environmental microbiology

Environmental Microbiology (2014) **16**(8), 2514–2524 doi:10.1111/1462-2920.12392

Rare but active taxa contribute to community dynamics of benthic biofilms in glacier-fed streams

Linda Wilhelm,[1] Katharina Besemer,[1,2] Christina Fasching,[1] Tim Urich,[3] Gabriel A. Singer,[1,2] Christopher Quince[4] and Tom J. Battin[1,2]*

Departments of [1]Limnology and Bio-Oceanography and [3]Ecogenomics and Systems Biology, University of Vienna, Althanstr. 14, A-1090 Vienna, Austria.

[2]WasserCluster Lunz, Dr Carl Kupelwieser Promenade 5, A-3293 Lunz am See, Austria.

[4]Department of Civil Engineering, University of Glasgow, Oakfield Avenue, Glasgow G12 8LT, UK.

Summary

Glaciers harbour diverse microorganisms, which upon ice melt can be released downstream. In glacier-fed streams microorganisms can attach to stones or sediments to form benthic biofilms. We used 454-pyrosequencing to explore the bulk (16S rDNA) and putatively active (16S rRNA) microbial communities of stone and sediment biofilms across 26 glacier-fed streams. We found differences in community composition between bulk and active communities among streams and a stronger congruence between biofilm types. Relative abundances of rRNA and rDNA were positively correlated across different taxa and taxonomic levels, but at lower taxonomic levels, the higher abundance in either the active or the bulk communities became more apparent. Here, environmental variables played a minor role in structuring active communities. However, we found a large number of rare taxa with higher relative abundances in rRNA compared with rDNA. This suggests that rare taxa contribute disproportionately to microbial community dynamics in glacier-fed streams. Our findings propose that high community turnover, where taxa repeatedly enter and leave the 'seed bank', contributes to the maintenance of microbial biodiversity in harsh ecosystems with continuous environmental perturbations, such as glacier-fed streams.

Received 25 July, 2013; accepted 1 January, 2014. *For correspondence. E-mail tom.battin@univie.ac.at; Tel. (+43) 1 4277 76410; Fax (+43) 1 4277 9572.

Introduction

Ice covers approximately 10% of the Earth's surface and represents the largest freshwater reservoir (Barnett *et al.*, 2005; Anesio and Laybourn-Parry, 2012; Stibal *et al.*, 2012). Glaciers and the polar caps are now recognized as a biome that harbours diverse microbial communities including viruses, bacteria, protozoa and algae (Anesio and Laybourn-Parry, 2012). These microorganisms are actively involved in a number of biogeochemical processes in ice ecosystems, mediating carbon fixation and respiration (Hood *et al.*, 2009; Singer *et al.*, 2012), nitrogen fixation (Telling *et al.*, 2011), iron cycling (Hodson *et al.*, 2008), and methanogenesis (Lanoil *et al.*, 2009; Wadham *et al.*, 2012). Ice-borne microorganisms also partially modulate a pool of biogeochemically diverse dissolved organic matter, which, upon ice melt, is released from glaciers and ice caps downstream into glacier-fed streams and the ocean, where it supports heterotrophic metabolism (Hood *et al.*, 2009; Singer *et al.*, 2012).

While the englacial environment offers relatively constant conditions for microbial life, glacier-fed streams are highly dynamic, and microorganisms therein are exposed to massive temporal fluctuations. Discharge, channel stability, sediment transport, water temperature and solute chemistry can significantly fluctuate both diurnally and seasonally with major implications for benthic life (Hannah *et al.*, 2007; Milner *et al.*, 2010). Despite this apparent environmental harshness, biofilms – surface-attached and matrix-enclosed microbial communities – dwell in glacier-fed streams (Battin *et al.*, 2001; 2004; Freimann *et al.*, 2013; Wilhelm *et al.*, 2013). In these streams, benthic sediments are continuously reworked and scoured, which is to some extent a function of grain size distribution. Differing susceptibility to physical disturbance and varying surface topology may affect structure and function of the microorganisms that are associated with differently sized sediment grains (DeFlaun and Mayer, 1983; Lock, 1993; Murray and Jumars, 2002).

There is evidence that microbial richness in glacier-fed streams may increase due to glacier retreat, while the turnover of communities among glacier-fed streams may decrease, ultimately reducing regional biodiversity (Wilhelm *et al.*, 2013). This study did not differentiate between putatively dormant and active taxa, which

according to theory may vary widely and contribute differently to community dynamics, the maintenance of biodiversity and ecosystem stability (Lennon and Jones, 2011). RNA has a much shorter lifespan than DNA and can therefore serve as a tool to capture putatively active microorganisms (Moeseneder *et al*., 2005). In fact, the growth of microorganisms is related to their ribosomal content, and while a single active cell can contain up to 10^4 ribosomes (Moeseneder *et al*., 2005), this number may markedly decrease during starvation (Fegatella *et al*., 1998). On the other hand, the detection of microorganisms on the 16S rDNA level mainly depends on their abundance. Rare microorganisms are not detectable at the DNA level but might still be metabolically active and thus identified at the RNA level. The 16S rRNA is therefore a valuable tool to investigate the potential activity of microbial communities and compare it with their abundance based on rDNA. Furthermore, this approach may be used to assess the relationship between abundance and activity, as was done for the coastal ocean and lakes (Jones and Lennon, 2010; Campbell *et al*., 2011; Campbell and Kirchman, 2013).

Microbial communities typically consist of a few dominant taxa, which contribute markedly to community functioning, and a large proportion of rare taxa, the microbial 'seed bank' (Pedrós-Alió, 2006). This 'seed bank' might constitute a reservoir of genetic diversity that could initiate changes in community composition as a response to environmental changes (Lennon and Jones, 2011). However, whether the rare microbial biosphere represents an active part of the community and its role for community dynamics and ecosystem functioning remains controversial (Pedrós-Alió, 2006; Kirchman *et al*., 2010; Lennon and Jones, 2011).

The present study reports on the community composition of benthic biofilms across a broad range of glacier-fed streams in the Austrian Alps. We deliberately sampled fine (i.e. epipsammic biofilms) and coarse sediments (i.e. epilithic biofilms) to cover the range of benthic microbial communities in glacier-fed streams and used 454-pyrosequencing on the 16S rDNA and the 16S rRNA to capture the bulk and the apparently active components of these communities.

We anticipated that the composition of the putatively active community reflects the contemporary environmental conditions in the glacier-fed streams. Furthermore, based on the assumption of varying susceptibility to environmental fluctuations, we also expected differences between epipsammic and epilithic biofilm communities. Finally, we hypothesized that a marked proportion of rare taxa is active in glacier-fed streams similar to the microbial communities in oceans and lakes (Jones and Lennon, 2010; Campbell *et al*., 2011; Campbell and Kirchman, 2013). We assumed that frequent shifts in relative abundances and resuscitation from dormancy as induced by environmental fluctuations, but also immigration from upstream communities, are all factors that result in a rare but active biosphere in glacier-fed streams.

Results

Alpha diversity and community composition of active and bulk communities

This study covered 26 glacier-fed streams located along the main chain of the Austrian Alps (Fig. 1, Supporting Information Tables S1 and S2) from where we collected epilithic and epipsammic biofilm samples and analysed the 16S rRNA and 16S rDNA using 454-pyrosequencing. After de-noising, our data set consisted of 734 124 reads (minimum 1074 reads per sample), which clustered into 13 610 operational taxonomic units (OTUs) at a 97% sequence similarity level (see rarefaction curves in Supporting Information Fig. S1).

Alpha diversity, estimated as rarefied richness and Shannon diversity, was significantly higher in bulk epipsammic biofilms than in active epipsammic biofilms (Wilcoxon test: Richness: $P < 0.05$, $n = 24$; Shannon: $P < 0.01$, $n = 24$; Supporting Information Fig. S2). Also, bulk epipsammic biofilms were significantly more diverse than bulk epilithic biofilms (Wilcoxon test: Richness, $P < 0.05$, $n = 24$; Shannon: $P < 0.001$, $n = 24$; Supporting Information Fig. S2). The 'true diversity', which provides a reliable estimate of the number of taxa in a community (Quince *et al*., 2008) did not differ significantly between groups (Supporting Information Fig. S2). We compared our results with data from pre-alpine stream biofilms (Besemer *et al*., 2013) and found that the alpha diversity, rarefied to the same number of reads, was comparable in both stream types with richness averaging 228 ± 75 (mean $\pm$ standard deviation) and Shannon diversity averaging 56 ± 39 in the glacier-fed streams, and with richness averaging 228 ± 67 and Shannon diversity averaging 82 ± 45 in the pre-alpine streams.

Community composition showed considerable variation among samples; for instance, no single OTU was found in all epilithic or all epipsammic samples. Non-metric multidimensional scaling analysis (nMDS) based on the presence-/absence-based Sørensen index and the abundance-based Morisita–Horn index (both calculated from rarefied data) was applied to illustrate community composition patterns (Fig. 2). A two-way permutational analysis of variance (PERMANOVA) with the factors 'epilithic-epipsammic' and 'DNA-RNA' revealed significant differences between both epilithic and epipsammic (Sørensen: pseudo-$F_{1,98} = 2.347$, $P < 0.001$; Morisita–Horn: pseudo-$F_{1,98} = 4.983$, $P < 0.001$), as well as between active and bulk communities (Sørensen: pseudo-$F_{1,98} = 8.636$, $P < 0.001$; Morisita–Horn: pseudo-

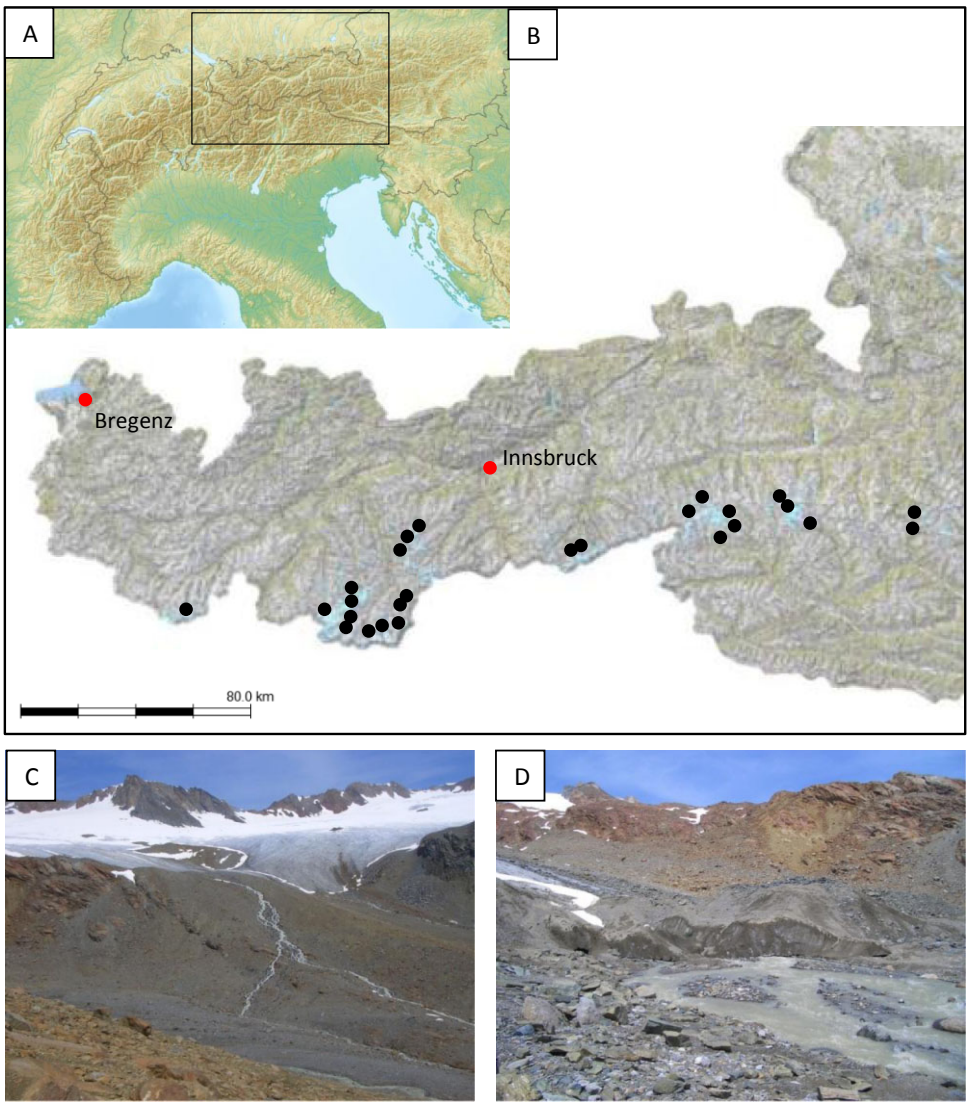

Fig. 1. Map and photographs of sampling sites. (A) European Alps (source: http://www.wikipedia.org) and (B) map of Western Austria including the 26 sampling sites (black circles). (C, D) Photographs show the Vernagtferner and its respective stream as sampled for this study.

$F_{1,98} = 8.578$, $P < 0.001$). Prompted by a significant interaction term for the Morisita–Horn index (pseudo-$F_{1,98} = 2.185$, $P < 0.01$), we performed pairwise comparisons, which revealed significant differences among all four groups ('epilithic-epipsammic': DNA: $F_{1,50} = 5.401$, $P < 0.001$; RNA: $F_{1,48} = 2.160$, $P < 0.01$; 'DNA-RNA': epilithic: $F_{1,48} = 3.567$, $P < 0.01$; epipsammic: $F_{1,50} = 7.314$, $P < 0.001$). We do recognize that the significant difference between active and bulk communities may be attributable to the use of different extraction methods. In fact, different extraction methods were required to yield sufficient nucleic acids from our samples. For the analysis based on the Morisita–Horn index, we found significantly different dispersions between active and bulk communities for both biofilm types (multivariate homogeneity of group dispersions: epilithic biofilm: $P < 0.01$; epipsammic biofilm: $P < 0.001$). Indeed, community composition is more variable for the active communities, and any difference in average composition between epilithic and epipsammic communities based on the Morisita-Horn index thus has to be interpreted with care.

Here, we used universal primers (Quince *et al.*, 2011) to capture a most comprehensive picture of the microbial diversity in the biofilms, including, besides Bacteria, also Archaea and chloroplasts. It is in fact well known that benthic biofilms in glacier-fed streams comprise abundant primary producers and can even include Archaea (Battin *et al.*, 2001). We found that 98.85% of all OTUs were classified as Bacteria, 0.40% as Archaea and 0.75% failed to be classified to any domain. At the

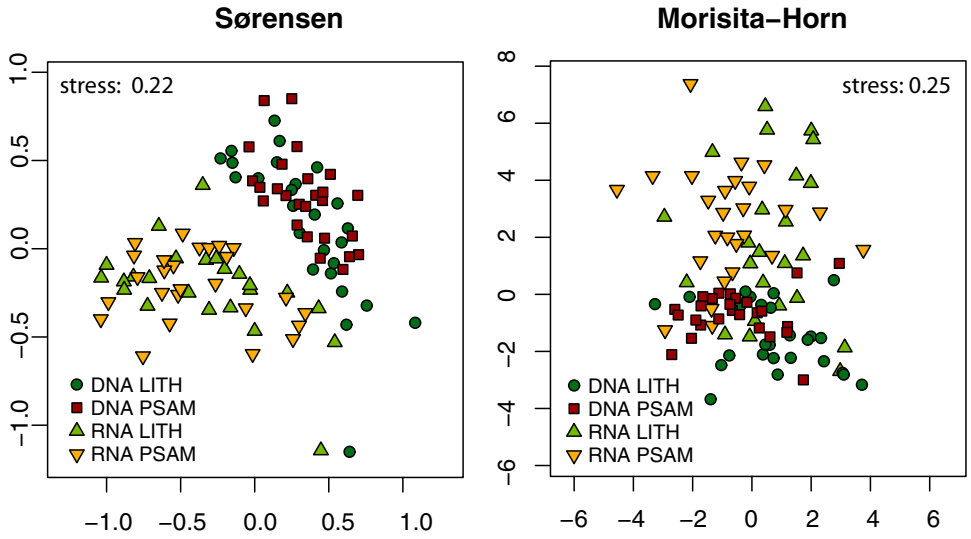

Fig. 2. Community composition of active and bulk biofilm communities. nMDS of microbial communities applying the presence-/absence-based Sørensen and the relative abundance-based Morisita–Horn similarity indices after resampling to 1074 reads. LITH, epilithic biofilm, PSAM, epipsammic biofilm.

bacterial phylum level, *Proteobacteria*, chloroplasts/*Cyanobacteria*, *Bacteroidetes* and *Actinobacteria* dominated both epilithic and epipsammic communities. The relative abundances of phyla in active and bulk communities were highly correlated (Fig. 3A; Spearman correlation: $r_s = 0.90$, $P < 0.001$ and $r_s = 0.79$, $P < 0.001$ in epilithic and epipsammic biofilms respectively). Among the most abundant phyla, *Firmicutes* deviated most from this relation with higher relative abundances in the active community (8.3% and 1.0% in the active and bulk communities respectively). At family level, the correlation of relative abundances between active and bulk communities was weaker (Fig. 3B; Spearman correlation: $r_s = 0.69$, $P < 0.001$ and $r_s = 0.65$, $P < 0.001$ in epilithic and epipsammic biofilms respectively). *Comamonadaceae* were highly abundant, both in active and in bulk communities (10.2% versus 14.7%), representing 305 OTUs affiliated to 14 different genera. Within this family, the most important genus was *Polaromonas*, which was represented by 50 OTUs. In epilithic biofilms, another dominant group was 'chloroplasts of *Bacillariophyta*', which was clearly less abundant in epipsammic biofilms.

Relationships between environmental variables and microbial communities

To identify possible relationships between microbial community composition and measured environmental variables, we used a canonical correspondence analysis (CCA; Supporting Information Fig. S3). In line with the ordination, a global test [analysis of variance (ANOVA)-like test; Oksanen *et al*., 2013] was computed to calculate the joint effect of constraints. Overall, the constraints of the CCA significantly explained community composition in all ordinations, except the active epilithic biofilm communities (ANOVA-like analysis: bulk epilithic biofilm: F-value = 1.140, $P < 0.05$, d.f. = 25; bulk epipsammic biofilm: F-value = 1.137, $P < 0.05$, d.f. = 25; active epilithic biofilm: F-value = 1.077, $P = 0.07$, d.f. = 24; active epipsammic biofilm: F-value = 1.238, $P < 0.05$, d.f. = 25). The global test was followed by a forward variable selection procedure to identify the particular contribution of each parameter (Oksanen *et al*., 2013). Streamwater pH and conductivity were significantly related to community composition in the bulk communities of epilithic (pH: $P < 0.05$, conductivity: $P < 0.01$) and epipsammic biofilms (pH: $P < 0.01$, conductivity: $P < 0.01$), while dissolved organic carbon (DOC) concentration explained a significant amount of variation in the community composition of active epipsammic biofilms ($P < 0.01$). Comparing the configurations of the CCA, we found significant congruence in the ordinations of active and bulk communities; this relationship was less pronounced for epipsammic than for epilithic biofilms (Procrustes test: epilithic biofilm: $r = 0.70$, $P < 0.001$; epipsammic biofilm: $r = 0.49$, $P < 0.001$).

Linking abundance and activity of microorganisms

To test for differences in the potential activity of abundant and rare taxa, we assigned the OTUs of each bulk community to an abundant (i.e. the most abundant 20%) and a rare group. OTUs were scored as 'active' if their rRNA relative abundance was greater than their rDNA relative

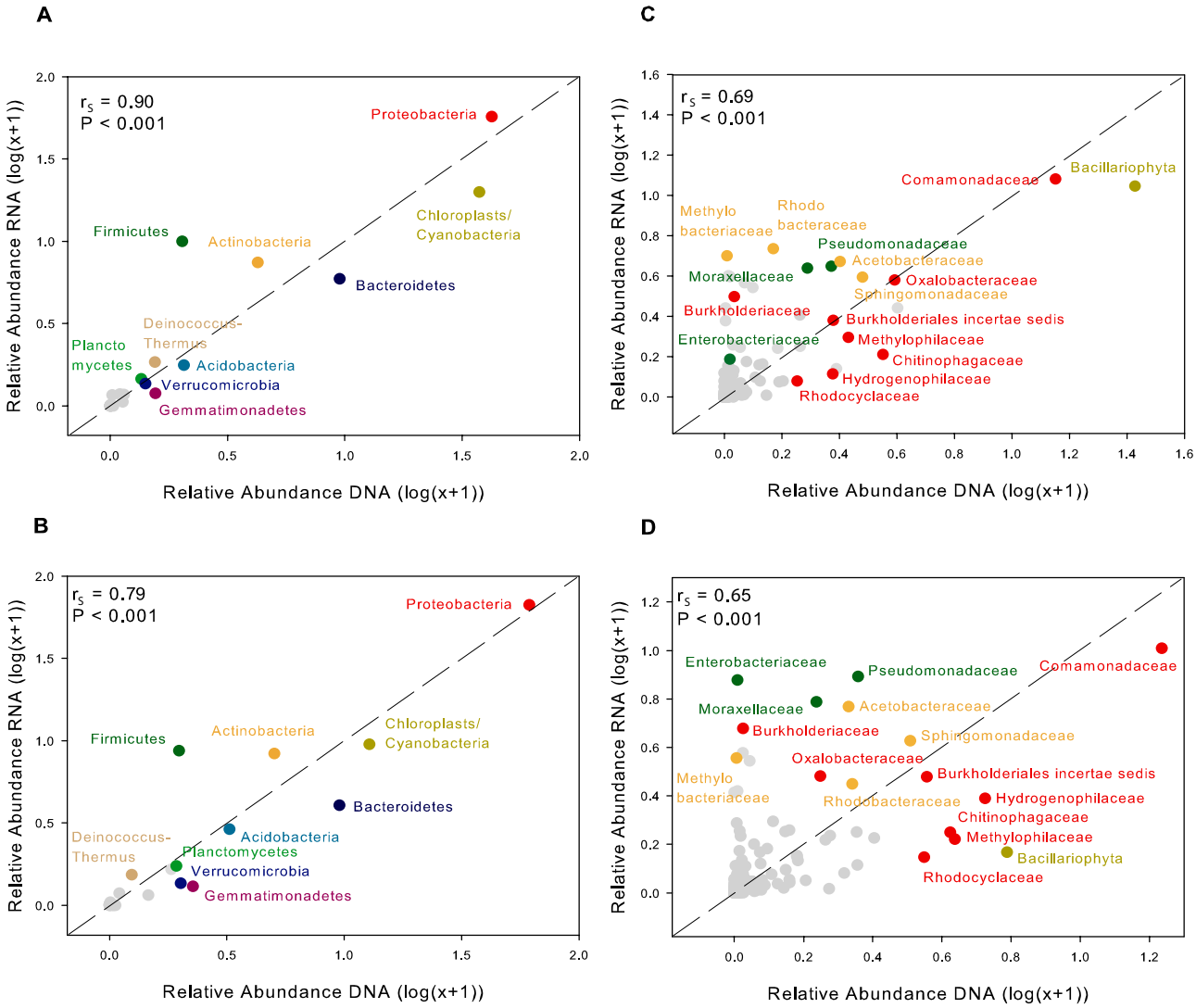

Fig. 3. Correlation of relative abundances of taxa in active and bulk communities at phylum (A, B) and at family (C, D) level. (A, C) represent epilithic biofilm communities, while (B, D) illustrate epipsammic biofilm communities. Relative abundances were log-transformed to reduce skewness of the data. Grey points indicate taxa, which occurred only in low numbers. At family level, taxa are coloured according to classes, where orange points represent *Alphaproteobacteria*, red points represent *Betaproteobacteria* and green points represent *Gammaproteobacteria*.

abundance, all others as 'inactive' (Fig. 4A); this is consistent with the approach proposed by Jones and Lennon (2010). The ratios of 'active' to 'inactive' OTUs were significantly higher in the rare group than in the abundant group of OTUs (Wilcoxon test: epilithic biofilm: $P < 0.001$, $n = 24$; epipsammic biofilm: $P < 0.001$, $n = 26$; Fig. 4B). This pattern was robust against changes in the threshold used to classify OTUs as abundant (see Experimental procedures and Supporting Information Table S3). The similarity between active and bulk communities from one glacier was generally low; after rarefaction to the same number of reads, active and bulk communities shared 11.5% ± 6.0% of the OTUs, constituting 45.9% ± 20.9% of the reads.

Discussion

Glaciers are now recognized as ecosystems harbouring diverse microbial communities (Simon *et al.*, 2009; Anesio and Laybourn-Parry, 2012; Wilhelm *et al.*, 2013). In order to capture the apparently active microbial community and thus provide a first insight beyond 'who is there' in glacier-fed streams, we contrasted community structure and biodiversity based on rRNA and rDNA. This strategy has been successfully applied in lakes and in the ocean (Jones and Lennon, 2010; Campbell *et al.*, 2011; Lanzén *et al.*, 2011; Campbell and Kirchman, 2013).

The significantly higher rRNA : rDNA ratios in rare taxa compared with abundant taxa in this study suggest that a

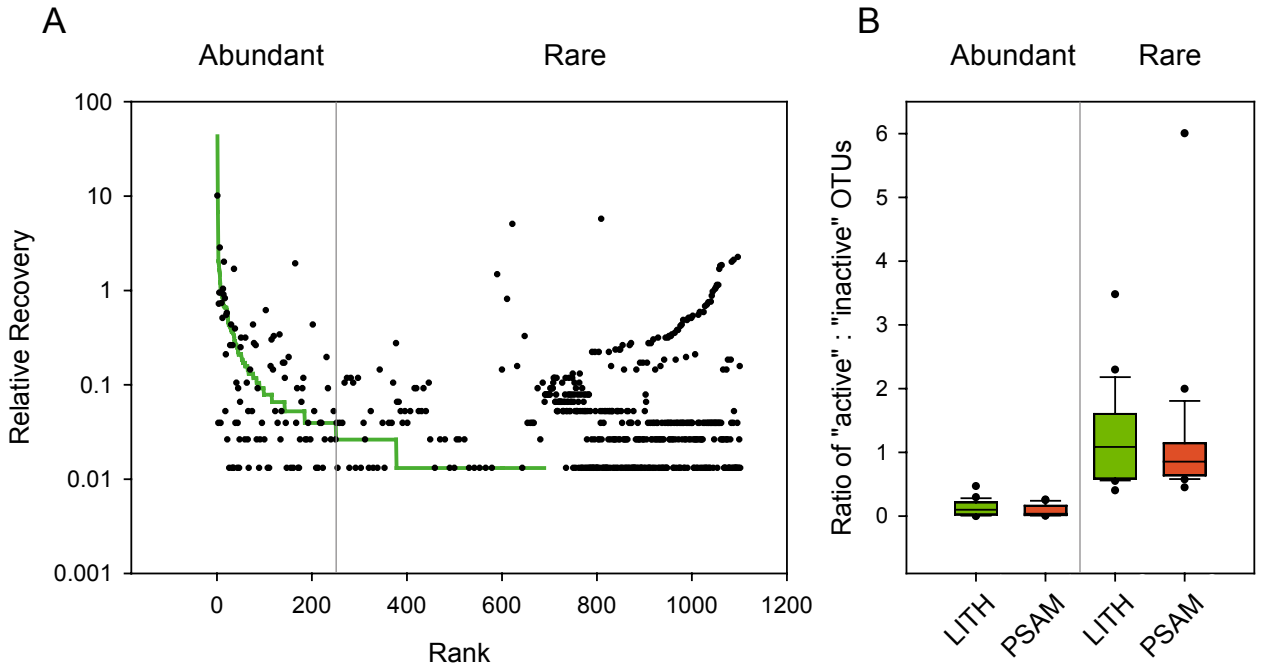

Fig. 4. Apparent activity of abundant and rare OTUs.
A. Example of a rank abundance curve for the bulk biofilm community (rDNA) of one glacier-fed stream. Points represent the relative recovery of the corresponding OTUs in the active community (rRNA). Points, which have the same rDNA rank, are shown in random order. Points above the curve are scored as 'active', while points below the curve are scored as 'inactive'. For better readability, we used a logarithmic scale for the *y*-axis. This scale, however, does not allow plotting of zero values, which are therefore not shown in the graph.
B. Boxplot representation of ratios of 'active' to 'inactive' scored OTUs among glacier-fed streams. Abundant and rare groups were defined by applying a 20% threshold according to the rDNA rank. LITH, epilithic biofilm; PSAM, epipsammic biofilm.

remarkable number of rare taxa in glacier-fed streams are disproportionately active. This metabolically active rare biosphere is potentially important for the functioning of microbial communities and the temporal dynamics of community composition (Jones and Lennon, 2010; Campbell *et al*., 2011; Campbell and Kirchman, 2013). Disproportionally active rare taxa might increase in abundance over time, resulting in microbial rank abundance curves that change accordingly (Jones and Lennon, 2010; Lennon and Jones, 2011).

In line with Lennon and Jones (2011), we argue that in fluctuating environments such as glacier-fed streams, where water chemistry and temperature can change seasonally and diurnally (Hannah *et al*., 2007; Milner *et al*., 2010), frequent transitions between active and dormant states may be advantageous to maintain biodiversity. DOC concentration is generally low in glacier-fed streams and labile fractions are often restricted to pulses during peak primary production (Battin *et al*., 2004). These dynamics may affect the physiology and metabolic machinery of microorganisms where superior competitors for resources may grow actively, while others may switch to transient dormancy. Furthermore, glacier retreat was recently proposed to affect upstream sources of microbial diversity with consequences for the microbial communities in glacier-fed streams (Wilhelm *et al*., 2013). Upstream sources could sporadically subsidize the communities in glacier-fed streams with novel taxa (Battin *et al*., 2001; Wilhelm *et al*., 2013), which would arguably contribute to the dynamics of a rare but active component of the local community.

We found alpha diversity in glacier-fed streams to be comparable with low-elevation streams as reported by Besemer and colleagues (2013). This is remarkable given the nutrient- and energy-poor environment of glacier-fed streams and their permanently low temperature. Transient dormancy could be a candidate mechanism that maintains high microbial diversity in these oligotrophic systems (Jones and Lennon, 2010).

Active and bulk communities were similar at the phylum level, indicating that the observed differentiation between these groups resulted from differences at lower phylogenetic levels. Relative abundances of families of active and bulk communities were generally related but with some notable exceptions. For instance, *Alphaproteobacteria* and *Gammaproteobacteria* had higher relative abundances in the rRNA-based communities, while *Betaproteobacteria* dominated rDNA-based communities indicating different apparent activity of these classes. These patterns may to some extent be caused by our use

of different extraction methods for DNA and RNA, which, however, is contrasted by the overall strong correlation at the phylum level. Moreover, both the number of rRNA operons and the number of ribosomes per cell vary between taxonomic groups (Moeseneder *et al*., 2005; Blazewicz *et al*., 2013), which may attenuate the correlation between microbial activity and rRNA abundance. *Firmicutes* or *Gammaproteobacteria*, for example, are known for overall high and heterogeneous 16S rRNA gene copy numbers, which may overestimate their abundance and impair diversity estimates (Kembel *et al*., 2012; Větrovský and Baldrian, 2013). *Comamonadaceae* was the most abundant family in both epilithic and epipsammic biofilms, and the genus *Polaromonas* was dominating within this family. There is growing evidence that *Polaromonas* is typical for ice ecosystems, as corroborated by its occurrence in glacier ice and subglacial sediments using cultivation (Margesin *et al*., 2012), 16S rDNA clone libraries (Skidmore *et al*., 2005; Lanoil *et al*., 2009) and metagenomics (Simon *et al*., 2009). The global distribution of *Polaromonas* among ice ecosystems has been proposed to be caused by significant transport of dormant cells via high-elevation air currents (Darcy *et al*., 2011), facilitating its colonization even of remote habitats directly exposed to atmospheric deposition (Hervas and Casamayor, 2009). Archaea comprised only a minor part of the community; this agrees with a recent metagenomic study in an alpine glacial system (Edwards *et al*., 2013) and with an earlier study using fluorescent *in situ* hybridization on biofilms in a glacier-fed stream (Battin *et al*., 2001).

We found differences in community composition between epilithic and epipsammic biofilms, which were particularly pronounced when considering the presence and absence of OTUs. This difference may be attributable to the physical properties of the sediment. For instance, sandy material is less stable and has higher turnover than larger grains including stones, causing more frequent abrasion of attached communities (Wimpenny *et al*., 2000) and consequently steady renewal of the community. Furthermore, the larger specific surface area of sandy sediment may foster the adhesion of microbial cells (even dormant) suspended in the streamwater, thereby generating a substantial 'seed bank' of dormant taxa. This notion is in fact supported by the elevated diversity in the bulk epipsammic community.

The environmental variables that we measured explained only a relatively small fraction of the variation in community composition. However, we identified streamwater pH as a major variable affecting both epilithic and epipsammic bulk community structure in the glacier-fed streams. This agrees with reports from soils where pH predicts microbial community composition (Fierer and Jackson, 2006). Streamwater DOC concentration was extremely low across all study sites and yet it was related to community composition of the putatively active microorganisms in epipsammic biofilms. Epipsammic biofilm communities typically comprise less algae than epilithic biofilms (Romani and Sabater, 2000), which is likely attributable to differences in sediment grain size. Our finding on a lower relative abundance of chloroplasts/*Cyanobacteria* in epipsammic than in epilithic biofilms supports this notion. The relation of DOC concentration with variation in community composition suggests that microbial heterotrophs in epipsammic biofilms rely more on external sources to satisfy their carbon demand, while microbial heterotrophs in epilithic biofilms may benefit from algal exudates. While we have not measured algal biomass in this study, this assumption is supported by observations that biofilms with low ratios of chlorophyll *a* to microbial cell counts indeed utilize more DOC from the streamwater (Battin *et al*., 2003). Moreover, epipsammic biofilms were shown to have the ability to degrade more organic carbon than epilithic biofilms, which may simply be due to enhanced accumulation of organic carbon in a sandy matrix (Romani and Sabater, 2001). The unexplained variance suggests that further factors, likely including unmeasured environmental variables but also dispersal dynamics, may influence community composition.

Concurrent sequencing of both 16S rRNA and 16S rDNA is now common to evaluate taxa that apparently differ in their metabolic activity (Jones and Lennon, 2010; Campbell *et al*., 2011; Besemer *et al*., 2012; Campbell and Kirchman, 2013). It is conducive to unravel various components of complex microbial communities and mechanisms underlying the maintenance of microbial biodiversity. However, the dual rRNA and rDNA approach also has its limitations, and results should be interpreted with caution (Blazewicz *et al*., 2013). Moreover, 454-pyrosequencing yields only relative abundances of microbial taxa, therefore precluding any conclusions on the absolute activity of the studied systems, and DNA extraction, cDNA synthesis and polymerase chain reaction (PCR) amplification are all sources of potential biases (Lanzén *et al*., 2011). The relation between bulk and active taxa at phylum level indicates that our methods yielded reproducible results at this phylogenetic level at least.

In conclusion, our study showed that many rare taxa were active in glacier-fed stream biofilms, which is in line with studies on lakes and marine microbial communities (Jones and Lennon, 2010; Campbell *et al*., 2011; Campbell and Kirchman, 2013). We suggest that high community turnover, where taxa repeatedly enter and leave the 'seed bank', is particularly relevant for the maintenance of microbial biodiversity in glacier-fed streams where environmental fluctuations are pronounced.

Experimental procedures

Study sites and sample collection

The study sites covered 26 glacier-fed streams located along the main chain of the Austrian Alps (Silvretta, Ötztaler Alps, Stubaier Alps, Zillertaler Alps, Venediger group, Granatspitz group, Glockner group, Hochalmspitze and Ankogel; geographically ranging from 10.162°E to 13.278°E and from 46.778°N to 47.135°N). We sampled within 45 days in July and August 2010 during a period where glaciers were largely snow-free. Streamwater pH, electrical conductivity and temperature were measured in the field using two different probes (pH 320, Cond340i, WTW, Weilheim, Germany). We collected epilithic and epipsammic biofilm samples from the streambed surface of the glacier-fed streams within 10 m downstream of the glacier terminus. Epilithic biofilms (visible by eye) were sampled by rigorously shaking approximately 80 stones (diameter ~2 cm) successively in 50 ml tubes using sterile-filtered streamwater. Removed cells in the supernatant were filtered onto sterile GSWP filters (0.22 μm mixed cellulose-ester filters; Millipore, Billerica, MA, USA). Epipsammic biofilms were collected with sandy sediment into 2 ml Cryo tubes with a sterile spatula. To avoid RNA degradation, sampling time was kept as short as possible (maximum 20 min), and all samples were immediately frozen in liquid nitrogen still in the field.

DOC analysis and biogeochemistry

Streamwater was analysed for DOC concentration using a Sievers 900 TOC Analyser (GE Analytical Instruments, Boulder, CO, USA) operated with an inorganic carbon removal unit as described previously (Wilhelm *et al*., 2013). Concentrations of streamwater $N\text{-}NH_4$, $N\text{-}NO_2$ and $N\text{-}NO_3$ were determined using Continuous Flow Analysis (Alliance Instruments, Salzburg, Austria); all three fractions were summed to dissolved inorganic nitrogen for further analyses.

RNA and DNA extractions, cDNA synthesis, and PCR amplification

RNA was extracted using a modification of the method described previously (Urich *et al*., 2008). Briefly, 0.5 ml of both cetyltrimethyl ammonium bromide buffer and phenol : chloroform : isoamylalcohol (25:24:1, pH 6.8) were added to a lysing matrix E tube (MP Biomedicals, Santa Ana, CA, USA) containing 0.4 g of epipsammic biofilms or one filter with the concentrated epilithic biofilm sample. The cells were lysed in a FastPrep machine (MP Biomedicals) at speed 5.0 for 45 s, followed by nucleic acid precipitation with PEG 8000. Total nucleic acids were subjected to DNase treatment using the RQ1 RNase-free DNase (Promega, Madison, WI, USA) and purified with the MEGA clear purification kit (Life Technologies, Carlsbad, CA, USA). The total RNA was used as a template for random hexamer-primed reverse transcription using Revert Aid Reverse Transcriptase (Thermo Scientific, Waltham, MA, USA), as indicated in the manufacturer's protocol. For higher yields, 0.2 μl of the T4gene32protein (New England Biolabs, Ipswich, MA, USA) were added to each reaction. DNA was extracted using the PowerSoil DNA extraction kit (MoBio, Carlsbad, CA, USA) according to the manufacturer's recommendations. As glacier-fed streams are highly oligotrophic with low microbial biomass, the use of two different methods for RNA and DNA extraction was necessary to obtain reasonable nucleic acid concentrations (> 2 ng μl^{-1}).

The V4-V5 region of the 16S rRNA and 16S rDNA was amplified in using the universal primers 515F 5′-GTGNC AGCMGCCGCGGTAA-3′ and 926R 5′-CCGYCAATTYMT TTRAGTTT-3′, containing the 454 Titanium A and B adaptors respectively (Quince *et al*., 2011). To reduce barcode-specific bias (Berry *et al*., 2011), we employed forward primers with two different sample-specific barcodes for each sample. Each 50 μl PCR reaction contained 0.5 μmol l^{-1} of each primer (Thermo Scientific), 0.2 mmol l^{-1} deoxyribonucleotide triphosphates (Thermo Scientific), 40 μg bovine serum albumin (Thermo Scientific), 2.5–4 mmol l^{-1} $MgCl_2$, 1 U Taq-DNA Polymerase (Thermo Scientific) or 2 U Pure Taq DNA Polymerase (PeqLab) with the recommended PCR buffer and 4 μl DNA or RNA extract. We used the Pure Taq DNA Polymerase and the Taq-DNA Polymerase for amplification of the 16S rRNA the 16S rDNA respectively. Samples were amplified using an initial denaturing step of 2 min at 94°C, followed by 30–35 cycles of 30 s denaturation at 94°C, 30 s annealing at 56°C, 1 min elongation at 72°C and a final elongation for 10 min at 72°C. Each PCR included a negative control. PCR products were run on a 1% agarose gel and purified using the Gel Extraction Kit (Qiagen, Hilden, Germany). The purified PCR products were quantified by gel using the Gel Doc XR + System (Bio-Rad, Hercules, CA, USA) and pooled equimolar for pyrosequencing.

Amplicons were sequenced on a GS FLX Titanium Sequencer in Liverpool (Centre for Genomic Research, University of Liverpool, Liverpool, UK).

Data processing

The full 454-pyrosequencing data set is available at the National Center for Biotechnology Information Sequence Read Archive under the accession number SRP016506. Raw output files were filtered and de-noised using the software package AmpliconNoiseV1.0 (Quince *et al*., 2011). First, we preclustered the sequences with PyroNoise (AmpliconNoiseV1.0), then corrected PCR single base errors using SeqNoise (AmpliconNoiseV1.0), and finally identified and removed chimeras with the Perseus algorithm at an intercept of $\alpha = -7.5$ and a coefficient of $\beta = 0.5$ (Quince *et al*., 2011). A complete linkage algorithm on a 97% identity level clustered the cleaned reads to OTUs (AmpliconNoiseV1.0), and taxonomic assignments were determined by using a naïve Bayesian rDNA classifier (Ribosomal Database Project; Wang *et al*., 2007) at an 80% confidence threshold. Given that the number of reads was too low in two samples of the rRNA-based communities, we excluded them from all analyses. Furthermore, we removed one OTU from the whole data set, which was assigned to the genus *Escherichia/Shigella* and was found in the RNA only. Although negative controls were included in each PCR, the low nucleic acid concentrations of our samples and potential contaminating bacterial DNA originating from the reverse transcriptase enzyme preparation (Spangler *et al*., 2009) may have accounted for this specific OTU.

Rarefaction curves for the 454-pyrosequencing data were computed using the AmpliconNoise software package V1.0 (Quince *et al*., 2011). Alpha diversity, estimated as richness and as the Shannon diversity (i.e. the number equivalents of the Shannon entropy; Jost, 2006), was calculated after rarefying communities to the lowest number of reads obtained for a sample (i.e. 1074) to correct for sample size. The 'true diversity' of the communities was estimated by Bayesian fitting of the OTU's abundance distributions to a Sichel distribution using the Diversity Estimation software according to Quince and colleagues (2008). To compare our results on microbial alpha diversity to other stream ecosystems, we rarefied the data of a recent study (Besemer *et al*., 2013), which used the same primers and followed the same bioinformatics pipeline, to 1074 reads and estimated richness and Shannon diversity.

To visualize patterns in community composition, we applied the presence-/absence-based Sørensen and the relative abundance-based Morisita–Horn dissimilarity index – the latter was chosen because of its relative robustness against undersampling (Chao *et al*., 2006) – and performed nMDS analyses. For the calculation of dissimilarity matrices, communities were randomly resampled to the sample with the lowest number of reads (1074 reads). We used a repetitive resampling procedure (drawing 1000 random samples from each community) to avoid the stochastic bias, which is inherent if only one random sample is drawn from a community (Besemer *et al*., 2013). Finally, we calculated the mean of 1000 dissimilarity values for each cell in the dissimilarity matrix.

A two-way non-parametric PERMANOVA (Anderson, 2001) was performed to test for differences in the location of communities. The two factors included were 'epilithic-epipsammic' and 'DNA-RNA'. As the analysis of the Morisita–Horn index revealed a significant interaction between these two factors, we used a post-hoc test to pairwise compare the groups. *P*-values were Bonferroni corrected to account for multiple comparisons. A concomitant test for dispersion homogeneity (the multivariate analogue of variance homogeneity) among all groups was done by computing the average distances of the communities to the respective group centroids in the ordination space, which were then compared using a permutation test (Anderson, 2006).

We applied CCA to determine which environmental variables explained patterns of community composition among sites best. CCA is an ordination technique that seeks the most prominent linear gradients in multivariate data sets under the constraint that the gradients are linear combinations of a set of explanatory variables and assuming a unimodal species–environment relationship (Ter Braak, 1987). To assess the significance of the constraints, we performed a permutation-based ANOVA-like test (Oksanen *et al*., 2013). In a next step, forward variable selection was used to generate the most efficient model from a set of potential explanatory variables (Oksanen *et al*., 2013).

The relative abundances of OTUs in rRNA and rDNA of each glacier-fed stream were compared to identify whether rare taxa were potentially active. We used a 20% threshold for defining the abundant group of OTUs for each sample, representing 20% of the most abundant OTUs respectively. Additionally we applied a 30%, 40% and 50% threshold to illustrate the robustness of this analysis. To account for sample size, we randomly resampled the larger sample of a rRNA/rDNA pair. Next, we calculated ratios between the relative abundances of each OTU in rRNA and rDNA, and counted the OTUs with ratios above 1 (scored as active) and below 1 (scored as inactive) for the abundant and rare group respectively (Jones and Lennon, 2010). This procedure was repeated 1000 times, and the results were averaged. The ratios of active to inactive OTUs were compared using a Wilcoxon test.

All analyses were performed in R using the packages *vegan* (Oksanen *et al*., 2013) and *MASS* (Venables and Ripley, 2002), and graphs were prepared in R (R Development Core Team, 2011) and Sigma Plot.

Acknowledgements

We are grateful to C. Preiler, L. Hartmann, B. Eichelberger, B. Preiler, I. Hödl and L. Nicklas for assistance in the field. We thank P. Offre for fruitful discussions regarding sample preparation. K. Wagner assisted in the laboratory and C. Kroisleitner provided glacier coverage data. Financial support came from the Austrian Science Fund (START Y420-B17) to T.J.B.

References

Anderson, M.J. (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26:** 32–46.

Anderson, M.J. (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **62:** 245–253.

Anesio, A.M., and Laybourn-Parry, J. (2012) Glaciers and ice sheets as a biome. *Trends Ecol Evol* **27:** 219–225.

Barnett, T.P., Adam, J.C., and Lettenmaier, D.P. (2005) Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature* **438:** 303–309.

Battin, T.J., Wille, A., Sattler, B., and Psenner, R. (2001) Phylogenetic and functional heterogeneity of sediment biofilms along environmental gradients in a glacial stream. *Appl Environ Microbiol* **67:** 799–807.

Battin, T.J., Kaplan, L.A., Newbold, J.D., Cheng, X., and Hansen, C. (2003) Effects of current velocity on the nascent architecture of stream microbial biofilms. *Appl Environ Microbiol* **69:** 5443–5452.

Battin, T.J., Wille, A., Psenner, R., and Richter, A. (2004) Large-scale environmental controls on microbial biofilms in high-alpine streams. *Biogeosciences* **1:** 159–171.

Berry, D., Ben Mahfoudh, K., Wagner, M., and Loy, A. (2011) Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Appl Environ Microbiol* **77:** 7846–7849.

Besemer, K., Peter, H., Logue, J.B., Langenheder, S., Lindström, E.S., Tranvik, L.J., *et al*. (2012) Unraveling assembly of stream biofilm communities. *ISME J* **6:** 1459–1468.

Besemer, K., Singer, G., Quince, C., Bertuzzo, E., Sloan, W., and Battin, T.J. (2013) Headwaters are critical reservoirs of microbial diversity for fluvial networks. *Proc R Soc B* **280:** 20131760. doi:10.1098/rspb.2013.1760.

Blazewicz, S.J., Barnard, R.L., Daly, R.A., and Firestone, M.K. (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J* **7:** 2061–2068. doi:10.1038.

Campbell, B.J., and Kirchman, D.L. (2013) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *ISME J* **7:** 210–220.

Campbell, B.J., Yu, L., Heidelberg, J.F., and Kirchman, D.L. (2011) Activity of abundant and rare bacteria in a coastal ocean. *Proc Natl Acad Sci U S A* **108:** 12776–12781.

Chao, A., Chazdon, R.L., Colwell, R.K., and Shen, T.J. (2006) Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* **62:** 361–371.

Darcy, J.L., Lynch, R.C., King, A.J., Robeson, M.S., and Schmidt, S.K. (2011) Global distribution of *Polaromonas* phylotypes-evidence for a highly successful dispersal capacity. *PLoS ONE* **6:** e23742.

DeFlaun, M.F., and Mayer, L.M. (1983) Relationships between bacteria and grain surfaces in intertidal sediments. *Limnol Oceanogr* **28:** 873–881.

Edwards, A., Pachebat, J.A., Swain, M., Hegarty, M., Hodson, A.J., Irvine-Fynn, T., *et al*. (2013) A metagenomic snapshot of taxonomic and functional diversity in an alpine glacier cryoconite ecosystem. *Environ Res Lett* **8:** 035003.

Fegatella, F., Lim, J., Kjelleberg, S., and Ricardo, C. (1998) Implications of rRNA Operon copy number and ribosome content in the marine oligotrophic ultramicrobacterium *Sphingomonas* sp strain RB2256. *Appl Environ Microbiol* **64:** 4433–4438.

Fierer, N., and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* **103:** 626–631.

Freimann, R., Bürgmann, H., Findly, S.E.G., and Robinson, C.T. (2013) Response of lotic microbial communities to altered water source and nutritional state in a glacated alpine floodplain. *Limnol Oceanogr* **58:** 951–965.

Hannah, D.M., Brown, L.E., Milner, A.M., Gurnell, A.M., McGregor, G.R., Petts, G.E., *et al*. (2007) Integrating climate – hydrology – ecology for alpine river systems. *Aquat Conserv* **17:** 636–656.

Hervas, A., and Casamayor, E.O. (2009) High similarity between bacterioneuston and airborne bacterial community compositions in a high mountain lake area. *FEMS Microbiol Ecol* **67:** 219–228.

Hodson, A., Anesio, A.M., Tranter, M., Fountain, A.G., Osborn, M., Priscu, J.C., *et al*. (2008) Glacial ecosystems. *Ecol Monogr* **78:** 41–67.

Hood, E., Fellman, J., Spencer, R.G.M., Hernes, P.J., Edwards, R., D'Amore, D., *et al*. (2009) Glaciers as a source of ancient and labile organic matter to the marine environment. *Nature* **462:** 1044–1047.

Jones, S.E., and Lennon, J.T. (2010) Dormancy contributes to the maintenance of microbial diversity. *Proc Natl Acad Sci U S A* **107:** 5881–5886.

Jost, L. (2006) Entropy and diversity. *Oikos* **113:** 363–375.

Kembel, S.W., Wu, M., Eisen, J.A., and Green, J.L. (2012) Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLoS Comput Biol* **8:** e1002743.

Kirchman, D.L., Cottrell, M.T., and Lovejoy, C. (2010) The structure of bacterial communities in the western Arctic Ocean as revealed by pyrosequencing of 16S rRNA genes. *Environ Microbiol* **12:** 1132–1143.

Lanoil, B., Skidmore, M., Priscu, J.C., Han, S., Foo, W., Vogel, S.W., *et al*. (2009) Bacteria beneath the West Antarctic ice sheet. *Environ Microbiol* **11:** 609–615.

Lanzén, A., Jørgensen, S.L., Bengtsson, M.M., Jonassen, I., Øvreås, L., and Urich, T. (2011) Exploring the composition and diversity of microbial communities at the Jan Mayen hydrothermal vent field using RNA and DNA. *FEMS Microbiol Ecol* **77:** 577–589.

Lennon, J.T., and Jones, S.E. (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* **9:** 119–130.

Lock, M.A. (1993) Attached microbial communities in rivers. In *Aquatic Microbiology – An Ecological Approach*. Ford, T.E. (ed.). Oxford, UK: Blackwell Scientific Publications, pp. 113–138.

Margesin, R., Spröer, C., Zhang, D., and Busse, H. (2012) *Polaromonas glacialis* sp. nov. and *Polaromonas cryoconiti* sp. nov., isolated from alpine glacier cryoconite. *Int J Syst Evol Microbiol* **62:** 2662–2668.

Milner, A.M., Brittain, J.E., Brown, L.E., and Hannah, D.M. (2010) Water sources and habitat of alpine streams. In *Alpine Waters*. Bundi, U. (ed.). Berlin Heidelberg, Germany: Springer, pp. 175–191.

Moeseneder, M.M., Arrieta, J.M., and Herndl, G.J. (2005) A comparison of DNA- and RNA-based clone libraries from the same marine bacterioplankton community. *FEMS Microbiol Ecol* **51:** 341–352.

Murray, J.L.S., and Jumars, P.A. (2002) Clonal fitness of attached bacteria predicted by analog modeling. *Bioscience* **52:** 343–355.

Oksanen, A.J., Blanchet, F.G., Kindt, R., Minchin, P.R., Hara, R.B.O., Simpson, G.L., *et al*. (2013) vegan: Community Ecology Package. R package version 2.0-5.

Pedrós-Alió, C. (2006) Marine microbial diversity: can it be determined? *Trends Microbiol* **14:** 257–263.

Quince, C., Curtis, T.P., and Sloan, W.T. (2008) The rational exploration of microbial diversity. *ISME J* **2:** 997–1006.

Quince, C., Lanzen, A., Davenport, R.J., and Turnbaugh, P.J. (2011) Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12:** 1–18.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Romani, A.M., and Sabater, S. (2000) Influence of algal biomass on extracellular enzyme activity in river biofilms. *Microb Ecol* **40:** 16–24.

Romani, A.M., and Sabater, S. (2001) Structure and activity of rock and sand biofilms in a Mediterranean stream. *Ecology* **82:** 3232–3245.

Simon, C., Wiezer, A., Strittmatter, A.W., and Daniel, R. (2009) Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. *Appl Environ Microbiol* **75:** 7519–7526.

Singer, G.A., Fasching, C., Wilhelm, L., Niggemann, J., Steier, P., Dittmar, T., *et al*. (2012) Biogeochemically diverse organic matter in Alpine glaciers and its downstream fate. *Nat Geosci* **5:** 710–714.

Skidmore, M., Anderson, S.P., Sharp, M., Foght, J., and Lanoil, B.D. (2005) Comparison of microbial community compositions of two subglacial environments reveals a possible role for microbes in chemical weathering processes. *Appl Environ Microbiol* **71:** 6986–6997.

Spangler, R., Goddard, N.L., and Thaler, D.S. (2009) Optimizing Taq polymerase concentration for improved signal-to-noise in the broad range detection of low abundance bacteria. *PLoS ONE* **4:** e7010.

Stibal, M., Šabacká, M., and Žárský, J. (2012) Biological processes on glacier and ice sheet surfaces. *Nat Geosci* **5:** 771–774.

Telling, J., Anesio, A.M., Tranter, M., Irvine-Fynn, T., Hodson, A., Butler, C., *et al*. (2011) Nitrogen fixation on Arctic glaciers, Svalbard. *J Geophys Res* **116:** G03039.

Ter Braak, C.J.F. (1987) Ordination. In *Data Analysis in Community and Landscape Ecology*. Jongman, R.H.G., Ter Braak, C.J.F., and Van Tongeren, O.F.R. (eds). New York, USA: Cambridge University Press, pp. 91–173.

Urich, T., Lanzén, A., Qi, J., Huson, D.H., Schleper, C., and Schuster, S.C. (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS ONE* **3:** e2527.

Větrovský, T., and Baldrian, P. (2013) The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS ONE* **8:** e57923.

Venables, W.N., and Ripley, B.D. (2002) *Modern Applied Statistics with S*, Fourth edn. New York, USA: Springer.

Wadham, J.L., Arndt, S., Tulaczyk, S., Stibal, M., Tranter, M., Telling, J., *et al*. (2012) Potential methane reservoirs beneath Antarctica. *Nature* **488:** 633–637.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73:** 5261–5267.

Wilhelm, L., Singer, G.A., Fasching, C., Battin, T.J., and Besemer, K. (2013) Microbial biodiversity in glacier-fed streams. *ISME J* **7:** 1651–1660. doi:10.103.

Wimpenny, J., Manz, W., and Szewzyk, U. (2000) Heterogeneity in biofilms. *FEMS Microbiol Rev* **24:** 661–671.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher' web-site:

Fig. S1. Rarefaction curves of epilithic and epipsammic communities. Rarefaction curves were computed by using the AmpliconNoise software package V1.0 at a 97% sequence similarity level (Quince *et al*., 2011). For most samples, the rarefaction curves did not reach an asymptote, indicating a significant amount of undetected diversity, especially for bulk epipsammic biofilms.

Fig. S2. Alpha diversity of epilithic and epipsammic communities. Richness and Shannon diversity were estimated after rarefying communities to the lowest number of reads obtained for a sample (i.e. 1074). The 'true diversity' was estimated by fitting Sichel distribution curves to the abundance distributions obtained from the 454-pyrosequencing data. Significant differences (Wilcoxon test, $n = 24$) between groups are indicated by asterisks: $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. RNA LITH, active epilithic biofilm; DNA LITH, bulk epilithic biofilm; RNA PSAM, active epipsammic biofilm; DNA PSAM, bulk epipsammic biofilm.

Fig. S3. Canonical correspondence analysis of active and bulk biofilm communities. (A) Bulk epilithic biofilm, (B) active epilithic biofilm, (C) bulk epipsammic biofilm, (D) active epipsammic biofilm. Streamwater pH and conductivity explained a considerable amount of variation in the bulk biofilm communities, while DOC concentration was related to the composition of active epipsammic biofilms. Streamwater temperature appeared less important for all ordinations.

Table S1. Characteristics of the study glaciers in the Austrian Alps. Longitude and latitude of each glacier were derived from the World Glacier Inventory (http://nsidc.org/data/glacier_inventory/browse.html). Glacier cover (%) was calculated from total drainage area (Austrian Map 2001) and glacier area (World Glacier Inventory). Elevation (metres above sea level) refers to the glacier terminus.

Table S2. Environmental variables of the glacier-fed streams. pH, electrical conductivity and streamwater temperature were measured in the field using probes (pH 320, Cond340i). DOC concentrations were measured with a Sievers 900 TOC Analyser. $N\text{-}NO_2$, $N\text{-}NO_3$, $N\text{-}NH_4$ (summed up to dissolved inorganic nitrogen – DIN) and $P\text{-}PO_4$ were measured using a Continuous Flow Analysis (Alliance instruments).

Table S3. Rare but active taxa in glacier-fed streams. In line with the use of a 20% threshold for assigning OTUs to an abundant group, we also stepwise increased this threshold to 50% (20%, 30%, 40% and 50%) to illustrate the robustness of this analysis. We calculated ratios of rRNA : rDNA and then tested whether the abundant and rare groups differed from each other using a Wilcoxon test. In a few samples, there was no single OTU in the 'rare rDNA' group, but all were assigned to the 'rare rRNA' group. As this fact makes the calculation of a ratio of rRNA : rDNA impossible, we excluded those specific samples from the analysis (notice lower *n*).