bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Stochastic microbial community assembly decreases biogeochemical function

Emily B. Graham[1,*] and James C. Stegen[2]

[1] Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352; emily.graham@pnnl.gov
[2] Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352; James.Stegen@pnnl.gov
* Correspondence: emily.graham@pnnl.gov; Tel.: +1-509-372-6049

Abstract: Much research has focused on improving ecosystem models by incorporating microbial regulation of biogeochemistry. However, models still struggle to predict biogeochemical function in future scenarios linked to accelerating global environmental change. Ecological mechanisms may influence the relationship between microbial communities and biogeochemistry, and here, we show that stochastic dispersal processes (e.g., wind-driven or hydrologic transport) can suppress biogeochemical function. Microbial communities are assembled by deterministic (e.g., selection) and stochastic (e.g., dispersal) processes, and the balance of these two processes is hypothesized to influence how microbial communities correspond to biogeochemical function. We explore the theoretical basis for this hypothesis and use ecological simulation models to demonstrate potential influences of assembly processes on ecosystem function. We assemble 'receiving' communities under different levels of dispersal from a source community (selection-only, moderate dispersal, and homogenizing dispersal). We then calculate the degree to which assembled individuals are adapted to their environment and relate the level of adaptation to biogeochemical function. We also use ecological null models to further link assembly the level of deterministic assembly to function. We find that dispersal can decrease biogeochemical function by increasing the proportion of maladapted taxa, outweighing selection. The niche breadth of taxa is also a key determinant of biogeochemical function, suggesting a tradeoff between the function of generalist and specialist species. Together, our results highlight the importance of considering ecological assembly processes to reduce uncertainty in predictions of biogeochemical cycles under future environmental scenarios.

Keywords: stochasticity, determinism, microbial ecology, biogeochemistry, null model, ecosystem function

---

1. Introduction

Environmental systems are undergoing rapid changes in the wake of heightened anthropogenic activity. Predictive ecosystem models have been developed and refined since the 1980s to forecast how these changes may influence future ecosystem health [e.g., 1,2]. While such models originally focused on inferring biogeochemical process rates from substrate pool sizes, recent efforts have improved forecasts by including properties of the microbial communities that mediate these cycles, such as microbial biomass [3-6], dormancy [7], carbon use efficiency [8], and community structure or functional traits [3,9,10]. Yet, these models still struggle to accurately represent biogeochemical cycling in dynamic natural ecosystems, and scientists have called for a better understanding of microbial ecology to increase predictive ability [11-13]. As models begin to consider microbial ecology, there is a need to decipher linkages among spatiotemporal microbial processes and ecosystem-level biogeochemical function. Understanding these relationships will help us reduce uncertainty in future ecosystem functioning [11,12,14-17].

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Recent attempts to link microbial communities and ecosystem-level processes have yielded mixed results [12,13,18-21]. A common theme of this research is that microbial community structure reflects the aggregation of historical processes, and such processes may in turn influence the function of a community [13,14,22-25]. For instance, Evans and Wallenstein [22] demonstrate that precipitation history constrains microbial biogeochemical responses to drying and rewetting cycles. Likewise, Fukami *et al.* [26] show legacy effects of microbial colonization history on decomposer communities, and historical dispersal events may be a driver of community function in certain circumstances [27,28].

Several researchers have proposed that the inclusion of community assembly mechanisms, such as dispersal and selection, may enhance our predictions of microbially-mediated biogeochemical cycling [13,29-31]. These processes operate through space and time to determine microbial community composition [12,23,29,32]. They fall into two predominate categories that can be summarized as correlated (i.e., deterministic) or uncorrelated (i.e., stochastic) to environmental conditions. Stochastic processes can be further classified into dispersal, evolutionary diversification, and ecological drift, while determinism is largely dictated by selection [29,33]. Experimental research has shown unpredictable relationships between microbial diversity and ecosystem function, leading to the hypothesis that differences in community assembly history—and thus the relative contributions of stochastic and deterministic processes—drives relationships between microbial community structure and biogeochemical function [30,31].

In particular, many systems have high rates of dispersal that may decrease community-level biogeochemical function. Aquatic systems such as streams, rivers, and marine communities experience hydrologic mixing that disperses microorganisms into new environments [34-36]. In the terrestrial biosphere, snowmelt and rainfall events facilitate hydrologic connectivity in soils [37,38] that can transport organisms, and wind can play a large role in dispersing spore-forming microorganisms [39-41]. High rates of dispersal can add organisms to a microbial community that are not well-suited to local environmental conditions, decreasing the community's overall efficiency of biogeochemical cycling. For example, pH [42] and salinity [43,44] are widely considered strong environmental filters on microbial community structure. If microorganisms disperse from a moderate pH or salinity environment to a more extreme environment, they may be maladapted and have to expend energy to express traits that maintain neutral internal pH (e.g., H+ pumps) or maintain cellular water content (e.g., osmotic stress factors). These cell maintenance activities detract from the energy available to transform biogeochemical constituents and may suppress overall community rates of biogeochemical function. In contrast, locally adapted species would putatively have more efficient mechanisms for cell maintenance in the local environment and possess more expendable energy for catalyzing biogeochemical reactions.

The cumulative impacts of ecological processes through time and how they relate to ecosystem-level processes is an emerging research frontier in ecosystem science [13,14,22,24]. Hawkes and Keitt [14] demonstrate through simplified ecological models that historical conditions can constrain the response of a microbial community to an environmental perturbation. They lay a solid theoretical framework in which community-level microbial functions are the accretion of individual life-histories that determine population growth, composition, and fitness. However, they acknowledge their exclusion of dispersal processes from their models and do not explicitly consider stochasticity in their analysis. Hawkes and Keitt [14] therefore provide a foundation for future research and call for a holistic understanding of historical processes on microbial function, with a particular emphasis on the underlying mechanisms generating these trends.

Building on the work of Hawkes and Keitt [14], we hypothesize that stochastic assembly processes strongly influence biogeochemical function. Our aim here is to formalize this hypothesis and provide a simulation-based demonstration of how stochastic assembly can influence function. To do so, we employ ecological simulation models to explicitly represent dispersal and selection-based processes, and we link the resulting communities to biogeochemical function through organismal adaptation.

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2. Materials and Methods

All simulations, null models, statistical analyses, and graphics were completed in *R* software (https://cran.r-project.org/). The simulation models consisted of two parts and were followed by statistical analysis. One hundred replicates were run for each parameter combination in the simulation models.

First, a regional species was constructed following the protocol outlined in Stegen *et al.* [32]. Regional species pools were constructed by simulating diversification in which new species arise through mutations in the environmental optima of ancestral organisms. Environmental optima evolve along an arbitrary continuum from 0 to 1, following a Brownian process. Regional species pools reach equilibria according to the constraints described by Stegen *et al.* [32] and Hurlbert and Stegen [45] and summarized here: (1) we define a maximum number of total individuals in the pool (2E+06) such that the population size of a given species declines with an increasing number of species, and (2) the probability of extinction for a given species increases as its population size decreases according to a negative exponential function [population extinction probability $\propto \exp(-0.001 * \text{population size})$].

The evolution of a regional species pool was initiated from a single ancestor with randomly chosen environmental optimum (initially comprising all 2E+06 individuals in the population). Mutation probability was set as 1.00E−05. A descendant's environmental optimum deviated from its ancestor by a quantity selected from a Gaussian distribution with mean 0 and SD 0.2. Following mutation, populations sizes were reduced evenly such that the total number of individuals remained 2 million. The simulation was run for 250 time steps, which was sufficient to reach equilibrium species richness.

The model's second component assembled 4 local communities from the regional species pool according to scenarios conceptualized to test our hypotheses. Species were drawn from the regional species pool to generate a source community under weak selection and three receiving communities with no dispersal, moderate dispersal, and high dispersal in which the strength of selection/niche breadth and environmental conditions were allowed to vary across simulations. All communities had 100 species and 10,000 individuals, drawn probabilistically from the regional species pool. Selection probabilities of species from the regional pool were set by a Gaussian function with standard deviation set as the 'niche breadth' and the deviation of each species optimum from the defined environment per the following equation:

$$\mathrm{SP} = e^{-\frac{d^2\sqrt{2n\pi}}{2n}}, \tag{1}$$

where SP is the selection probability, d is the deviation from the environmental optima, and n is the niche breadth.

In the source community, we used one niche breadth for all simulations, which was the maximum value used for receiving communities (0.175). This value represents weak selection, which allows for assembly of a broader range of species than under strong selection. The environmental conditions in the source community was also set to a single value using the following procedure: We generated 10 regional species pools and combined species abundances and environmental optima from these pools to generate one aggregate pool representative of the probable distributions of environmental optima yielded by our simulations. We set the environmental optimum of the source community to one end of this spectrum (5th percentile) to allow for comparisons with receiving communities that had the same or larger environmental values. This maximized the range of environmental differences between the source and receiving communities.

For receiving communities, we allowed the environmental optimum and niche breadth to vary across simulations. Environmental optima ranged from 0.05 to 0.95 by intervals of 0.04736842 to yield 20 conditions. Niche breadth ranged from 0.0075 to 0.175 by 0.008815789 to yield 20 conditions. Receiving communities were assembled under all possible combinations of environmental optima and niche breadths. Communities for our selection-only case were assembled based only on the selection

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

probabilities as defined by equation 1, using the same approach as for the assembly of the source community. For moderate and homogenizing dispersal, we modified selection probabilities to incorporate species dispersing from the source community as defined by the following equation:

$$SP_{disp} = SP + 0.05(S_{source}{}^{D}), \quad (2)$$

where SP_{disp} is the selection probability accounting for dispersal, S_{source} is species abundances in the source community, and D a parameter reflecting dispersal rate. Parameter D was set to 1 for moderate dispersal and 2 for homogenizing dispersal. All possible communities were simulated with 100 replicate regional species pools such that all possible combination of parameters were used once with each regional species pool. Biogeochemical function was calculated for all communities as the sum of species abundances multiplied by their selection probabilities.

Following the assembly of communities, the role of selection in structuring communities was determined via null models developed by Stegen *et al.* [32,46]. Each receiving community was compared to an associated source community that was assembled from the same regional species pool. The approach uses pairwise phylogenetic turnover between communities, calculated using the mean-nearest-taxon-distance (βMNTD) metric [47,48], to infer the relative influence of selection. Communities were evaluated for significantly less turnover than expected (βNTI < −2, homogeneous selection) or more turnover than expected (βNTI > 2, variable selection) by comparing observed βMNTD values to the mean of a null distribution of βMNTD values – and normalizing by its standard deviation – to yield the β-nearest taxon index (βNTI) [49]. Significance levels for βNTI are based on standard deviations: |βNTI| = 2 denotes two standard deviations from the mean of the null distribution. Inferences from βNTI have previously been shown to be robust [13,23,32,50].

Finally, we analyzed differences in model outputs using standard statistical approaches. We used pairwise Kolmogorov-Smirnov tests to compare distributions of species optima between simulations. To compare biogeochemical function of the three dispersal cases, we used ANOVA followed by post-hoc Tukey's HSD tests. We also analyzed changes in biogeochemical function across changes in the environmental parameter using quadratic regressions. Finally, correlations between βNTI and function were assessed with linear regression.

3. Results and Discussion

As ecosystem process models become more sophisticated [e.g., 7,9,11], there is a need to improve these models by better understanding the linkages among spatiotemporal microbial processes and ecosystem-level functions. Here, we use ecological simulation models to demonstrate the importance of stochastic microbial assembly in predicting future biogeochemical functioning. We propose that (1) stochastically assembled communities are composed of species that are not well adapted to their environment and, in turn, that (2) stochastic assembly processes decrease biogeochemical function (Figure 1).

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

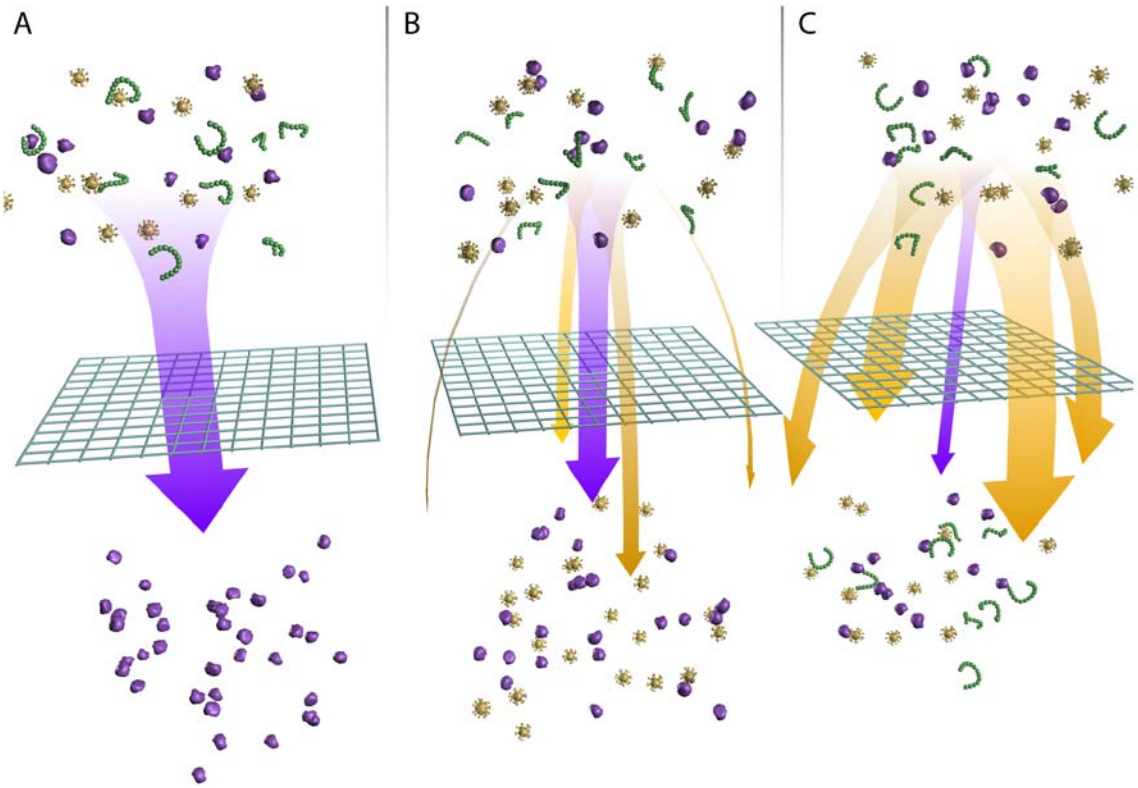


Figure 1. We propose a conceptual model in which stochastic assembly processes decrease biogeochemical function. (A) In a community structured entirely by determinism, environmental filtering restricts community composition to species that are well-adapted to prevailing conditions, resulting in enhanced biogeochemical function. (B) In communities with moderate stochasticity (conceptualized here as dispersal), environmental filters are weaker, and stochasticity increases the abundance of maladapted organisms in the community. In turn, the community is less efficient and exhibits lower biogeochemical function. (C) Under high levels of stochasticity, environmental filters are minimal, allowing for the assembly of a heterogeneous community. The community is unsuited to the environment and mediates low rates of biogeochemical function.

We found that microbial community assembly history impacted biogeochemistry by altering the fitness of organisms to a given environment. The niche optima of taxa in selection-only communities more closely matched their simulated environmental conditions compared to communities assembled with dispersal (Figure 2, $p < 0.001$). These selection-only communities also experienced the highest rates of biogeochemical function on average (Figure 3a, $p < 0.0001$).

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

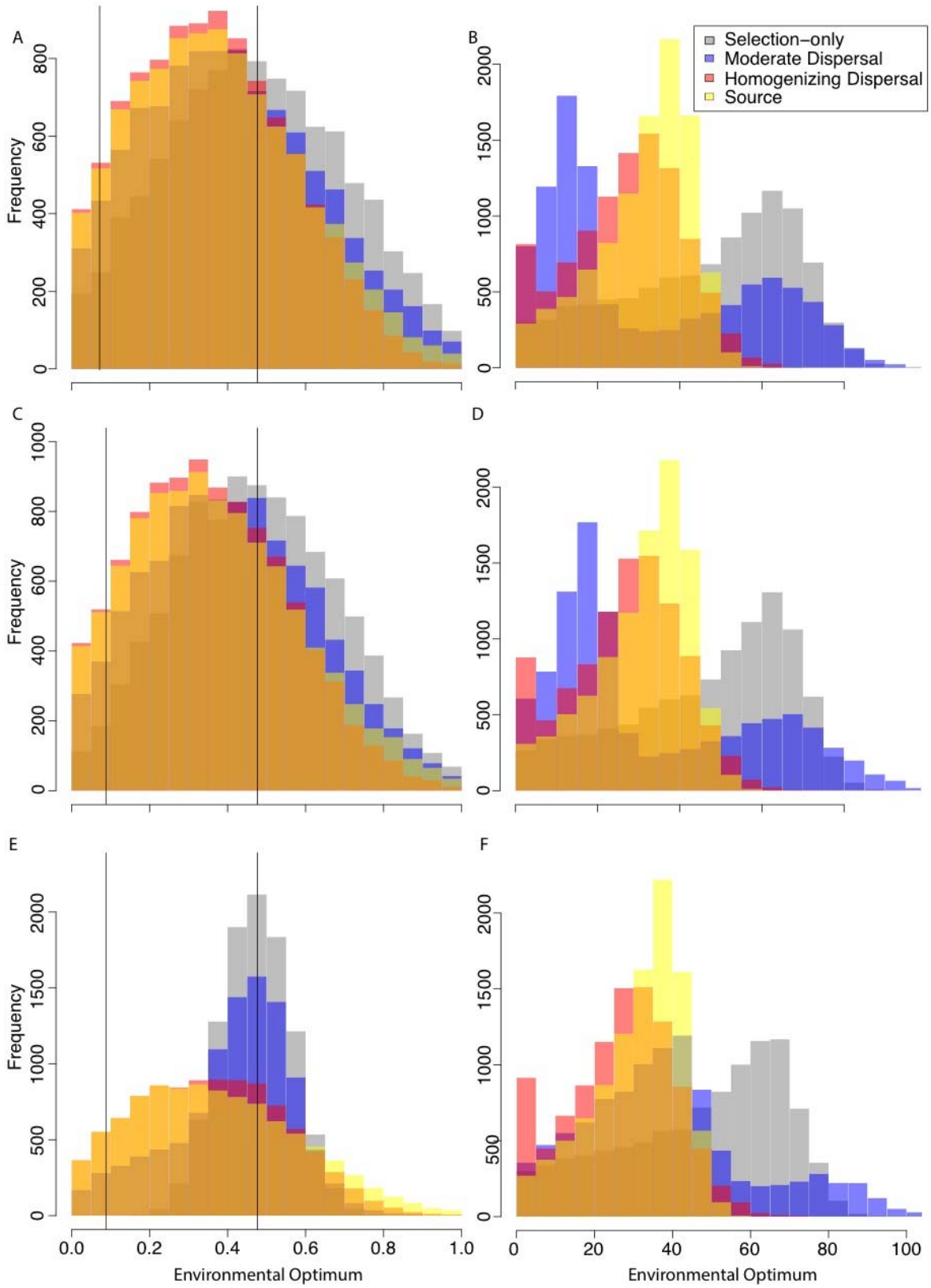


Figure 2. Histograms of species optima are shown under narrow (A-B, niche breadth = 0.0075), moderate (C-D, niche breadth = 0.086842105), and wide (E-F, niche breadth = 0.175) niches in an example

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

environment (0.476315789). Column 1 displays optima without accounting for species abundances and column 2 displays abundance-weighted values. The source community and its environment is displayed in yellow. Regardless of niche breadth, selection-only scenarios (gray) displayed the tightest distribution around the defined environment (vertical black line), followed by moderate (blue) and homogenizing (red) dispersal. Moderate dispersal scenarios had bimodal abundance-weighted distributions, signifying contributions from both the source and local communities. Further, narrow niche breadth supported tighter distributions in all dispersal cases.

In natural systems, microbial community composition differences can be due to competitive dynamics that select for organisms in different niches (determinism) [51,52] and to immigration of new taxa from the regional species pool (stochasticity) [29,49,53]. Strong local selective pressures can lead to more fit species and enhanced biogeochemistry [29]. Due to the lack of immigrating maladapted species in the selection-only simulations, biogeochemical rates were maintained regardless of the difference between source and receiving community environments. This indicates that maintenance of biogeochemical function can be due solely to species adaptation to local conditions (as opposed to kinetic or resource limitations that are not included in our models). Indeed, a plethora of literature demonstrates that environmental features such as pH [42], nutrients [54], and salinity [43,44] impact microbial community structure and biogeochemical function, and our results support linkages among these ecosystem properties via microbial adaptation.

We also indicate that microorganisms from dissimilar environments, such as those transported from longer distances or across steep physical or geochemical gradients, are more likely to suppress biogeochemical function than those from more similar environments. When we included dispersal from a source community, greater differences between the source and receiving communities led to decreases in biogeochemical function in the receiving communities (Figure 3b, $p < 0.0001$), and this effect became more pronounced as dispersal increased. Natural systems are influenced by some combination of dispersal and selection and our results indicate that function is maximized when dispersal is minimized and selection is maximized.

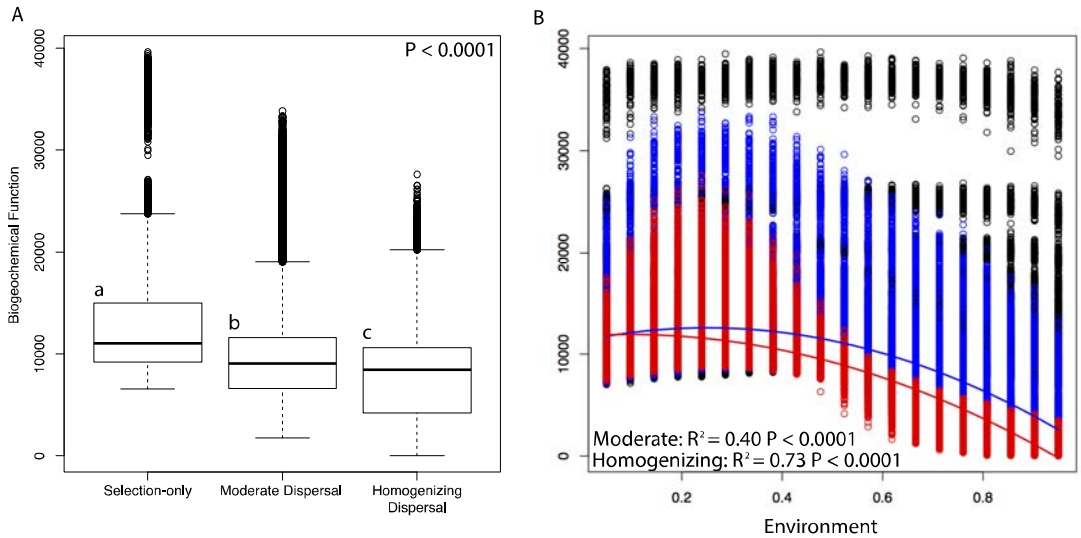
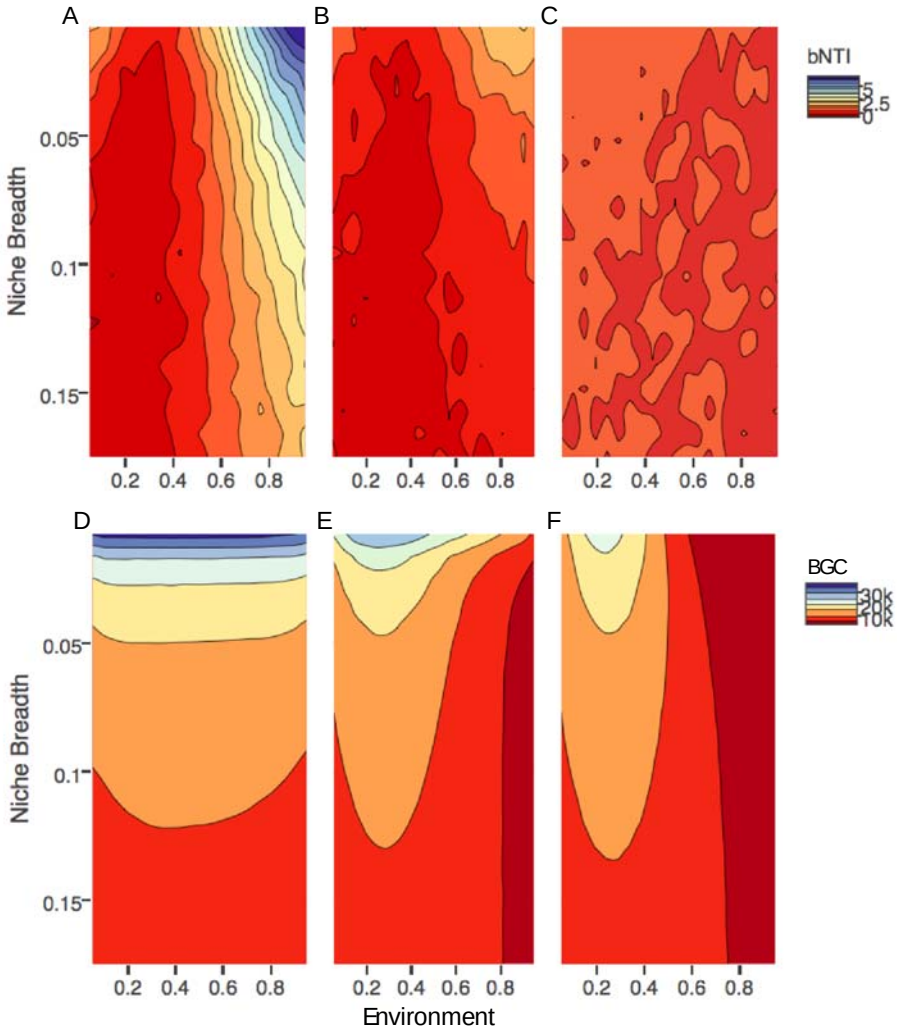


Figure 3. (a) Biogeochemical function varied across dispersal cases, with the lowest function in the case of homogenizing dispersal. Upper and lower hinges of the box plot represent the 75th and 25th percentiles and whiskers represent 1.5 times the 75th and 25th percentiles, respectively. Letters denote statistical significance. **(b)** In the selection-only case (black), biogeochemical function did not vary with changes in

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the environment. With moderate (blue) and homogenizing (red) dispersal, biogeochemical maxima were consistent with environmental optima of species in the source community (Figure 3). Beyond these maxima, function declined as the environment became more dissimilar to the source community.

We also observed that niche breadth of taxa in the receiving community was a key parameter in dictating biogeochemical function when selective pressures differed between source and receiving communities. In cases without dispersal, biogeochemical function was dictated entirely by niche breadth regardless of differences in selective pressures between source and receiving communities, as inferred from null models (Figure 4a, 4d). Selective pressures in the selection-only receiving communities were most dissimilar to the source community (βNTI > 2) in simulations with both narrow niche breadth and environmental conditions that were very different from the source community (Figure 4d). This relationship was also apparent (but weaker) in simulations with an intermediate amount of dispersal (Figure 4b). In receiving communities with high rates of dispersal, stochasticity (|βNTI| < 2) was the dominant process regardless of the niche breadth or environmental condition (Figure 4f). When environmental conditions were dissimilar to the source community, biogeochemical function increased with the relative influence of deterministic assembly, as inferred by βNTI (Figure 5). This indicates that βNTI may be an effective tool for predicting biogeochemical function across environments.



bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4. Average biogeochemical function **(a–c)** and selective environments **(d–f)** are displayed for each dispersal case across all parameter combinations. Biogeochemical function was entirely explained by niche breadth in selection-only scenarios **(a)** vs. a combination of niche breadth and environment in dispersal cases **(b–c)**. In dispersal cases **(b–c)**, biogeochemical maxima were observed when environmental conditions matched the source community and niche breadth was narrow. Selective pressures were primarily stochastic ($|\beta NTI| = 2$) in both dispersal cases **(e–f)**. Without dispersal **(d)**, the selective environment was most dissimilar to the source community when niche breadth was smallest and the environment was most dissimilar to the source community.

Further, broader niches and thus weaker selection led to lower rates of biogeochemical function, supporting a tradeoff between communities comprised of specialist vs. generalist species [14,55,56]. Previous work in microbial systems has posited life-history tradeoffs between specialist vs. generalist species, whereby specialists expend more energy to establish their niches but function at higher levels once established [57]. Specialist species have also been found to be more sensitive to changes in the environment due to strong adaptation to their local environment, with generalists being more resilient but functioning at lower rates [58]. While we do not address temporal dynamics in our model, the separation of biogeochemical function based on niche breadth indicates a central role for the balance of specialist vs. generalist microorganisms within a community in determining function, regardless of prevailing environmental conditions.

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

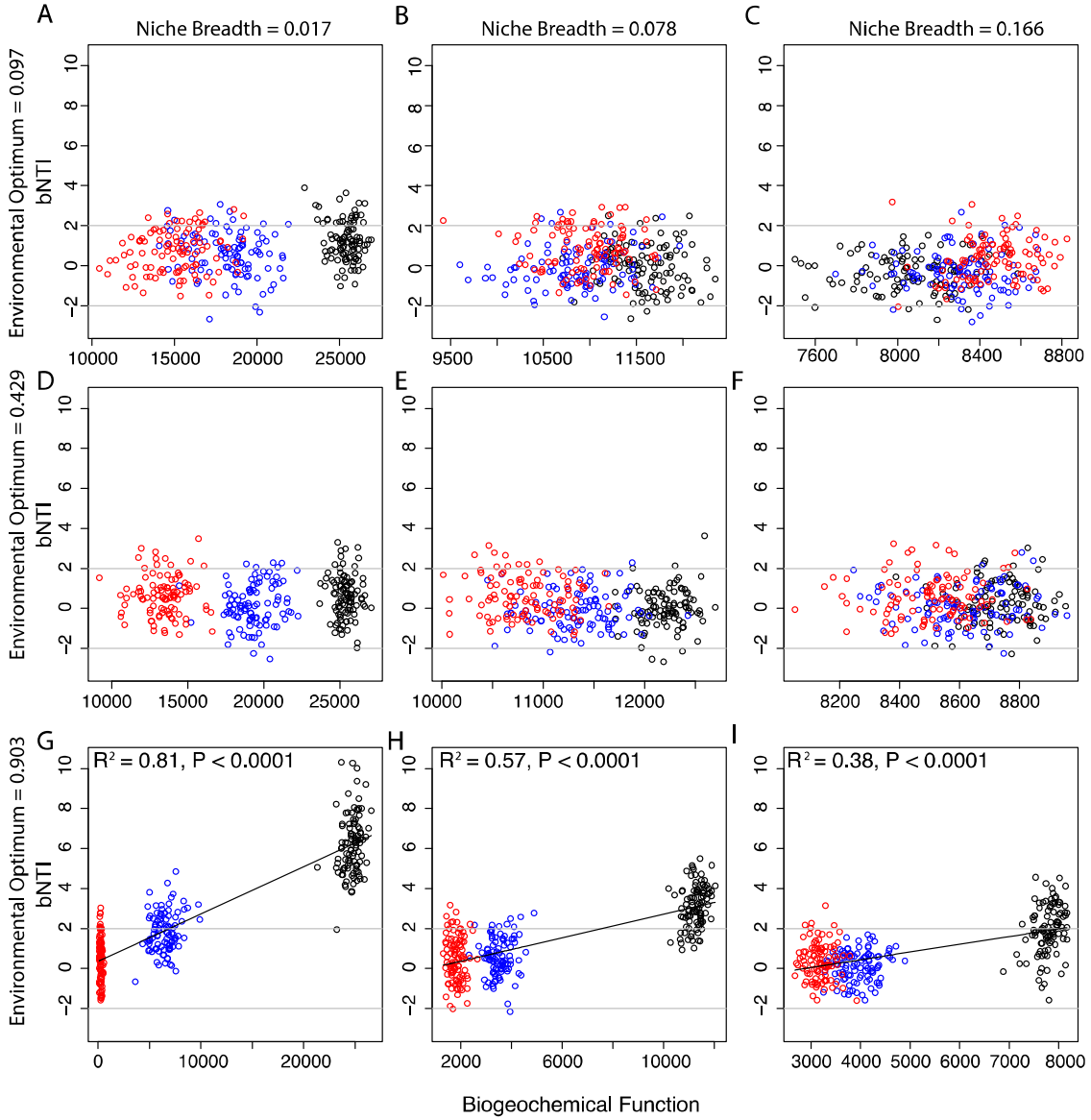


Figure 5. The relationship between βNTI and biogeochemical function varied across dispersal cases and parameter space. We show nine representative relationships in **(a–i)** to depict the range of variability in each parameter. Dispersal case is shown with color — selection-only is black, moderate dispersal is blue, and homogenizing dispersal is red. Niche breadth varies by column from small (1st) to large (3rd), and environment varies by row from more (top) to less (bottom) similar to the source community. βNTI values are mostly stochastic under similar and moderately different environmental conditions to the source community and do not correlate with biogeochemical function in these scenarios **(a–f)**. However, βNTI correlated with function when the environment was most dissimilar to the source community **(g–i)**, a relationship that declined with increasing niche breadth.

While our results suggest that maladapted immigrating organisms decrease biogeochemical function, it is important to note that stochasticity may offer buffering capacity that maintains or increases biogeochemical function relative to well-adapted deterministic communities in the context of future environmental perturbations not simulated here [58]. Stochastic spatial processes, such as dispersal, may

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lead to coexistence of species with different environmental optima resulting in a community that can rapidly adapt to changing environment conditions maintain biogeochemical function in the face of perturbation. Researchers have long demonstrated positive relationships between biodiversity and ecosystem function in both macrobial [59,60] and microbial [61-63] systems, and new work has highlighted the role of stochasticity in maintaining this connection [64]. Conversely, a lack of stochasticity may result in species so well-optimized to their environment that they are vulnerable to environmental change [58]. While these communities would putatively exhibit high rates of biogeochemical function at steady state, their function would plummet in response to perturbation, akin to observations of a tradeoff between function and vulnerability in plant communities [65,66].

Taken together, we reveal how stochastic community assembly can decrease adaptation to local environments and, in turn, decrease biogeochemical function. Our modelling approach demonstrates plausible outcomes of microbial assembly processes on ecosystem functioning, and integrating this knowledge with factors such as historical abiotic conditions, competitive dynamics, and life-history traits could substantially improve ecosystem model predictions. Microbially-explicit models (e.g., MIMICS, MEND) are rapidly becoming more sophisticated and are readily amenable to modules that represent ecological assembly processes [3,4]. We propose that new microbially-explicit models should go beyond microbial mechanisms at a given point in time or space, and building upon the foundation laid by Hawkes and Keitt (2015), incorporate ecological dynamics that operate across longer time scales to influence biogeochemical function.

Some models already offer the capability to incorporate ecological dynamics with only slight adjustments to their model structures. For instance, hydrologists have used PFLOTRAN to predict patterns of ground-water surface water mixing in river corridors [67], and this model has been recently extended to incorporate aerobic respiration [68]. Since the extended version of PFLOTRAN includes transport and biogeochemistry, additional structural elements could be included that represent the influence of dispersal (e.g., through advective transport) on biogeochemical rates. Several diffusion-based model frameworks also exist for terrestrial modelling efforts [e.g., 69] and could be integrated into larger scale community land models (CLM's) to adjust rate parameters. Alternatively, one could add a minimal amount of complexity to address dispersal processes in mechanistic ecological models such as Hawkes and Keitt's (2015) model by allowing organisms to be added from source communities. Resulting information could then be used either to offset rate calculations in biogeochemical models or to create simplified ecological modules to insert into larger scale ecosystem models.

Although there are many available avenues to merge modelling efforts in microbial ecology and ecosystem science, there is little debate that integrated models will increase the accuracy of predictions in novel environments. Future research should focus on identifying the full suite of ecological processes that are critical to biogeochemical cycling and incorporating these mechanisms into emerging model frameworks.

Acknowledgments: We thank Andrew Pitman for help editing manuscript text and Nathan Johnson for graphical assistance. This research was supported by the US Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of Subsurface Biogeochemical Research Program's Scientific Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). PNNL is operated for DOE by Battelle under contract DE-AC06-76RLO 1830. This research was performed using Institutional Computing at PNNL.

Author Contributions: E.B.G and J.C.S conceived and designed this work; E.B.G. performed the simulations, analyzed the data, and wrote the manuscript; J.C.S. contributed to manuscript revisions and code development.

Conflicts of Interest: The authors declare no conflicts of interest.

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Parton, W.J.; Stewart, J.W.; Cole, C.V. Dynamics of c, n, p and s in grassland soils: A model. *Biogeochemistry* **1988**, *5*, 109-131.
2. Schimel, D.; Parton, W.; Kittel, T.; Ojima, D.; Cole, C. Grassland biogeochemistry: Links to atmospheric processes. *Climatic Change* **1990**, *17*, 13-25.
3. Wieder, W.; Grandy, A.; Kallenbach, C.; Bonan, G. Integrating microbial physiology and physio-chemical principles in soils with the microbial-mineral carbon stabilization (mimics) model. *Biogeosciences* **2014**, *11*, 3899-3917.
4. Wang, G.; Post, W.M.; Mayes, M.A. Development of microbial-enzyme-mediated decomposition model parameters through steady-state and dynamic analyses. *Ecological Applications* **2013**, *23*, 255-272.
5. Davidson, E.A.; Samanta, S.; Caramori, S.S.; Savage, K. The dual arrhenius and michaelis–menten kinetics model for decomposition of soil organic matter at hourly to seasonal time scales. *Global Change Biology* **2012**, *18*, 371-384.
6. Sulman, B.N.; Phillips, R.P.; Oishi, A.C.; Shevliakova, E.; Pacala, S.W. Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated co2. *Nature Climate Change* **2014**, *4*, 1099.
7. Wang, G.; Jagadamma, S.; Mayes, M.A.; Schadt, C.W.; Steinweg, J.M.; Gu, L.; Post, W.M. Microbial dormancy improves development and experimental validation of ecosystem model. *The ISME journal* **2015**, *9*, 226.
8. Wieder, W.R.; Bonan, G.B.; Allison, S.D. Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change* **2013**, *3*, 909-912.
9. Wieder, W.; Grandy, A.; Kallenbach, C.; Taylor, P.; Bonan, G. Representing life in the earth system with soil microbial functional traits in the mimics model. *Geoscientific Model Development* **2015**, *8*, 1789-1808.
10. Allison, S. A trait-based approach for modelling microbial litter decomposition. *Ecology letters* **2012**, *15*, 1058-1070.
11. Buchkowski, R.W.; Bradford, M.A.; Grandy, A.S.; Schmitz, O.J.; Wieder, W.R. Applying population and community ecology theory to advance understanding of belowground biogeochemistry. *Ecology letters* **2017**, *20*, 231-245.
12. Graham, E.B.; Knelman, J.E.; Schindlbacher, A.; Siciliano, S.; Breulmann, M.; Yannarell, A.; Beman, J.; Abell, G.; Philippot, L.; Prosser, J. Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? *Frontiers in microbiology* **2016**, *7*.
13. Graham, E.B.; Crump, A.R.; Resch, C.T.; Fansler, S.; Arntzen, E.; Kennedy, D.W.; Fredrickson, J.K.; Stegen, J.C. Coupling spatiotemporal community assembly processes to changes in microbial metabolism. *Frontiers in Microbiology* **2016**, *7*.
14. Hawkes, C.V.; Keitt, T.H. Resilience vs. Historical contingency in microbial responses to environmental change. *Ecology letters* **2015**, *18*, 612-625.
15. Prosser, J.I.; Bohannan, B.J.; Curtis, T.P.; Ellis, R.J.; Firestone, M.K.; Freckleton, R.P.; Green, J.L.; Green, L.E.; Killham, K.; Lennon, J.J. The role of ecological theory in microbial ecology. *Nature reviews. Microbiology* **2007**, *5*, 384.
16. Hall, E.; Bernhardt, E.; Bier, R.; Bradford, M.; Boot, C.; Cotner, J.; del Giorgio, P.; Evans, S.; Graham, E.; Jones, S. Understanding how microbiomes influence the systems they inhabit: Insight from ecosystem ecology. *bioRxiv* **2016**, 065128.
17. Evans, S.E.; Wallenstein, M.D. Climate change alters ecological strategies of soil bacteria. *Ecology letters* **2014**, *17*, 155-164.
18. Wallenstein, M.D.; Hall, E.K. A trait-based framework for predicting when and where microbial adaptation to climate change will affect ecosystem functioning. *Biogeochemistry* **2012**, *109*, 35-47.
19. Martiny, J.B.; Jones, S.E.; Lennon, J.T.; Martiny, A.C. Microbiomes in light of traits: A phylogenetic perspective. *Science* **2015**, *350*, aac9323.
20. Bier, R.L.; Bernhardt, E.S.; Boot, C.M.; Graham, E.B.; Hall, E.K.; Lennon, J.T.; Nemergut, D.R.; Osborne, B.B.; Ruiz-González, C.; Schimel, J.P. Linking microbial community structure and microbial processes: An empirical and conceptual overview. *FEMS microbiology ecology* **2015**, *91*.

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

21. Rocca, J.D.; Hall, E.K.; Lennon, J.T.; Evans, S.E.; Waldrop, M.P.; Cotner, J.B.; Nemergut, D.R.; Graham, E.B.; Wallenstein, M.D. Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *The ISME journal* **2015**, *9*, 1693.
22. Evans, S.E.; Wallenstein, M.D. Soil microbial community response to drying and rewetting stress: Does historical precipitation regime matter? *Biogeochemistry* **2012**, *109*, 101-116.
23. Graham, E.B.; Crump, A.R.; Resch, C.T.; Fansler, S.; Arntzen, E.; Kennedy, D.W.; Fredrickson, J.K.; Stegen, J.C. Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. *Environmental Microbiology* **2017**, *19*, 1552-1567.
24. Fukami, T. Historical contingency in community assembly: Integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics* **2015**, *46*.
25. Martiny, J.B.; Martiny, A.C.; Weihe, C.; Lu, Y.; Berlemont, R.; Brodie, E.L.; Goulden, M.L.; Treseder, K.K.; Allison, S.D. Microbial legacies alter decomposition in response to simulated global change. *The ISME journal* **2017**, *11*, 490.
26. Fukami, T.; Dickie, I.A.; Paula Wilkie, J.; Paulus, B.C.; Park, D.; Roberts, A.; Buchanan, P.K.; Allen, R.B. Assembly history dictates ecosystem functioning: Evidence from wood decomposer communities. *Ecology Letters* **2010**, *13*, 675-684.
27. France, K.E.; Duffy, J. Diversity and dispersal interactively affect predictability of ecosystem function. *NATURE-LONDON-* **2006**, *441*, 1139.
28. Finlay, B.J. Global dispersal of free-living microbial eukaryote species. *Science* **2002**, *296*, 1061-1063.
29. Nemergut, D.R.; Schmidt, S.K.; Fukami, T.; O'Neill, S.P.; Bilinski, T.M.; Stanish, L.F.; Knelman, J.E.; Darcy, J.L.; Lynch, R.C.; Wickey, P. Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews* **2013**, *77*, 342-356.
30. Nemergut, D.R.; Shade, A.; Violle, C. When, where and how does microbial community composition matter? *Frontiers in microbiology* **2014**, *5*.
31. Pholchan, M.K.; Baptista, J.d.C.; Davenport, R.J.; Sloan, W.T.; Curtis, T.P. Microbial community assembly, theory and rare functions. *Frontiers in microbiology* **2013**, *4*.
32. Stegen, J.C.; Lin, X.; Fredrickson, J.K.; Konopka, A.E. Estimating and mapping ecological processes influencing microbial community assembly. *Frontiers in microbiology* **2015**, *6*.
33. Vellend, M. Conceptual synthesis in community ecology. *The Quarterly review of biology* **2010**, *85*, 183-206.
34. Bie, T.; Meester, L.; Brendonck, L.; Martens, K.; Goddeeris, B.; Ercken, D.; Hampel, H.; Denys, L.; Vanhecke, L.; Gucht, K. Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology letters* **2012**, *15*, 740-747.
35. Beisner, B.E.; Peres-Neto, P.R.; Lindström, E.S.; Barnett, A.; Longhi, M.L. The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology* **2006**, *87*, 2985-2991.
36. Heino, J.; Melo, A.S.; Siqueira, T.; Soininen, J.; Valanko, S.; Bini, L.M. Metacommunity organisation, spatial extent and dispersal in aquatic systems: Patterns, processes and prospects. *Freshwater Biology* **2015**, *60*, 845-869.
37. Bracken, L.; Wainwright, J.; Ali, G.; Tetzlaff, D.; Smith, M.; Reaney, S.; Roy, A. Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews* **2013**, *119*, 17-34.
38. Pringle, C.M. Hydrologic connectivity and the management of biological reserves: A global perspective. *Ecological Applications* **2001**, *11*, 981-998.
39. Gonzalez-Martin, C.; Teigell-Perez, N.; Valladares, B.; Griffin, D.W. The global dispersion of pathogenic microorganisms by dust storms and its relevance to agriculture. *Advances in Agronomy* **2014**, *127*, 1-41.
40. Smith, D.J.; Timonen, H.J.; Jaffe, D.A.; Griffin, D.W.; Birmele, M.N.; Perry, K.D.; Ward, P.D.; Roberts, M.S. Intercontinental dispersal of bacteria and archaea by transpacific winds. *Applied and environmental microbiology* **2013**, *79*, 1134-1139.
41. Bottos, E.M.; Woo, A.C.; Zawar-Reza, P.; Pointing, S.B.; Cary, S.C. Airborne bacterial populations above desert soils of the mcmurdo dry valleys, antarctica. *Microbial ecology* **2014**, *67*, 120-128.
42. Fierer, N.; Jackson, R.B. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 626-631.
43. Hollister, E.B.; Engledow, A.S.; Hammett, A.J.M.; Provin, T.L.; Wilkinson, H.H.; Gentry, T.J. Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments. *The ISME journal* **2010**, *4*, 829.

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

44. Casamayor, E.O.; Massana, R.; Benlloch, S.; Øvreås, L.; Díez, B.; Goddard, V.J.; Gasol, J.M.; Joint, I.; Rodríguez-Valera, F.; Pedrós-Alió, C. Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern. *Environmental Microbiology* **2002**, *4*, 338-348.
45. Hurlbert, A.H.; Stegen, J.C. When should species richness be energy limited, and how would we know? *Ecology letters* **2014**, *17*, 401-413.
46. Stegen, J.C.; Lin, X.; Fredrickson, J.K.; Chen, X.; Kennedy, D.W.; Murray, C.J.; Rockhold, M.L.; Konopka, A. Quantifying community assembly processes and identifying features that impose them. *The ISME journal* **2013**, *7*, 2069.
47. Webb, C.O.; Ackerly, D.D.; Kembel, S.W. Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* **2008**, *24*, 2098-2100.
48. Fine, P.V.; Kembel, S.W. Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western amazonian tree communities. *Ecography* **2011**, *34*, 552-565.
49. Stegen, J.C.; Lin, X.; Konopka, A.E.; Fredrickson, J.K. Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME journal* **2012**, *6*, 1653.
50. Dini-Andreote, F.; Stegen, J.C.; van Elsas, J.D.; Salles, J.F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences* **2015**, *112*, E1326-E1332.
51. DeAngelis, K.M.; Silver, W.L.; Thompson, A.W.; Firestone, M.K. Microbial communities acclimate to recurring changes in soil redox potential status. *Environmental Microbiology* **2010**, *12*, 3137-3149.
52. Cregger, M.A.; Schadt, C.W.; McDowell, N.G.; Pockman, W.T.; Classen, A.T. Response of the soil microbial community to changes in precipitation in a semiarid ecosystem. *Applied and environmental Microbiology* **2012**, *78*, 8587-8594.
53. Green, J.; Bohannan, B.J. Spatial scaling of microbial biodiversity. *Trends in ecology & evolution* **2006**, *21*, 501-507.
54. Schimel, D.S.; Braswell, B.; McKeown, R.; Ojima, D.S.; Parton, W.; Pulliam, W. Climate and nitrogen controls on the geography and timescales of terrestrial biogeochemical cycling. *Global Biogeochemical Cycles* **1996**, *10*, 677-692.
55. Wilson, D.S.; Yoshimura, J. On the coexistence of specialists and generalists. *The American Naturalist* **1994**, *144*, 692-707.
56. Kassen, R. The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of evolutionary biology* **2002**, *15*, 173-190.
57. Lennon, J.T.; Aanderud, Z.T.; Lehmkuhl, B.; Schoolmaster, D.R. Mapping the niche space of soil microorganisms using taxonomy and traits. *Ecology* **2012**, *93*, 1867-1879.
58. Shade, A.; Peter, H.; Allison, S.D.; Baho, D.L.; Berga, M.; Bürgmann, H.; Huber, D.H.; Langenheder, S.; Lennon, J.T.; Martiny, J.B. Fundamentals of microbial community resistance and resilience. *Frontiers in microbiology* **2012**, *3*.
59. Cardinale, B.J.; Matulich, K.L.; Hooper, D.U.; Byrnes, J.E.; Duffy, E.; Gamfeldt, L.; Balvanera, P.; O'Connor, M.I.; Gonzalez, A. The functional role of producer diversity in ecosystems. *American journal of botany* **2011**, *98*, 572-592.
60. Hooper, D.U.; Adair, E.C.; Cardinale, B.J.; Byrnes, J.E.; Hungate, B.A.; Matulich, K.L.; Gonzalez, A.; Duffy, J.E.; Gamfeldt, L.; O'Connor, M.I. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* **2012**, *486*, 105-108.
61. Bell, T.; Newman, J.A.; Silverman, B.W.; Turner, S.L.; Lilley, A.K. The contribution of species richness and composition to bacterial services. *Nature* **2005**, *436*, 1157.
62. Langenheder, S.; Bulling, M.T.; Solan, M.; Prosser, J.I. Bacterial biodiversity-ecosystem functioning relations are modified by environmental complexity. *PloS one* **2010**, *5*, e10834.
63. Levine, U.Y.; Teal, T.K.; Robertson, G.P.; Schmidt, T.M. Agriculture's impact on microbial diversity and associated fluxes of carbon dioxide and methane. *The ISME journal* **2011**, *5*, 1683.
64. Knelman, J.E.; Nemergut, D.R. Changes in community assembly may shift the relationship between biodiversity and ecosystem function. *Frontiers in microbiology* **2014**, *5*.
65. Tilman, D. The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* **1999**, *80*, 1455-1474.

bioRxiv preprint doi: https://doi.org/10.1101/183897; this version posted September 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

66. Tilman, D.; Reich, P.B.; Knops, J.; Wedin, D.; Mielke, T.; Lehman, C. Diversity and productivity in a long-term grassland experiment. *Science* **2001**, *294*, 843-845.
67. Mills, R.T.; Lu, C.; Lichtner, P.C.; Hammond, G.E. In *Simulating subsurface flow and transport on ultrascale computers using pflotran*, Journal of physics: conference series, 2007; IOP Publishing: p 012051.
68. Song, X.; Chen, X.; Stegen, J.; Hammond, G.; Song, H.-S.; Dai, H.; Graham, E.; Zachara, J.M. Drought conditions maximize the impact of high-frequency flow variations on thermal regimes and biogeochemical function in the hyporheic zone. **in review**.
69. Manzoni, S.; Moyano, F.; Kätterer, T.; Schimel, J. Modeling coupled enzymatic and solute transport controls on decomposition in drying soils. *Soil Biology and Biochemistry* **2016**, *95*, 275-287.