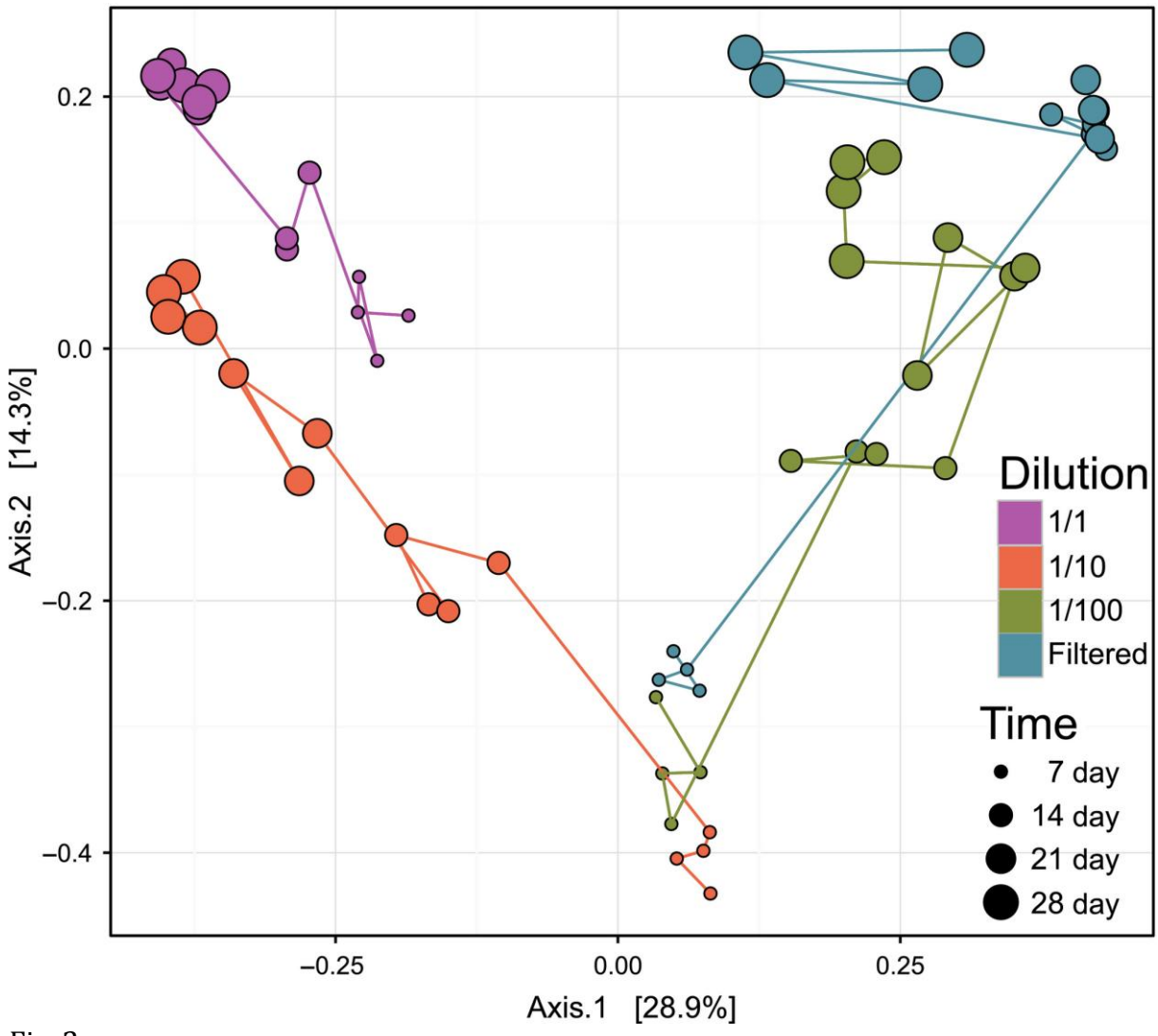
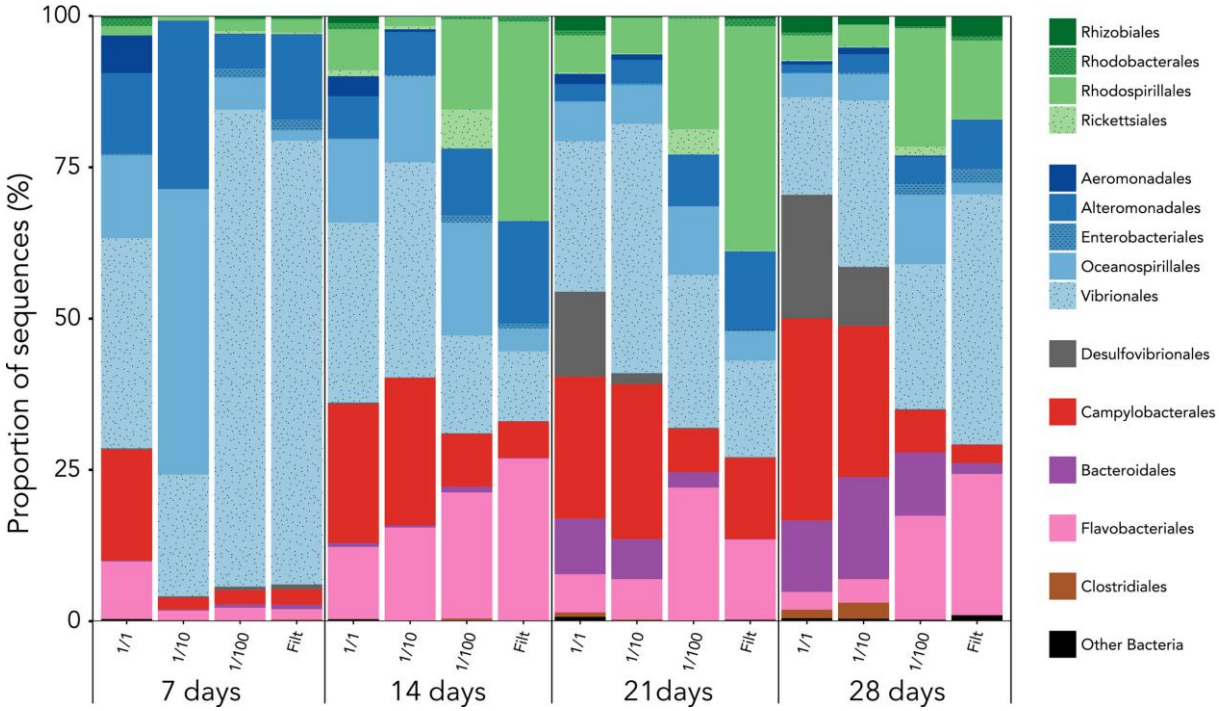
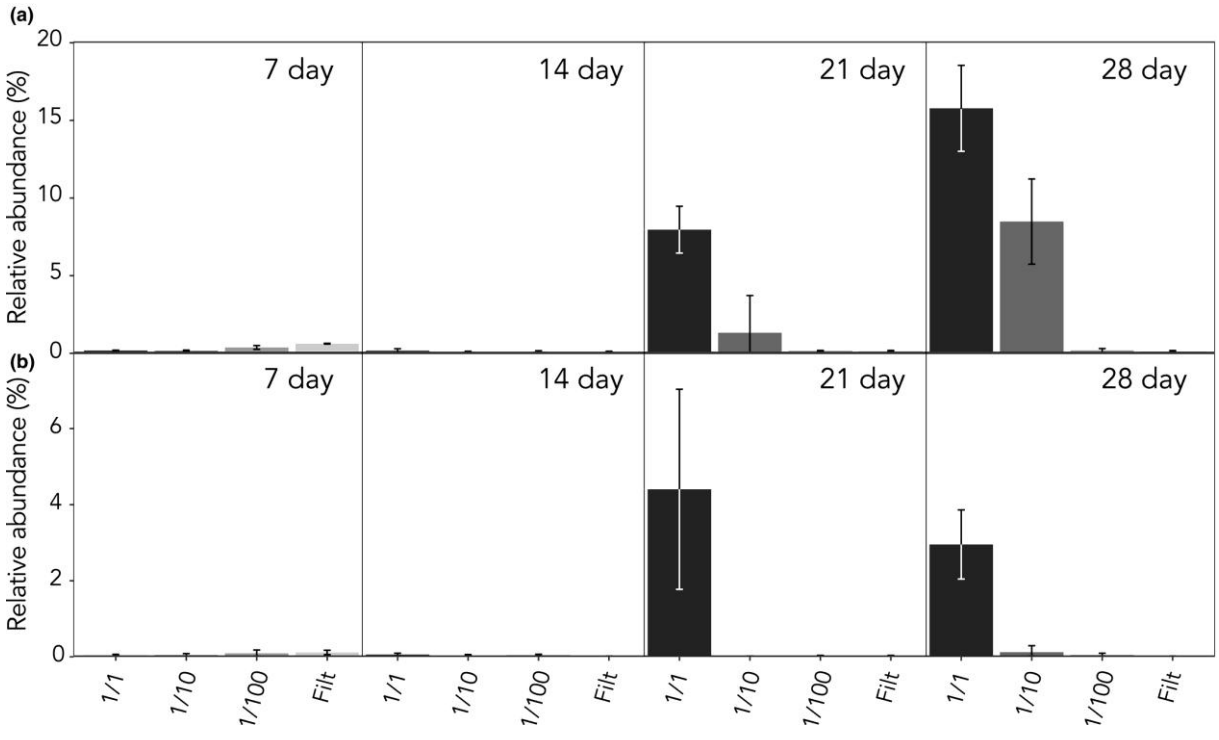

0.2
0.0
−0.2
−0.4
Axis.2 [14.3%]
−0.25
0.00
0.25
Axis.1 [28.9%]
Dilution
1/1
1/10
1/100
Filtered
Time
7 day
14 day
21 day
28 day

Fig. 2



646
647 Fig. 3

648



649
650

651 Fig. 4

652

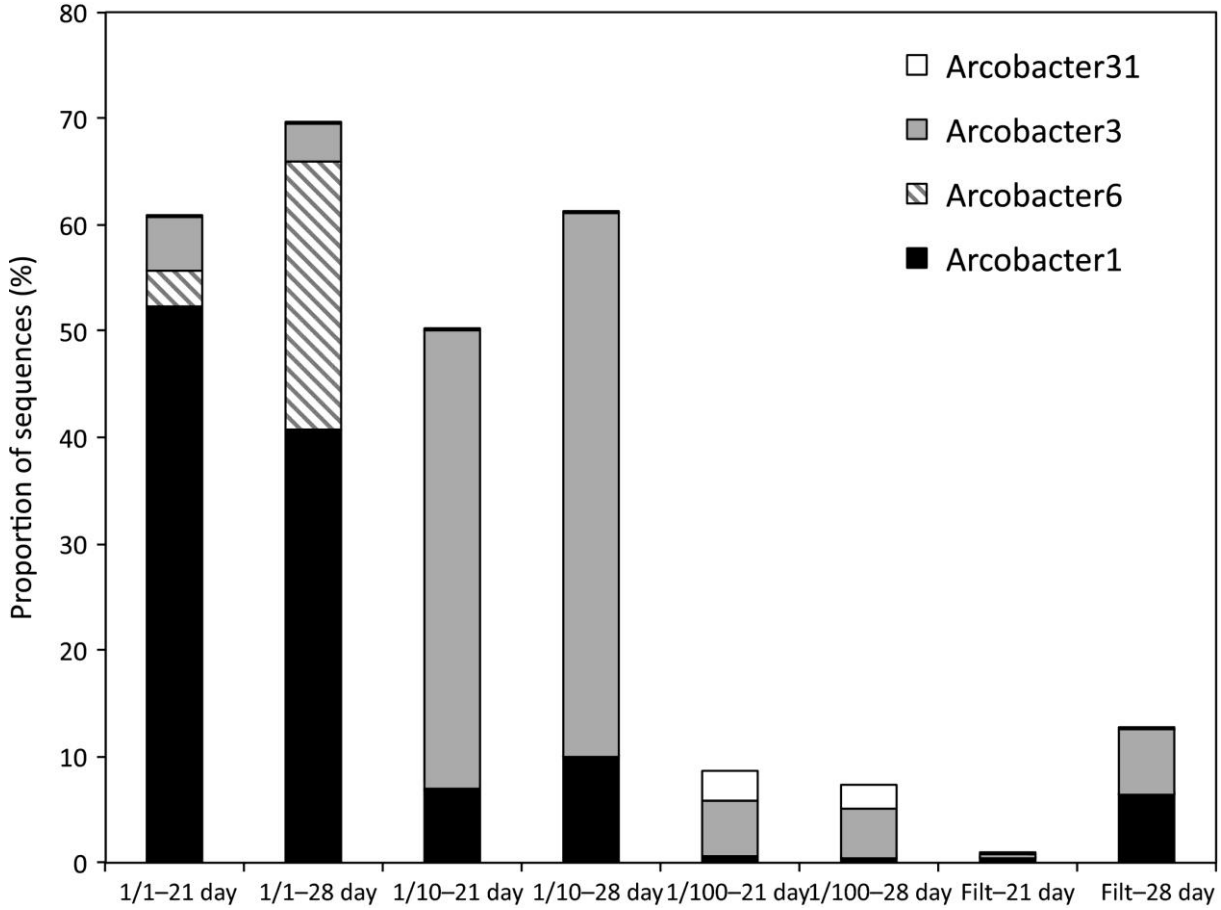

Arcobacter31
Arcobacter3
Arcobacter6
Arcobacter1
Proportion of sequences (%)
0
10
20
30
40
50
60
70
80
1/1–21 day
1/1–28 day
1/10–21 day
1/10–28 day
1/100–21 day
1/100–28 day
Filt–21 day
Filt–28 day

Fig. 5