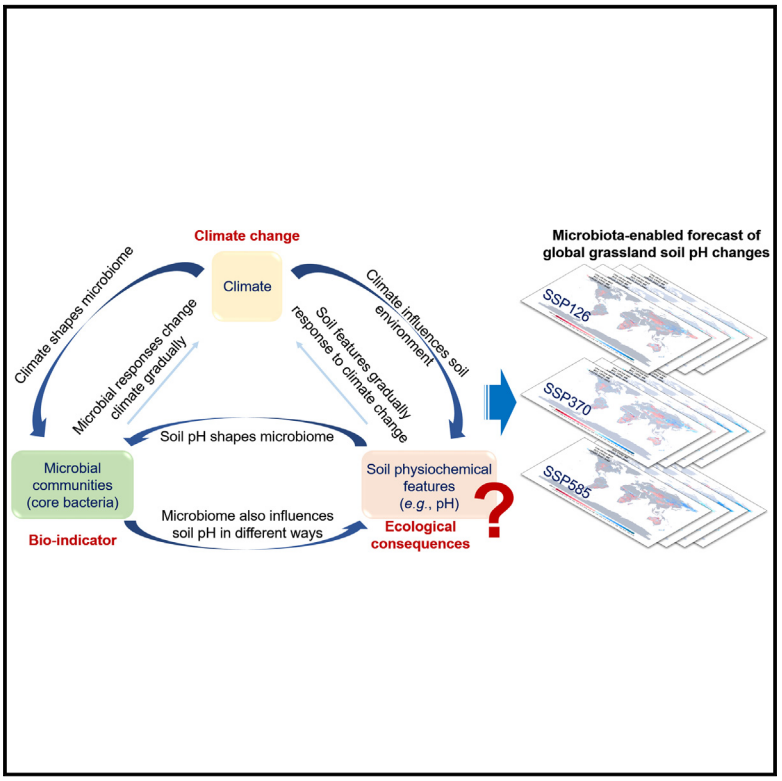
One Earth

Article

CoBacFM: Core bacteria forecast model for global grassland pH dynamics under future climate warming scenarios

Graphical abstract



Highlights

- A microbe-central model to predict future soil pH changes through bioindicators
- 32.5%–32.9% of grassland regions will become more alkaline by 2100
- The alkaline grassland regions expand in future scenarios
- Field experiments performed similar alkalization tendency with model prediction

Authors

Kai Feng, Shang Wang, Qing He, ..., Yunfeng Yang, Jizhong Zhou, Ye Deng

Correspondence

yedeng@rcees.ac.cn

In brief

Soil pH affects a myriad of soil physiochemical properties that further influence ecosystem productivity and diversity. However, in this century, due to the rapidly changing climate, the biogeochemical cycles driven by soil microorganisms could change, with severe effects on soil pH and ecosystem functions. We leveraged a global grassland soil microbiota dataset to forecast the dynamic changes in soil pH under climate change. Our study provides new insight into forecasting soil pH by incorporating soil microbiota into Earth models.

Feng et al., 2024, One Earth 7, 1275–1287
July 19, 2024
https://doi.org/10.1016/j.oneear.2024.06.002

CellPress

One Earth

CellPress
OPEN ACCESS

Article

CoBacFM: Core bacteria forecast model for global grassland pH dynamics under future climate warming scenarios

Kai Feng,[1] Shang Wang,[1] Qing He,[1,2] Michael Bonkowski,[3] Mohammad Bahram,[4,5,6] Etienne Yergeau,[7] Zhujun Wang,[1,8] Xi Peng,[1,2] Danrui Wang,[1,2] Shuzhen Li,[1,9] Yingcheng Wang,[1,10] Zhicheng Ju,[1,2,11] Xiongfeng Du,[1,2] Chengliang Yan,[1,2]

(Author list continued on next page)

[1]CAS Key Laboratory of Environmental Biotechnology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China
[2]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China
[3]Terrestrial Ecology, Institute of Zoology, Cluster of Excellence on Plant Sciences (CEPLAS), University of Cologne, Cologne, Germany
[4]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden
[5]Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia
[6]Department of Agroecology, Aarhus University, Slagelse, Denmark
[7]Institut national de la recherche scientifique, Centre Armand-Frappier Santé Biotechnologie, Laval, Quebec, Canada

(Affiliations continued on next page)

SCIENCE FOR SOCIETY Soil pH is a critical factor in sustaining species growth and survival on Earth, yet with ongoing climate warming, it is difficult to predict the soil pH changes due to the complexity of soil systems. Moreover, the soil pH generally maintains a relatively stable state in response to external disturbances due to its buffering capacity (e.g., excess ions can be adsorbed by minerals or organic compounds). Soil microbiota are highly correlated with soil pH and quickly respond to climate warming, bridging climates and soil systems; thus, they are good bioindicators of soil pH changes. According to the developed microbe-central model, the grassland soils will become more alkaline in future climate change scenarios, as evidenced by 14 field warming simulation experiments and featured in northeastern Asia, Africa, and Oceania. The hotspot regions and predicted results by this microbial model could be considered in future decision-making strategies and policies in facing climate change.

SUMMARY

Soil microbes regulate various biogeochemical cycles on Earth and respond rapidly to climate change, which is accompanied by changes in soil pH. However, the long-term patterns of these changes under future climate scenarios remain unclear. We propose a core-bacteria-forecast model (CoBacFM) to model soil pH changes by shifts of core bacterial groups under future scenarios using a curated soil microbiota dataset of global grasslands. Our model estimates that soil pH will increase in 63.8%–67.0% of grassland regions and decrease in 10.1%–12.4% of regions. Approximately 32.5%–32.9% of regions will become more alkaline by 5.6%, and these areas expand in all future scenarios. These results were supported by 14 warming simulation experiments. Using bacterial responses as bioindicators of soil pH, the CoBacFM method can accurately forecast pH changes in future scenarios, and the changing global climate is likely to result in the alkalization of grasslands.

INTRODUCTION

Soil, as the foundation of the worldwide terrestrial ecosystem, provides life-sustaining resources and harbors immense biodiversity on Earth.[1] Ongoing climate changes, characterized by warmer conditions, accelerate organic matter decline and biodiversity loss. These shifts threaten global soil functions and processes.[2] However, predicting how soil processes will change

CellPress
OPEN ACCESS

One Earth
Article

Songsong Gu,[1,2] Tong Li,[1,2] Xingsheng Yang,[1,2] Wenli Shen,[1] Ziyan Wei,[1,12] Qiulong Hu,[13] Pengfei Li,[14] Yanmei Zhu,[14] Guangxin Lu,[10] Clara Qin,[15] Gengxin Zhang,[16] Chunwang Xiao,[17] Yunfeng Yang,[18] Jizhong Zhou,[19,20] and Ye Deng[1,2,21,*]

[8]College of Tropical Crops, Hainan University, Haikou, China
[9]Aquatic EcoHealth Group, Key Laboratory of Urban Environment and Health, Fujian Key Laboratory of Watershed Ecology, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China
[10]College of Agriculture and Animal Husbandry, Qinghai University, Xining, China
[11]Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong SAR, China
[12]State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China
[13]College of Horticulture, Hunan Agricultural University, Changsha, China
[14]Wenshan Tobacco Company of Yunnan Province, Wenshan, Yunnan, China
[15]Department of Environmental Studies, University of California, Santa Cruz, Santa Cruz, CA, USA
[16]Key Laboratory of Alpine Ecology and Biodiversity, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China
[17]College of Life and Environmental Sciences, Minzu University of China, Beijing, China
[18]Institute of Environment and Ecology, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, China
[19]Institute for Environmental Genomics, Department of Microbiology and Plant Biology, School of Civil Engineering and Environmental Sciences, and School of Computer Science, University of Oklahoma, Norman, OK, USA
[20]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[21]Lead contact
*Correspondence: yedeng@rcees.ac.cn
https://doi.org/10.1016/j.oneear.2024.06.002

under future climate scenarios remains challenging due to the inherent complexity of soils. Climate-driven changes in matter, energy, and biodiversity interact in complex ways with soil functions, thus hindering efforts to accurately predict changes in soil functioning under future climate changes.

Soil pH, as one of the significant variables for soil chemistry, regulates biochemical processes and the capacity of available nutrients to organisms in terrestrial ecosystems[1] and is a good predictor of the extent to which climate change will affect the Earth. A specific range of soil pH conditions is one of the critical environmental factors in the colonization and growth of plants, animals, and microbes in ecosystems, influencing biodiversity and ecosystem functions, such as carbon storage and emission.[3] Generally, soil buffering effects maintain a relatively stable state of soil pH, with aluminum compounds at low pH and carbonates at high pH immediately consuming excess ions.[1] However, the equilibrium of acid and nonacid cations in soils may be disrupted by various climatic conditions, for example, alkaline soils in arid regions and acid soils in wet areas.[1] Microorganisms involved in biogeochemical processes may change soil pH by generating protons and hydroxyl ions.[4] Therefore, soil pH could be altered by external disturbances. For instance, changing precipitation regimes and the decomposition of soil organic matter via microbial activities in response to climate change alter the profile of organic acid compounds in soils,[5,6] thus altering soil pH. Organic acid metabolisms such as oxalate catabolism by bacteria could increase the soil pH by up to 2.5 units.[7] The microbial metabolic processes in nitrogen cycles, such as ammonification releasing hydroxyl ions and denitrification consuming protons, could also increase soil pH.[4] Subsequently, the soil pH changes, in turn, drive the succession and assembly of microbial community structure.[8,9] Therefore, soil pH and the microbial community are intertwined in response to climate change. Following changes within a microbial community, the soil pH is expected to change accordingly, but the magnitude and direction of future soil pH changes remain unclear.

Soil microbiota are sensitive to changes in annual temperature and annual precipitation,[10–13] exhibiting essentially non-linear correlations and time-dependent dynamics.[14] Owing to those complexities, terrestrial models for predicting soil pH changes usually neglect microbial communities.[15] Soil microorganisms are the most diverse lives in the soil biosphere,[16] demonstrating keystone roles in sustaining ecosystem functions and potential mitigation of global climate change.[17,18] Numerous studies demonstrate that the soil microbiota drive the biogeochemical processes through metabolic activities,[18,19] regulate soil carbon storage[20] and greenhouse gases,[21] and facilitate soil nutrition transformations for plant productivity.[22] The changes in soil microbiota intertwine with ecosystem functions. Under the current global climate change background, the biodiversity, species turnover rate, and community stability of soil microbiota are changed accordingly,[10,23,24] suggesting the responses of microbial communities to climate change should be considered in climate adaptation planning. However, we still lack an integrated framework to depict the responses of soil microbiota to climate change and its downstream effects, such as soil pH changes. Due to the high diversity of microbial species in terrestrial ecosystems (e.g., decades of the order of magnitude of species richness),[25] it is challenging to directly include microbial species as individual variables in prediction models. Instead, microbial species can be grouped into ecological clusters (eco-clusters), which refer to specific microbial groups that potentially share similar functions. The eco-cluster framework offers new insights into linking terrestrial ecosystems and climate change scenarios.[26] Specifically, eco-clusters reduce the complexity of microbial communities but retain information on microbial species' responses to environmental perturbations, making them both indicators and predictors of environmental change.[26,27]

Here, we propose a core-bacteria-forecast model (CoBacFM) (Figure 1) to investigate the biogeography of core soil microbiota under current and future climate conditions in global grasslands. Grasslands, covering over 40% of the global land area, represent the largest terrestrial ecosystem.[28,29] To date, it remains challenging to find common response patterns by grasslands to climate change (e.g., acidification, alkalization) in local- or regional-scale field experiments.[29] In this study, three shared socioeconomic pathways, namely, SSP126 (sustainable way, 2°C to end of 2100), SSP370 (moderate scenario), and SSP585

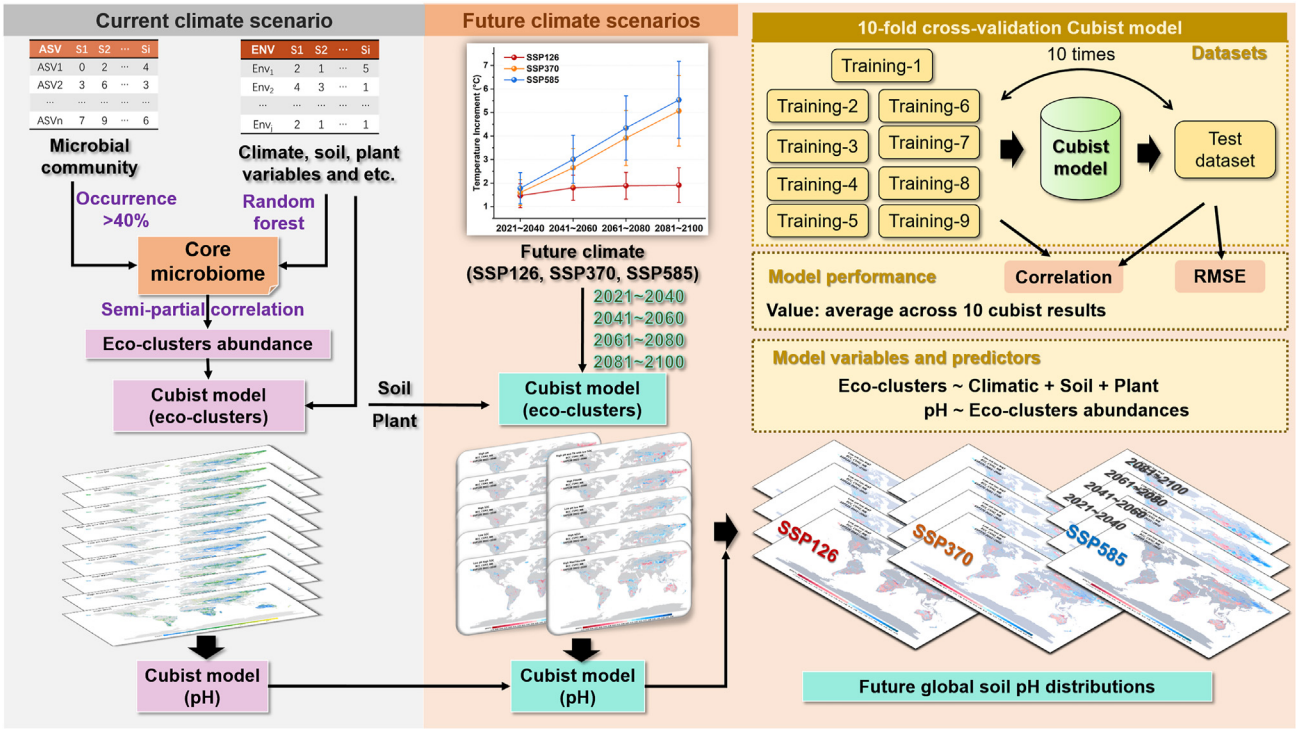


Figure 1. Schematic approaches for CoBacFM leveraging the bacteria dynamics to project the soil pH change under future climate scenarios

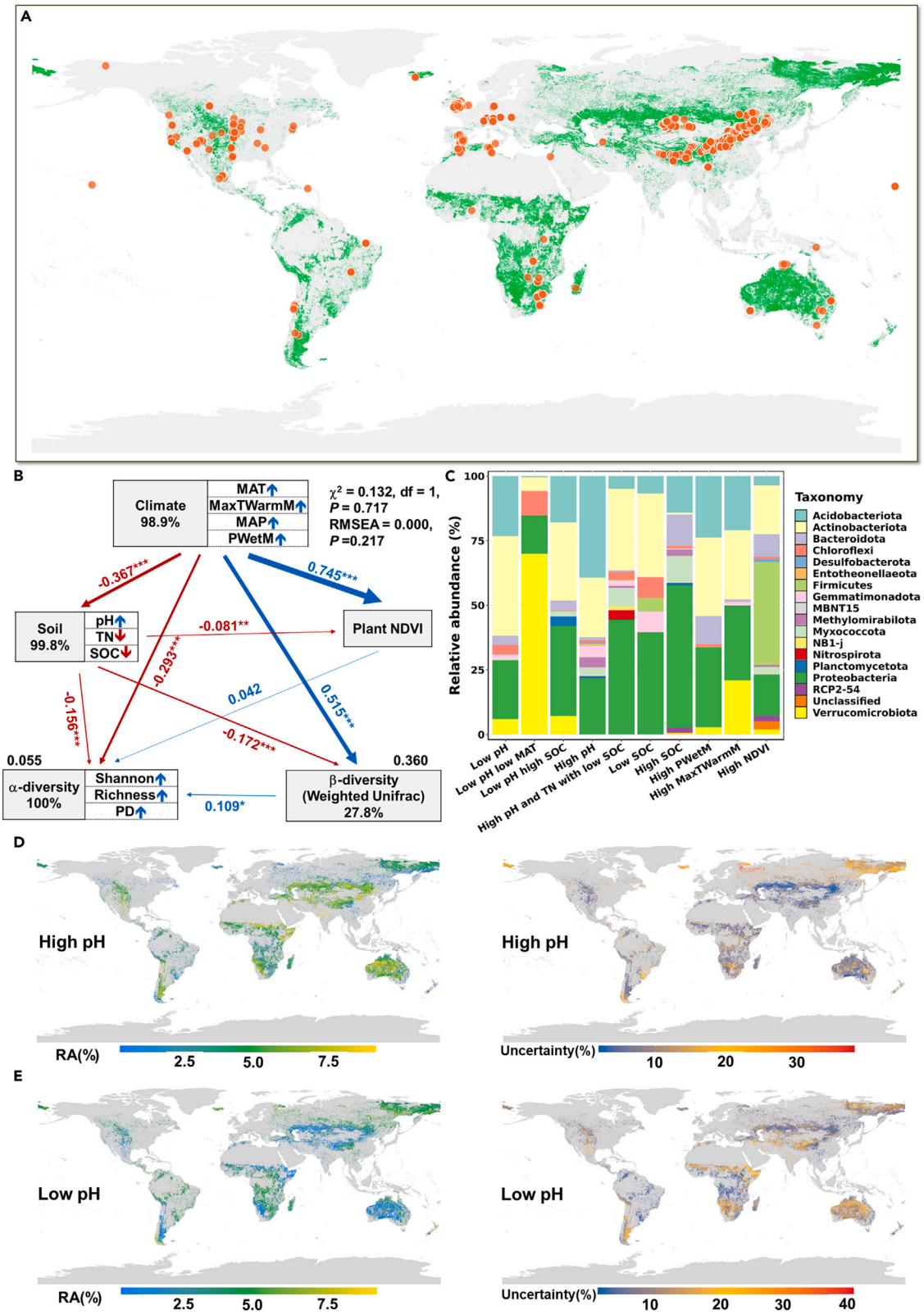
By filtering the global microbial dataset and using random forest to explain species distribution based on multiple environmental variables, the core microbiota (e.g., bacteria) and, subsequently, eco-clusters were assigned. Eco-clusters are a group of core microbial taxa with a similar response to specific environmental dynamics (e.g., the eco-cluster of low-pH may prefer acidic conditions). Then, the relative abundance of each eco-cluster was predicted by multiple environmental variables with the Cubist model, including climate (current scenario to build model and future scenario for predictions), soil, and plants. Considering microbial activities in biogeochemical processes, the shifts of eco-clusters can be used by Cubist models to predict soil pH changes in current and future scenarios. After comparing the current and future soil pH shifts, the model can generate the global soil pH shifting maps under multiple climate scenarios at different periods. For each Cubist model, a 10-fold cross-validation approach was used to evaluate the model performance with correlation coefficient and RMSE value. Details are described in Experimental procedures.

(upper boundary) for 4 periods (2021–2040, 2041–2060, 2061–2080, and 2081–2100) were chosen to model future climate conditions. This core bacteria-driven model links changes in eco-clusters to changes in climate conditions and soil functioning features such as soil pH. To this end, we analyzed high-throughput sequencing datasets of the 16S rRNA gene from 3,703 samples across 1,251 grassland sites located around the world (Figure 2A; Table S1), including natural grasslands from tropical, temperate, and tundra regions.[28] Our main objectives were to depict the future biogeographical distributions of grassland core microbiota and to forecast dynamic changes in soil pH under future climate change via changes in bacterial eco-clusters. Specifically, this CoBacFM model estimated that soil pH will increase in 63.8%–67.0% of grassland regions, especially in northeastern Asia, Africa, and Oceania grasslands. Moreover, 10.1%–12.4% of grassland regions will experience a decrease in soil pH, including in central North America, southern Africa, and eastern Asia. Approximately one-third of regions will become more alkaline by 5.6%, and the area of alkaline soils expand in all future scenarios. More important, the 14 independent warming simulation field experiments support the model forecasted results (i.e., the soil pH of 4 sites increased significantly under warming conditions and 9 sites became more alkaline as predicted by the model). Using bacterial responses as bioindicators of soil pH, our model forecasted the potential variabilities of soil pH under future climate changes, and identified hotspot regions could be considered in future decision-making strategies and policies for mitigating future climate warming.

RESULTS AND DISCUSSION

Identification of grassland core microbiota and classification of eco-clusters

The curated soil microbiome dataset covered multiple types of grasslands distributed across all continents except for Antarctica (Figure 2A), and only 989 phylotypes (amplicon sequence variants, occupancy of 40% within each continent, and a total relative abundance of 29.4%) were selected as the core microbiota (globally prevalent and abundant species) for the global grassland ecosystem (Figure S1; Data S1). Among the environmental variables significantly correlated to biodiversity (Table S2), climatic factors such as mean annual temperature (MAT), the maximum temperature of the warmest month (MaxTWarmM), mean annual precipitation (MAP), and precipitation of the wettest month (PWetM), as well as some soil physicochemical parameters, such as soil pH, soil organic carbon (SOC), and total nitrogen (TN) content, were directly associated with the α-diversity and β-diversity of the core microbiota



(legend on next page)

(Figures 2B and S2). After evaluating the coverage of each environmental variable to the remaining regions of the global grasslands, these environmental variables could individually cover 87.9%–100% of grasslands and together cover more than 94.7% of the grasslands if only a single variable was missing (Figure S3). Subsequently, a total of 625 phylotypes possessing significant semi-partial correlations with environmental variables ($p < 0.001$) were categorized into 10 eco-clusters (Figure 2C)—(1) high-pH, (2) low-pH, (3) high-SOC, (4) low-SOC, (5) high-pH and TN with low-SOC, (6) low-pH-high-SOC, (7) low-pH-low-MAT, (8) high-PWetM, (9) high-MaxTWarmM, and (10) high-NDVI (normalized difference vegetation index)—according to microbial preference to specific environmental factors (Figure S4). For example, taxa of the high-pH eco-cluster showed higher relative abundance under alkaline soil conditions (pH >7) and those of the low-pH eco-cluster preferred acidic soil conditions (pH <7) (Figure S5). In addition, the taxonomic groups differed between eco-clusters; for example, Actinobacteria sp. was more abundant in the cluster of low-SOC and less abundant at high-SOC (Figures 2C and S6). The high relative abundance of Actinobacteria in the low-SOC cluster might relate to their ability to decompose recalcitrant carbon under nutrient-poor conditions,[30] while copiotrophic-associated Alphaproteobacteria and Gammaproteobacteria showed higher relative abundances in nutrient-rich conditions (Figure S6).

Because the relative abundances of the eco-clusters were significantly associated with specific environmental conditions, we established the biogeographic distributions of the eco-clusters at a global scale, using the Cubist model with 10-fold cross-validation (Figure 1). Our analysis resulted in 0.786–0.928 correlations for training datasets and 0.810–0.922 correlations in test datasets ($p < 0.001$; Table S3), indicating the comparable and reliable performance of this Cubist forecast model. The global biogeographic distributions of the 10 eco-clusters were interpolated, and the maps clearly showed a dominance of specific clusters in certain regions (i.e., hotspot regions) (Figures 2D, 2E, and S7). In particular, the eco-clusters of high-pH (alkaline skewed) and low-pH (acid skewed) occupied contrasting hotspot regions (e.g. the relative abundance of the high-pH cluster was higher in central Asia [Figure 2D], while the low-pH cluster was lower [Figure 2E]). This region-specific distribution of eco-clusters may be related to microbial adaptations to the local environmental features, suggesting environmental selection of eco-clusters. Similar to animals and plants, the abundant core microorganisms are reported to be essential drivers of the energy flow and biogeochemical processes in ecosystems.[18,31] Therefore, the shifts among the core bacteria could, to some extent, serve as bioindicators for macroscopic ecosystem function changes under future climate scenarios.

Core microbiota projections under future climate change scenarios

By focusing on the predominant eco-clusters and their environmental drivers, we next examined the temporal responses of eco-clusters to future climate conditions (Figure 1). The PiecewiseSEM models showed that climatic variables, mainly MAT, MaxTWarmM, MAP, and PWetM, made large contributions to explaining the relative abundances of the 10 eco-clusters (Figures S8 and S9). The relative abundances of the 10 eco-clusters under future climate scenarios were determined by changing the 4 climatic variables in the Cubist prediction models. Here, we mainly considered 3 climatic scenarios, SSP126, SSP370, and SSP585, for 4 time periods (2021–2040, 2041–2060, 2061–2080, and 2081–2100) (Figure S10; Video S1). The predicted spatiotemporal distributions of the eco-clusters showed great differences worldwide, especially with different scenarios. For example, the cluster low-pH-low-MAT, which was mainly dominated by Verrucomicrobiota (Figure 2C), showed a tendency to decrease in the low latitudes and temperate regions under all climate change scenarios, with large variations in the amount of change in central and northeastern Asia, central North America, and southern Africa, indicating a change of soil pH in these regions (Figure S11). Therefore, based on the future distributions of eco-clusters, it is possible to assign the hotspot regions for global grasslands, where the microbial activities and soil properties are sensitive to future climate change.

To validate the model predictions, we curated grassland soil microbiota datasets from 14 globally distributed field experiments that simulated climate change, with a focus on simulated warming vs. ambient conditions (Table S4). The experiment sites were distributed in Asia (11 sites), Europe (1 site), and North America (2 sites), the majority simulating elevated temperatures of 1.5°C –2.0°C. We used soil microbiota results from these 14 sites to validate the observed dynamics of eco-clusters under future climate warming conditions. Taking the eco-cluster low-pH-low-MAT, which showed a direct preference for temperature, as an example, most sites in Iceland and on the Qinghai-Tibetan Plateau in Asia showed an abundance increase under warming compared to ambient conditions (the response ratio with a 90% confidence interval) (Figure S12). This experimental result was in line with the model predictions, which showed that the relative abundance of low-pH-low-MAT would increase in approximately 78.6% of the sites until 2100 (Figure S13). Therefore, the temporal dynamics of eco-clusters could be projected from their biogeographical distribution in global grasslands and the variance of climate-relevant

Figure 2. Global distribution of grassland microbiome dataset and relationships between core microbiota and environmental variables, and the relative abundance of eco-clusters (e.g., low-pH and high-pH) in global grasslands

(A) The distribution of unique samples from the curated grassland microbiome dataset.

(B) SEM shows the relationship between the core microbiota community and environmental variables. The first principal component axes of climate, soil, and α-diversity were included with explained proportions in the box, and blue and red arrows within the box indicate the positive and negative correlations, respectively, for the first principal component and components. The first principal coordinate axis of weighted Unifrac distance was used to indicate β-diversity. PD, Faith's phylogenetic diversity.

(C) Relative abundance of species within the 10 eco-clusters at the phylum level.

(D and E) The relative abundance for high-pH (D) and low-pH (E) in global grasslands, respectively (left) and their uncertainty values (right) were calculated by the ratio of SD to mean values with 10-fold cross-validation results.

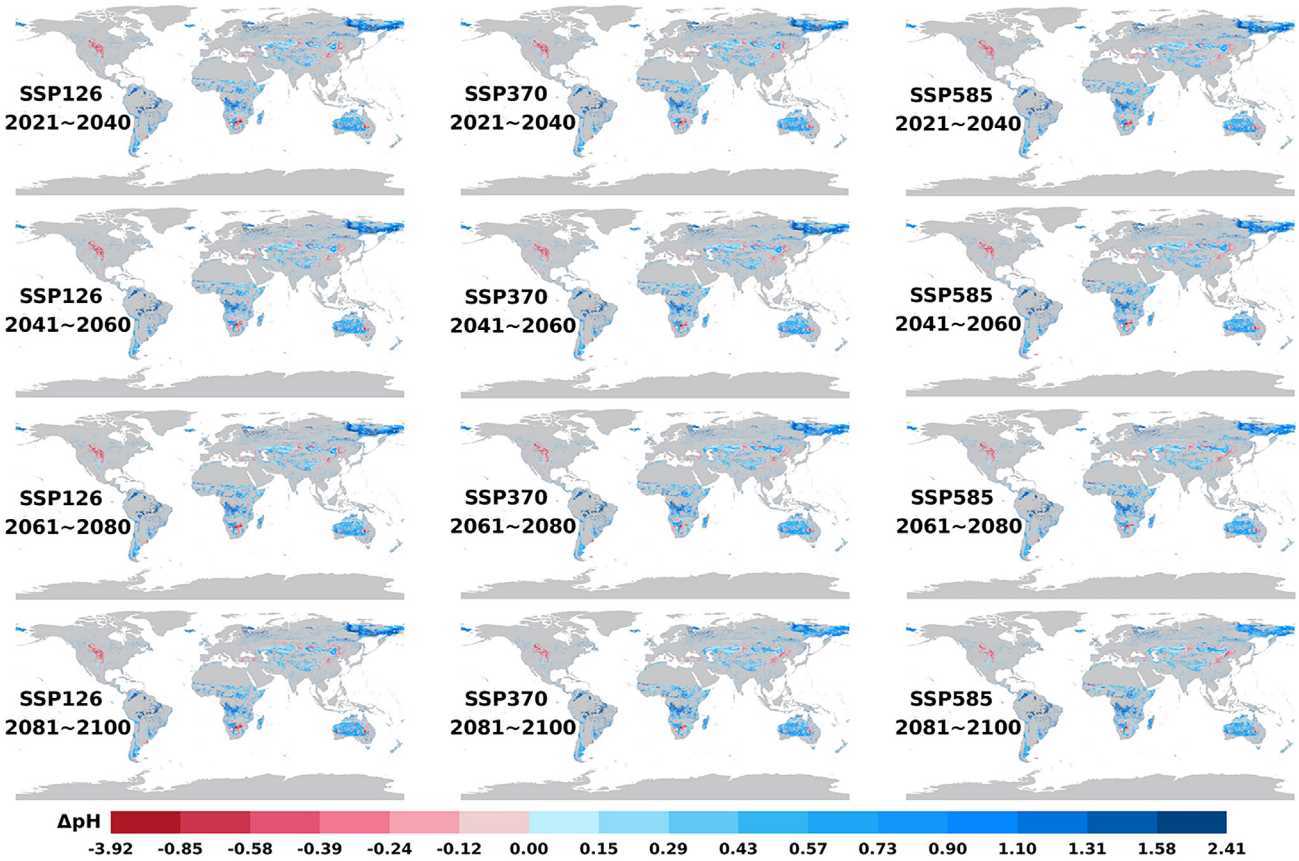


Figure 3. The differences of projected edaphic pH to current pH under future climate scenarios (SSP126, SSP370, and SSP585) over 4 time periods (2021–2040, 2041–2060, 2061–2080, 2081–2100)

The blue and red values indicate increased and decreased pH, respectively.

clusters could be demonstrated by simulated warming experiments in natural grasslands.

Soil pH change under future climate change

Many studies have shown that microbial activities can alter the surrounding pH conditions.[32,33] In line with these studies, our results showed that the relative abundance of eco-clusters was a better predictor of soil pH compared to SOC (Figures S8 and S14). In addition, the Cubist model for current grassland soil pH (Figure 1) had high confidences of 0.940 ± 0.000 and 0.928 ± 0.016 for both the training and test datasets, respectively (Table S5). Biogeographic dynamics of edaphic pH in global grasslands showed a large variability in response to future climate change under three scenarios (SSP126, SSP370, and SSP585). The soil pH in central North America, southern Africa, and eastern Asia decreased, whereas that in northeastern Asia, Africa, and Oceania increased (Figure 3). Moreover, at the grid-point level (0.1° resolution), 10.1%–12.4% of the regions with a significant decreasing pH trend shrank with time (4.3%–4.9%, $p < 0.001$) (Figures 4A and S15). A total of 5.5%, 5.4%, and 5.5% of total grassland regions trended to being more acidic, with decreasing magnitudes of 5.7%, 5.6%, and 5.6% under SSP126, SSP370, and SSP585 scenarios, respectively (Figures 4B and S15). Nevertheless, the soil pH of 63.8%–67.0% of grassland regions showed a significantly enhanced tendency to increase (10.8%–11.0%, $p < 0.001$) under future temporal dynamics (Figures 4C and S15). Importantly, approximately 32.5%, 32.6%, and 32.9% of grassland soils would become more alkaline under the three scenarios by 5.6%, 5.6%, and 5.6%, respectively (Figures 4D and S15). These results were verified by another climate model, IPSL-CM6A-LR, which showed that about one-third of grassland regions became more alkaline, and the areas expanded with time (Figure S16). Additionally, there was an inflection point under SSP126 indicating the resilience of soil pH by the end of 2100, which may support sustainable development by mitigation policies to climate change.

Soil pH changes from the 14 warming experiments showed different responses under warming conditions (Figure 5; Table S4). Four experimental sites from Asia (Haibei1, Tiebujia, Hulun, and Siziwang) showed significant edaphic pH increases under warming conditions, while most other sites showed high variance without significant change (Figure 5A). This was consistent with the predictions from the modeling, that more than 30% of areas will become more alkaline under future climate change scenarios (Figure 4). Moreover, the predictions showed that pH for most of the field sites tended to increase (11 of 14 sites) and become more alkaline (9 of 14 sites) under future climate change scenarios, while only the two sites from North America

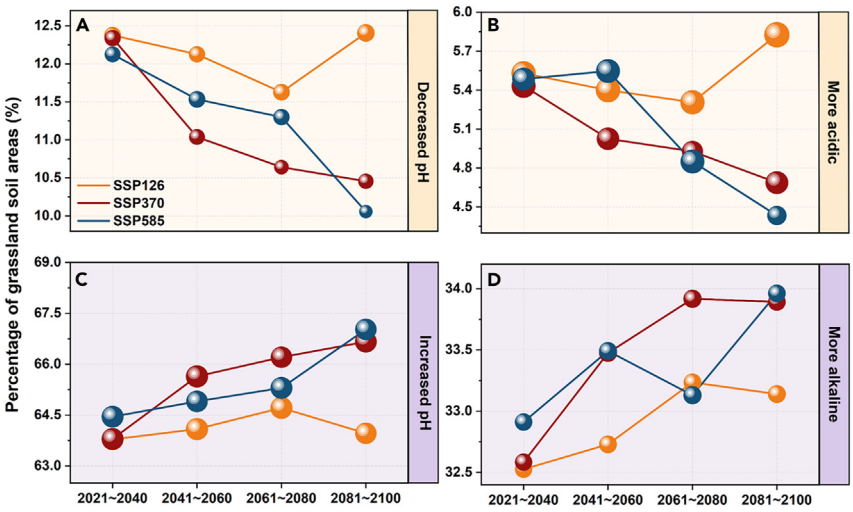


Figure 4. Proportion of grassland areas related to pH change in the future under 3 scenarios (SSP126, SSP370, and SSP585) compared to the current scenario

(A–D) The soil pH values in the grassland regions were significantly changed, including decreasing (A), acidic tendency (B), increasing (C), and alkaline tendency (D).

and one site from Asia showed acidic conditions with comparable stability in pH and little variability (Figure 5B). By contrast, the field regions from the Qinghai-Tibetan Plateau in Asia showed high variability, especially under higher emission scenarios (SSP585), indicating large uncertainties in soil responses to future climate change. In addition, based on a smaller coefficient of variance for pH values (Figure S17), most field regions showed less variability of soil pH under a lower SSP scenario (SSP126) compared to a higher SSP scenario (SSP585) (Figure 5B), which was further cross-validated by the IPSL-CM6A-LR climate change model (Figures S18 and S19).

The biogeographic distribution of predicted grassland soil pH showed pronounced temporal dynamic changes. Greater changes in soil pH under future climate scenarios were found in northeastern Asia, Africa, and Oceania, but fewer changes were found in central North America, southern Africa, and eastern Asia (Figure 3). Therefore, the former regions should be given special consideration in future policy decisions, as large changes may have significant feedback on plant growth and productivity.[34] It was reported that the abundance of plants can be reduced due to increased soil alkaloid concentrations.[35] Although soil buffering systems maintain equilibriums of anion and cation concentrations in response to disturbances,[1] microbial activities are essential for maintaining soil pH. The metabolic pathway processes related to proton and hydroxyl ion generation and consumption could affect pH conditions, such as ammonification and denitrification. Similarly, the 10 identified eco-clusters, characterized by the CorBacFM framework to be functionally associated with nitrogen cycling and carbon degradation processes, potentially influence soil pH changes (Table S6). For example, the functions related to denitrification and respiration that could raise soil pH were significantly different than other functions in the eco-cluster of high-pH. Through these biogeochemical processes (Table S6), microbial activities and their enzymes may regulate soil pH via acid or alkaline substrates under warming conditions.[5] For example, it has been shown that ammonia concentration increased under warming treatment in a temperate grassland,[36] and long-term warming may enrich some microbially produced alkaline biochemicals, including saccharides and amino acids.[33] Moreover, molecular experiments have shown that bacteria may decrease environmental pH by glucose metabolism to regulate their swarming motility,[32] supporting the microbial regulation of surrounding pH conditions. In addition to microbial functions, other factors may contribute to the increase in soil pH, including nutrient uptake by plants, water content under warming conditions, and formation of carbonates due to a lower partial pressure of CO_2.[34] These cautious implications support soil ecological forecasts mediated by microbial functions.

Our prediction via the CoBacFM framework (Figure 1) and field validation (Figure 5) depicted the dynamic soil pH changes in global grasslands under future climate change, but some limitations remain. The relationships between soil pH and soil microbial communities are complex, and it remains difficult to quantitatively estimate the causal relationships of soil bacteria and soil pH changes in response to climate warming. PiecewiseSEM estimated the path sizes between soil pH and eco-clusters (Figures S8 and S14) and found that the bi-directed paths changed for different eco-clusters and that it is unclear for soil bacterial community. This is consistent with the traditional viewpoint that soil pH is generally a primary driver of bacterial community composition and microbial activities. Meanwhile, microbial activities can further alter soil pH, and vice versa (e.g., microbial nitrification is a major driver of soil acidification). Thus, the soil pH and soil microbial communities are correlated, which is also a prerequisite of the CoBacFM framework to predict soil pH changes in grasslands through microbial eco-clusters as bioindicators in response to climate change. Another limitation is that the CoBacFM only presented potential consequences at the shallow surface of grassland soils in natural ecosystems, while the soil pH and microbial responses to climate change may vary with soil depth[37] and anthropogenic effects[23] (e.g., fertilization, mowing, grazing). The integration of multiple factors, including soil depth, soil topography, and parent material,[38] as well as other terrestrial ecosystem types and land use changes, might contribute to more comprehensive insights into global soil pH functions under future climate change. Moreover, other microbial communities, such as fungal communities involved in nutrient cycling, play important roles in soil functioning,[4] and should be considered in future model predictions. Considering the biases of the curated datasets, a global standardized survey campaign with uniform grids may improve the predictions as well, and future studies should try to cover understudied regions such as northeastern Asia and Africa. In addition, integrating genomic and metagenomic data, and thus potential microbial functions,[39] can provide more information about biological

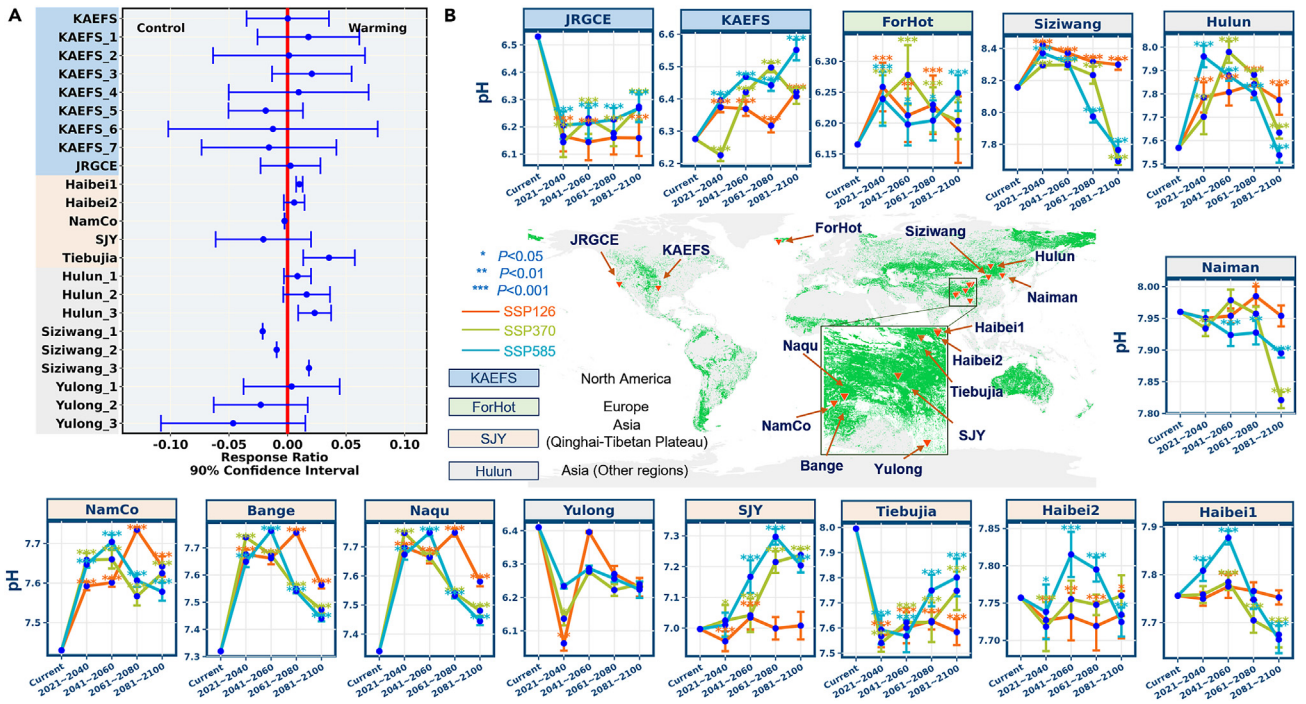


Figure 5. Soil pH changes in 14 global grassland sites at simulated warming in current and future scenarios

(A) Response ratio for the soil pH in warming conditions compared to control conditions in different experimental sites. The different numbers next to each site name indicate different sampling times.

(B) Global distribution of the 14 experimental sites related to simulated warming in grasslands, and the shifts of their edaphic pH with the predicted values under 3 future scenarios (BCC-CSM2-MR model) compared to the current values. Data are shown as mean ± SD, and the significance was tested using a 1-sample Student's t test for different levels (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

pathways with directed or undirected effects of ecological consequences and then strengthen the microbial predictions with 16S rRNA dataset to macro-scale ecosystem functions. Integrating and developing more robust methodology techniques for Earth systems models may improve the forecasts of future climate and find potential mitigation strategies, thus helping human society.[14] Specifically, expanding to less-studied grassland regions (Figure 5), long-term ecological surveys and simulation experiments are needed to better understand microbial states and future ecosystem functions.[40]

Conclusion

Our proposed CoBacFM framework enables the prediction of soil pH in global grassland regions via microbial responses to climate change, as done here to forecast the dynamic temporal changes of soil pH until 2100 under various climate change scenarios. Soil pH would increase in over 60% of grassland regions, almost half of which would be more alkaline under future climate change, with the area of alkaline soils expanding over time. This is important for designing future grassland management strategies, especially for regions in northeastern Asia, Africa, and Oceania, as soil pH determines the ecological niches of organisms and grass productivity.[1,41] Moreover, these results were supported by our cross-validation data from 14 warming experiments that showed that soil pH was affected by warming, including a significant increase in two sites, and the predicted soil changes at most sites also showed alkalization tendency under future climate scenarios. However, the traditional view holds that soil pH is less affected by external environmental disturbances,[1] and that chronic changes in soil features urgently require long-term monitoring systems. The validation by field experiments bridges the ecological model prediction and the observatory monitoring of terrestrial ecosystems, promoting the ecological modeling development, and demonstrating insights into decision-making strategies in response to climate change for ecosystem functioning and sustainable stability. In addition, the present study focused only on the soil pH across global grasslands because of less artificial intervention and more sensitivity to large-scale climate changes than farmlands and forests. However, with the accumulation of soil microbiome and human activity data, we believe our microbe-central model could be extended to other global ecosystems focusing on more microbe-relevant parameters, such as SOCs and storage,[42] because we are all living on a microbe-dominated Earth.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Information related to this manuscript can be obtained from the lead contact, Ye Deng (yedeng@rcees.ac.cn).

Materials availability

The materials and dataset used in this study can be found in the supplemental information and the publicly archived repository (https://doi.org/10.5281/zenodo.8004812).

Global grassland microbiome data collection and processing

We systematically searched for peer-reviewed articles that were published from 2006 to 2020 (accessed on October 10, 2020) and focused on microbial community analysis in natural grassland fields in the Web of Science by using keyword combinations such as environment (soil), organism (microb* or microorganism*), techniques (OTU* or ASV* or phylotypes* or bacteria* or archae* or fung* or 16S or 18S or ITS), habitat (grassland or steppe* or meadow* or savanna* or prairie*), and excluded some words (DGGE or PLFA or spp. or microcosm* or enzyme* activit*). After including some manually curated papers, the final total was 6,159 articles. We then filtered each article manually using the following criteria: (1) the samples were collected from topsoil (0–20 cm) within natural grassland fields, excluded lab-cultured samples; (2) the raw sequences were obtained using the Illumina platform, and covered the V4 region of the 16S rRNA gene; and (3) the raw sequencing reads of the research were complete, without data missing, and could be downloaded from public repositories. The final dataset retained 3,703 samples from 63 studies, covering 1,251 grassland sites across all continents except for Antarctica (Figure 2, listed in Table S1).

After obtaining the raw sequencing datasets, the processing of sequences was conducted on an in-house pipeline,[43] the detailed workflow of which has been previously described.[26] Briefly, the FLASH program was used to merge the pair-end sequencing reads,[44] and the primer set of 515F/806R was used to extract the targeted V4 region of 16S rRNA gene amplicons.[45] Btrim was used to filter for high-quality reads,[46] retaining only the non-N base reads within the range of 245–260 bps, resulting in 259,719,239 clean reads. The amplicon sequence variants (ASVs) table was generated based on UNOISE using Vsearch after removing chimeras *de novo*,[47] resulting in the creation of nearly 1.26 million ASVs. After assigning the taxonomy of each ASV against the SILVA database (version 138.1)[48] and retaining only "bacteria" sequences, we merged samples from within the same sites and rarefied the ASV table with the 20,000 reads per site. MAFFT[49] and FastTree[50] were used to align the representative sequences and construct a phylogenetic tree for the ASVs.

The soil features of each sample were collected from the published articles, including soil pH, SOC, and TN. The current climatic variables and NDVI were extracted from WorldClim (version 2.1)[51] and NASA Earth observations (https://neo.gsfc.nasa.gov/) according to the geographic locations, respectively.

CoBacFM

Core grassland microbiota and biodiversity analysis

The core grassland microbiota was identified according to the prevalence of ASVs—in other words, the ASVs that occurred in more than 40% of the sites within a continent[26] and were observed on more than 3 continents—resulting in a total of 989 ASVs (Figure 1). The core community comprised 29.4% of all sequence reads in the global dataset (relative to 20,000 reads per sample) and its relative sequence abundance varied from 20.0% to 36.8% within each continent (Figure S1). Representing the soil microbiota with a fraction of all sequences is a necessary trade-off for using dimension-reduced groups of species.[27] The α-diversity (Shannon, Richness, and Phylogenetic diversity) and β-diversity (weighted UniFrac distance) were calculated to measure the biodiversity profiles of the core bacterial community. Structural equation models (SEMs) and Mantel tests were used to explore the relationships between biodiversity and environmental variables for the core bacterial community. The first axis of the principal-component analyses of climate, soil, and α-diversity features was included in the SEM analysis, and the Pearson correlation coefficient was used to show the correlations between each variable and its respective principal component.

Eco-cluster assignments

The core microbiota were clustered into 10 eco-clusters (specific microbial groups potentially sharing similar functions) using the following steps with similar workflows[26,27] (Figure 1; Data S1): (1) 8 environmental variables were selected for this analysis, including MAT, MaxTWarmM, MAP and PWetM, soil pH, SOC, TN, and NDVI; (2) random forest modeling was used to predict the ASV relative abundance by the selected environmental variables, retaining only the ASVs with more than 30% explanations; (3) semi-partial correlation coefficients were calculated between ASV relative abundance and each variable; and (4) ASVs with significant environmental associations ($p < 0.001$) were assigned to eco-clusters using cluster analysis based on Euclidean distance with the ward.D algorithm, which resulted in 10 eco-clusters associated with high-pH, low-pH, high-SOC, low-SOC, high-pH and TN with low-SOC, low-pH-high-SOC, low-pH-low-MAT, high-PWetM, high-MaxTWarmM, and high-NDVI. Piecewise SEM[52] was used to evaluate the relationships between environmental variables and eco-cluster relative abundances, and then to quantify the contribution of each eco-cluster relative abundance to infer soil parameters (e.g., soil pH [Figure S8]). The best model of PiecewiseSEM was determined by backward selection with the lowest Akaike information criterion values. The relative abundances of the assigned 10 clusters showed linear or unimodal patterns with the preferred environmental variables (Figure S5), and each eco-cluster contained certain taxonomic groups (Figures 2 and S6). The relationship between soil parameters and the relative abundance of eco-clusters was bi-directionality; in other words, soil parameters, such as soil pH, showed directed paths to eco-clusters, and there were paths from eco-clusters to soil pH (Figure S8). The standardized total effects for each environmental variable to explain eco-clusters and for eco-clusters to predict soil parameters were estimated (Figures S9 and S13). To further explore the functional potentials of the eco-clusters, Functional Annotation of Prokaryotic Taxa (FAPROTAX) (version 1.2.7)[53] was used to link each eco-cluster to different ecological processes (Table S6). The Fisher's exact test or chi-square with Yates' correction test was used to test whether the assigned function potential to each eco-cluster was significantly different from other functions. The analyses and plots were conducted in R (version 4.1.1) with "vegan" (version 2.5–7), "randomforest" (version 4.7–1), "piecewiseSEM" (version 2.1.2), "ppcor" (version 1.1), "pheatmap" (version 1.0.12), "parallel" (version 4.1.1) and "ggplot2" (version 3.3.5) packages.

Geospatial modeling

A gridded grassland map with a resolution of 0.1° × 0.1° was extracted from the 2020 version of GlobeLand30,[54] containing 284,756 gridded points. For each point, three groups of environmental variables—climate (MAT, MaxTWarmM, MAP, PWetM), soil (pH, SOC, TN) and plant (NDVI)—were extracted from the public databases of WorldClim 2.1,[51] SoilGrid 2.0,[15] and NASA, respectively, and considered as an environmental variable dataset for the current scenario. Furthermore, the extent of extrapolation of the collected grassland dataset was calculated using a convex hull approach to evaluate the global-scale environmental variables based on the above training dataset.[55] Based on the prediction methods for eco-cluster relative abundances and soil pH below, each gridded point was assigned a predicted value, and subsequently, the prediction was mapped to a finer grassland map using an automatic kriging approach from the "automap" (version 1.0-14) package. An automatically determined semivariogram model was used for spatial interpolation. The geospatial modeling and map plots were conducted by "gstat" (version 2.0-7), "maptools" (version 1.1-1), "raster" (version 3.4-13), and "ggplot2" (version 3.3.5) packages in R.

Cubist models for current eco-clusters

The Cubist model is a decision-based regression tree to estimate target variable values from the explanatory variables, and the model attempts to minimize the biased predictions by using committee models and nearest-neighbor models.[56] The Cubist model is widely used in geospatial analysis.[26,27] To simplify, the dataset used to build the Cubist model for eco-clusters was randomly divided into 10 sub-datasets—9 as training datasets and 1 as a test dataset—and then the Cubist committee model was constructed according to Equations 1–4 based on the training datasets (Figure 1). In Equation 1, the subscript i indicates the 10 assigned eco-clusters, and the function is built by the Cubist method. The environmental variables under the current scenario, indicated by the subscript "current," were extracted from the public databases.

$$Ecoclusterabundance(i)_{current} = f_{Cubist}(Climate_{current}, Soil_{current}, Plant_{current}) \quad \text{(Equation 1)}$$

$$Climate_{current} \sim f(MAT_{current}, MaxTWarmM_{current}, MAP_{current}, PWetM_{current}) \quad \text{(Equation 2)}$$

$$Soil_{current} \sim f(pH_{current}, SOC_{current}, TN_{current}) \quad \text{(Equation 3)}$$

$$Plant_{current} \sim f(\mathrm{NDVI}_{current}) \quad \text{(Equation 4)}$$

Once the Cubist model was constructed, the average error and correlation coefficient were used to evaluate model performance on the training datasets, and the root-mean-square error (RMSE, Equation 5) and correlation coefficient were used to evaluate model performance on the test dataset.

$$RMSE = \sqrt{\frac{\sum_{j=1}^{m}(predicted_j - original_j)^2}{m}} \quad \text{(Equation 5)}$$

The m indicates the total number of data records used in the Cubist models, and j belongs to the range of 1–m. Using this constructed Cubist model, the relative abundance of each eco-cluster was predicted for the grassland gridded points (284,756 points). The analysis was conducted using the "Cubist" (version 0.4.0) package in R.

The 10-fold cross-validation of cubist models

We created the global distribution map for each eco-cluster using an ensemble approach, by taking average values from the 10-fold cross-validated Cubist models (Figure 1). The grassland soil microbiota dataset was randomly divided into 10 sub-datasets. Then, 9 sub-datasets were included as a training dataset, and the remaining sub-dataset was used as a test dataset for the Cubist model. Subsequently, the Cubist model was constructed as outlined above and was used to predict the grassland gridded points globally at the current climatic scenario. This step was repeated 10 times by selecting different combinations of sub-datasets as the training dataset. The final predicted value (relative abundance) for each eco-cluster was averaged from the 10 prediction results (Equation 6), and the coefficient of variation was used to indicate the uncertainty of the predicted values following Equations 7 and 8.

$$\overline{Ecoclusterabundance(i)} = \left(\sum_{k=1}^{n} Ecoclusterabundance_k\right) \Big/ n, n = 10 \quad \text{(Equation 6)}$$

$$Standarddeviation_{Ecoclusterabundance(i)} = \sqrt{\frac{\sum_{k=1}^{10}(Ecoclusterabundance_k - \overline{Ecoclusterabundance(i)})^2}{n-1}}, n = 10 \quad \text{(Equation 7)}$$

$$Uncertainty_{10foldcrossvalidation} = Standarddeviation_{Ecoclusterabundance(i)} \Big/ \overline{Ecoclusterabundance(i)} \quad \text{(Equation 8)}$$

n indicates the number of cross-validations (10 in this case), and k belongs to the range of 1–n.

Cubist models for future eco-clusters

The climatic variables under future climate scenarios for the grassland gridded points were extracted separately from the WorldClim 2.1 database[51] for specific scenarios (SSP126, SSP370, and SSP585), specific periods (2021–2040, 2041–2060, 2061–2080, and 2081–2100), and climatic models (BCC-CSM2-MR from Beijing Climate Center[57] and IPSL-CM6A-LR from Institut Pierre-Simon Laplace[58]). The selected three climate scenarios estimated different air temperature increments—0.6°C–1.8°C for SSP126, 3.0°C–6.5°C for SSP370, and 3.8°C–8.6°C for SSP585 by the end of 2100.[59] We then predicted the distribution of eco-cluster relative abundances under future climate scenarios using Equations 9 and 10 with an approach similar to that used for current conditions. The subscript "future" in the equations indicates all combinations of the three SSP scenarios, four time periods, and two climatic models, and each case was conducted separately.

$$Ecoclusterabundance(i)_{future} = f_{Cubist}(Climate_{future}, Soil_{current}, Plant_{current}) \quad \text{(Equation 9)}$$

$$Climate_{future} \sim f(MAT_{future}, MaxTWarmM_{future}, MAP_{future}, PWetM_{future}) \quad \text{(Equation 10)}$$

Cubist models for soil pH

Using the assigned eco-cluster relative abundances at current conditions, the Cubist model for soil pH was built (Equation 11) and was used to predict soil pH under future climate change in grassland ecosystems using Equation 12. Also, the Cubist model performance was evaluated using the correlation coefficient based on the 10-fold cross-validated results.

$$pH_{current} = \frac{\sum_{k=1}^{10} f_{Cubist,k}(Ecoclusterabundance(1)_{current}, \ldots, Ecoclusterabundance(i)_{current})}{10} \quad \text{(Equation 11)}$$

$$pH_{future} = \frac{\sum_{k=1}^{10} f_{Cubist,k}(Ecoclusterabundance(1)_{future}, \ldots, Ecoclusterabundance(i)_{future})}{10} \quad \text{(Equation 12)}$$

The areas of pH-changed regions were summarized based on predicted pixel-gridded points, and the significance of the soil pH change was evaluated using the paired Student's t test corrected with the Bonferroni method.

Cross-validation of field experiments

In addition to the cross-validation of statistical strategies with the ensemble approach, we included 14 warming experiments (1°C–6°C increasing

temperature versus ambient condition) carried out on grasslands (Table S4) to validate model predictions. These 14 experimental sites were collected and filtered out following the same procedure and data inclusion criterion in the above section "Global grassland microbiome data collection and processing" (updated October 5, 2023), mainly distributed in Asia, Europe, and North America. Within each site, the response ratio was used to show whether significant differences existed in eco-clusters and soil pH features under warming versus ambient conditions (at a 90% confidence level). In addition, the predicted soil pH for each experimental site was an average of modeling results based on the specific geographic location of each site and its surrounding regions ($0.1^\circ \times 0.1^\circ$), and the significance levels were determined using the one-sample Student's t test.

DATA AND CODE AVAILABILITY

The raw dataset was downloaded from public repositories of published articles (Table S1). The dynamic change of images, scripts and relevant datasets are accessible on the download page (https://denglab.org.cn) and publicly archived repository (https://doi.org/10.5281/zenodo.8004812).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.oneear.2024.06.002.

ACKNOWLEDGMENTS

This project was supported by the National Key Research and Development Program of China (no. 2022YFE0114000), the National Science Foundation of China (nos. 42207141 and U23A2043), the China Postdoctoral Science Foundation (no. 2021M703410), and the Wenshan Tobacco Company of Yunnan Province of China (no. 2021530000241033). We thank Dr. James W. Voordeckers for correcting the wording and grammar of this paper. We also thank Paul B.L. George, Sanghoon Kang, and He Hu for corrections to their public data.

AUTHOR CONTRIBUTIONS

K.F. and Y.D. designed the study and the framework. K.F., Q. He, Z. Wang, X.P., D.W., S.L., Y.W., Z.J., X.D., C.Y., S.G., T.L., X.Y., W.S., Z. Wei, Q. Hu, P.L., and Y.Z. contributed to the dataset collection. G.L., C.Q., G.Z., C.X., Y.Y., and J.Z. provided the simulated warming samples (SJY, JRGCE, NamCo, Hulun, Haibei2, and KAEFS). K.F. conducted the data processing, ecological modeling, and statistical analysis. K.F. and Y.D. wrote the manuscript, with revisions from S.W., M. Bonkowski, M. Bahram, E.Y., and J.Z. All authors read and approved the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 27, 2023
Revised: November 26, 2023
Accepted: June 11, 2024
Published: July 5, 2024

REFERENCES

1. Weil, R.R., and Brady, N.C. (2016). The Nature and Properties of Soils, Fifteenth edition (Pearson).
2. Banwart, S.A., Nikolaidis, N.P., Zhu, Y.-G., Peacock, C.L., and Sparks, D.L. (2019). Soil Functions: Connecting Earth's Critical Zone. Annu. Rev. Earth Planet Sci. *47*, 333–359. https://doi.org/10.1146/annurev-earth-063016-020544.
3. Xu, T., Zhang, M., Ding, S., Liu, B., Chang, Q., Zhao, X., Wang, Y., Wang, J., and Wang, L. (2021). Grassland degradation with saline-alkaline reduces more soil inorganic carbon than soil organic carbon storage. Ecol. Indicat. *131*, 108194. https://doi.org/10.1016/j.ecolind.2021.108194.
4. Philippot, L., Chenu, C., Kappler, A., Rillig, M.C., and Fierer, N. (2024). The interplay between microbial communities and soil properties. Nat. Rev. Microbiol. *22*, 226–239. https://doi.org/10.1038/s41579-023-00980-5.
5. Fanin, N., Mooshammer, M., Sauvadet, M., Meng, C., Alvarez, G., Bernard, L., Bertrand, I., Blagodatskaya, E., Bon, L., Fontaine, S., et al. (2022). Soil enzymes in response to climate warming: Mechanisms and feedbacks. Funct. Ecol. *36*, 1378–1395. https://doi.org/10.1111/1365-2435.14027.
6. Vet, R., Artz, R.S., Carou, S., Shaw, M., Ro, C.-U., Aas, W., Baker, A., Bowersox, V.C., Dentener, F., Galy-Lacaux, C., et al. (2014). A global assessment of precipitation chemistry and deposition of sulfur, nitrogen, sea salt, base cations, organic acids, acidity and pH, and phosphorus. Atmos. Environ. *93*, 3–100. https://doi.org/10.1016/j.atmosenv.2013.10.060.
7. Syed, S., Buddolla, V., and Lian, B. (2020). Oxalate Carbonate Pathway—Conversion and Fixation of Soil Carbon—A Potential Scenario for Sustainability. Front. Plant Sci. *11*, 591297.
8. Zhou, Z., Wang, C., and Luo, Y. (2020). Meta-analysis of the impacts of global change factors on soil microbial diversity and functionality. Nat. Commun. *11*, 3072. https://doi.org/10.1038/s41467-020-16881-7.
9. Tripathi, B.M., Stegen, J.C., Kim, M., Dong, K., Adams, J.M., and Lee, Y.K. (2018). Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. ISME J. *12*, 1072–1083. https://doi.org/10.1038/s41396-018-0082-4.
10. Guo, X., Feng, J., Shi, Z., Zhou, X., Yuan, M., Tao, X., Hale, L., Yuan, T., Wang, J., Qin, Y., et al. (2018). Climate warming leads to divergent succession of grassland microbial communities. Nat. Clim. Chang. *8*, 813–818. https://doi.org/10.1038/s41558-018-0254-2.
11. de Vries, F.T., Griffiths, R.I., Bailey, M., Craig, H., Girlanda, M., Gweon, H.S., Hallin, S., Kaisermann, A., Keith, A.M., Kretzschmar, M., et al. (2018). Soil bacterial networks are less stable under drought than fungal networks. Nat. Commun. *9*, 3033. https://doi.org/10.1038/s41467-018-05516-7.
12. Zhou, J., Deng, Y., Shen, L., Wen, C., Yan, Q., Ning, D., Qin, Y., Xue, K., Wu, L., He, Z., et al. (2016). Temperature mediates continental-scale diversity of microbes in forest soils. Nat. Commun. *7*, 12083. https://doi.org/10.1038/ncomms12083.
13. Wang, S., Wang, X., Han, X., and Deng, Y. (2018). Higher precipitation strengthens the microbial interactions in semi-arid grassland soils. Glob. Ecol. Biogeogr. *27*, 570–580. https://doi.org/10.1111/geb.12718.
14. Tiedje, J.M., Bruns, M.A., Casadevall, A., Criddle, C.S., Eloe-Fadrosh, E., Karl, D.M., Nguyen, N.K., Zhou, J., and Whiteley, M. (2022). Microbes and Climate Change: a Research Prospectus for the Future. mBio *13*, 800222–e100822. https://doi.org/10.1128/mbio.00800-22.
15. Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., and Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL *7*, 217–240. https://doi.org/10.5194/soil-7-217-2021.
16. Guerra, C.A., Bardgett, R.D., Caon, L., Crowther, T.W., Delgado-Baquerizo, M., Montanarella, L., Navarro, L.M., Orgiazzi, A., Singh, B.K., Tedersoo, L., et al. (2021). Tracking, targeting, and conserving soil biodiversity. Science *371*, 239–241. https://doi.org/10.1126/science.abd7926.
17. Coban, O., De Deyn, G.B., and van der Ploeg, M. (2022). Soil microbiota as game-changers in restoration of degraded lands. Science *375*, abe0725. https://doi.org/10.1126/science.abe0725.
18. Sokol, N.W., Slessarev, E., Marschmann, G.L., Nicolas, A., Blazewicz, S.J., Brodie, E.L., Firestone, M.K., Foley, M.M., Hestrin, R., Hungate, B.A., et al. (2022). Life and death in the soil microbiome: how ecological processes influence biogeochemistry. Nat. Rev. Microbiol. *20*, 415–430. https://doi.org/10.1038/s41579-022-00695-z.
19. Anthony, M.A., Crowther, T.W., Maynard, D.S., van den Hoogen, J., and Averill, C. (2020). Distinct Assembly Processes and Microbial Communities Constrain Soil Organic Carbon Formation. One Earth *2*, 349–360. https://doi.org/10.1016/j.oneear.2020.03.006.

20. Tao, F., Huang, Y., Hungate, B.A., Manzoni, S., Frey, S.D., Schmidt, M.W.I., Reichstein, M., Carvalhais, N., Ciais, P., Jiang, L., et al. (2023). Microbial carbon use efficiency promotes global soil carbon storage. Nature *618*, 981–985. https://doi.org/10.1038/s41586-023-06042-3.

21. Kuypers, M.M.M., Marchant, H.K., and Kartal, B. (2018). The microbial nitrogen-cycling network. Nat. Rev. Microbiol. *16*, 263–276. https://doi.org/10.1038/nrmicro.2018.9.

22. van der Heijden, M.G.A., Bardgett, R.D., and Van Straalen, N.M. (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecol. Lett. *11*, 296–310.

23. Rillig, M.C., Ryo, M., Lehmann, A., Aguilar-Trigueros, C.A., Buchert, S., Wulf, A., Iwasaki, A., Roy, J., and Yang, G. (2019). The role of multiple global change factors in driving soil functions and microbial biodiversity. Science *366*, 886–890. https://doi.org/10.1126/science.aay2832.

24. Yuan, M.M., Guo, X., Wu, L., Zhang, Y., Xiao, N., Ning, D., Shi, Z., Zhou, X., Wu, L., Yang, Y., et al. (2021). Climate warming enhances microbial network complexity and stability. Nat. Clim. Chang. *11*, 343–348. https://doi.org/10.1038/s41558-021-00989-9.

25. Locey, K.J., and Lennon, J.T. (2016). Scaling laws predict global microbial diversity. Proc. Natl. Acad. Sci. USA *113*, 5970–5975. https://doi.org/10.1073/pnas.1521291113.

26. Wang, S., Bao, X., Feng, K., Deng, Y., Zhou, W., Shao, P., Zheng, T., Yao, F., Yang, S., Liu, S., et al. (2021). Warming-driven migration of core microbiota indicates soil property changes at continental scale. Sci. Bull. *66*, 2025–2035. https://doi.org/10.1016/j.scib.2021.01.021.

27. Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., Maestre, F.T., Singh, B.K., and Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. Science *359*, 320–325. https://doi.org/10.1126/science.aap9516.

28. Bardgett, R.D., Bullock, J.M., Lavorel, S., Manning, P., Schaffner, U., Ostle, N., Chomel, M., Durigan, G., L. Fry, E., Johnson, D., et al. (2021). Combatting global grassland degradation. Nat. Rev. Earth Environ. *2*, 720–735. https://doi.org/10.1038/s43017-021-00207-2.

29. Jansson, J.K., and Hofmockel, K.S. (2020). Soil microbiomes and climate change. Nat. Rev. Microbiol. *18*, 35–46. https://doi.org/10.1038/s41579-019-0265-7.

30. Ho, A., Di Lonardo, D.P., and Bodelier, P.L.E. (2017). Revisiting life strategy concepts in environmental microbial ecology. FEMS Microbiol. Ecol. *93*, fix006. https://doi.org/10.1093/femsec/fix006.

31. Enquist, B.J., Abraham, A.J., Harfoot, M.B.J., Malhi, Y., and Doughty, C.E. (2020). The megabiota are disproportionately important for biosphere functioning. Nat. Commun. *11*, 699. https://doi.org/10.1038/s41467-020-14369-y.

32. Dharanishanthi, V., Orgad, A., Rotem, N., Hagai, E., Kerstnus-Banchik, J., Ben-Ari, J., Harig, T., Ravella, S.R., Schulz, S., and Helman, Y. (2021). Bacterial-induced pH shifts link individual cell physiology to macroscale collective behavior. Proc. Natl. Acad. Sci. USA *118*, e2014346118. https://doi.org/10.1073/pnas.2014346118.

33. Gargallo-Garriga, A., Sardans, J., Ayala-Roque, M., Sigurdsson, B.D., Leblans, N.I., Oravec, M., Klem, K., Janssens, I.A., Urban, O., and Peñuelas, J. (2021). Warming affects soil metabolome: The case study of Icelandic grasslands. Eur. J. Soil Biol. *105*, 103317. https://doi.org/10.1016/j.ejsobi.2021.103317.

34. Wong, V.N.L., Greene, R.S.B., Dalal, R.C., and Murphy, B.W. (2010). Soil carbon dynamics in saline and sodic soils: a review. Soil Use Manag. *26*, 2–11. https://doi.org/10.1111/j.1475-2743.2009.00251.x.

35. McCulley, R.L., Bush, L.P., Carlisle, A.E., Ji, H., and Nelson, J.A. (2014). Warming reduces tall fescue abundance but stimulates toxic alkaloid concentrations in transition zone pastures of the U.S. Front. Chem. *2*, 88. https://doi.org/10.3389/fchem.2014.00088.

36. Zhou, X., Chen, C., Wang, Y., Xu, Z., Han, H., Li, L., and Wan, S. (2013). Warming and increased precipitation have differential effects on soil extracellular enzyme activities in a temperate grassland. Sci. Total Environ. *444*, 552–558. https://doi.org/10.1016/j.scitotenv.2012.12.023.

37. Wang, Z., Feng, K., Lu, G., Yu, H., Wang, S., Wei, Z., Dang, N., Wang, Y., and Deng, Y. (2022). Homogeneous Selection and Dispersal Limitation Dominate the Effect of Soil Strata Under Warming Condition. Front. Microbiol. *13*, 801083. https://doi.org/10.3389/fmicb.2022.801083.

38. Jenny, H. (1941). Factors of Soil Formation. Soil Sci. *52*, 415.

39. Ramoneda, J., Stallard-Olivera, E., Hoffert, M., Winfrey, C.C., Stadler, M., Niño-García, J.P., and Fierer, N. (2023). Building a genome-based understanding of bacterial pH preferences. Sci. Adv. *9*, eadf8998. https://doi.org/10.1126/sciadv.adf8998.

40. Bardgett, R.D., and Caruso, T. (2020). Soil microbial community responses to climate extremes: resistance, resilience and transitions to alternative states. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190112. https://doi.org/10.1098/rstb.2019.0112.

41. Žurovec, O., Wall, D.P., Brennan, F.P., Krol, D.J., Forrestal, P.J., and Richards, K.G. (2021). Increasing soil pH reduces fertiliser derived N2O emissions in intensively managed temperate grassland. Agric. Ecosyst. Environ. *311*, 107319. https://doi.org/10.1016/j.agee.2021.107319.

42. Bai, Y., and Cotrufo, M.F. (2022). Grassland soil carbon sequestration: Current understanding, challenges, and solutions. Science *377*, 603–608. https://doi.org/10.1126/science.abo2380.

43. Feng, K., Zhang, Z., Cai, W., Liu, W., Xu, M., Yin, H., Wang, A., He, Z., and Deng, Y. (2017). Biodiversity and species competition regulate the resilience of microbial biofilm community. Mol. Ecol. *26*, 6170–6182. https://doi.org/10.1111/mec.14356.

44. Magoc, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.

45. Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., Jansson, J.K., Caporaso, J.G., Fuhrman, J.A., et al. (2016). Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems *1*, e00009-15–e00015. https://doi.org/10.1128/mSystems.00009-15.

46. Kong, Y. (2011). Btrim: A fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. Genomics *98*, 152–153. https://doi.org/10.1016/j.ygeno.2011.05.009.

47. Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ *4*, e2584. https://doi.org/10.7717/peerj.2584.

48. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. *41*, D590–D596.

49. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

50. Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Mol. Biol. Evol. *26*, 1641–1650. https://doi.org/10.1093/molbev/msp077.

51. Fick, S.E., and Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. *37*, 4302–4315. https://doi.org/10.1002/joc.5086.

52. Lefcheck, J.S. (2016). piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. Methods Ecol. Evol. *7*, 573–579. https://doi.org/10.1111/2041-210X.12512.

53. Louca, S., Parfrey, L.W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. Science *353*, 1272–1277. https://doi.org/10.1126/science.aaf4507.

54. Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., et al. (2015). Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS J. Photogrammetry Remote Sens. *103*, 7–27. https://doi.org/10.1016/j.isprsjprs.2014.09.002.

55. Ma, H., Mo, L., Crowther, T.W., Maynard, D.S., van den Hoogen, J., Stocker, B.D., Terrer, C., and Zohner, C.M. (2021). The global distribution

and environmental drivers of aboveground versus belowground plant biomass. Nat. Ecol. Evol. *5*, 1110–1122. https://doi.org/10.1038/s41559-021-01485-1.

56. Quinlan, J.R.; Learning with continuous classes (1992). 5th Australian Joint Conference on Artificial Intelligence (World Scientific).

57. Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., et al. (2019). The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6. Geosci. Model Dev. (GMD) *12*, 1573–1600. https://doi.org/10.5194/gmd-12-1573-2019.

58. Boucher, O., Servonnat, J., Albright, A.L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., et al. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. J. Adv. Model. Earth Syst. *12*, e2019MS002010. https://doi.org/10.1029/2019MS002010.

59. Meinshausen, M., Nicholls, Z.R.J., Lewis, J., Gidden, M.J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., et al. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. Geosci. Model Dev. (GMD) *13*, 3571–3605. https://doi.org/10.5194/gmd-13-3571-2020.

One Earth, Volume *7*

Supplemental information

CoBacFM: Core bacteria forecast model for global grassland pH dynamics under future climate warming scenarios

Kai Feng, Shang Wang, Qing He, Michael Bonkowski, Mohammad Bahram, Etienne Yergeau, Zhujun Wang, Xi Peng, Danrui Wang, Shuzhen Li, Yingcheng Wang, Zhicheng Ju, Xiongfeng Du, Chengliang Yan, Songsong Gu, Tong Li, Xingsheng Yang, Wenli Shen, Ziyan Wei, Qiulong Hu, Pengfei Li, Yanmei Zhu, Guangxin Lu, Clara Qin, Gengxin Zhang, Chunwang Xiao, Yunfeng Yang, Jizhong Zhou, and Ye Deng

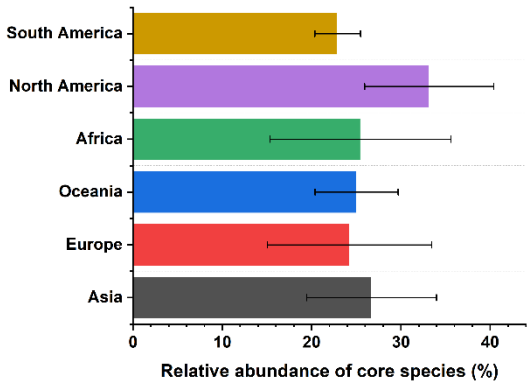


Figure S1 Relative abundance of the core community across six continents. The values are mean ±SD.

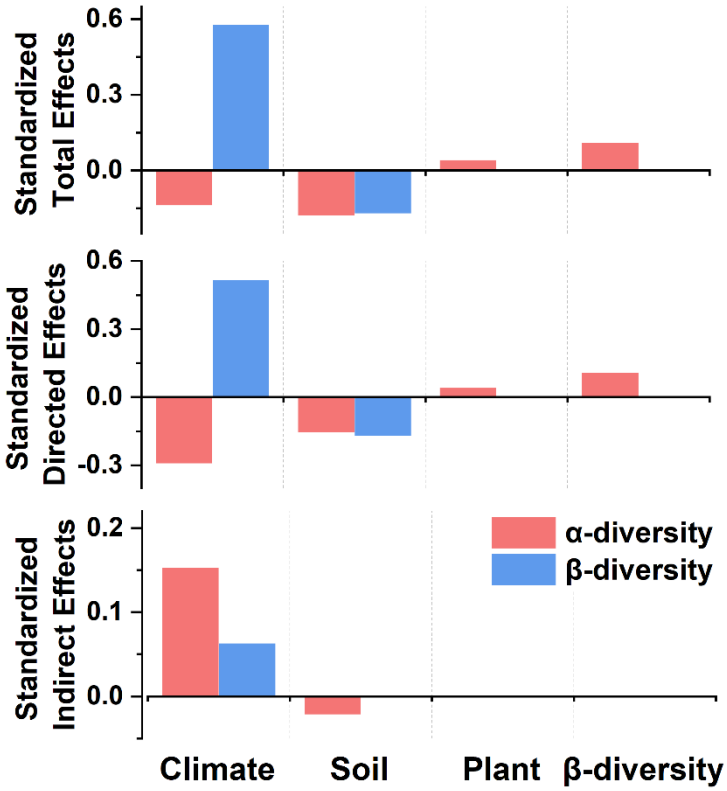


Figure S2 Standardized effects of structure equation models for core community analysis.

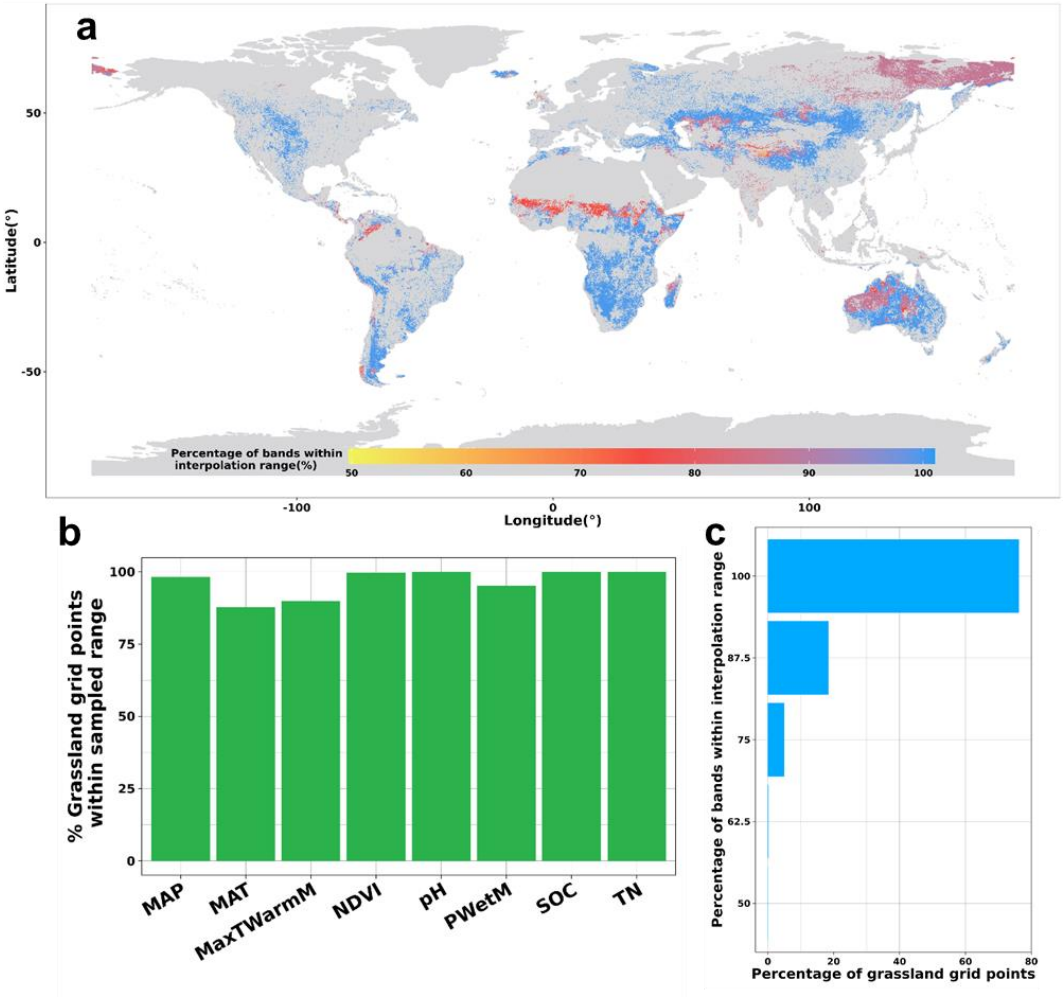


Figure S3 Percentage of interpolation environmental variable ranges within global grassland using the collected dataset at global scale (a), for the grid-based points of each variable (b), and the grid-based points summary (c).

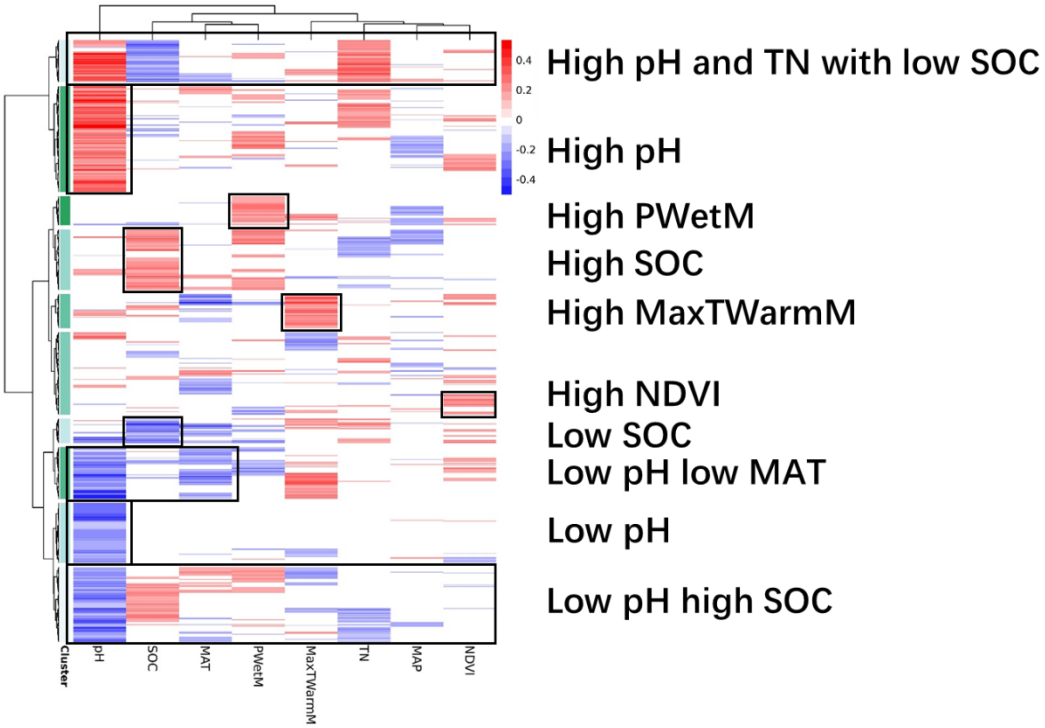


Figure S4 Clustering analysis of semi-partial correlations of ASV relative abundance and environmental variables. Red and blue colors indicate positive and negative correlation coefficients, respectively.

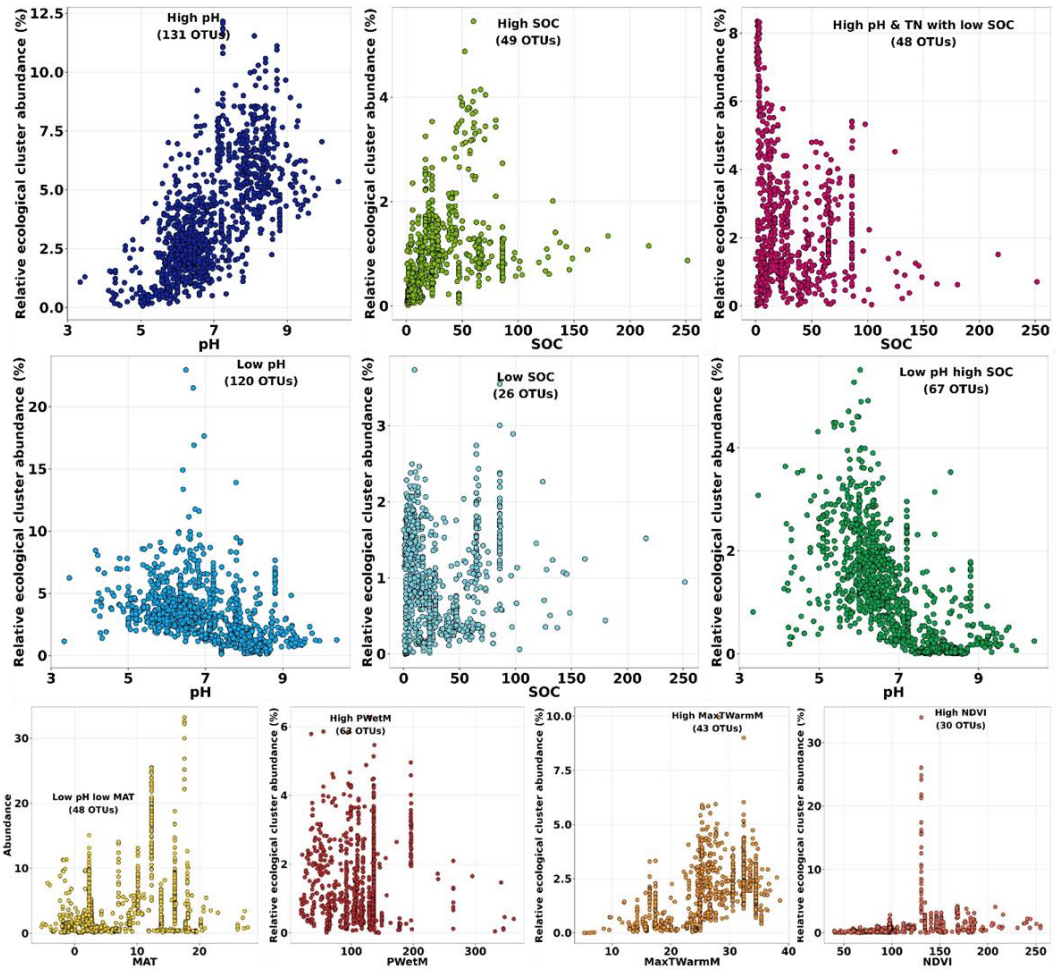


Figure S5 Relationships between relative abundance of core eco-clusters and their preferred environmental variables.

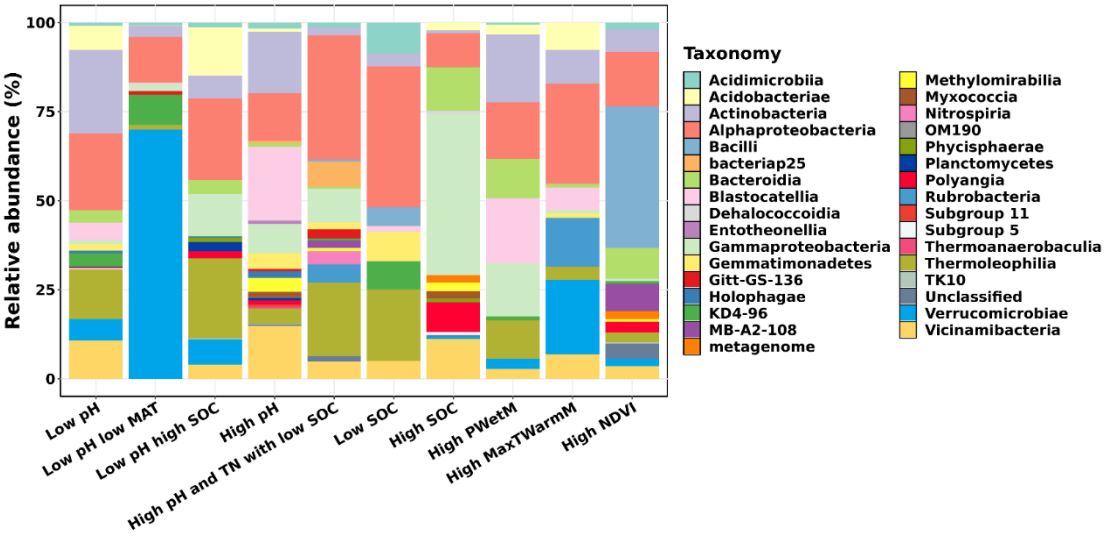


Figure S6 Taxonomic information of eco-clusters at the class level.

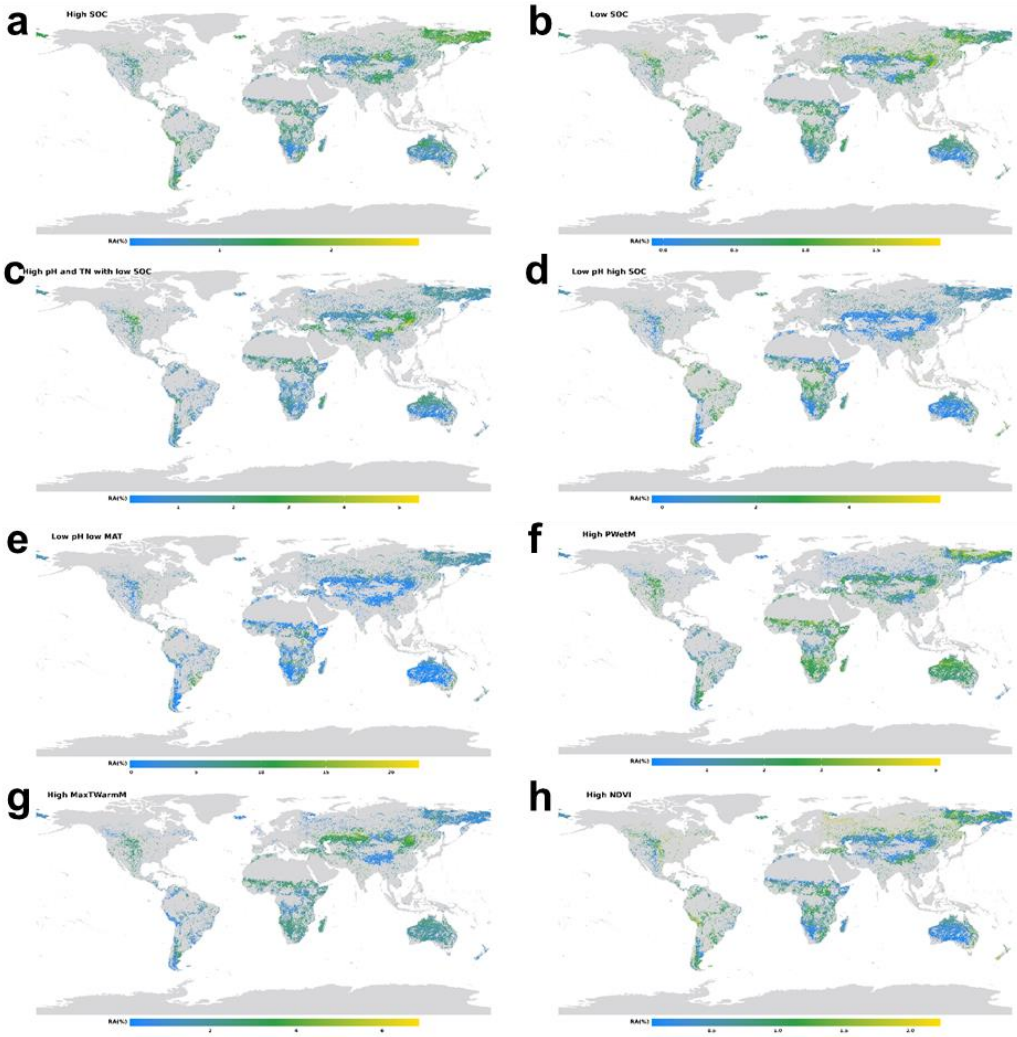


Figure S7 The biogeographic distributions for assigned eco-clusters of the core microbiota for global grasslands: (a) High SOC, (b) Low SOC, (c) Low pH and TN with low SOC, (d) Low pH high SOC, (e) Low pH low MAT, (f) High PWetM, (g) High MaxTWarmM, (h) High NDVI.

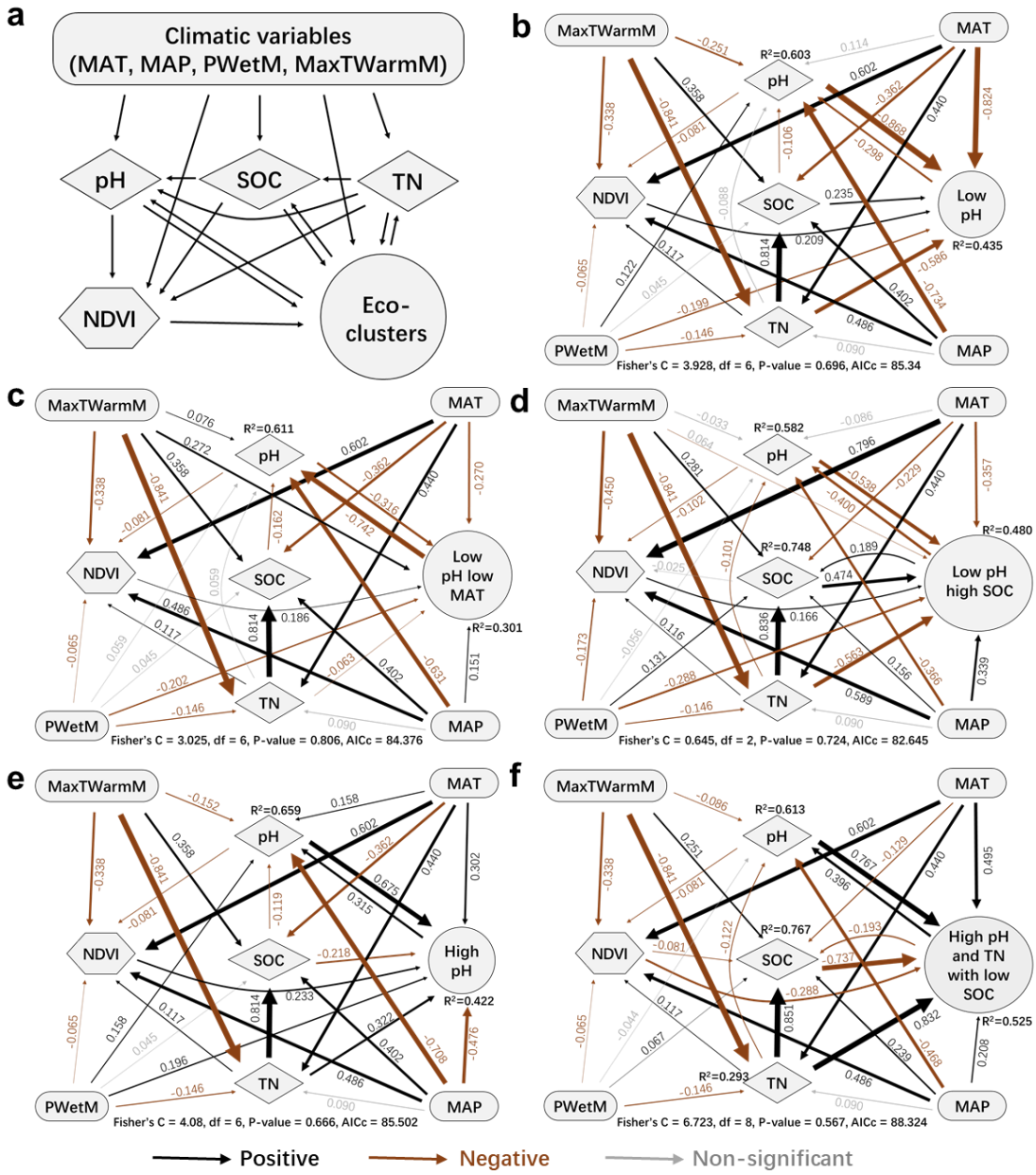


Figure S8 Piecewise structure equation models (PiecewiseSEMs) to show the relationships between environmental features and relative abundance of eco-clusters. The four climate variables (MAT, MaxTWarmM, MAP, and PWetM), three edaphic variables (pH, SOC, and TN), and plant index (NDVI) were included.

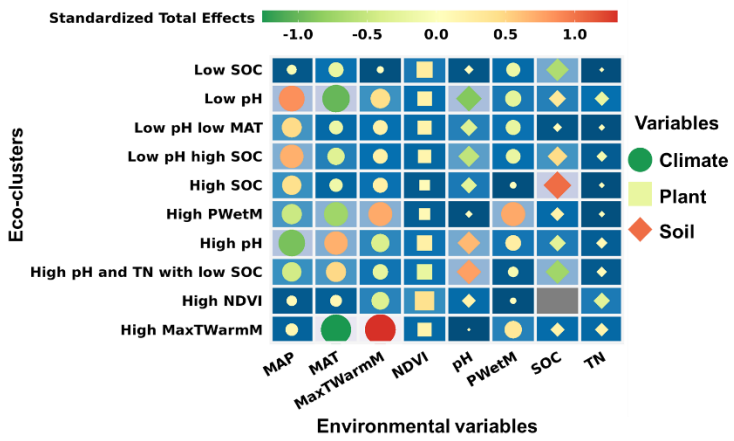


Figure S9 Standardized total effects of each environmental variable to the relative abundance of eco-clusters by using piecewise structure equation models (PiecewiseSEMs).

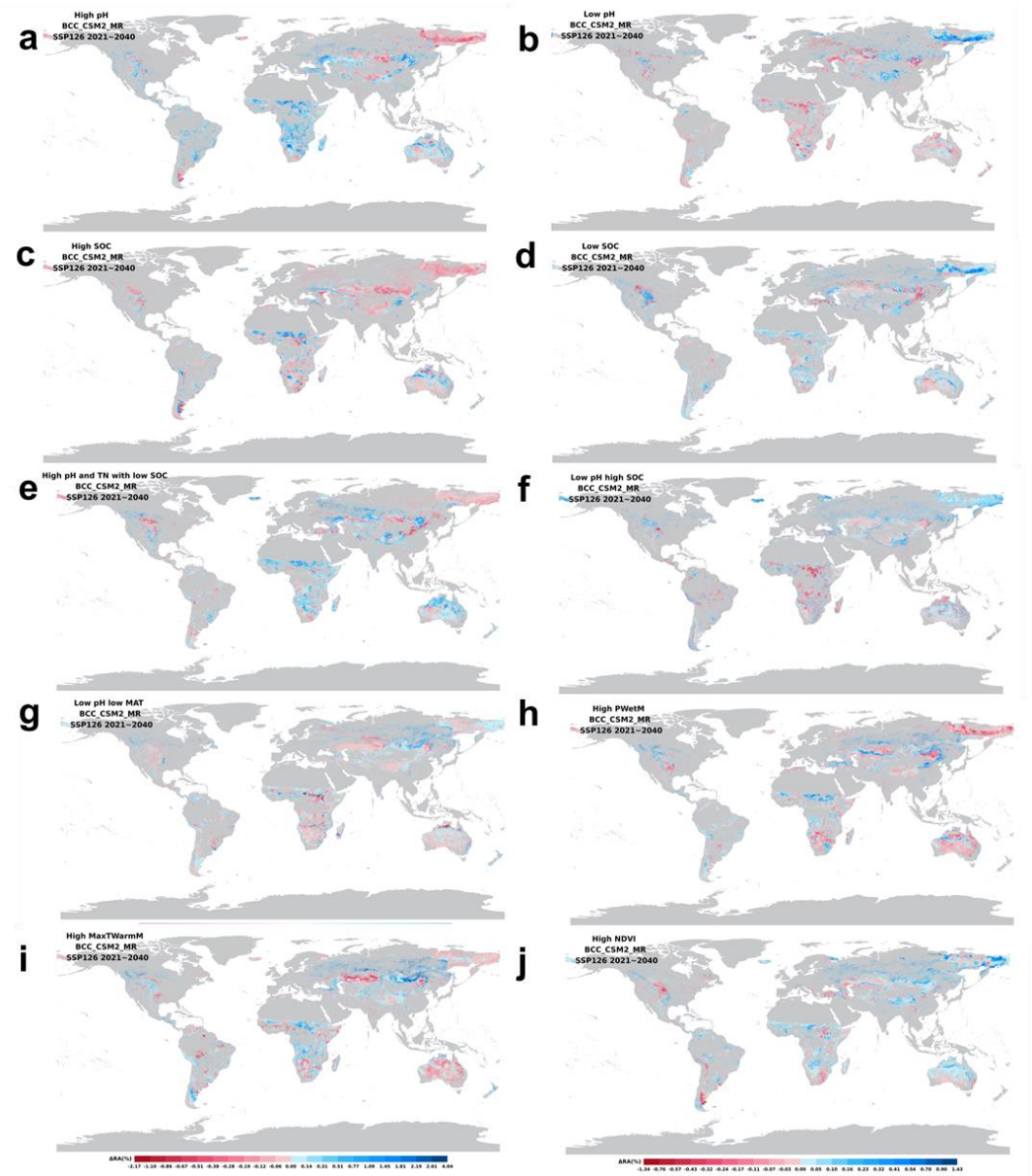


Figure S10 The predicted biogeographic distributions for assigned eco-clusters of the core microbiota for global grasslands with SSP126, SSP370 and SSP585 scenarios of temporal dynamics (2021-2040, 2041-2060, 2061-2080, 2081-2100) of the BCC-CSM2-MR climate model. The dynamics of the pictures can be viewed at https://denglab.org.cn/index.php/download/.

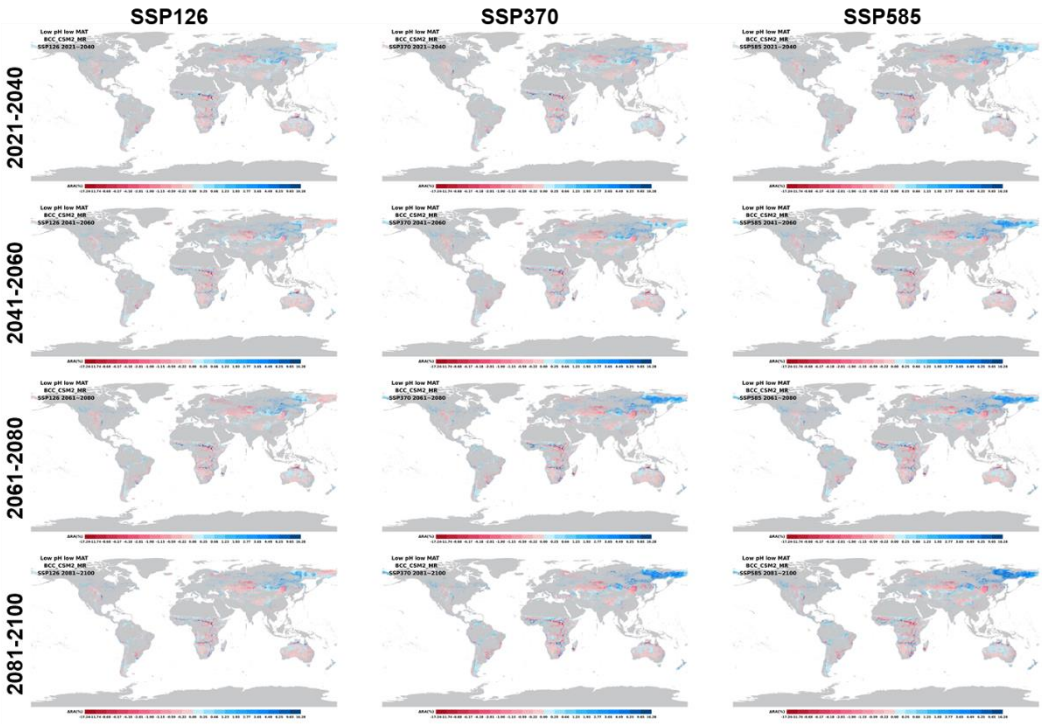


Figure S11 The differences of projected relative abundance to current abundance for the low pH low MAT eco-cluster under future climate scenarios (SSP126, SSP370, and SSP585) during four time periods. The blue and red values indicate increased and decreased relative abundances, respectively.

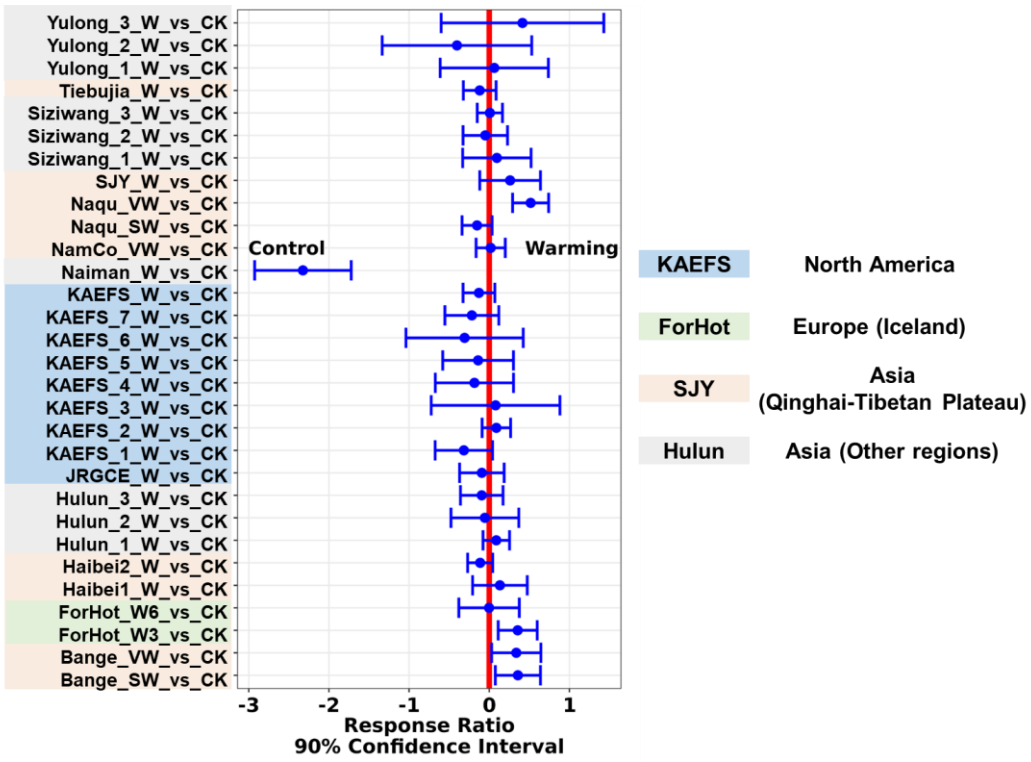


Figure S12 Response ratios for the relative abundance of low pH low MAT eco-cluster under warming condition compared to control condition within ten simulated warming sites. W, warming; CK, non-warming; VW, variable warming; SW, stable warming. The colors indicate different geographic regions.

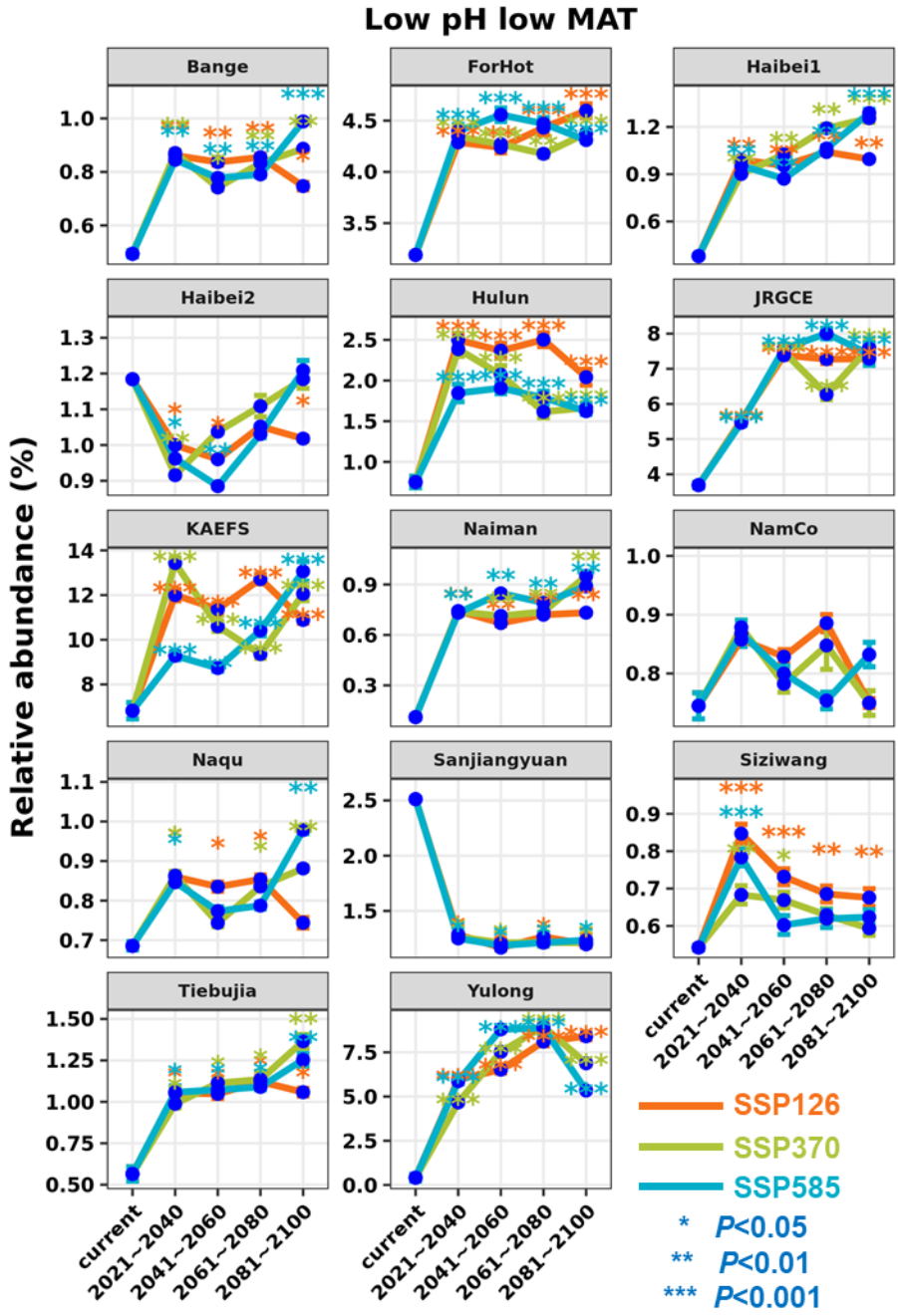


Figure S13 The relative abundance shifts for the low pH low MAT eco-cluster with the predicted values under three future scenarios (SSP126, SSP370, and SSP585) compared to the current values at fourteen simulated warming sites. The significance was tested using a one-sample Student's *t*-test for different levels.

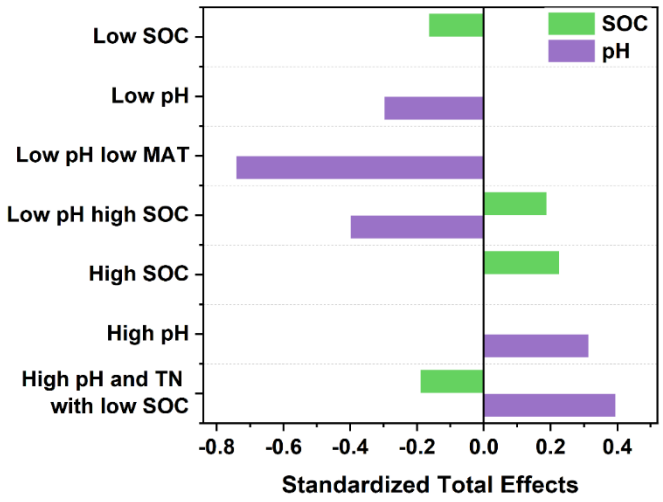


Figure S14 Standardized total effects of the assigned clusters to predict soil pH and SOC features based on piecewise structure equation models.

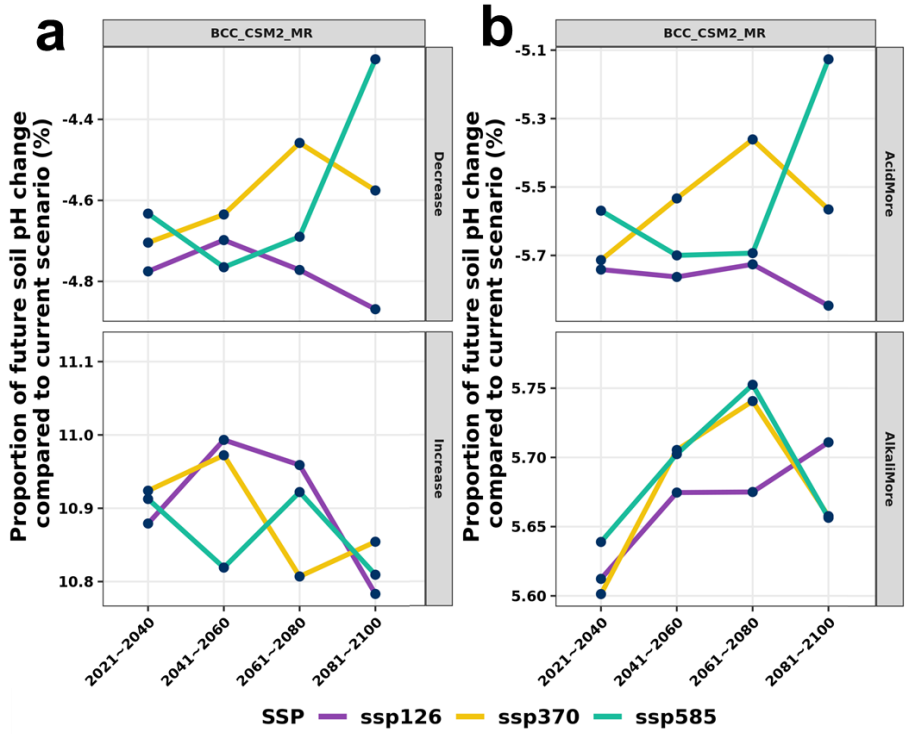


Figure S15 Proportion of average pH change in future under three scenarios (SSP126, SSP370, and SSP585) compared to current scenario in global grassland soils. (a) Extent of pH increase and decrease. (b) Tendency of pH to be acid and alkaline.

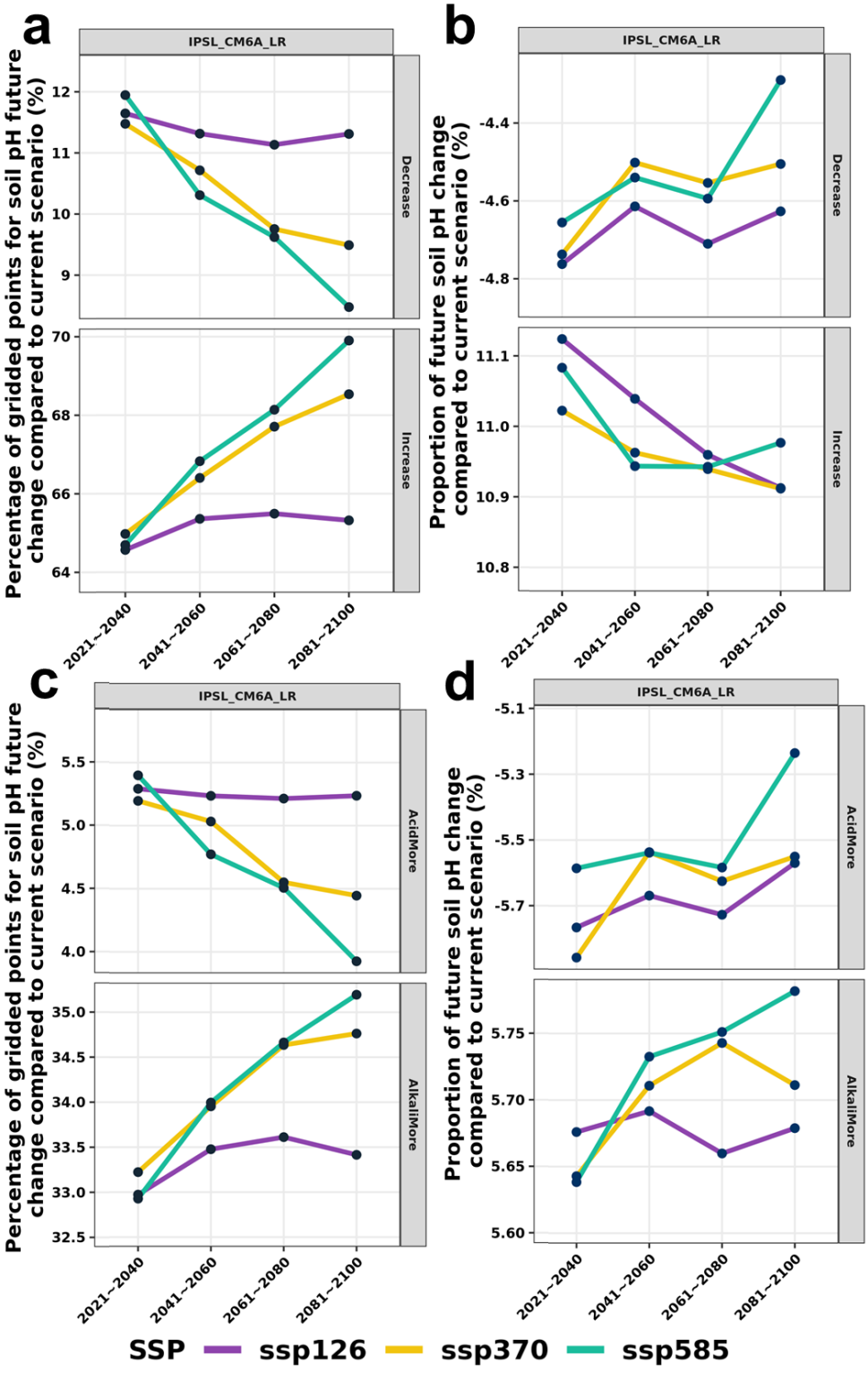


Figure S16 Proportion of gridded grassland areas related to pH change in future under three scenarios (SSP126, SSP370, and SSP585 within the IPSL-CM6A-LR climate model) compared to current scenario. (a) The gridded points related to increasing and decreasing pH. (b) The gridded points related to pH change of acidic and alkaline conditions. (c) Extent of pH increase and decrease. (d) Tendency of pH to be acid and alkaline.

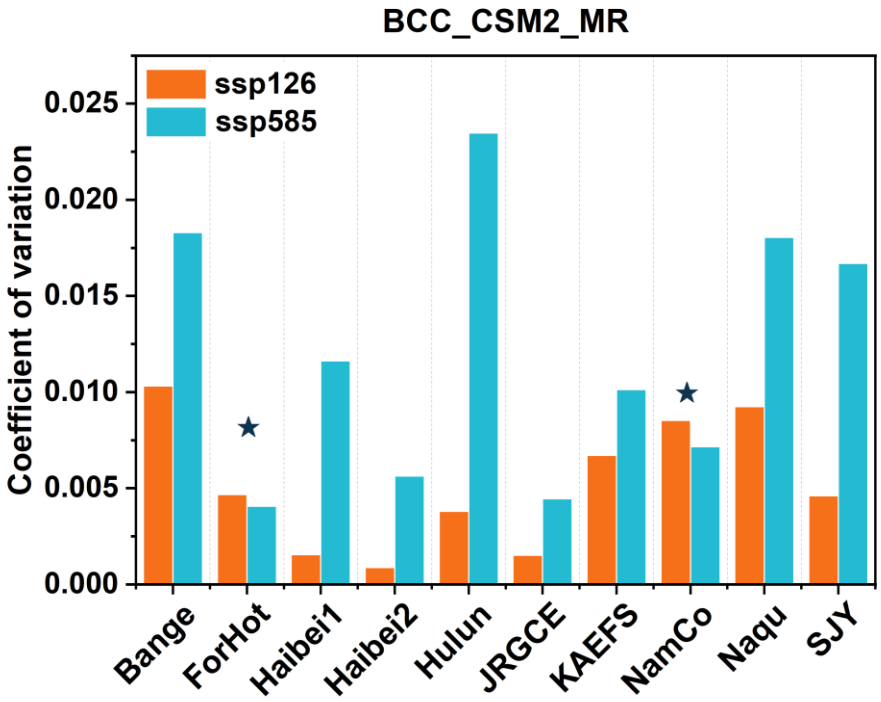


Figure S17 The coefficient of variance (CV) for predicted future pH changes under SSP126 and SSP585 (BCC-CSM2-MR model) in ten simulated warming sites. The stars indicate that the CV values under SSP126 were larger than SSP585.

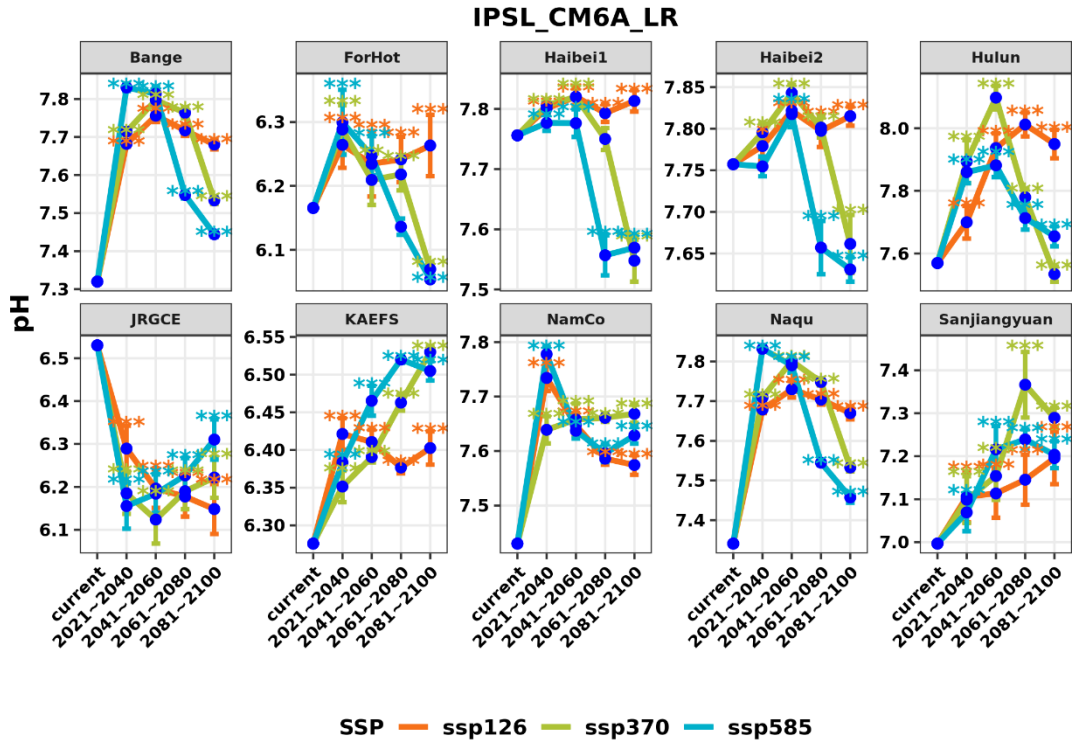


Figure S18 The shifts of edaphic pH with the predicted values under three future scenarios (IPSL-CM6A-LR model) compared to the current values at ten simulated warming sites. The significance was tested using a one-sample Student's *t*-test for different levels (***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$).

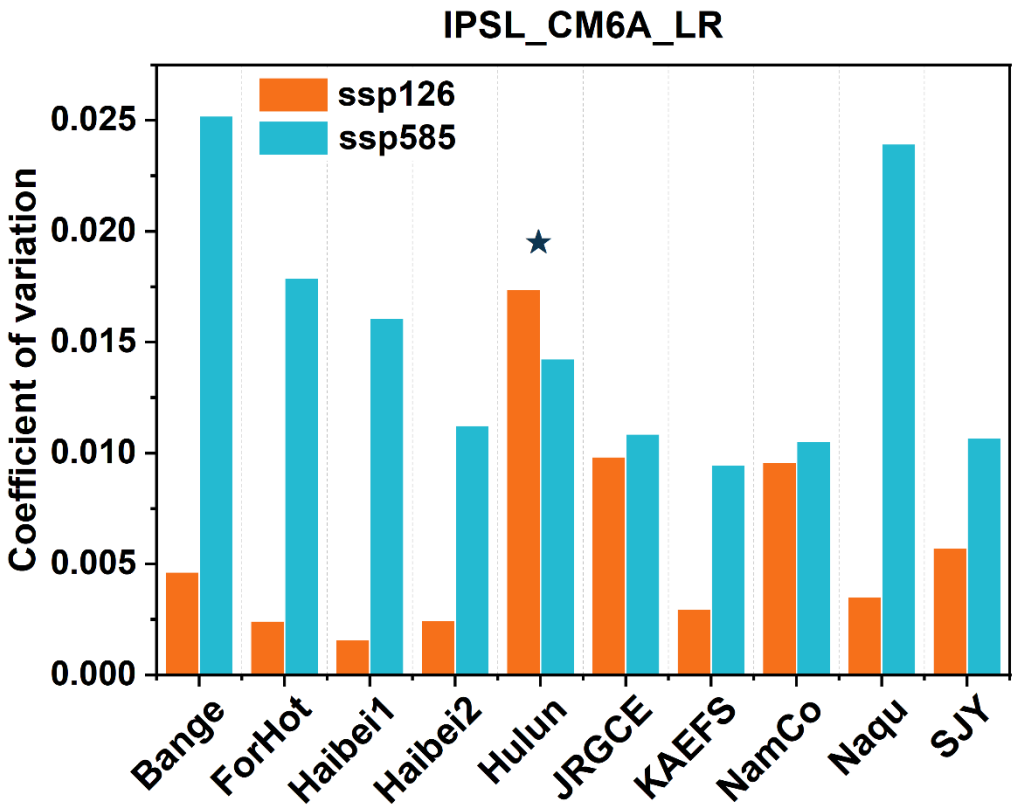


Figure S19 The coefficient of variance (CV) for predicted future pH changes under SSP126 and SSP585 (IPSL-CM6A-LR model) in ten simulated warming sites. The star indicates that the CV value under SSP126 was larger than SSP585.

Table S1. (separate file)
List of collected dataset for soil microbiome in grasslands

Table S2 Mantel test for the relationships between the core community and environmental factors

Envs	Full names	$r_{Bray\text{-}Curtis}$	$r_{Jaccard}$
pH	Soil pH	**0.306*** **	**0.323*** **
SOC	Soil organic carbon	**0.087*** **	0.016
TN	Total nitrogen	**0.096*** **	**0.164*** **
MAT	Mean Annual Temperature	**0.078*** **	-0.015
MDR	Mean Diurnal Range	**0.109*** **	**0.153*** **
Isothermality	Isothermality	-0.043	0.007
TempSeason	Temperature Seasonality	-0.155	-0.153
MaxTWarmM	Max Temperature of Warmest Month	**0.157*** **	**0.071*** **
MinTColdM	Min Temperature of Coldest Month	-0.100	-0.184
TAR	Temperature Annual Range	-0.167	-0.112
MeanTWetQ	Mean Temperature of Wettest Quarter	**0.163*** **	**0.113*** **
MeanTDryQ	Mean Temperature of Driest Quarter	-0.106	-0.176
MeanTWarmQ	Mean Temperature of Warmest Quarter	**0.190*** **	**0.086*** **
MeanTColdQ	Mean Temperature of Coldest Quarter	-0.078	-0.151
MAP	Mean Annual Precipitation	**0.096*** **	**0.083*** **
PWetM	Precipitation of Wettest Month	**0.105*** **	**0.130*** **
PDryM	Precipitation of Driest Month	0.019	0.032
PrecSeason	Precipitation Seasonality	**0.018***	-0.091
PWetQ	Precipitation of Wettest Quarter	**0.076*** **	**0.120*** **
PDryQ	Precipitation of Driest Quarter	**0.035****	0.030
PWarmQ	Precipitation of Warmest Quarter	**0.157*** **	**0.167*** **
PColdQ	Precipitation of Coldest Quarter	-0.122	-0.132
NDVI	Normalized Difference Vegetation Index	**0.174*** **	**0.147*** **

*****, P < 0.001; **, P < 0.01; *, P < 0.05**

Table S3 Cubist model performance summary of 10-fold cross validations for each eco-clusters

	Training dataset		Test dataset		
	Average error	Correlation coefficient (*r*)	RMSE	Correlation coefficient (*r*)	*P* value
Low pH	0.780±0.022	0.786±0.014	1.281±0.275	0.825±0.048	<0.001
High pH and TN with low SOC	0.435±0.005	0.928±0.004	0.649±0.042	0.922±0.016	<0.001
Low SOC	0.195±0.003	0.863±0.005	0.293±0.048	0.862±0.035	<0.001
High SOC	0.222±0.003	0.908±0.004	0.329±0.044	0.900±0.024	<0.001
High PWetM	0.364±0.005	0.871±0.003	0.557±0.038	0.853±0.020	<0.001
Low pH high SOC	0.307±0.004	0.890±0.005	0.469±0.046	0.890±0.026	<0.001
High MaxTWarmM	0.461±0.005	0.807±0.005	0.701±0.097	0.810±0.039	<0.001
Low pH low MAT	1.242±0.018	0.916±0.005	2.098±0.150	0.910±0.024	<0.001
High pH	0.745±0.004	0.909±0.003	1.050±0.100	0.899±0.022	<0.001
High NDVI	0.361±0.008	0.876±0.011	1.120±0.438	0.893±0.076	<0.001

Table S4 Fourteen simulated warming sites across global grasslands

Locations	Full names	Longitude	Latitude	Continents (regions)	Plot information	Temperature increment	Warming Samples	Control Samples
Bange	Bange county, Nagqu Prefecture of the Tibetan Autonomous Region, China	92.03	31.39	Asia (Qinghai-Tibetan Plateau)	Open-top chamber	1.5~2℃	18	9
ForHot	ForHot experiment, Iceland	-21.19	64.00	Europe	Geothermal heat	3℃, 6℃	10	5
Haibei1	Haibei Alpine Meadow Ecosystem Research Station, China	101.20	37.61	Asia (Qinghai-Tibetan Plateau)	Infrared radiator	1.2~1.7℃ (growth season) 1.5~2℃ (non-growth season)	4	4
Haibei2	Haibei Alpine Grassland Ecosystem Research Station, China	101.20	37.50	Asia (Qinghai-Tibetan Plateau)	Infrared heater	2℃	6	6
Hulun	Hulun Lake Reserve of Inner Mongolia, China	113.21	48.75	Asia (Other region)	Open-top chamber	1.8~2.1℃	15	15
JRGCE	Jasper Ridge Global Change Experiment, USA	-122.23	37.40	North America	Infrared heat lamp	1~2℃	8	16
KAEFS	Kessler Atmospheric and Ecological Field Station, USA	-97.52	34.98	North America	Infrared radiator	3℃	28	28
NamCo	Nam Co Station, China	90.99	30.77	Asia (Qinghai-Tibetan Plateau)	Infrared lamp	2℃	4	4
Naqu	Nagqu county, Nagqu Prefecture of the Tibetan Autonomous Region, China	92.02	31.44	Asia (Qinghai-Tibetan Plateau)	Open-top chamber	1.5~2℃	18	9

SJY	Sanjiangyuan Alpine Grassland Ecosystem Field Observation Station, China	97.30	33.41	Asia (Qinghai-Tibetan Plateau)	Open-top chamber	2.00±0.24℃	3	3
Naiman	Naiman Desertification Research Station, China	120.70	42.93	Asia (Other region)	Open-top chamber	1.33-4.8℃	3	3
Yulong	Yulong Snow Mountain, China	100.17	27.00	Asia (Other region)	Open-top chamber	1.15℃, 1.28℃, 1.92℃	17	18
Siziwang	Siziwang Banner County, China	111.88	41.77	Asia (Other region)	Infrared heater	1℃	15	15
Tiebujia	Tiebujia Town of Gonghe County, China	99.58	37.03	Asia (Qinghai-Tibetan Plateau)	Open-top chamber	2℃	3	3

Table S5 Cubist model performance summary of 10-fold cross validations for predicting soil pH with eco-clusters

	Items	Values with Standard Deviation[#]
Training dataset	Average error	0.252±0.003
	Correlation coefficient (***r***)	0.940±0.000
Test dataset	RMSE	0.416±0.043
	Correlation coefficient (***r***)	0.928±0.016
	P value	<0.001

The average values and deviations were obtained by 10-fold cross-validation.

Table S6 Functional potentials of the ten eco-cluster groups predicted by FAPROTAX

Functional potential	Influence on pH	High MaxTW armM	High NDVI	High pH	High pH and TN with low SOC	High PWetM	High SOC	Low pH	Low pH high SOC	Low pH low MAT	Low SOC
Aerobic chemoheterotrophy	Undetermined	*** c	* c	*** c	** c	*** c	NS c	*** c	*** c	*** c	** c
Aromatic compound degradation	Undetermined		NS f					NS f	NS f	NS f	NS f
Chemoheterotrophy	Undetermined	*** c	** c	*** c	*** c	*** c	*** c	*** c	*** c	*** c	*** c
Denitrification	Increase			*** f		* f		NS f	NS f	NS f	
Nitrate denitrification	Increase			*** f		* f		NS f	NS f		
Nitrate reduction	Increase/None	* f	NS f	* c	NS f	** f		NS c	NS f		
Nitrate respiration	Increase			*** f		* f		NS f	NS f		
Nitrite denitrification	Increase			*** f		* f		NS f	NS f		
Nitrite respiration	Increase			*** f		* f		NS f	NS f		
Nitrogen fixation	Undetermined		NS f	NS f		NS f	NS f	NS f	NS f	NS f	
Nitrogen respiration	Increase/None			*** f		* f		NS f	NS f	NS f	
Nitrous oxide denitrification	Increase			*** f		* f		NS f	NS f		
Ureolysis	Increase		NS f	NS f			NS f	NS f			
Anoxygenic photoautotrophy	None			*** f		* f		NS f	NS f		
Anoxygenic photoautotrophy S oxidizing	None			*** f		* f		NS f	NS f		
Photoautotrophy	None			*** f		* f		NS f	NS f		
Photoheterotrophy	Undetermined			*** f		* f		NS f	NS f		
Phototrophy	Undetermined			*** f		* f		NS f	NS f		

The Fisher's exact test (f) or Chi-square with Yates' correction test (c) was used to test whether the assigned function potential to each eco-cluster was significantly different to other functions. The chosen method depended on the expected value for each cell of the contingency table, *i.e.*, Fisher' exact test was used when minimum expected value less than 1, and Chi-square with Yates' correction test for the remaining condition.

***: $P < 0.001$, **: $P < 0.01$, *: $P < 0.05$, NS: non-significance.

Movie S1. (separate file)
Dynamic version of Fig. S10. The predicted biogeographic distributions for assigned eco-clusters of the core microbiota for global grasslands with SSP126, SSP370 and SSP585 scenarios of temporal dynamics (2021-2040, 2041-2060, 2061-2080, 2081-2100) of the BCC-CSM2-MR climate model.

Data S1. (separate file)
Relative abundances, taxonomy, preferences to environmental factors and eco-clusters of core bacterial species of global soil grasslands.