



Experimentally induced habitat filtering in marine bacterial communities

Mikael Pontarp^{1,*}, Johanna Sjöstedt², Per Lundberg¹

¹Theoretical Population Ecology and Evolution Group, Department of Biology, Ecology Building, Lund University, Lund 22362, Sweden

²Department of Natural Science, Linnaeus University, Kalmar 39182, Sweden

ABSTRACT: We tested the habitat filtering hypothesis by measuring the phylogenetic structure in marine bacterial communities before and after experimentally induced stress. The habitat filtering hypothesis predicts that phylogenetic clustering (mean relatedness) should increase as the environment becomes suitable for only a subset of the original community. We show that community composition and phylogenetic structure were considerably changed with changes in salinity and dissolved organic carbon. Community composition showed no consistent patterns, while the phylogenetic relatedness between species consistently increased with treatment. We have no information about species interactions in our system, but the phylogenetic signal is strong enough to suggest that habitat filtering is the dominant assembly process. Our results support the hypothesis that habitat characteristics and environmental stress can 'filter' a community so that only closely related species can persist. This non-random phylogenetic signal also implies a relationship between ecologically relevant characteristics and species relatedness.

KEY WORDS: Habitat filtering · Phylogenetic structure · Niche conservatism · Microbial ecology

—Resale or republication not permitted without written consent of the publisher—

INTRODUCTION

Natural communities are dictated by ecological and evolutionary assembly processes, historical contingencies and biogeography (Ricklefs & Schluter 1993, Losos & Schluter 2000, Hubbell 2001, Leibold et al. 2004, Ackerly et al. 2006, Urban & Skelly 2006, Wiens et al. 2011). These processes are manifested in signals of community structure, such as species diversity, species abundance and species relatedness (Vellend 2010). The phylogenetic signal, in particular, has received considerable attention (Webb et al. 2002, Emerson & Gillespie 2008, Cavender-Bares et al. 2009, Vamosi et al. 2009, Pausas & Verdu 2010). A local community that is assembled non-randomly from closely related species (with a random assembly from the available regional or global species pool as a reference) is said to be 'clustered'. In contrast, the members of an 'overdispersed' community are less

related to each other than expected by chance. Based on the assumption of a mapping between ecological characteristics and relatedness between species, these patterns have been interpreted in 2 ways. Habitat filtering implies that the environment has properties which ensure that only related species with a certain set of similar traits persist, leading to phylogenetic clustering. Competitive exclusion, in contrast, implies that closely related species that also share similar traits have limited room for coexistence, and therefore, only more distantly related species can co-occur locally.

Several studies on various organisms have inferred possible assembly processes from the observed phylogenetic structure (reviewed by Vamosi et al. 2009). This approach is especially rewarding when studying microbial communities where the phylogenetic resolution and coverage is often high and other methods are lacking. Recent studies have used this

*Email: mikael.pontarp@biol.lu.se

approach on bacterial communities and suggested that habitat filtering is an important assembly process in these systems (Horner-Devine & Bohannan 2006, Newton et al. 2007, Pontarp et al. 2012). These conclusions are also supported by studies showing that closely related microbial taxa are more ecologically similar than distantly related taxa (Nübel et al. 1999) and that different bacterial taxa dominate in different environments (Pommier et al. 2007, Schatzenhofer et al. 2009). There are, however, limitations to this approach, especially when applied to microbial communities. Results are dependent on the taxonomic, temporal and spatial scales (e.g. Silvertown et al. 2001, Cavender-Bares et al. 2006, Kraft et al. 2007, Swenson 2009, Leibold et al. 2010, Ricklefs & Jenkins 2011). There is no formal theory for the exact nature of the relationship between ecologically relevant traits (shared traits and trait values) and relatedness, but such a relationship is often implicitly assumed (Wiens et al. 2010). A formal theory is also lacking for understanding how the phylogenetic signal is created by the interplay between ecology and evolution (Pontarp et al. 2012b), and there are, to our knowledge, no empirical studies that have experimentally shown the link between process and patterns in large natural communities.

Here, we used an available data set (Sjöstedt et al. 2012) of marine bacterial communities to test the ‘habitat filtering’ hypothesis. Natural communities were sampled in the Baltic Sea and Skagerrak and subjected to experimental environmental stress by altered salinity and nutrient level concentrations in a chemostat setup. Salinity is an important factor forming the bacterial community (Glöckner et al. 2000), and different bacterial groups specialize in either low or high molecular weight DOC (Covert & Moran 2001). Specific genera dominate at different nutrient concentrations (Pinhassi & Berman 2003), and correlation between the evolutionary relationship and the similarity of the habitat has been shown (von Mering et al. 2007). Consequently, we hypothesized that the treatment would ‘filter’ the naturally sampled community to contain a subset of species that share ecological traits that are suitable for the confinement of the chemostats and the change in salinity and nutrient concentrations. Given a mapping between ecological characters and relatedness, suggested by e.g. Nübel et al. (1999), Horner-Devine & Bohannan (2006), Newton et al. (2007), Boyd et al. (2010), and Pontarp et al. (2012), we also predicted an increase in the degree of community clustering with environmental stress. Further, we predicted that the different stressors have different effects on species from

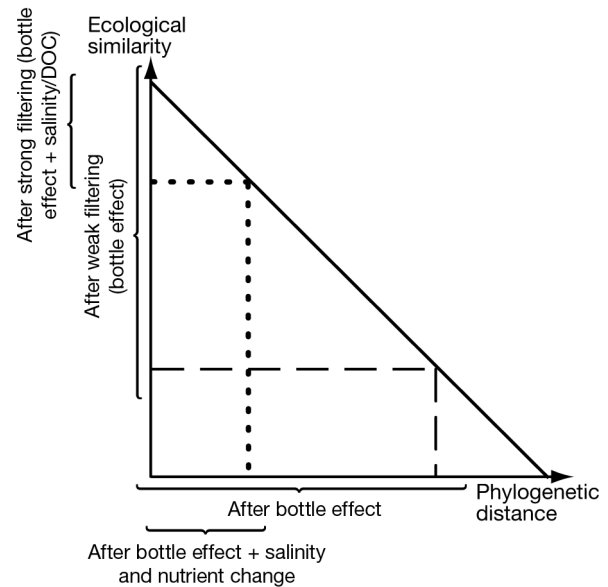


Fig. 1. Hypothetical illustration of our predictions for the community response to different stressors. We predict treatment to ‘filter’ species with similar ecological characteristics. Given a mapping between ecological characteristics and phylogenetic similarity, illustrated by the diagonal line, filtering should result in a community with higher phylogenetic relatedness (higher clustering) than the naturally sampled community. We predict chemostat effects to be less than, and additive to, the effects of salinity and nutrient stressors. Predicted causalities between filtering and phylogenetic patterns are illustrated by long dashed lines (bottle effects only) and short-dashed lines (bottle effects + changes in salinity and nutrient concentrations)

different parts of the phylogeny. Consequently, the clustering effect should be larger in the transplant treatments (chemostat effects + changed salinity and nutrient concentration) than in the control treatment (chemostat effects only). Fig. 1 presents a hypothetical illustration of our predictions.

MATERIALS AND METHODS

We did an extensive reanalysis of a data set previously produced by Sjöstedt et al. (2012), in which natural bacterial communities, sampled in the Baltic Sea and Skagerrak, Sweden, were inoculated into chemostats and exposed to habitat change by means of a gradual medium transplant (the Baltic community was exposed to Skagerrak medium and vice versa). Three main stress factors could potentially affect the communities under these treatments: bottle effects from the chemostat itself (Ferguson et al. 1984), a change in salinity (Bouvier & del Giorgio 2002, Langenheder et al. 2003, Crump et al. 2004) and a change in nutrient composition and concentra-

tion (Cottrell & Kirchman 2000, Covert & Moran 2001). We investigated the effect of this habitat change on the phylogenetic structure and community composition of the communities by analyzing 16S rRNA gene data before and after the transplant.

Sampling and experimental setup

Sampling was conducted on 2 occasions (June and August) in 2007 in both the Baltic Sea and the Skagerrak (referred to as exp1 and exp2 by Sjöstedt et al. 2012). Baltic and Skagerrak seawater was used as a medium after it was filtered (3.0 and 0.2 μm) and autoclaved. The inocula were also filtered (0.6 μm) to ensure that no Protozoa were present (see Sjöstedt et al. 2012 for details about the sampling procedure and preparation). Each sample was inoculated into a chemostat system and run at 20°C and in a 12 h light:12 h dark cycle. The transplants started with an initial (first 6 d) inflow of medium collected at the same sample sites as the inoculated community. Thereafter, over a period of 11 d, the Baltic Sea communities experienced a gradual change to the medium collected in the Skagerrak (referred to as BAL-T). The Skagerrak community experienced a similar transplant from the Skagerrak to the Baltic medium (referred to as SKA-T). After the transplant, the communities were exposed for another 5 d to a constant inflow of fully transplanted medium. Parallel to the transplant experimental setup, 1 community each from the Baltic and Skagerrak (both sampled in August) was treated with a constant inflow of non-transplanted medium (referred to as BAL-C and SKA-C). In total, 10 communities were retrieved from Sjöstedt et al. (2012): 2 sampled communities each from the Baltic Sea and Skagerrak (BALexp1, BALexp2, SKAexp1 and SKAexp2), the transplanted communities (BAL-T exp1, BAL-T exp2, SKA-T exp1 and SKA-T exp2) and the controls (BAL-C exp2 and SKA-C exp2). For simplicity, we exclude the abbreviation 'exp' from the names and only use BAL1, BAL2, SKA1 and SKA2 for the naturally sampled communities. Also, to emphasize that we are studying the change in community composition and community structure as a result of the water transplant, in the present paper, we refer to the 2 Baltic treatments as BAL→SKA1 and BAL→SKA2. Similarly, we refer to the 2 Skagerrak treatments as SKA→BAL1 and SKA→BAL2. The 2 communities, 1 Baltic and 1 Skagerrak community, that were treated with constant environments follow the Sjöstedt et al. (2012) naming method and are referred to as BAL-C 2 and SKA-C 2.

DNA extraction, PCR, clone libraries and 454-pyrosequencing

DNA was extracted using an enzyme/phenol-chloroform protocol.

Clone libraries and 454-sequencing of 16S rRNA genes were used to compare community composition and structure before and after chemostat treatment. One clone library was constructed from each of the 10 communities. 454-pyrosequencing was performed from the Baltic inocula sampled in August and its transplanted counterpart (equivalent to BAL→SKA2).

Sequence analysis

The PHRED base-calling software (Ewing & Green 1998, Ewing et al. 1998) was used on the Sanger data, resulting in 960 sequences (Sjöstedt et al. 2012). Reads <200 bp long and reads with undetermined nucleotides were removed from the 454 data set (initially containing 16 793 sequences). Primer and adaptor sequences were trimmed using the SeqClean software (<http://compbio.dfci.harvard.edu/tgi/software/>) with UniVec_Core as database (www.ncbi.nlm.nih.gov/VecScreen/UniVec.html). Sanger sequences have been deposited in the EMBL database under accession numbers FR647485 to FR648326, while pyrotag sequences have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive under accession number SRA012481.2.

The Greengenes on-line tool (<http://greengenes.lbl.gov>) and ChimeraSlayer were used to remove chimeric sequences, align sequences and assign a taxon for each sequence (DeSantis et al. 2006, Haas et al. 2011, Sjöstedt et al. 2012). Greengenes currently supports 3 different taxonomies: Ribosomal Database Project, NCBI and Hugenholtz. If ≥ 2 of these taxonomies showed the same taxonomic lineage at the phylum/class level to a particular sequence, the sequence was kept for further analysis. Sequences shorter than 750 nucleotides were removed from the Sanger data set. In total, 848 Sanger sequences and 16 513 454-sequences were used in the analysis. Maximum likelihood trees were inferred on the full Sanger data sets (all sequences irrespective of phylum or library) and on data subsets defined by taxonomic lineage at the phylum level on Sanger data and 454 data, respectively. Because of the large size of the 454 data set, phylogenetic trees were inferred on data subsets (phylum level) only.

RAxML software (version 7.0.4) (GTRMIX model and default settings) was used to construct maximum likelihood trees (Stamatakis 2006). The output from RAxML was used as input to the RAMI software, which clusters sequences based on their patristic distance (branch lengths) in the phylogenetic tree into operational taxonomic units (OTUs) (Pommier et al. 2009). The produced clusters were here viewed as OTUs, a proxy for bacterial species, and sequences within OTUs were considered as identical and/or unique 'individuals'. We did 3 separate analyses in which we used a patristic distance threshold of 0.06, 0.04 and 0.02, respectively, as the similarity cut-off. To reduce potential sequencing errors commonly associated with pyrosequencing (Kunin et al. 2010), we chose to do the analysis on the 454 libraries including OTUs containing >1 sequence (singletons were removed). Finally, consensus sequences for each of the OTUs were constructed with the RAMI software, and phylogenetic trees, based on the consensus sequences, were inferred with RAxML as above.

Community/phylogenetic patterns

The procedure of first classifying each sequence, dividing the data into subsets defined by phyla and then clustering the data subsets separately ensured that each OTU output from RAMI was a distinct entity, from here on referred to as a 'species'. These RAMI clusters were, however, composed of sequences potentially originating from different data libraries. By keeping track of the origin of each sequence in each OTU, we were able to define species that are represented in one or several libraries with different corresponding relative abundances. This information was used to infer community similarity based on species abundances with the Morisita index of similarity (Krebs 1999). The distinct species and the numbers of sequences associated with it were also used when analyzing a variety of community composition metrics: the number of species and mean species abundance was calculated, Margalef's index was used for species richness analysis, and Simpson's index was calculated for dominance/evenness analysis (Krebs 1999, Magurran 2004). Phylogenetic diversity, defined as mean pair-wise phylogenetic distance (MPD), was calculated for each of the communities in the Sanger data set (Webb 2000, Swenson 2009).

The phylogenetic structure, in terms of the degree of phylogenetic clustering (mean relatedness), was analyzed for the different treatments both on the full Sanger data set (including all phyla) and on the

Sanger- and 454-subsets (subdivided at the phylum level). This results in a value of the degree of clustering for each of the 10 Sanger libraries (including all phyla or for each phylum separately) in relation to the full Sanger data set, which was viewed as the species pool. We also generated a value for the degree of clustering per phylum for each of the two 454 libraries. It has been suggested that the degree of phylogenetic structure is dependent on the taxonomic level (Silvertown et al. 2001). Consequently, as we know little about the taxonomic level at which the phylogenetic signal of marine bacteria is really meaningful, we did the analysis based on single sequences (terminal leaves in the input phylogenetic tree as single sequences) and OTUs (terminal leaves in the input phylogeny as consensus sequences defining the species). The phylogenetic structure of the communities was analyzed with the *comstruct* tool implemented in the Phylocom software (Webb et al. 2008). This tool calculates the net relatedness index (NRI), which quantifies the structure of a sample phylogeny derived from the mean phylogenetic distance, consequently capturing the degree of clustering of the phylogeny from root to terminal leaves. In addition, the nearest taxa index (NTI) quantifies the terminal structure of the sample phylogeny, hence capturing the clustering of the terminal nodes in the tree. The Phylocom null model used here (#2) constructs a random phylogenetic tree by assigning sequences/OTUs to the localities (localities found in the focal phylogeny) by random draws from the phylogeny pool (all OTUs available in the input data). This model maintains the sequence richness of each sample, but the identities of the species occurring in each sample are randomized. Positive NRI and NTI values indicate a clustered phylogeny in which coexisting taxa are assembled non-randomly from a restricted, in terms of relatedness, part of the available species pool. A negative NRI and NTI value, in contrast, indicates overdispersion such that coexisting taxa within a phylogeny are less related to each other than would be expected by chance. The statistical significance of our results was evaluated by the use of a 2-tailed significance test for the NRI and NTI results based on the rank high and rank low values in the Phylocom output (Webb et al. 2008). These rank values are the number of runs where the null model has lower or higher NRI and NTI values than the focal phylogeny. Positive NRI and NTI values from the 1000 permutations associated with rank low values >975 and rank high values <25 and negative NRI and NTI associated with rank low values <25 and rank high values >975 were significant at $p < 0.05$.

RESULTS

Community composition

Here, we present results produced by analyses made at a RAMI threshold value of 0.04; additional results from analysis, including those using threshold values of 0.02 and 0.06, can be seen in the supplementary material (www.int-res.com/articles/suppl/m477p077_supp.pdf). Pair-wise phylogenetic distance analysis showed that the bacterial communities differed phylogenetically among all libraries, including the different transplants, controls and the 2 natural non-treated communities that were sampled on different occasions (Sjöstedt et al. 2012). These differences were confirmed by our results on community species composition, which showed Morisita's index values to be low (Table S1 in the supplement). No pattern was found in relation to treatment, geography or temporal sampling occasion in the community similarity analysis. Further, we could not find any consistent changes in community evenness with experimental treatment from the Simpson's evenness values, calculated from the species data (Table 1; see Table S2 in the supplement). The analysis of mean species abundance between treatments gave no conclusive results. The mean abundance values ranged from 2 to 4 species, and the standard deviation around each mean value was large, ranging from 1.2 to 8.1. The patterns in species richness (Margalef's species richness index) and MPD largely followed the patterns seen for the number of OTUs (Table 1). As shown by Sjöstedt et al. (2012), the number of species in the naturally sampled communities was larger in the Baltic than in the Skagerrak.

Phylogenetic structure

Similar patterns were found when analyzing the phylogenetic structure for the Sanger data on OTU level and sequence data (apart from generally larger values for the sequence-level analysis compared to the OTU-level analysis). The relative differences in NRI and NTI values between communities were similar for the different approaches. Similar results have also been shown in a previous study on globally distributed natural marine bacterial communities (Pontarp et al. 2012). Consequently, we choose to present the results from our OTU level analysis only, as they are comparable to previously published phylogenetic structure analyses.

When analyzing the phylogenetic structure of each naturally sampled community (non-transplanted) including all phyla, BAL1 was found to have a significantly negative NRI value, and BAL1, SKA1 and SKA2 were found to have significantly positive NTI values in the Sanger data (Table 2). Four phyla (*Alphaproteobacteria* [Alpha], *Bacteroidetes* [Bacteroid], *Betaproteobacteria* [Beta] and *Gammaproteobacteria* [Gamma]) were common in a majority of the 4 naturally sampled communities. In addition to these 4, an additional 5 phyla were found (*Actinobacteria* [Actino], *Cyanobacteria* [Cyano], *Deltaproteobacteria* [Delta], *Firmicutes* [Firmi] and *Verrucomicrobia* [Verruco]) in the two 454 communities. When analyzing the structure of each phylum in the non-transplanted communities, 3 and 7 of the possible 15 library/phylum combinations showed significant positive NRI and NTI values, respectively. No significant negative values were found at the phylum level for the Sanger data (Table 2). For the 454 analysis, 2 and 4 of the 9 possible library/phylum combinations showed significantly positive clustering NRI and NTI values, respectively (Table 2). See Table S3 in the

Table 1. Number of operational taxonomic units, OTU (species), Simpson's evenness values and Margalef's index of species richness for the Sanger and 454 (non-singletons) libraries, respectively; mean pair-wise phylogenetic distance (MPD) for Sanger libraries. Analysis made on OTUs constructed with a RAMI software threshold value = 0.04. Note that MPD values are unavailable for 454-data; the data size exceeded the limit for the methodological approach

	No. of OTUs	Evenness	Richness	MPD
Sanger data				
BAL→SKA1				
Start	39	0.67	9.36	1.26
Transferred	34	0.29	7.77	0.81
BAL→SKA2				
Start	42	0.39	9.09	1.11
Transferred	23	0.61	5.14	0.66
SKA→BAL1				
Start	29	0.18	6.21	1.05
Transferred	34	0.42	7.30	1.18
SKA→BAL2				
Start	25	0.94	6.81	1.23
Transferred	26	0.39	6.08	0.86
Controls				
BAL-C 2	21	0.22	4.49	1.13
SKA-C 2	24	0.21	5.10	0.81
454 data				
BAL→SKA2				
BAL 454	275	0.05	30.35	–
BAL-T 454	69	0.09	7.58	–

Table 2. The degree of clustering (net relatedness index [NRI] and nearest taxa index [NTI]) for the non-treated bacterial communities sampled in the Baltic Sea and Skagerrak. Positive and negative values indicate clustered and overdispersed communities, respectively. n: number of OTUs in the sample. NRI_gt and NTI_gt: number of random permutations (999 in total) that resulted in higher values than the analyzed sample. RAMI value for clustering = 0.04. *p < 0.1; **p < 0.05

Sample	n	NRI	NRI_gt	NTI	NTI_gt
Sanger data					
All bacteria					
BAL1	42	-1.96**	978	2.88**	3
BAL2	42	0.47	322	-0.25	611
SKA1	30	1.20	114	1.98**	13
SKA2	27	-1.19	875	4.56**	0
Alphaproteobacteria					
BAL1	9	1.08	135	2.14**	7
BAL2	10	-0.71	779	1.17	114
SKA1	12	-1.27	892	1.52	54
SKA2	7	-0.72	781	2.50**	1
Bacteroidetes					
BAL1	12	0.69	260	1.74*	36
BAL2	12	-0.62	725	-0.44	672
SKA1	4	1.25	87	1.60*	40
SKA2	6	1.79**	14	2.01**	9
Betaproteobacteria					
BAL1	3	0.11	696	-0.41	803
BAL2	2	0.25	521	0.25	521
SKA2	3	0.88*	37	1.25**	16
Gammaproteobacteria					
BAL1	4	-0.92	829	-0.45	702
BAL2	6	-1.27	868	-1.62	926
SKA1	5	1.15	74	1.00	147
SKA2	5	1.51**	20	1.51*	35
454 data					
BAL2					
<i>Actinobacteria</i>	42	5.43*	27	5.53*	27
<i>Alphaproteobacteria</i>	42	0.56	255	0.91	217
<i>Bacteroidetes</i>	51	1.40	103	2.37**	25
<i>Betaproteobacteria</i>	25	-0.06	520	1.04	171
<i>Cyanobacteria</i>	18	4.06*	43	3.44*	43
<i>Deltaproteobacteria</i>	17	2.14	96	2.01*	42
<i>Firmicutes</i>	22	0.69	97	0.74	97
<i>Gammaproteobacteria</i>	44	-2.06	987	0.03	479
<i>Verrucomicrobia</i>	14	3.37	73	2.89	73

supplement for results using RAMI values of 0.02 and 0.06.

After treatment in the chemostat, 82% of the treated communities, irrespective of whether they were Baltic transplants, Skagerrak transplants or control treatments, had a higher degree (measured as the difference in NRI and NTI values between samples) of clustering compared to the original community in the Sanger data (Fig. 2; see Fig. S1 in the

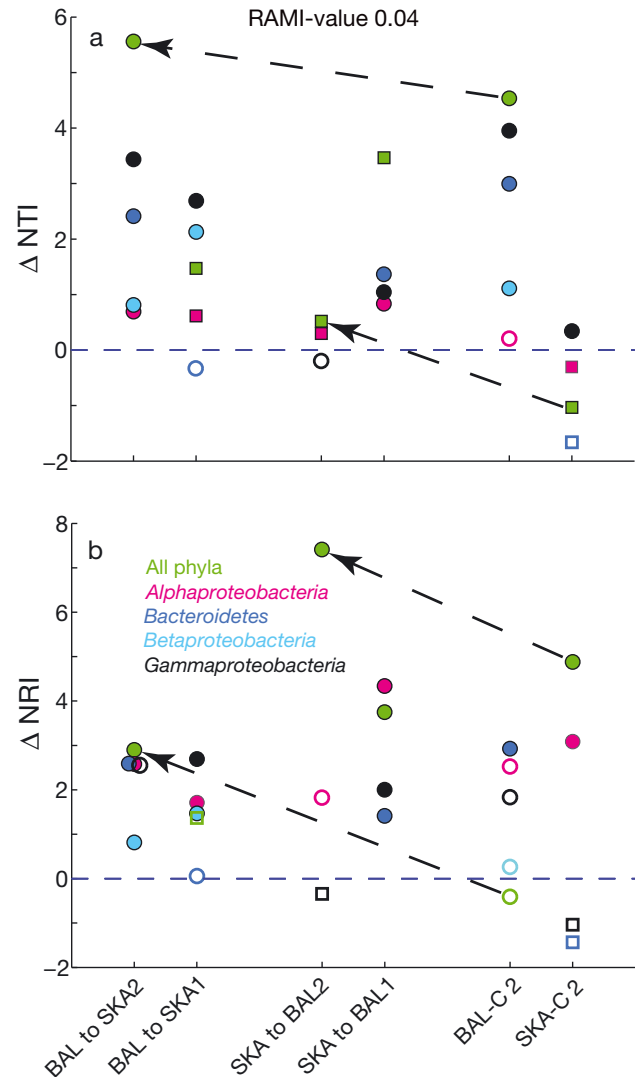


Fig. 2. Change in (a) NTI and (b) NRI for the different Sanger-data treatments, measured as the difference in NTI and NRI values between a natural and its corresponding transplanted community. BAL to SKA1: Baltic community (sampled in June) transplanted into Skagerrak water; BAL to SKA2: Baltic community (sampled in August) transplanted into Skagerrak water; SKA to BAL1: Skagerrak community (June) transplanted into Baltic water; SKA to BAL2: Skagerrak community (August) transplanted into Baltic water; BAL-C 2: Baltic community (August), no water transplant; SKA-C 2: Skagerrak community (August), no water transplant. Color coding denotes analyses done on data subsets defined by OTUs (species) that were assigned to the same phylum (except green, which denotes all phyla). Filled squares: significant values for both the naturally sampled (untreated) and treated (transplant and control) communities; filled circles: significant values for the treated community only; open squares: significant values for the natural community only; open circles: no significant structuring. Dashed arrows: NRI and NTI differences between transplanted (bottle effect + changes in salinity and DOC concentrations) and control treatments (bottle effects) for data including all phyla

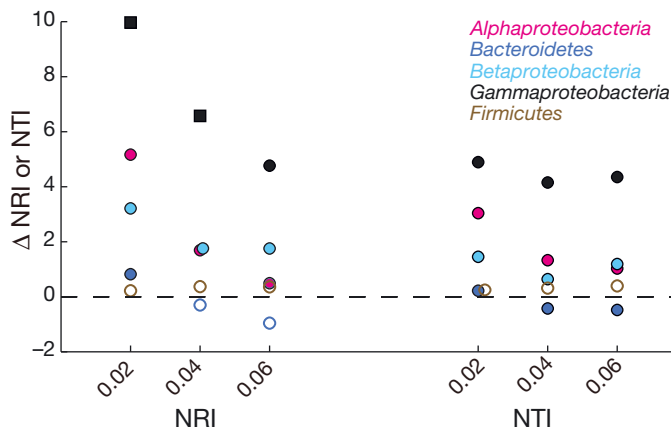


Fig. 3. The change in NTI and NRI (measured as differences) for the Baltic community (sampled in August) transplanted into Skagerrak water, measured on 454 data. Color coding denotes phylum subsets. See Fig. 2 for further explanation of symbols. Analyses were done for RAMI threshold values 0.02, 0.04 and 0.06

supplement). In addition, several of the communities that showed no significant structure before treatment became significantly clustered after treatment. Note also that these results were robust over the full range of phyla as well as for the full Sanger data set including all phyla. The same pattern was shown when analyzing the phylogenetic structure of the Baltic community, before and after treatment, based on the 454-sequencing results (Fig. 3). Note also that these results were robust over different RAMI threshold values. This was especially evident in the high-resolution 454 data, where we also found a decrease in clustering with increased RAMI value (Fig. 3). Finally, clustering, in terms of NRI and NTI values, was greater in both transplanted treatments (BAL→SKA2 and SKA→BAL2) than in the control treatments (BAL→BAL-C 2 and SKA→SKA-C 2) in data including all phyla (see arrows in Fig. 2).

DISCUSSION

Our results confirm the 'habitat filtering' hypothesis by showing that short-term environmental stress has the potential to 'filter' an already clustered community to consist of even more closely related species or types. It is perhaps not very surprising that only a certain subset of the marine bacteria coped with changes in salinity and nutrient levels and the confinement to laboratory chemostats. It is, however, intriguing that this subset was phylogenetically clustered.

In agreement with previous studies (Horner-Devine & Bohannan 2006, Newton et al. 2007, Pontarp et al. 2012), we interpret our results on phylogenetic structure in the naturally sampled (non-transplanted) communities according to the traditional dichotomy, introduced by Webb et al. (2002), as a signal of habitat filtering. This hypothesis is supported by our results showing an increased phylogenetic community clustering with treatment. Some closely related species decreased in abundance and even went extinct, while others increased in abundance possibly because of resilience to the abiotic change and secondary biotic effects (see Sjöstedt et al. 2012 for details). Although 'new' species may have been recruited from the 'rare biosphere' (Sogin et al. 2010, Sjöstedt et al. 2012), this did not, in most cases, compensate for species loss. Further, although the community response to environmental change is unpredictable (e.g. Martiny et al. 2006, Fuhrman et al. 2008, Andersson et al. 2010), our results showed that the response, whether it was extinction or recruitment from the 'rare biosphere', was non-random in terms of species relatedness.

We expected the community structure to be primarily driven by salinity and nutrient levels and less by the bottle effect (Fig. 1). Unfortunately, we cannot explicitly tease apart the relative effect of the different stressors on the communities. However, our results (Fig. 2) supported our predictions (Fig. 1) about additive clustering effects between the chemostat confinement and changed salinity and DOC concentrations in the transplant-control comparisons. This indicates that there are indeed relevant effects from the ecological stressors of changed salinity and nutrient concentration. Possible mechanisms dictating these communities would, for example, be physiological resilience to changed salinity concentrations or relatedness-specific nutrient requirements. It should, however, be noted that our results also show that the bottle effect may cause environmental stress that 'filters' the communities. This does not, however, contradict our conclusion that environmental stress causes 'habitat filtering' and phylogenetic clustering in marine bacterial communities.

Although our conclusions are supported by previous studies, we do not exclude alternative structuring processes. It has, for example, been shown that competition within a community can, under certain circumstances, produce patterns of community clustering (Mayfield & Levine 2010). We have no means of distinguishing between signals driven by the abiotic effect of the treatment and secondary biotic effects from community dynamics. However, because our

medium was filtered and no obvious cell aggregations (investigated using microscopy) were detected in the chemostats during the experiment, we suspect that direct cell–cell interactions were limited. The relative role of primary abiotic effects and secondary biotic interactions in bacterial communities is still an open question that needs to be investigated further. It is also important to consider biases that may be introduced by our methodological approach. Some bacterial taxa (e.g. SAR 11 and SAR 86) are found frequently and in high abundance in clone libraries and sequences from 454-pyrosequencing from environmental samples (Mullins et al. 1995, Morris et al. 2002), but cultivation has not been possible with current laboratory techniques or only for a few strains in dilute cultures (Rappé et al. 2002, Cho & Giovannoni 2004, Simu & Hagström 2004). This may be one of the reasons for the diversity loss with treatment. This can also explain aspects of the increased clustering seen in the full data set (including all phyla) with treatment. However, it does not explain the generality in our results showing clear effects on both diversity and phylogenetic structure across the full range of phyla.

Our analyses are based on the assumption that closely related organisms are more ecologically similar than distantly related taxa. However, horizontal gene transfer (HGT) may introduce novel functions, including complex physiological capabilities encoded by operons, and thereby change the ecological niche of the recipient organism. HGT is not limited by sequence divergence and can occur between distantly related organisms (Lawrence 2002). It has been shown that ecologically relevant genes (operational genes, involved in housekeeping) are exchanged more often than phylogenetically informative genes (informational genes, involved in transcription, translation and related processes (Jain et al. 1999). Genetic exchange has also been shown to occur frequently among co-occurring bacteria (Caro-Quintero et al. 2011) and could therefore affect the phylogenetic structure of the community, reducing phylogenetic clustering and overdispersion. However, the level of phylogenetic clustering observed among the studied communities suggests that HGT did not have a substantial effect on the phylogenetic pattern.

Our approach is robust and general as we make no explicit assumptions about the bacterial characteristics, community structure or requirements. We only assume that different species have different requirements (ecological niches) and that environmental stress has the potential to rearrange community com-

position. By restricting the time frame of the experiment, we isolated ecological effects only. The use of OTUs, which were defined by the patristic distance RAMI threshold, as a proxy for species may have unforeseen effects on the results. In addition, the method has been shown to be scale-dependent (e.g. Silvertown et al. 2001, Cavender-Bares et al. 2006, Kraft et al. 2007, Swenson 2009, Leibold et al. 2010, Ricklefs & Jenkins 2011). However, we used different RAMI thresholds from the quite relaxed value of 0.06, which dampens possible noise in data, to the rather strict value of 0.02 to test the generality of our results. Previous studies have shown that the specific RAMI threshold does not have any qualitative effects on the results regarding phylogenetic structure; if anything, a relaxed RAMI value reduces the possibility of detecting significant structuring (Pontarp et al. 2012). Our results are also robust over the range of RAMI values, and the signals seem to be ubiquitous throughout the terminal parts of the phylogenies (shown by the increased clustering with decreased RAMI value; Fig. 3).

As we know little of the relevant ecological scale in bacteria, we did the analysis on different taxonomic scales. The clear phylogenetic signals found together with the change in community composition lead us to conclude that we used relevant scales for the analysis of our data. The consistencies in Δ NRI and Δ NTI with treatment, irrespective of, for example, the time of sampling, also imply generality in our results. One potential flaw of the present study is the lack of replicates and multiple controls. Despite this, we argue that our conclusions are justified. The results are robust over different taxonomies and RAMI threshold values, and the data contain multiple treatments. However, we encourage future data sampling and experimental work that may provide better microbial data that can be used as a model system to study ecological theory.

Finally, the link between ecological processes and phylogenetic patterns hinges on a relationship between ecologically relevant traits and species (here defined by our RAMI-threshold) relatedness (illustrated as a straight line in Fig. 1) (Wiens & Graham 2005, Wiens et al. 2010). It has been largely unknown if such a relationship exists in bacteria (see however Newton et al. 2007, Boyd et al. 2010). Previous studies on functional traits have also shown that microbial communities from similar environments exhibited functional profiles that were more similar than profiles from communities from divergent environments (Tringe et al. 2005, DeLong et al. 2006). Therefore, the fact that we found this non-random

phylogenetic signal, as a result of environmental stress, suggests a mapping between ecologically relevant traits (niches) and relatedness across these bacterial communities. Also, this mapping seems to be consistent over the full range of phyla studied here. We do, however, recognize that this relationship may not be valid for all taxa and traits, and it may be scale-dependent or nonlinear. Consequently, we invite further investigations of the distribution of traits over the phylogeny of bacteria to strengthen the implications of phylogenetic signal analysis.

Acknowledgements. The present work was supported by grants from the Swedish Research Council (to P.L.). We thank L. Riemann and A. Tunlid for invaluable comments and thoughts throughout the compilation of the manuscript.

LITERATURE CITED

- Ackerly DD, Schilck DW, Webb CO (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology* 87:S50–S61
- Andersson AF, Riemann L, Bertilsson S (2010) Pyrosequencing reveals contrasting seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME J* 4: 171–181
- Bouvier TC, del Giorgio PA (2002) Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. *Limnol Oceanogr* 47: 453–470
- Boyd ES, Hamilton TL, Spear JR, Lavin M, Peters JW (2010) [FeFe]-hydrogenase in Yellowstone National Park: evidence for dispersal limitation and phylogenetic niche conservatism. *ISME J* 4:1485–1495
- Caro-Quintero A, Deng J, Auchtung J, Brettar I, Hofle MG, Klappenbach J, Konstantinidis KT (2011) Unprecedented levels of horizontal gene transfer among spatially co-occurring *Shewanella* bacteria from the Baltic Sea. *ISME J* 5:131–140
- Cavender-Bares J, Keen A, Miles B (2006) Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. *Ecology* 87:S109–S122
- Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW (2009) The merging of community ecology and phylogenetic biology. *Ecol Lett* 12:693–715
- Cho JC, Giovannoni SJ (2004) Cultivation and growth characteristics of a diverse group of oligotrophic marine *Gammaproteobacteria*. *Appl Environ Microbiol* 70: 432–440
- Cottrell MT, Kirchman DL (2000) Natural assemblages of marine Proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl Environ Microbiol* 66:1692–1697
- Covert JS, Moran MA (2001) Molecular characterization of estuarine bacterial communities that use high- and low-molecular weight fractions of dissolved organic carbon. *Aquat Microb Ecol* 25:127–139
- Crump BC, Hopkinson CS, Sogin ML, Hobbie JE (2004) Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. *Appl Environ Microbiol* 70:1494–1505
- DeLong EF, Preston CM, Mincer T, Rich V and others (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311:496–503
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M and others (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–5072
- Emerson BC, Gillespie RG (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol Evol* 23:619–630
- Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8:186–194
- Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175–185
- Ferguson RL, Buckley EN, Palumbo AV (1984) Response of marine bacterioplankton to differential filtration and confinement. *Appl Environ Microbiol* 47:49–55
- Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL, Brown JH (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proc Natl Acad Sci USA* 105:7774–7778
- Glöckner FO, Zaichikov E, Belkova N, Denissova L, Pernthaler J, Pernthaler A, Amann R (2000) Comparative 16S rRNA analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of Actinobacteria. *Appl Environ Microbiol* 66: 5053–5056
- Haas BJ, Gevers D, Earl AM, Feldgarden M and others (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 21:494–504
- Horner-Devine MC, Bohannan BJM (2006) Phylogenetic clustering and overdispersion in bacterial communities. *Ecology* 87:S100–S108
- Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, NJ
- Jain R, Rivera MC, Lake JA (1999) Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci USA* 96:3801–3806
- Kraft NJB, Cornwell WK, Webb CO, Ackerly DD (2007) Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *Am Nat* 170: 271–283
- Krebs CJ (1999) *Ecological methodology*. Addison Wesley Longman, Menlo Park, CA
- Kunin V, Engelbrektsen A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* 12:118–123
- Langenheder S, Kisand V, Wikner J, Tranvik LJ (2003) Salinity as a structuring factor for the composition and performance of bacterioplankton degrading riverine DOC. *FEMS Microbiol Ecol* 45:189–202
- Lawrence JG (2002) Gene transfer in bacteria: speciation without species? *Theor Popul Biol* 61:449–460
- Leibold MA, Holyoak M, Mouquet N, Amarasekare P and others (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* 7:601–613
- Leibold MA, Economo EP, Peres-Neto P (2010) Metacommunity phylogenetics: separating the roles of environmental

- filters and historical biogeography. *Ecol Lett* 13: 1290–1299
- Losos JB, Schluter D (2000) Analysis of an evolutionary species-area relationship. *Nature* 408:847–850
- Magurran AE (2004) *Measuring biological diversity*. Blackwell Scientific Publications, Oxford
- Martiny JBH, Bohannan BJM, Brown JH, Colwell RK and others (2006) Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* 4:102–112
- Mayfield MM, Levine JM (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol Lett* 13: 1085–1093
- Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, Giovannoni SJ (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420:806–810
- Mullins TD, Britschgi TB, Krest RL, Giovannoni SJ (1995) Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. *Limnol Oceanogr* 40:148–158
- Newton RJ, Jones SE, Helmus MR, McMahan KD (2007) Phylogenetic ecology of the freshwater *Actinobacteria* acI lineage. *Appl Environ Microbiol* 73:7169–7176
- Nübel U, Garcia-Pichel F, Kuhl M, Muyzer G (1999) Quantifying microbial diversity: morphotypes, 16S rRNA genes, and carotenoids of oxygenic phototrophs in microbial mats. *Appl Environ Microbiol* 65:422–430
- Pausas JG, Verdu M (2010) The jungle of methods for evaluating phenotypic and phylogenetic structure of communities. *Bioscience* 60:614–625
- Pinhassi J, Berman T (2003) Differential growth response of colony-forming α - and γ -proteobacteria in dilution culture and nutrient addition experiments from Lake Kinneret (Israel), the eastern Mediterranean Sea, and the Gulf of Eilat. *Appl Environ Microbiol* 69:199–211
- Pommier T, Canbäck B, Riemann L, Boström KH and others (2007) Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* 16:867–880
- Pommier T, Canbäck B, Lundberg P, Hagström Å, Tunlid A (2009) RAMI: a tool for identification and characterization of phylogenetic clusters in microbial communities. *Bioinformatics* 25:736–742
- Pontarp M, Canbäck B, Tunlid A, Lundberg P (2012a) Phylogenetic analysis suggests that habitat filtering is structuring marine bacterial communities across the globe. *Microb Ecol* 64:8–17
- Pontarp M, Ripa J, Lundberg P (2012b) On the origin of phylogenetic structure in competitive metacommunities. *Evol Ecol Res* 14:1–16
- Rappé MS, Connon SA, Vergin KL, Giovannoni SJ (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 418:630–633
- Ricklefs RE, Jenkins DG (2011) Biogeography and ecology: towards the integration of two disciplines. *Philos Trans R Soc Lond B* 366:2438–2448
- Ricklefs RE, Schluter D (1993) *Species diversity in ecological communities: historical and geographical perspectives*. University of Chicago Press, Chicago, IL
- Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA, Pernthaler J (2009) Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ Microbiol* 11:2078–2093
- Silvertown J, Dodd M, Gowing D (2001) Phylogeny and the niche structure of meadow plant communities. *J Ecol* 89: 428–435
- Simu K, Hagström A (2004) Oligotrophic bacterioplankton with a novel single-cell life strategy. *Appl Environ Microbiol* 70:2445–2451
- Sjöstedt J, Koch-Schmidt P, Pontarp M, Canbäck B and others (2012) Recruitment of members from the rare biosphere of marine bacterioplankton communities after environmental disturbance. *Appl Environ Microbiol* 78: 1361–1369
- Sogin ML, Morrison H, McLellan S, Welch DM, Huse S (2010) The rare biosphere: sorting out fact from fiction. *Genome Biol* 11:119
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690
- Swenson NG (2009) Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS ONE* 4:e4390
- Tringe SG, von Mering C, Kobayashi A, Salamov AA and others (2005) Comparative metagenomics of microbial communities. *Science* 308:554–557
- Urban MC, Skelly DK (2006) Evolving metacommunities: toward an evolutionary perspective on metacommunities. *Ecology* 87:1616–1626
- Vamosi SM, Heard SB, Vamosi JC, Webb CO (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Mol Ecol* 18:572–592
- Vellend M (2010) Conceptual synthesis in community ecology. *Q Rev Biol* 85:183–206
- von Mering C, Hugenholtz P, Raes J, Tringe SG and others (2007) Quantitative phylogenetic assessment of microbial communities in diverse environments. *Science* 315: 1126–1130
- Webb CO (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *Am Nat* 156:145–155
- Webb CO, Ackerly DD, McPeck MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annu Rev Ecol Syst* 33:475–505
- Webb CO, Ackerly DD, Kembel SW (2008) Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24:2098–2100
- Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu Rev Ecol Syst* 36:519–539
- Wiens JJ, Ackerly DD, Allen AP, Anacker BL and others (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol Lett* 13: 1310–1324
- Wiens JJ, Pyron RA, Moen DS (2011) Phylogenetic origins of local-scale diversity patterns and the causes of Amazonian megadiversity. *Ecol Lett* 14:643–652