Microbial communities in terrestrial surface soils are not widely limited by carbon

Yongxing Cui\textsuperscript{1} | Shushi Peng\textsuperscript{1} | Manuel Delgado-Baquerizo\textsuperscript{2,3} | Matthias C. Rillig\textsuperscript{4} | César Terrer\textsuperscript{5} | Biao Zhu\textsuperscript{6} | Xin Jing\textsuperscript{7} | Ji Chen\textsuperscript{8} | Jinquan Li\textsuperscript{9} | Jiao Feng\textsuperscript{10} | Yue He\textsuperscript{1} | Linchuan Fang\textsuperscript{11} | Daryl L. Moorhead\textsuperscript{12} | Robert L. Sinsabaugh\textsuperscript{13} | Josep Peñuelas\textsuperscript{14,15}\

\textsuperscript{1}Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing, China\
\textsuperscript{2}Laboratorio de Biodiversidad y Funcionamiento Ecosistémico, Instituto de Recursos Naturales y Agrobiología de Sevilla (IRNAS), CSIC, Sevilla, Spain\
\textsuperscript{3}Unidad Asociada CSIC-UPO (BioFun), Universidad Pablo de Olavide, Sevilla, Spain\
\textsuperscript{4}Institute of Biology, Freie Universität Berlin, Berlin, Germany\
\textsuperscript{5}Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Boston, Massachusetts, USA\
\textsuperscript{6}Institute of Ecology, College of Urban and Environmental Sciences, and Key Laboratory for Earth Surface Processes of the Ministry of Education, Peking University, Beijing, China\
\textsuperscript{7}State Key Laboratory of Grassland Agro-Ecosystems, and College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou, Gansu, China\
\textsuperscript{8}Department of Agroecology, Aarhus University, Tjele, Denmark\
\textsuperscript{9}Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, National Observations and Research Station for Wetland Ecosystems of the Yangtze Estuary, Institute of Biodiversity Science and Institute of Eco-Chongming, School of Life Sciences, Fudan University, Shanghai, China\
\textsuperscript{10}College of Resources and Environment, Huazhong Agricultural University, Wuhan, China\
\textsuperscript{11}School of Resource and Environmental Engineering, Wuhan University of Technology, Wuhan, China\
\textsuperscript{12}Department of Environmental Sciences, University of Toledo, Toledo, Ohio, USA\
\textsuperscript{13}Department of Biology, University of New Mexico, Albuquerque, New Mexico, USA\
\textsuperscript{14}CSIC, Global Ecology Unit CREAF-CSIC-UAB, Bellaterra, Catalonia, Spain\
\textsuperscript{15}CREAF, Cerdanyola del Vallès, Catalonia, Spain

Abstract

Microbial communities in soils are generally considered to be limited by carbon (C), which could be a crucial control for basic soil functions and responses of microbial heterotrophic metabolism to climate change. However, global soil microbial C limitation (MCL) has rarely been estimated and is poorly understood. Here, we predicted MCL, defined as limited availability of substrate C relative to nitrogen and/or phosphorus to meet microbial metabolic requirements, based on the thresholds of extracellular enzyme activity across 847 sites (2476 observations) representing global natural ecosystems. Results showed that only about 22% of global sites in terrestrial surface soils show relative C limitation in microbial community. This finding challenges the conventional hypothesis of ubiquitous C limitation for soil microbial metabolism. The limited geographic extent of C limitation in our study was mainly attributed to plant litter, rather than soil organic matter that has been processed by microbes, serving as the dominant C source for microbial acquisition. We also identified a significant latitudinal pattern of predicted MCL with larger C limitation at mid- to high latitudes,
1 | INTRODUCTION

Global soils store at least twice as much carbon (C) as Earth’s atmosphere (Bond-Lamberty et al., 2018; Friedlingstein et al., 2022), with C pool size determined mainly by the balance of C inputs from plants and C losses via soil microbial heterotrophic respiration ($R_h$). Compared with plants, whose production is generally limited by soil nitrogen (N) and phosphorus (P) (Du et al., 2020; Elser et al., 2007), heterotrophic microbial communities in soil are considered to be inherently C limited (Hobbie & Hobbie, 2013; Soong et al., 2020). The C constraints on microbial catabolic processes thus potentially determine the retention of detrital organic C and, ultimately, the size of the soil C pool (Liang et al., 2017; Sinsabaugh et al., 2009; Sinsabaugh & Follstad Shah, 2012) and have consequently gained increasing attention in the context of global climate change (Soong et al., 2020).

Metabolic theory predicts that labile C, as a source of both energy and structural material to microbes, is ubiquitously limiting to microbial metabolism in soils (Hobbie & Hobbie, 2013; Soong et al., 2020). Empirical studies, however, have not provided strong evidence for it (Cheng et al., 1996; Demoling et al., 2007; Ekblad & Nordgren, 2002; Hursh et al., 2017), partly because almost all of these studies involved manipulation experiments. Although such experiments can infer the extent of C limitation on microbial activities by directly measuring responses of microbial metabolism to added C, the addition of C at any level can induce cascading responses of microbial metabolism to other resources (Chen et al., 2020; Cui et al., 2022; Spohn & Kuzyakov, 2013), thus creating uncertainty in identifying the real C-limiting status. Meta-analyses based on manipulation experiments have also not found robust spatial patterns of microbial C limitation (MCL), potentially due to combining different microbial response indicators or using effect-size metrics not standardized by the level of C addition. On a global scale, C limitation probably constrains microbial functions in detrital food webs and their responses to climate change (Sinsabaugh & Follstad Shah, 2012; Soong et al., 2020). Understanding the patterns and mechanisms of MCL in global soils, therefore, remains a high priority.

Both theory and empirical evidence notably raise a fundamental question: are heterotrophic microbial communities widely limited by soil C, even in C-rich surface soils? According to the principles of matter and energy conservation and ecological stoichiometry (Allen & Gillooly, 2002; Peñuelas & Balandi, 2019; Sterner & Elser, 2002), MCL depends on the availability of substrate C relative to other nutrients (generally N and/or P), a notion also supported by many empirical studies, from molecules to ecosystems (Chakravali et al., 2022; Tamale et al., 2021). The degree of MCL may also be highly variable among substrates due to the broad stoichiometric gradient across diverse C components that range from plant litter to complex soil organic matter (SOM) (Fierer, 2017; Liang et al., 2017; Xu et al., 2013). We used the concept of the threshold element ratio (TER), defined as the elemental C:N or C:P ratio that controls the switch of microbial metabolism from the acquisition of energy (C) to limiting nutrients (N and/or P) (Allen & Gillooly, 2009; Sinsabaugh et al., 2009), to calculate an approximate C limitation of soil microbial communities fed by two broad categories of substrates, plant litter and SOM (Figure 1). Our calculations indicated that the supply of C for microbial metabolism should be generally adequate, relative to the supplies of N and P, when plant litter is the dominant substrate. In contrast, C limits microbial communities more frequently when SOM is the dominant substrate. In fact, microbes can acquire C more readily from plant litter than SOM with recalcitrant molecular structures protected by minerals and aggregates, when both are available (Cotrufo et al., 2019; Fierer, 2017; Fontaine et al., 2007). We thus hypothesized that C would not be the primary limiting resource to microbes in terrestrial surface soils.

To test our hypothesis and investigate the patterns and mechanisms of MCL in soils globally, we employed our proposed metric (MCL) that integrates four C-, N-, and P-acquiring ecoenzymatic activities (Table S1) based on TER. Liebig’s Law of the Minimum, and ecological stoichiometric theory (Allen & Gillooly, 2009; Cui et al., 2023; Sinsabaugh & Follstad Shah, 2012; von Liebig, 1840). The MCL, defined as limited availability of substrate C relative to N and/or P to meet microbial metabolic requirements, was calculated using a newly compiled global ecoenzyme database to estimate the global distribution of MCL in natural ecosystems (see Method section 2.2 for details). The predicted results were further compared against two independent global datasets of available soil nutrients and field-diagnosed microbial responses to plant litter manipulations (Table S2). We then assessed 29 environmental variables, including variables of climate, vegetation and soil, to identify potential mechanisms responsible for patterns of MCL. $R_h$ is a critical pathway of terrestrial C release and may depend on the levels of MCL, so we also linked the predicted MCLs to the average daily and annual rates of soil $R_h$. It should be noted that the responses of microorganisms directly utilizing dissolved organic matter, such as

whereas this limitation was generally absent in the tropics. Moreover, MCL significantly constrained the rates of soil heterotrophic respiration, suggesting a potentially larger relative increase in respiration at mid- to high latitudes than low latitudes, if climate change increases primary productivity that alleviates MCL at higher latitudes. Our study provides the first global estimates of MCL, advancing our understanding of terrestrial C cycling and microbial metabolic feedback under global climate change.

**KEYWORDS**

decomposer community, ecological stoichiometry, global climate change, heterotrophic respiration, resource limitations, soil carbon cycling, soil-climate feedback
rhizosphere symbionts, may not be captured by the ecoenzyme-based approach. Thus, this study focuses on the detritus food web only, not on microbes feeding on dissolved substrates.

2 METHODS

2.1 Data collection

2.1.1 Global data of community-level ecoenzymatic activities

We compiled a global database of ecoenzymatic activities at the community level from a survey of the literature using the Web of Science (http://isiknowledge.com) and the Google Scholar Resource Integrated Database (https://scholar.google.com). Combinations of keywords such as "extracellular enzyme," "exo-enzyme," "ecoenzyme," and "microbial nutrient limitation" were used to search studies published from 1980 to 2021. The criteria for inclusion were: (1) the studies included the activities of C-, N-, and P-acquiring enzyme (BG (β-1, 4-glucosidase), NAG (β-1, 4-N-acetylglucosaminidase) and/or LAP (L-leucine aminopeptidase), and AP (acid or alkaline phosphatase)) (Table S1), and (2) the ecoenzymatic activities were measured fluorometrically using a 200 μM solution of substrate labeled with 4-methylumbelliferone or 7-amino-4-methylcoumarin. The data were extracted from tables or figures (using GetData Graph Digitizer software v.2.25) of selected studies. We excluded data for intensively managed...
ecosystems (e.g., agroforests, fertilized plantations, sown pastures, croplands, and urban forests) to avoid misattributing natural nutrient limitations due to anthropogenic disturbances. Our database contained 2476 observations from 290 studies across 847 sites (Figure S51). We also recorded corresponding information from the literature on site locations (longitude and latitude) and vegetation types.

2.1.2 | Datasets of available soil nutrients and microbial responses

To test our predictions of MCL based on ecoenzyme activities, we simultaneously compiled two other datasets: (1) concentrations of labile soil C, N, and P across the globe (DOC, NO₃-N and/or NH₄-N, and Olsen-P) and (2) the responses of microbial biomass (including microbial biomass C [MBC] and phospholipid fatty acids [PLFAs]) and Rₘ to experimental additions of plant litter in natural ecosystems. For the first dataset, we converted indices into ratios of labile soil C and N (Lc:n) and labile soil C and P (Lc:p). We obtained totals of 1159 and 1147 data points of Lc:n and Lc:p, respectively. These data were extracted from the same pool of studies used to build the global ecoenzyme database, so their coordinates were identical. Published experiments were required to satisfy two criteria for inclusion in the second dataset from 290 publications: (1) the plant litter-addition experiments were conducted in the field and included litter-manipulation and control treatments in the same ecosystem under the same environmental conditions and included measurements of MBC, PLFAs, or Rₘ in both the litter-addition and control treatments and (2) the control treatment was not fertilized either before the start of the experiment or during the experiment. We excluded experiments in which litter and inorganic nutrients (N and/or P) were added in combination for assessing the effects of nutrients on microbes. Of course, plant litter contains not only C but also N and P, but here, we just consider litter effects as the C source on heterotrophic microbial metabolism because of much higher C concentrations than other nutrients. In total, we obtained 28 studies of paired observations of the responses of MBC to litter addition, 18 studies of paired observations of the responses of PLFAs to litter addition, and 17 studies of paired observations of the responses of Rₘ to litter addition, at a global scale (Table S2; Figure 3a).

2.1.3 | Environmental variables

We retrieved a variety of variables for climate, vegetation, topography, and soil from multiple sources of data at relatively fine spatial resolution (see Table S5 for details) to identify the effects of potential environmental drivers on MCL and the relationship between the limitation and geographical pattern. We obtained data for a total of 29 environmental variables at each of the spatial coordinates in the predicted MCL data set.

2.1.4 | Global data of average daily and annual Rₘ

We collected data for two types of Rₘ (average daily and annual Rₘ) from relevant publications or existing databases to identify their relationships to MCL. We collected data for average daily Rₘ from the same studies as those used for the global ecoenzyme database. These data were derived from field measurements of natural soils that had not been disturbed or manipulated by humans. After the data were filtered, we obtained a total of 517 observations of average daily Rₘ at 83 sites from 43 published studies (Figure S59) and expressed as mg C g⁻¹ day⁻¹. We extracted average annual Rₘ from the Soil Respiration Database (SRDB, Version 5.0, https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1827) contributed by Jian et al. (2021). We obtained a total of 2476 data points of average annual Rₘ at corresponding spatial coordinates with the MCL data set.

2.2 | Quantifying microbial C limitation using the ecoenzyme-based method

The MCL was quantified via our recently proposed ecoenzyme-based approach. The rationale and details can be found in Cui et al. (2023):

$$MCL = \ln \left( \frac{EEA_{C:N} \times EEA_{C:P}}{2.25 (n_0 \times p_0)} \right)$$

EEA_{C:N} = \frac{C – acquiring enzymes}{N – acquiring enzymes}

and EEA_{C:P} = \frac{C – acquiring enzymes}{P – acquiring enzymes}

The normalization constants n_0 and p_0 for the model were obtained from standardized major axis regressions of the ecoenzyme activities (Table S3), that is, n_0 = e^{intercept} in the regressions for ln(BG) versus ln(NAG+LAP) and p_0 = e^{intercept} in the regressions for ln(BG) versus ln(AP).

Finally, our predicted global values of MCL were normally distributed (Anderson-Darling normality test, A = 0.30 and p = 0.57) and had a global mean of -1.92 (ranging from -9.64 to 3.66) (Figure S3).

2.3 | Meta-analysis

We used a meta-analysis to evaluate the responses of MBC, PLFAs, and Rₘ to litter-addition at each paired observations (Hedges et al., 1999). For these analyses, our databases included means, standard deviations (SDs), and sample sizes (n) extracted from the
studies. We calculated SDs when standard errors (SEs) rather than SDs were reported:

$$SD = SE \sqrt{n}$$  \hspace{1cm} (2)

If neither SD nor SE was reported, we approximated the missing SDs by multiplying the reported means by the average coefficient of variance of our complete data set. We estimated the SDs for the control and the litter-addition treatments.

We evaluated the microbial responses to litter-addition using the natural logarithm of the response ratio (LRR):

$$LRR = \ln (\frac{X_t}{X_c}) = \ln (X_t) - \ln (X_c)$$  \hspace{1cm} (3)

where \(X_t\) and \(X_c\) are the arithmetic means in the litter-addition and control treatments, respectively.

The data for the meta-analysis were evaluated using the "rma. mv" function in the “metafor” R package (Viechtbauer, 2010). Several studies contributed more than one response ratio, so we included the variable "Article" as a random factor.

## 2.4 Environmental drivers of geographic variation in microbial C limitation

We assessed the effects of climate, vegetation, topography, and soil properties on MCL by training random-forest models with all 29 environmental variables using the “RandomForest” R package. The default node size, a key parameter, was accepted throughout the training based on parameter optimization (Figure S5). A 10-fold cross-validation was conducted to identify the best models and the potential importance of the variables using the "rfcv" function. Model significance and cross-validated \(R^2\) were assessed using 1500 permutations of the response variable and the "A3" package. The random-forest regression models constructed for predicting MCL explained 76.8% of the variation.

We calculated the values of increased node purity for each potential predictor in our constructed random-forest models using the "rfPermut" R package to identify the relative importance of all environmental variables in explaining the variations of MCL. The values were further transformed into relative importance (%) to facilitate interpretation (Figure 4a). Similarly, the significance of each predictor on the response variables was assessed using the "rfPermute" R package.

Finally, we divided 11 significant variables into eight categories (latitude, elevation), climate (mean annual temperature [MAT] and mean annual precipitation [MAP]), temperature seasonality (TSEA), gross primary production (GPP), plant-derived C (aboveground biomass [AGB], plant-litter production [Litter], and belowground biomass [BGB]), CaCO\(_3\) concentration, and pH) to explore cascading relationships between MCL and these key drivers. We used partial least squares path models to identify possible pathways where variables controlled MCL. A variance inflation factor (VIF) <10 for all variables in the model was required for removing the internal factor collinearity of the block (module). Loading in the outer models required VIFs >0.7 for removing independent variables with small contributions. The models were constructed using the "innerplot" function in the "plspm" R package (Sanchez et al., 2016).

## 2.5 Relationships of microbial C limitation with \(R_h\)

We identified correlations between MCL and \(R_h\) by first performing a linear-regression analysis between MCL and average daily and annual \(R_h\) (Figure 5a,b) and then conducted a partial correlation analysis of MCL with average daily and annual \(R_h\) by controlling the other two key factors (MAT and MAP) using the "pcor.test" function in the "ppcor" R package (Figure 5c). We also used a variation-partitioning analysis to quantify the independent and joint effects of MAT, MAP, and MCL on average daily and annual \(R_h\) using the "varpart" function in the "vegan" R package (Figure 5d,e). To simplify interpretation, we calculated the independent and joint relative influences of MAT, MAP, and MCL on \(R_h\):

$$\text{Independent relative influence} = \frac{\text{Independent explanation of each variable}}{\text{Total explanation of all variables}}$$

$$\text{Joint relative influence} = \frac{\text{Joint explanation of one with others}}{\text{Total explanation of all variables}}$$

## 2.6 Other statistical analyses

We used a quadratic function model to describe latitudinal trends in MCL (Figure 2a), and then identified the potential latitudinal breakpoints from a piecewise linear-regression analysis (Figure 2b; Table S4). The regression relationships between MCL and latitude were fitted and tested with linear models using the “segmented” R package (Muggeo, 2008). The confidence intervals of the breakpoints were calculated using 1000 bootstrap samples and the "SiZer" R package (Toms & Lesperance, 2003). A linear mixed-effects model was used to identify the differences in MCL among climatic zones (Figure 2c) and the main types of natural vegetation (Figure S4). The model was constructed using the "lme" function in the "nlme" R package, with "climatic zones" and "vegetation types" as fixed factors and "sampling site" as a random factor. Pearson correlation coefficients were used to identify correlations between the response variable (MCL) and potential predictor variables (the 29 environmental factors) (Figure 5f). Generalized linear models examined relationships of MCL with LRR, logarithmically transformed ratios of available soil nutrients, and 10 key predictor variables (Figure 3; Figure 57); of logarithmically transformed ratios of available soil nutrients with plant-derived C production and soil organic C (SOC) concentration (Figure 58); and of SOC concentration with average daily and annual \(R_h\) (Figure S11). We also tested and visualized the interactions of each pair of variables...
among MAT, MAP, and MCL on soil $R_h$ in regression models using the "lm" and "visreg2d" functions in the "visreg" R package (Figure S10). All statistical analyses were performed using R software (v.3.3.2) (Development Core Team R, 2016).

3 | RESULTS

3.1 | Microbial C limitation in global terrestrial surface soils

We found that only 22% of 847 globally distributed sites (2476 observations) showed microbial C limitation (>0) (Figure 2a; Figure S1). A unimodal latitudinal pattern of MCLs was identified, with minimal C limitation near the equator (Figure 2a,b). Significant differences in MCLs among five climatic zones indicated that MCL was lowest in the tropics, with a mean of $-3.76$, SD = 1.80; Arid, mean = $-0.801$, SD = 1.76; Temperate, mean = $-1.09$, SD = 1.88; Cold, mean = $-0.870$, SD = 1.59; and Polar, mean = 0.334, SD = 2.06. The numbers in brackets below the violin charts represent the sample sizes (n). Different letters indicate significant differences ($p < .001$) among the climatic zones based on an analysis of linear mixed-effect models followed by Tukey’s tests. [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 2 Spatial patterns of soil microbial C limitation (MCL) at the global scale. (a) Spatial patterns of MCL across latitude fitted by a quadratic model. (b) The latitudinal breakpoint for MCL was estimated using piecewise regression analyses, and correlations of MCL with latitude before and after the latitudinal breakpoint were identified using generalized linear models. The shaded circle indicates breakpoint latitude, and the shaded area is the 97.5% confidence interval of the breakpoint. Solid black lines indicate model fits between MCL and latitude. (c) Differences in MCL among five climatic zones. Tropical, mean = $-3.76$, SD = 1.80; Arid, mean = $-0.801$, SD = 1.76; Temperate, mean = $-1.09$, SD = 1.88; Cold, mean = $-0.870$, SD = 1.59; and Polar, mean = 0.334, SD = 2.06. The numbers in brackets below the violin charts represent the sample sizes (n). Different letters indicate significant differences ($p < .001$) among the climatic zones based on an analysis of linear mixed-effect models followed by Tukey’s tests. [Colour figure can be viewed at wileyonlinelibrary.com]

Two independent approaches were used to verify the effectiveness of the ecoenzyme-based method of MCL. First, we identified the responses of MBC, PLFAs, and associated $R_h$ to the addition of plant litter across experimental sites that matched MCL locations (Figure 3a). Significant and positive correlations of MCLs with these responses indicated that the predicted MCL can be supported by the results of field manipulation experiments ($p < .05$). Second, microbial resource limitation theoretically responds non-linearly to the availability of soil resources as described by the Michaelis–Menten function (Sinsabaugh & Follstad Shah, 2012), so we tested whether the estimated MCLs were associated with relative availability of soil C across sites using available data. We found significant negative relationships of MCLs with $\ln(Lc:n)$ and $\ln(Lc:p)$ (ln-transformed DOC (dissolved organic C): (NO$_3^-$-N+NH$_4^+$-N) and DOC:Olsen-P, respectively) at the global scale ($p < .001$; Figure 3b,c), which further demonstrates the utility of MCL to identify microbial metabolic limitation imposed by the availability of soil C relative to N or P. Overall, the MCL metric was a generally useful metric for quantifying MCL.

3.2 | Environmental drivers of microbial C limitation

To identify the factors driving MCL, we used a random-forest approach to construct predictive models using 29 potential predictors suggested in the literature (Ma et al., 2021; Soong et al., 2020). Climate and vegetation properties were generally the most important drivers affecting MCLs in our predictive models ($p < .001$;
explaining 77% of the variation (cross-validated $R^2$). Figure 4a). The models included the following 10 most important predictors for describing MCLs: GPP, TSEA, Litter, MAT, MAP, Depth to Bedrock, Elevation, Slope, BGB, and AGB. Pearson correlation analysis and generalized linear models further supported our random-forest results, and also identified soil CaCO$_3$ concentration and pH as important predictors of MCLs ($p < .001$; Figures S6 and S7). However, SOC concentration that is generally considered as the measure of C availability was not significantly correlated to MCLs ($p > .05$; Figure S6). These analyses indicated that MCLs significantly increased with TSEA, Depth to Bedrock, Elevation, CaCO$_3$ concentration, and pH, but decreased with MAT, MAP, GPP, AGB, Litter, and BGB.

We then used a partial least-squares path model to investigate the direct and indirect associations between selected environmental factors and MCLs (Figure 4b,c). AGB, Litter, and BGB, regulated positively by latitude-dependent MAT and MAP, played negative central roles in MCL. This finding suggested that plant litter rather than SOM was the main source of C for the microbes in terrestrial surface soils. The significant and positive relationships of ln(Lc:n) and ln(Lc:p) with GPP, AGB, Litter, and BGB ($p < .001$; Figure S8), and having no or weak correlations with SOC concentration, provide further support for this argument. These results confirmed our hypothesis based on theoretical calculations that C availability is sufficient for microbial metabolism when plant litter was the dominant C source (Figure 1).

3.3 | The effect of microbial C limitation on $R_h$

By linking MCLs with soil $R_h$, we further found that MCL was significantly and negatively correlated with both average daily and annual $R_h$ ($p < .01$; Figure 5a,b). This result was supported by a partial correlation analysis showing controls exerted by MAT and MAP ($p < .01$; Figure 5c) because climatic conditions dominate soil $R_h$ on a large spatial scale (Bond-Lamberty et al., 2018; Hursh et al., 2017; Steidinger et al., 2019). Taking these three factors as the primary drivers of soil $R_h$, we found that their joint influences played more important roles in determining both soil average daily and annual $R_h$ than their independent influences (Figure 5d,e), consistent with interactions between any pair of variables among MAT, MAP, and MCL on $R_h$ (Figure S10). In contrast, SOC concentration had minor effects on average daily and annual soil $R_h$ (Figure S11), despite being traditionally considered as a critical control (Bond-Lamberty et al., 2018; Hursh et al., 2017). This finding indicated that SOM would not likely be the dominant C source for microbial heterotrophic metabolism in the surface soil, partly because SOM, includes more microbial-processed products with recalcitrant molecular structures that are protected by soil minerals and aggregates (Cotrufo et al., 2015, 2019; Jastrow, 1996). These results suggested that C availability relative to N and/or P for microbial metabolism instead of total soil C concentration was the key control determining soil $R_h$, and thus MCL could be a promising functional trait linking microbes to the release of soil C.
FIGURE 4  Key environmental factors affecting microbial C limitation (MCL). Random-forest models identify the relative influences of 29 environmental factors on MCL (a). Partial least squares path modeling determines the major pathways of the influences of climate, vegetation, topography, and soil properties on MCL (b). (a) The annular chart represents the percentages of cumulative relative contributions of four categories of environmental factors (climate, vegetation, topography, and soil). The climatic factors are mean annual temperature (MAT), mean annual precipitation (MAP), precipitation seasonality (PSEA), temperature seasonality (TSEA), and aridity index (AI). The vegetation factors are the Shannon diversity of enhanced vegetation index (Shannon EVI), the leaf area index (LAI), gross primary production (GPP), aboveground biomass (AGB), litter production (Litter), belowground biomass (BGB), and root depth. The topographical factors are elevation and slope. The soil factors are depth to bedrock, bulk density (BD), soil moisture (Moisture), cation-exchange capacity (CEC), soil pH (pH), particle composition (percentages of Sand, Silt, and Clay), CaCO$_3$ concentration, soil organic C concentration (SOC), total N concentration (TN), total P concentration (TP), SOC:TN ratio (Sc:n), SOC:TP ratio (Sc:p), and TN:TP ratio (Sn:p). * $p < .05$; ** $p < .01$. (b) The pathways of the influences of climate, vegetation, topography, and soil properties on MCL. Blue solid and orange dotted arrows indicate positive and negative flows of causality ($p < .05$), respectively. The numbers near the arrows indicate significant standardized path coefficients. $R^2$ indicates the variance of the dependent variable explained by the model. *** $p < .001$. Note that the gray dashed lines indicate potential relationships of Latitude with CaCO$_3$ concentration and Elevation, but with no theoretical supporting rationale that latitude could affect them, so we did not connect them in the figure. [Colour figure can be viewed at wileyonlinelibrary.com]
4 | DISCUSSION

4.1 | Pattern of microbial C limitation and its mechanisms

Our global estimates indicated that MCL did not dominate surface soils of natural terrestrial ecosystems at least from the ecoenzymatic perspective, contradicting the long-standing dogma that soil heterotrophs are predominately C-limited (Conant et al., 2011; Hobbie & Hobbie, 2013; Soong et al., 2020). The classic paradigm in terrestrial soils is that heterotrophic microbes are always constrained by the C source that they need but cannot produce on their own. However, we found that plant litter C was the dominant source of C for heterotrophic microbes in surface soils and could be sufficient for microbial metabolism relative to N and P supply (Figure S2). Because large quantities of polymeric plant detritus are produced annually both above- and belowground, and require catalysis by extracellular enzymes for microbial acquisition (Sinsabaugh & Follstad Shah, 2012), profiles of enzymatic activity integrating microbial demands across available resources reflect long-term temporal trajectories of microbial metabolic limitation in the detritus food web.

Two underlying mechanisms may determine the patterns of MCL we found in global surface soils. First, surface soils (Figure S2) are naturally C-rich due to the sustained input of plant litter with high C:N:P ratios, especially at low latitudes where primary productivity is high (Lehmann & Kleber, 2015; Xu et al., 2013). In particular, C inputs from C-rich root detritus and rhizodeposition may be further increased via root exudates stimulated by competition between plants and microbes for N and P in the rhizosphere (Kuzyakov & Xu, 2013). Moreover, plant-derived carbohydrate-rich inputs are easily decomposed by ecoenzymes, especially under favorable environmental conditions at low latitudes (Cotrufo et al., 2019; Fontaine et al., 2007; Jastrow, 1996).

Second, in contrast to low latitudes, soils at high latitudes with low primary production and relatively high MCL are dominated by SOM that contains transformed compounds with complex molecular structures that are difficult to decompose (Cotrufo et al., 2015; Fontaine et al., 2007; Liang et al., 2017). The physical protection of organic matter with the soil matrix also hinders microbial access to C from these substrates (Cotrufo et al., 2019; Lehmann & Kleber, 2015). As a result, microbes must allocate more resources to enzyme production for C acquisition in SOM-dominated food webs.
The energetic trade-offs between use of litter versus SOM lead to differences in respiration and growth that may promote dormancy or minimal maintenance metabolism in SOM-dominated communities (Fontaine et al., 2007). As a result, relatively high C limiting status indicated by ecoenzymatic patterns at high latitudes is associated with low metabolic activities, as indicated by \( R_h \).

### 4.2 Implications of microbial C limitation for soil C cycling

MCL is not ubiquitous in terrestrial surface soils (Figure 2), but as a physiological measure of soil C availability, it would essentially constrain \( R_h \) (Figure 5). The relatively high constraint we found at high latitudes, however, may be mitigated if primary productivity increases with climate change, which will lead to higher soil \( R_h \) and C release. For example, climatic warming, atmospheric N deposition, and CO\(_2\) fertilization may provide more plant-derived C for microbes by promoting plant growth (Huang et al., 2019; Peñuelas et al., 2017; Terrer et al., 2019; Zhang et al., 2022). Given a generally high N limitation at high latitudes (Du et al., 2020), these global change factors could also increase microbial C demand by reducing soil N limitation through direct N deposition or an indirect increase in biological N fixation and mineralization at high latitudes (Meunier et al., 2016; Peñuelas & Sardans, 2022; Terrer et al., 2021). As a result, soils at high latitudes, despite receiving more plant C input, could also lose more C via higher microbial metabolic activities under climate change.

In contrast, more-widespread terrestrial surface soils (about 78% of the geographic area) without MCL, especially at low latitudes (even the deep rhizosphere), may have the potential to sequester more surplus plant-derived C with an increase in primary productivity under climate change. This argument is supported by a recent study indicating that N deposition can increase soil C sequestration in tropical forests (Lu et al., 2021). This potential C sequestration benefit may be further strengthened by the inherent P limitation at low latitudes (Du et al., 2020; Vitousek et al., 2010), as P limitation can suppress microbial growth and metabolism (Cui et al., 2022). From the perspective of microbial resource limitation, climate change may thereby decouple the potential benefits of MCL from the soil C sink at high latitudes but increase the soil C sink at low latitudes.

Our findings also highlight that MCL may be one of the important reasons for soil C accumulation at high latitudes, which generally have larger stocks of SOC than ecosystems at low latitudes. The relationships between MCL and \( R_h \) suggested that high C limitation at high latitudes means low \( R_h \) (Figure 5). Also, microbial communities generally have a relatively high C use efficiency at high latitudes, because they maintain low metabolic rates at low environmental temperatures (Allison et al., 2010). Higher C use efficiency generally means more production of microbial biomass and other microbial products when the total uptake of C remains constant, potentially beneficial for the accumulation of microbially derived C in soils and contributing to soil C storage (Cotrufo et al., 2015; Wang et al., 2021). However, the high stocks of SOC at high latitudes may be increasingly unstable under global warming due to high-temperature sensitivity of microbial growth coupled to increased availability of plant-derived C.

### 4.3 Uncertainties

Previous studies, ranging from theoretical predictions to experiments, provide important but limited insights into patterns and mechanisms of C limitation to soil heterotrophic microorganisms (Cheng et al., 1996; Cui et al., 2021; Demoling et al., 2007; Ekblad & Nordgren, 2002; Soong et al., 2020). Our global predictions, combining an ecoenzyme-based method with geographically extensive databases, provide mechanistic understandings of global patterns of MCL in surface soils. However, there are several uncertainties in this study. First, our use of ecoenzymatic activities as indicators of microbial nutrient acquisition cannot unequivocally define the status of C demand in all microbial communities in all systems. For example, ecoenzymes are rarely involved in the microbial acquisition of C when simple plant-derived compounds such as rhizospheric carboxylic acid are directly used by microbes (Ding et al., 2021), because they do not require enzyme catalysis to form assimilable products (Schimel et al., 2022). Similarly, inputs of inorganic nutrients, such as atmospheric N deposition, may also skew patterns of relative C-, N-, and P-acquiring enzyme activities (Moorhead et al., 2023). Nonetheless, turnover of resulting microbial biomass is driven by catalytic enzymes so that impacts of these effects are constrained.

A second area of uncertainty in the present study is that we only investigated surface soils. We suspect that MCL increases along depth profiles due to sharply reduced C input of plant litter and generally lower C:N and C:P ratios in deeper soils (Don et al., 2013). However, C limitation in hotspots of deep soils such as in the rhizospheres of deep roots may be relieved by exudates or debris (Cheng et al., 1996; Kuzyakov & Xu, 2013). In addition, different community composition and functional taxa such as r-versus K-strategist-dominated communities may generate distinct resource limiting status. Deep soils are functionally different from surface soils because the limited supply of C drives microbial communities to conserve C, including a higher relative abundance of autotrophic microorganisms that increases re-fixation of respired CO\(_2\) (Šantrůčková et al., 2018) and a lower relative abundance of heterotrophic microorganisms, specifically saprotrophic fungi (Gittel et al., 2014). Third, we have not accounted for the widespread symbiotic relationships between microbes and plants, whereby microbes can directly acquire C from plants in exchange for nutrients (N and P) without the need to use ecoenzymes (Martin et al., 2016; Zak et al., 2019). Of course, the turnover of these microbial symbionts in turn requires catalytic activities of extracellular enzymes. In the future, studies would thus benefit from analyses of function-specific gene expressions in microbial communities and isotopic labeling of...
polymeric substrates, to provide greater mechanistic insights into ecoenzyme-based methods.

Finally, the interpretation and extrapolation of microbial-mediated biochemical reactions across global scales integrate multiple levels of resolution, which can obscure differences in controls at finer grains (Bradford et al., 2021). For example, in this study we found different patterns of enzyme activities in polar versus tropical systems, likely due to different strengths of underlying controls (above). Indeed, Prescott (2010) described different thresholds for dominant influences of key factors on decomposition, including a pervasive temperature limitation emerging at a MAT below 10°C, whereas wet tropical systems are often most limited by P availability (Cui et al., 2021). Thus, global patterns of ecoenzymatic responses to global environmental controls may not reflect the patterns of dominant controls operating within specific ecosystems or shifts in the dominant controls across environmental thresholds between ecosystems. These may be examples of Simpson’s Paradox inherent to characterizing microbial-mediated global processes (Bradford et al., 2021).

Despite these uncertainties, our findings of global patterns in MCL have a robust base in both metabolic theories and stoichiometry methods, are mechanistically interconnected with key environmental drivers, and are strongly supported by results from both field experiments and labile nutrient ratios (Figures 3 and 4). These findings provide a novel view of how soil C pools may respond to environmental change, given varying patterns of plant-litter utilization by heterotrophic microorganisms along environmental gradients. The microbially mediated release of soil C is one of the largest sources of uncertainty in predicting terrestrial C dynamics, so our identification of a strong link between a predictable metric of MCL and the highly variable patterns of $R_{h}$ observed across global scales represents a potentially valuable step toward better quantifying soil C release in a way that allows integration into current models of terrestrial C cycling.

AUTHOR CONTRIBUTIONS
Yongxing Cui conceived and developed the study, Yongxing Cui compiled the database. Yongxing Cui and Shushi Peng performed analysis of the data. Yongxing Cui, Shushi Peng, Manuel Delgado-Baquerizo, Matthias C. Rillig, César Terrer, Daryl L. Moorhead, Robert L. Sinsabaugh, and Josep Peñuelas contributed to the interpretation of the results. Yongxing Cui, Shushi Peng, Manuel Delgado-Baquerizo, Matthias C. Rillig, César Terrer, Biao Zhu, Xin Jing, Ji Chen, Jinquan Li, Jiao Feng, Yue He, Linchuan Fang, Daryl L. Moorhead, Robert L. Sinsabaugh, and Josep Peñuelas wrote and revised the manuscript.

ACKNOWLEDGMENTS
This study was financially supported by the National Natural Science Foundation of China (32101378) and Project funded by the China Postdoctoral Science Foundation (2022M710004). We are grateful to B. Zhu, X. Jing, H. Li, W. Zheng, H. M. Siljaren, D. A. Cowan, H. Kang, J. Feng, X. L. Cheng, C. Wang, J. Q. Li, and X. K. Lu for sharing their data. We thank Y. He, M. Y. Feng, and M. J. Wang for their assistance during data acquisition and analysis.

CONFLICT OF INTEREST STATEMENT
The authors declare no competing financial interests.

DATA AVAILABILITY STATEMENT
The database that supports the findings of this study is available at Figshare.com (https://doi.org/10.6084/m9.figshare.22801163).

ORCID
Yongxing Cui https://orcid.org/0000-0002-8624-2785
Shushi Peng https://orcid.org/0000-0001-5098-726X
Manuel Delgado-Baquerizo https://orcid.org/0000-0002-6499-576X
Matthias C. Rillig https://orcid.org/0000-0003-3541-7853
César Terrer https://orcid.org/0000-0002-5479-3486
Biao Zhu https://orcid.org/0000-0001-9858-7943
Xin Jing https://orcid.org/0000-0002-7146-7180
Ji Chen https://orcid.org/0000-0001-7026-6312
Jinquan Li https://orcid.org/0000-0002-2219-2571
Jiao Feng https://orcid.org/0000-0003-2879-3498
Yue He https://orcid.org/0000-0002-7754-7026
Linchuan Fang https://orcid.org/0000-0003-1923-7908
Daryl L. Moorhead https://orcid.org/0000-0002-3460-5302
Josep Peñuelas https://orcid.org/0000-0002-7215-0150

REFERENCES


Guerrero-Ramírez, N. R., Pizarro, V., & Turner, B. L. (2020). Soil and microbial nutrient status are heterogeneous within an elevational belt on a neotropical mountain. Pedobiologia, 82, 150689.


Li, J., Nie, M., & Pendall, E. (2020). Soil physico-chemical properties are more important than microbial diversity and enzyme activity in controlling carbon and nitrogen stocks near Sydney, Australia. Geoderma, 366, 114201.


Waring, B. G. (2013). Exploring relationships between enzyme activities and leaf litter decomposition in a wet tropical forest. Soil Biology and Biochemistry, 64, 89–95.


**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.