A Journal of Macroecology

Global Ecology and Biogeography

Global drivers and patterns of microbial abundance in soil

Hector M. Serna-Chavez[1,2]†, Noah Fierer[3] and Peter M. van Bodegom[1]*

[1]*Department of Ecological Science, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands,* [2]*Department of Life Sciences, University of Coimbra, P3004-517 Coimbra, Portugal,* [3]*Department of Ecology and Evolutionary Biology and the Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA*

*Correspondence: Peter M. van Bodegom, Department of Ecological Science, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.
E-mail: p.m.van.bodegom@vu.nl
†Current Address: Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, PO Box 94248, 1090 GE Amsterdam, The Netherlands.

ABSTRACT

Aim While soil microorganisms play key roles in Earth's biogeochemical cycles, methodological constraints and sparse data have hampered our ability to describe and understand the global distribution of soil microbial biomass. Here, we present a comprehensive quantification of the environmental drivers of soil microbial biomass.

Location Global.

Methods We used a comprehensive global dataset of georeferenced soil microbial biomass estimates and high-resolution climatic and soil data.

Results We show that microbial biomass carbon (C_{Mic}) is primarily driven by moisture availability, with this single variable accounting for 34% of the global variance. For the microbial carbon-to-soil organic carbon ratio (C_{Mic}/C_{Org}), soil nitrogen content was an equally important driver as moisture. In contrast, temperature was not a significant predictor of microbial biomass patterns at a global scale, while temperature likely has an indirect effect on microbial biomass by influencing rates of evapotranspiration and decomposition. As our models explain an unprecedented 50% of the global variance of C_{Mic} and C_{Mic}/C_{Org}, we were able to leverage gridded environmental information to build the first spatially explicit global estimates of microbial biomass and quantified the global soil microbial carbon pool to equal 14.6 Pg C.

Main Conclusions Our unbiased models allowed us to build the first global spatially explicit predictions of microbial biomass. These patterns show that soil microbial biomass is not primarily driven by temperature, but instead, biomass is more heterogeneous through the effects of moisture availability and soil nutrients. Our global estimates provide important data for integration into large-scale carbon and nutrient models that may imply a major step forward in our ability to predict the global carbon balance, now and in a future climate.

Keywords
Biogeochemical cycles, global carbon cycling, moisture limitation, nitrogen limitation, soil microbial abundance, soil microbial biomass, soil microbial carbon.

INTRODUCTION

Ecologists have long been intrigued by the distribution of life on Earth. Information on the distribution and abundance of aboveground plant biomass has been available for decades which improved our understanding of the processes structuring plant communities, driving species abundances and influencing ecosystem processes. In contrast, efforts to estimate global patterns in abundance and the size of the soil microbial biomass pool have lagged considerably. Given the key role soil microorganisms play in ecosystem processes, elucidating such patterns will represent a critical step as we seek to improve our ability to predict and understand microbial controls on biogeochemical cycles.

Local- and regional-scale studies have identified environmental drivers of microbial biomass carbon (C_{Mic}) and of the biologically active fraction of the soil organic carbon pool (C_{Mic}/C_{Org}) (Insam, 1990; Wardle, 1992; Franzluebbers *et al.*, 2001; Bachar *et al.*, 2010; Drenovsky *et al.*, 2010; Dequiedt *et al.*, 2011). However, the effect and direction of these drivers are rarely consistent across studies. Further identification and quantification of drivers of soil microbial abundance and activity have been hampered by lack of representative data, collected using uniform (standard) methodologies. Global patterns may be presumed, however, as evidenced by stoichiometric and metabolic constraints determining patterns of microbial abundance and activity across all soils (Cleveland & Liptzin, 2007; Manzoni *et al.*, 2008; Sinsabaugh *et al.*, 2008). Fierer *et al.* (2009) presented results suggesting that the variability in C_{Mic} across major biomes is strongly correlated to soil organic carbon (SOC) content and plant productivity. However, by using biome level means they omitted intrabiome variation, which may hinder the identification of those factors influencing microbial biomass levels across and within biomes. To our knowledge, no other attempts have been made to quantify, let alone map, global patterns of soil microbial biomass. Such quantifications are critical given that current global estimates of the soil microbial biomass pool vary widely from 13.9 and 23.2 to 26 Pg C (Wardle, 1992; Whitman *et al.*, 1998; Xu *et al.*, 2013, respectively; $1 \, Pg = 1 \times 10^{15} \, g$). Given that the data used to generate these estimates were largely derived from extrapolations based on temperate forest and cropland soils, a more accurate determination of the size of the microbial biomass pool is clearly needed.

The uncertainties in previous estimates leave an important gap in our understanding of the soil microbial world and soil carbon dynamics. Elucidating the environmental drivers that determine the size of the soil microbial pool is essential to better evaluate global fluxes and cycling of soil carbon and soil nutrients alike, as all are strongly influenced by soil microbial communities. Also, the interactions between above- and below-ground communities are increasingly understood; however, it is still unknown if the spatial patterns of soil microbial abundance mirror those of plant biomass. Therefore, we focus here on quantifying and explaining global patterns of soil microbial biomass and its relationship to climate and soil drivers. By subsequently applying our multivariate models to gridded climate and soil information, we derive and present the first spatially explicit global estimates of C_{Mic} in the soil profile and the topsoil C_{Mic}/C_{Org} ratio. Our spatially explicit estimates of the soil microbial pool can substantially improve predictions of global biogeochemical and vegetation models and, thus, allowing better estimations of the carbon balance.

MATERIALS AND METHODS

Soil C_{Mic} estimates

We based our study on a dataset of soil microbial biomass estimates across all major biomes (Cleveland & Liptzin, 2007, as extended and modified by Fierer *et al.*, 2009). To improve latitudinal representation, we complemented the dataset with estimates from arctic and tropical regions. The final dataset comprised 414 georeferenced estimates (in mg $C_{Mic} \, kg^{-1}$ soil), most of which had been obtained from mineral soils and only a few from organic soils. All estimates were from the A horizons obtained under control (non-manipulated) conditions. No estimates from litter layers were included, as drivers of microbial biomass in litter layers are presumably different from those in soil. Also, because seasonal variation in microbial biomass was seldomly reported, we only present annual mean microbial biomass. The majority of estimates (92%) had been obtained using the chloroform fumigation-extraction (CFE) method (Vance *et al.*, 1987) and a minor proportion (8%) from chloroform fumigation-incubation (CFI) (Jenkinson & Powlson, 1976). For both, microbial biomass in the depth-interval sampled was calculated as:

$$C_{Mic} = E_C / k_{EC} \tag{1}$$

where E_C represents the difference between fumigated and unfumigated soil extractable carbon or CO_2 produced; and k_{EC} the extraction efficiency accounting for incomplete mineralization of biomass (CFI) or incomplete extraction of the killed biomass (CFE). While CFE and CFI are subject to biases in soils with high organic matter content and high acidity, respectively (Martens, 1995), no other method to quantify microbial biomass is bias-free or as widely used (Martens, 1995; Joergensen *et al.*, 2011). Moreover, we found no systematic deviations in microbial biomass estimates under potentially biasing conditions for CFE or CFI (see Appendix S1).

To facilitate comparisons across studies, we calculated E_C using the reported C_{Mic}, the applied extraction coefficient (k_{EC}) and equation 1. Afterwards, a k_{EC} of 0.4 was applied to standardize all estimates. In addition, considering soil bulk density and the depth interval sampled, we extrapolated microbial biomass content to the first metre of soil profile (C_{Mic}, in g C m^{-2}). This extrapolation was made assuming that soil conditions change continuously and similarly across all soils down to the first metre (Fierer *et al.*, 2009) (Appendix S1). Because bulk density was usually not given for the individual studies, we applied the bulk density reported in the global database for the specific soil type (Batjes, 1995) if the estimate was from a mineral soil. Given that the bulk density of organic soils varies much more strongly, we used the value reported in the original publication (or from other publications describing the same site). In addition, because all depth intervals sampled for microbial biomass represented soil conditions in the upper 30 cm, we calculated the topsoil C_{Mic}/C_{Org}, considered an index of microbial activity. First, we normalized all estimates in the dataset to the amount of microbial biomass (g $C_{Mic} \, m^{-2}$) in the topsoil (Appendix S1). Then for mineral soils, we extracted topsoil SOC from the soil profile database (Batjes, 1995), whereas for organic soils, SOC was obtained from the original publication.

Soil data

We used the ISRIC-WISE 0.5°-lat × 0.5°-long global database of soil profiles (Batjes, 1995) to extract soil variables, including bulk density (g cm^{-3}), SOC (% mass), cation exchange capacity (CEC, meq 100 g^{-1}), C-N ratio (g C g^{-1} N), pH (in H$_2$O solution), total nitrogen (% mass) and texture variables (% of sand and clay). These parameters reflect the physical-chemical characteristics of mineral soils (0–30 cm). Parameters for organic soils (0–30 cm) were extracted from the original publications or from related publications. Most commonly, studies reported SOC, total nitrogen, bulk density and pH. The missing parameters for organic soils were excluded from the analysis. These actions were a pragmatic solution to minimize biases because of missing organic soil data. Not considering microbial biomass estimates from organic layers may have led to underrepresentation or misrepresentation of particular global regions.

We used topsoil parameters to assess environmental drivers of soil profile C$_{Mic}$ because they had higher correlation coefficients and accounted for more of the global variation in microbial biomass. We also ran analyses calculating soil parameters as (1) arithmetic mean of topsoil (0–30 cm) and subsoil (30–100 cm) data as reported in the ISRIC-WISE database, and (2) as weighted-means according to their relative importance to microbial biomass (87% topsoil, 13% subsoil). In both cases, neither correlation coefficients nor the explanatory power of multivariate models was higher than when using topsoil parameters alone.

Climate data

We used the climate research unit 0.5°-lat × 0.5°-long surface mean monthly climatology database of global land areas excluding Antarctica (New *et al.*, 1999). Using the georeferences of the biomass sampling locations, we calculated mean annual precipitation (MAP, mm H$_2$O yr^{-1}); maximum monthly frequencies of wet and frost days (MaxWD and MaxFD, respectively), lowest annual minimum monthly temperature (MinAT, in °C), highest annual maximum monthly temperature (MaxAT, °C), mean annual temperature (MAT, °C), and annual temperature range (T$_{range}$ = MaxAT – MinAT, in °C). MaxWD is defined as the highest monthly number of days where precipitation exceeded 1.0 or 0.1 mm. Similarly, MaxFD represents the highest monthly number of days with gross minimum temperature below 0°C. MinAT and MaxAT are defined as the lowest minimum and highest maximum monthly temperatures in a year, respectively. Additionally, we calculated mean monthly and annual evaporative demand (ETo, mm H$_2$O day^{-1} and mm H$_2$O yr^{-1}), mean annual soil moisture deficit (MASMD, mm H$_2$O) and the annual ratio of MAP and ETo (MAP/ETo) for each biomass estimate. MASMD is the sum of differences between mean monthly ETo and mean monthly precipitation (Fierer & Jackson, 2006). ETo was calculated according to the Food and Agriculture Organization Penman–Monteith equation using climate data and assuming constant vegetation and roughness attributes, which makes the estimate independent of the standing vegetation (Allen *et al.*, 1998).

Statistical analysis

Multivariate regression analyses were performed to quantify global environmental drivers and to model soil microbial biomass as a linear function of multiple explanatory variables. Estimates of soil profile C$_{Mic}$ (g C$_{Mic}$ m^{-2}) and topsoil C$_{Mic}$/C$_{Org}$ ratio were used as response variables and analysed individually. Log$_{10}$ transformations were applied to both response variables to approach normal data distribution. Subsequently, all soil and climate variables were plotted against the transformed response variables to determine which to include in the regression analysis. In each case, Pearson's correlation coefficients (r, $\alpha = 0.05$) were calculated. We log$_{10}$-transformed some soil and climate variables to approach normal data distribution, enhance r, and improve linear relations between the explanatory and response variables. These transformations were performed separately for each analysis, as both response variables showed different patterns with explanatory variables. Upon the log$_{10}$ transformations, essentially linear relations were obtained.

To identify and select the most parsimonious set of environmental variables that explain C$_{Mic}$ and C$_{Mic}$/C$_{Org}$ variance, we performed one-by-one backward stepwise regressions. To manage multicollinearity between environmental variables, we only included in our models explanatory variables of which the r among them varied between 0.7 and −0.7. Additionally, variance inflation factors (VIFs) were calculated for all regressors in multivariate models, and a limit of VIF ≤ 4 was set. Criteria to drop variables were: collinearity with other explanatory variables, significance of regression coefficients, single-term deletion *F-test* and the overall improvement of model's fit after removal. The best multivariate models were selected according to the adjusted coefficient of determination (R_{adj}^2), Akaike's information criteria (AIC) and regression diagnostic plots (studentized residuals versus fitted values and distribution of studentized residuals). For the developed models, we assessed the relative variance accounted for by each explanatory variable using a hierarchical partitioning metric that accounts for the direct effect of each regressor and the adjusted effect for all other regressors (lmg metric, Groemping, 2006).

As a control, we developed a null model to explain soil profile C$_{Mic}$ and topsoil C$_{Mic}$/C$_{Org}$ using 'biome' as sole explanatory variable. All estimates in our dataset were classified as boreal forest, desert, temperate coniferous forest, temperate deciduous forest, temperate grassland, tropical forest or tundra. This simple biome model served to analyse interbiome differences and to compare further multivariate models with, that is, to assess whether or not we improved the proportion of explained variance by using continuous relationships with, environmental variables. The proportion of explained variance by interbiome differences alone was calculated as:

$$\eta^2 = \left(\frac{SS_{biome}}{SS_{biome} + SS_{residuals}} \right) \times 100 \qquad (2)$$

Furthermore, to cross-validate the developed multivariate models we compared our estimates to a dataset of substrate-induced respiration (SIR) estimates ($\mu g\ C\ g^{-1}\ h^{-1}$) from 77 soil samples. This dataset comprised soil samples from an array of temperate soils in North America and a few tropical soils in Puerto Rico, as described in Fierer & Jackson (2006). SIR estimates were obtained for the top 30 cm of mineral soil using the methodology described in Fierer *et al.* (2003). For each location where SIR had been measured, we predicted soil profile C_{Mic} based on its georeferences to derive soil and climate variables using the $0.5°$-lat $\times 0.5°$-long global soil profile (Batjes, 1995) and climate (New *et al.*, 1999) databases, and our multivariate models. We then assessed the significance of correlations (Pearson's r, $\alpha = 0.05$) between our C_{Mic} estimates and the SIR-responsive biomass data. Finally, we mapped the global soil C_{Mic} and C_{Mic}/C_{Org} ratio applying our multivariate models to the gridded climate and the soil profile information for each $0.5° \times 0.5°$ grid, following appropriate back-transformations. All statistical analyses were conducted in R v.2.13.0 (R Development Core Team, 2011), while the application of the developed regression models to derive the world maps was conducted in Fortran using Force 2.0 v.2.0.9p.

RESULTS

Quantification of global environmental drivers

Overall, moisture variables were the main individual drivers of soil profile C_{Mic}. Correlation with the MAP-ETo ratio (MAP/ETo) was the strongest ($r = 0.46$, $P < 0.0001$) followed by MAP ($r = 0.45$, $P < 0.0001$), both common measures of moisture availability and supply. Also, the interaction between moisture availability and MAT (MAP/ETo $\times$ MAT) exhibited a strong linear relationship with C_{Mic} ($r = 0.57$, $P < 0.0001$). This interaction indicated that with increasing MAT, the linear relationship between moisture availability (MAP/ETo) and microbial biomass becomes stronger. From our temperature-related variables only annual T_{range} had a strong correlation with C_{Mic} ($r = -0.33$, $P < 0.0001$). All other variables, including MAT, exhibited low correlation coefficients. Soil characteristics also showed an important influence; while the relation between SOC and C_{Mic} was not as strong as expected ($r = 0.14$, $P = 0.004$), total nitrogen and pH were quite strongly correlated to C_{Mic} ($r = 0.25$ and -0.26, respectively, $P < 0.0001$).

The analysis of topsoil C_{Mic}/C_{Org} data suggests that soil and temperature variables exert a stronger influence on this ratio than on C_{Mic}. Still, variables associated with soil moisture availability had the strongest correlation coefficients. However, it was MAP instead of MAP/ETo $\times$ MAT that exhibited the strongest correlation ($r = 0.43$ and 0.32, respectively, $P < 0.0001$). The influence of nutrients was exemplified by the strong correlation with total nitrogen content ($r = -0.37$, $P < 0.0001$) and, to a lesser extent, with the carbon-nitrogen (C-N) ratio ($r = -0.25$, $P < 0.0001$). MAT and the MinAT exhibited strong correlation with C_{Mic}/C_{Org} ($r = 0.33$, $P < 0.0001$, in both cases). Also, the MaxFD strongly constrained

C_{Mic}/C_{Org} ($r = -0.36$, $P < 0.0001$). For both microbial biomass variables, all other environmental correlates are presented in Appendix S2.

Multivariate models to predict global patterns of microbial biomass

We selected the best multivariate linear models that explained soil profile C_{Mic} and topsoil C_{Mic}/C_{Org} (Table 1). Selected models had a distribution of studentized residuals close to normal, with no distinct patterns between residuals and the explanatory variables in the models or those variables not included in the model. Also, no true outliers were detected (Cook's $D < 0.08$; Hat-values < 0.09; Bonferroni's adjusted $P > 0.05$).

Soil profile C_{Mic} was best explained by a combination of MAP/ETo $\times$ MAT, the MaxWD, soil pH and total nitrogen, accounting for 39.0% of the total variance (Model 1, $F_{4,386} = 63.2$, $P < 0.0001$, $AIC_{M1} = -875.6$; Fig. 1a). A relative importance analysis of regressors supported the finding that moisture availability variables are the main drivers of C_{Mic} alone accounting for more than 80% of the explained variance (Fig. 2a). For topsoil C_{Mic}/C_{Org}, the best model included MAP/ETo $\times$ MAT, MaxWD, soil total nitrogen, pH, C-N ratio and CEC. It explained 50.7% of the total variance in C_{Mic}/C_{Org} (Model 4, $F_{6,360} = 63.7$, $P < 0.0001$; $AIC_{M4} = -802.9$; Fig. 1b). Here, the relative importance analysis suggested a more balanced contribution between moisture availability and soil nutrient variables as main drivers, together accounting for more than half of the explained variance (Fig. 2b). Alternative multivariate models 2 and 3, representing the second best fits, are presented in Appendix S1.

For comparison, we also assessed interbiome differences for our dataset (see Methods). Biome level means of C_{Mic} were significantly different (ANOVA $F_{6,407} = 10.6$, $P < 0.0001$) particularly because of tropical forests exhibiting higher microbial biomass than several other biomes (Fig. 3a). A similar result was obtained for the biome level means of C_{Mic}/C_{Org} (ANOVA $F_{6,390} = 11.2$, $P < 0.0001$; Fig. 3b). These interbiome differences reflect inherent differences in environmental conditions among biomes and account for 13.5% (7–19%) and 14.7% [8–20%; η^2 and 95% confidence interval (CI)] of the total C_{Mic} and C_{Mic}/C_{Org} variation, respectively. This level of variance is about three times less than the variance explained when directly coupling environmental drivers to biomass estimates. By directly estimating the impacts of environmental drivers on soil microbial biomass, we were able to account for the interbiome and intrabiome differences in microbial biomass.

We cross-validated our model estimates using microbial biomass estimated using a different methodology: SIR ($\mu g\ C\ g^{-1}\ h^{-1}$; see Methods). Predicted C_{Mic} and SIR were better related to estimates from our best model ($r = 0.39$; $P < 0.001$) than to our alternative second-best model ($r = 0.34$; $P < 0.05$), although correlation coefficients were not statistically different from each other ($Z_{calculated} = 0.35$; $P > 0.05$). Given that this cross-validation compares local SIR estimates with grid-average microbial biomass, these correlation coefficients are reasonably

Table 1 Environmental drivers of soil profile C_{Mic} (Models 1 and 2) and the topsoil C_{Mic}/C_{Org} ratio (Models 3 and 4)

|	Drivers	*F*-value	β (± 1SEM)	*b*	*P*	*VIF*
Model 1	MAP/ETo† × MAT	230.0	0.042 (± 0.004)	0.61	***	2.6
	MaxWD	4.8	0.013 (± 0.005)	0.15	**	2.1
	pH	14.1	0.090 (± 0.022)	0.21	***	1.7
	Total N†	4.5	0.131 (± 0.062)	0.09	*	1.1
Model 2	MAP/ETo† × MAT	237.3	0.039 (± 0.004)	0.57	***	2.8
	MaxWD	7.4	0.020 (± 0.005)	0.23	***	2.3
	pH	11.4	0.071 (± 0.026)	0.17	**	2.3
	C-N ratio†	20.7	−1.353 (± 0.432)	−0.14	**	1.3
	CEC	5.5	0.008 (± 0.004)	0.10	*	1.1
Model 3	MAP†	131.4	0.652 (± 0.082)	0.46	***	2.4
	ETo†	24.9	−0.471 (± 0.129)	−0.18	***	1.7
	MaxWD	0.5	0.023 (± 0.006)	0.24	***	2.5
	Total N†	113.2	−0.817 (± 0.103)	−0.50	***	2.8
	C-N ratio†	61.4	−2.943 (± 0.448)	−0.28	***	1.3
	pH†	21.5	1.227 (± 0.343)	0.19	***	2.1
	CEC†	4.6	0.376 (± 0.176)	0.12	*	2.1
Model 4	MAP/ETo† × MAT	83.5	0.041 (± 0.004)	0.55	***	2.6
	MaxWD	1.7	0.022 (± 0.005)	0.23	***	2.4
	Total N†	225.0	−0.961 (± 0.095)	−0.59	***	2.5
	C-N ratio†	43.1	−2.067 (± 0.441)	−0.19	***	1.3
	pH†	24.1	1.268 (± 0.334)	0.20	***	2.0
	CEC†	5.2	0.396 (± 0.173)	0.12	*	2.1

*$P < 0.05$, **$P < 0.001$, ***$P < 0.001$.

†Variable is $\log_{10}$-transformed.

β, regression coefficients; *b*, standardized coefficients; CEC, soil cation exchange capacity; C_{Mic}, microbial biomass carbon; C_{Mic}/C_{Org}, microbial carbon-to-soil organic carbon ratio; C-N ratio, soil carbon-nitrogen ratio; ETo, annual evaporative demand; MAP, mean annual precipitation; MAT, mean annual temperature; MaxWD, maximum monthly frequency of wet days; pH, soil pH; SEM, standard error of the mean; Total N, soil total nitrogen content; VIF, variance inflation factors.
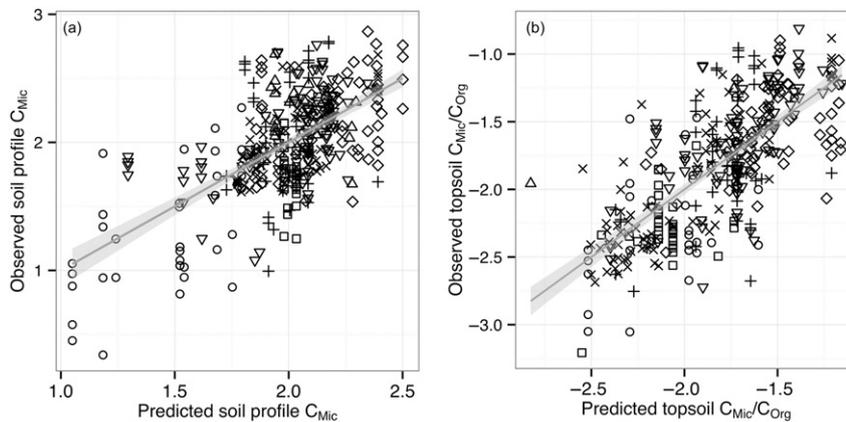
Figure 1 Observed versus predicted values. (a) Soil profile microbial biomass carbon (C_{Mic}) estimates from Model 1 (as provided in Table 1). Regression line: slope of 1.000 ± 0.062; β ± 1 standard error of the mean (SEM)]; R^2 of 39.8%, $P < 0.0001$ ($n = 391$). (b) Topsoil C_{Mic}/soil organic carbon (C_{Org}) ratio estimated using Model 4 (as provided in Table 1). Regression line: slope of 1.000 ± 0.051; β ± 1SEM); adjusted coefficient of determination (R_{adj}^2) of 50.7%, $P < 0.0001$ ($n = 367$). For both panels, shadowed area represents the 95% confidence interval (CI), and all axes are on a $\log_{10}$ scale. Legend and biome classification as follows: boreal forest (□, BF), desert (○, D), tundra (△, T), temperate coniferous forest (+, TCF), temperate deciduous forest (×, TDF), tropical forest (◊, TF), and temperate grassland (▽, TG).
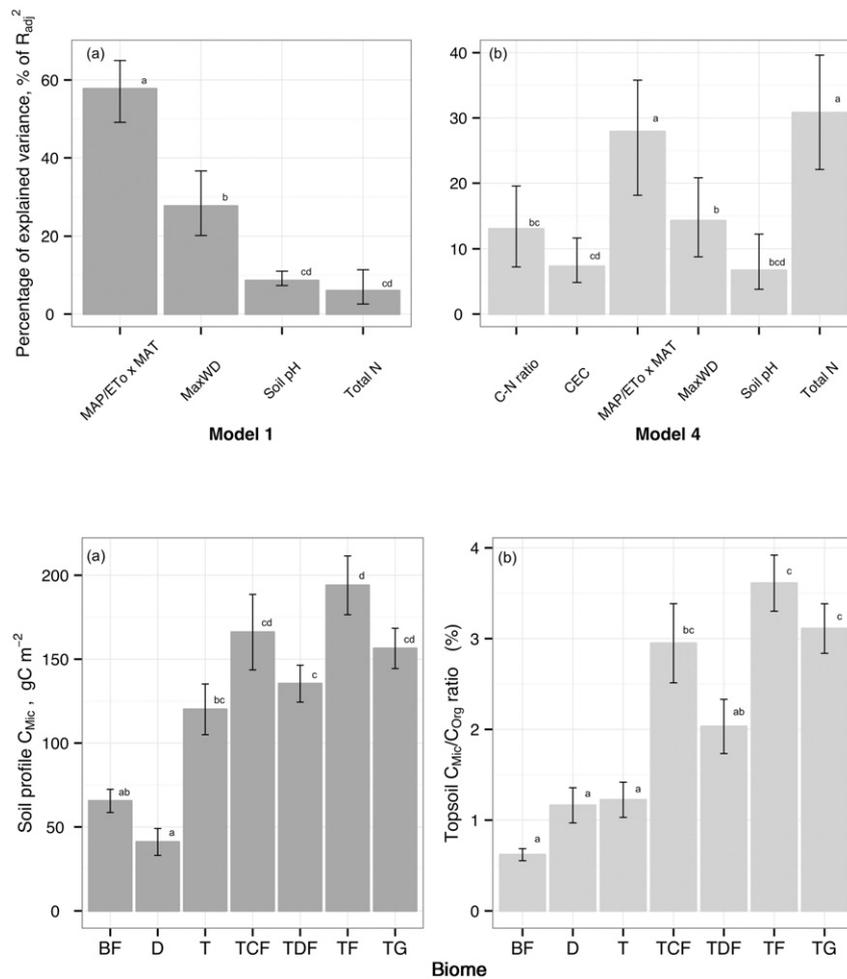
Figure 2 Relative variance component analysis of explanatory variables for (a) Model 1 estimating global soil profile microbial biomass carbon (C_{Mic}) [adjusted coefficient of determination (R_{adj}^2) = 39%] and (b) Model 4 estimating global topsoil C_{Mic}/soil organic carbon (C_{Org}) ratio (R_{adj}^2 = 50.7%). In both cases, metrics have been normalized to sum 100%. Confidence intervals (95%) were obtained after 1000 bootstrap runs. Equal lower-case letter indicate that confidence intervals for the difference between the contributions of each regressor includes zero ($P < 0.05$). Abbreviations as explained in Table 1.

Figure 3 Interbiome differences in (a) soil profile microbial biomass carbon (C_{Mic}) ($n = 414$) and (b) topsoil C_{Mic}/soil organic carbon (C_{Org}) ($n = 397$). Values presented are arithmetic means ± 1 standard error of the mean (SEM). Similar lower-case letter indicate that means are not significantly different from each other (Tukey's post-hoc test, $\alpha = 0.05$). Biome classification as follows: boreal forest (BF), desert (D), tundra (T), temperate coniferous forest (TCF), temperate deciduous forest (TDF), tropical forest (TF), and temperate grassland (TG).

high. More importantly, the residual variance of the cross-validation did not show deviations from null patterns, indicating unbiased estimates.
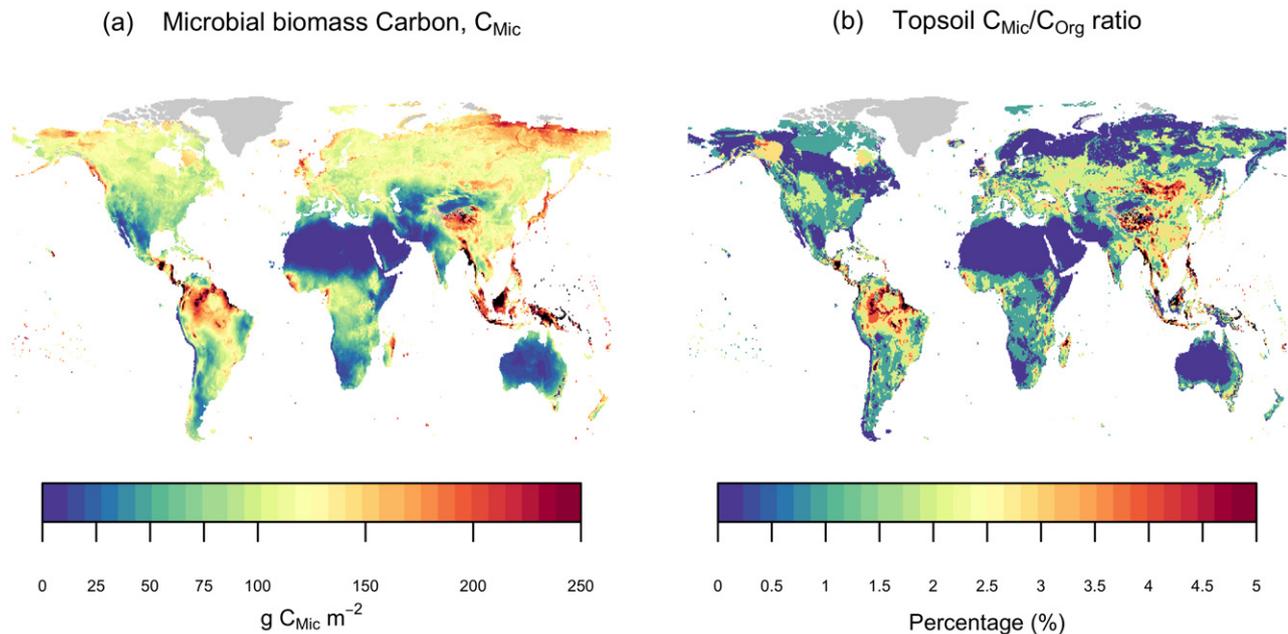
Global estimates of microbial biomass

The spatially explicit global patterns of microbial biomass derived from our multivariate models are presented in Fig. 4. Hotspots of soil profile C_{Mic} were in tropical regions; however, comparable abundances were evident in some subtropical, temperate and boreal regions. The lowest C_{Mic} was estimated for arid and semi-arid regions (Fig. 4a). Our multivariate models led to an estimated global soil profile microbial biomass pool of 14.6 Pg C_{Mic} (± 0.007 Pg; 95% CI). For the topsoil C_{Mic}/C_{Org}, patterns indicated highest ratios in tropical and subtropical regions where soil mineralization rates are typically higher (Aerts & Chapin, 2000). Conversely, low ratios persisted in arid regions but also in most boreal and tundra regions (Fig. 4b). For the latter regions, low ratios were expected given substrate accumulation because of temperature and moisture constraints (Deyn *et al.*, 2008), as was also apparent in our models. Finally, our model estimated the global topsoil C_{Mic}/C_{Org} at 1.2% (0–2; 75% CI).

DISCUSSION

Global environmental drivers of microbial abundance

By analysing the continuous relations between environmental drivers and microbial biomass, we were able to explore the variations between and within biomes without any *a priori* classification and reduced the residual variance in biomass estimates. In our analysis, moisture-related variables exhibited the strongest linear relationships with C_{Mic} and C_{Mic}/C_{Org}. This may be no surprise, as moisture availability can strongly affect microbial activity in both wet and dry climates (Bachar *et al.*, 2010; Blankinship *et al.*, 2011) while also influencing nutrient availability. Although high soil moisture is known to restrict microbial activity (Drenovsky *et al.*, 2010; Wu *et al.*, 2011), we found no evidence for such limitations, likely because anaerobiosis is sufficiently infrequent to limit microbial biomass in surface soils. Furthermore, because both C_{Mic} and C_{Mic}/C_{Org} were positively related to annual moisture availability, this suggests that microbial biomass increases more with precipitation than SOC. By including the inherent influence of temperature on annual moisture availability (MAP/ETo $\times$ MAT interaction) we were able to better explain the variability in microbial biomass.

(a) Microbial biomass Carbon, C$_{Mic}$

(b) Topsoil C$_{Mic}$/C$_{Org}$ ratio

Figure 4 Global 0.5°-lat × 0.5°-long spatially explicit estimates of soil microbial abundance. (a) Patterns depict soil profile microbial biomass carbon (C$_{Mic}$) derived from Model 1 (as provided in Table 1). Areas in black represent estimates greater than 250 g C$_{Mic}$ m^{-2}. The average C$_{Mic}$ content was 97.2 g C$_{Mic}$ m^{-2} [± 0.45; 95% confidence interval (CI)]. (b) Patterns depict global estimates for topsoil C$_{Mic}$/C$_{Org}$ derived from Model 4 (as provided in Table 1). Areas in black represent estimates greater than 5%. For both panels, grey areas represent ice-covered surfaces.

Because MAT alone showed little correlation with C$_{Mic}$, while annual ETo, MAP/ETo and the MASMD were all strongly correlated (Appendix S2), this suggests that temperature predominantly affects C$_{Mic}$ indirectly by influencing evapotranspiration rates. Conversely, both temperature variables and moisture were important drivers of C$_{Mic}$/C$_{Org}$, in contrast with effects found for C$_{Mic}$. This suggests that the influence of temperature on C$_{Mic}$/C$_{Org}$ is mainly through its effects on SOC, which is known to be strongly dependent on temperature (Parton *et al.*, 2007). Furthermore, any variable that expressed the seasonality of precipitation (e.g. coefficient of variation or precipitation in the wettest quarter), as well as of temperature (e.g. temperature of the coldest quarter), was less able to account for the variance in either C$_{Mic}$ or C$_{Mic}$/C$_{Org}$ than mean annual variables (Appendix S1).

Soil parameters were also important drivers of C$_{Mic}$ and especially of C$_{Mic}$/C$_{Org}$, suggesting that soil conditions have a stronger influence on the amount of organic carbon immobilized in microbial biomass than on biomass itself. The positive relationships we found between C$_{Mic}$, SOC and total nitrogen content are in line with previous evidence (Wardle, 1992; Cleveland & Liptzin, 2007; Fierer *et al.*, 2009). Multivariate models that included total nitrogen instead of SOC explained a higher variance of C$_{Mic}$ and were a better fit despite total nitrogen being highly correlated with SOC ($R^2 = 95\%$; $F_{1,392} = 7281$; $P < 0.0001$). Thus, while the relation between C$_{Mic}$ and total nitrogen also reflects an influence of SOC, our results suggest that effects of nitrogen limitation are more consistent. However,

these effects seem to decrease at high substrate concentrations because low microbial biomass was found in tundra soils at high total nitrogen (and coinciding high SOC; Appendix S2). These areas are characterized by climatic stress, such as prolonged low temperatures and long periods of moisture shortage because of frost, which can restrict decomposition leading to accumulation of organic matter (Wardle, 1992; Cleveland & Liptzin, 2007; Fierer *et al.*, 2009). Also, this accumulated carbon may be more recalcitrant (lower substrate quality; Franzluebbers *et al.*, 2001; Fierer *et al.*, 2009). As depicted in our maps, low C$_{Mic}$/C$_{Org}$ ratios are evident for these regions under climatic stress, while the estimated C$_{Mic}$ is similar to that found in tropical regions.

Unlike for C$_{Mic}$, the effect of soil nitrogen on C$_{Mic}$/C$_{Org}$ was negative. Given that C$_{Mic}$/C$_{Org}$ is generally considered an index of microbial activity, these results suggest that in addition to limitations by soil nitrogen on microbial biomass, nitrogen may also hamper microbial activity. A biome-deviation analysis revealed that for all biomes, the relationship was indeed negative except for desert soils (see Appendix S1). The latter is expected because in arid regions, soil nitrogen is considered a limiting resource (Gallardo & Schlesinger, 1992). Thus, increased nitrogen availability could allow for an increase in C$_{Mic}$/C$_{Org}$ if other environmental conditions are favourable. A negative relationship does not seem to be due to impaired decomposition by anoxic conditions with accumulated soil organic matter (as occurs in e.g. humid tropical forest, Pett-Ridge *et al.*, 2006), given that this cannot explain the negative impacts in other biomes. However, the negative relationship is consistent with nitrogen fertilization

studies (Dijkstra *et al.*, 2005; Treseder, 2008; Ramirez *et al.*, 2010). This may be explained by a shift in the soil microbial community to one with lower standing biomass but higher turnover rates given higher nutrient availabilities (Allison *et al.*, 2008; Ramirez *et al.*, 2010; Eisenhauer *et al.*, 2012) or a reduction in the allocation of belowground labile carbon with increasing soil nitrogen contents (Dijkstra *et al.*, 2005; Allison *et al.*, 2008; Eisenhauer *et al.*, 2012), which may subsequently increase the carbon demand of microbial biomass and induce shifts in community structure (Allison *et al.*, 2008). Given that aforementioned processes are intimately linked, they may represent complementary responses (with potentially different importance to subsets of our data) and may together help explain the overall negative global relationship between total nitrogen and C_{Mic}/C_{Org}.

Soil resource quality (C-N ratio) also exhibited an important influence on both microbial variables. High C-N ratios tend to indicate low organic matter decomposability, whereas higher substrate quality enhances microbial activity and growth (Parton *et al.*, 2007; Dequiedt *et al.*, 2011). Also, environmental stress can limit decomposition causing an accumulation of soil carbon and a higher C-N ratio because of accumulation of undecomposed plant litter. This might restrain microbial biomass abundance by reducing substrate quality. These mechanisms account for the negative relationship between our response variables and soil C-N ratio. Furthermore, according to our models, C_{Mic}/C_{Org} and C_{Mic} increased with pH. Acidic conditions are known to restrict organic matter decomposition and microbial (enzymatic) activity. As pH increases, microbial activity follows, allowing the immobilization of a bigger proportion of organic carbon in microbial biomass (Wardle, 1992; Sinsabaugh *et al.*, 2008). In accordance with previous studies (Anderson & Domsch, 1993; Baath & Anderson, 2003), this can translate into an overall positive effect of pH on microbial biomass. Soil texture variables also exhibited important influence on C_{Mic}/C_{Org} and on C_{Mic} (Appendix S2). It is likely that the relation between CEC and C_{Mic}/C_{Org} reflects an enhanced retention of organic matter and reduced quality (Wardle, 1992; Six *et al.*, 2002).

The patterns mentioned earlier might be the result of complex interacting effects in the soil environment and with climate (e.g. the climate, soil organic matter and pH relationship; Insam, 1990; Jobbagy & Jackson, 2000; Franzluebbers *et al.*, 2001). Nonetheless, the drivers described here are likely the main actors because of the imposed restraints on collinearity between environmental variables in our models (see Methods).

Future improvements in soil microbial biomass estimates

We recognize two important sources of uncertainty. First, given the scale of the environmental information used to build our models, local heterogeneity was not included, which likely explains an important proportion of residual variance in our models, particularly in the validation outcomes. However, we expect that this did not affect our predicted global patterns

because our predictions accurately portrayed average observations (see 1:1 patterns in Fig. 1). And, in spite of this discrepancy in scale, our validation showed similar patterns to the observed SIR as reported earlier for local datasets (Anderson & Joergensen, 1997; Fierer *et al.*, 2003, 2009). Considering uncertainties in our models, the correlation between our biomass estimates and SIR data are in line with previous studies (thus further supporting the contention that our estimates and spatial patterns are unbiased). Second, to standardize measurements, we extrapolated microbial biomass estimates from the depth-interval sampled to the first meter. This might have biased estimates for specific sites, e.g. where data were obtained from organic layers on top of mineral soils. However, even though depth distributions will differ among sites, and biomes, it presents a shape similar to the global average SOC depth distribution (Jobbagy & Jackson, 2000) and likely only contributed to residual variance in our models. Moreover, including soil parameters of subsoil conditions did not improve our models (see Methods). Future work elucidating microbial biomass depth distribution in different soil environments will help making suitable extrapolations and, thus, better comparisons. A better characterization of soil parameters along the profile might help improving the resolution of our multivariate models.

Spatial patterns and estimates of microbial biomass

Our unbiased multivariate models allowed us to build global spatially explicit predictions of microbial biomass in soils. These patterns show that soil microbial biomass is not primarily driven by temperature (as reflected in latitude), but instead biomass is more heterogeneous through the effects of moisture availability and soil nutrients (Fig. 4). Similarly, precipitation and soil nutrients have been shown to influence patterns of plant biomass (Moles *et al.*, 2009; Vicca *et al.*, 2012). To assess whether global patterns in plant and soil microbial abundance coincide, we tested the relation between net primary production of potential vegetation (NPP in g m^{-2} year^{-1}, from the Lund-Potsdam-Jena-Dynamic Global Vegetation Model, Sitch *et al.*, 2003) and soil C_{Mic}. Their correlation coefficient was comparable ($r = 0.5$, $P < 0.0001$) to those of annual moisture variables (Appendix S2). However, a large portion of the variance in NPP is already accounted for by annual moisture supply and availability variables (e.g. MAP and MAP/ETo× MAT, in bivariate regressions: $R^2 = 0.52$ and 0.45, respectively; $P < 0.0001$). Indeed, including NPP in our models did not increase the amount of explained variance or their fit. Instead, replacing the main driver (i.e. MAP/ETo × MAT for the C_{Mic} model) with NPP led to a decrease in explained variance from $R_{adj}^2 = 0.39$ ($P < 0.0001$; AIC −876), to $R_{adj}^2 = 0.28$ ($P < 0.0001$; AIC = −816). Compared with soil microbial biomass, plant biomass seems to be more constrained by temperature (Law *et al.*, 2002), as MAT was more strongly correlated to NPP ($r = 0.39$; $P < 0.0001$) than to C_{Mic} ($r = 0.07$; $P > 0.05$). Together, this seems to reveal that (1) the extent and direction of these relationships with environmental drivers is different for plant NPP and soil microbial biomass, and

(2) that annual climatic variables are better able than NPP to account for the global variance in soil microbial abundance.

Our comprehensive analysis provides a global estimate of 14.6 Pg C_{Mic} in the soil profile. This estimate falls within the range of previous estimates (Wardle, 1992; Whitman *et al.*, 1998; Xu *et al.*, 2013), which had been hampered by low representation in particular biomes and lack of spatially explicit analysis of drivers. In addition, according to the well-conserved C : N : P ratio of microbial biomass (Cleveland & Liptzin, 2007), our findings suggest that the global soil microbial nitrogen and phosphorous pools are around 2.0 and 0.6 Pg, respectively. Our study also estimate the global topsoil C_{Mic}/C_{Org} ratio at 1.2% (0–2%; 75% CI), which is in line with previous studies that estimated it for topsoil (cumulatively, 0.3–5%; Insam, 1990; Wardle, 1992; Zak *et al.*, 1994; Xu *et al.*, 2013) and for the soil profile (0.6–1.1%; Fierer *et al.*, 2009), respectively.

Microbial biomass estimates used in our study are mostly representative for the active season, which matters most for ecosystem fluxes and services. Thereto, our estimates are well suited to understand and to further improve quantitative estimates of ecosystem fluxes and services, now and in the face of global environmental change. In fact, our quantitative estimates may serve many applications as soil microorganisms mediate the rates of carbon and nutrient cycling in terrestrial ecosystems (Wardle, 1992; Falkowski *et al.*, 2008). Consequently, microbial biomass serves as a critical control on the feedbacks between global change, organic matter decomposition and ecosystem productivity (Manzoni & Porporato, 2009; Todd-Brown *et al.*, 2012). Current global biogeochemical and vegetation models have only a coarse, if any, representation of soil microbial properties and most commonly frame decomposition of organic matter as first-order decay, that is, only as a function of environment (temperature and moisture) and not of microbial biomass (Parton *et al.*, 1993; Sitch *et al.*, 2003), hampering the predictive ability of these models (Manzoni & Porporato, 2009; Todd-Brown *et al.*, 2012). The problems with scaling up information on microbial abundance and activity from sparse local data (Todd-Brown *et al.*, 2012) have made it difficult to integrate soil microbial biomass into global scale decomposition models. However, our global estimates of soil microbial biomass provide important data for integration into large-scale carbon and nutrient models, which may imply a major step forward in our ability to predict the global carbon balance, now and in a future climate. Thereto, our global maps of C_{Mic} and topsoil C_{Mic}/C_{Org} are available in Appendix S3.

In summary, our study shows how using average environmental variables allows us to explain a large portion of the global variance in soil microbial biomass. By identifying generic environmental constraints on soil microbial abundance, we show the stronger constraints exerted by soil variables and temperature on C_{Mic}/C_{Org} than on C_{Mic}, suggesting that topsoil microbial activity is further constrained by local soil conditions and low temperatures. Also, by deriving a spatially explicit global map of soil profile C_{Mic} and topsoil C_{Mic}/C_{Org} distribution, we provide an improved spatially explicit estimate of the global biological soil carbon pool. Our estimates, in combination with new under-standing of substrate and decomposer stoichiometry (Cleveland & Liptzin, 2007; Manzoni *et al.*, 2008; Sinsabaugh *et al.*, 2008), will together improve our ability to predict the controls and rates of organic matter decomposition, nutrient mineralization and soil-atmosphere carbon exchange.

REFERENCES

Aerts, R. & Chapin, F.S. (2000) The mineral nutrition of wild plants revisited: a re-evaluation of processes and patterns. *Advances in Ecological Research*, **30**, 1–67.

Allen, R.G., Pereira, L.S., Raes, D. & Smith, M. (1998) *Crop evapotranspiration – guidelines for computing crop water requirements – FAO Irrigation and drainage paper 56*. Food and Agriculture Organization, Rome, Italy.

Allison, S.D., Czimczik, C.I. & Treseder, K.K. (2008) Microbial activity and soil respiration under nitrogen addition in Alaskan boreal forest. *Global Change Biology*, **14**, 1156–1168.

Anderson, T.H. & Domsch, K.H. (1993) The metabolic quotient for CO_2 (QCO_2) as a specific activity parameter to assess the effects of environmental conditions, such as pH, on the microbial biomass of forest soils. *Soil Biology and Biochemistry*, **25**, 393–395.

Anderson, T.H. & Joergensen, R.G. (1997) Relationship between SIR and FE estimates of microbial biomass C in deciduous forest soils at different pH. *Soil Biology and Biochemistry*, **29**, 1033–1042.

Baath, E. & Anderson, T.H. (2003) Comparison of soil fungal/bacterial ratios in a pH gradient using physiological and PLFA-based techniques. *Soil Biology and Biochemistry*, **35**, 955–963.

Bachar, A., Al-Ashhab, A., Soares, M.I.M., Sklarz, M.Y., Angel, R., Ungar, E.D. & Gillor, O. (2010) Soil microbial abundance and diversity along a low precipitation gradient. *Microbial Ecology*, **60**, 453–461.

Batjes, N.H. (ed.) (1995) *A homogenized soil data file for global environmental research: a subset of FAO, ISRIC and NCRS profiles (version 1.0)*. Working Paper and Preprint 95/10b. International Soil Reference and Information Centre, Wageningen.

Blankinship, J.C., Niklaus, P.A. & Hungate, B.A. (2011) A meta-analysis of responses of soil biota to global change. *Oecologia*, **165**, 553–565.

Cleveland, C.C. & Liptzin, D. (2007) C:N:P stoichiometry in soil: is there a 'Redfield ratio' for the microbial biomass? *Biogeochemistry*, **85**, 235–252.

Dequiedt, S., Saby, N.P.A., Lelievre, M., Jolivet, C., Thioulouse, J., Toutain, B., Arrouays, D., Bispo, A., Lemanceau, P. & Ranjard, L. (2011) Biogeographical patterns of soil molecular microbial biomass as influenced by soil characteristics and management. *Global Ecology and Biogeography*, **20**, 641–652.

Deyn, G.B.D., Cornelissen, J.H.C. & Bardgett, R.D. (2008) Plant functional traits and soil carbon sequestration in contrasting biomes. *Ecology Letters*, **11**, 516–531.

Dijkstra, F.A., Hobbie, S.E., Reich, P.B. & Knops, J.M.H. (2005) Divergent effects of elevated CO_2, N fertilization, and plant diversity on soil C and N dynamics in a grassland field experiment. *Plant and Soil*, **272**, 41–52.

Drenovsky, R.E., Steenwerth, K.L., Jackson, L.E. & Scow, K.M. (2010) Land use and climatic factors structure regional patterns in soil microbial communities. *Global Ecology and Biogeography*, **19**, 27–39.

Eisenhauer, N., Cesarz, S., Koller, R., Worm, K. & Reich, P.B. (2012) Global change belowground: impacts of elevated CO_2, nitrogen, and summer drought on soil food webs and biodiversity. *Global Change Biology*, **18**, 435–447.

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

Fierer, N. & Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences USA*, **103**, 626–631.

Fierer, N., Schimel, J.P. & Holden, P.A. (2003) Variations in microbial community composition through two soil depth profiles. *Soil Biology and Biochemistry*, **35**, 167–176.

Fierer, N., Strickland, M.S., Liptzin, D., Bradford, M.A. & Cleveland, C.C. (2009) Global patterns in belowground communities. *Ecology Letters*, **12**, 1238–1249.

Franzluebbers, A.J., Haney, R.L., Honeycutt, C.W., Arshad, M.A., Schomberg, H.H. & Hons, F.M. (2001) Climatic influences on active fractions of soil organic matter. *Soil Biology and Biochemistry*, **33**, 1103–1111.

Gallardo, A. & Schlesinger, W.H. (1992) Carbon and nitrogen limitations of soil microbial biomass in desert ecosystems. *Biogeochemistry*, **18**, 1–17.

Groemping, U. (2006) Relative importance for linear regression in R: the package relaimpo. *Journal of Statistical Software*, **17**, issue 1.

Insam, H. (1990) Are the soil microbial biomass and basal respiration governed by the climatic regime? *Soil Biology and Biochemistry*, **22**, 525–532.

Jenkinson, D.S. & Powlson, D.S. (1976) The effects of biocidal treatments on metabolism in soil – I. Fumigation with chloroform. *Soil Biology and Biochemistry*, **8**, 167–177.

Jobbagy, E.G. & Jackson, R.B. (2000) The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications*, **10**, 423–436.

Joergensen, R.G., Wu, J. & Brookes, P.C. (2011) Measuring soil microbial biomass using an automated procedure. *Soil Biology and Biochemistry*, **43**, 873–876.

Law, B.E., Falge, E., Gu, L. *et al.* (2002) Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation. *Agricultural and Forest Meteorology*, **113**, 97–120.

Manzoni, S. & Porporato, A. (2009) Soil carbon and nitrogen mineralization: theory and models across scales. *Soil Biology and Biochemistry*, **41**, 1355–1379.

Manzoni, S., Jackson, R.B., Trofymow, J.A. & Porporato, A. (2008) The global stoichiometry of litter nitrogen mineralization. *Science*, **321**, 684–686.

Martens, R. (1995) Current methods for measuring microbial biomass C in soil: potentials and limitations. *Biology and Fertility of Soils*, **19**, 87–99.

Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E., Pitman, A., Hemmings, F.A. & Leishman, M.R. (2009) Global patterns in plant height. *Journal of Ecology*, **97**, 923–932.

New, M., Hulme, M. & Jones, P. (1999) Representing twentieth century space-time climate variability. Part I: development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate*, **12**, 829–856.

Parton, W., Silver, W.L., Burke, I.C., Grassens, L., Harmon, M.E., Currie, W.S., King, J.Y., Adair, E.C., Brandt, L.A., Hart, S.C. & Fasth, B. (2007) Global-scale similarities in nitrogen release patterns during long-term decomposition. *Science*, **315**, 361–364.

Parton, W.J., Scurlock, J.M.O., Ojima, D.S., Gilmanov, T.G., Scholes, R.J., Schimel, D.S., Kirchner, T., Menaut, J.C., Seastedt, T., Moya, E.G., Kamnalrut, A. & Kinyamario, J. (1993) Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global Biogeochemical Cycles*, **7**, 785–809.

Pett-Ridge, J., Silver, W.L. & Firestone, M.K. (2006) Redox fluctuations frame microbial community impacts on N-cycling rates in a humid tropical forest soil. *Biogeochemistry*, **81**, 95–110.

R Development Core Team (2011) *R: a language and environment for statistical computing*. R Development Core Team.

Ramirez, K.S., Lauber, C.L., Knight, R., Bradford, M.A. & Fierer, N. (2010) Consistent effects of nitrogen fertilization on soil bacterial communities in contrasting systems. *Ecology*, **91**, 3463–3470.

Sinsabaugh, R.L., Lauber, C.L., Weintraub, M.N. *et al.* (2008) Stoichiometry of soil enzyme activity at global scale. *Ecology Letters*, **11**, 1252–1264.

Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J.O., Levis, S., Lucht, W. & Sykes, M.T. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, **9**, 161–185.

Six, J., Conant, R.T., Paul, E.A. & Paustian, K. (2002) Stabilization mechanisms of soil organic matter: implications for C-saturation of soils. *Plant and Soil*, **241**, 155–176.

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.

Treseder, K.K. (2008) Nitrogen additions and microbial biomass: a meta-analysis of ecosystem studies. *Ecology Letters*, **11**, 1111–1120.

Vance, E.D., Brookes, P.C. & Jenkinson, D.S. (1987) An extraction method for measuring soil microbial biomass C. *Soil Biology and Biochemistry*, **19**, 703–707.

Vicca, S., Luyssaert, S., Peñuelas, J., Campioli, M., Chapin, F.S., Ciais, P., Heinemeyer, A., Högberg, P., Kutsch, W.L., Law, B.E., Malhi, Y., Papale, D., Piao, S.L., Reichstein, M., Schulze, E.D. & Janssens, I.A. (2012) Fertile forests produce biomass more efficiently. *Ecology Letters*, **15**, 520–526.

Wardle, D.A. (1992) A comparative assessment of factors which influence microbial biomass Carbon and Nitrogen levels in soil. *Biological Reviews*, **67**, 321–358.

Whitman, W.B., Coleman, D.C. & Wiebe, W.J. (1998) Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences USA*, **95**, 6578–6583.

Wu, Z., Dijkstra, P., Koch, G.W., Penuelas, J. & Hungate, B.A. (2011) Responses of terrestrial ecosystems to temperature and precipitation change: a meta-analysis of experimental manipulation. *Global Change Biology*, **17**, 927–942.

Xu, X., Thornton, P.E. & Post, W.M. (2013) A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. *Global Ecology and Biogeography*, **22**, 737–749.

Zak, D.R., Tilman, D., Parmenter, R.R., Rice, C.W., Fisher, F.M., Vose, J., Milchunas, D. & Martin, C.W. (1994) Plant production and soil microorganisms in late-successional ecosystems: a continental-scale study. *Ecology*, **75**, 2333–2347.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1 Secondary multivariate models that predict patterns of C_{Mic} and C_{Mic}/C_{Org}. (A) Observed and predicted soil profile C_{Mic} estimates from Model-2: slope = 1.0 ($\pm 5.9 \times 10^{-2}$, $\beta \pm 1SEM$), intercept: -1.6×10^{-8} ($\pm 1.2 \times 10^{-2}$), $R^2 = 43.9\%$, $F_{1-365} = 285$; $P < 0.0001$. (B) Observed and predicted topsoil C_{Mic}/C_{Org} ratios from Model-3: Slope = 1.0($\pm 5.2 \times 10^{-2}$; $\beta \pm 1SEM$), intercept: -1.8×10^{-5}, $R^2 = 49.9\%$, $F_{1-365} = 363$; $P < 0.0001$. Shadowed area is the 95% C.I and all axes are in $\log_{10}$ scale. Panels (C) and (D) are the relative importance of regressors included in Model 2 and 3, respectively. In both, panels equal lower-case letter indicate that confidence intervals for the difference between the contributions of each regressor includes zero ($P < 0.05$). Also, metrics are normalized to sum 100% of explained variance and 95% C.I. were obtained after 1000 bootstrap runs. Legend and biome classification as follows: *boreal forest* (□, BF), *desert* (○, D), *tundra* (△, T), *temperate coniferous forest* (+, TCF), *temperate deciduous forest* (✕, TDF), *tropical forest* (◊, TF), and *temperate grassland* (▽, TG). MAP: mean annual precipitation; ETo: annual evaporative demand; MAT: mean annual temperature; MaxWD: maximum monthly frequencies of wet days; Total N: soil total nitrogen content; CEC: soil cation exchange capacity; pH: soil pH; C-N ratio: soil Carbon-Nitrogen ratio.

Figure S2 Relationship between topsoil total nitrogen and C_{Mic}/C_{Org}. Biome category is presented on top of each panel, label is as follows: (BF) boreal forest, (D) desert, (T) tundra, (TCF) temperate coniferous forest, (TDF) temperate deciduous forest, (TF) tropical forest and (TG) temperate grassland. Regression coefficients and standard error ($\beta \pm 1SEM$) from the corresponding bivariate regressions are as follows: BF: $-0.79(\pm 0.23)$, $P < 0.001$; D: $+1.22(\pm 0.44)$, $P < 0.001$; T: $-0.58(\pm 0.13)$, $P < 0.0001$; TCF: $-0.40(\pm 0.41)$, $P > 0.05$; TDF: $-1.14(\pm 0.13)$, $P < 0.0001$; TF: $-0.38(\pm 0.13)$, $P < 0.001$; TG: $-0.76(\pm 0.14)$, $P < 0.0001$. In all graphs, grey lines represent the bivariate regression line and the shaded-area the 95% C.I.

Table S1 Environmental drivers of soil profile microbial biomass carbon (g C_{Mic} m^{-2}). Bivariate regressions of annual and seasonal climatic variables versus microbial biomass.

Appendix S1 Supplementary information on materials, methods, and results.

Appendix S2 Scatterplots and Person's correlation factors of both soil microbial biomass carbon (C_{Mic}) and soil microbial biomass carbon-soil organic carbon (C_{Mic}/C_{Org}), and environmental variables.

Appendix S3 ASCII maps of soil profile microbial biomass carbon (C_{Mic}), and for the topsoil soil microbial biomass carbon-soil organic carbon (C_{Mic}/C_{Org}).

BIOSKETCHES

Hector Serna-Chavez' main research interests focus on the underlying role of functional diversity on ecosystem services, including the spatiotemporal characteristics of delivery and its resilience under global environmental change.

Noah Fierer's research focuses on the ecology and biogeography of microorganisms in a range of systems including the atmosphere, soi, and the human body, exploring linkages between microbial diversity and the functioning of microbial communities.

Peter van Bodegom works at the interface between community ecology, macroecology and earth system modelling. With his group, he aims to quantify globally applicable functional relationships between vegetation, soil microorganisms, and their environment by targeted experiments, meta-analyses and process-based modelling.

Editor: Martin Sykes