Soil Biology & Biochemistry 68 (2014) 279–282

Contents lists available at ScienceDirect

Soil Biology & Biochemistry

journal homepage: www.elsevier.com/locate/soilbio

Short communication

Do we need to understand microbial communities to predict ecosystem function? A comparison of statistical models of nitrogen cycling processes

Emily B. Graham [a, b], William R. Wieder [b, c], Jonathan W. Leff [d, e], Samantha R. Weintraub [b, e], Alan R. Townsend [b, e], Cory C. Cleveland [f], Laurent Philippot [g], Diana R. Nemergut [a, b, *]

[a] Environmental Studies Program, University of Colorado, UCB 397, Boulder, CO 80309-0397, USA
[b] Institute for Arctic and Alpine Research, University of Colorado, UCB 450, Boulder, CO 80309-0450, USA
[c] University Corporation for Atmospheric Research, TSS, CGD/NCAR, PO Box 3000, Boulder, CO 80307-3000, USA
[d] Cooperative Institute for Research in Environmental Sciences, University of Colorado, UCB 216, Boulder, CO 80309-0216, USA
[e] Department of Ecology and Evolutionary Biology, University of Colorado, UCB 334, Boulder, CO 80309-0334, USA
[f] Department of Ecosystems and Conservation Sciences, University of Montana, 32 Campus Dr., Missoula, MT 59812, USA
[g] INRA, UMR 1347 Agroecology, 21000 Dijon, France

ARTICLE INFO

Article history:
Received 23 March 2013
Received in revised form
22 August 2013
Accepted 28 August 2013
Available online 10 September 2013

Keywords:
Modeling
Functional genes
Biogeochemistry
qPCR
Nitrification
Dissimilatory nitrate reduction to ammonium
N_2O flux
Structure–function relationship
Ecosystem processes

ABSTRACT

Despite the central role of microorganisms in biogeochemistry, process models rarely explicitly account for variation in communities. Here, we use statistical models to address a fundamental question in ecosystem ecology: do we need to better understand microbial communities to accurately predict ecosystem function? Nitrogen (N) cycle process rates and associated gene abundances were measured in tropical rainforest soil samples collected in May (early wet season) and October (late wet season). We used stepwise linear regressions to examine the explanatory power of edaphic factors and functional gene relative abundances alone and in combination for N-cycle processes, using both our full dataset and seasonal subsets of the data. In our full dataset, no models using gene abundance data explained more variation in process rates than models based on edaphic factors alone, and models that contained both edaphic factors and community data did not explain significantly more variation in process rates than edaphic factor models. However, when seasonal datasets were examined separately, microbial predictors enhanced the explanatory power of edaphic predictors on dissimilatory nitrate reduction to ammonium and N_2O efflux rates during October. Because there was little variation in the explanatory power of microbial predictors alone between seasonal datasets, our results suggest that environmental factors we did not measure may be more important in structuring communities and regulating processes in October than in May. Thus, temporal dynamics are key to understanding the relationships between edaphic factors, microbial communities and ecosystem function in this system. The simple statistical method presented here can accommodate a variety of data types and should help prioritize what forms of data may be most useful in ecosystem model development.

Information on microbial communities is rarely explicitly considered in large-scale ecosystem models. Instead, most such models implicitly assume that microbial activity can be represented by mathematical equations that apply across diverse environments (Todd-Brown et al., 2012). However, recent work supports predictive relationships between microbial traits and ecosystem function (Follows et al., 2007; Allison, 2012). Thus, a fundamental question for ecosystem ecology remains widely debated: When do we need to understand details about microbial communities to accurately predict process (*e.g.*, Carney and Matson, 2005; van der Heijden et al., 2008; Leff et al., 2012; Petersen et al., 2012)?

In particular, the *added* value of data on microbial traits – or the predictive power of data on microbial community traits above and

* Corresponding author. Present address: Institute for Arctic and Alpine Research, University of Colorado, UCB 450, Boulder, CO 80309-0450, USA. Tel.: +1 303 492 6387.
E-mail addresses: Emily.Graham@colorado.edu (E.B. Graham), Diana.Nemergut@colorado.edu (D.R. Nemergut).

0038-0717/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.soilbio.2013.08.023

Table 1
Multiple linear regressions were constructed with three sets of predictors for each process, (1) edaphic variables, (2) relative gene abundances, and (3) both edaphic variables and relative gene abundances. Predictors and statistical results of best-fit models are presented in Table 1. Edaphic models versus overall models were compared using ANOVA, and the results are also presented below.

Process		Edaphic model	Relative gene abundance model	Overall model	Edaphic – overall model comparison
Nitrification	Predictors	NH_4^+_pool, NH_4^+_pool*pH, pH	amoA	amoA, NH_4^+_pool, pH*amoA, pH	$F(19,18) = 3.37$ $p = 0.08$
	Statistical Results	Model $p = 0.007$ Adj. $R^2 = 0.38$	Model $p = 0.06$ Adj. $R^2 = 0.12$	Model $p = 0.005$ Adj. $R^2 = 0.45$	
DNRA	Predictors	ln(soil_moist), pH, NO_3^-_pool, ln(soil_moist)*pH	napA, narG	ln(soil_moist), NO_3^-_pool, NO_3^-_pool*pH, pH, narG, pH*ln(soil_moist)	$F(20,18) = 3.29$ $p = 0.06$
	Statistical Results	Model $p = 0.001$ Adj. $R^2 = 0.49$	Model $p = 0.002$ Adj. $R^2 = 0.39$	Model $p = 0.0008$ Adj. $R^2 = 0.59$	
$^{15}N_2O$ Efflux	Predictors	NO_3^-_pool, pH, ln(soil_moist)	nirS*nosZ, nirS, nosZ, nirK	pH, nirS, NO_3^-_pool*nirK, NO_3^-_pool, nirK, ln(soil_moist)*nosZ, ln(soil_moist), nosZ	$F(19,14) = 2.08$ $p = 0.13$
	Statistical Results	Model $p = 0.0008$ Adj. $R^2 = 0.51$	Model $p = 0.0008$ Adj. $R^2 = 0.55$	Model $p = 0.003$ Adj. $R^2 = 0.62$	

beyond that of environmental factors alone – has not yet been explicitly considered. Schimel (2001) noted that many process-based models *implicitly* consider microorganisms by accounting for variation in factors that regulate microbial community composition, such as pH (Fierer et al., 2007), moisture (Nemergut et al., 2010), substrate availability (Legg et al., 2012), temperature (Shade et al., 2012) and salinity (Lozupone et al., 2007). Yet, communities are not entirely determined by abiotic variables, as factors including dormancy (Jones and Lennon, 2010; Lennon and Jones, 2011), priority effects (Fukami, 2004), and neutral community assembly processes (Ferrenberg et al., 2013; Nemergut et al., 2013) can also structure communities. The degree to which such factors affect the composition, functional traits, and activity of a given microbial community will affect the value of microbial data in predicting ecosystem processes beyond that of environmental factors alone.

Here, we used statistical models to compare the power of edaphic factors to predict soil microbial processes with and without data on microbial traits. Because data on microbial traits can provide a more accurate representation of functional potential than data on overall community structure (Polz et al., 2006; Burke et al., 2011), we used quantitative polymerase chain reaction (qPCR) data on functional genes for this analysis. We focused on genes involved in nitrogen (N) cycling as well as measurements of nitrification, dissimilatory nitrate reduction to ammonium (DNRA), and nitrous oxide (N_2O) emission rates determined using 15-N tracers. All data were generated from soils collected in May (early wet season) and October (late wet season) from a lowland tropical forest on the Osa Peninsula, Costa Rica (8°43′ N, 83°37′ W; Wieder et al., 2013). Abundances of genes involved in nitrification (bacterial and *Thaumarchaeota amoA*), nitrate reduction (*narG* and *napA*), nitrogen fixation (*nifH*), and denitrification (a likely source of N_2O emissions; *nirS*, *nirK*, and *nosZ*) were used as proxies of microbial trait abundances, as described by Wieder et al. (2013). Edaphic factors, including pH, moisture, NO_3^- and NH_4^+ pools, and total C and N content were collected to describe environmental conditions (Wieder et al., 2013). Because only a subset of the data for which we had qPCR data were used, some of the relationships identified here vary slightly from those presented in Wieder et al. (2013).

Three sets of multiple linear regressions were fit to the data to explain rates of each N-cycle process: (1) models with edaphic predictors only; (2) models with gene abundance predictors (*narG*, *napA*, *nifH*, *nirS*, *nirK*, and *nosZ* relative to bacterial 16S rRNA gene abundance and *amoA* relative to bacterial + *Thaumarchaeota* 16S rRNA gene abundance) only; and (3) models with both edaphic and gene relative abundance predictors. Comparisons between edaphic and overall models were conducted using a partial ANOVA to compare the sum of squared errors for each model, and to determine if models with different predictors were significantly different ($\alpha = 0.05$). Finally, we used linear regression to compare the residuals of the best-fit edaphic models and individual gene relative abundance predictors to determine if microbial predictors explained a different proportion of the variance in process than edaphic factors alone. We performed analyses on samples collected during May and October separately as well as on the entire dataset together to examine the effect of temporal dynamics on the relationships between edaphic factors, microbial communities and N-cycling processes.

When examining our data across both seasons combined, we found that edaphic factors yielded more explanatory power than

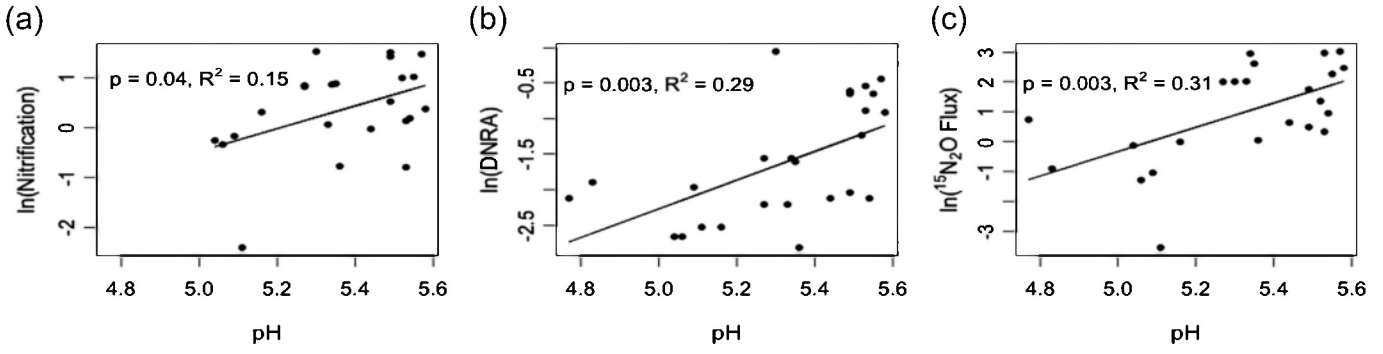

Fig. 1. Univariate models were constructed evaluate the effect of pH alone on (a) nitrification, (b) DNRA, and (c) $^{15}N_2O$ efflux rates. pH was the strongest individual edaphic predictor of all N-cycle processes.

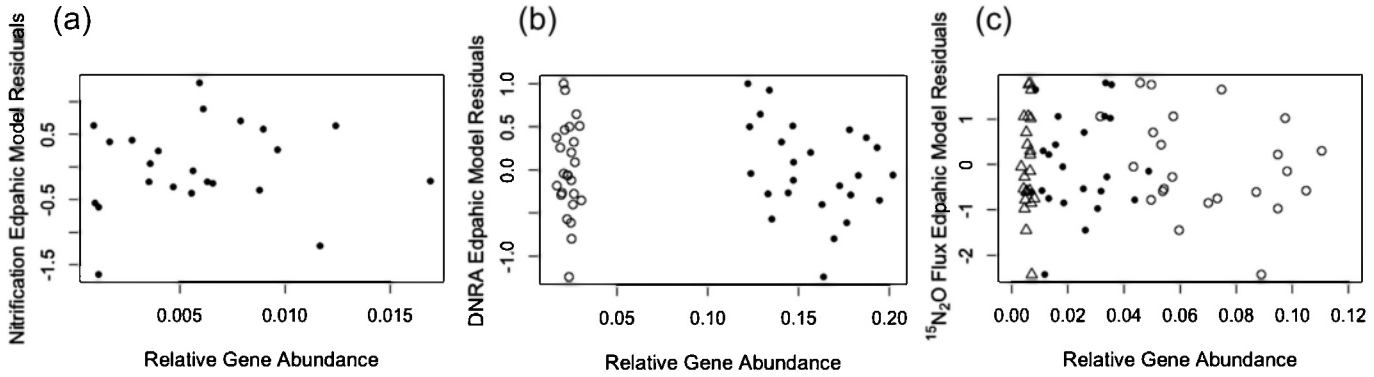

Fig. 2. The explanatory power of relative gene abundances on edaphic model residuals was examined using linear regressions in our full dataset for (a) nitrification, (b) DNRA, and (c) $^{15}N_2O$ efflux rates. In panel (a), *amoA* abundance is represented by filled dots. In panel (b), *narG* abundance is represented by filled dots and *napA* abundance is represented by unfilled dots. In panel (c), filled dots represent *nirS* abundance, unfilled dots represent *nirK* abundance, and unfilled triangles represent *nosZ* abundance. Relative gene abundances did not explain any residual variation in edaphic models, suggesting that gene abundances explained the same portion of variance in process as edaphic predictors.

microbial predictors for nitrification (Adj. R^2 of 0.38 vs. 0.12), but that the explanatory power of edaphic versus microbial predictors did not vary for DNRA or N_2O emissions (Table 1). The explanatory power of both edaphic and gene abundance predictors on rates was lowest for nitrification, which could indicate the presence of unmeasured edaphic factors and/or undetectable nitrification genes (Hatzenpichler, 2012). In the full dataset, we did not observe any significant increase in explanatory power when microbial predictors were added to edaphic models for any process (Table 1). A similar relationship was observed by Attard et al. (2011), who showed that denitrification rates in an agricultural system were primarily determined by soil characteristics rather than microbial community structure or functional composition. Furthermore, when we considered the effect of each variable independently, we found that pH was the strongest individual predictor of all N-cycle processes and was correlated with all N-cycling gene abundances, suggesting that pH was the most important factor in determining both process rates and microbial functional traits across seasons (Fig. 1). Finally, data on gene relative abundances did not explain any residual variation in process rates constructed from edaphic factors alone (Fig. 2).

Interestingly, when we analyzed samples from each season separately, we found that all models yielded higher explanatory power in the late wet season than we observed in the full dataset (Table 2 vs. Table 1). Moreover, including both functional gene relative abundances and edaphic factors in DNRA and N_2O efflux models significantly improved our explanatory power relative to models with only edaphic predictors for the October dataset (Table 2). Our statistical power was much lower in the early wet season than in the late wet season ($n = 8$ and $n = 17$, respectively), and only one model – an edaphic model for DNRA rates – explained significant variation in process rates in May (Adj. $R^2 = 0.72$); no other model for any process or predictor set yielded a significant relationship in the May dataset (data not shown).

While some of the observed seasonal differences are undoubtedly related to differences in sample size, our analysis shows a difference between the explanatory power of models for the late wet season data as compared to the entire dataset. Furthermore, the explanatory power of microbial predictors alone for process rates did not vary between our October subset and the full dataset (Tables 1 and 2), suggesting that a decrease in the importance of edaphic factors, not an increase in the importance of microbial communities, drove the added value of microbial data in October. Together our results suggest that environmental factors that we did not measure are more important in structuring communities in the late wet season than in May. October soils featured higher total abundances of DNRA and denitrification genes (data not shown), supporting previous work that shows that anoxic processes are more important in this season than in the early wet season (Wieder et al., 2011). It is possible that soil oxygen or some other

Table 2
Multiple linear regressions were constructed with three sets of predictors for each process in October and May, (1) edaphic variables, (2) relative gene abundances, and (3) both edaphic variables and relative gene abundances. Predictors and statistical results of best-fit models for our October subset are presented in Table 2, as only one model was significant in May. Edaphic models versus overall models were compared using ANOVA, and the results are also presented below.

Process		Edaphic model	Relative gene abundance model	Overall model	Edaphic – overall model comparison
Nitrification	Predictors	pH, NH_4^+_pool	amoA	pH*ln(amoA), pH, ln(amoA), NH_4^+_pool	$F(14,12) = 1.18$ $p > 0.20$
	Statistical Results	Model $p < 0.001$ Adj. $R^2 = 0.70$	Model $p = 0.06$ Adj. $R^2 = 0.16$	Model $p < 0.001$ Adj. $R^2 = 0.70$	
DNRA	Predictors	soil_moist*NO_3^-_pool, soil_moist, NO_3^-_pool	napA, narG, napA*narG	soil_moist*NO_3^-_pool, soil_moist, NO_3^-_pool, napA, narG, napA*narG	$F(13,9) = 5.07$ $p = 0.02$
	Statistical Results	Model $p < 0.001$ Adj. $R^2 = 0.83$	Model $p = 0.03$ Adj. $R^2 = 0.38$	Model $p < 0.001$ Adj. $R^2 = 0.92$	
$^{15}N_2O$ Efflux	Predictors	pH*NO_3^-_pool, pH, NO_3^-_pool, soil_moist	nirS*nosZ, nirK, nirS, noZ	pH*NO_3^-_pool, pH, NO_3^-_pool, soil_moist*nosZ, soil_moist, nosZ, nirK, nirS	$F(12,8) = 6.56$ $p = 0.01$
	Statistical Results	Model $p = 0.004$ Adj. $R^2 = 0.60$	Model $p = 0.004$ Adj. $R^2 = 0.62$	Model $p < 0.001$ Adj. $R^2 = 0.86$	

unmeasured factor (e.g., Ehrenfeld, 2003; Fortuna et al., 2012) is a more important driver of communities as well as process in October as compared to May. Thus, had we measured all potential regulatory edaphic factors, the added value of microbial data in October may not have been significant.

Our results suggest that the added value of microbial data for explaining function over edaphic factors alone varies seasonally and by N-cycle process in this tropical rainforest ecosystem. Broad conclusions on the need for directly modeling community structure will require further examination across ecosystem types and processes, at a range of temporal scales. Specifically, ecosystem models may need to incorporate seasonal changes in the drivers of processes to most accurately predict ecosystem function. Thus, the simple statistical method presented here, which can be readily extended to accommodate a wide variety of ecological data across spatial and temporal scales, can be used to evaluate the utility of data types in disparate ecosystems.

Acknowledgments

This work was supported by grant numbers DEB-0910012, DEB-0515744, DEB-0852916 and DEB-1221215 from the National Science Foundation to W.W., A.T., C. C. and D. N.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.soilbio.2013.08.023.

References

Allison, S.D., 2012. A trait-based approach for modelling microbial litter decomposition. Ecol. Lett. 15, 1058–1070.
Attard, E., Recous, S., Chabbi, A., De Berranger, C., Guillaumaud, N., Labreuche, J., Philippot, L., Schmid, B., Le Roux, X., 2011. Soil environmental conditions rather than denitrifier abundance and diversity drive potential denitrification after changes in land uses. Glob. Change Biol. 17, 1975–1989.
Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., Thomas, T., 2011. Bacterial community assembly based on functional genes rather than species. Proc. Natl. Acad. Sci. U. S. A. 108, 14288–14293.
Carney, K.M., Matson, P.A., 2005. Plant communities, soil microorganisms, and soil carbon cycling: does altering the world belowground matter to ecosystem functioning? Ecosystems 8, 928–940.
Ehrenfeld, J.G., 2003. Effects of exotic plant invasions on soil nutrient cycling processes. Ecosystems 6, 503–523.
Ferrenberg, S., O'Neill, S.P., Knelman, J.E., Todd, B., Duggan, S., Bradley, D., Robinson, T., Schmidt, S.K., Townsend, A.R., Williams, M.W., 2013. Changes in assembly processes in soil bacterial communities following a wildfire disturbance. ISME J 7 (6), 1102–1111.
Fierer, N., Morse, J.L., Berthrong, S.T., Bernhardt, E.S., Jackson, R.B., 2007. Environmental controls on the landscape-scale biogeography of stream bacterial communities. Ecology 88, 2162–2173.
Follows, M.J., Dutkiewicz, S., Grant, S., Chisholm, S.W., 2007. Emergent biogeography of microbial communities in a model ocean. Science 315, 1843–1846.
Fortuna, A.M., Honeycutt, C.W., Vandemark, G., Griffin, T.S., Larkin, R.P., He, Z.Q., Wienhold, B.J., Sistani, K.R., Albrecht, S.L., Woodbury, B.L., Torbert, H.A., Powell, J.M., Hubbard, R.K., Eigenberg, R.A., Wright, R.J., Alldredge, J.R., Harsh, J.B., 2012. Links among nitrification, nitrifier communities, and edaphic properties in contrasting soils receiving dairy slurry. J. Environ. Qual. 41, 262–272.
Fukami, T., 2004. Assembly history interacts with ecosystem size to influence species diversity. Ecology 85, 3234–3242.
Hatzenpichler, R., 2012. Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. Appl. Environ. Microbiol. 78, 7501–7510.
Jones, S.E., Lennon, J.T., 2010. Dormancy contributes to the maintenance of microbial diversity. Proc. Natl. Acad. Sci. U. S. A. 107, 5881–5886.
Leff, J.W., Nemergut, D.R., Grandy, A.S., O'Neill, S.P., Wickings, K., Townsend, A.R., Cleveland, C.C., 2012. The effects of soil bacterial community structure on decomposition in a tropical rain forest. Ecosystems 15, 284–298.
Legg, T.M., Zheng, Y., Simone, B., Radloff, K.A., Mladenov, N., Gonzalez, A., Knights, D., Siu, H.C., Rahman, M.M., Ahmed, K.M., 2012. Carbon, metals, and grain size correlate with bacterial community structure in sediments of a high arsenic aquifer. Front. Microbiol. 3.
Lennon, J.T., Jones, S.E., 2011. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat. Rev. Microbiol. 9, 119–130.
Lozupone, C.A., Hamady, M., Kelley, S.T., Knight, R., 2007. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Appl. Environ. Microbiol. 73, 1576–1585.
Nemergut, D.R., Cleveland, C.C., Wieder, W.R., Washenberger, C.L., Townsend, A.R., 2010. Plot-scale manipulations of organic matter inputs to soils correlate with shifts in microbial community composition in a lowland tropical rain forest. Soil Biol. Biochem. 42, 2153–2160.
Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Legg, T.M., Stanish, L., Knelman, J.E., Darcy, J.L., Lynch, R., Wickey, P., Ferrenberg, S., 2013. Microbial community assembly: patterns and processes. Microbiol. Mol. Biol. Rev. 77, 342–356.
Petersen, D.G., Blazewicz, S.J., Firestone, M., Herman, D.J., Turetsky, M., Waldrop, M., 2012. Abundance of microbial genes associated with nitrogen cycling as indices of biogeochemical process rates across a vegetation gradient in Alaska. Environ. Microbiol. 14, 993–1008.
Polz, M.F., Hunt, D.E., Preheim, S.P., Weinreich, D.M., 2006. Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. Philos. Trans. R. Soc. Lond. B Biol. Sci. 361, 2009–2021.
Schimel, J.P., 2001. In: Schulze, E.D. (Ed.), Global Biogeochemical Cycles in the Climate System. Academic, San Diego, CA, pp. 239–254.
Shade, A., Read, J.S., Youngblut, N.D., Fierer, N., Knight, R., Kratz, T.K., Lottig, N.R., Roden, E.E., Stanley, E.H., Stombaugh, J., Whitaker, R.J., Wu, C.H., McMahon, K.D., 2012. Lake microbial communities are resilient after a whole-ecosystem disturbance. ISME J. 6, 2153–2167.
Todd-Brown, K.E.O., Hopkins, F.M., Kivlin, S.N., Talbot, J.M., Allison, S.D., 2012. A framework for representing microbial decomposition in coupled climate models. Biogeochemistry 109, 19–33.
van der Heijden, M.G.A., Bardgett, R.D., van Straalen, N.M., 2008. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecol. Lett. 11, 296–310.
Wieder, W.R., Cleveland, C.C., Townsend, A.R., 2011. Throughfall exclusion and leaf litter addition drive higher rates of soil nitrous oxide emissions from a lowland wet tropical forest. Glob. Change Biol. 17, 3195–3207.
Wieder, W.R., Cleveland, C.C., Taylor, P.G., Nemergut, D.R., Hinckley, E.-L., Philippot, L., Bru, D., Weintraub, S.R., Martin, M., Townsend, A.R., 2013. Experimental removal and addition of leaf litter inputs reduces nitrate production and loss in a lowland tropical forest. Biogeochemistry 113, 629–642.