

# Mycorrhiza-mediated competition between plants and decomposers drives soil carbon storage

Colin Averill<sup>1</sup>, Benjamin L. Turner<sup>2</sup> & Adrien C. Finzi<sup>3</sup>

Soil contains more carbon than the atmosphere and vegetation combined<sup>1</sup>. Understanding the mechanisms controlling the accumulation and stability of soil carbon is critical to predicting the Earth's future climate<sup>2,3</sup>. Recent studies suggest that decomposition of soil organic matter is often limited by nitrogen availability to microbes<sup>4–6</sup> and that plants, via their fungal symbionts, compete directly with free-living decomposers for nitrogen<sup>6,7</sup>. Ectomycorrhizal and ericoid mycorrhizal (EEM) fungi produce nitrogen-degrading enzymes, allowing them greater access to organic nitrogen sources than arbuscular mycorrhizal (AM) fungi<sup>8–10</sup>. This leads to the theoretical prediction that soil carbon storage is greater in ecosystems dominated by EEM fungi than in those dominated by AM fungi<sup>11</sup>. Using global data sets, we show that soil in ecosystems dominated by EEM-associated plants contains 70% more carbon per unit nitrogen than soil in ecosystems dominated by AM-associated plants. The effect of mycorrhizal type on soil carbon is independent of, and of far larger consequence than, the effects of net primary production, temperature, precipitation and soil clay content. Hence the effect of mycorrhizal type on soil carbon content holds at the global scale. This finding links the functional traits of mycorrhizal fungi to carbon storage at ecosystem-to-global scales, suggesting that plant–decomposer competition for nutrients exerts a fundamental control over the terrestrial carbon cycle.

Nitrogen (N) availability influences biosphere–atmosphere exchanges of carbon (C) by limiting C inputs to the soil from net primary production<sup>12</sup> (NPP), and C outputs associated with the activity of decomposer microbes<sup>4</sup>. Most plant species on the Earth associate with symbiotic mycorrhizal fungi to acquire nutrients from soil. EEM fungi produce a wide range of enzymes that release N from soil organic matter<sup>13</sup>, whereas AM fungi lack these enzyme systems<sup>10,14</sup>. Accordingly, EEM-associated plants (EEM plants) acquire substantially more organic N from the soil than do AM-associated plants (AM plants)<sup>9,10,15</sup>, and also compete directly for organic N with other free-living decomposer microbes in the soil. A recent theoretical model suggests that uptake of organic N by EEM plants slows the rate of decomposition and increases soil C storage by exacerbating the nitrogen limitation of free-living decomposer activity and their production of enzymes that degrade soil organic matter<sup>11</sup>. There is as yet little support for this contention.

We used a mixed effects model to test the hypothesis that ecosystems dominated by EEM fungi (EEM ecosystems) store significantly more soil C than do ecosystems dominated by AM fungi (AM ecosystems) after accounting for variation in soil N and other drivers of soil C storage. We assembled a global data set containing soil C, N and clay content to a depth of one metre, as well as site-specific vegetation descriptions to determine biome and mycorrhizal type (Table 1). We then used global data products to assign mean annual temperature (MAT), mean annual precipitation (MAP)<sup>15</sup>, and NPP<sup>16</sup> to determine whether the effect of mycorrhizal type on soil C storage was statistically significant after accounting for variations in biome type and biophysical properties assumed to control decomposition in ecosystem and Earth system models<sup>17</sup> (see Methods Summary). The statistical analysis included soil clay content because such secondary minerals have the potential to sorb and stabilize soil organic carbon<sup>18</sup>.

We found that EEM ecosystems store 1.7 times more C per unit of soil N than do AM ecosystems (Fig. 1). The most parsimonious, corrected Akaike Information Criterion (AICc)-selected model (mycorrhizal type  $\times$  N interaction,  $P < 0.0001$ , AICc-selected model  $R^2_{LR} = 0.91$ ; Fig. 1 and Methods) supported the removal of climate variables, NPP and clay content, which were weakly correlated with soil C content, though the biome type remained in the model as a random effect (Fig. 2). The mycorrhizal  $\times$  N interaction remained significant even when all predictors were included in the model. This result shows that mycorrhizal status exerts a far larger control over soil C content than do climate variables, NPP or clay content. Weak relationships between NPP, climate and soil C storage at the global scale have also been reported elsewhere<sup>19</sup>.

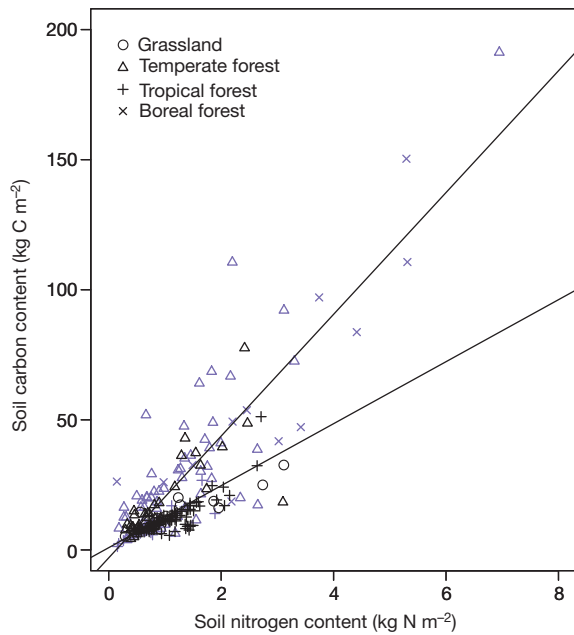
We conducted a sensitivity analysis of the model to ensure the findings were robust to the exclusion of biomes that contained only a single mycorrhizal type, surface organic horizons (which represent a large fraction of surface soil C content) in cold climates and EEM data points whose soil N content was greater than the largest observation found in AM ecosystems (see Extended Data Figs 1–5 and Extended Data Tables 1–5). In all cases a mycorrhizal effect was retained. This suggests that current model formulations of the terrestrial C cycle—that is, NPP-driven accumulations of C in soil pools that turn over based on

**Table 1 | Biome data**

Biome	<i>n</i>	AM	EEM	C stock (kg C m <sup>-2</sup> )	N stock (kg N m <sup>-2</sup> )	MAT (°C)	MAP (mm)	NPP (kg C m <sup>-2</sup> yr <sup>-1</sup> )	Clay (%)
Boreal forest	12	0	12	61.4 (11.7)	2.9 (0.5)	-2.5 (1.3)	497 (76)	319 (27)	5.8 (0.8)
Temperate forest	99	41	58	24.5 (2.6)	1.1 (0.1)	8.6 (0.4)	1,544 (133)	633 (28)	13.8 (1.1)
Tropical forest	104	83	21	11.7 (0.6)	1.1 (0.1)	24.7 (0.3)	2,697 (58)	956 (23)	47.9 (1.9)
Grassland	12	12	0	14.5 (2.2)	1.4 (0.3)	10.8 (1.7)	857 (145)	576 (99)	20.8 (3.1)

*n*, Number of observations; AM, number of *n* that are AM; EEM, number of *n* that are EEM; C stock, mean soil C content; N stock, mean soil N content; MAP (mm), MAT (°C), Clay, soil clay content. All values are means within a biome type, with the associated standard error given in parentheses.

<sup>1</sup>Department of Integrative Biology, Graduate Program in Ecology, Evolution and Behavior, University of Texas at Austin, Austin, Texas 78712, USA. <sup>2</sup>Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancon, Republic of Panama. <sup>3</sup>Department of Biology, Boston University, Boston, Massachusetts 02215, USA.

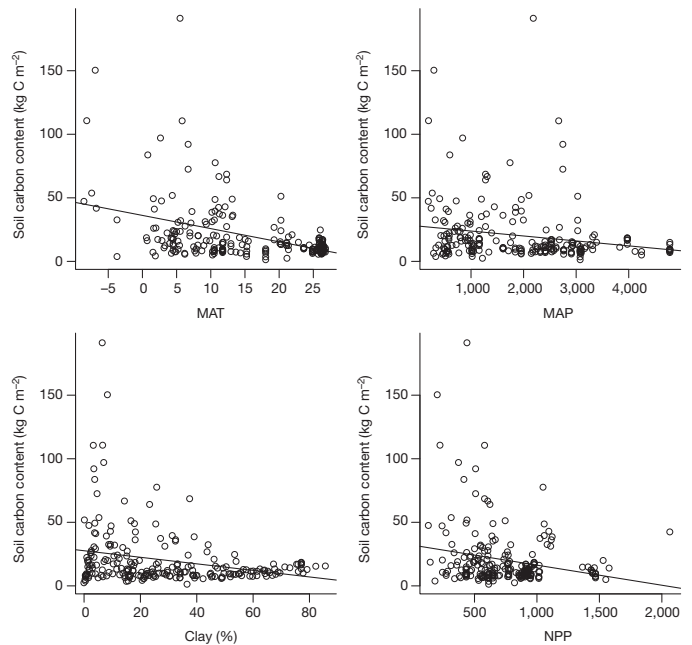


**Figure 1 | The relationship between soil carbon and nitrogen content to a depth of one metre in AM and EEM ecosystems.** The difference between the slopes is significant at the  $P < 0.0001$  level, based on the best AICc-selected full model output, after starting with all predictors ( $n = 227$ ). EEM systems store 1.7 times more C per unit N than do AM systems. Symbol shape reflects biome categorization. Symbol colour reflects mycorrhizal type, with purple symbols for EEM observations and black symbols for AM observations. Plotted lines represent univariate regression lines of the respective subsets of the data and are included for visualization purposes only.

Arrhenius temperature kinetics and a soil moisture multiplier<sup>20</sup>—lack a major driver of the decomposition process, namely, mycorrhizal type.

The results reported here support the recent theoretical contention that competition for organic N between EEM fungi and free-living microbes increases soil C storage<sup>11</sup>, and we show that this effect holds from tropical to high-latitude ecosystems (Fig. 1). Competition-induced declines in decomposition rate in EEM ecosystems are further supported by natural abundance and <sup>15</sup>N-labelling studies that show that EEM plants acquire more organic N from the soil than do AM plants<sup>13,21</sup>, and that experimental exclusion of EEM fungi increases the rate of organic matter decomposition<sup>7</sup> by increasing the biomass of free-living microbes and the activity of their C-degrading enzymes<sup>6</sup>. In contrast, the exclusion of AM fungi from the soil reduces the rate of decomposition by reducing the substrate supply to free-living decomposers<sup>22</sup>. Greater soil C storage in EEM ecosystems than in AM ecosystems at the large spatial scale reported here demonstrates that fine-scale mechanistic studies on the functionality of mycorrhizal symbioses—including N uptake preferences<sup>13,21</sup>, partitioning of plant-C belowground<sup>23</sup>, productivity<sup>24</sup> and decomposition<sup>7,22</sup>—can be scaled up to predict the consequences of AM versus EEM symbioses at the ecosystem-to-global scale.

It is possible that mycorrhizal effects on soil C pools may be confounded by differences in litter chemistry between EEM and AM plants. Compared to AM plants, litter from EEM plants can contain wider C:N ratios and greater concentrations of lignin and polyphenolic compounds, all of which are negatively correlated with short-term rates of litter decay<sup>25</sup>. However, lignin and polyphenols represent only a fraction of the soil C pool and compound-specific <sup>13</sup>C labelling studies show that these compounds decompose as fast or faster than 'labile' soil C compounds such as proteins and polysaccharides<sup>26</sup>. Moreover, recent theoretical<sup>27</sup> and empirical<sup>28</sup> work suggests that more recalcitrant (that is, more difficult to decompose) or higher C:N plant inputs may lead to less, rather than more, soil C storage than labile inputs because of lower microbial C-use efficiency (that is, the fraction of C assimilated that is allocated to



**Figure 2 | The relationships between soil carbon content to a depth of one metre and MAT, MAP, clay content and NPP.** Univariate regressions show that MAT in degrees celcius (a), MAP in mm precipitation per year (b), depth weighted clay content (c) and NPP in  $\text{kg C m}^{-2}$  per year (d) are significantly correlated with soil C storage ( $R^2 = 0.18, 0.04, 0.06$  and  $0.04, P < 0.0001, 0.0022, < 0.0001$  and  $0.0009$ , respectively;  $n = 227$ ). Regression lines represent univariate relationships rather than the output of the full model and are for visualization purposes only. None of these predictors were significant in the full model and were removed from the model after AICc selection.

growth rather than respiration). Therefore, we discount the possibility of a direct effect of litter chemistry on the observed variation in soil C storage among mycorrhizal types, although we cannot discount a potential indirect effect of litter chemistry owing to variations in microbial C-use efficiency.

This analysis shows that mycorrhizal functional traits are as important a control over decomposition and soil C storage as are soil chemical properties and the physical protection of organic matter<sup>26</sup>. More broadly, we demonstrate that the identity and functional traits of soil microorganisms exert a fundamental control over the terrestrial C cycle. This implies that global changes (for example, atmospheric N deposition, climate warming) that alter competitive interactions for N between EEM fungi and free-living microbial decomposers will affect soil C storage at regional to global scales.

## METHODS SUMMARY

We collected data on soil C, N and percentage clay to a depth of one metre in soil profiles spanning tropical, temperate and boreal forests and grasslands. Data were obtained from a variety of sources, including direct observations and both published and unpublished data (see Acknowledgements and Methods). MAT, MAP and NPP were assigned on the basis of latitude and longitude using global data products. Data are summarized by biome in Table 1. Data were analysed in a mixed effects framework using the *lme* function implemented in the *nlme* package for R statistical software<sup>29</sup>. We tested for a main effect of mycorrhizal status on soil C as well as an interaction between mycorrhizal type and soil N with biome coded as a random effect and all other variables coded as fixed effects. Model selection was performed using AICc selection criteria to prevent over-fitting the model. Results discussed in the text are based on the full model output based on the best AICc-selected model starting with all predictors. Reported correlation factor  $R^2$  values are based on the log ratio  $R^2$ . The mycorrhizal effect size reported in the text is determined by comparing the parameter estimate of the interaction between mycorrhizal type and soil N to the parameter estimate of the main effect of soil N, based on model outputs from the best AICc-selected full model.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 17 December 2012; accepted 25 November 2013.**

**Published online 8 January 2014.**

- Tarnocai, C. *et al.* Soil organic carbon pools in the northern circumpolar permafrost region. *Glob. Biogeochem. Cycles* **23**, GB2023 (2009).
- Jenkinson, D. S., Adams, D. E. & Wild, A. Model estimates of CO<sub>2</sub> emissions from soil in response to global warming. *Nature* **351**, 304–306 (1991).
- Knorr, W., Prentice, I. C., House, J. I. & Holland, E. A. Long-term sensitivity of soil carbon turnover to warming. *Nature* **433**, 298–301 (2005).
- Schimel, J. P. & Weintraub, M. N. The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: a theoretical model. *Soil Biol. Biochem.* **35**, 549–563 (2003).
- Allison, S. D., Gartner, T. B., Mack, M. C., McGuire, K. & Treseder, K. Nitrogen alters carbon dynamics during early succession in boreal forest. *Soil Biol. Biochem.* **42**, 1157–1164 (2010).
- Lindahl, B. D., de Boer, W. & Finlay, R. D. Disruption of root carbon transport into forest humus stimulates fungal opportunists at the expense of mycorrhizal fungi. *ISME J.* **4**, 872–881 (2010).
- Gadgil, R. & Gadgil, P. Mycorrhiza and litter decomposition. *Science* **233**, 133 (1971).
- Näsholm, T. *et al.* Boreal forest plants take up organic nitrogen. *Nature* **392**, 914–916 (1998).
- Hodge, A., Campbell, C. D. & Fitter, A. H. An arbuscular mycorrhizal fungus accelerates decomposition and acquires nitrogen directly from organic material. *Nature* **413**, 297–299 (2001).
- Read, D. J. & Perez-Moreno, J. Mycorrhizas and nutrient cycling in ecosystems—*a journey towards relevance?* *New Phytol.* **157**, 475–492 (2003).
- Orwin, K. H., Kirschbaum, M. U. F., St John, M. G. & Dickie, I. A. Organic nutrient uptake by mycorrhizal fungi enhances ecosystem carbon storage: a model-based assessment. *Ecol. Lett.* **14**, 493–502 (2011).
- LeBauer, D. S. & Treseder, K. K. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology* **89**, 371–379 (2008).
- Averill, C. & Finzi, A. Increasing plant use of organic nitrogen with elevation is reflected in nitrogen uptake rates and ecosystem  $\delta^{15}\text{N}$ . *Ecology* **92**, 883–891 (2011).
- Lindahl, B. D., Finlay, R. D. & Cairney, J. W. G. Enzymatic activities of mycelia in mycorrhizal fungal communities. *The Fungal Community: its Organization and Role in the Ecosystem* 3rd edn, 331–348 (CRC Press, 2005).
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. E. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
- Zhao, M. & Running, S. W. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* **329**, 940–943 (2010).
- Randerson, J. T. *et al.* Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. *Glob. Change Biol.* **15**, 2462–2484 (2009).
- Torn, M. S., Trumbore, S. E., Chadwick, O. A., Vitousek, P. M. & Hendricks, D. M. Mineral control of soil organic carbon storage and turnover. *Nature* **389**, 170–173 (1997).
- Cebrian, J. & Duarte, C. M. Plant growth-rate dependence of detrital carbon storage in ecosystems. *Science* **268**, 1606–1608 (1995).
- Davidson, E. A. & Janssens, I. A. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* **440**, 165–173 (2006).
- Gallet-Budynek, A. *et al.* Intact amino acid uptake by northern hardwood and conifer trees. *Oecologia* **160**, 129–138 (2009).
- Cheng, L. *et al.* Arbuscular mycorrhizal fungi increase organic carbon decomposition under elevated CO<sub>2</sub>. *Science* **337**, 1084–1087 (2012).
- Rygiel, P. T. & Andersen, C. P. Mycorrhizae alter quality and quantity of carbon allocated below ground. *Nature* **369**, 58–60 (1994).
- van der Heijden, M. G. A. *et al.* Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**, 69–72 (1998).
- Talbot, J. M. & Finzi, A. C. Differential effects of sugar maple, red oak, and hemlock tannins on carbon and nitrogen cycling in temperate forest soils. *Oecologia* **155**, 583–592 (2008).
- Schmidt, M. W. I. *et al.* Persistence of soil organic matter as an ecosystem property. *Nature* **478**, 49–56 (2011).
- Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Deneff, K. & Paul, E. The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Glob. Change Biol.* **19**, 988–995 (2013).
- Frey, S. D., Lee, J., Melillo, J. M. & Six, J. The temperature response of soil microbial efficiency and its feedback to climate. *Nature Clim. Change* **3**, 395–398 (2013).
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & the R Development Core Team *nlme: Linear and Nonlinear Mixed Effects Models* R package version 3. 1–109, <http://cran.r-project.org/web/packages/nlme/nlme.pdf> (2013).

**Acknowledgements** We thank L. Nave and the International Soil Carbon Network for access to their database. C. Hawkes provided feedback during data collection and initial analyses of C storage. C. Iversen, J. Powers and M. Vadeboncouer provided unpublished data that contributed to this analysis. D. Jacquier provided the Australian soil database and E. Carlston helped to extract data from the Australian soil database. C. Shaw provided the Siltanen soil carbon database and the Forest Ecosystem Carbon Database of Canadian soils. T. Baisden provided scans of the California Soil-Vegetation Survey. E. Brzostek, N. Fowler, P. Groffman, E. Hobbie, B. Schlesinger and B. Waring provided feedback on earlier versions of this manuscript. The Center for Tropical Forest Science (CTFS) and Smithsonian Institution Geo-observatories (SIGEO) provided funding for the collection and analysis of soil profile data at large forest dynamics plots, and we thank the many collaborators, field assistants and laboratory technicians who assisted in the collection and analysis of soil profile data. This work benefited from extensive data contributions to the International Soil Carbon Network from both the USDA Natural Resources Conservation Service, National Cooperative Soil Survey, and the US Geological Survey. C.A. was supported by a fellowship from the University of Texas at Austin and by the National Science Foundation Graduate Research Fellowship Program (grant DGE-1110007). A.C.F. was supported by NSF grant number DEB 07-43564 and DOE grants 10-DOE-1053 and DE-SC0006916. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

**Author Contributions** C.A. and B.L.T. collected the data. C.A. performed all statistical analyses. C.A. and A.C.F. conceptualized the work and wrote the manuscript. All authors contributed to revisions.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.A. ([colin.averill@utexas.edu](mailto:colin.averill@utexas.edu)).

## METHODS

**Data collection.** We accessed the International Soil Carbon database in February 2012 (<http://www.fluxdata.org/NSCN/SitePages/ISCN.aspx>), which contained a substantial amount of data from the United States Department of Agriculture and the National Resources Conservation Service (<http://ssldata.nrcs.usda.gov>). Plant species descriptions were supplemented with the National Soil Information System (NASIS) database (available from the United States Department of Agriculture, National Resources Conservation Service on request). We supplemented this data with unpublished data on 103 soil profiles from temperate and tropical forests. We then further supplemented this database by accessing the Canadian Forest Ecosystem Carbon Database and the Siltanen Database provided by C. Shaw, the Australian National Soil Database provided by D. Jacquier and a subset of the California Soil-Vegetation Survey provided by T. Baisden. We further supplemented the tropical sites by performing a Google Scholar search for articles containing all of the words “bulk density” + “clay” + “nitrogen” + “meter”, with the exact phrase: “soil carbon storage”; and with at least one of the words “tropics” and “tropical”. We also put out calls for these data on the ESA Biogeosciences Listserv and the National Soil Carbon Network listserv. J. Powers and M. Vadeboncouer provided unpublished data for this analysis. C. Iversen provided unpublished data from sites described in ref. 30, with permission.

**Calculating soil C and N content and clay concentration.** Soil C and N content were estimated as the sum of bulk density weighted soil C or N values to a depth of one metre or bedrock. Soil C and N content to a depth of one metre included organic horizons. Percentage clay to a depth of one metre was calculated as the depth-weighted average percentage clay concentration across all depth increments. Organic horizons were not included in the clay calculation. For our unpublished data set, soil C and N concentrations were determined by combustion and gas chromatography with thermal conductivity detection on a Thermo Flash 1112 Analyzer (CE Elantech), bulk density was determined by the excavation method<sup>31</sup>, and particle size distribution (including clay content) was determined by the pipette method following pretreatment to remove soluble salts, organic matter and iron oxides<sup>32</sup>.

**Assigning NPP, MAT and MAP.** NPP, MAT and MAP were assigned using global data products and the latitude and longitude of each site. NPP was determined from a ten-year average Moderate Resolution Imaging Spectroradiometer (MODIS) NPP product, MOD17A3 (ref. 16). MAT and MAP were taken from the WorldClim global climate data product<sup>15</sup>.

**Mycorrhizal classification.** Mycorrhizal type was assigned based on a site-specific vegetation observation of dominant species. We excluded observations that described a mixed mycorrhizal composition when no relative abundance data was available, although for forest biomes we ignored understory plants. When relative abundance data were available we required at least 70% of basal area of trees exceeding 10 cm in diameter at breast height to be one mycorrhizal type or the other. Some vegetation descriptions merely said “grassland” or “plains”, which we classified as AM. Seven observations had a vegetation description of “Sierran mixed coniferous forest”, which we classified as EEM. We note that forest classifications did not always include information on understory species. Biomass of understory plants is quite small by contrast to canopy trees, so it is unlikely that the understory plants had an important effect on patterns of soil C storage.

**Biome classification.** Biomes were assigned using the Whittaker Biome Diagram<sup>33</sup> and the MAT and MAP observations generated for each site from the WorldClim data product<sup>15</sup>. There were 32 instances in which Whittaker biome classifications were reassigned. Each of these soil observations was from a data set which had a description of vegetation and its United States Environmental Protection Agency ecoregion. If the Whittaker biome classification was not consistent with the vegetation description and EPA ecoregion, the biome was assigned to best match the vegetation description. For example, if the Whittaker MAT and MAP classification placed an observation into the temperate forest biome, but the vegetation description listed a “grassland” and the EPA ecoregion was “Great Plains,” then we classified the observation as the grassland biome. Finally, 12 observations from temperate rain forest fell far outside the Whittaker biome diagram as they had exceptionally high MAP values for a temperate forest (> 3,000 mm); however, we included them within the temperate forest category.

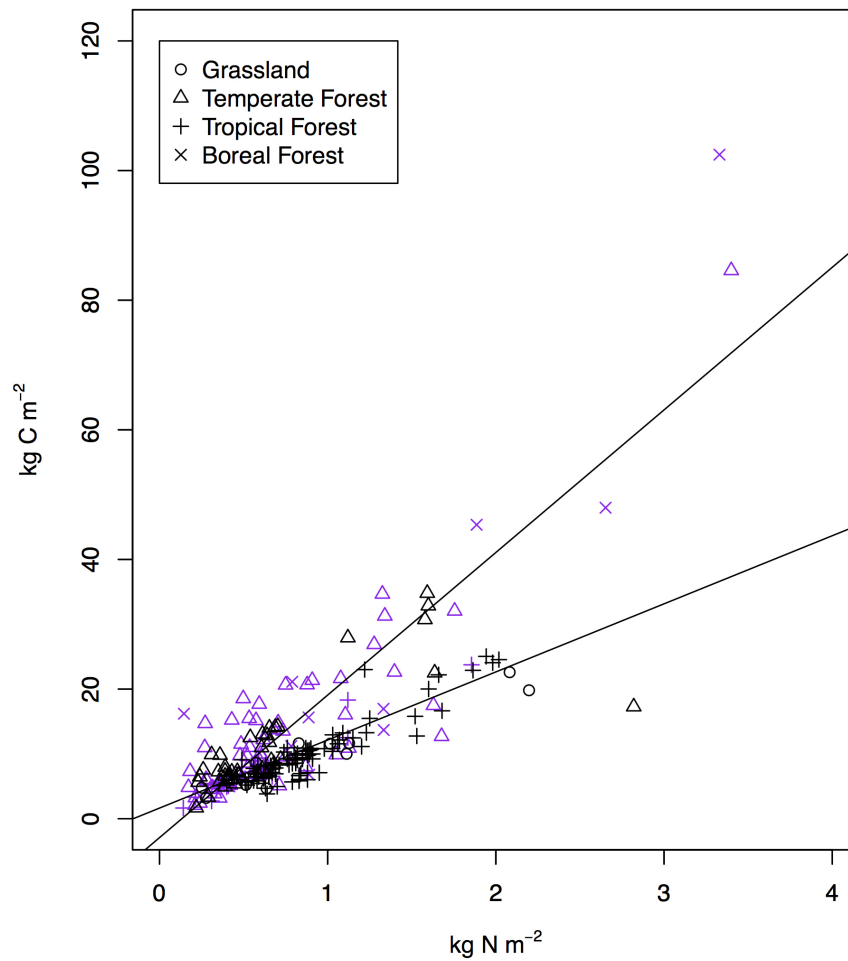
**Statistical approach.** We sought to model C storage as a function of mycorrhizal status and soil N while simultaneously accounting for variations in MAT, MAP, NPP and soil clay content. The data were analysed using a mixed effects framework, with MAT, MAP, NPP, clay, soil N, mycorrhizal status, and the interaction between N and mycorrhizal status as fixed effects and biome as a random effect

because the number of AM and EM observations was not evenly distributed among biomes. Model selection was performed using corrected AIC (AICc) criteria, using the *AICcmodavg* package in R<sup>34</sup>. We required a minimum of a one-point AICc improvement to justify removing a term from the model. Final linear models did not have normally distributed residuals and were strongly heteroscedastic. The heteroscedasticity in the model probably arises from a sampling error that is a constant percentage of total observed soil C and N, rather than a constant value (that is,  $\pm 10\%$  rather than  $\pm 10$  kg). We therefore fitted models by percentage least squares by weighting each observation by the inverse of the dependent variable (soil C stock) (as in ref. 35) and implemented using the weights function in *lme*<sup>29</sup>. Normality and homoscedasticity were inspected using plots of the normalized residuals. Because the  $R^2$  metric has different properties in linear mixed effects models and in standard linear models, we report an  $R^2$  statistic based on the likelihood ratio of the model ( $R^2_{LR}$ ) that has properties similar to those of the  $R^2$  implemented in linear models, as presented in ref. 36 and implemented in R using the *lmmfit* package<sup>37</sup>. The mycorrhizal effect size reported in the main text and Methods was determined by comparing the parameter estimate of the interaction between mycorrhizal type and soil N to the parameter estimate of the main effect of soil N, based on model outputs from the best AICc-selected full model.

Multicollinearity was assessed using variance inflation factors. These were calculated using the *vif* function in the *car* package in R<sup>38</sup>. The factors were calculated from the linear ordinary least-squares regression of all fixed effects without interactions. Collinearity was determined to be not a problem because all variance inflation factor values were less than ten for all independent variables<sup>39</sup>.

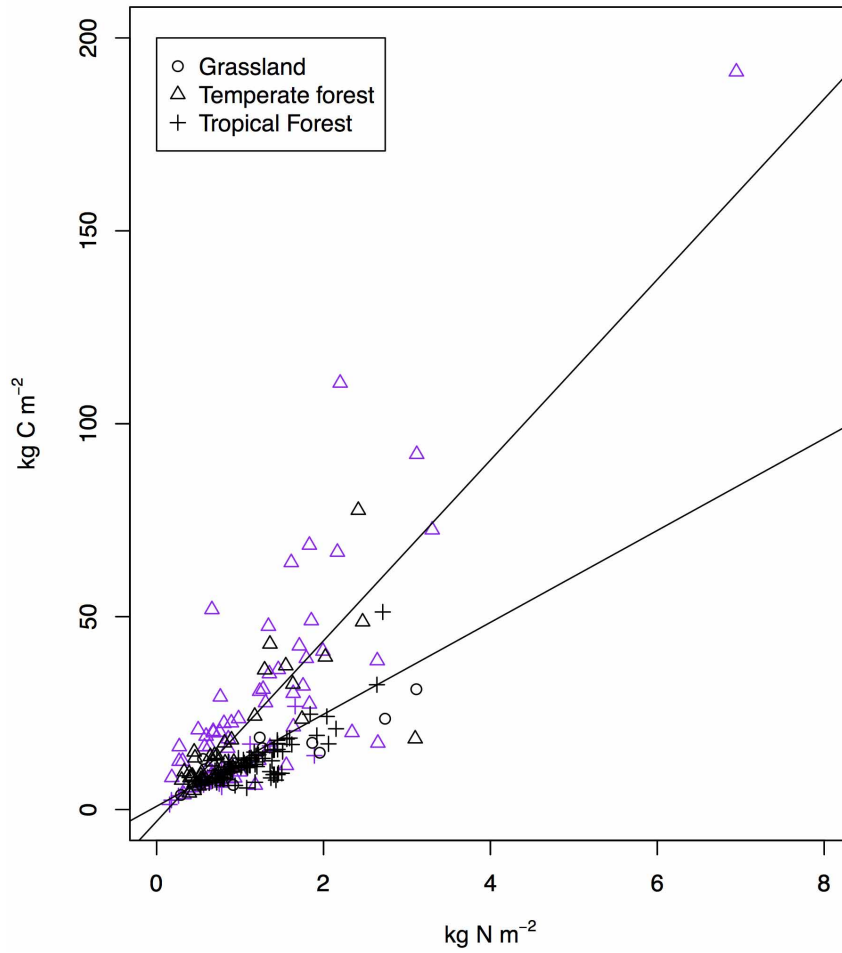
We further assessed the validity of the mixed effects framework by conducting a Monte Carlo simulation of the data set and analysis. Soil C distributions within biome and in most data sets are skewed distributions. To generate simulated data we determined the skew parameters of the soil C distribution (location, shape and scale), and then simulated a new biome-specific soil C distribution using the *rsnorm* function from The VGAM package in R<sup>40</sup>. The number of new data points drawn from each biome was set to the number of observations in the original data set. So, for example, a total of 99 observations were drawn for temperate forests, 41 of which were randomly assigned AM status and 58 EEM status. This data therefore assumes no difference in the biome-specific distributions of AM versus EEM observations, and our analysis should therefore detect no effect of mycorrhizal type on soil C storage. We then modelled soil C content as a function of mycorrhizal type with biome coded as a random effect and counted the number of times the mycorrhizal effect was significant at the 0.05 level (that is, a 1 in 20 random chance). We found that only 6.55% of the 10,000 simulations generated a statistically significant mycorrhizal effect, very slightly more frequently than expected by random chance. Furthermore, they were just as likely to be positive as negative, which means a positive effect of EEM status on soil C was detected less than 5% of the time. In contrast, our analysis finds a positive effect of EEM fungi on soil C storage at the level of  $P \leq 0.0001$  (that is, a <1 in 10,000 chance the effect is due to random chance). Therefore, our sampling and data analysis approach did not make us more likely to detect a positive effect of mycorrhizal type on soil C storage.

30. Mayes, M. A., Heal, K. R., Brandt, C. C., Phillips, J. R. & Hardine, P. M. Relation between soil order and sorption of dissolved organic carbon in temperate subsoils. *Soil Sci. Soc. Am. J.* **76**, 1027–1037 (2012).
31. Grossman, R. B. & Reinsch, T. G. 2002. in *Methods of Soil Analysis, Part 4: Physical Methods* (eds Dane, J. H. & Topp, C.) 201–203 (Soil Society of America, 2002).
32. Gee, G. W. & Or, D. 2002. in *Methods of Soil Analysis, Part 4: Physical Methods* (eds Dane, J. H. & Topp, C.) 255–293 (Soil Society of America, 2002).
33. Whittaker, R. H. *Communities and Ecosystems* 2nd edn, 111–191 (Macmillan, 1975).
34. Mazerolle, M. J. *AICcmodavg: Model selection and Multimodel Inference based on (Q)AIC(c)* R package version 1.31, <http://cran.r-project.org/web/packages/AICcmodavg/index.html> (2013).
35. Tofallis, C. Least squares percentage regression. *J. Mod. Appl. Stat. Methods* **7**, 526–534 (2008).
36. Magee, L.  $R^2$  measures based on wald and likelihood ratio joint significance tests. *Am. Stat.* **44**, 250–253 (1990).
37. Aleksandra, M. *lmmfit: Goodness-of-Fit-Measures for Linear Mixed Models with One-Level-Grouping* R package version 1.0, <http://cran.r-project.org/web/packages/lmmfit/index.html> (2011).
38. Fox, J. & Weisberg, H. S. *An {R} Companion to Applied Regression* 2nd edn (Sage Publications, 2011).
39. Zuur, A. F., Ieno, E. N. & Elphick, C. S. A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* **1**, 3–14 (2010).
40. Yee, T. W. The VGAM package for categorical data analysis. *J. Stat. Softw.* **32**, 1–34 (2010).



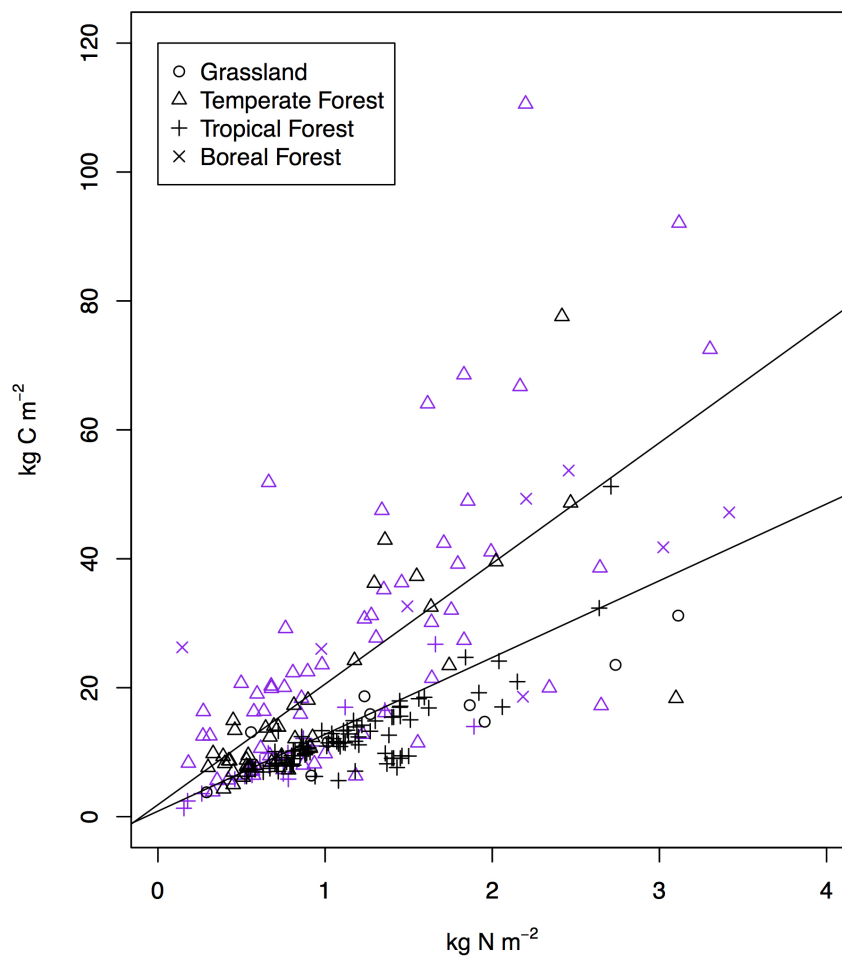
**Extended Data Figure 1 | Soil C versus N in the first 50 cm of mineral soil.** Purple symbols are EEM observations and black symbols are AM observations. Plotted lines represent univariate regression lines of the respective subsets of the data. We note that plotted lines are univariate regressions of data subsets and are included for visualization purposes only. Removal of the surface

organic horizon did not qualitatively change the interpretation of the data. Both the full model and the best AICc-selected model had a significant interactive effect between mycorrhizal type and soil N on soil C storage, with EEM systems storing 1.6 times more C per unit N than AM systems ( $P < 0.0001$ ).



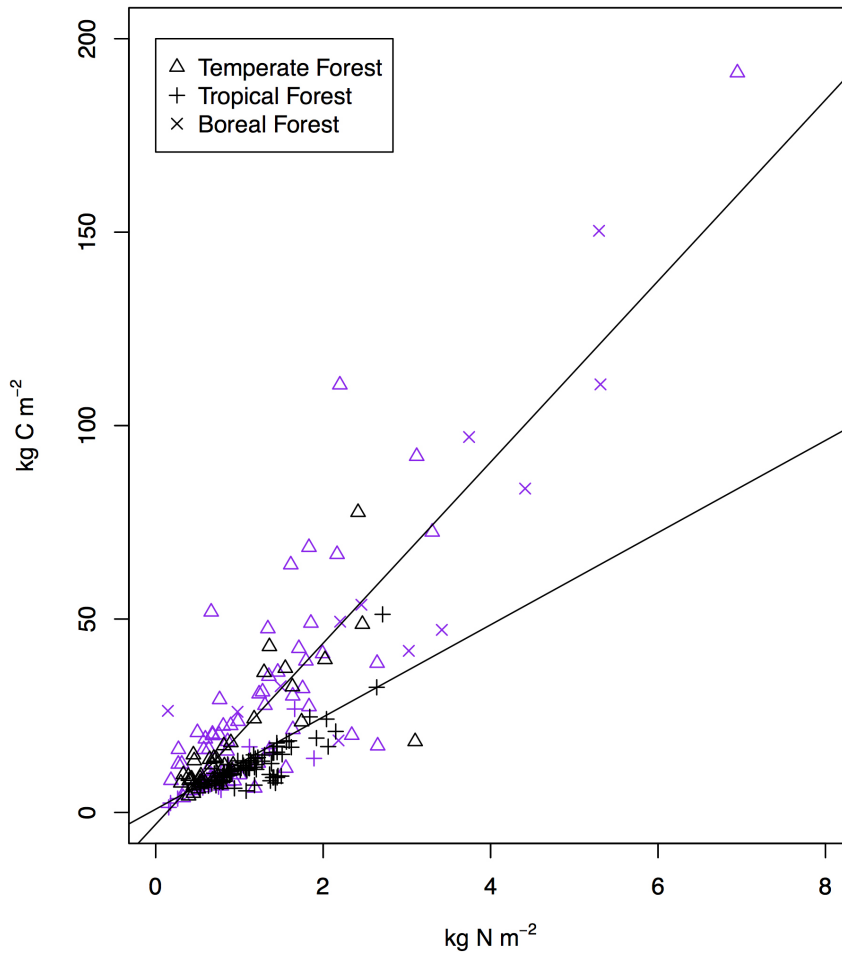
**Extended Data Figure 2 | Soil C versus N excluding boreal observations.** Purple symbols are EEM observations and black symbols are AM observations. Plotted lines represent univariate regression lines of the respective subsets of the data. We note that plotted lines are univariate regressions of data subsets

and are included for visualization purposes only. Both the full model and the best AICc-selected model showed a significant interactive effect of mycorrhizal type and soil N on soil C storage, with EEM systems storing 1.6 times more C per unit N than AM systems ( $P = 0.0014$ ).



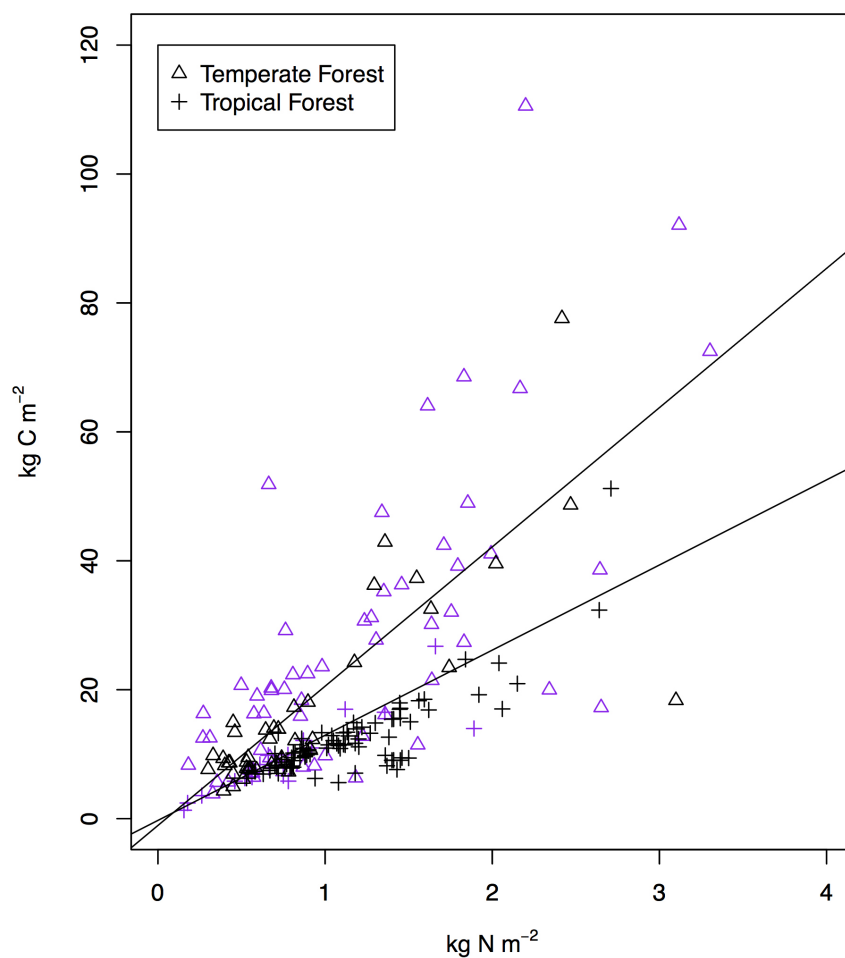
**Extended Data Figure 3 | Soil C versus N limiting data set to observations with less than  $3.5 \text{ kg N m}^{-2}$ .** Purple symbols are EEM observations and black symbols are AM observations. Plotted lines represent univariate regression lines of the respective subsets of the data. We note that plotted lines are univariate regressions of data subsets and are included for visualization

purposes only. Both the full model and the best AICc-selected model found a significant interactive effect of mycorrhizal type and soil N on soil C storage, with EEM systems storing 1.4 times more C per unit N than AM systems ( $P = 0.0304$ ).



**Extended Data Figure 4 | Soil C versus N excluding grassland observations.** Purple symbols are EEM observations and black symbols are AM observations. Plotted lines represent univariate regression lines of the respective subsets of the data. We note that plotted lines are univariate regressions of data subsets

and are included for visualization purposes only. Both the full model and the best AICc-selected model found a significant interactive effect of mycorrhizal type and soil N on soil C storage, with EEM systems storing 1.5 times more C per unit N than AM systems ( $P = 0.0023$ ).



**Extended Data Figure 5 | Soil C versus N restricting the analysis to temperate and tropical forest observations only.** Purple symbols are EEM and black symbols are AM observations. Plotted lines represent univariate regression lines of the respective subsets of the data. We note that plotted lines are univariate regressions of data subsets and are included for visualization purposes only. Both the full model and the best AICc-selected model

incorporated the interactive effect of mycorrhizal type and soil N on soil C storage, with EEM systems storing 1.3 times more C per unit N than AM systems, although the effect was marginally not significant ( $P = 0.0690$ ). We re-emphasize that the full model incorporates biome type, and weights observations by the inverse of their C values, to prevent undue influence of large observations on the estimated effect size.

Extended Data Table 1 | Mineral soil (0–50 cm) analysis regression output from the best AICc-selected model

Parameter	Estimate	Standard error	d.o.f.	<i>t</i> -value	<i>P</i> -value
(Intercept)	1.239237	0.8238489	218	1.504204	0.1340
EEM	-2.019305	0.8460802	218	-2.386659	0.0179
N	10.362798	0.8404036	218	12.330739	0.0000
EEM:N	5.749893	1.2659686	218	4.541892	0.0000

C ≈ mycorrhizal status × N, random = biome,  $R^2_{LR} = 0.091$ . C, soil carbon in  $\text{kg m}^{-2}$ ; EEM, the effect of EEM fungi on soil carbon. The *t*-value is from Student's test. d.o.f., degrees of freedom.

Extended Data Table 2 | Removing boreal forests analysis from the best AICc-selected model

Parameter	Estimate	Standard error	d.o.f.	<i>t</i> -value	<i>P</i> -value
(Intercept)	2.027935	1.71859	209	1.179999	0.2393
EEM	-2.3091	1.594658	209	-1.448022	0.1491
N	9.73505	1.134723	209	8.57923	0.0000
EEM:N	5.580011	1.725335	209	3.23416	0.0014

C ≈ mycorrhizal status × N, random = biome,  $R^2_{LR} = 0.89$ . C, soil carbon in kg m<sup>-2</sup>; EEM, the effect of EEM fungi on soil carbon. The *t*-value is from Student's test. d.o.f., degrees of freedom.

Extended Data Table 3 | Restricting range of N content analysis from the best AICc-selected model

Parameter	Estimate	Standard error	d.o.f.	<i>t</i> -value	<i>P</i> -value
(Intercept)	2.797858	1.915448	215	1.460681	0.1456
EEM	-0.8853	1.560126	215	-0.567454	0.5710
N	9.85441	1.120904	215	8.791487	0.0000
EEM:N	3.615833	1.658742	215	2.179866	0.0304

C ≈ mycorrhizal status × N, random = biome,  $R^2_{LR} = 0.83$ . C, soil carbon in kg m<sup>-2</sup>; EEM, the effect of EEM fungi on soil carbon. The *t*-value is from Student's test. d.o.f., degrees of freedom.

Extended Data Table 4 | Removing grasslands analysis from the best AICc-selected model

Parameter	Estimate	Standard error	d.o.f.	<i>t</i> -value	<i>P</i> -value
(Intercept)	3.745027	2.927873	209	1.279095	0.2023
EEM	-2.215022	1.735368	209	-1.276399	0.2032
N	10.265126	1.371112	209	7.486716	0.0000
EEM:N	5.63454	1.82723	209	3.08365	0.0023

C ≈ mycorrhizal status × N, random = biome,  $R^2_{LR} = 0.91$ . C, soil carbon in kg m<sup>-2</sup>; EEM, the effect of EEM fungi on soil carbon. The *t*-value is from Student's test. d.o.f., degrees of freedom.

Extended Data Table 5 | Temperate and tropical biomes only, from the best AICc-selected model

Parameter	Estimate	Standard error	d.o.f.	<i>t</i> -value	<i>P</i> -value
(Intercept)	1.887213	2.443358	197	0.772385	0.4408
EEM	-0.856856	1.702001	197	-0.50344	0.6152
N	10.357411	1.283155	197	8.071831	0.0000
EEM:N	3.438798	1.881089	197	1.828089	0.0690

C ≈ mycorrhizal status × N, random = biome,  $R^2_{LR} = 0.83$ . C, soil carbon in  $\text{kg m}^{-2}$ ; EEM, the effect of EEM fungi on soil carbon. The *t*-value is from Student's test. d.o.f., degrees of freedom.