ORIGINAL ARTICLE

Macroecological patterns of marine bacteria on a global scale

Anthony S. Amend[1,2]*, Tom A. Oliver[3], Linda A. Amaral-Zettler[4,5], Antje Boetius[6,7], Jed A. Fuhrman[8], M. Claire Horner-Devine[9], Susan M. Huse[4], David B. Mark Welch[4], Adam C. Martiny[1], Alban Ramette[6,7], Lucie Zinger[6], Mitchell L. Sogin[4] and Jennifer B. H. Martiny[1]

[1]*Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA,* [2]*Department of Botany, University of Hawaii at Manoa, Honolulu, HI, USA,* [3]*Hawaiian Institute of Marine Biology, University of Hawaii at Manoa, Kaneohe, HI, USA,* [4]*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Woods Hole Marine Biological Laboratory, Woods Hole, MA, USA,* [5]*Department of Geological Sciences, Brown University, Providence, RI, USA,* [6]*Microbial Habitat Group, Max Planck Institute for Marine Microbiology, Bremen, Germany,* [7]*HGF MPG Joint Research Group on Deep Sea Ecology and Technology, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany,* [8]*Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA,* [9]*School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA*

ABSTRACT

Aim To test whether within-species and among-species patterns of abundance and latitudinal range in marine bacteria resemble those found for macro-organisms, and whether these patterns differ along latitudinal clines.

Location Global pelagic marine environments.

Methods Taxon-specific sequence abundance and location were retrieved from the open-access V6-rRNA pyrotag sequence data base VAMPS (http://vamps.mbl.edu/), which holds a massive collection of marine bacterial community data sets from the International Census of Marine Microbes sampling effort of global ocean water masses. Data were randomly subsampled to correct for spatial bias and for differences in sampling effort.

Results We show that bacterial latitudinal ranges are narrower than expected by chance. When present in both Northern and Southern hemispheres, taxa occupy restricted ranges at similar latitudes on both sides of the equator. A significant and positive relationship exists between sequence abundance and latitudinal range, although this pattern contains a large amount of variance. Abundant taxa in the tropics and in the Northern Hemisphere generally have smaller ranges than those in the Southern Hemisphere. We show that the mean latitudinal range of bacterial taxa increases with latitude, supporting the existence of a Rapoport effect in marine bacterioplankton. Finally, we show that bacterioplankton communities contain a higher proportion of abundant taxa as they approach the poles.

Main conclusions Macroecological patterns such as the abundance–range relationship, in general, extend to marine bacteria. However, differences in the shape of these relationships between bacteria and macro-organisms call into question whether the processes and their relative importance in shaping global marine bacteria and macro-organism distributions are the same.

Keywords
Abundance–range relationship, bacterioplankton, latitudinal gradient, macro-ecology, macro-organisms, marine microbes, micro-organisms, Rapoport's rule, rarity.

*Correspondence: Anthony S. Amend, Department of Botany, University of Hawaii at Manoa, 3190 Maile Way, Honolulu, HI 96822, USA.
E-mail: amend@hawaii.edu*

INTRODUCTION

A growing body of evidence suggests that micro-organisms exhibit large-scale biogeographical patterns that resemble those of macro-organisms. For example, using DNA sequencing techniques to identify molecular operational taxonomic units (OTUs), comparisons of microbial communities often show positive taxa–area relationships (e.g. Horner-Devine *et al.*, 2004; Bell *et al.*, 2005), latitudinal species richness gradients (e.g. Pommier *et al.*, 2007; Fuhrman *et al.*, 2008), and isolation-by-distance patterns (e.g. Bell, 2010; Martiny *et al.*, 2011) comparable (though not identical) to those of macro-organisms (Soininen, 2012).

Until recently, limited throughput and the high cost of DNA sequencing constrained molecular-based descriptions of microbial biogeography. The emergence of high-throughput

pyrosequencing techniques now makes it possible to construct global-scale data sets that capture information about distributions of high- and low-abundance microbes. Just as natural historians have mapped the distribution of plants and animals for centuries, microbial ecologists now have the opportunity to develop data sets that will inform us about microbial biogeography.

Among the most basic biogeographical patterns – noted as early as Darwin (1859) – is the positive relationship between a species' abundance and the size of that species' range (i.e. the extent of its distribution). This pattern is prevalent among a variety of distinct organisms, including plants, birds, insects, mammals and bacteria (Gaston & Lawton, 1990; Nemergut et al., 2011). Furthermore, the abundance–range relationship is generally robust to different metrics of abundance (e.g. counts of individuals summed over a species' range, relative abundance of a species, mean abundance where present) and to different metrics of distribution (e.g. range extent, area of occupancy, latitudinal extent).

There are at least three non-exclusive explanations for a positive abundance–range relationship. First, differences in the range of a species' habitat or resource use can drive distribution patterns. Generalist species with flexibility in resource use and wide environmental tolerances might be more abundant at a local scale and occur in a broader range of locations than specialist species (Brown, 1984). Second, the abundance–range relationship could result from metapopulation dynamics. Locally abundant species are likely to produce more dispersal agents, thereby increasing the distribution of a species regardless of its resource use (Hanski, 1982, 1991; Gaston et al., 2000). Finally, the pattern might reflect ascertainment bias: species that occur in low abundance are less likely to be detected at a given location, even if their true distribution is, in fact, widespread. While this last bias is generally unavoidable regardless of survey method, statistical methods can partly disentangle this artefact from a biological signal (Bock, 1987; Gaston et al., 2000).

The positive relationship between a species' abundance and its range tends to encompass a great deal of variation, although the fit to a linear model can increase as the functional and taxonomic breadth included in the analysis is narrowed (Brown, 1984). This variation appears to be constrained in large organisms even when relatively broad groups are considered (e.g. all flowering plants, or all birds). In particular, within any broadly defined and contiguous habitat, there are generally few macro-organisms that are locally abundant but narrowly distributed, such that the data points of the abundance–range relationship are constrained within a triangular area when plotted on a log-transformed scale (Gaston, 1994; Brown, 1995). This missing 'abundant local' category is one of three forms of within-habitat rarity codified by Rabinowitz (1981), the other two being locally rare species with a large range size ('sparse') and locally rare species with a small range size ('rare endemics'). Within the microbial literature, rarity, particularly OTUs of the rare endemic category, have recently garnered much attention for

contributing to estimates of species richness (Curtis & Sloan, 2005; Sogin et al., 2006). Less attention has been paid to how often micro-organisms fall within these other rarity categories.

Here, we use a global data set from the International Census of Marine Microbes (Amaral-Zettler et al., 2010) to examine whether marine bacteria on the ocean surface follow similar abundance–range patterns to those found in larger organisms. Recent studies have noted that bacterial sequence abundance is positively correlated with 'occurrence' (the fraction of samples in which a taxon is detected; Östman et al., 2010; Nemergut et al., 2011; Barberán et al., 2012). However, studies of occurrence patterns do not account for potential spatial autocorrelation and thus do not directly address range size. In this study, we use latitudinal range as a proxy for range size, as latitude is tightly correlated with numerous environmental clines (such as sea surface temperature and productivity) known to impact marine life, particularly in the near-surface stratum. Because we rely on molecular methods to sample microbial communities, we use the number of pertinent sequences in a sample as a proxy for abundance. These data are semi-quantitative and, although not perfect, are appropriate for comparing relative abundance within a taxon (Amend et al., 2010a).

Given this extensive data set, we addressed four aspects of the abundance–range relationships in bacterioplankton. First, using the Rabinowitz rarity categories as a reference, we tested whether rare and abundant taxa are randomly distributed across latitudes. Second, we examined overall patterns of the abundance and latitudinal range relationship of OTUs, and how the breadth of taxonomic focus affects this pattern. We expected a coarser taxonomic resolution to encompass greater functional diversity and therefore to exhibit greater variance in this relationship (Brown, 1995). Third, we examined how marine bacterial abundance and range vary by latitude. A commonly observed pattern in macro-organisms is that the average range size of taxa within a community correlates positively with latitude. Range sizes are, on average, larger within communities located near the poles. This relationship, also known as Rapoport's rule (Stevens, 1989), has been observed in both marine and terrestrial systems, although its generality as a 'rule' remains contentious (Rohde, 1996; Gaston et al., 1998). Finally, we hypothesized that if bacterioplankton conform to both an abundance–range relationship and to Rapoport's rule, we might predict how patterns of rarity and abundance themselves follow latitudinal gradients.

MATERIALS AND METHODS

Data set construction and metric calculations

The ICoMM data set

We used the International Census of Marine Microbes (ICoMM) sequence collection of *16S* rRNA gene V6 hypervariable region

pyrotags (aka barcode sequences) obtained by 454 pyrosequencing of DNA extracted and amplified from ocean samples. The data are publicly available through the VAMPS website (http://icomm.mbl.edu/). The ICoMM researchers originally surveyed microbial diversity from 575 environmental samples, including water samples from the major water masses in the world's ocean, resulting in one of the largest microbial sample-by-taxon matrices produced to date. This type of molecular sampling does not distinguish active organisms from those that are dormant (or recently dead), so samples may integrate data from communities with distinct temporal turnover. This data set is particularly amenable to global analyses because processing of the samples was relatively standardized and DNA amplification and sequencing was performed in a single laboratory using identical methods. Among some samples, pre-filters (from 3 to 0.8 μm) eliminated the largest particles, and the microbial fraction was ultimately extracted from 0.2-μm filters in all cases. For samples from the Max Planck Institue (MPI; see Appendix S1 in Supporting Information), where large- and small-size fractions were amplified by polymerase chain reaction (PCR) and sequenced separately, the two fractions were combined and analysed as a single sample here. Sequence quality control and processing are discussed at length elsewhere (Huse *et al.*, 2007, 2008, 2010; Zinger *et al.*, 2011). Whereas the observed accuracy of next-generation sequencing technologies has been a matter of some debate in the literature (e.g. Kunin *et al.*, 2010), potential sequencing errors should manifest evenly throughout the geographical range of our study so as not to bias comparative analyses.

Sample selection criteria

We used only data from pelagic samples of the upper water column within 200 m of the ocean surface. This depth represents the potential limit of the photic zone, and has been used as a near-surface depth cut-off for other studies (Zinger *et al.*, 2011). The mean depth in our study was 18.6 ± 38.6 m (standard deviation, SD). Although this depth range potentially contains a large amount of environmental heterogeneity, many factors (such as turbidity or thermocline depth) vary widely throughout the ocean such that no single depth cut-off will achieve environmental constancy throughout the world's oceans. There was no apparent latitudinal bias for sample depth (Fig. S1 in Appendix S2), and no significant difference in mean depth between hemispheres (independent two-tailed t-test, d.f. $= 255$, $P = 0.85$). As in Zinger *et al.* (2011), samples were classified as 'near-shore' if they were located < 200 nautical miles from a coastline. There were no significant differences between the proportions of near-shore and offshore samples in the Northern and Southern hemispheres (Fisher exact test, d.f. $= 1$, $P = 1$).

The resulting data set contained 3,893,852 individual sequences, derived from 257 pelagic samples (see Appendix S1). We binned sequences into OTUs at the 97% sequence identity level. At this cut-off level, our data set contained 28,153 total OTUs, 12,502 of which (44%) occurred more than once in our retrieved data set (i.e. non-singletons). Unless noted otherwise, all analyses were computed with the non-singleton matrix.

Calculation of abundance and range metrics

In some locations, multiple samples were retrieved at different times or depths. To account for the non-independence of these samples, we combined samples collected at identical latitude and longitude by summing their OTU counts. As a result, the 257 samples collapsed into 84 samples from unique locations (Fig. 1). To account for varying numbers of sequences in each unique location, each sample was standardized by randomly resampling without replacement to the number of sequences in the least sampled location (4280 sequences).

Latitudinal range, latitudinal centre (the mean of latitudes where an OTU is present, weighted by sequence abundance) and the rarefied abundance of each OTU used in subsequent analyses were calculated as the means from 100 such random sub-samples. Because any particular subsample may not include all rare OTUs, the mean abundance for some OTUs was < 1.
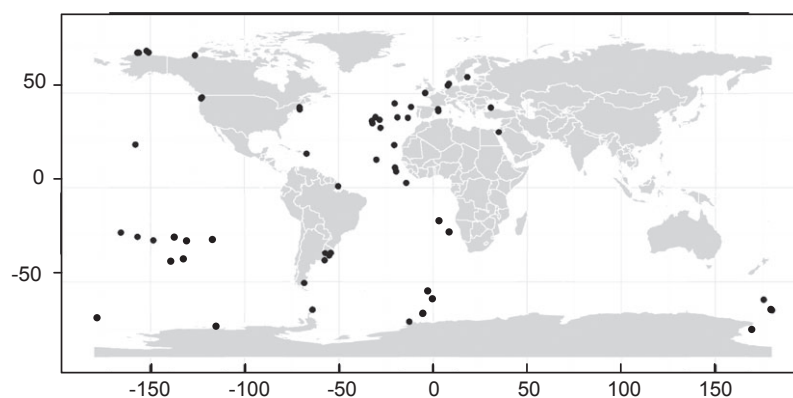

Figure 1 Locations of 84 unique sampling points of marine bacteria included in this analysis. Map projection is equirectangular.

For each OTU, we calculated the latitudinal range in three ways. First, we subtracted its southernmost latitude from the northernmost latitude (literal latitudinal range; an OTU detected at 70° N and 50° S would have a literal latitudinal range of 120°). The distribution of individuals within a range tends to be both dynamic and irregular (Brown, 1995), so while this measure reflects the uncorrected spatial extent of an OTU's occurrence, it does not account for occupancy or vacancy within that extent. Second, we calculated the difference between the maximum and minimum of the absolute values of latitudes where present (absolute latitudinal range); thus, an OTU occurring at 70° N, 50° N and 60° S would have an absolute latitudinal range of 20°. Under the general assumption that most abiotic conditions (such as temperature or day length) correlate with latitude, we interpret this measure as a rough reflection of an OTU's environmental breadth. Finally, we calculated the latitudinal range within each hemisphere (north and south latitudinal range). We interpret this last measure also as a reflection of an OTU's environmental breadth, but accounting for differences or asymmetries between hemispheres.

Definition of rarity and abundance types

To investigate how OTU abundance was distributed across its range, we classified OTUs into three rarity categories – rare endemic, abundant local and sparse – based on latitudinal range and relative sequence abundance minima and maxima. Abundant and widespread OTUs were classified in a separate category ('common'). The upper and lower limits of abundance and range for each of these categories were selected arbitrarily to produce approximately equal numbers of OTUs. These methods differ somewhat from those of Rabinowitz (1981), who used naturalist knowledge rather than numerical criteria to define rarity categories, and who also considered both within- and among-habitat effects.

Statistics and randomization procedures

Range restriction

We used a resampling procedure to test whether OTU ranges were more restricted than expected given random occurrence patterns. For each of the 100 subsamples generated above, the observed mean latitudinal range was compared to 100 randomizations in which sampling effort and OTU abundance counts were maintained but OTU locations were randomly assigned. Statistical significance for whether observed mean values of OTU ranges differed from null expectations were calculated via this bootstrapping procedure.

Northern versus southern ranges

To test for latitudinal complementary (whether location in one hemisphere predicts an OTU's location in the other), we calculated a northern and southern latitudinal centre for each

OTU. We then examined whether these centres were positively correlated across OTUs. We used Fisher's exact test to determine whether there were equivalent proportions of near-shore and offshore samples between hemispheres.

Abundance versus latitudinal range

Locally weighted scatterplot smoothing curves (LOESS) indicated that the relationship between log-transformed OTU abundance and literal latitudinal range were approximately linear. Thus, a linear regression model was implemented to evaluate these data. The smallest abundance values were rounded up to the nearest 0.1 for clearer visualization on a log-transformed scale. To determine whether the taxonomic breadth considered influenced the fit of this relationship, linear regression models were fitted for each distinct taxonomic group at decreasing breadth (within phylum, class, order, family or genus). At each resolution, taxonomic groups were removed from the analysis if they contained < 100 OTUs. Taxonomic ranks are nested by definition, precluding statistical comparisons among them here.

Range size versus latitude

To test Rapoport's rule using a method similar to that of Stevens (1989), we calculated the average latitudinal range of all OTUs present in each sample (Stevens tested the average within a latitudinal band) using each of the three latitudinal range metrics. We then tested the correlation between the mean range of the sample's OTUs versus the latitude of the sample. This method counts each OTU at every sample where present.

We also tested a statistic related to Rapoport's rule in a manner similar to that of Rohde *et al.* (1993), in which the average literal latitudinal range of all OTUs centred within a 5° latitudinal band are correlated with latitude. Here, we calculate the centre of each OTU in each hemisphere independently so that disjunct species found near the poles are not statistically centred near the equator. The means of OTU ranges in the Northern and Southern hemispheres were compared using a Mann–Whitney *U*-test.

In order to test whether the end-member groups – rare and abundant bacterial types – correlate with latitude, we calculated the proportion of each sample composed of the rarest OTUs (singletons: present as a single sequence in the data set) and the most abundant reads (the upper 99th quantile of the abundance distribution, averaging 484.2 ± 8.0 sequences in the data set per resample) and correlated this with the latitude of the sample.

Statistical packages

All data manipulations and statistics described above were performed using R 2.11.1 (R Development Core Team, 2010) and the VEGAN 1.17-6 (Oksanen *et al.*, 2011), FIELDS 6.3 (Furrer *et al.*, 2009), MAPTOOLS 0.7-38 (Lewin-Koh & Bivand,

2009), GGPLOT2 0.8.9 (Wickham, 2009) and SPAM 0.23-0 (Furrer & Sain, 2008) packages as well as custom scripts.

RESULTS

Rarity classes and distribution patterns of marine bacterioplankton

Three of the four rarity/abundance categories were parameterized to contain approximately equivalent numbers of OTUs (78–92). The rare endemic category contained many overlapping species, so 1836 OTUs were selected (85 were randomly selected for display in Fig. 2a). Occurrences of the rare endemic and common bacterial OTUs are distributed relatively evenly across latitudes (Fig. 2a; yellow and blue, respectively). However, the other categories display distinctly non-random patterning. The vast majority of abundant local OTUs were derived from a single sample taken in the Amazon river delta (Fig. 2b, purple). All other abundant local OTUs were from other coastal samples such as the Baltic Sea and Black Sea. Thus, rather than displaying a latitudinal pattern, the abundant locals appear to be coastal specialists, reflecting the habitat distribution of our samples. In contrast, sparse OTUs (Fig. 2a, green) were clearly aggregated in space, more often displaying 'hotspots' of abundance at higher latitudes compared to tropical latitudes. All of the rarity groups contained members of the seven most abundant phyla (Fig. S2 in Appendix S2).
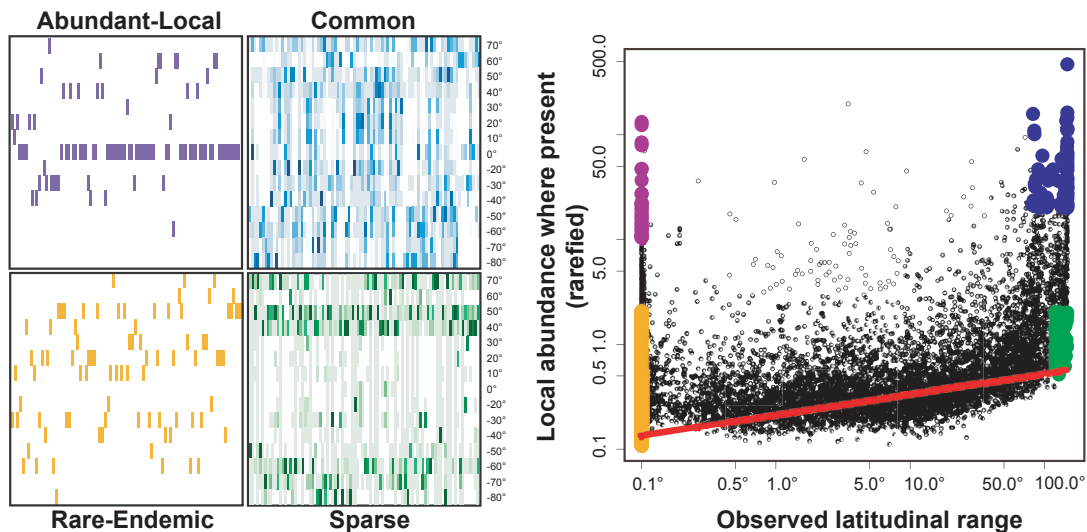
Marine bacterial OTUs as a whole were significantly more restricted in their range than would be expected by random occurrence. The literal latitudinal range expected from our randomization procedures was 31.7°, while that observed from the ICoMM data set was 13.45° (bootstrap $P = 0.01$). OTUs also appeared to be significantly limited in environmental breadth (expected absolute latitudinal range = 13.14°; observed absolute latitudinal range = 5.38°, bootstrap $P = 0.01$).

Latitudinal complementarity – the extent to which location in one hemisphere predicts location in the other – also appeared to be a pattern among the widest-ranging bacterial OTUs. Approximately 30% (3832 of 12,502) of non-singleton OTUs were detected in both Northern and Southern hemispheres. The northern and southern locations of these OTUs were positively correlated (found within a similar range of latitudes on both sides of the equator). Latitudinal complementarity is strongest when only the most wide-ranging OTUs (upper 99th percentile, $n = 124$, range > 142.6°) are included (Fig. 3a, $r^2 = 0.77$, $P < 0.0001$), and diminished as more narrowly distributed taxa were added (Fig. 3b; complete data set, $n = 3832$, $r^2 = 0.33$, $P < 0.0001$).

Abundance–range patterns in marine bacterioplankton

Similar to patterns found in macro-organisms, the relationship between the local abundance of a marine bacterial OTU
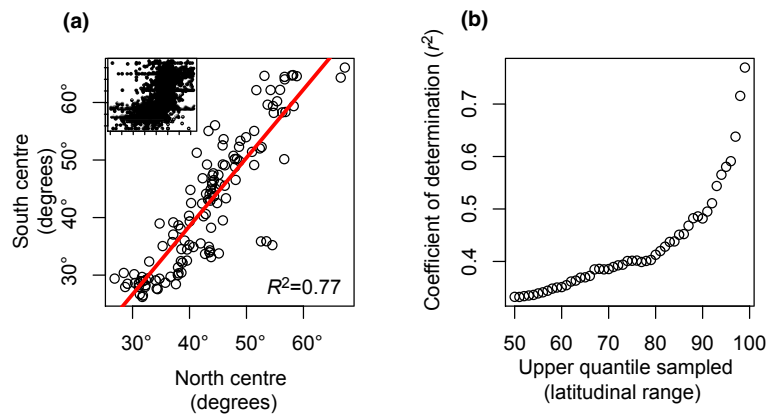

Figure 2 Distributions for operational taxonomic units (OTUs) of marine bacteria corresponding to the three rarity categories of Rabinowitz (1981) as well as common species as depicted in a heat map (a). Each column is an OTU, and darker colours represent greater proportional abundance. Colours correspond to OTUs highlighted with the same colours in the figure at right. Occurrences of rare-endemic and common OTUs are evenly distributed across latitudes. Sparse OTUs are at highest abundance at opposite and complementary latitudes > 50°. Abundant-local OTUs are found in coastal samples, particularly near the Amazon delta (near latitude 0°). For rare-endemic OTUs (yellow), 85 of 1836 were selected at random for display. OTUs were normalized by sampling depth, and by abundance within each 10° bin (rows); column sums are 1. The abundance range relationship (b) is positive ($r^2 = 0.114$, $P < 0.0001$) similar to that found in large-bodied organisms. Coloured points refer to OTUs classified under rarity types as in the figure at left. All values were arbitrarily rounded up to 0.1 for display.

Figure 3 Northern operational taxonomic unit (OTU) centre location versus southern centre location (weighted mean of latitude where present) of marine bacteria. (a) The correlation within the upper 99th quantile of the OTUs with the greatest latitudinal extent is high ($r^2 = 0.77$, $P < 0.0001$), showing that marine bacterial distributions are complementary in opposite hemispheres. Inset figure is all near-surface bacteria ($r^2 = 0.33$, $P < 0.0001$). (b) This complementarity is highest (as measured by a correlation coefficient) among species with the greatest latitudinal range, and decreases as less wide-ranging OTUs are included in the analysis.

and its latitudinal extent is significantly positive (log-transformed; $r^2 = 0.114$, $P < 0.0001$; Fig. 2b). Despite this trend, the relationship is not triangular when plotted on a log-transformed scale. Notable is the prominent group of abundant local OTUs (Fig. 2b, purple), although these are a small minority of all OTUs. In particular, with the Amazon plume sample removed, Fig. 2a assumes a more classical triangle shape.

Local abundance is tightly correlated with total abundance and relative abundance ($r^2 = 0.83$, $P < 0.0001$), and thus the shape of the abundance–range relationship is robust to the abundance metric employed (plots not shown). This relationship was also robust to resolution of taxonomic breadth considered, i.e. when the relationship was constrained to within-group correlations (within genus, family, order, class or phylum; Fig. S3 in Appendix S2). Median values of r^2 within groups varied from 0.06 at the phylum level to 0.16 at the genus level, seemingly trending upward as breadth was narrowed.

Abundance and range relationships along a latitudinal gradient

As observed for larger organisms, the range sizes of marine bacteria vary along a latitudinal gradient. The mean (literal) latitudinal range size of all non-singleton OTUs in each sample is strongly correlated with latitude, supporting the existence of Rapoport's pattern in marine bacterial communities ($r^2 = 0.60$, $P < 0.0001$; Fig. 4a). Rapoport's rule is typically assessed within a single hemisphere, because samples from both hemispheres are not available. Because the ICoMM data set includes samples from both hemispheres, the positive relationship that we found could be attributed to latitudinal complementarity of the OTUs. For instance, taxa that are found at the poles would have a very large range (across all latitudes), even if they were restricted to narrow absolute latitudinal ranges. To separate the effects of latitudinal

complementarity and latitudinal (and likely, environmental) breadth on this pattern, we also calculated the mean absolute latitude range of OTUs within samples (Fig. 4b). Whereas the correlation of absolute latitudinal range maintains a significant positive relationship ($r^2 = 0.14$, $P = 0.001$), the fit is much weaker than that of literal latitudinal range. Single-hemisphere correlations (Fig. 4c,d) show different patterns: highly significant in the Southern Hemisphere ($R^2 = 0.60$, $P < 0.0001$), but not significantly different from zero in the Northern Hemisphere.

Environmental changes can be abrupt along marine depth gradients. For instance, Stevens (1996), found that the Rapoport relationship in marine fishes diminished with water depth. Indeed, many of the outlier values from our Rapoport's analysis occur at the lowest depths. When we recalculated the range size–latitude relationship to include only those samples within the top 15 m of the water column, the fit (r^2) of all of the correlations improved (from $n = 257$ samples to $n = 196$; Fig. S4 in Appendix S2).

Whereas Stevens' method examines the impact of latitudinal clines on *community* composition, the 'midpoint' measurement of Rapoport's rule more explicitly examines the relationship between range size of *individual taxa* and latitude. Using the midpoint method, we find a significant though weak correlation with latitude ($r^2 = 0.03$, $P < 0.001$). Notably, there are distinct differences between tropical and temperate OTUs. Species centred in the tropics, even the most abundant, are seldom detected in polar regions (Fig. 5).

We also note some significant differences between the Northern and Southern Hemisphere patterns of marine bacterial ranges (Fig. 5). The mean ranges of OTUs in the Southern Hemisphere ($36.61° \pm 40.92$ SD) are more than twice those of OTUs in the Northern Hemisphere ($17.04° \pm 32.02$ SD); Mann–Whitney *U*-test: $P < 0.0001$.

Finally, not only did geographical range size vary by latitude, but so did the relative abundance of the rarest and
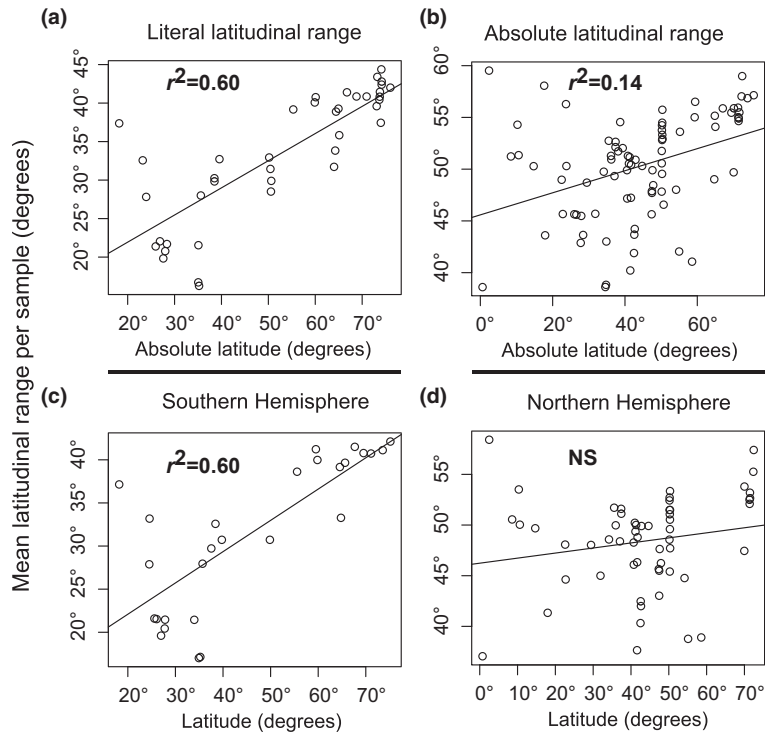
Figure 4 Rapoport's rule (mean latitudinal range size of all non-singletons per sample) is significant for marine bacterial communities when latitudinal range is calculated using literal, absolute and Southern Hemisphere calculations. The relationship is not significant when restricted to the Northern Hemisphere. Differences between literal and absolute range relationships indicate that complementarity is more deterministic of the literal Rapoport relationship than is environmental or niche breadth. All correlation coefficients increase as analyses are restricted to samples < 15 m in depth only (see Fig. S2 in Appendix S2).
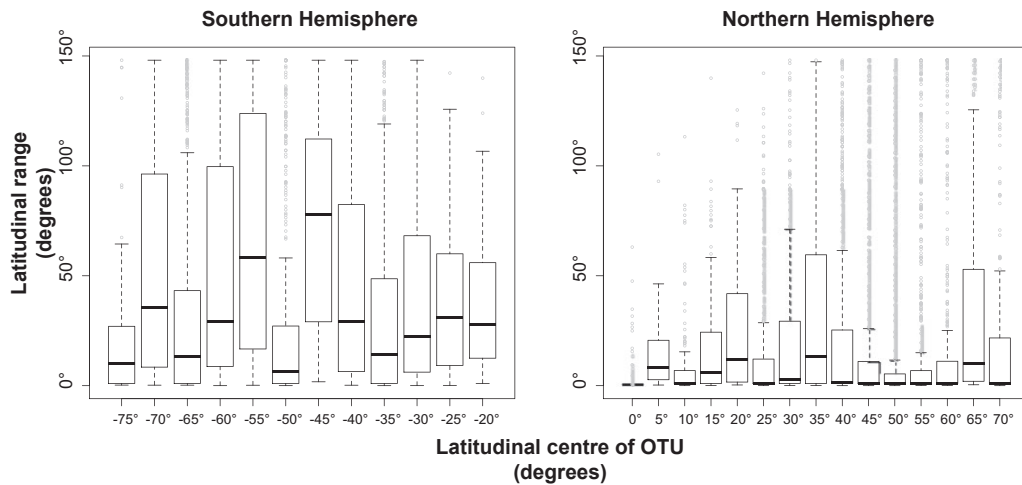

Figure 5 The relationship between the centre location of an operational taxonomic unit (OTU), and that OTU's latitudinal range plotted at 5° bands (the 'midpoint' method). Marine bacteria occurring in both Northern and Southern hemispheres are plotted twice in this figure. Boxes contain the upper and lower quartiles of latitudinal ranges for each 5° bin, and the median is displayed as the band within. Outliers (grey circles) are more than 1.5 times the interquartile range, and whiskers display the extent of non-outlier values. Latitudinal ranges are significantly greater in the Southern Hemisphere. OTUs centred in the tropics are seldom found in temperate regions.

most abundant OTUs. Having shown that communities contain more narrow-ranging bacterial OTUs near the equator, and that there was a positive relationship between abundance and latitudinal range size, we might expect that communities at lower latitudes also contain more rare taxa. Indeed, there

is a significant negative relationship between latitude and the proportion of the sample composed of sequences found only once in the data set (singletons; $r^2 = 0.05$, $P = 0.001$; Fig. 6a). Conversely, our expectation that communities near the poles contain more abundant taxa was correct: there is a
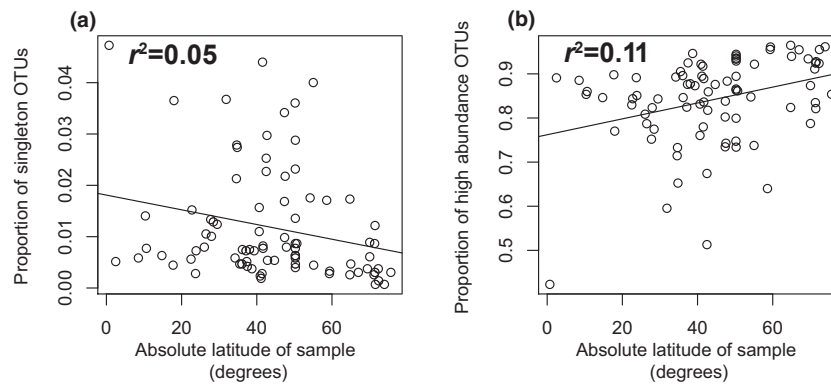
Figure 6 The least abundant operational taxonomic units (OTUs) comprise a greater proportion of samples at lower latitudes (a; $r^2 = 0.05$, $P < 0.0001$), whereas the most abundant OTUs comprise a greater proportion of samples at higher latitudes (b; $r^2 = 0.11$, $P < 0.0001$). For samples of marine bacteria with identical locations (lat and long) the mean value is calculated and presented here as a single point. Summary statistics of 100 such randomizations are reported here. Singleton OTUs were present in a single read; high abundance sequences were the upper 99th percentile (rarefied abundance $> 484.2 \pm 8.0$).

positive relationship between latitude and the proportion of samples composed of the most abundant taxa (upper 99th quantile; abundance $> 484.2 \pm 8.0$; $r^2 = 0.11$, $P = 0.001$; Fig. 6b). Removal of the Amazon plume sample (the apparent outlier near the equator) diminishes the coefficient (singletons: $r^2 = 0.03$; upper 99th quantile: $r^2 = 0.06$) but the trends are unchanged. Thus, we conclude that the frequency distribution of bacterioplankton communities reflected latitudinal clines.

DISCUSSION

Direct comparison of patterns of macro-organism and micro-organism diversity is challenging, as the scales of evolutionary history and taxonomic diversity vary greatly. For example, the domain Bacteria is estimated to be c. 3.5 billion years old, more than 20 times older than the class containing the entire diversity of Earth's birds. Nevertheless, we find here that many of the macroecological patterns of abundance and range observed in larger taxa are evident among marine bacteria in the upper level of the ocean.

Rarity classes and ranges of marine bacterioplankton

Like macro-organisms (Rabinowitz, 1981; Gaston, 1994), marine bacterioplankton display different forms of rarity. Recent studies of marine micro-organisms often highlight the large number of rare endemics (e.g. Sogin *et al.*, 2006), but other rarity categories may well deserve attention. Habitat heterogeneity among our samples clearly accounts for the prominence of abundant local taxa. Zinger *et al.* (2011) showed that bacterioplankton community composition is maximally heterogeneous near the coast, where environmental gradients can be abrupt. Most of our abundant local OTUs are found in association with the Amazon river plume, a unique and liminal environment. All of the other abundant local OTUs relate to near-shore habitats,

suggesting that these taxa may be specially adapted to idiosyncratic local conditions, which are themselves rare.

More intriguing, however, is the latitudinal distribution of sparse taxa, which appear to be more abundant near the poles. This pattern may be the result of how our categories were parameterized (a wide distribution requires a polar occurrence and a low abundance requires absence elsewhere). An analytical artefact does not, however, account for the 'hotspot' clumping at adjacent high latitudes, which suggests a biological basis for range limitation.

We found that marine bacterial OTUs, on the whole, are significantly restricted in their range, suggesting that OTUs are spatially 'clumped' with regard to latitude. As summarized by Hengeveld & Haeck (1982), a typical pattern among a wide range of organisms is for maximal abundance towards the centre of a species' range, and diminished abundance towards the edges. This pattern generally results in a single occurrence 'hotspot'. However, the generality of this pattern and the statistical rigour with which it has been tested has been contested (Sagarin & Gaines, 2002).

A variation on this 'hotspot' distribution for marine bacteria, particularly among the most widespread OTUs, is their tendency to form aggregations in opposite hemispheres at complementary latitudes. These 'bipolar' (Ekman, 1953) or 'antitropical' (Hubbs, 1952) distributions are known among marine organisms particularly when the distributions of genera or families are considered. The overall frequency of antitropicality at the species level, however, tends to be quite low. Such patterns are found among some species of coldwater algae (Peters & Breeman, 1992; Bischoff & Wiencke, 1995), invertebrates with pelagic dispersal mechanisms (Hilbish *et al.*, 2000), and a variety of other marine organisms including fish and mammals (Lindberg, 1991). Observations of terrestrial macro-organisms with antitropical distributions are rarer (Raven, 1963). Although comparable studies are generally lacking for terrestrial microbes, similar patterns were found in a global study of indoor fungi (Amend *et al.*,

2010b). Latitudinal complementarity in marine bacteria tends to manifest in temperate and polar areas, with diminished abundance (rather than true absence) in the tropics.

Abundance–range patterns in marine bacterioplankton

As in the majority of previous studies of plants and animals, marine bacteria show a positive relationship between taxon abundance and range size, encompassing a large amount of variance. Above, we described three mechanisms that might partly drive this pattern: niche breadth, metacommunity dynamics and ascertainment bias. While none of these mechanisms may be discounted outright, two lines of evidence suggest that the pattern is not solely attributable to under-sampling the rarest OTUs (ascertainment bias). First, despite correcting for both variation in sample size and bias in sampling locations, we find that range limitation is geographically patterned and more prevalent near the equator. Tropical species, even the most abundant, are not found in temperate waters. This geographical patterning points to either biological or geographical determinants as opposed to a systematic methodological bias. Second, we find that bacterial OTUs which are present in more than one location are restricted in their ranges compared to a random distribution: OTUs tend to have significantly clustered 'hotspots' of abundance. Thus, the abundance–range relationship is not just a matter of undersampling bacterial diversity.

The fit of a linear model to the abundance–range relationship appeared to improve with increasing taxonomic resolution, although our data are not amenable to assessing this trend statistically. Whereas a phylogenetically explicit examination of the abundance–range relationship is beyond the scope of this study, it has been suggested elsewhere that phylogenetic conservatism of abundance and range may be confounded by historical biogeographical determinants (such as dispersal events) and may, therefore, not be heritable in the same sense as other ecological traits (Webb & Gaston, 2003).

Abundance and range relationships along a latitudinal gradient

Marine bacteria appear to strongly follow Rapoport's rule. On average, tropical OTUs have smaller ranges than temperate OTUs. Further, OTUs centred in the tropics, even those that are most abundant, are almost entirely absent from samples at higher latitudes. In contrast, OTUs centred at higher latitudes display both narrow and broad distributions and are often found in tropical samples.

Rapoport's rule is often attributed in part to niche-based determinants; for example, climatic stability at low latitudes may select for specialist species with relatively narrow niche breadth (Klopfer & MacArthur, 1960). This explanation and others, such as diminished competition at higher latitudes, remain to be tested explicitly for marine bacteria. Alternative explanations offered for terrestrial systems such as decreasing

land area at lower latitudes, or an increase in extinctions due to glaciation at higher latitudes (Gaston et al., 1998) may also be important in marine environments.

Several other mechanisms that are not usually considered in larger organisms may also contribute to the Rapoport effect in bacterioplankton. In particular, complementarity at least partly drives this pattern, since high-latitude taxa with mirrored north–south distributions will have a greater latitudinal range than low-latitude taxa. This effect is evident in differences in explanatory power between literal and absolute or hemispheric latitudinal range measurements. Stevens originally attributed the Rapoport effect to clines in species' environmental tolerances or niche breadth (Stevens, 1989, 1996). The seasonal and diurnal environmental extremes endured by species at high latitudes, he argued, relates to the environmental variance that a species would endure over a large latitudinal range. In our study, environmental variance would best correlate with the absolute latitudinal range metric, which, although significant, explains relatively little variance.

A positive Rapoport's pattern may also be the consequence of sampling a latitudinal richness gradient of such a high-diversity group. Even when sampling intensity is normalized throughout the extent of a study, a Rapoport effect would manifest if community structures near the equator were less even than those at temperate latitudes (Fuhrman et al., 2008). This gradient would result in tropical species being less proportionally abundant in a community and less likely to be enumerated, thereby resulting in underestimates of range sizes (Colwell & Hurtt, 1994).

We also found unexpected differences between bacterial OTU ranges between the Northern and Southern hemispheres. OTUs in the Northern Hemisphere tend to have smaller ranges than those in the Southern Hemisphere, and the Rapoport relationship is stronger in the Southern Hemisphere. This latter result runs marginally contrary to the results of a meta-analysis of 246 studies testing for a Rapoport effect in larger organisms (Ruggiero & Werenkraut, 2007). The authors report significantly stronger correlations in studies of terrestrial Northern Hemisphere latitudinal gradients than in Southern Hemisphere when the pattern is measured using the Stevens method, although other methods failed to detect this difference. Presumably, in marine environments, we would expect the inverse of any land mass effects, and a greater Southern Hemisphere ocean surface area may contribute to our different results. In contrast, a study of Rapoport patterns among global parasitic microbial groups only found significant correlations in the Southern Hemisphere for bacteria and indirectly transmitted viruses, whereas Rapoport patterns were significant in both hemispheres for helminths, protozoans and fungi (Guernier & Guégan, 2009). In a study of marine molluscs (Fortes & Absalão, 2004), the authors found greater ranges among species in the Southern Hemisphere, and a concomitant steeper slope of the Rapoport pattern, although these differences were principally restricted to coastal Pacific species (no differences were found between northern and southern species on the Atlantic coast). The authors attributed this difference partly to

smaller northern biogeographical provinces, which restricts the range size of the species within them.

While biophysical asymmetry may account for differences between patterns in the Northern and Southern hemispheres, sampling asymmetry may play a role as well. The latitudinal distributions of samples were, however, roughly equivalent in the two hemispheres, and there were no significant differences between the proportions of near-shore and offshore samples.

CONCLUSIONS

The range and abundance relationships of marine bacterioplankton observed here both conform to and differ from those found in macro-organisms. Whereas some macroecological patterns such as Rapoport's rule and spatial autocorrelation of abundance extend to marine bacteria, important differences such as latitudinal complementarity also exist.

The similarities between macroecological *patterns* do not necessarily signify similarities between *processes*, and much work is still needed to understand the factors shaping marine bacterial abundance–range patterns and how these differ from those of macro-organisms. Examining the extent to which various evolutionary and ecological processes contribute to microbial abundance–range patterns will no doubt be a fruitful avenue for future research.

REFERENCES

Amaral-Zettler, L., Artigas, L.F., Baross, J., Loka Bharathi, P.A., Boetius, A., Chandramohan, D., Herndl, G., Kogure, K., Neal, P., Pedrós-Alió, C., Ramette, A., Schouten, S., Stal, L., Thessen, A., de Leeuw, J. & Sogin, M. (2010) A global census of marine microbes. *Life in the world's oceans: diversity, distribution, and abundance* (ed. by A.D. McIntyre), pp. 221–245. Wiley-Blackwell, Oxford.

Amend, A.S., Seifert, K.A. & Bruns, T.D. (2010a) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Molecular Ecology*, **19**, 5555–5565.

Amend, A.S., Seifert, K.A., Samson, R. & Bruns, T.D. (2010b) Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proceedings of the National Academy of Sciences USA*, **107**, 13748–13753.

Barberán, A., Bates, S.T., Casamayor, E.O. & Fierer, N. (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal*, **6**, 343–351.

Bell, T. (2010) Experimental tests of the bacterial distance–decay relationship. *The ISME Journal*, **4**, 1357–1365.

Bell, T., Ager, D., Song, J.-I., Newman, J.A., Thompson, I.P., Lilley, A.K. & van der Gast, C.J. (2005) Larger islands house more bacterial taxa. *Science*, **308**, 1884.

Bischoff, B. & Wiencke, C. (1995) Temperature ecotypes and biogeography of Acrosiphoniales (Chlorophyta) with Arctic–Antarctic disjunct and Arctic/cold-temperature distributions. *European Journal of Phycology*, **30**, 19–27.

Bock, C.E. (1987) Distribution–abundance relationships of some Arizona landbirds: a matter of scale? *Ecology*, **68**, 124–129.

Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The American Naturalist*, **124**, 255–279.

Brown, J.H. (1995) *Macroecology*. University of Chicago Press, Chicago, IL.

Colwell, R.K. & Hurtt, G.C. (1994) Nonbiological gradients in species richness and a spurious Rapoport effect. *The American Naturalist*, **144**, 570–595.

Curtis, T.P. & Sloan, W.T. (2005) Exploring microbial diversity – a vast below. *Science*, **309**, 1331–1333.

Darwin, C. (1859) *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life.* John Murray, London.

Ekman, S. (1953) *Zoogeography of the sea*. Sidgwick & Jackson, London.

Fortes, R.R. & Absalão, R.S. (2004) The applicability of Rapoport's rule to the marine molluscs of the Americas. *Journal of Biogeography*, **31**, 1909–1916.

Fuhrman, J.A., Steele, J.A., Hewson, I., Schwalbach, M.S., Brown, M.V., Green, J.L. & Brown, J.H. (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences USA*, **105**, 7774–7778.

Furrer, R. & Sain, S.R. (2008) spam: a sparse matrix R package with emphasis on MCMC methods for Gaussian Markov random fields. *Journal of Statistical Software*, **36**, 1–25.

Furrer, R., Nychka, D. & Sain, S. (2009) *Fields – tools for spatial data*. R Package Version 6.3. Available at: http://www.image.ucar.edu/GSP/Software/Fields/.

Gaston, K.J. (1994) *Rarity*. Chapman & Hall, London.

Gaston, K.J. & Lawton, J.H. (1990) Effects of scale and habitat on the relationship between regional distribution and local abundance. *Oikos*, **58**, 329–335.

Gaston, K.J., Blackburn, T.M. & Spicer, J.I. (1998) Rapoport's rule: time for an epitaph? *Trends in Ecology and Evolution*, **13**, 70–74.

Gaston, K.J., Blackburn, T.M., Greenwood, J.J.D., Gregory, R.D., Quinn, R.M. & Lawton, J.H. (2000) Abundance–occupancy relationships. *Journal of Applied Ecology*, **37** (Suppl. 1), 39–59.

Guernier, V. & Guégan, J.-F. (2009) May Rapoport's rule apply to human associated pathogens? *EcoHealth*, **6**, 509–521.

Hanski, I. (1982) Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos*, **38**, 210–221.

Hanski, I. (1991) Single-species metapopulation dynamics: concepts, models and observations. *Biological Journal of the Linnean Society*, **42**, 17–38.

Hengeveld, R. & Haeck, J. (1982) The distribution of abundance. I. Measurements. *Journal of Biogeography*, **9**, 303–316.

Hilbish, T.J., Mullinax, A., Dolven, S.I., Meyer, A., Koehn, R.K. & Rawson, P.D. (2000) Origin of the antitropical distribution pattern in marine mussels (*Mytilus* spp.): routes and timing of transequatorial migration. *Marine Biology*, **136**, 69–77.

Horner-Devine, M.C., Lage, M., Hughes, J.B. & Bohannan, B.J.M. (2004) A taxa–area relationship for bacteria. *Nature*, **432**, 750–753.

Hubbs, C.L. (1952) Antitropical distribution of fishes and other organisms. *Proceedings of the 7th Pacific Science Congress of the Pacific Science Association. Vol. III: Meteorology and Oceanography*, pp. 324–329. R.E. Owen, Wellington, New Zealand.

Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L. & Welch, D.M. (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology*, **8**, R143.

Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. & Sogin, M.L. (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*, **4**, e1000255.

Huse, S.M., Welch, D.M., Morrison, H.G. & Sogin, M.L. (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology*, **12**, 1889–1898.

Klopfer, P.H. & MacArthur, R.H. (1960) Niche size and faunal diversity. *The American Naturalist*, **94**, 293–300.

Kunin, V., Engelbrektson, A., Ochman, H. & Hugenholtz, P. (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*, **12**, 118–123.

Lewin-Koh, N.J. & Bivand, R. (2009) *Maptools: tools for reading and handling spatial objects*. R package version 0.7-38. Available at: http://cran.r-project.org/package=maptools.

Lindberg, D.R. (1991) Marine biotic interchange between the northern and southern hemispheres. *Paleobiology*, **17**, 308–324.

Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D. & Horner-Devine, M.C. (2011) Drivers of bacterial β-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences USA*, **108**, 7850–7854.

Nemergut, D.R., Costello, E.K., Hamady, M., Lozupone, C., Jiang, L., Schmidt, S.K., Fierer, N., Townsend, A.R., Cleveland, C.C., Stanish, L. & Knight, R. (2011) Global patterns in the biogeography of bacterial taxa. *Environmental Microbiology*, **13**, 135–144.

Oksanen, J., Blanchet, J.F., Kindt, R., Legendre, P., O'Hara, R.B. et al. (2011) *The vegan package*. R package version 1.17-6. Available at: http://cran.r-project.org/package=vegan.

Östman, Ö., Drakare, S., Kritzberg, E.S., Langenheder, S., Logue, J.B. & Lindström, E.S. (2010) Regional invariance among microbial communities. *Ecology Letters*, **13**, 118–127.

Peters, A.F. & Breeman, A.M. (1992) Temperature responses of disjunct temperate brown algae indicate long-distance dispersal of microthalli across the tropics. *Journal of Phycology*, **28**, 428–438.

Pommier, T., Canbäck, B., Riemann, L., Boström, K., Simu, K., Lundberg, P., Tunlid, A. & Hagström, Å. (2007) Global patterns of diversity and community structure in marine bacterioplankton. *Molecular Ecology*, **16**, 867–880.

R Development Core Team (2010) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rabinowitz, D. (1981) Seven forms of rarity. *The biological aspects of rare plant conservation* (ed. by H. Synge), pp. 205–217. Wiley, Chichester, UK.

Raven, P.H. (1963) Amphitropical relationships in the floras of North and South America. *Quarterly Review of Biology*, **38**, 151–177.

Rohde, K. (1996) Rapoport's Rule is a local phenomenon and cannot explain latitudinal gradients in species diversity. *Biodiversity Letters*, **3**, 10–13.

Rohde, K., Heap, M. & Heap, D. (1993) Rapoport's rule does not apply to marine teleosts and cannot explain latitudinal gradients in species richness. *The American Naturalist*, **142**, 1–16.

Ruggiero, A. & Werenkraut, V. (2007) One-dimensional analyses of Rapoport's rule reviewed through meta-analysis. *Global Ecology and Biogeography*, **16**, 401–414.

Sagarin, R.D. & Gaines, S.D. (2002) The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecology Letters*, **5**, 137–147.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M. & Herndl, G.J. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences USA*, **103**, 12115–12120.

Soininen, J. (2012) Macroecology of unicellular organisms – patterns and processes. *Environmental Microbiology Reports*, **4**, 10–22.

Stevens, G.C. (1989) The latitudinal gradient in geographical range: how so many species coexist in the tropics. *The American Naturalist*, **133**, 240–256.

Stevens, G.C. (1996) Extending Rapoport's rule to Pacific marine fishes. *Journal of Biogeography*, **23**, 149–154.

Webb, T.J. & Gaston, K.J. (2003) On the heritability of geographic range sizes. *The American Naturalist*, **161**, 553–566.

Wickham, H. (2009) *ggplot2: elegant graphics for data analysis.* Springer, New York.

Zinger, L., Amaral-Zettler, L.A., Fuhrman, J.A., Horner-Devine, M.C., Huse, S.M., Welch, D.B.M., Martiny, J.B.H., Sogin, M., Boetius, A. & Ramette, A. (2011) Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS ONE*, **6**, e24570.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Data table including names, locations and depths of samples included in this analysis.
Appendix S2 Additional figures describing sample distribution, taxonomic affinities, and the effects of taxonomic resolution and depth cutoffs on analyses.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

This paper was written in collaboration with the Beta-Diversity working group of the International Census of Marine Microbes, a group of researchers focused on community-level dynamics of ocean microbes on a global scale.

Author contributions: M.L.S., D.B.M.W. and S.M.H. conducted the sequencing and sequence analysis; A.S.A., T.A.O. and J.B.H.M. conceived the ideas and wrote the manuscript; A.S.A. and T.A.O. conducted analyses; all authors contributed critical feedback and editorial oversight.

Editor: John Lambshead