Diversity begets diversity in microbiomes

One Sentence Summary: Making use of the most comprehensive biodiversity catalogue, the Earth Microbiome Project, we show that prokaryote diversification significantly correlates with microbiome diversity, pointing to the importance of species interactions in driving microbial evolution and community structure.

Authors: Naïma Madi[1], Michiel Vos[2], Pierre Legendre[1] and B. Jesse Shapiro[1*]

1. Departement de sciences biologiques, Universite de Montreal, Canada
2. European Centre for Environment and Human Health, University of Exeter, Penryn, UK

*correspondence: jesse.shapiro@umontreal.ca

keywords: microbiome, diversification, evolution, ecology, Earth Microbiome Project, 16S rRNA

Abstract

Microbes are embedded in complex microbiomes where they engage in a wide array of interspecific interactions. However, how these interactions shape diversification, and ultimately biodiversity, is not well understood. Two competing hypotheses have been put forward to explain how species interactions could influence diversification rates. Ecological Controls (EC) predicts a negative diversity-diversification relationship, where the evolution of novel types becomes constrained as available niches become filled. Diversity Begets Diversity (DBD) predicts a positive relationship, with diversity promoting diversification via niche construction and other species interactions. Using the Earth Microbiome Project, the largest standardized survey of global biodiversity to date, we provide support for DBD as the dominant driver of microbiome diversity. Only in the most diverse microbiomes does DBD reach a plateau, presumably because of increasingly saturated niche space. Genera that are strongly associated with particular biomes show a stronger DBD relationship than non-residents, consistent with prolonged evolutionary interactions driving diversification. Genera with larger genomes also experience a stronger DBD response, which could be due to a higher potential for metabolic interactions and niche construction. Our results provide evidence that microbiome diversity – and its potential for future diversification – is crucially shaped by species interactions.

40 **Main text**

41 The majority of the Tree of Life consists of microbes (*1–3*) and the functioning of
42 all Earth's ecosystems is reliant on diverse microbial communities (*4*). High-throughput
43 16S rRNA gene amplicon sequencing studies continue to yield unprecedented insight into
44 the taxonomic richness of microbiomes (e.g. (*5*, *6*)), and abiotic drivers of community
45 composition (e.g. pH (*7*, *8*)) are increasingly uncovered. Microbe-microbe (biotic)
46 interactions can also be important in determining community composition (*9*), but
47 comparatively little is known about how such interactions (e.g. cross-feeding (*10*) and
48 toxin-mediated interference competition (*11*, *12*)) shape microbiome diversity.

49 The dearth of studies exploring how microbial interactions influence
50 diversification stands in marked contrast to a long research tradition on biotic controls of
51 plant and animal diversity (*13*, *14*). In an early study of 49 animal (vertebrate and
52 invertebrate) community samples, Elton plotted the number of species versus the number
53 of genera and observed a ~1:1 ratio in each individual sample, but a ~4:1 ratio when all
54 samples were pooled (*14*). Elton took this observation as evidence for competitive
55 exclusion, preventing related species, more likely to overlap in niche space, to co-exist.
56 This concept, more recently referred to as niche filling or "ecological controls" (EC) (*15*)
57 predicts speciation (or, more generally, diversification) rates to decrease with standing
58 species diversity because available niches become filled over time (*16*). In contrast, the
59 Diversity Begets Diversity (DBD) model predicts that diversity favors further
60 diversification when species interactions create novel niches (*17*, *18*). For example, niche
61 construction (i.e. the physical, chemical or biological alteration of the environment by
62 microbes) could influence the evolution of the species constructing the niche, and/or that

63 of co-occurring species (*19*, *20*). Evidence for the action of EC vs. DBD in plant and
64 animal communities has been mixed (*18*, *21-24)*, and models suggest that intraspecific
65 competition can promote diversification, while interspecific competition and predation
66 inhibit diversification (*25*). Laboratory evolution experiments tracking the diversification
67 of a focal bacterial lineage in communities of varying complexity have also yielded
68 contradicting results, with support for EC, DBD, or intermediate scenarios (*26-30*).
69 To test whether natural microbial communities are shaped by EC or DBD
70 dynamics, we used 2,000 microbiome samples each rarefied to 5,000 16S Amplicon
71 Sequence Variants (ASVs; exact, error-corrected DNA sequences, directly comparable
72 across samples (*31*)) archived in the Earth Microbiome Project (EMP), the largest
73 available repository of biodiversity based on standardized sampling and sequencing
74 protocols (*32*). All samples used were subjected to rarefaction, as diversification rate
75 estimates are highly sensitive to sampling effort (*33,34*). In the absence of time series or a
76 fossil record, microbial diversification must be inferred from extant diversity. Rather than
77 taking a phylogenetic approach requiring complex assumptions (*35*, *36*), we use the
78 equivalent of the Species:Genus (S:G) ratios that Elton used three quarters of a century
79 ago (*14*) to infer bacterial diversification rates. We use a range of taxonomic ratios
80 (ASV:Genus, Genus:Family, Family:Order, Order:Class, and Class:Phylum) as proxies
81 for diversification of a focal lineage of interest, from shallow to deep evolutionary time,
82 and plot these as a function of the number of non-focal lineages (Genera, Families,
83 Orders, Classes, and Phyla, respectively) with which the focal lineage could interact. A
84 negative relationship is consistent with the EC hypothesis, whereas a positive relationship
85 is consistent with the DBD hypothesis (**Figure 1**).

86 We used generalized linear mixed models (GLMMs) (*37*) to determine how the
87 diversification of a focal lineage (*e.g.* its ASV:Genus ratio) is affected by the diversity of
88 other lineages (*e.g.* non-focal genera) in the community. The effects of environment (as
89 defined by the EMP Ontology 'level 3 biomes;' Methods) and the identity of the focal
90 lineage were included by fitting these as random effects on the slope and intercept. The
91 submitting laboratory (identified by the principal investigator) and the EMP unique
92 sample identifier (i.e. if two taxa were part of the same sample) were also included as
93 random effects. The DBD model was supported across all taxonomic ratios, which all had
94 significantly positive slopes fitting the diversity-diversification relationship (**Table 1,**
95 **Supplementary Data file 1 Section 1**). In some cases, the variation due to lineage,
96 lineage*environment, or environment* laboratory interactions could reverse the sign of
97 the slope (**Table 1**), but the vast majority of slope estimates are positive, supporting DBD
98 (**Figure S1**).
99 As an illustration, **Figure 2** shows diversification within the Proteobacteria
100 (Classes:Phylum) as a function of all other Phyla in each sample across 17 biomes. All
101 slopes were significantly positive ($P < 0.001$ after Bonferroni correction), except in
102 hypersaline and non-saline sediments. The strongest DBD effect was found in animal
103 proximal gut microbiomes (slope=0.3, P=1.03e−08). For each taxonomic ratio, the three
104 most prevalent taxa followed positive slopes in most environments (**Figures S2-S6**), with
105 only a few instances of significantly negative slopes (Alphaproteobacteria, Rhizobiales,
106 Flavobacteriales, Flavobacteriaceae, Planctomyces in the plant rhizosphere, and
107 Clostridium in non-saline sediment). Overall, the majority of significant slope estimates

108 were positive (81.7-99.5% depending on the taxonomic ratio considered), favoring DBD
109 as the dominant (but not universal) model supported by the data (**Table S1**).
110 To test for any potential confounding effects of data structure or sampling bias,
111 we sought to remove any patterns of co-occurrence between ASVs in the same sample
112 via permutation. We took 2,000 simulated samples by selecting from the overall
113 distribution of 155,002 unique ASVs across all samples, weighted by their abundance
114 (total number of sequence counts). This resulted in a slightly negative diversity-
115 diversification relationship (slope = −0.002; Pearson correlation = −0.61; $P<2.2.e-16$;
116 **Figure S7**), indicating that the observed positive relationships (**Table 1; Figure 2**) are
117 not the effect of data structure.
118 The taxonomic approach considered thus far has certain limitations. First, not all
119 taxonomic ranks have the same phylogenetic depth (*38*) and not all named taxa are
120 monophyletic (*39*). Second, unclassified ASVs had to be excluded from taxonomic
121 analysis. Although these limitations are not expected to bias the results toward either EC
122 or DBD, we sought to validate the results with a taxonomy-independent approach,
123 clustering ASVs at decreasing levels of nucleotide identity, from 100% identical ASVs
124 down to 75% identity (roughly equivalent to phyla (*40*)). We estimated diversification as
125 the mean number of descendants per cluster (e.g. number of 100% clusters per 97%
126 cluster) and plotted this against the total number of non focal clusters (97% identity in
127 this example). For each of the six nucleotide divergence ratios tested, the relationship
128 between diversity and diversification remained positive (**Figure S8**), consistent with
129 DBD and suggesting that the taxonomic analyses were largely unbiased.

130 To exclude the possibility that our results were driven by abiotic confounders, we
131 repeated the taxonomic analysis on a subset of 192 EMP samples for which
132 measurements of four important abiotic drivers of diversity, temperature, pH, latitude,
133 and elevation (*7*, *8*, *15*, *41*) were available. We fitted a GLMM with diversification rate
134 as the dependent variable, and with the number of non-focal lineages, the four abiotic
135 factors and their interactions as predictors (fixed effects). As in the full dataset (**Table 1**),
136 diversification was positively associated with diversity at all taxonomic ratios (**Table S2**).
137 As expected, certain abiotic factors, alone or in combination with diversity, had
138 significant effects on diversification. However, the effects of abiotic factors were always
139 weaker than the effect of community diversity (**Table S2**; **Supplementary Data file 1**
140 **Section 2**). Although only a small subset of abiotic factors was considered, this analysis
141 suggests that the DBD trend is unlikely to be mainly driven by variation in the abiotic
142 environment.
143 The DBD hypothesis rests on the premise that species interactions drive
144 diversification (*15*, *18*). We therefore expect that lineages that are more tightly associated
145 with a specific biome (i.e. long-term residents) are more likely to have had a long history
146 of interaction and thus are more likely to experience DBD than lineages that are not
147 tightly associated with that biome (i.e. poorly adapted migrants or broadly adapted
148 generalists). To test this prediction, we clustered environmental samples by their genus-
149 level community composition using fuzzy *k*-means clustering (**Figure 3A**). Three clusters
150 were identified, which we refer to as 'animal-associated', 'saline,' and 'non-saline'
151 clusters. All three clusters included some unexpected outliers (*e.g.* plant corpus grouping
152 with animals), but were generally intuitive and consistent with known distinctions

153 between host-associated vs. free-living (*32*), and saline vs. non-saline communities (*42*).
154 Resident genera were defined as those with a strong preference for a particular
155 environment cluster, using indicator species analysis (permutation test, $P<0.05$; **Figure**
156 **3A**; **Figure S9**; **Supplementary Data file 2**), and genera without a strong preference
157 were considered generalists. For each environment cluster, we ran a GLMM with resident
158 genus-level diversity (number of non-focal genera) as a predictor of diversification
159 (ASV:Genus ratio) for residents, generalists, or migrants (residents of one cluster found
160 in a different cluster) (**Supplementary Data file 1 Section 3**). Resident diversity had no
161 significant effect on the diversification of generalists ($z=0.646$, $P=0.518$; $z=0.279$,
162 $P=0.780$; $z=0.347$, $P=0.729$, respectively for animal-associated, saline and non-saline
163 clusters), but did significantly increase resident diversification ($z=7.1$, $P= 1.25e-12$;
164 $z=3.316$, $P=0.0009$; $z=7.109$, $P=1.17e-12$, respectively). Resident diversity significantly
165 decreased migrant diversification in saline ($z=-3.194$, $P=0.0014$) and non-saline
166 environment clusters ($z=-2.840$, $P=0.0045$), but had no significant effect in the animal-
167 associated cluster ($z=-0.566$, $P=0.571$) (**Figure 3B**). These results suggest that diversity
168 begets diversification among lineages sharing the same environment over a long
169 evolutionary time period, but that this is not the case for lineages that do not consistently
170 occur in the same microbiome and presumably interact less frequently. The
171 diversification of migrants in a new environment might even be impeded, presumably
172 because most niches are already occupied by residents.
173 The positive effect of diversity on diversification should eventually reach a
174 plateau as niches, including those constructed by biotic interactions, become saturated. In
175 the animal distal gut, a relatively low-diversity biome, we observed a strong linear DBD

176 relationship at most sequence identity ratios; in contrast, the more diverse soil biome
177 clearly attained a plateau (**Figure S10**). To further test the hypothesis that increasingly
178 diverse microbiomes experience weaker DBD due to saturated niche space, we used a
179 GLMM to model Family:Order, Order:Class and Class:Phylum diversification (as DBD
180 slope variation by environment was statistically significant for these taxonomic ratios;
181 **Table 1**). We included the interaction between diversity and environment type as a fixed
182 effect. Consistent with our hypothesis, DBD slopes were significantly more positive in
183 less diverse (often host-associated) biomes (**Figure 4A**, **Figure S11, Supplementary**
184 **Data file 1 Section 4**). We also found evidence for a DBD plateau in the nucleotide
185 identity analyses, in which cubic models provided slightly better fits than linear models
186 across all biomes (**Figure S8**).
187 The Black Queen hypothesis posits that microbes embedded in complex
188 communities can exploit the production of extracellular public goods produced by other
189 species, resulting in selection for loss of genes encoding these goods – as long as the
190 essential trait is not lost from the community as a whole (*43*). Lineages that interact more
191 frequently with other lineages through such public good exploitation can be expected to
192 experience greater genome reduction and also to experience stronger DBD, because their
193 survival and diversification is dependent on other community members. To test whether
194 genome size is negatively correlated with DBD, we assigned genome sizes to 576 genera
195 for which at least one whole-genome sequence was available and added an interaction
196 term between genome size and diversity as a fixed effect to the GLMM (Methods).
197 Contrary to the Black Queen prediction, we observed a slight but significant positive
198 effect of genome size on the slope (z=2.5, P=0.01; **Figure 4B, Supplementary Data file**

199 **1 Section 5**). The positive relationship may even be stronger than estimated, because
200 genus-level genome size estimates are likely quite noisy. This result supports a model in
201 which biotic interactions (and resulting diversification) drive genome expansion (*e.g.*
202 through the accumulation of toxin- and resistance-gene diversity during antagonistic
203 coevolution (*11*)). Alternatively (or additionally), species with larger biosynthetic gene
204 repertoires and greater opportunity to engage in niche construction (*19*) could be more
205 prone to interact with other species, driving DBD.

206 Using 10 million individual marker sequences, we demonstrate a pervasive
207 positive relationship between prokaryotic diversity and diversification, which holds
208 across a broad range of microbiome types. The strength of the DBD relationship
209 dissipates with increasing microbiome diversity which might be due to niche saturation,
210 or potentially due to the fact that highly diverse communities prevent species from
211 reliably interacting with each other. DBD appears to be particularly strong among deeply
212 diverged lineages (*e.g.* phyla), suggesting the importance of DBD in the ancient
213 diversification of bacterial lineages and supporting the view that high taxonomic ranks
214 are ecologically coherent (*44*, *45*). The very early stages of diversification are
215 inaccessible at the resolution of 16S ASVs, but could be addressed in the future using
216 genomic and metagenomic methods. Due to the correlational nature of our data, it is not
217 possible to test whether the positive relationship between diversification and diversity is
218 primarily due to the creation of novel niches via biotic interactions and niche construction
219 (*20*), or potentially due to increased competition leading to adaptation to underexploited
220 resources (*12*, *29*). Regardless of the underlying mechanisms, our results clearly
221 demonstrate the importance of biotic interactions in shaping microbiome diversity, which

222 has important implications for their function and stability (*46*, *47*). The answer to the

223 question 'why are microbiomes so diverse?' (*48*) might in a large part be because

224 microbiomes are so diverse (*23*).

Acknowledgements

We thank Luke Thompson for assistance obtaining EMP data and Zofia Ecaterina Taranu, Vincent Fugère and Guillaume Larocque for advice on Generalized Linear Mixed Models. We are also grateful to Steven Kembel and Tom Battin for critical comments that improved the manuscript. **Funding:** This project was made possible by an NSERC Discovery Grant and Canada Research Chair to BJS. **Authors contributions:** Conceptualization: BJS, MV. Data curation: NM. Formal analysis: NM, MV, BJS. Funding acquisition: BJS. Investigation: NM, MV, PL, BJS. Methodology: NM, MV, PL, BJS. Resources: BJS, PL. Supervision: PL, BJS. Software: NM. Visualization: NM. Writing original draft: NM, MV, BJS. Writing - review & editing: NM, MV, PL, BJS. **Competing interests:** none to declare. **Data and materials availability:** All data is available from the Earth Microbiome Project (ftp.microbio.me), as detailed in the Methods. All computer code used for analysis are available at https://github.com/Naima16/dbd.git.

Supplementary Materials

Materials and Methods

Tables S1 – S2

Fig S1 – S11

File 1. Full GLMM outputs.

File 2. Indicator species analysis.

Table 1. Diversity has a positive effect on diversification across taxonomic ratios. The GLMMs showed statistically significant positive effect of diversity on diversification. Each row reports the effect of diversity on diversification, as well as its standard deviation, Wald z-statistic for its effect size and the corresponding *P*-value (left section), or standard deviation on the slope for the significant random effects (right section). SE=standard error, Env=environment type, Lin=lineage type, Lab=Principal Investigator ID, Sample=Sample ID. Interactions are denoted as '*'. n.s.=not significant (likelihood-ratio test).

	Slope (fixed effects)				**Standard deviation on the slope (random effects)**				
	Diversity	**SE**	**z**	***P***	**Env**	**Lin**	**Lin*Env**	**Env*Lab**	**Sample**
ASV: Genus	0.091	0.016	5.792	6.95e-09	n.s.	0.074	0.142	0.114	0.067
Genus: Family	0.047	0.008	5.911	3.41e-09	n.s.	0.071	0.07	0.039	n.s.
Family: Order	0.119	0.017	7.001	2.54e-12	0.023	0.094	0.092	0.106	n.s.
Order: Class	0.109	0.020	5.447	5.13e-08	0.05	0.141	0.078	0.051	n.s.
Class: Phylum	0.272	0.043	6.341	2.29e-10	0.119	0.174	0.119	0.114	n.s.

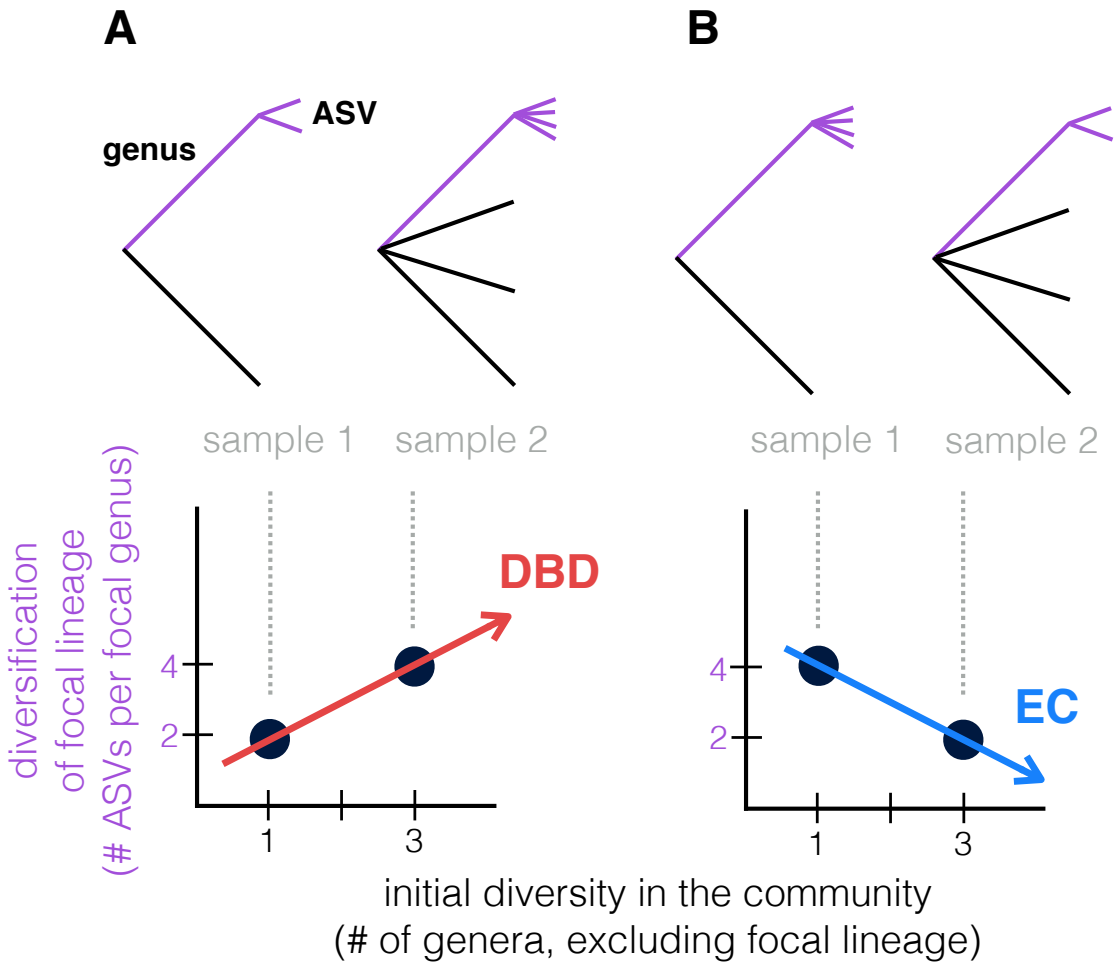


Figure 1. Contrasting the Diversity Begets Diversity (DBD) and Ecological Controls (EC) models of diversification. We consider the diversification of a focal lineage as a function of initial diversity present at the time of diversification. (A) For example, sample 1 contains one non-focal genus, and two ASVs diversify within the focal genus (point at x=1, y=2 in the plot). Sample 2 contains three non-focal genera, and four ASVs diversify within the focal genus (point at x=3, y=4). Tracing a line through these points yields a positive slope, supporting the Diversity Begets Diversification (DBD) model (red). (B) Alternatively, a negative slope would support the Ecological Controls (EC) model (blue line).

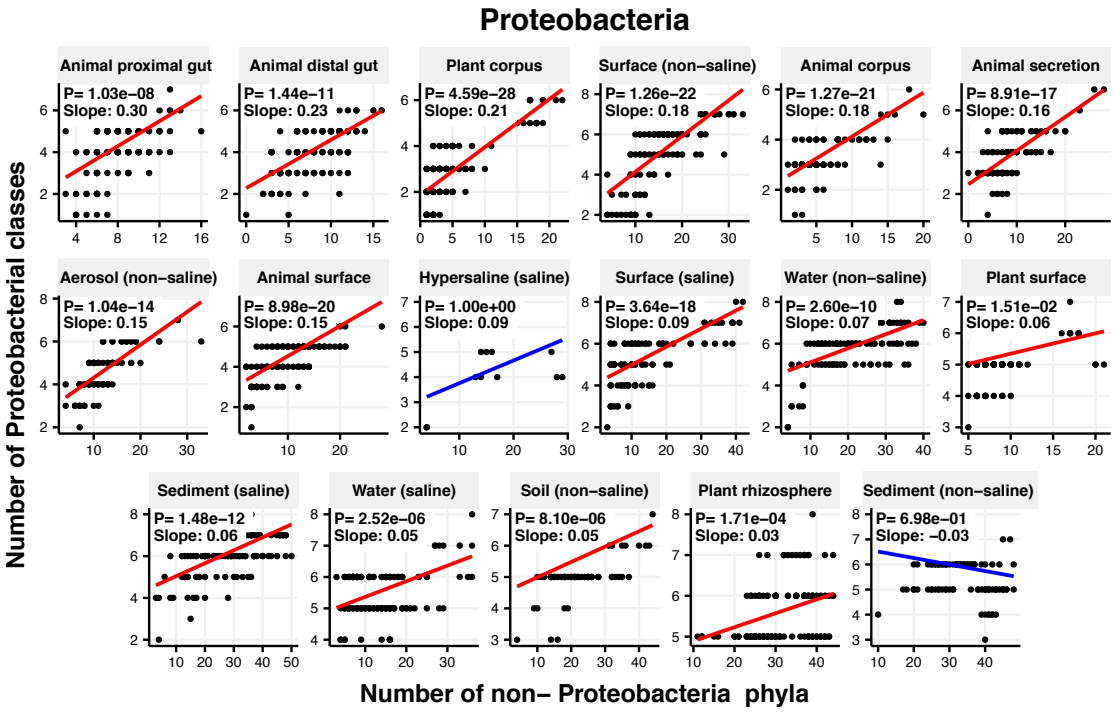


Figure 2. Diversification as a function of diversity across biomes in the phylum Proteobacteria. Linear models are shown for diversification (the number of classes within Proteobacteria, y-axis) as a function of diversity (the number of non-proteobacterial phyla, x-axis) in each of the 17 environments (EMPO3 biomes). P-values are Bonferroni corrected for 17 tests. Significant (P <0.05) models are shown with red trend lines; non-significant (P > 0.05) trends are shown in blue.

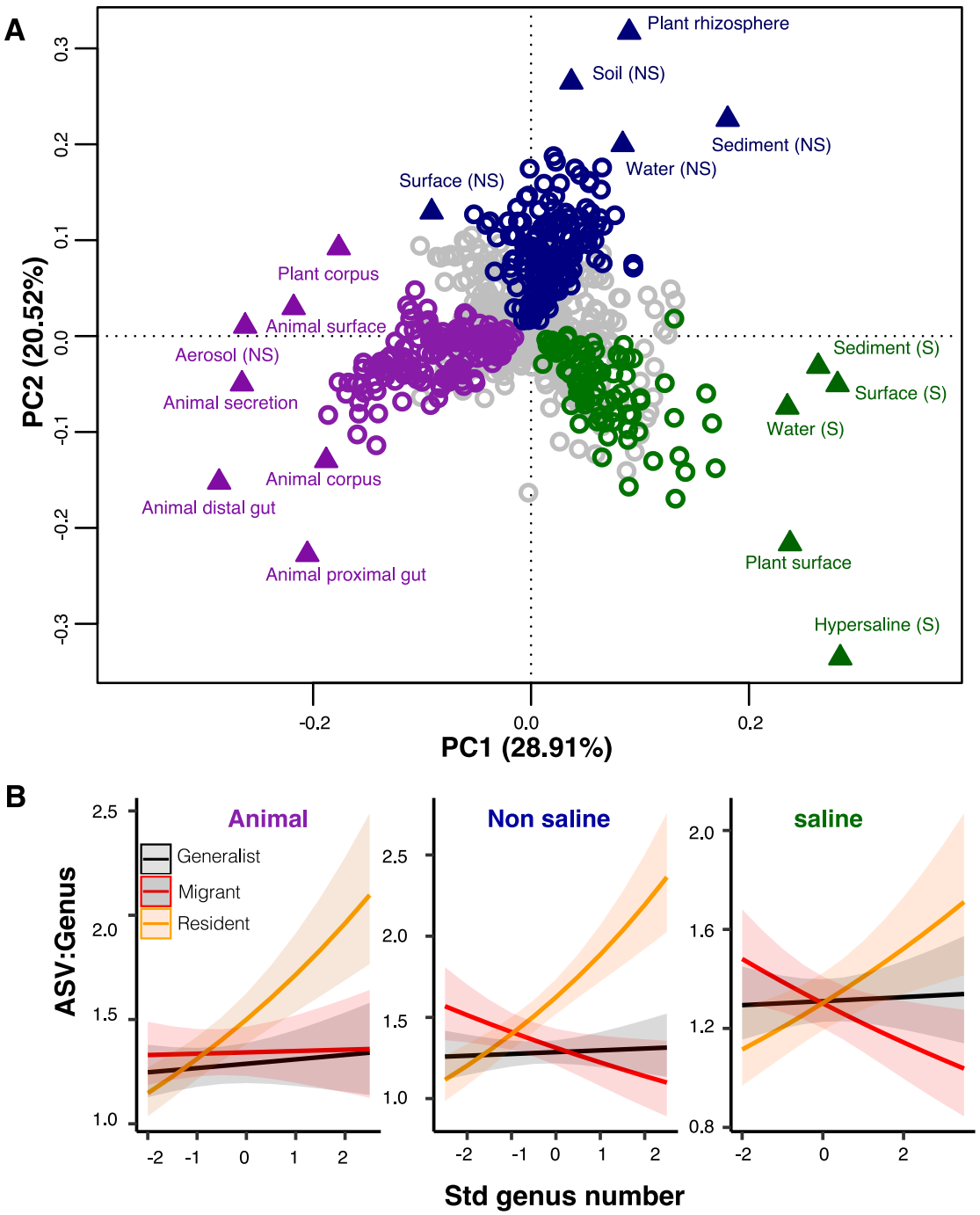
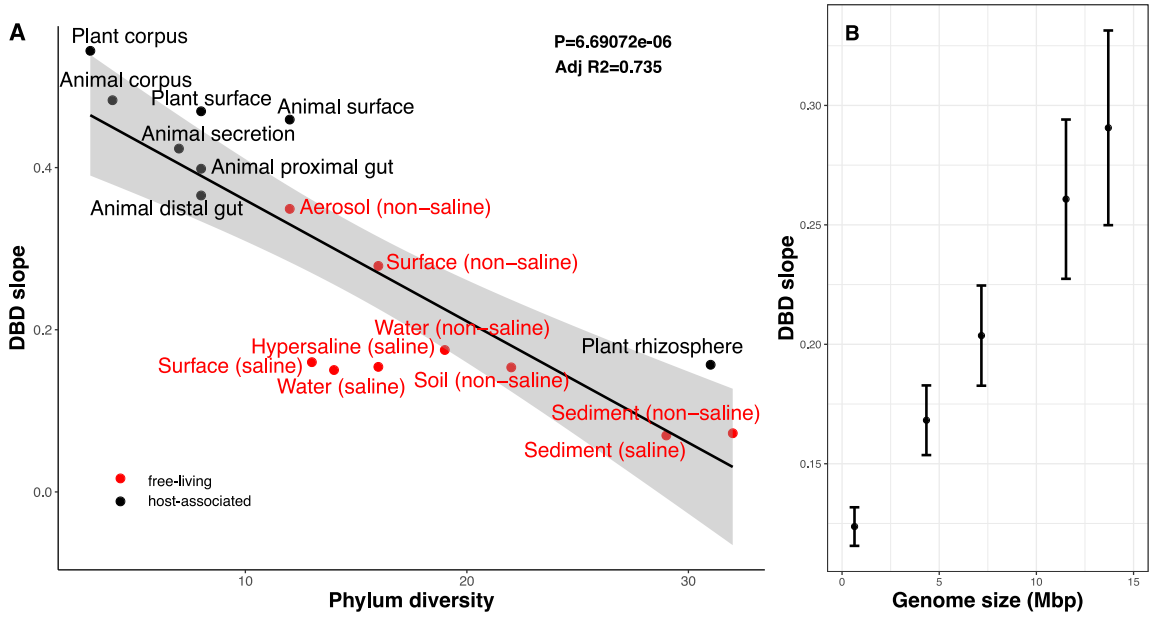


Figure 3. Diversity begets diversification in resident versus non resident genera. (A) PCA showing genera clustering into their preferred environment clusters. Circles indicate genera and triangles indicate environments (EMPO 3 biomes). The three environment clusters identified by fuzzy *k*-means clustering are: Non-saline (NS, blue),

280 saline (S, green) and animal-associated (purple). Resident genera were identified by
281 indicator species analysis. **(B) DBD in resident versus migrant genera across**
282 **environment clusters.** Results of GLMMs modeling diversification as a function of
283 diversity in resident, migrant, or generalist groups. The x-axis shows the standardized
284 number of non-focal resident genera (diversity); the y-axis shows the number of ASVs
285 per focal genus (diversification). Resident focal genera are shown in orange, migrant
286 focal genera in red, and generalist focal genera in black.



287

288

289 **Figure 4. Ecological and evolutionary mechanisms to explain variation in the**

290 **strength of DBD. (A) DBD slope is higher in low-diversity (often host-associated)**

291 **microbiomes.** The x-axis shows the mean number of phyla in each biome. On the y-axis,

292 DBD slope was estimated by the GLMM predicting diversification as a function of the

293 interaction between diversity and environment type at the Class:Phylum ratio

294 (**Supplementary Data file 1 Section 4.3**). The line represents a regression line; the

295 shaded area depicts 95% confidence limits of the fitted values. **(B) Positive correlation**

296 **between genome size and DBD slope.** Results are shown from a GLMM predicting

297 diversification as a function of the interaction between diversity and genome size at the

298 ASV:Genus ratio (**Supplementary Data file 1 Section 5**). The x-axis is genus-level

299 genome size in Mbp (min=0.97, max=14.78); the y-axis is DBD slope (the effect of

300 diversity on diversification). Vertical bars indicate 95% confidence limits of the fitted

301 values.

References

1. S. Sunagawa *et al.*, Ocean plankton. Structure and function of the global ocean microbiome. *Science*. **348**, 1261359 (2015).

2. P. Lapierre, J. P. Gogarten, Estimating the size of the bacterial pan-genome. *Trends Genet.* **25**, 107–110 (2009).

3. L. A. Hug *et al.*, A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016).

4. P. G. Falkowski, T. Fenchel, E. F. Delong, The microbial engines that drive Earth's biogeochemical cycles. *Science*. **320**, 1034–1039 (2008).

5. M. L. Sogin *et al.*, Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12115–12120 (2006).

6. S. Louca, F. Mazel, M. Doebeli, L. W. Parfrey, A census-based estimate of Earth's bacterial and archaeal diversity. *PLoS Biol.* **17**, e3000106 (2019).

7. C. L. Lauber, M. Hamady, R. Knight, N. Fierer, Soil pH as a predictor of soil bacterial community structure at the continental scale: a pyrosequencing-based assessment. *Appl. Environ. Microbiol.* **75**, 5111-5120 (2009).

8. J. F. Power *et al.*, Microbial biogeography of 925 geothermal springs in New Zealand. *Nat. Commun.* **9**, 2876 (2018).

9. D. M. Needham, J. A. Fuhrman, Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nature Microbiology*, 1, 16005 (2016).

10. E. C. Seth, M. E. Taga, Nutrient cross-feeding in the microbial world. *Front. Microbiol.* **5**, 350 (2014).

11. T. L. Czárán, R. F. Hoekstra, L. Pagie, Chemical warfare between microbes promotes biodiversity. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 786–790 (2002).

12. M. E. Hibbing, C. Fuqua, M. R. Parsek, S. B. Peterson, Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* **8**, 15–25 (2010).

13. G. F. Gause, *The Struggle for Existence* (Courier Corporation, 2003).

14. C. Elton, Competition and the Structure of Ecological Communities. *J. Anim. Ecol.* **15**, 54–68 (1946).

15. D. Schluter, M. W. Pennell, Speciation gradients and the distribution of biodiversity. *Nature*. **546**, 48–55 (2017).

16. D. L. Rabosky, A. H. Hurlbert, Species richness at continental scales is dominated by ecological limits. *Am. Nat.* **185**, 572–583 (2015).

17. R. H. Whittaker, Evolution and Measurement of Species Diversity. *Taxon*. **21**, 213–251 (1972).

18. V. Calcagno, P. Jarne, M. Loreau, N. Mouquet, P. David, Diversity spurs diversification in ecological communities. *Nat. Commun.* **8**, 15810 (2017).

19. M. San Roman, A. Wagner, An enormous potential for niche construction through bacterial cross-feeding in a homogeneous environment. *PLoS Comput. Biol.* **14**, e1006340 (2018).

20. K. N. Laland, F. J. Odling-Smee, M. W. Feldman, Evolutionary consequences of niche construction and their implications for ecology. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10242–10247 (1999).

21. T. D. Price *et al.*, Niche filling slows the diversification of Himalayan songbirds. *Nature* **509**, 222-225 (2014).

22. D. L. Rabosky *et al.*, An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*. **559**, 392–395 (2018).

23. B. C. Emerson, N. Kolm, Species diversity can drive speciation. *Nature*. **434**, 1015–1017 (2005).

24. M. W. Palmer, T. A. Maurer, Does Diversity Beget Diversity? A Case Study of Crops and Weeds. *J. Veg. Sci.* **8**, 235–240 (1997).

25. Pontarp Mikael, Petchey Owen L., Ecological opportunity and predator–prey interactions: linking eco-evolutionary processes and diversification in adaptive radiations. *Proceedings of the Royal Society B: Biological Sciences*. **285**, 20172550 (2018).

26. M. A. Brockhurst, N. Colegrave, D. J. Hodgson, A. Buckling, Niche occupation limits adaptive radiation in experimental microcosms. *PLoS One*. **2**, e193 (2007).

27. P. Gómez, A. Buckling, Real-time microbial adaptive diversification in soil. *Ecol. Lett.* **16**, 650–655 (2013).

28. J. R. Meyer, R. Kassen, The effects of competition and predation on diversification in a model adaptive radiation. *Nature*. **446**, 432–435 (2007).

29. A. Jousset, N. Eisenhauer, M. Merker, N. Mouquet, S. Scheu, High functional diversity stimulates diversification in experimental microbial communities. *Sci Adv*. **2**, e1600124 (2016).

30. S. F. Bailey, J. R. Dettman, P. B. Rainey, R. Kassen, Competition both drives and impedes diversification in a model adaptive radiation. *Proc. Biol. Sci.* **280**, 20131253 (2013).

31. B. J. Callahan, P. J. McMurdie, S. P. Holmes, Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).

32. L. R. Thompson *et al.*, A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, nature24621 (2017).

33. N. J. Gotelli, R. K. Colwell, Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* **4**, 379–391 (2001).

34. A. Z. Krug, D. Jablonski, J. W. Valentine, Species-genus ratios reflect a global history of diversification and range expansion in marine bivalves. *Proc. Biol. Sci.* **275**, 1117–1123

(2008).

35. R. S. Etienne, A. L. Pigot, A. B. Phillimore, How reliably can we infer diversity-dependent diversification from phylogenies? *Methods Ecol. Evol.* **7**, 1092–1099 (2016).

36. S. Louca *et al.*, Bacterial diversification through geological time. *Nat Ecol Evol*. **2**, 1458–1467 (2018).

37. B. M. Bolker *et al.*, Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* **24**, 127–135 (2009).

38. M. Vos, A species concept for bacteria based on adaptive divergence. *Trends Microbiol.* **19**, 1–7 (2011).

39. D. H. Parks *et al.*, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).

40. K. T. Konstantinidis, J. M. Tiedje, Towards a genome-based taxonomy for prokaryotes. *J. Bacteriol.* **187**, 6258–6264 (2005).

41. M. Delgado-Baquerizo *et al.*, A global atlas of the dominant bacteria found in soil. *Science*. **359**, 320–325 (2018).

42. C. A. Lozupone, R. Knight, Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11436–11440 (2007).

43. J. J. Morris, R. E. Lenski, The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio* **3**, e00036-12 (2012).

44. L. Philippot *et al.*, The ecological coherence of high bacterial taxonomic ranks. *Nat. Rev. Microbiol.* **8**, 523–529 (2010).

45. J. B. H. Martiny, S. E. Jones, J. T. Lennon, A. C. Martiny, Microbiomes in light of traits: A phylogenetic perspective. *Science*. **350**, aac9323 (2015).

46. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science*. **350**, 663–666 (2015).

47. F. Pennekamp *et al.*, Biodiversity increases and decreases ecosystem stability. *Nature*. **563**, 109–112 (2018).

48. D. E. Dykhuizen, Santa Rosalia revisited: why are there so many species of bacteria? *Antonie Van Leeuwenhoek*. **73**, 25–33 (1998).

49. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, Articles*. **67**, 1–48 (2015).

50. X. A. Harrison *et al.*, A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*. **6**, e4794 (2018).

51. W. Forstmeier, H. Schielzeth, Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behav. Ecol. Sociobiol.* **65**, 47–55 (2011).

52. M. Dufrene, P. Legendre, Species Assemblages and Indicator Species: The Need for a Flexible Asymmetrical Approach. *Ecol. Monogr.* **67**, 345–366 (1997).