

Bacterial diversity in a glacier foreland of the high Arctic

URSEL M.E. SCHÜTTE*,§ ZAID ABDO,†,‡,§ JAMES FOSTER*,§ JACQUES RAVEL,¶ JOHN BUNGE,** BJØRN SOLHEIM†† and LARRY J. FORNEY*,§

Departments of *Biological Sciences, †Mathematics and ‡Statistics, §Initiative for Bioinformatics and Evolutionary Studies (IBEST), University of Idaho, Moscow, ID 83844, USA, ¶School of Medicine, University of Maryland, Baltimore, MD 21201, USA, **Department of Statistical Science, Cornell University, Ithaca, NY 14853, USA, ††Department of Biology, University of Tromsø, 9037 Tromsø, Norway

Abstract

Over the past 100 years, Arctic temperatures have increased at almost twice the global average rate. One consequence is the acceleration of glacier retreat, exposing new habitats that are colonized by microorganisms whose diversity and function are unknown. Here, we characterized bacterial diversity along two approximately parallel chronosequences in an Arctic glacier forefield that span six time points following glacier retreat. We assessed changes in phylotype richness, evenness and turnover rate through the analysis of 16S rRNA gene sequences recovered from 52 samples taken from surface layers along the chronosequences. An average of 4500 sequences was obtained from each sample by 454 pyrosequencing. Using parametric methods, it was estimated that bacterial phylotype richness was high, and that it increased significantly from an average of 4000 (at a threshold of 97% sequence similarity) at locations exposed for 5 years to an average of 7050 phylotypes per 0.5 g of soil at sites that had been exposed for 150 years. Phylotype evenness also increased over time, with an evenness of 0.74 for 150 years since glacier retreat reflecting large proportions of rare phylotypes. The bacterial species turnover rate was especially high between sites exposed for 5 and 19 years. The level of bacterial diversity present in this High Arctic glacier foreland was comparable with that found in temperate and tropical soils, raising the question whether global patterns of bacterial species diversity parallel that of plants and animals, which have been found to form a latitudinal gradient and be lower in polar regions compared with the tropics.

Keywords: bacteria, community, diversity, glacier, High Arctic

Received 29 May 2009; revision received 2 August 2009; accepted 8 August 2009

Introduction

The Arctic is disproportionately affected by climate change and arctic temperatures have increased at almost twice the global average rate over the past 100 years (Bernstein *et al.* 2007). The acceleration of glacier retreat is one consequence of these temperature increases and the net mass balance of most Arctic glaciers has been negative over the past few decades (Dowdeswell *et al.* 1997). Predictions are that the rate of loss will increase in the future (Kintisch 2009). The newly exposed rocks and minerals represent new habitats for microorganisms that

have key roles in soil development, nutrient cycling and plant growth. Bacterial communities may be key determinants of Arctic ecosystem stability and function because of their important roles in soil development, biogeochemical cycles and heterotrophic activities. As fundamental knowledge on bacterial communities in High Arctic glacier forelands is lacking, studies on bacterial community diversity and composition in these extreme environments are needed. The design of these studies should take into account the extraordinary spatial heterogeneity of these habitats and patchy distribution of bacteria in soils (Green *et al.* 2004; Horner-Devine *et al.* 2004; Green & Bohannan 2006).

Valuable information has been gained from studies carried out to assess bacterial diversity and succession

Correspondence: Dr Larry J. Forney, Fax: 1-208-885-7905; E-mail: lforney@uidaho.edu

in montane glacier forelands (Schipper *et al.* 2001, Sigler & Zeyer 2002; Sigler *et al.* 2002, Nemergut *et al.* 2007; Schmidt Reed *et al.* 2008). These studies that were based on the analysis of 16S rRNA gene sequences provide conflicting information pertaining to changes in phylotype diversity over time following glacier retreat. Nemergut *et al.* (2007) found that it increased while others found the opposite (Sigler & Zeyer 2002; Sigler *et al.* 2002). These findings suggest that the development of bacterial communities in soils following glacier retreat may be less predictable than in plant communities, where species richness consistently increases over time (Huston & Smith 1987; Matthews 1992; Hodkinson *et al.* 2003; Walker & del Moral 2003). Although the methods used were appropriate for the aims of these studies, the sampling schemes, numbers of samples and the use of low resolution analytical methods (Bent & Forney 2008) precluded statistically meaningful estimates of bacterial species richness (α -diversity), evenness or turnover rate (β -diversity). Furthermore, the extrapolation of findings from studies of montane glacier forelands to those of the High Arctic may be problematic because the dynamics of alpine and High Arctic glaciers differ. High Arctic glaciers are commonly polythermal and they gather a large reservoir of meltwater beneath the glacial mass during winter, which is then released abruptly as a surging breakout in spring. These thermal and hydrological differences result in comparatively unstable forelands in the High Arctic (Harland 1997; Hodkinson *et al.* 2003). Thus, fundamental differences might exist between the montane and High Arctic glacier forelands.

The few studies carried out provide only limited insight to microbial diversity in polar soils as only a small number of samples were examined. Nonetheless, the findings suggest that bacterial diversity in polar habitats is high despite the extreme environmental conditions that exist (Broady 1996; Vishniac 1996, Shi *et al.* 1997; Zhou *et al.* 1997; Lawley *et al.* 2004, Shivaji *et al.* 2004; Neufeld & Mohn 2005; Steven *et al.* 2007; González-Toril *et al.* 2008), and a large proportion of the bacterial taxa are unlike those previously described. An analysis of Siberian tundra soil showed that mainly α -, δ -Proteobacteria were present, but 7% of the 16S rRNA gene sequences retrieved from the samples were more than 20% different from those previously reported (Zhou *et al.* 1997). The findings were in agreement with those of Shivaji *et al.* (2004) who found members of seven different phyla including α -, β -, γ -Proteobacteria and *Actinobacteria*, and culturable bacteria that belonged to genera previously not reported to be in Antarctic soils such as *Sphingomonas* and *Acidovorax*. Neufeld & Mohn (2005) did probably the most extensive study on bacterial diversity in this environment. The authors

sampled geographically distant boreal and arctic soils, and diversity was studied using ~13 000 serial analysis of ribosomal sequence tags and denaturing gradient gel electrophoresis of 16S rRNA genes. They reported that diversity in arctic soils was higher than in boreal soils with the highest diversity found at 82 °N, an extreme northern location. Until recently, investigations of bacterial diversity based on cultivation-independent methods were limited by the lack of high throughput methods for sequencing 16S rRNA genes. The advent of 'next generation' DNA sequencing technologies such as pyrosequencing has obviated this problem and provided the means necessary to conduct intensive and extensive analyses of bacterial diversity in different environments (Sogin *et al.* 2006; Roesch *et al.* 2007; Dethlefsen *et al.* 2008; Huse *et al.* 2008).

We recently showed that bacterial succession occurred in the glacier foreland of Midtre Lovén glacier (Schütte *et al.* 2009), and here we report the results of studies carried out to assess species (phylotype) richness, evenness and turnover rate along two chronosequences in this foreland. These approximately parallel chronosequences span six times of exposure following retreat of the Midtre Lovén glacier. The bacterial diversity in surface soils was assessed by pyrosequencing of bar-coded 16S rRNA gene amplicons present in samples taken from surface soils. We did not expect high diversity in glacier foreland soils because environmental conditions are harsh including UV radiation, low water content and low nutrient contents. Furthermore, increase in plant richness and higher soil organic matter and higher water content (Hodkinson *et al.* 2003) most probably results in an increase in the number of spatial niches in the soil thus, leading to the expectation that bacterial phylotype richness increases over time in the foreland of Midtre Lovén glacier. Using different statistical methods to estimate species richness, evenness and the species turnover rate, we were able to show a remarkably high bacterial diversity at all times since glacier retreat with an increase over time.

Materials and methods

Study site and sampling

The foreland of Midre Lovén glacier near the settlement Ny-Ålesund, West Spitsbergen (74–81 °N; 10–35 °E) was chosen as a field site to assess bacterial diversity. Samples were taken in the beginning of August 2003 along the chronosequence established by Hodkinson *et al.* (2003) and along a second approximately parallel chronosequence located about 25 m away. The second chronosequence was designed to account for differences in bacterial community structure caused by horizontal var-

iation and the two chronosequences were assumed to be independent of each other (Schütte *et al.* 2009). Both chronosequences included six sites with a total distance of about 1 km and the distance between sites was approximately 100–200 m. The total time period covered was 150 years with site 1–6 representing times since glacier retreat of 5, 19, 40, 63, 100 and 150 years respectively (Schütte *et al.* 2009). Hodkinson *et al.* (2003) had previously determined the time since deglaciation of each sampling location using photographs and radiocarbon analysis.

At each sampling location, a transect 5 m long was established along which five samples were taken 1 m apart. Each sample had a surface area of 10 × 10 cm². Larger gravel (> ~3–5 cm) was manually removed before the samples were taken. In this study, only the samples taken from the surface were used, which contained a mixture of vegetation, rhizosphere and bulk soil and did not vary in thickness among sampling sites (about 1 cm). In total, 60 samples were collected, placed in plastic bags and kept frozen on ice during transport to the laboratory and stored at –80 °C until analysed.

454 Pyrosequencing

Genomic DNA was isolated from 0.5 g of the soil samples, which had been mixed in their bag, using a modified procedure based on the UltraClean™ Fecal DNA Kit (MoBio). For each sample, soil was weighed into a sterile 2 mL tube containing 0.5 g of silicate glass beads (Glenn Mills) and 750 µL of TE (1 mM Tris and 50 mM EDTA) was added. The cells were treated with 50 µL lysozyme (10 mg/mL; Sigma-Aldrich) and 25 µL of mutanolysin (2 mg/mL; Sigma-Aldrich). Afterwards, 60 µL of the solution S1 and 200 µL of the IRS solution (MoBio) were added followed by a bead beating step for 3 min at full speed (Biospec Products Inc). The tubes were centrifuged for 4 min at 13 000 g and 450 µL of the supernatants was transferred to a new 2 mL tube before following the protocol from MoBio. The columns were incubated for 5 min at room temperature before eluting the DNA using 50 µL deionized water with pH 7, heated to 90 °C. As positive controls, we used cells from pure cultures of *Escherichia coli* K12 MG1655 and *Lactobacillus aviarius* ATCC 43232. The DNA from all samples was further cleaned using the Wizard PCR purification kit (Promega).

We amplified the bacterial 16S rRNA gene and the primers used flanked the variable regions 1 and 2 (*Escherichia coli* positions 27F–338R). The primer sequences were as follows: 454_27F 5'-GCCTTGCCAGCCCCGCTC-AGTCAGAGTTTGATCCTGGCTCAG and 454_338R 5'-GCCTCCCTCGGCCATCAGNNNNNNNNNCATGCTGCCCTCCCGTAGGAGT; where the underlined

sequences are 454 Life Sciences® adapters B and A in 27F and 338R respectively, and the bold font denotes the universal 16S rRNA primers 27F and 338R. The 338R primer included a unique sequence tag to barcode each of the samples denoted by the eight N. This allowed us to sequence the amplicons from all samples simultaneously, and afterwards assign each sequence to the sample they were obtained from. Using a specific barcode for each reverse primer resulted in each PCR reaction containing a different reverse primer, thus, a negative control was run for each reaction. Each PCR reaction contained 35.2 µL of PCR-grade water, 5.0 µL of 10× buffer (Applied Biosystems), 6.0 µL of 25 mM MgCl₂ (Applied Biosystems), 1.2 µL of 20 mg/mL BSA (Roche Applied Science), 0.4 µL of 25 mM dNTP (Amersham Bioscience), 0.5 µL of 20 µM forward primer 454_27f, 1 µL of 10 µM reverse primer 454_338r, 0.2 µL of 5 U/µL Taq DNA polymerase (Applied Biosystems) and 1.0 µL of DNA template, in a total volume of 50 µL. Amplification of 16S rRNA genes was performed using an initial denaturation step at 94 °C for 4 min, followed by 34 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min and an extension at 72 °C for 2 min. The final extension was 10 min step at 72 °C. The concentrations of amplicons were estimated using a GelDoc quantification system (BioRad), and approximately equal amounts (~100 ng) of all amplicons were mixed in a single tube. Amplification primers and reaction buffer were removed by processing the amplicon mixture with the AMPure Kit (Agencourt). Emulsion PCRs were performed as described in Margulies *et al.* (2005). Sequences were obtained using a Roche 454 GS-FLX (Roche-454 Life Sciences).

Quality control of sequencing

We followed the criteria previously described to assess the quality of sequence reads (Sogin *et al.* 2006; McKenna *et al.* 2008). To pass, a sequence read had to (i) include a perfect match to the sequence tag (bar-code) and the 16S rRNA gene primer; (ii) be at least 100 bp in length; (iii) have no more than four undetermined bases; and (iv) have at least 60% match to a previously determined 16S rRNA gene sequence after alignment with NAST (DeSantis *et al.* 2006). Sequence reads with the same bar-code were grouped using a script that integrates the whole genome alignment tool MUMmer (Kurtz *et al.* 2004) with the 16S rRNA gene reverse primer and bar-code sequences as references.

The error rate in pyrosequencing is reportedly 0.25% per read (Huse *et al.* 2007), suggesting that most of the sequence reads obtained after this filtering contained

either 0 or 1 error. As demonstrated previously (McKenna *et al.* 2008), this potential single nucleotide error will not interfere with data analysis and is unlikely to have caused a sequence to be assigned to the incorrect taxon because more than 12–14 errors would be required to form a new taxon at the species level using an operational taxonomic unit threshold of 97% sequence similarity (McKenna *et al.* 2008). The assignment of sequences to taxonomic units was carried out using the Ribosomal Database Project Bayesian classifier (Wang *et al.* 2007).

Statistical analysis

Estimation of species richness. To investigate changes in phylotype richness along the chronosequences, both parametric and nonparametric richness estimation methods were used. This was carried out for all 52 individual samples taken along the chronosequence. Sequences were aligned within each sample using the secondary-structure alignment program INFERNAL that is accessible through the RDP pyrosequencing webpage (Cole *et al.* 2009). We then computed evolutionary distances between all sequences within each sample using DNAdist from the Phylip 3.68 package (Felsenstein 1993), with default parameters (F84 distance correction, transition/transversion ratio of 2.0). Given these distances, we used DOTUR (Schloss & Handelsman 2005) to cluster all sequences using furthest-neighbour clustering with a 1%, 3% and 10% similarity thresholds. The resulting abundance distributions were then converted into frequency counts to determine how many sequences occurred once (singletons), twice (doubletons) and so on. The approaches for parametric richness estimation were previously described by Bunge & Barger (2008). In brief, different models were fitted to the frequency counts for all possible truncation points of the frequency counts. Different goodness of fit tests (GF0–GF5) and the Akaike Information Criterion (AIC) were then used to determine how well each model fit the frequency count data (see Bunge & Barger (2008) for details). For a model to be chosen to fit the data best, GF0–GF5 had to be greater than 0.01, and the AIC had to be lower than that for other models. We used estimates based on finite mixture models, because these models performed generally better than others (see Bunge & Barger (2008) for details). For example, the Inverse Gaussian model frequently resulted in unreasonably large estimates with large standard errors. For nonparametric estimators, the ACE and ACE1 estimates of richness were used (Chao & Bunge 2002). If the coefficient of variation (CV) of ACE was greater than 0.8, the estimate from ACE1 was used

and if CV less than 0.8, the estimate from ACE was used (Chao & Bunge 2002). Curves were fitted to the parametric and nonparametric richness estimates observed for each different threshold of sequence similarity to describe the change in richness over time. The significance of the trend was tested using analysis of variance (ANOVA) based on stepwise model selection starting with the most parameter-rich model (fourth order model). Significance of the model was then determined correcting for multiple testing. To ensure that richness estimation was not influenced by low number of sequences per individual sample, the analysis was repeated, and this time the sequences from all the replicates for the corresponding time of exposure since glacier retreat were pooled.

Calculation of evenness. The evenness index E_{var} (Smith & Wilson 1996) was used to quantify evenness at different locations along the chronosequences. E_{var} is mathematically independent from richness and it has the property of being symmetric, that is, it weighs abundant and rare taxa equally (Beisel *et al.* 2003). Although this index may have an overall lower sensitivity (Beisel *et al.* 2003), symmetry is an essential characteristic, particularly when measuring evenness in bacterial communities that have large numbers of rare members.

Calculation of turnover rate. The changes in species composition over time (species turnover rate or beta diversity) along the chronosequences were calculated using a two-step procedure because it was computationally challenging to determine pairwise distances for all 230 000 sequences. First, all sequences from the time of exposure were pooled and pairwise distances between sequences were calculated using DNAdist (F84 correction with default transition transversion ratio, Felsenstein 1993). Sequences were then clustered based on a threshold of 97% sequence similarity using the furthest-neighbour algorithm (DOTUR, Schloss & Handelsman 2005). Second, for each cluster, we selected the first FASTA sequence identifier as a proxy. These sequences representing the different clusters from the six times since glacier retreat were then pooled. To determine the similarity of these sequences, all pairwise distances were calculated using DNAdist (Felsenstein 1993), and the sequences were clustered into phylotypes based on a threshold of 97% sequence similarity using DOTUR (Schloss & Handelsman 2005). After the clustering, a table was made that showed which phylotype was present at which time since glacier retreat. Based on this matrix, the turnover rate was then calculated using the measures described by Whittaker (1972) and Diamond & May (1977) using the R package (R Foundation for Statistical Computing).

Results

In this study, we assessed bacterial diversity along two approximately parallel chronosequences in a glacier foreland in the High Arctic based on phylotype richness, evenness and turnover rate. This is challenging because different methods use different metrics that lead to disparities in the estimates. For example, non-parametric methods to estimate richness consistently provide lower estimates of richness than do parametric methods (Bunge & Barger 2008). Here, we used multiple approaches for each aspect of diversity investigated and these provided a range of values for phylotype richness, evenness and turnover rates. Various thresholds of sequence similarity among 16S rRNA gene sequences are commonly used as a proxy for different taxonomic levels in studies on microbial diversity with 90%, 97% and 99% corresponding to differences between genera, species and strains respectively (Hagström *et al.* 2002, Oakley *et al.* 2008).

Changes in phylotype richness

Phylotype richness was estimated and changes were determined along the chronosequences using both non-parametric (ACE and ACE1, Chao & Lee 1992) and parametric richness estimates (mixture models, Bunge & Barger 2008). Estimates of phylotype richness were obtained for three thresholds of sequence similarity for the replicate soil samples taken from each time since glacier retreat (Fig. 1). Richness estimates obtained from the corresponding samples from two chronosequences were in agreement (data not shown). For all times since glacier retreat, nonparametric estimates were consistently lower than parametric estimates, but both approaches indicated that phylotype richness was high. For example, using a threshold of 97% sequence similarity phylotype richness at the different locations ranged from 2700 to 5800 per 0.5 g of soil, whereas the parametric estimates ranged on average from 4100 to 7000 per 0.5 g of soil. As expected, estimates of richness changed as the threshold for sequence similarity varied. For example, at a threshold of 90% sequence similarity, there was an average of 2700 phylotypes per 0.5 g of soil at 5 years since glacier retreat (nonparametric estimate), but with a threshold of 99% sequence similarity, the richness estimate averaged 6000 phylotypes per 0.5 g of soil (nonparametric estimation, Fig. 1). Furthermore, phylotype richness varied among replicates within each time of glacier retreat (Fig. 1). For example, at 63 years after glacier retreat, the lowest estimate for richness at 97% sequence similarity was approximately 4000 phylotypes per 0.5 g of soil and the largest was about 7400 phylotypes per 0.5 g of soil (parametric esti-

mation). Using regression analysis in conjunction with ANOVA based on stepwise model selection (data not shown), we found that phylotype richness increased significantly over time despite spatial heterogeneity (Fig. 1). This was the case for all three thresholds of sequence similarity, which suggests that phylotype richness changed not only on a fine taxonomic scale (99% sequence similarity), but also on a broader taxonomic scale (90% sequence similarity). As the number of sequences in each sample was low, the standard errors, not among the replicates, but obtained when fitting the parametric and nonparametric models to the frequency count data, were large for some of the individual richness estimates. By pooling the sequences from all the replicates of the corresponding time since glacier retreat, we had about 35 000 sequences per sample, standard errors for each of the richness estimates were low and we were able to confirm that phylotype richness increased over time (data not shown). Using this approach, we were able to show that for all three thresholds of sequence similarity, phylotype richness was high and increased significantly over time independent of the method used to estimate richness.

Changes in phylotype evenness

The structure of a bacterial community is characterized also by the evenness of the phylotype rank-abundance distribution. Thus, we determined changes in evenness along the chronosequence in the foreland of Midre Lovén glacier using the index E_{var} . This index is mathematically independent from richness (Smith & Wilson 1996) and weighs variation in rare phylotypes equally to variation in abundant members. This characteristic of symmetry seemed to be essential for determining evenness of these communities given that no phylotypes in these bacterial communities were abundant (about 9% was 'abundant' in our samples, but 40% is considered abundant in Smith & Wilson (1996) and Beisel *et al.* (2003)). As for phylotype richness, the evenness values were determined for sequence similarity thresholds of 90%, 97% and 99% for each of the replicate soil samples taken from each location along the chronosequences (Fig. 2). Evenness did not differ between the two chronosequences (data not shown). Overall, evenness was high considering that log-normal distributions of bacterial species are thought to be common for bacterial communities in natural environments (Curtis *et al.* 2002). For example, at 150 years since glacier retreat, the evenness values of E_{var} for 99% sequence similarity were on average 0.74. Evenness increased with finer taxonomic resolution. This is illustrated by the values obtained at 5 years since glacier retreat; the evenness values of E_{var} for 90% sequence similarity were on

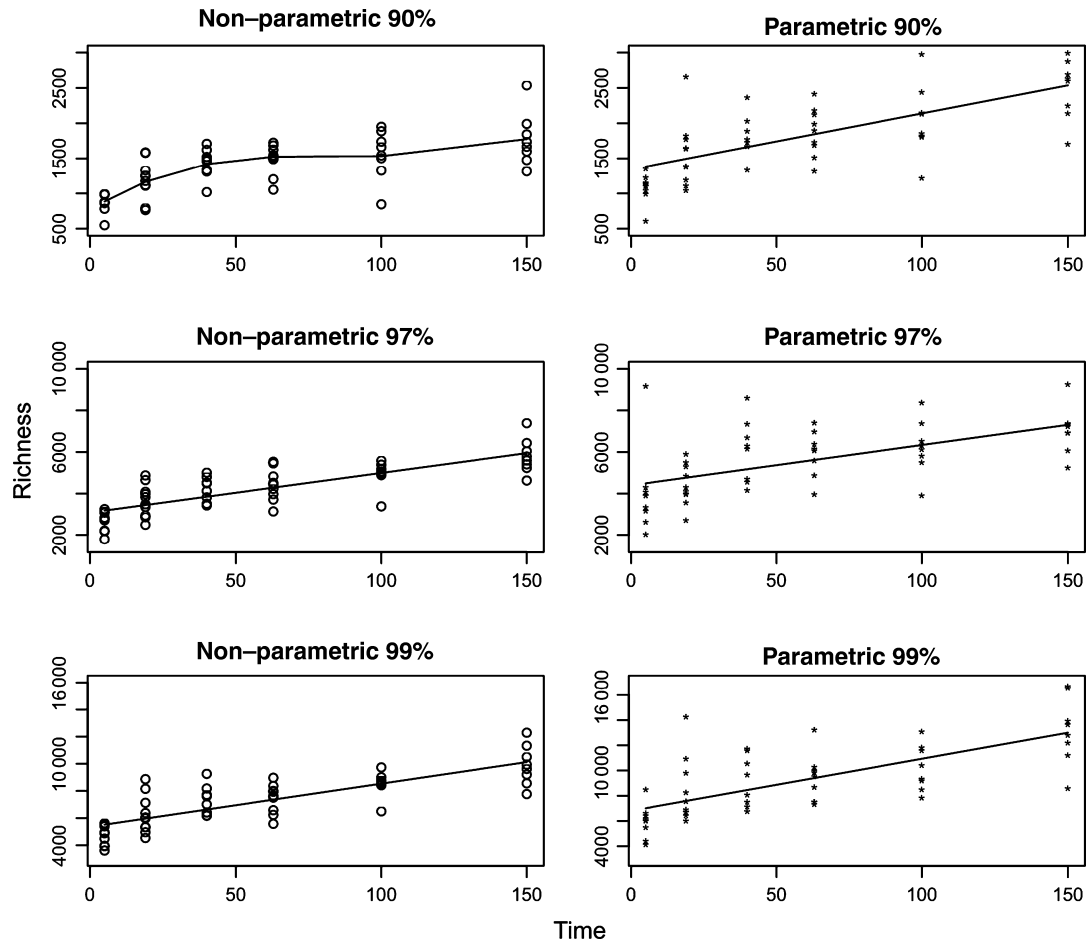


Fig. 1 Estimation of phylotype richness of bacterial communities sampled from six times since glacier retreat along the chronosequence based on data from 454 pyrosequencing of 16S rRNA genes. Richness estimation was carried out for three levels of sequence similarities (90%, 97% and 99%) using nonparametric and parametric richness estimators (Bunge & Barger 2008). Estimates from the corresponding samples from the two chronosequences were in good agreement and thus were not shown separately. If the coefficient of variance (CV) of the frequency data was greater than 0.8, the estimate ACE1 was used; otherwise ACE was used. For a parametric model (either two-mixed exponential or three-mixed exponential) to be chosen to fit the data best, GF0–GF5 had to be greater than 0.01 and the AIC had to be the lowest compared with the other models for that particular truncation point. Regression analysis was performed for each percentage sequence similarity to describe the shape of change in richness over time. The significance of the trend was tested using analysis of variance (ANOVA) based on stepwise model selection starting with the most parameter-rich model and correcting for multiple testing. Linear models were significant except for nonparametric estimation based on 90% sequence similarity in which case a third order model was significant.

average 0.39, at 97% sequence similarity the average was 0.55 and evenness increased to an average of 0.64 at 99% sequence similarity (Fig. 2). This suggests that phylotypes that were more dominant than other phylotypes on a broader taxonomic scale consisted of a number of relatively evenly distributed 'species' or 'strains', which resulted in an overall higher evenness at a finer taxonomic scale. As for phylotype richness, there was some within-site variation in evenness for each time since glacier retreat. For example, at 40 years since glacier retreat and a threshold of 90% sequence similarity, evenness values of E_{var} of the individual replicates ranged from 0.4 to above 0.5 (Fig. 2). Evenness values

were plotted over time (Fig. 2), but curves were not fitted to the data because the assumption of homoscedasticity (also known as constant variance) was violated even after data transformation. However, although the absolute values of the two evenness indices differ, they are consistent in that both suggest an increase in evenness over time.

The trends in Figs 1 and 2 suggest a positive correlation between phylotype richness and evenness. To assess whether phylotype evenness was significantly correlated with richness, Pearson correlation analysis was carried out for both combinations of richness estimators and evenness measure used in this study. Phylo-

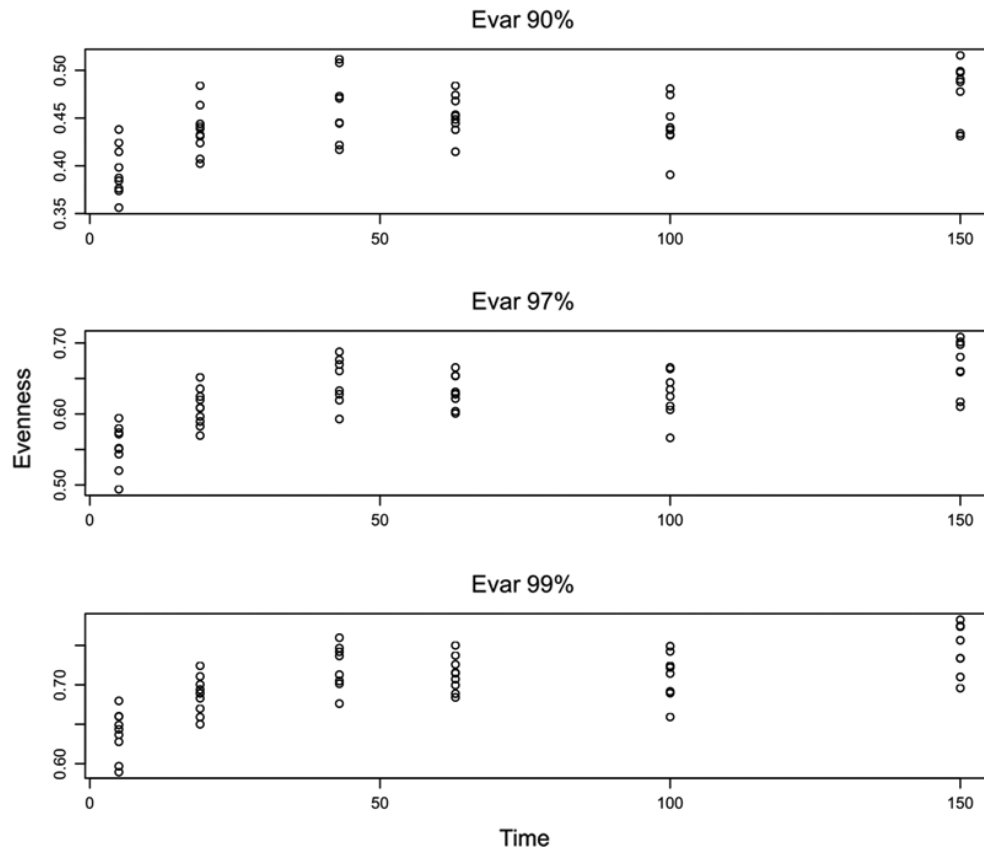


Fig. 2 Phylotype evenness of bacterial communities sampled from six times since glacier retreat along the chronosequence based on data from 454 pyrosequencing of 16S rRNA genes. Evenness was determined for three levels of sequence similarities (90%, 97% and 99%) using the $E_{1/D}$ and E_{var} index respectively. Estimates from the corresponding samples from the two chronosequences were in good agreement and thus were not shown separately.

type richness was significantly positively correlated with evenness for all comparisons and for all thresholds of sequence similarities (Table 1). This positive correlation suggests that in these soils, phylotype richness and evenness may both be meaningful descriptors of bacterial diversity.

Table 1 Pairwise Pearson correlation coefficients between phylotype richness and evenness for three levels of sequence similarities of 16S rRNA genes

Sequence similarity (%)	Type of richness estimation	Evenness
		E_{var}
90	Nonparametric	0.624*
	Parametric	0.528*
97	Nonparametric	0.718*
	Parametric	0.537*
99	Nonparametric	0.700*
	Parametric	0.622*

* $P < 0.001$.

Phylotype turnover rates

To measure changes in compositions of bacterial communities along the chronosequence, we determined the turnover rate using the measure by Whittaker (1972) and Diamond & May (1977) based on a 97% sequence similarity. These indices are measures of turnover and they describe the dissimilarity of community compositions. The measure by Diamond & May (1977) additionally accounts for time and thus, it represents a rate of change. Possible values range from zero to one and higher values suggest larger turnover. As a result of computational complexity, these calculations were only carried out for the data sets in which sequences of all replicate soil samples from each time since glacier retreat were pooled. Based on the assumption of the 'chronosequence approach' that space can be substituted for time (Walker & del Moral 2003), the results based on the measure by Whittaker (1972) suggest a high turnover in community composition for all pairwise comparisons of bacterial communities along the

Table 2 Turnover rate of bacterial communities along the chronosequence determined using dissimilarity measure of Whittaker (1972)

Years since glacier retreat	5	19	40	63	100	150
5	0	0.978	0.978	0.97	0.984	0.96
19		0	0.986	0.982	0.988	0.976
40			0	0.977	0.977	0.973
63				0	0.981	0.964
100					0	0.977
150						0

chronosequence (Table 3). For example, the turnover in phylotypes between 5 and 19 years since glacier retreat suggests that about 98% of the phylotypes between the communities differed. This value was similar to the comparison of the community composition obtained from 5 and 150 years since glacier retreat (0.96, Table 2). The analysis of turnover using the Whittaker index did not take into account that the difference in times since glacier retreat was not constant among all sites sampled. For example, the difference in time since glacier retreat between the first and second location was 14 years, but between locations five and six it was 50 years, and from the first location to the location exposed the longest it was 145 years. Using the method by Diamond & May (1977), the results showed that in the first 14 years turnover rate was about 7% per year (0.071, Table 3) and that this rate decreased to about 0.7% per year from 100 to 150 years since glacier retreat (0.006, Table 3). Overall, these results suggest that there was tremendous turnover of bacterial communities along the chronosequence and that changes in bacterial community compositions were particularly rapid in the early times following glacier retreat.

Bacterial community composition

Similar to the study by Roesch *et al.* (2007), a large proportion of the 16S rRNA gene sequences was not simi-

Table 3 Change in bacterial community composition per year (turnover rate per year) along the chronosequence determined using measure employed by Diamond & May (1977)

Years since glacier retreat	5	19	40	63	100	150
5	0	0.071	0.029	0.017	0.010	0.007
19		0	0.047	0.023	0.012	0.007
40			0	0.043	0.016	0.009
63				0	0.027	0.011
100					0	0.02
150						0

lar to anything previously reported; 15% of the sequences could not be assigned to any known phylum and on average 70% of the sequences in a sample could not be classified into a genus. The sequences that could be classified were assigned to 20 phyla (Acidobacteria, Chlamydiae, BRC1, Nitrospira, OD1, Chloroflexi, WS3, Deinococcus-Thermus, Bacteroidetes, Proteobacteria, OP11, Firmicutes, Planctomycetes, Spirochaetes, OP10, TM7, Gemmatimonadetes, Actinobacteria, Cyanobacteria, Verrucomicrobia) and 284 genera (see Tables S1 and S2, Supporting information). The 10 most abundant genera were *GP4*, *Spirosoma*, *Sphingomonas*, *GP6*, *GP3*, *Terromonas*, *Hymenobacter*, *Gemmatimonas*, *Brevundimonas* and *Sphingopyxis*. Many sequences found belonged to genera that grow oligotrophically and chemoorganotrophically such as 19 genera belonging to the phylum Acidobacteria and 54 genera belonging to the Actinobacteria. Diverse genera of nitrogen fixing genera were present including the cyanobacterial genera *GPVI*, *GP11a*, *GPI* and *GPIV*, as well as *Rhizobium*, *Mesorhizobium*, *Bradiorhizobium*, *Hydrogenophaga*, *Burkholderia*, *Herbaspirillum*, *Azohydromonas*, *Blastobacter*, *Methylocystis*, *Agromonas*, *Methylosinus*, *Ensifer*, *Acetobacterium*, *Pelomonas*, *Devosia* and *Methylocella*. The large proportion of the unassigned sequences underlines the large diversity and how little we know about the bacterial diversity in these High Arctic soils.

Discussion

This is the first time that bacterial diversity in a terrestrial ecosystem in the High Arctic has been intensively and extensively assessed using cultivation-independent methods. Replicate soil samples from six locations along two parallel chronosequences were collected in the foreland of Midtre Lovén glacier that span a 150 years period since deglaciation. The diversity was characterized by phylogenetic analysis of 16S rRNA sequences that had been amplified from genomic DNA extracted from the soil samples. By using high throughput pyrosequencing methods, about 35 000 sequences were obtained for each of the six times following deglaciation. Despite the spatial heterogeneity of these habitats, the experimental design and statistical methods used allowed us to demonstrate that bacterial phylotype richness and evenness were higher than we expected even at sites that had been exposed for as little as 5 years and increased more or less linearly over subsequent years.

There was a statistically significant positive correlation between changes in phylotype richness and evenness along the chronosequence. This positive correlation is consistent with the models describing relationships among various diversity indices by deBenedictis (1973)

and Hill (1973), but stand in contrast to multiple studies that found either no correlation or a negative relationship between species richness and evenness (Stirling & Wisley 2001; Ma 2005; Bock *et al.* 2007). The index used in this study to calculate evenness is, in contrast to other indices such as NHC, mathematically independent from phylotype richness (Smith & Wilson 1996). The bacterial communities present in soils of the Midtre Lovén glacier foreland showed high evenness with no dominant members and a large proportion of rare taxa. Using a threshold of 97% sequence similarity, about 25% of the sequences in each sample occurred only once or twice in a sample, while common types such as Acidobacteria only accounted for about 9% of all the sequences. Thus, the increase in species richness reflects the addition of uncommon phylotypes, resulting in a more flat rank-abundance distribution of phylotypes (higher evenness) and thus, a positive correlation between richness and evenness. However, a sampling bias may also partly account for this correlation (Bent & Forney 2008). To illustrate this, consider two bacterial communities that have comparable phylotype richness, but in one the phylotypes are equally abundant, while in the other the rank abundance of phylotypes is skewed with one taxon accounting for 90% of all the members, and additional rare populations accounting for the remaining 10%. DNA sequencing constitutes a 'random draw' from the pool of 16S rRNA gene sequences in a sample (Bent & Forney 2008). Hence, in the latter community it is more probably that sequences from the dominant population will be sampled, while in the former community there is a higher probability that a 'draw' will be from a taxon not previously sampled and the apparent richness of the community would be higher (Bent & Forney 2008). Fortunately, the sampling bias in this study may be of less importance as the distribution of all populations in these bacterial communities was relatively even.

Perhaps the most striking finding from this study was the extraordinarily high species turnover (beta diversity) that occurred soon after glacier retreat and that continued thereafter. This was consistent with our previous findings that succession occurred along the chronosequences (Schütte *et al.* 2009). Although this result is based on the assumption of the 'chronosequence approach' that space can be substituted for time (Walker & del Moral 2003) and the cause of this species turnover is of course unknown, it is tempting to speculate that it was driven by substantial changes in the soil environment following deglaciation. Alternatively, it could reflect the fact that bacterial populations initially present were not well suited to the extant conditions in newly exposed soils or that immigrant populations could occupy ecological niches that developed follow-

ing deglaciation and weathering of the rocks and minerals. At later times along the chronosequence, the turnover rate decreased to 0.006, indicating that the soil community composition had stabilized somewhat and the organisms present effectively prevented colonization by invading organisms. It is difficult to compare the turnover rates observed to those found in other studies of terrestrial habitats, because such measures are influenced by the sampling design, the area sampled and the metrics employed (Koleff *et al.* 2003). Some measures of turnover rate weigh the number of species present in both communities more, while others weigh more the number of species exclusive to one or the other location (Koleff *et al.* 2003). In addition, studies such as that of Noguez *et al.* (2005) employed a nested quadrat design, while we sampled along transects. Nonetheless, our results were generally consistent with those of Noguez *et al.* (2005) in which the turnover rate based on the species area relationship (Whittaker 1972; Arita & Rodriguez 2002) was determined to be 0.42 and 0.47 at two different locations in a biosphere reserve at the western coast in Mexico. These values were higher than those observed in animals, invertebrates and protists (Noguez *et al.* 2005).

The extraordinary diversity found in this glacial forefield is consistent with the results from other studies on bacterial diversity in polar environments and adds to the growing list of habitats, in which bacterial diversity has been shown to be unfathomable (Borneman *et al.* 1996; Torsvik *et al.* 1998; Zhou *et al.* 2004; Neufeld & Mohn 2005; Sogin *et al.* 2006; Elshahed *et al.* 2008). Additionally, similar to other studies (Roesch *et al.* 2007;), significant fractions of the phylotypes in these communities were potentially novel. About 15% of all sequences could not be classified into known bacterial phyla and about 70% of all the sequences could not be classified into a known genus. Thus, with each additional study, we gain a better understanding of the full extent of prokaryotic diversity and its distribution on Earth. The study by Roesch *et al.* (2007) is the most comparable to ours, because the authors estimated bacterial richness in soils by analysis of 16S rRNA gene sequences obtained by pyrosequencing. Interestingly, Roesch *et al.* (2007) found that for a threshold of 97% sequence similarity phylotype richness in a Canadian forest soil (13 000 per g) was higher than that of three agricultural soils from temperate (Florida and Illinois, USA) and tropical (Brazil) climates, which were similar to one another (5–6 thousand per g). This was true for all thresholds of sequence similarity examined. Moreover, the species accumulation (rarefaction) curves for the Canadian soil markedly differed in shape from those determined for the other soils, and suggested there was greater phylotype evenness. Thus, the general

finding of higher phylotype richness and evenness in the Canadian soil is reminiscent of that observed in the Arctic glacial forefield. Broad generalizations based on so few data are admittedly risky, but it is tempting to speculate that the latitudinal gradient in species diversity observed for plants and animals in which diversity is the greatest in the tropics and decreases towards the poles may not hold true for bacteria.

The difficulty faced in adequately sampling the high bacterial diversity in soils and differences in the computational methods used to estimate richness, evenness and species turnover rates hamper the assessment of prokaryotic species diversity and make it difficult to compare the results of different studies. For example, nonparametric methods used to estimate phylotype richness are known to be more conservative (Bunge 2008) and consistently result in lower phylotype richness estimates than parametric methods. This was the case in this study. Nonetheless, the estimates of phylotype richness both indicated that richness was high and followed the same trend over time, although the absolute values sometimes differed as much as two-fold (Fig. 1). It is not possible to know which estimates were closer to the true values, because the true underlying abundance distributions in the microbial communities were unknown. Similarly, different evenness indices may suggest different evenness of the members of the community, and in this study, we chose E_{var} because this index weighs the proportion of predominant members and rare members equally (Smith & Wilson 1996). Such discrepancies in estimates of diversity metrics are not uncommon because the means of estimating these metrics have limitations and assumptions that may not always be appropriate (for more detailed reviews see Beisel *et al.* 2003; Bunge & Barger 2008; Koleff *et al.* 2003; Smith & Wilson 1996). We have no easy solution for this problem and instead used multiple measures to obtain a range of possible values for each diversity metric.

As with other efforts to characterize bacterial diversity based solely on cataloguing 16S rRNA gene sequences, it is not possible to reliably infer the functions of the bacterial populations present. However, the general characteristics of some of the populations present in the glacial foreland soils are known and worth noting. Two phyla, Acidobacteria and Actinobacteria, were found in larger numbers along the chronosequences. Acidobacteria are commonly found in soils (Neufeld & Mohn 2005; Roesch *et al.* 2007) and are known to grow under oligotrophic conditions and at low pH (Dion 2008). Thus, the abundance of genera belonging to Acidobacteria in samples from the entire chronosequence and their increase in number over time may reflect the low levels of nutrients present in High Arctic soils and the decrease in soil pH that occurs along the

chronosequence (Hodkinson *et al.* 2003). Several genera belonging to the phylum Actinobacteria were also found; in particular *Nocardia* found in our soil samples are known to form mycelia. The formation of mycelia is suggested to be advantageous in oligotrophic and desert environments, because the hyphae allow the bacteria to reach water and nutrients in pores further away (Dion 2008). More striking was the presence of N-fixing genera besides Cyanobacteria such as *Rhizobium*, *Mesorhizobium* and *Bradiorhizobium*. Most studies in terrestrial polar environments on nitrogen fixation focused on Cyanobacteria (Whitton & Potts 2000); however, our data suggest that additional genera should be considered to contribute to N-fixation in High Arctic soils. In addition, a larger proportion of chemoorganotrophic phylotypes (32%) than photoautotrophs (10%) was observed at the youngest sites since glacier retreat, which contradicts the community assembly rule that photoautotrophic organisms are the first in early succession. These results are consistent with Skidmore *et al.* (2000), who detected viable chemoorganotrophic bacteria under the John Evans glacier, Canada and the suggestion of Hodkinson *et al.* (2002) that heterotrophic organisms may precede autotrophs in community assembly.

The mechanisms that work to maintain high bacterial diversity in soils of glacial forelands in the Arctic are unknown. We postulate that three mechanisms working individually or in some combination seem plausible. First, all soils are highly structured environments composed of soil aggregates that have a plethora of different microenvironments. These microenvironments differ from one another in many ways including pore water chemistry, the composition of particulates, the kind and amounts of organic matter, gas exchange, water holding capacity, the efficiency of binding bacterial cells and other features (Ehrlich & Newman 2008). Moreover, differences in micro-topography are associated with variation in temperature and wind. The matrix of pertinent environmental parameters creates a high order of ecological niches that can select for genotypes that are particularly well suited for local conditions. Second, bacterial populations in glacier foreland soils experience disturbances of intermediate strength on a range of temporal scales. For example, on a diurnal scale soil, temperatures can fluctuate from -10 to 25 °C within a few hours (Torsvik & Ovreas 2008), while over longer periods, there are changes in precipitation and in the glacial drainage system. The intermediate-disturbance hypothesis (Morin 1999) posits that such disturbances prevent the domination of a few organisms and the resulting community is even and species rich. Finally, competitive exclusion may be ineffective because of low bacterial growth rates that restrict the ability of one or more populations to exploit existing conditions and increase

in number. Bacterial growth rates in High Arctic environments may be restricted because of lack of water, solar radiation, fluctuations in temperature and continuous freeze–thawing (Torsvik & Ovreas 2008). Moreover, temperatures are generally low and there is a paucity of nutrients that would support the growth of heterotrophs. The exceptionally slow growth of organisms in polar environments is illustrated by the fact that cyanobacterial cells in Antarctica were estimated to divide once every other year (Fritsen & Priscu 1996). Given the conditions that prevail, potential competitive advantages are counteracted by other factors that limit growth. Thus, individual populations are left to coexist and are unable to competitively exclude other phylotypes in the bacterial community.

Acknowledgements

The field work was funded by the Amundsen Center at the University of Tromsø and the Norwegian Polar Institute. The DNA Sequence Analysis Core Facility at the University of Idaho is supported by an NIH Center of Biomedical Research Excellence grant (P20 RR016448) from the National Center for Research Resources to L.J.F. Phylotype richness estimation was funded by NSF grant #0816638 'Estimating Biological Diversity and its Patterns'. This research was conducted using the resources of the Cornell University Center for Advanced Computing, which receives funding from Cornell University, New York State, the National Science Foundation, and other leading public agencies, foundations and corporations. This research was (partially) supported by Grant No. AB 2046 from the Student Grant Program at the University of Idaho, Moscow, ID. We thank Dr Rolf A. Olsen and the University Centre in Svalbard (UNIS) for facilitating our field research and Dr Ian Hodkinson and Dr Steve Coulson for information on their sampling locations and valuable discussions. We also wish to acknowledge Dr Eva Top for advice on methods for DNA isolation, Dr Chris Williams for valuable discussions on the experimental design, Dr Stefano Ventura and Silvia Turichia for their assistance in the field. Finally, we appreciate helpful comments on the manuscript provided by Jacob Pierson.

Conflicts of interest

The authors have no conflict of interest to declare and note that the sponsors of the issue had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Arita HT, Rodriguez P (2002) Geographic range, turnover rate and the scaling of species diversity. *Ecography*, **25**, 541–550.
 Beisel JN, Usseglio-Polatera P, Bachmann V, Moreteau JC (2003) A comparative analysis of evenness index sensitivity. *International Review of Hydrobiology*, **88**, 3–15.

deBenedictis PA (1973) On the correlations between certain diversity indices. *American Naturalist*, **107**, 295–302.
 Bent SJ, Forney LJ (2008) The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. *ISME J*, **2**, 689–695.
 Bernstein L, Bosch P, Canziani O, Chen Z, Christ R, Davidson O *et al.* (2007) Climate change 2007: synthesis report. *Synthesis report*. Available at: <http://www.ipcc.ch/>.
 Bock CE, Jones ZF, Bock JH (2007) Relationships between species richness, evenness, and abundance in a southwestern savanna. *Ecology*, **88**, 1322–1327.
 Borneman J, Skroch PW, O'Sullivan KM *et al.* (1996) Molecular microbial diversity of an agricultural soil in Wisconsin. *Applied and Environmental Microbiology*, **62**, 1935–1943.
 Broady PA (1996) Diversity, distribution and dispersal of Antarctic terrestrial algae. *Biodiversity and Conservation*, **5**, 1307–1335.
 Bunge J (2008) Statistical estimation of uncultivated microbial diversity. In: *Uncultivated Microorganisms* (ed. Epstein S), pp. 1–18. Springer, Berlin, Heidelberg.
 Bunge J, Barger K (2008) Parametric models for estimating the number of classes. *Biometrical Journal*, **50**, 971–982.
 Chao A, Bunge J (2002) Estimating the number of species in a stochastic abundance model. *Biometrics*, **58**, 531–539.
 Chao A, Lee SM (1992) Estimating the number of classes via sample coverage. *Journal of the American Statistical Association*, **87**, 210–217.
 Cole JR, Wang Q, Cardenas E *et al.* (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research*, **37**, D141–D145.
 Curtis TP, Sloan WT, Scannell JW (2002) Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 10494–10499.
 DeSantis TZJ, Hugenholtz P, Keller K *et al.* (2006) NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Research*, **34**, W394–W399.
 Dethlefsen L, Huse S, Sogin ML, Relman DA (2008) The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biology*, **6**, 2383–2400.
 Diamond JM, May RM (1977) Species turnover rates on islands: dependence on census interval. *Science*, **197**, 266–270.
 Dion P (2008) Extreme views of prokaryote evolution. In: *Microbiology of Extreme Soils* (eds Dion P, Nautiyal CS), pp. 45–70. Springer, Berlin Heidelberg.
 Dowdeswell JA, Hagen JO, Björnsson H *et al.* (1997) The mass balance of circum-Arctic glaciers and recent climate change. *Quaternary Research*, **48**, 1–14.
 Ehrlich HL, Newman DK (2008) *Geomicrobiology*. Taylor & Francis Group, LLC, Boca Raton, FL.
 Elshahed MS, Youssef NH, Spain AM *et al.* (2008) Novelty and uniqueness patterns of rare members of the soil biosphere. *Applied and Environmental Microbiology*, **74**, 5422–5428.
 Felsenstein J (1993) Phylogeny inference package (version 3.2). *Cladistics*, **5**, 164–166.
 Fritsen CH, Priscu JC (1996) Photosynthetic characteristics of cyanobacteria in permanent ice covers on lakes in the McMurdo Dry Valleys, Antarctica. *Antarctic J US*, **31**, 216–218.

- González-Toril E, Amils R, Delmas RJ, Petit JR, Komárek J, Elster J (2008) Diversity of bacteria producing pigmented colonies in aerosol, snow and soil samples from remote glacial areas (Antarctica, Alps and Andes). *Biogeosciences*, **5**, 1607–1630.
- Green J, Bohannan JM (2006) Spatial scaling of microbial biodiversity. *Trends in Ecology & Evolution*, **21**, 501–507.
- Green JL, Holmes AJ, Westoby M *et al.* (2004) Spatial scaling of microbial eukaryote diversity. *Nature*, **432**, 747–750.
- Hagström A, Pommier T, Rohwer F *et al.* (2002) Use of 16S ribosomal DNA for delineation of marine bacterioplankton species. *Applied and Environmental Microbiology*, **68**, 3628–3633.
- Harland WB (1997) *The Geology of Svalbard*. The Geological Society, Bath, UK.
- Hill MO (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**, 427–432.
- Hodkinson ID, Webb NR, Coulson SJ (2002) Primary community assembly on land – the missing stages: why are the heterotrophic organisms always there first? *Journal of Ecology*, **90**, 569–577.
- Hodkinson ID, Coulson SJ, Webb NR (2003) Community assembly along proglacial chronosequences in the high Arctic: vegetation and soil development in north-west Svalbard. *Journal of Ecology*, **91**, 651–663.
- Horner-Devine MC, Lage M, Hughes JB, Bohannan BJ (2004) A taxa-area relationship for bacteria. *Nature*, **432**, 750–753.
- Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology*, **8**, R143.
- Huse SM, Dethlefsen JAHL, Welch DM, Relman DA, Sogin ML (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*, **4**, 1–10.
- Huston M, Smith T (1987) Plant succession: life history and competition. *American Naturalist*, **130**, 168–198.
- Kintisch E (2009) Global warming: projections of climate change go from bad to worse, scientists report. *Science*, **323**, 1546–1547.
- Koleff P, Gaston KJ, Lennon JJ (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367–382.
- Kurtz S, Phillippy A, Delcher AL *et al.* (2004) Versatile and open software for comparing large genomes. *Genome Biology*, **5**, R12.
- Lawley B, Ripley S, Bridge P, Convey P (2004) Molecular analysis of geographic patterns of eukaryotic diversity in Antarctic soils. *Applied and Environmental Microbiology*, **70**, 5963–5972.
- Ma M (2005) Species richness vs evenness: independent relationship and different responses to edaphic factors. *Oikos*, **111**, 192–198.
- Margulies M, Egholm M, Altman WE *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437**, 376–380.
- Matthews JA (1992). *The Ecology of Recently Deglaciated Terrain: A Geocological Approach to Glacier Forelands and Primary Succession*. Cambridge University Press, New York, NY.
- McKenna P, Hoffmann C, Minkah N *et al.* (2008) The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathogens*, **4**, 1–12.
- Morin PJ (1999) *Community Ecology*. Blackwell Science, Inc, Malden, MA.
- Nemergut DR, Anderson SP, Cleveland CC *et al.* (2007) Microbial community succession in an unvegetated, recently deglaciated soil. *Microbiol Ecol*, **53**, 110–122.
- Neufeld JD, Mohn WW (2005) Unexpectedly high bacterial diversity in arctic tundra relative to boreal forest soils, revealed by serial analysis of ribosomal sequence tags. *Applied and Environmental Microbiology*, **71**, 5710–5718.
- Noguez AM, Arita HT, Escalante AE, Forney LJ, García-Oliva F, Souza V (2005) Microbial macroecology: highly structured prokaryotic soil assemblages in a tropical deciduous forest. *Global Ecology and Biogeography*, **14**, 241–248.
- Oakley BB, Fiedler TL, Marrazzo JM, Fredricks DN (2008) Diversity of human vaginal communities and associations with clinically defined bacterial vaginosis. *Applied and Environmental Microbiology*, **74**, 4898–4909.
- Roesch LF, Fulthorpe RR, Riva A *et al.* (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME Journal*, **1**, 283–290.
- Schipper LA, Degens BP, Sparling GP, Duncan LC (2001) Changes in microbial heterotrophic diversity along five plant successional sequences. *Soil Biology and Biochemistry*, **33**, 2093–2103.
- Schloss PD, Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and Environmental Microbiology*, **71**, 1501–1506.
- Schmidt SK, Reed SC, Nemergut DR *et al.* (2008) The earliest stages of ecosystem succession in high-elevation (5000 metres above sea level), recently deglaciated soils. *Proceedings. Biological Sciences/The Royal Society*, **275**, 2793–2802.
- Schütte UME, Abdo Z, Bent SJ *et al.* (2009) Bacterial succession in a glacier foreland of the High Arctic. *The ISME Journal*, **3**, 1253–1268.
- Shi T, Reeves RH, Gilichinsky DA, Friedmann EI (1997) Characterization of viable bacteria from Siberian permafrost by 16S rDNA sequencing. *Microbial Ecology*, **33**, 169–179.
- Shivaji S, Reddy GSN, Aduri RP, Kutty R, Ravensschlag K (2004) Bacterial diversity of a soil sample from Schirmacher Oasis, Antarctica. *Cellular and Molecular Biology*, **50**, 525–536.
- Sigler WV, Zeyer J (2002) Microbial diversity and activity along the forefields of two receding glaciers. *Microbial Ecology*, **43**, 397–407.
- Sigler WV, Crivii S, Zeyer J (2002) Bacterial succession in glacial forefield soils characterized by community structure, activity and opportunistic growth dynamics. *Microbial Ecology*, **44**, 306–316.
- Skidmore ML, Foght JM, Sharp MJ (2000) Microbial life beneath a high arctic glacier. *Applied and Environmental Microbiology*, **66**, 3214–3220.
- Smith B, Wilson JB (1996) A consumer's guide to evenness indices. *Oikos*, **76**, 70–82.
- Sogin ML, Morrison HG, Huber J *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 12115–12120.
- Steven B, Pollard WH, Greer CW, Whyte LG (2007) Microbial diversity and activity through a permafrost/ground ice core

- profile from the Canadian high Arctic. *Environmental Microbiology*, **10**, 3388–3403.
- Stirling G, Wisley B (2001) Empirical relationships between species richness, evenness, and proportional diversity. *American Naturalist*, **158**, 286–299.
- Torsvik V, Ovreas L (2008) Microbial diversity, life strategies, and adaptation to life in extreme soils. In: *Microbiology of Extreme Soils* (eds Dion P, Nautiyal CS), pp. 15–43. Springer, Berlin Heidelberg.
- Torsvik V, Daae FL, Sandaa RA, Ovreas L (1998) Novel techniques for analysing microbial diversity in natural and perturbed environments. *Journal of Biotechnology*, **64**, 53–62.
- Vishniac HS (1996) Biodiversity of yeasts and filamentous microfungi in terrestrial Antarctic ecosystems. *Biodiversity and Conservation*, **5**, 1365–1378.
- Walker LR, del Moral R (2003) *Primary Succession and Ecosystem Rehabilitation*. Cambridge University Press, New York, NY.
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.
- Whittaker RH (1972) Evolution and measurement of species diversity. *Taxon*, **21**, 213–251.
- Whitton BA, Potts M (2000) *The Ecology of Cyanobacteria: Their Diversity in Time and Space*. Kluwer Academic Publishers, NL.
- Zhou J, Davey ME, Figueras JB, Rivkina E, Gilichinsky D, Tiedje JM (1997) Phylogenetic diversity of a bacterial community determined from Siberian tundra soil DNA. *Microbiology*, **143**, 3913–3919.
- Zhou J, Xia B, Huang H, Palumbo AV, Tiedje JM (2004) Microbial diversity and heterogeneity in sandy subsurface soils. *Applied and Environmental Microbiology*, **70**, 1723–1734.

Ursel Schütte is interested in understanding patterns of biogeography for microorganisms by applying macroecological concepts to the microbial world. Zaid Abdo is a professor of Statistics and Mathematics and he combines mathematical modeling and statistical analysis using computationally inten-

sive methods and simulations to address current problems in biology such as the analysis of microbial community composition and diversity, experimental evolution, and the barcoding of life. James Foster explores and attempts to understand both natural and simulated evolution by developing and analyzing genetic algorithms. He also investigates biological problems such as the emergence and structure of microbial ecosystems. Jacques Ravel's research interests extend from the application of microbial genomics to explore the human microbiome, to comparative analyses of microbial genome sequences and microbial forensics. John Bunge's research focuses on the problem of estimating the species richness or biodiversity in a community. Bjørn Solheim focuses on Arctic terrestrial ecosystems, in particular cyanobacterial communities and their role in nitrogen fixation, and the impact of climate change on these types of communities. Larry Forney's research centers on the diversity and distribution of prokaryotes. Both field and laboratory studies are done to explore the temporal and spatial patterns of community diversity, as well as factors that influence the dynamics of inter- and intra-species competition and how environmental conditions might influence the tempo of adaptive evolution.

Supporting Information

Additional supporting information may be found in the online version of this article

Table S1 List of phyla and genera present in any of the soil samples analyzed

Table S2 List of genera present at the different times since glacier retreat

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.