

Habitat species pools for phylogenetic structure in microbes

Jianjun Wang,^{1*} Janne Soininen² and Ji Shen¹

¹State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China.

²Department of Geosciences and Geography, University of Helsinki, Helsinki FIN-00014, Finland.

Summary

The processes underlying the phylogenetic structure of the biotic communities are typically scale-dependent and thus often poorly resolved. Illustrated by the study of macroorganisms, it is suggested that the relative influence of ecological processes on the phylogenetic structure of the communities can be inferred by the geographical definition of the species pools. However, given the high dispersal ability of microbes, the spatial delineation of the species pool may not be that practical for microbial taxa. This idea is supported by the observational data on bacteria along an elevational gradient. Significant negative values of standardized effect size of the mean nearest taxon distance were consistently observed for different sized species pools considered. Reviewing the reports on microbial phylogenetic structure so far, we suggested that the ‘habitat species pools’ are perhaps more important for microbes than spatial delineated ‘regional species pools’.

A recent decade has witnessed an increasing interest in using phylogenetic analyses with the null models to infer evolutionary and ecological processes underlying community structure, but the scale dependency is still a major challenge for these analyses (e.g. Swenson *et al.*, 2006). This is because in the community phylogenetics research, the null models involve subjectively defined ‘species pools’ (Webb *et al.*, 2002; Swenson *et al.*, 2006), while different processes typically act on different scales. Moreover, relevant scales may differ between the organisms concerned. Illustrated by the study of hummingbirds in the Andes (Graham *et al.*, 2009; González-Caro *et al.*,

2012), Lessard and colleagues (2012) advised that the local ecological processes should be studied using regional species pools truly influential on focal communities thus excluding the taxa that are unlikely to disperse to a focal site. For instance, one can define ‘the pool to species that are found within the same elevation band as the focal community’ (Lessard *et al.*, 2012). However, we expect that such a strict spatial delineation of the species pools may not be practical for microbes due to their high dispersal rate across sites and the pivotal role of species sorting by the environment in structuring community compositions (Wang *et al.*, in press). This would be especially true when the samples are collected within relatively small spatial scales, yet covering long environmental gradients (e.g. in mountain regions).

To test for the idea that the spatial delineation of species pools is not practical for microbes, we analyse bacterial phylogenetic structure with different sized species pools on mountainsides. Briefly, stream biofilm bacteria were sampled at 24 sites along a stony mountainside stream extending from 1820 to 4050 m in elevation (Wang *et al.*, 2011; 2012a). Bacterial 16S rRNA genes were amplified and sequenced using a Roche 454 FLX pyrosequencer. We separated the assemblages into two groups: (i) whole bacterial communities and (ii) phylum *Proteobacteria*. For both assemblage groups, we defined species pools using different spatial grains by dividing the whole elevational gradient into one to four bins. The significant phylogenetic signal in ecological niches in the data allows one to infer the underlying ecological processes using phylogenetic information (Wang *et al.*, in press). For phylogenetic community structure, we calculated the observed mean nearest taxon distance (MNTD) of all species pairs occurring in a community based on an observed dataset of a community (Webb *et al.*, 2002). MNTD is an estimate for the mean phylogenetic relatedness between each operational taxonomic unit (OTU) in a bacterial community and its nearest relative, and it provides useful ecological information for understanding the phylogenetic structure of the communities as well as underlying drivers (i.e. species interactions and environmental filtering). Further, the standardized effect size of MNTD (ses.MNTD) was calculated for each community by comparing the observed MNTD with 1000 null communities (Webb *et al.*, 2002). The obtained standardized effect size measure (ses.MNTD)

Received 20 October, 2012; accepted 23 December, 2012.
*For correspondence. E-mail jjwang@niglas.ac.cn; Tel. (+86) 25 86882030; Fax (+86) 25 57714759.

can be used to test for whether communities are phylogenetic clustered or overdispersed (see below and Webb, 2000). We generated the null communities with the assumption that all species are equally able to colonize the sites within the defined species pool, and thus each species have the same expected prevalence (Gotelli, 2000; Kembel and Hubbell, 2006). The total species richness of each elevation was maintained at the same level, and the species for each elevation were chosen randomly without a replacement from the corresponding pool of species present along the elevation. Negative ses.MNTD values and low quantiles ($P < 0.05$) suggest that communities are phylogenetically clustered due to environmental filtering, that is, co-occurring species are more closely related than expected by chance. Positive values and high quantiles ($P > 0.95$) indicate that the co-occurring species are less closely related than expected by chance (i.e. communities are overdispersed) (Webb, 2000).

For one species pool covering the whole elevational gradient, ses.MNTD and elevation showed a significant negative relationship with each other for the whole bacterial communities (Fig. 1A) as well as for *Proteobacteria* (Fig. 1E). When different spatial sizes in species pools

were considered, the trends for the ses.MNTD along elevation were however not always significantly negative for each bin (Fig. 1). Species pools may thus affect the elevational trends in phylogenetic clustering due to the varying length of the covered environmental gradients. This view is supported by the steeper slopes in the relationships between phylogenetic structure and elevation for the whole bacterial communities at low elevations in the smaller elevational bins (Fig. 1A–D). This may be due to the large environmental heterogeneity (e.g. in total phosphorus) between sites at low elevations (Wang *et al.*, 2011) and due to the increased human activities below 2580 m in elevation (Wang *et al.*, 2012b). Due to the limited number for sampling sites, we cannot reliably test the influence of environmental heterogeneity or human activity on the relationship between phylogenetic structure and elevation, but it warrants future investigations. Furthermore, the ses.MNTD values for the whole bacterial community were higher than for *Proteobacteria* only (Fig. 1E). Generally, these results, agreeing with the previous reports on macroorganisms (Swenson *et al.*, 2006; Lessard *et al.*, 2012), suggest that species pools should not be ignored when examining the influence of local ecological processes for microbes.

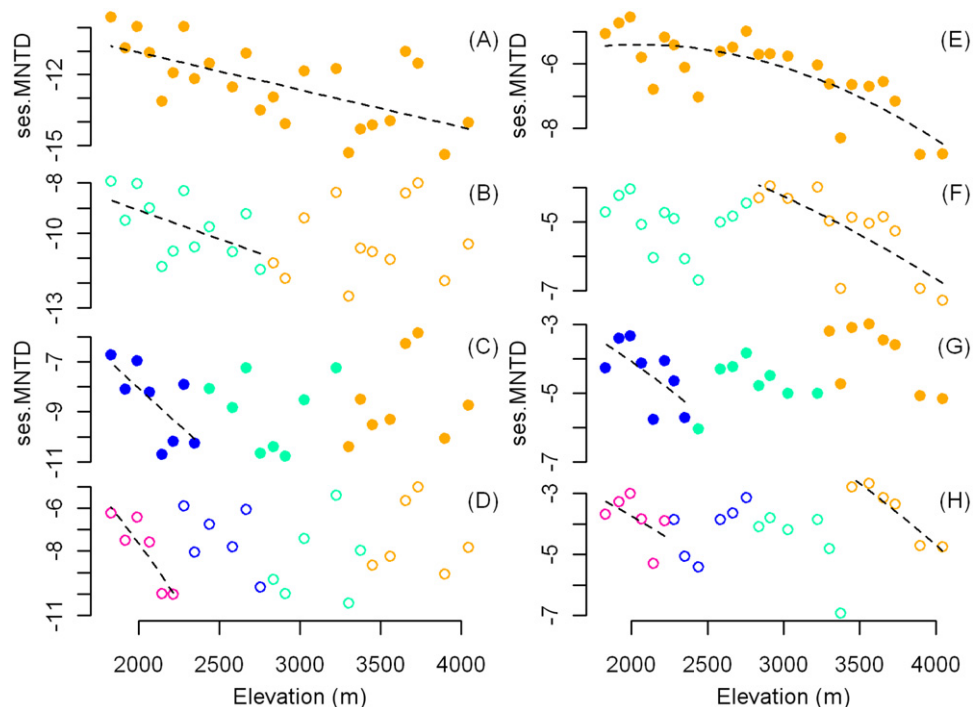


Fig. 1. The phylogenetic structure of bacteria with the different definitions of the species pools with one (A and E), two (B and F), three (C and G) and four (D and H) bins along the elevational gradient, which are indicated by the colour numbers in each panel. The left panels (A), (B), (C) and (D) represent the whole bacterial communities and the right panels (E), (F), (G) and (H) represent *Proteobacteria*. (A) and (E) were adopted from (Wang *et al.*, 2012a). All ses.MNTD were significant ($P < 0.01$). Negative ses.MNTD values and low quantiles ($P < 0.05$) suggest that communities are phylogenetically clustered due to environmental filtering, that is, co-occurring species are more closely related than expected by chance. The dotted lines indicate the significant trends of ses.MNTD as a function of elevation using either linear or quadratic model depending of the value of Akaike's information criterion.

All values of ses.MNTD were significantly negative when different sized species pools were considered although the values approached the null prediction when the species pool sizes decreased (Fig. 1). The negative values indicated that bacterial communities were phylogenetically clustered regardless of the spatial species pools or assemblage groups considered. It is interesting to note that for microbes, phylogenetic clustering is likely to be evident at various spatial scales ranging from millimetres (Armitage *et al.*, 2012), 1 km (Horner-Devine and Bohannan, 2006; Stegen *et al.*, 2012), 30–40 km (Bryant *et al.*, 2008; Wang *et al.*, 2012a), hundreds of kilometres (Horner-Devine and Bohannan, 2006), to a globe scale (Barberán and Casamayor, 2010). Furthermore, at regional scales, persistent seed bank may explain the seasonal dynamics for microbial communities (Caporaso *et al.*, 2011). These evidences indicate that the spatial delineation of 'regional species pool' is not that important for microbes and will not greatly affect whether communities are phylogenetically clustered or overdispersed. That is, the microbial communities are more likely to show phylogenetic clustering than overdispersion regardless of the spatial scales considered. The length of the environmental gradients within and across habitat types studied, i.e. the environmental variation within the 'habitat species pool', is perhaps more important than the spatial scale when local ecological processes are examined for microbes. For instance, one should first consider the degree of environmental heterogeneity even when defining the species pool within the elevational bands. Furthermore, 'regional species pool' needs to be delineated only when there is obvious dispersal limitation in the system, e.g. in the subsurface environments.

Compared with hummingbirds (Graham *et al.*, 2009) or plants (Swenson *et al.*, 2006), microbial communities show a long evolutionary history. Along with a rapid development of high-throughput sequencing techniques and experimental manipulation in microbial communities, it may be promising to conduct more studies within a phylogenetic framework. For instance, dividing microbial communities into functional groups or evolutionary bins would be useful when inferring the drivers of local community structure. Rather than detecting significant phylogenetic clustering at the microbial community level, one would like to uncover the strength of local ecological processes from the phylogenetic structure along the defined environmental gradient gradients (e.g. Horner-Devine and Bohannan, 2006).

Acknowledgements

J.W. appreciates 973 (2012CB956100), NIGLAS (NIGLAS2012135004), NSFC (40903031, 41273088),

Jiangsu NSF (BK2010605), CPSF (2011 M500397) and CAS oversea visiting scholarship (2011-115). Thanks to Dr Brian Hayden for language proofreading.

References

- Armitage, D.W., Gallagher, K.L., Youngblut, N.D., Buckley, D.H., and Zinder, S.H. (2012) Millimeter-scale patterns of phylogenetic and trait diversity in a salt marsh microbial mat. *Front Microbiol* **3**: 293.
- Barberán, A., and Casamayor, E. (2010) Global phylogenetic community structure and b-diversity patterns in surface bacterioplankton metacommunities. *Aquat Microb Ecol* **59**: 1–10.
- Bryant, J.A., Lamanna, C., Morlon, H., Kerkhoff, A.J., Enquist, B.J., and Green, J.L. (2008) Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proc Natl Acad Sci USA* **105**: 11505–11511.
- Caporaso, J.G., Paszkiewicz, K., Field, D., Knight, R., and Gilbert, J.A. (2011) The Western English Channel contains a persistent microbial seed bank. *ISME J* **6**: 1089–1093.
- González-Caro, S., Parra, J.L., Graham, C.H., McGuire, J.A., and Cadena, C.D. (2012) Sensitivity of metrics of phylogenetic structure to scale, source of data and species pool of hummingbird assemblages along elevational gradients. *PLoS ONE* **7**: e35472.
- Gotelli, N.J. (2000) Null model analysis of species co-occurrence patterns. *Ecology* **81**: 2606–2621.
- Graham, C.H., Parra, J.L., Rahbek, C., and McGuire, J.A. (2009) Phylogenetic structure in tropical hummingbird communities. *Proc Natl Acad Sci USA* **106**: 19673–19678.
- Horner-Devine, C.M., and Bohannan, B.J.M. (2006) Phylogenetic clustering and overdispersion in bacterial communities. *Ecology* **87**: 100–108.
- Kembel, S.W., and Hubbell, S.P. (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology* **87**: 86–99.
- Lessard, J.-P., Belmaker, J., Myers, J.A., Chase, J.M., and Rahbek, C. (2012) Inferring local ecological processes amid species pool influences. *Trends Ecol Evol* **27**: 600–607.
- Stegen, J.C., Lin, X., Konopka, A.E., and Fredrickson, J.K. (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* **6**: 1653–1664.
- Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J., and Zimmerman, J.K. (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology* **87**: 2418–2424.
- Wang, J., Soininen, J., Zhang, Y., Wang, B., Yang, X., and Shen, J. (2011) Contrasting patterns in elevational diversity between microorganisms and macroorganisms. *J Biogeogr* **38**: 595–603.
- Wang, J., Soininen, J., He, J., and Shen, J. (2012a) Phylogenetic clustering increases with elevation for microbes. *Environ Microbiol Rep* **4**: 217–226.
- Wang, J., Soininen, J., Zhang, Y., Wang, B., Yang, X., and Shen, J. (2012b) Patterns of elevational beta diversity in

- micro- and macroorganisms. *Glob Ecol Biogeogr* **21**: 743–750.
- Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J.C., *et al.* (in press) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J.*
- Webb, C.O. (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *Am Nat* **156**: 145–155.
- Webb, C.O., Ackerly, D.D., McPeck, M.A., and Donoghue, M.J. (2002) Phylogenies and community ecology. *Annu Rev Ecol Syst* **33**: 475–505.