



OPEN

Biomass addition alters community assembly in ultrafiltration membrane biofilms

Marisa O. D. Silva & Jakob Pernthaler✉

Freshwater biofilms assemble from a pool of rare water column genotypes. Random density fluctuations and temporal species turnover of functionally equivalent potential colonizers result in compositional variability of newly formed biofilm communities. We hypothesized that stronger environmental filtering as induced by enhanced substrate levels might reduce the impact of a locally variable pool of colonizers and instead select for more universal habitat specialists. Our model were heterotrophic biofilms that form on membranes during gravity-driven ultrafiltration of lake water. In four separate experiments, biomass of the cyanobacterium *Microcystis* was added to the feed water of one set of treatments (BM) and the resulting biofilm communities were compared to unamended controls (CTRL). Biomass addition led to a significant shift of community assembly processes: Replicate BM biofilms were more similar to each other than by chance in 3 of 4 experiments, whereas the opposite was the case for CTRL communities. Moreover, BM communities were more stochastically assembled across experiments from a common 'regional' pool of biofilm colonizers, whereas the composition of CTRL communities was mainly determined by experiment-specific 'local' genotypes. Interestingly, community assembly processes were also related to both, physiology (aerobic vs. anaerobic lifestyle) and the phylogenetic affiliation of biofilm bacteria.

Virtually every submerged natural and artificial surface attracts the spontaneous formation of biofilms¹. This biofouling process has important ecological, hygienic and economic consequences^{2,3}. Many if not most microbes in lakes and streams are capable of surface attachment, as reflected by the high diversity in the various types of freshwater biofilms^{4–8}. Biofilm-forming bacteria have a dual life style and undergo profound physiological transitions between the planktonic and sessile phase⁹. However, there are differences in the respective affinity of lacustrine bacteria to the attached life style, and some abundant genera are even exclusively planktonic^{10,11}. As a consequence, the free-living and biofilm-associated freshwater microbial assemblages substantially differ in composition^{12,13}. Population dynamics and community structure in multispecies biofilms are strongly shaped by interspecific interactions¹⁴, and environmental filtering has been proposed as a major assembly mechanism in, e.g., stream biofilm⁴.

The relationship between the community composition of biofilms with that of the source assemblage is poorly understood. With the possible exception of highly productive systems such as wastewater¹⁵, the dominant genotypes of aquatic biofilms appear to be rare in the surrounding water. These infrequent biofilm forming bacteria often have a 'tychoplanktic' life style in that they originate from a variety of non-pelagic habitats, such as suspended organic aggregates¹⁶, metazoan guts⁶ or body surfaces¹⁷, epiphytic or epilithic biofilms⁷, or terrestrial influx¹⁸. Community composition within these biofilm types may spatially vary or change over time^{13,19}, thereby leading to an ever-shifting pool of potential colonizers of novel surfaces. These mechanisms add an element of chance to the basically niche-driven⁴ selection of biofilm bacteria: For one, the functional redundancy of co-existing taxa in combination with priority effects²⁰ will instigate a 'competitive lottery' of potential colonizers from the 'local' species pool²¹, i.e. those species that are present in the individual source communities. Secondly, larger scale spatial heterogeneity²² or temporal (e.g., seasonal) species turnover²³ in the water column will create additional variation at the level of the 'regional' pool, defined as the totality of species in all source communities.

While the conditions in biofilms always differ from that of the surrounding water phase¹², their growth on inert surfaces such as rocks, glass slides, sand grains or porous membranes arguably depends on external substrates such as dissolved and particulate organic matter²⁴. It is conceivable that the quantity or quality of externally

Limnological Station, Department of Plant and Microbial Biology, University of Zurich, Seestrasse 187, 8802 Kilchberg, Switzerland. ✉email: pernthaler@limnol.uzh.ch

provided substrates may also shift the respective importance of the ‘local’ and ‘regional’ pools of potential colonizers during biofilm community assembly, e.g., by affecting functional redundancy and the balance between competitors in the source assemblage²⁵. We addressed this hypothesis using biofilms in experimental devices designed to investigate the potential of gravity-driven membrane filtration (GDM) for decentralized drinking water production^{26,27}. We synoptically analysed microbial community composition and assembly processes in four independent experiments that were performed over the course of several years^{27–29} with feed water from a single lake.

Results

Diversity and composition of biofilm communities. The complete sequencing dataset comprised 3,495 OTUs and approximately 10^6 reads. It was rarefied to the sample with the lowest read number (1.34×10^4) by random read exclusion²⁷. The normalized data set used for all subsequent analyses consisted of 2,721 OTUs with a total of 0.21×10^6 reads. The CTRL treatment had significantly higher OTU richness (734 ± 165 , mean \pm standard deviation) than the BM treatment (357 ± 63) (Student’s *t* test, $n = 8$, $p < 0.001$). The two treatments shared 786 genotypes that together formed 87% of total reads. Sixteen OTUs, representing almost one quarter of total reads, were present in every single community (Table 1).

In general, only very few OTUs on GDM biofilms were affiliated to the genotypes that are typically found in freshwater bacterioplankton³⁰. Proteobacteria represented 51% of the total amount of OTUs in biofilm communities and 61% of the total amount of reads (Fig. 1). However, many OTUs within all phylogenetic groups of biofilm bacteria were treatment-specific, i.e., were either absent or had considerably lower read numbers in one of the treatments. A total of 63 OTUs significantly contributed to the distance-based separation of the BM and CTRL treatments (SIMPER analysis, $p < 0.05$). The numbers in brackets after the colour codes of the individual phylogenetic groups in Fig. 1 report their relative distribution. The majority of the treatment-discriminating OTUs were affiliated with Betaproteobacteria (31%), Alphaproteobacteria (24%), and Bacteroidetes (20%).

Community structure and assembly processes. More than 40% of OTUs from either treatment only occurred in a single biofilm community (Fig. 2). By contrast, only 49 and 75 OTUs were present in all eight samples of the BM and CTRL treatments, respectively. The subset of ‘single sample’ OTUs had significantly more reads in the CTRL than in the BM treatments, whereas the opposite was the case for those OTUs that occurred in all communities of a treatment type (Student’s *t* tests, $n = 16$, $p < 0.01$).

Average linkage clustering of Bray–Curtis dissimilarities indicated a general separation between treatment types (Fig. 3), i.e., the communities of the BM treatments from all 4 experiments were more similar to each other than to the corresponding samples from the CTRL treatment. Similarity profile analysis showed that all phylogenetic groups except Firmicutes significantly contributed to this separation, albeit to a variable extent (Fig. 1). The clustering of only those OTUs that occurred in all biofilms (Table 1) also resulted in the perfect separation of samples according to treatment type (Bray–Curtis dissimilarity: 78%, data not shown). Five of these 16 ‘universal colonizers’ significantly contributed to this separation, all of them with higher read numbers in the BM treatment (Table 1).

The separation of communities according to treatment type was further corroborated by an error-free classification by Random Forest (RF) analysis (Out-of-Bag Error (OOB) = 0%, number of trees = 1,000, number of samples = 16, categories = 2). However, a RF classification of samples according to experiment was also remarkably accurate, and only 1 out of 16 samples was misclassified (OOB = 6.25%, number of trees = 1,000, number of samples = 16, categories = 4). The 20 OTUs that were most relevant for classification according to treatment type and experiment are listed in suppl. Tables S1 and S2, respectively.

Biological replicates of the BM treatments had significantly lower beta diversity (dissimilarity values: 0.39 ± 0.07) than those of the CTRL treatment (0.51 ± 0.07 ; one-tailed Student’s *t*-test, $n = 8$, $p < 0.05$) (Fig. 3). Moreover, Raup–Crick indices showed that the biological replicates of the BM treatments were significantly more similar to each other than by chance ($RC > 0.95$) in 3 out of 4 experiments, whereas the opposite was the case for replicates of the CTRL treatments ($RC < -0.95$).

Null model analysis of community assembly processes indicated that the biofilms of the CTRL treatments were predominantly shaped by the ‘local’ pool of OTUs from the corresponding experiments (NST: 26%). By contrast, a significantly higher ($p < 0.001$) influence of the ‘regional’ pool of OTUs (i.e., that were present in several or all experiments) was found for biofilms of the BM treatments (NST: 66%). A graphic interpretation of these results is given in Fig. 4. In addition, assembly processes within the two treatments were also assessed for each of the phylogenetic lineages depicted in Fig. 1. Striking differences in treatment-specific NST values to those of the total community were observed for some phylogenetic groups, most prominently for Bacteroidetes and Firmicutes (Fig. 4, suppl. Table S3).

Aerobic and anaerobic subcommunities. 39% of all OTUs representing 78% of total reads could be assigned to either aerobic or anaerobic taxa, based on information on their closest cultured relatives. If possible, anaerobic taxa were further split into strict (272 OTUs) and facultative (62 OTUs) types. However, since both groups behaved similarly in the below described analyses, this distinction was abandoned for the sake of increased statistical power.

Both, the aerobic and anaerobic sub-communities were clearly distinct between treatments, as reflected by average linkage clustering of Bray–Curtis distances (data not shown) and highly accurate RF classification ($OOB_{\text{aerobic}} = 0\%$; $OOB_{\text{anaerobic}} = 6.25\%$; number of trees = 1,000, number of samples = 16, categories = 2). Null model analysis performed on the sub-community of aerobic genotypes yielded similar results to the complete dataset, i.e., a significantly higher impact of OTUs from the local pool on the CTRL (NST: 26%) than on the

| Taxonomy | Accession number (% identity with closest known sequence) | Proportion of total reads (%) | BM treatment: % of reads | CTRL treatment: % of reads | Contribution (%) |
|--|---|-------------------------------|--------------------------|----------------------------|------------------|
| <i>Rhodoferax</i> sp. | LN870966 (100) | 4.6 | 4.2 | 0.4 | 21.28** |
| <i>Paucibacter</i> sp. | KM187599 (99.6) | 3.9 | 3.8 | 0.1 | 18.41*** |
| <i>Pseudoxanthomonas</i> sp. | LN560679 (100) | 2.9 | 2.6 | 0.4 | 14.62** |
| <i>Undibacterium seohonense</i> | KC735151 (100) | 2.6 | 2.4 | 0.2 | 11.97*** |
| <i>Ideonella</i> sp. | JF176654 (100) | 1.6 | 0.9 | 0.7 | 3.73 |
| <i>Variovorax paradoxus</i> | MN684277 (100) | 1.6 | 1.0 | 0.5 | 4.56 |
| Uncultured <i>Rhodobacter</i> sp. | MN493576 (100) | 1.6 | 0.3 | 1.3 | 5.27 |
| <i>Pseudorhodoferax</i> sp. | FPLS01023064 (100) | 1.1 | 1.0 | 0.1 | 4.89** |
| <i>Halicomonobacter hydrossis</i> | NR_074420 (100) | 0.9 | 0.3 | 0.6 | 2.88 |
| <i>Pelomonas</i> sp. | KX508949 (100) | 0.8 | 0.5 | 0.2 | 2.52 |
| <i>Dongia</i> sp. | FPLS01061553 (100) | 0.8 | 0.02 | 0.7 | 4.15 |
| <i>Hydrogenophaga</i> sp. | HAFE01076884 (100) | 0.5 | 0.2 | 0.4 | 1.79 |
| <i>Methyloversatilis</i> sp. | FJ660513 (100) | 0.5 | 0.1 | 0.4 | 1.58 |
| <i>Paucibacter</i> sp. | FPLS01019111 (100) | 0.4 | 0.2 | 0.2 | 1.33 |
| <i>Candidatus Methylopusillus planktonicus</i> | FN668046 (100) | 0.2 | 0.03 | 0.2 | 0.94 |
| Uncultured <i>Myxococcales bacterium</i> | DQ646306 (97.9) | 0.03 | 0.02 | 0.01 | 0.07 |

Table 1. OTUs found in all biological replicates of all treatments and experiments. The respective contribution of these 16 OTUs to the distance-based distinction between treatment types was estimated by 'simper' analysis (1000 permutations), with asterisks denoting significance. ** $p < 0.01$; *** $p < 0.001$.

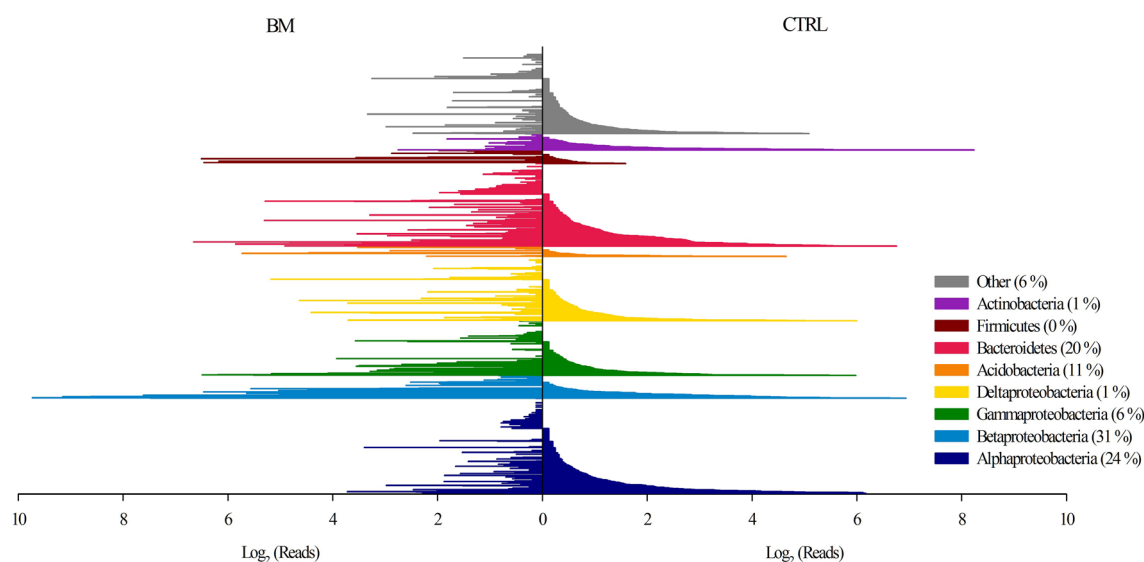


Figure 1. Distribution of phylogenetic groups in GDM biofilms with (BM) and without (CTRL) addition of cyanobacterial biomass. The right side of the graph plots OTUs in CTRL in ascending order of read numbers, and the left side the corresponding read numbers of the same OTUs in BM. Average read numbers (\log_2 transformed) of all 8 samples per treatment are depicted. A total of 63 OTUs significantly contributed to the distance-based separation of the two treatments (SIMPER analysis, $p < 0.05$). The distribution of these treatment-discriminating OTUs across the different phylogenetic groups is denoted in brackets after the name of the respective group.

BM (NST: 65%) biofilm communities ($p < 0.005$). By contrast, the sub-communities of anaerobic genotypes were always predominantly recruited from the regional OTU pool irrespective of treatment type (NST_{CTRL}: 64%; NST_{BM}: 91%; $p > 0.05$) (Fig. 4). This conclusion was further supported by the poor RF classification of anaerobic genotypes according to experiment (OOB = 44%; number of trees = 50,000, number of samples = 16, categories = 4).

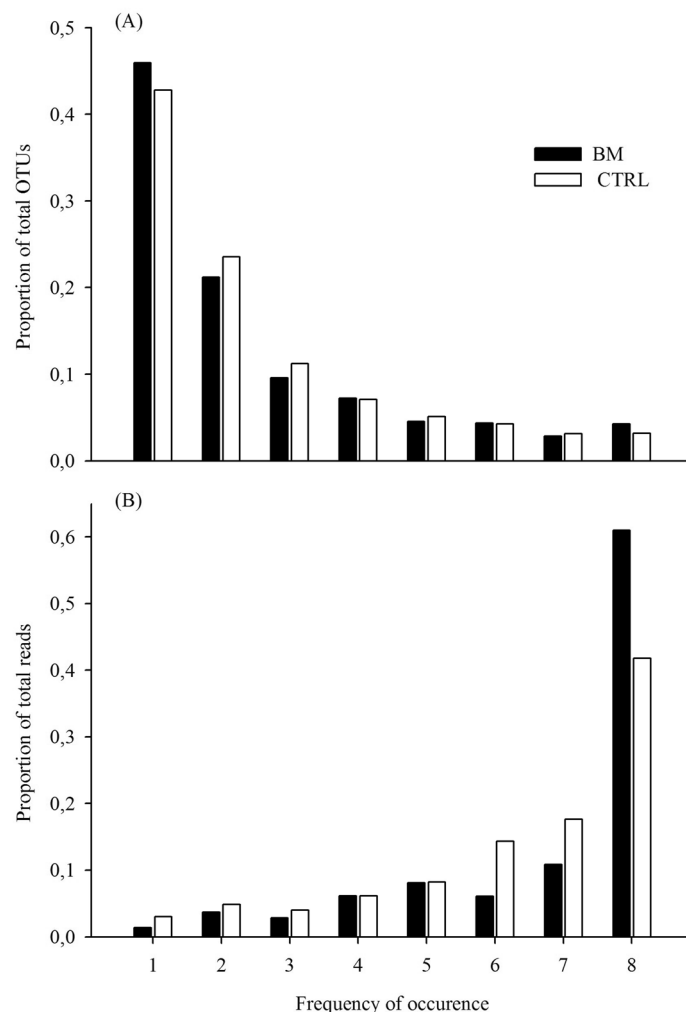


Figure 2. (A) Numbers of OTUs that were present in 1 to 8 biofilm communities of either treatment. (B) Summed read numbers of these OTUs. OTUs that occurred in all 8 communities of the CTRL treatment formed a significantly smaller proportions of total reads than in the BM treatment (Student's *t* test, $p < 0.05$).

Discussion

It is still challenging to accurately quantify the 'rare biosphere' of potential biofilm colonizers in the bacterioplankton^{31,32} and, thus, to study biofilm assembly processes by direct comparison with the source community. However, other approaches help to identify factors that modulate the respective importance of species sorting, immigration and stochasticity in biofilms, e.g., the analysis of communities in the context of the biotic and abiotic variation of the surrounding water^{33,34}, or in their response to habitat manipulation³⁵.

The diversity (richness) of the GDM biofilm communities was significantly lower in the substrate amended BM treatments than in the CTRL. This is the opposite of what has been described for planktonic diversity in oligotrophic ponds³⁶, or groundwater microbial assemblages³⁵, but agrees with observations on microbial assemblages in eutrophied lake sediments or from a marine phytoplankton bloom^{37,38}. Such seemingly contradictory findings might potentially be reconciled by assuming a hump-shaped relationship between microbial diversity and productivity, as observed in soil³⁹. On the one hand, severe substrate limitation will impose restrictions on the survival of many genotypes and select for a small set of oligocarbophilic specialists. On the other hand, extremely high substrate levels will specifically favour the most rapidly growing 'opportunistic' genotypes, which in turn might negatively affect the growth of others by releasing waste products or toxins.

BM treatments also featured a significantly higher NST score than the CTRL. At a first glance, it appears paradoxical that these communities should be mainly shaped by neutral processes such as immigration and ecological drift at arguably more selective conditions. However, null-model-based approaches are very sensitive to regional pool pre-selection: Biofilm communities consist of a subset of the bacterioplankton source assemblage that is already strongly filtered by deterministic processes⁴. Our analysis thus assessed if the individual communities tended to be formed by bacteria that were able to colonize the biofilms and at the same time were specific to individual experiments ('deterministic' with respect to the impact of the local species pool) or by such genotypes that were present in all source assemblages ('stochastic', i.e., mainly formed by the regional species pool) (Fig. 4).

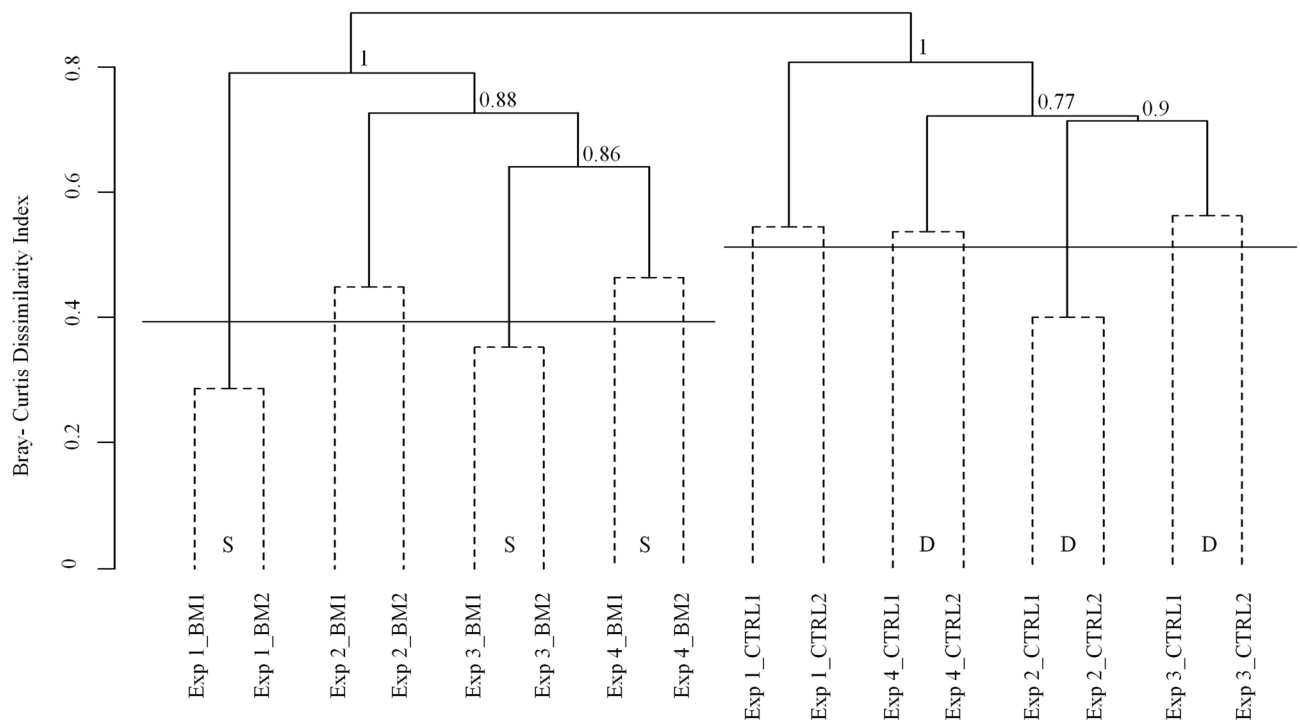


Figure 3. Average linkage clustering of all samples from the four experiments according to Bray–Curtis dissimilarity. The stability of the main branches was tested by bootstrapping (1,000 interactions). Similarity profile analyses ($\alpha = 0.001$) identified 8 significantly different groups that corresponded to the individual experiments (solid lines). The horizontal lines are the average dissimilarity of biological replicates within individual experiments (BM: 39%; CTRL: 51%), which is significantly smaller in the BM treatment (1-sided t test, $p < 0.05$). S, D: experimental replicates that were significantly more similar or dissimilar than expected by chance (modified Raup-Crick index > 0.95 or < -0.95 , respectively).

Thus, the high NST value of the BM treatments indicates that colonization of these biofilms was largely unrelated to priority effects, the compositional variability or the species turnover in the source water between experiments, and mainly driven by niche-related processes. The importance of environmental filtering in BM treatments is also suggested by the significantly higher similarity of the biological replicates within single experiment, i.e., lower Bray–Curtis distances (Fig. 3) and high Raup-Crick indices. Our findings agree with observations from parallel bioreactors, where higher substrate supply led to the formation of less divergent communities⁴⁰.

Competitive exclusion might be directly related to the higher availability of organic C and nutrients, e.g., by favouring more opportunistically growing genotypes⁴¹. Eutrophication and high organic load increased the importance of niche-driven selection in coastal bacterioplankton assemblages⁴². However, two indirect effects of biomass addition also need to be considered: For one, more extreme habitat conditions, in particular the more pronounced oxygen limitation due to higher respiration^{43,44}, might also have favoured niche-driven assembly processes in BM treatments. Environmental filtering was most prominent at conditions that most strongly deviated from those of the source assemblages (i.e., draught) in the sediments of a salt pan⁴⁵. Secondly, due to the added biomass the transmembrane flux in the BM treatments was on average between 2–3 times lower than in the CTRL treatments^{27–29,43}. Thus, the observed assembly patterns might at least in parts be due to ‘mass effects’, i.e., the successful colonizers of BM biofilms from the regional species pool (several of which are listed in Table 1) might have simply been more abundant in the source water than their competitors from the local pool.

The respective importance of assembly processes such as immigration and local selection is usually assessed at the level of entire microbial communities, implicitly assuming that these processes indiscriminately act on different community components. This has, e.g., been challenged by a separate assessment of the selection forces acting on common vs. rare members of microbial assemblages^{31,32,46}. Biofilms are spatially heterogeneous habitats, featuring a complex three-dimensional architecture with pronounced horizontal and vertical gradients of oxygen or substrates^{43,47}. It is conceivable that there might be differences in the assembly processes of distinct subcommunities that are either limited to particular microniches, such as obligate anaerobic bacteria⁴⁸, or that share fundamental life style related traits⁴⁹.

We found that niche-driven and neutral processes may selectively act on different compartments of GDM biofilm assemblages: In contrast to their aerobic counterparts, OTUs with an anaerobic metabolism were not selected from the local (i.e., experiment-specific) species pool in the CTRL treatment. Following the argument made above, we conclude that their establishment in GDM biofilms was thus mainly governed by niche-driven processes. Moreover, the clear difference in the composition of the anaerobic subcommunities between the two treatments suggests that anaerobic genotypes originated from several sources with different habitat properties, e.g. from sinking ‘lake snow’ particles⁵⁰ vs. from food pellets within zooplankton carcasses⁵¹. The contrasting

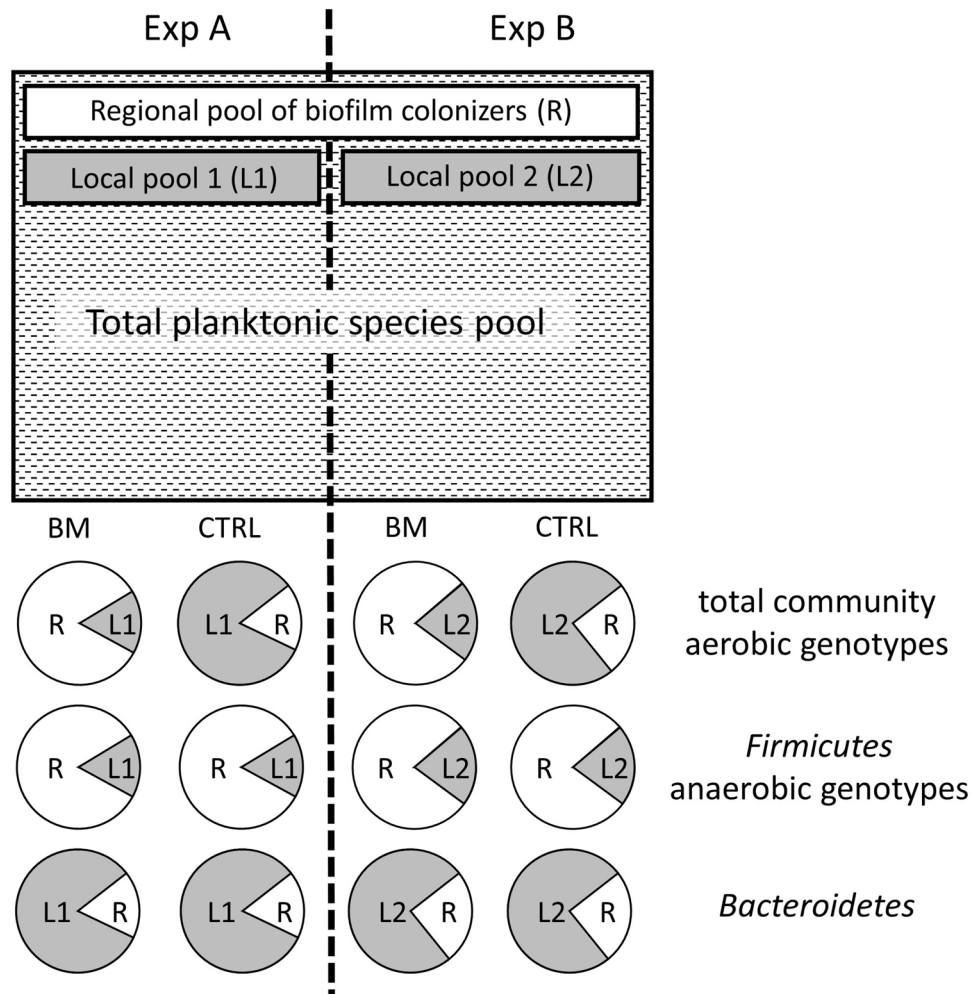


Figure 4. Conceptual depiction of biofilm community assembly in individual GDM experiments with (BM) and without (CTRL) addition of cyanobacterial biomass, as deduced by null model analysis (Normalized Stochasticity Ratio). The ‘regional pool’ of biofilm colonizers is defined as those species that were present in the feed water of all experiments, whereas the ‘local pool’ refers to genotypes that were specific to the feed water of a single experiment.

availability of organic C in the two treatments might have been a direct selective factor, yet it probably also shaped biofilm communities indirectly by defining the spatial patterns of oxygen distribution: On the one hand, interspersed anoxic microniches within a largely oxygenated matrix would lead to high habitat heterogeneity and variable selection in the CTRL biofilms⁴⁴. This agrees with the more than twice as high number of anaerobic OTUs that were exclusive to this treatment (159 vs. 70 OTUs). Moreover, 98% of the 605 reads affiliated to the *Nitrosomonadaceae* (microaerophilic lithoautotrophic ammonia oxidizers) originated from the CTRL treatment, suggesting an interplay of oxygen gradients and C limitation. On the other hand, the high loads of organic material in the BM treatments likely led to considerable lower redox conditions in large areas of these biofilms. This conclusion is supported by the observation that > 90% of reads from OTUs affiliated with sulphate reducing bacteria (total: 1536 reads) and with the strictly anaerobic *Clostridiales* (total: 6,413 reads) were found in this treatment. BM biofilms also featured tenfold higher read numbers of the most common ‘globally occurring’ OTU (Table 1) that is closely related to the facultative iron reducer *Rhodoferrax ferrireducens* (NR_074760, 98.8% sequence identity). In addition, there was evidence that microbes from individual phylogenetic groups did not adhere to the overall pattern of community assembly (Suppl. Table S3). In particular, OTUs affiliated with Bacteroidetes always tended to be recruited from the local species pool irrespective of treatment type, whereas the opposite was the case for Firmicutes (mainly consisting of Clostridiales classified as anaerobes). It has been argued that higher bacterial taxonomic ranks may to some extent also be ecologically meaningful units⁵², possibly because complex life style traits encoded by many genes are also phylogenetically conserved⁴⁹.

The addition of cyanobacterial biomass led to a significant shift towards a small subset of ‘globally occurring’ treatment-specific colonizers of GDM biofilms (Fig. 2) both, within and between experiments. Moreover, the treatment types could be unambiguously separated by only considering 5 ubiquitous genotypes that all had 5 to 10 times higher read numbers in the BM treatment (Table 1). Our observation of a few ‘universally occurring’

| Experiment | Start date of experiment | Duration (days) | Dry cell weight of the added biomass (BM treatment) (mg day ⁻¹) |
|--------------------------------|--------------------------|-----------------|---|
| 1. Kohler et al. ²⁹ | 16.06.2011 | 21 | 14.8 |
| 2. Silva et al. ²⁷ | 04.06.2014 | 23 | 15.2 |
| 3. This study | 16.10.2014 | 23 | 13.9 |
| 4. Silva et al. ²⁷ | 17.02.2017 | 30 | 17 |

Table 2. Information on the 4 experiments Citations are given for published data sets.

indicator species for the addition of cyanobacterial biomass to GDM biofilms is derived from experiments conducted in 2011, 2014 and 2017, and thus appears to be rather robust with respect to variability of the source assemblages. Moreover, all 5 of the above mentioned OTUs were also present in biomass-amended GDM biofilms fed with water from a stream²⁸. Two of them, *Rhodospirillum rubrum* sp. (LN870966) and *Undibacterium seohonense* (KC735151), also proliferated in GDM biofilms after addition of starch⁴³, suggesting that their role as indicators might go beyond a particular lake and carbon source. Metabarcoding of pro- and eukaryotic marker genes from different habitats -including biofilms⁵³- is increasingly regarded as a powerful tool for future environmental monitoring^{54,55}. Recently, Keeley et al. reported that a small number of abundant microbial genotypes within operationally defined 'eco-groups' could accurately predict the enrichment of benthic habitats in salmon farms⁵⁶. It is conceivable that bacterial biofilms in various technical or natural systems might also feature a small set of omnipresent reliable meters for fundamental aspects of habitat conditions such as substrate levels. Pronounced cyanobacterial blooms are a frequent phenomenon in many eutrophied lacustrine systems, resulting in concentrations of particulate organic carbon that are in the range of our experimental addition in the BM treatments⁵⁷. Our study might thus help to develop a better understanding of the microbial communities that develop on ultrafiltration membranes or in sand filters fed with water from such systems.

Materials and methods

Experimental systems. Microbial biofilms grew on the ultrafiltration membranes (150 kDa nominal cutoff, polyethersulfone membrane; Microdyn Nadir, Wiesbaden, Germany) of experimental gravity-driven water filtration (GDM) systems. Water from Lake Zurich, Switzerland (location: 47°19'13.24" N, 8°33'11.86" E) obtained from the aerobic layer at 5 m depth was used as continuous feed in four independent experiments conducted over a period of 6 years (Table 2)^{27–29}. Lake water was collected in sedimentation tanks and the GDM systems were supplied from the overflow. More details and a graphic depiction of the experimental systems are given in Silva et al.²⁷ and ²⁸.

In each experiment, two treatments were run in parallel with the same feed water: One set of GDMs (2–3 biological replicates per experiment) received a daily dose of destroyed biomass from an axenic culture of *Microcystis aeruginosa* PCC7806 (BM treatment, Table 2). Another set of GDMs was operated without additional manipulation (CTRL treatment) until biofilm maturation (2–3 weeks). The input of particulate organic carbon from lake water to the CTRL treatments was by approximately 2 orders of magnitude smaller than that of the BM treatments, as estimated from the concentrations of particulate organic matter (POC) in Lake Zurich⁵⁸ and daily flux rates through GDM systems^{27,28,43}.

DNA extraction and sequencing. At the end of each experiment, the biofilms were collected and DNA was extracted using the DNeasy PowerBiofilm Kit (Qiagen, Germany) according to the manufacturer's specification (but extending the removal step for inhibitors to one hour). The recovered DNA was dissolved in 10 mM Tris buffer and stored at -20 °C until further processing.

Partial 16S rRNA genes were amplified with a primer pair that excludes chloroplasts and cyanobacteria (799F-1115R)^{59,60}, but using a modified reverse primer for increased coverage²⁷. Sequences were obtained by Illumina MiSeq paired end (2 × 300 bp) sequencing. Sequencing data from two of the four experiments (Exp 2, 4) have been published before, albeit separately^{27,28}. The stored DNA extracts from another published experiment (Exp 1)²⁹ were reamplified and resequenced on the Illumina platform for the purpose of this study. Samples from all experiments were processed by the same company (LGC Genomics, Germany).

Amplicon sequence data from two biological replicates of each treatment type from all four experiments were collectively re-analysed in order to produce a single, coherent set of operational taxonomic units (OTUs). The raw sequence data processing, the definition of OTUs and their taxonomic assignment according to the SILVA taxonomy (version 132)⁶¹ were performed by an in-house pipeline as described previously²⁷.

Data treatment and statistical analysis. Data analysis was carried out in R⁶². Biofilm communities were clustered by the average linkage method based on their Bray–Curtis dissimilarity. This was done for the complete dataset and separately for each of the phylogenetic groups depicted in Fig. 1. Similarity profile analysis was used to establish significant clusters (1,000 simulations, $\alpha = 0.001$) and the stability of the main branches was estimated by bootstrapping (1,000 bootstraps). These calculations were performed with the R packages *clustsig*, *vegan*, and *fpc*^{63–65}. Similarity percentage (Simpser) analysis based on Bray–Curtis dissimilarity (R package *vegan*) was used to examine which of the 'core' genotypes that were present in all samples (Table 2) significantly contributed to the separation between treatments.

Differences in various community parameters between the BM and CTRL treatments were tested for significance by Student's t-tests, following a prior testing for normality (Shapiro–Wilk test) and homoscedasticity (Levene's test) with the R package *car*^{66–68}. Proportional data were logit transformed prior to statistical testing⁶⁹. In order to account for a possible confounding effect of differences in alpha diversity between treatments⁷⁰ we also calculated the pairwise Raup-Crick indices (RC) of the biological replicates using the R package *NST*⁷¹.

A Random Forest (RF) classification (R package *randomForest*) was performed to assess how accurately the samples could be assigned to treatments or experiments by means of their respective community composition, and to identify the responsible OTUs^{72,73}. The RF algorithm is a supervised machine learning model based on decision trees to classify data into pre-defined categories. The number of required decision trees (range: 1,000 to 50,000) was set to a value where the out-of-bag (OOB) error was stable. The Gini Impurity Metric was used as criterion to identify OTUs that were most relevant for the classification.

The influence of the 'local' and 'regional' pools of OTUs (i.e., OTUs that tended to be more or less specific for a single experiment or treatment) on biofilm community assembly in the two treatments was estimated by a null model analysis using the Normalized Stochasticity Ratio (NST; R package *NST*)⁷¹. NST values > 50% generally point to a predominance of stochastic processes on community assembly⁷¹. In our specific context, high NST values speak for the importance of biofilm colonizers from the 'regional pool' of OTUs that were present in similar read numbers across several or all experiments. Differences between NST values were assessed for statistical significance by bootstrap analyses (1,000 bootstraps).

Received: 3 February 2020; Accepted: 15 June 2020

Published online: 14 July 2020

References

- Flemming, H.-C. Why microorganisms live in biofilms and the problem of biofouling. In *Marine and Industrial Biofouling* (eds Flemming, H.-C. *et al.*) 3–12 (Springer, Berlin, 2018).
- Wahl, M., Goecke, F., Labes, A., Dobretsov, S. & Weinberger, F. The second skin: Ecological role of epibiotic biofilms on marine organisms. *Front. Microbiol.* **3**, 292 (2012).
- Coetser, S. E. & Cloete, T. E. Biofouling and biocorrosion in industrial water systems. *Crit. Rev. Microbiol.* **31**, 213–232 (2005).
- Besemer, K. *et al.* Unraveling assembly of stream biofilm communities. *ISME J.* **6**, 1459–1468 (2012).
- Eckert, E. M. & Pernthaler, J. Bacterial epibionts of *Daphnia*: a potential route for the transfer of dissolved organic carbon in freshwater food webs. *ISME J.* **8**, 1808–1819 (2014).
- Freese, H. M. & Schink, B. Composition and stability of the microbial community inside the digestive tract of the aquatic crustacean *Daphnia magna*. *Microb. Ecol.* **62**, 882–894 (2011).
- Hempel, M., Blume, M., Blindow, I. & Gross, E. M. Epiphytic bacterial community composition on two common submerged macrophytes in brackish water and freshwater. *BMC Microbiol.* **8**, (2008).
- Rahalkar, M., Bahulikar, R. A., Deutzmann, J. S., Kroth, P. G. & Schink, B. *Elstera litoralis* gen. nov., sp. nov., isolated from stone biofilms of lake constance, Germany. *Int. J. Syst. Evol. Microbiol.* **62**, 1750–1754 (2012).
- Berlanga, M. & Guerrero, R. Living together in biofilms: The microbial cell factory and its biotechnological implications. *Microb. Cell Fact.* **15**, 165 (2016).
- Salcher, M. M., Pernthaler, J. & Posch, T. Seasonal bloom dynamics and ecophysiology of the freshwater sister clade of SAR11 bacteria that rule the waves (LD12). *ISME J.* **5**, 1242–1252 (2011).
- Neuenschwander, S. M., Ghai, R., Pernthaler, J. & Salcher, M. M. Microdiversification in genome-streamlined ubiquitous freshwater Actinobacteria. *ISME J.* **12**, 185–198 (2018).
- Tsuchiya, Y. *et al.* Nutrient-rich microhabitats within biofilms are synchronized with the external environment. *Microbes Environ.* **24**, 43–51 (2009).
- Tang, X. *et al.* Characterization of bacterial communities associated with organic aggregates in a large, shallow, eutrophic freshwater Lake (Lake Taihu, China). *Microb. Ecol.* **58**, 307–322 (2009).
- Cordero, O. X. & Datta, M. S. Microbial interactions and community assembly at microscales. *Curr. Opin. Microbiol.* **31**, 227–234 (2016).
- Matar, G. K., Bagchi, S., Zhang, K., Oerther, D. B. & Saikaly, P. E. Membrane biofilm communities in full-scale membrane bioreactors are not randomly assembled and consist of a core microbiome. *Water Res.* **123**, 124–133 (2017).
- Bižić-Ionescu, M. *et al.* Comparison of bacterial communities on limnic versus coastal marine particles reveals profound differences in colonization. *Environ. Microbiol.* **17**, 3500–3514 (2015).
- Grossart, H. P., Dziallas, C., Leunert, F. & Tang, K. W. Bacteria dispersal by hitchhiking on zooplankton. *Proc. Natl. Acad. Sci.* **107**, 11959–11964 (2010).
- Monard, C., Gantner, S., Bertilsson, S., Hallin, S. & Stenlid, J. Habitat generalists and specialists in microbial communities across a terrestrial-freshwater gradient. *Sci. Rep.* **6** (2016).
- Allgair, M. & Grossart, H. P. Seasonal dynamics and phylogenetic diversity of free-living and particle-associated bacterial communities in four lakes in northeastern Germany. *Aquat. Microb. Ecol.* **45**, 115–128 (2006).
- Olsen, N. M. C. *et al.* Priority of Early Colonizers but No Effect on Cohabitants in a Synergistic Biofilm Community. *Front. Microbiol.* **10**, article 1949 (2019).
- Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S. & Thomas, T. Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. USA* **108**, 14288–14293 (2011).
- Salcher, M. M., Pernthaler, J., Frater, N. & Posch, T. Vertical and longitudinal distribution patterns of different bacterioplankton populations in a canyon-shaped, deep prealpine lake. *Limnol. Oceanogr.* **56**, 2027–2039 (2011).
- Eiler, A., Heinrich, F. & Bertilsson, S. Coherent dynamics and association networks among lake bacterioplankton taxa. *ISME J.* **6**, 330–342 (2012).
- Olapade, O. A. & Leff, L. G. Influence of dissolved organic matter and inorganic nutrients on the biofilm bacterial community on artificial substrates in a northeastern Ohio, USA, stream. *Can. J. Microbiol.* **52**, 540–549 (2006).
- Zhou, J. *et al.* Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. *MBio* **4** (2013).
- Pronk, W. *et al.* Gravity-driven membrane filtration for water and wastewater treatment: A review. *Water Res.* **149**, 553–565 (2019).
- Silva, M. O. D., Blom, J. F., Yankova, Y., Villiger, J. & Pernthaler, J. Priming of microbial microcystin degradation in biomass-fed gravity driven membrane filtration biofilms. *Syst. Appl. Microbiol.* **41**, 221–231 (2018).

28. Silva, M. O. D., Desmond, P., Derlon, N., Morgenroth, E. & Pernthaler, J. Source community and assembly processes affect the efficiency of microbial microcystin degradation on drinking water filtration membranes. *Front. Microbiol.* **10**, article 843 (2019).
29. Kohler, E. *et al.* Biodegradation of microcystins during gravity-driven membrane (GDM) ultrafiltration. *PLoS ONE* **9**, e111794 (2014).
30. Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D. & Bertilsson, S. A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 14–49 (2011).
31. Mo, Y. *et al.* Biogeographic patterns of abundant and rare bacterioplankton in three subtropical bays resulting from selective and neutral processes. *ISME J.* **12**, 2198–2210 (2018).
32. Yamamoto, K. *et al.* Diversity and geochemical community assembly processes of the living rare biosphere in a sand-and-gravel aquifer ecosystem in the Midwestern United States. *Sci. Rep.* **9** (2019).
33. Langenheder, S. *et al.* Bacterial metacommunity organization in a highly connected aquatic system. *FEMS Microbiol. Ecol.* **93**, 225 (2017).
34. Zhao, D. *et al.* Community composition and assembly processes of the free-living and particle-attached bacteria in Taihu Lake. *FEMS Microbiol. Ecol.* **93**, 1–10 (2017).
35. Zhou, J. *et al.* Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc. Natl. Acad. Sci. USA* **111**, E836–E845 (2014).
36. Chase, J. M. Stochastic Community Assembly Causes Higher Biodiversity in More Productive Environments. *Science (80-)* **328**, 1388–1391 (2010).
37. Wan, Y., Ruan, X., Zhang, Y. & Li, R. Illumina sequencing-based analysis of sediment bacteria community in different trophic status freshwater lakes. *Microbiologyopen* **6**, 1–15 (2017).
38. Yang, C. *et al.* Illumina sequencing-based analysis of free-living bacterial community dynamics during an Akashiwo sanguine bloom in Xiamen sea, China. *Sci. Rep.* **5**, 1–11 (2015).
39. Song, W., Kim, M., Tripathi, B. M., Kim, H. & Adams, J. M. Predictable communities of soil bacteria in relation to nutrient concentration and successional stage in a laboratory culture experiment. *Environ. Microbiol.* **18**, 1740–1753 (2016).
40. van der Gast, C. J., Ager, D. & Lilley, A. K. Temporal scaling of bacterial taxa is influenced by both stochastic and deterministic ecological factors. *Environ. Microbiol.* **10**, 1411–1418 (2008).
41. Pernthaler, A., Pernthaler, J., Eilers, H. & Amann, R. Growth patterns of two marine isolates: Adaptations to substrate patchiness?. *Appl. Environ. Microbiol.* **67**, 4077–4083 (2001).
42. Dai, W. *et al.* Bacterioplankton assembly and interspecies interaction indicating increasing coastal eutrophication. *Chemosphere* **177**, 317–325 (2017).
43. Silva, M. O. D. & Pernthaler, J. Priming of microcystin degradation in carbon-amended membrane biofilm communities is promoted by oxygen limited conditions. *FEMS Microbiol. Ecol.* **95**, fiz157 (2019).
44. Dini-Andreote, F., Stegen, J. C., Van Elsland, J. D. & Salles, J. F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. USA* **112**, E1326–E1332 (2015).
45. Valverde, A., Makhallanyane, T. P. & Cowan, D. A. Contrasting assembly processes in a bacterial metacommunity along a desiccation gradient. *Front. Microbiol.* **5**, article 668 (2014).
46. Zhang, H. *et al.* Interaction and assembly processes of abundant and rare microbial communities during a diatom bloom process. *Environ. Microbiol.* **00**, 1–13 (2019).
47. Stewart, P. S. & Franklin, M. J. Physiological heterogeneity in biofilms. *Nat. Rev. Microbiol.* **6**, 199–210 (2008).
48. Jackson, C. R. Changes in community properties during microbial succession. *Oikos* **101**, 444–448 (2003).
49. Martiny, A. C., Treseder, K. & Pusch, G. Phylogenetic conservatism of functional traits in microorganisms. *ISME J.* **7**, 830–838 (2013).
50. Ploug, H., Kühl, M., Buchholz-Cleven, B. & Jørgensen, B. B. Anoxic aggregates—An ephemeral phenomenon in the pelagic environment?. *Aquat. Microb. Ecol.* **13**, 285–294 (1997).
51. Glud, R. N. *et al.* Copepod carcasses as microbial hot spots for pelagic denitrification. *Limnol. Oceanogr.* **60**, 2026–2036 (2015).
52. Philippot, L. *et al.* The ecological coherence of high bacterial taxonomic ranks. *Nat. Rev. Microbiol.* **8**, 523–529 (2010).
53. Pu, Y., Ngan, W. Y., Yao, Y. & Habimana, O. Could benthic biofilm analyses be used as a reliable proxy for freshwater environmental health?. *Environ. Pollut.* **252**, 440–449 (2019).
54. Pawlowski, J. *et al.* The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Sci. Total Environ.* **637–638**, 1295–1310 (2018).
55. Vasselon, V., Rimet, F., Tapolczai, K. & Bouchez, A. Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). *Ecol. Indic.* **82**, 1–12 (2017).
56. Keeley, N., Wood, S. A. & Pochon, X. Development and preliminary validation of a multi-trophic metabarcoding biotic index for monitoring benthic organic enrichment. *Ecol. Indic.* **85**, 1044–1057 (2018).
57. Ye, L. *et al.* Dynamics of dissolved organic carbon after a cyanobacterial bloom in hypereutrophic Lake Taihu (China). *Limnologia* **41**, 382–388 (2011).
58. Van den Wyngaert, S., Salcher, M. M., Pernthaler, J., Zeder, M. & Posch, T. Quantitative dominance of seasonally persistent filamentous cyanobacteria (*Planktothrix rubescens*) in the microbial assemblages of a temperate lake. *Limnol. Oceanogr.* **56**, 97–109 (2011).
59. Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y. & Fierer, N. The ecology of the phyllosphere: Geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ. Microbiol.* **12**, 2885–2893 (2010).
60. Chelius, M. K. & Triplett, E. W. The diversity of archaea and bacteria in association with the roots of *Zea mays* L. *Microb. Ecol.* **41**, 252–263 (2001).
61. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).
62. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2019).
63. Hennig, C. *fpclust: Flexible Procedures for Clustering*. R Packag. version 2.2–3. <https://CRAN.R-project.org/package=fpclust> (2019).
64. Whitaker, D. & Christman, M. *clustsig: Significant Cluster Analysis*. R Packag. version 1.1. <https://CRAN.R-project.org/package=clustsig> (2014).
65. Oksanen, J. *et al.* *vegan: Community Ecology Package*. R Packag. version 1.1. <https://CRAN.R-project.org/package=vegan> (2019). doi:10.4135/9781412971874.n145
66. Shapiro, S. S. & Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).
67. Fox, J. & Weisberg, S. *An R Companion to Applied Regression* (Sage, Thousand Oaks, 2019).
68. Levene, H. Robust testes for equality of variances. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (ed. Olkin, I.) 278–292 (Stanford University Press, Stanford, 1960).
69. Warton, D. I. & Hui, F. K. C. The arcsine is asinine: the analysis of proportions in ecology. *Ecology* **92**, 3–10 (2011).
70. Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M. & Inouye, B. D. Using null models to disentangle variation in community dissimilarity from variation in α -diversity. *Ecosphere* **2**, art24 (2011).
71. Ning, D., Deng, Y., Tiedje, J. M. & Zhou, J. A general framework for quantitatively assessing ecological stochasticity. *Proc. Natl. Acad. Sci.* **116**, 16892–16898 (2019).

72. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).

73. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2**, 18–22 (2002).

Acknowledgements

We would like to thank Dr. Daliang Ning for his insightful comments on our results. This study was supported by the Swiss National Fund (Proj. Nr. CR32I2_149648 and 31003A_182336).

Author contributions

J.P. and M.S. planned the experiments. M.S. conducted the experimental work. M.S. and J.P. together analyzed the data. M.S. produced all figures and tables, and J.P. wrote the main manuscript text. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-68460-x>.

Correspondence and requests for materials should be addressed to J.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020