Environment drives high phylogenetic turnover among oceanic bacterial communities

Thomas Pommier1,*, Emmanuel J. P. Douzery2 and David Mouillot3,4

1Université de Lyon, INRA, CNRS, Université Lyon 1, Écologie Microbienne (UMR 5557, USC 1193), 43 Boulevard du 11 Novembre 1918, 69622 Villeurbanne, France
2UMR 5554 Institut des sciences de l’Évolution, and 3UMR 5119 ECOSYM, Université Montpellier 2, CNRS, Ifremer, IRD. Place E. Bataillon, 34095 Montpellier Cedex 5, France
3ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland 4811, Australia
*Author for correspondence (thomas.pommier@univ-lyon1.fr).

Although environmental filtering has been observed to influence the biodiversity patterns of marine bacterial communities, it was restricted to the regional scale and to the species level, leaving the main drivers unknown at large biogeographical scales and higher taxonomic levels. Bacterial communities with different species compositions may nevertheless share phylogenetic lineages, and phylogenetic turnover (PT) among those communities may be surprisingly low along any biogeographic or environmental gradient. Here, we investigated the relative influence of environmental filtering and geographical distance on the PT between marine bacterial communities living more than 8000 km apart in contrasted abiotic conditions. PT was high between communities and was more structured by local environmental factors than by geographical distance, suggesting the predominance of a lineage filtering process. Strong phenotype-environment mismatches observed in the ocean may surpass high connectivity between marine microbial communities.

Keywords: diversity; phylogenetic turnover; biogeography

1. INTRODUCTION

Species turnover patterns, the amount of difference in species communities across space, are central in the debate concerning what drives biodiversity across wide geographical scales. When local communities are structured according to environmental factors, e.g. resource availability, abiotic gradients or disturbance frequency, the concept of niche stands as a central explanation for species turnover [1]. Conversely, the neutral theory explicitly suggests ignoring differences among individuals, populations and species in response to local environments, and posits that dispersal limitation underlies species turnover between local communities: more spatially distant communities should contain more distinct species [2]. It has thus been of primary importance to disentangle the niche versus the neutral influence on large-scale biodiversity patterns by assessing the respective contribution of environmental versus geographical factors on species turnover [3].

While this issue has been investigated thoroughly in eukaryotes [4], it has been scarcely studied in microorganisms such as bacteria. Only three pioneering works have suggested that species turnover between bacterial communities was strongly driven by local environmental factors instead of geography [5–7]. However, such filtering effects of species by environmental factors have only been observed at limited geographical scales and is subject to the definition of bacterial species itself, i.e. operational taxonomic unit (OTU) mainly built from the 16S ribosomal RNA (rRNA), which remains a controversial topic among microbiologists [8,9]. Thus, two fundamental issues remain unexplored. First, the main drivers of biological turnover between bacterial communities remain unknown at large biogeographic scales [10]. Second, even if the environment drives turnover at the OTU level, this influence may decline and be outweighed by other factors (e.g. dispersal, historical contingencies) when expanding the phylogenetic scale towards higher taxonomic levels. Communities with totally different species compositions may still share a large amount of phylogenetic diversity and thus more biological similarity than previously considered [11].

To infer the main drivers of biodiversity on the Earth, currently used indices assessing the degree of species turnover among communities have been complemented by accurate measures of the corresponding phylogenetic turnover (PT). Recently, Bryant et al. [12] proposed such a measure and showed that acidobacterial lineages were not randomly distributed across an altitudinal gradient. However, PT among bacterial communities has never been investigated at large spatial scales where the influence of geography may be stronger than at regional scales. In the open ocean where communities withstand weak or no barriers to dispersal, we may expect a stronger influence of environmental filters than that of geography, according to the hypothesis that ‘everything is everywhere, but the environment selects’ (see Baas Becking, L. G. M. 1934 in [13]). We here examined how both environmental factors (niche filter) and geography (neutral hypothesis) drive PT between 25 marine bacterial communities sampled along a large-scale oceanic gradient. We relied on the Global Ocean Survey (GOS) dataset [14], which presents three main advantages: (i) to date, it represents the largest available record of bacterial communities sampled from various ecosystems that are linked by natural dispersion over more than 8000 km; (ii) it guarantees the reproducibility of data acquisition with the same sampling methodology and whole genome shotgun protocols, thus bypassing potential PCR biases that may overlook particular lineages [15]; and (iii) beyond a large geographical scale, it offers a wide range of environmental conditions from brackish to saline, from cold to warm, and from surface to deep waters. We therefore used this dataset to explore the link between PT, environmental factors and geography.

2. MATERIAL AND METHODS

16S rRNA gene sequences were retrieved from 25 sea water samples collected on the Sorcerer II as part of the GOS expedition [14]. We
Figure 1. Illustration of the phylogenetic turnover (PT) between bacterial communities. (a) Highest-likelihood phylogeny connecting 3228 sequences of 16S rDNA. Taxa unambiguously assigned to a given clade are interconnected by coloured branches (turquoise, Bacteroidetes edges; orange, γ-Proteobacteria; purple, α-Proteobacteria; black, remaining edges). Horizontal branch lengths are proportional to the amount of molecular change (scale bar inferred nucleotide substitution per site). The tree is the same as in panel (a) but displayed as a chronogram, i.e. branch lengths are proportional to relative time (scale bar inferred arbitrary time units; root = 100 units). (b) Focus on the highest (89%) PT between bacterial communities GS12, from the Chesapeake bay (red branches), and GS16, from the Gulf of Mexico (blue branches). Shared internal branches containing members of both communities are shown in green. Other branches are left in grey. The tree is the same as in panel (a). (c) Focus on the lowest (55%) PT between bacterial communities GS02, from the Gulf of Maine (red branches) and GS35, from the Galapagos Islands (blue branches). Tree plotting and colour options are the same as in panel (b).

acknowledge that environmental shotgun sequencing only discloses the most abundant phylotypes of local surface communities. After quality control of their alignments, a 3228-sequence maximum-likelihood tree was inferred from 1285 nucleotide sites, and then rendered ultrametric. The amount of PT between pairs of samples was estimated as the fraction of branch lengths that were unique (not shared) to each of the two microbial communities considered. The level of PT and the examination of their main drivers were investigated for the whole bacterial community and for two predominant taxonomic groups: the alpha (α) and gamma (γ) classes of Proteobacteria. Environmental distances were estimated using salinity, temperature, depth, and chlorophyll a measures which are assumed to be the main filters of marine bacterial diversity [10], and were estimated using the Gower distance that standardizes variables. Geographical distances between sampling locations were also calculated. The respective effect of geography and environment on PT between all pairs of microbial communities was assessed using multiple regressions on distance matrices. Additional details are available in the electronic supplementary material.

3. RESULTS AND DISCUSSION
High PT was observed between pairs of communities regardless of the class considered (for all classes: 53% < PT < 89%; for α-Proteobacteria: 37% < PT < 82% and for γ-Proteobacteria: 24% < PT < 83%). The two extreme values for this turnover are illustrated in figure 1 where communities shared either the lowest amount of branches (11% in (b)) or the highest amount of branches (47% in (c)) among all community pairs. At the OTU level, this range of PT was higher when compared with the PT measured in terrestrial systems for Acidobacteria (30–70%) and angiosperms (40–80%) [12]. This discrepancy suggests either that marine bacterial communities experience higher selective filters on lineages than do terrestrial bacteria (niche filtering), or that larger geographical distances between oceanic samples produced more phylogenetically distinct communities (dispersal limitation).

Overall, bacterial lineages were more structured by local environmental (i.e. temporal) factors than by geographical distances (table 1 and figure 2a). Surpassing the niche-filtering hypothesis, our results suggest the predominance of a lineage filtering process: bacterial communities experiencing different environments contain different lineages. This global trend was homogeneous among classes of Proteobacteria (figure 2b,c) although PT was poorly explained for α-Proteobacteria (2%). This particular class, including the SAR11 cluster and the genus Roseobacter, comprises the most common bacterial species in surface marine waters, which particularly thrive in worldwide oligotrophic offshore conditions [16]. However, complex patterns in the distributions of deep branches of SAR11 have also been shown previously at an open ocean site [17] and other environmental variables, such as water column mixing, is an important factor that is not very well accounted for in the GOS data.

Table 1. Results of multiple regressions on distance matrices using a permutation method with the matrix of pairwise PT (Phylosor index) between locations as the dependent matrix and pairwise geographical distance and environmental dissimilarity between locations as independent matrices. (Additional partial multiple regressions ($r^2$ without an independent matrix) allowed us to estimate the 'pure' effect of each independent matrix expressed as a percentage of the variation explained by the whole model. *$p$-value > 0.05, *$p$-value < 0.05, **$p$-value < 0.01.)

<table>
<thead>
<tr>
<th></th>
<th>$r^2$ of the total model</th>
<th>$r^2$ without an independent matrix</th>
<th>$r^2$ of the pure effect</th>
<th>% explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>all bacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>geographical</td>
<td>0.344**</td>
<td></td>
<td>0.007</td>
<td>2.06</td>
</tr>
<tr>
<td>distance</td>
<td></td>
<td>0.337**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>environmental</td>
<td></td>
<td>0.006n.s.</td>
<td>0.336</td>
<td>97.94</td>
</tr>
<tr>
<td>dissimilarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-Proteobacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>whole model</td>
<td>0.029n.s.</td>
<td></td>
<td>0.0005</td>
<td>2.45</td>
</tr>
<tr>
<td>geographical</td>
<td></td>
<td>0.029n.s.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>distance</td>
<td></td>
<td>0.0007n.s.</td>
<td>0.019</td>
<td>97.55</td>
</tr>
<tr>
<td>environmental</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dissimilarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ-Proteobacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>whole model</td>
<td>0.179*</td>
<td></td>
<td>0.053</td>
<td>30.03</td>
</tr>
<tr>
<td>geographical</td>
<td></td>
<td>0.125*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>distance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>environmental</td>
<td></td>
<td>0.013n.s.</td>
<td>0.166</td>
<td>69.97</td>
</tr>
<tr>
<td>dissimilarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conversely, many culturable marine bacteria belong to the \(\gamma\)-subclass of Proteobacteria [18], indicating a strong preference for nutrient rich conditions in the ocean (i.e. copiotrophy), and therefore a tendency to niche filtering. Also the \(\gamma\)-proteobacterial Vibronaceae has traditionally been regarded to be one central family of marine bacteria because of its strong interactions with other organisms, which are considered as niche specialized [18]. Accordingly, the PT for \(\gamma\)-Proteobacteria was explained by environmental distance but not by geographical distance. An alternative explanation for the discrepancy between the two groups could be a stronger mass effect for \(\alpha\)-Proteobacteria than for \(\gamma\)-Proteobacteria, i.e. the result of a higher, continuous supply of propagules from neighbouring areas [19], which would blur any community structure related to spatial and environmental factors.

With a production of \(8.10^{29}\) cells per year in marine surface waters the turnover of bacteria is estimated as ca 16 days [20]. Such a large amount of biomass flow at the surface layer implies marine bacteria may mask any spatial patterns at large scale. Indeed, ocean microbial communities may be considered as highly interconnected and thus largely invariant in their phylogenetic composition. Counter intuitively, we observed a high PT between bacterial communities suggesting that entire lineages are filtered out by environment. Under the hypothesis of niche conservatism [21], i.e. that tolerated environments are phylogenetically conserved and thus consistent within lineages, we may explain our results by strong phenotype-environment matches, i.e. a mechanism by which the phenotypes of immigrants are less likely to be suited to the local environment than the phenotypes of locally derived individuals. Consequently, the potential high connectivity between microbial communities of the sea, which would lead to invariance in community structure, may be substantially reduced by high post-dispersal mortality owing to vulnerability of small-size organisms to new environments. Hence, the actual biotic connectivity between bacterial oceanic communities and then their ecological dissimilarity would be more related to environmental match than to geographical distance. This pattern was already observed at the species level where genetic relationships among populations do not always scale with distance [22,23]. Finally, our study demonstrates that, at the level of microbial communities, the PT scales with environmental distance but not with geographical distance.

The observed pattern of niche filtration at the lineage level raises another key question: can the filtered lineages fulfill all core functions played by the marine bacterioplankton community? If all functions were fulfilled, niche filtration would not alter the overall ecosystem functioning. Conversely, ecosystems with no or low connectivity to the global ocean, associated with higher mortality of immigrating cells, would endure a strong decrease of functioning. As our study suggests that microbial communities of the open ocean are less similar between each other than previously thought [24], even if we consider the whole bacterial tree of life, the next challenge is certainly to investigate the functional diversity patterns that sustain ecosystem processes.


