Fossil-fuel-dependent scenarios could lead to a significant decline of global plant-beneficial bacteria abundance in soils by 2100

Pengfa Li[1,2,9], Leho Tedersoo[3,9], Thomas W. Crowther[4], Alex J. Dumbrell[5], Francisco Dini-Andreote[6], Mohammad Bahram[3], Lu Kuang[1], Ting Li[1], Meng Wu[2], Yuji Jiang[2], Lu Luan[2], Muhammad Saleem[7], Franciska T. de Vries[8], Zhongpei Li[2], Baozhan Wang[1,*], Jiandong Jiang[1,*]

[1] Department of Microbiology, College of Life Sciences, Nanjing Agricultural University, Key Laboratory of Agricultural and Environmental Microbiology, Ministry of Agriculture and Rural Affairs, Nanjing, China

[2] State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, China

[3] Mycology and Microbiology Center, University of Tartu, Tartu, Estonia

[4] Institute of Integrative Biology, ETH Zürich, 8092, Zürich, Switzerland

[5] School of Life Sciences, University of Essex, Colchester, Essex, United Kingdom

[6] Department of Plant Science, The Pennsylvania State University, University Park, PA 16802, USA

[7] Department of Biological Sciences, Alabama State University, Montgomery, AL 36104, USA

[8] Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands

[9] These authors contributed equally: Pengfa Li, Leho Tedersoo.

* Corresponding author: Baozhan Wang (bzwang@njau.edu.cn), Jiandong Jiang (jiang_jjd@njau.edu.cn).

Abstract

Exploiting the potential benefits of plant-associated microbes represents a sustainable approach to enhancing crop productivity. Plant-beneficial bacteria (PBB) provide multiple benefits to plants. However, the biogeography and community structure remain largely unknown. Here, we constructed a PBB database to couple microbial taxonomy with their plant-beneficial traits and analyzed the global atlas of potential PBB from 4,245 soil samples. We show that the diversity of PBB peaks in low-latitude regions, following a strong latitudinal diversity gradient. The distribution of potential PBB was primarily governed by environmental filtering, which was mainly determined by local climate. Our projections showed that fossil-fuel-dependent future scenarios would lead to a significant decline of potential PBB by 2100, especially biocontrol agents (-1.03%) and stress resistance bacteria (-0.61%), which may potentially threaten global food production and (agro)ecosystem services.

Introduction

To meet the increasing food demand, chemical fertilizers and pesticides have long been overused in agriculture. This occurs often more pronounced in developing regions, which has resulted in enormous damage to the environment and human health[1, 2]. Thus, meeting the growing demand for food, while reducing environmental impacts, and reversing trends in chemical overuse are predominant global challenges of the 21st century[3]. In this scenario, there has been a sustainable relationship between plant species and beneficial bacteria (i.e., biocontrol, plant growth-promoting and stress resistance bacteria) that have co-evolved over the last ~480M years[4-6]. The nature of this relationship provides valuable avenues for enhancing plant productivity in different agroecosystems around the world.

The plant-beneficial bacteria (PBB) in soils provide multiple benefits to plants and these can be didactically divided into three major categories according to their plant-beneficial traits[7]: (i) biocontrol capacity - the ability to reduce impacts from plant pathogens that would otherwise limit plant development[8, 9]; (ii) plant growth-promoting (PGP) - the ability to fix nitrogen, solubilize phosphorus and potassium, or produce siderophores and phytohormones[10, 11]; (iii) stress resistance provision - the ability to ameliorate plant water stress (e.g. from floods, drought or increased salinity)[12, 13]. Considering the ecological and environmental sustainability of PBB, relative to traditional chemical fertilizers, the application of PBB represents a promising strategy to realize the One Health concept[14]. Thus, developing strategies for the effective use of PBB requires a detailed understanding of the ecological drivers regulating their global distribution.

In recent decades, a growing body of research has explored biogeography of different microbial groups (bacteria and archaea[15-17], fungi[18, 19], and protists[20]), as well as distribution of

different functional groups such as phytopathogens[21] and nitrogen-cycling microbes[22]. However, despite their ecological, economical, and agricultural importance, large-scale distribution patterns of PBB as taken separately have never been examined. This imposes a limitation on our understanding about the role of environmental predictors shaping their diversity and response to global change drivers.

Given the tightly coupled relationship between temperature and biological processes, climate change is ubiquitously altering global biodiversity[23]. For example, climate change is expected to greatly influence the distribution of belowground microorganisms by accelerating species turnover and promoting a higher proportion of soil-borne pathogens[21], leading to increased incidences and severity of the diseases they cause[24]. However, the responses of PBB communities to climate change remain poorly understood, restricting future efforts to meet global food demand by exploiting the plant-beneficial microbiome.

Despite the widespread appreciation for the multiple functions performed by PBB, the vast array of microbial taxa and functions that benefit plant growth and/or promote plant protection remain largely uncharacterized or unidentified. To address this knowledge gap, we conducted a global survey aiming at understanding the biogeography of PBB and the ecological drivers regulating their global distribution, by linking microbial taxonomy to plant-beneficial traits in a comprehensive way. We first constructed a PBB database based on the species identified in documented literature mainly at the genus level, because most functions are conversed at genus level[25-27] (Fig. 1). Based on our PBB database and the microbiome data generated by the Earth Microbiome Project (EMP)[16], we determined the plant-beneficial traits of microbes in 4,245 soils distributed across 7 continents and 9 land cover types (Supplementary Fig. 1). Through a series of

77 theoretical and modeling approaches, we (i) identified the taxonomic composition of potential PBB,
78 (ii) mapped the distribution of potential PBB and revealed their underlying regulating factors, and
79 (iii) predicted changes in their relative abundances under future climate change scenarios.

80 **Results and Discussion**

81 **Global taxonomy of PBB in soils**

82 We constructed a PBB database that links taxa identity to plant-beneficial traits (Fig. 1a,
83 Supplementary Data 1). For this approach, we considered plant-beneficial traits to include
84 measured positive impacts through one of the three mechanisms: biocontrol capacity, plant
85 growth-promoting (PGP), or stress resistance. The PBB database uses the taxonomy information
86 from the Silva database (Silva v.138)[28], which provides the broad taxonomic coverage required. Our
87 main principles for PBB database construction were as follows. First, similar to other taxonomy-
88 based function-annotation databases (e.g., FAPROTAX[25], FUNGuild[29], and FungalTraits[30]), our PBB
89 database operates mainly at the genus level. Second, each taxon must have at least one plant-
90 beneficial trait. Third, the plant-beneficial trait should have been experimentally tested, either *in*
91 *situ* or *ex vivo*. Fourth, as some PBB may be potentially phytopathogenic, PBB taxa also identified
92 as pathogens were removed. For this purpose, we generated a comprehensive list of
93 phytopathogenic bacteria (Supplementary Data 2), which covers almost all known phytopathogens
94 recorded by 2022.

95 Our PBB database comprised 396 bacterial genera from 17 phyla, 27 classes, 76 orders, and
96 135 families (Fig. 1b). Of these, 92 have a potential biocontrol capacity, 368 have PGP functions,
97 and 51 support plant stress resistance. Genera belonging to the phyla Proteobacteria,

Actinobacteria, Firmicutes, and Bacteroidetes account for most PBB in our database (Fig. 2b). In the EMP dataset, after excluding phytopathogens, *Massilia* (1.83%), *Bacillus* (1.49%), *Sphingomonas* (1.43%), *Pseudomonas* (1.26%), *Bryobacter* (1.25%), *Bradyrhizobium* (0.83%), *Flavobacterium* (0.74%), *Arthrobacter* (0.68%), *Sphingobium* (0.65%), *Gemmatimonas* (0.48%) and *Flavisolibacter* (0.44%) were the most abundant PBB genera in soils globally (Fig. 1b). Our PBB database is freely available, and we also provide an R script to use with this database (Supplementary Data 3).

We conducted multiple field experiments to assess the applicability of the PBB database in explaining crop production. The results showed consistently positive correlations between the relative abundance of potential PBB and yield/biomass of maize (Pearson $r = 0.845$, $P < 0.001$), rice (Pearson $r = 0.534$, $P = 0.009$), and peanut (Pearson $r = 0.747$, $P = 0.005$; Supplementary Fig. 2), validating the assumption of potential PBB in effectively promoting the production of various crops.

Global biogeography and diversity of PBB

After excluding the 1327 phytopathogenic bacterial OTUs, our database recognized 13,979 potential plant-beneficial OTUs in the Earth Microbiome Project (EMP) soil samples. Globally, PBB represented 2.35% to 99.85% (mean = 21.54%) of all bacterial 16S rRNA gene sequences. The average relative abundances of biocontrol, PGP bacteria, and stress resistance categories were 10.85%, 21.07%, and 7.11%, respectively (Supplementary Fig. 3). At the continental scale, the relative abundance of potential PBB was highest in Oceania (38.55%) and Europe (29.56%), while both North and South America had below 15% of potential PBB (Fig. 2a). Oceania also occupied highest relative abundances of biocontrol (27.72%), PGP (37.36%) and stress resistance (23.84%) categories (Fig. 2a). With respect to distinct habitats and biomes, freshwater (33.78%) and grassland (33.55%) had the highest total potential PBB relative abundance, while tundra (15.38%) had the

lowest. Grassland (biocontrol: 23.85%; PGP: 33.03%; stress resistance: 19.22%) and tundra (biocontrol: 4.34%; PGP: 15.04%; stress resistance: 2.67%) soils had the highest and lowest relative abundances of all three types of potential plant-beneficial bacteria, respectively (Fig. 2a). In cropland, the relative abundances of all potential PBB, biocontrol, PGP, and stress-resistance bacteria were 22.49%, 12.43%, 21.96%, 10.51%, respectively (Fig. 2a). Globally, we found a weak – albeit significant – relationship between latitude and the relative abundance of potential PBB. This correlation suggests a faint latitudinal gradient in potential PBB relative abundance occurrence in global soils (R^2 = 0.005, P < 0.001; Fig. 2b, Supplementary Fig. 3).

We then investigated the global richness patterns of PBB (defined as the number of observed potential PBB OTUs). At the continental scale, Africa (mean richness = 76.73±17.94), Asia (67.21± 24.38), and North America (58.66±25.12) had the higher PBB richness, while Antarctica (27.75± 9.27) had the lowest (Supplementary Fig. 4). With respect to land cover types, rangeland (76.76± 14.00), grassland (75.67±26.71), freshwater (75.02±16.15), and cropland (67.74±22.19) had higher PBB richness, while desert (35.95±22.48) and tundra (33.28±15.04) had the lowest (Supplementary Fig. 4). We also found higher PBB richness in lower latitude regions, and determined a significant linear relationship between PBB richness and absolute latitude (Pearson's r = -0.320, P < 0.001; Fig. 2c, Supplementary Fig. 4). This supports the existence of a latitudinal diversity gradient (LDG) for potential PBB, which is consistent with the LDG observed for plants, arthropods, vertebrates[1], total fungi[18], and some bacteria[16]. To ensure this finding was robust and not biased by unbalanced sampling, we conducted a random resampling from densely sampled areas, and repeated the resampling for 100 times (Supplementary Fig. 5). The consistently negative latitude-diversity relationship confirmed that a latitudinal diversity gradient (LDG) for potential PBB

is not driven by sampling effects (Supplementary Fig. 5).

Principal coordinates analysis combined with three-way PERMANOVA showed that potential PBB communities were compositionally distinct across continents and land cover types (Fig. 2d, Supplementary Fig. 6). Land cover type was the primary factor explaining the composition of PBB communities (R^2 = 0.191, P < 0.001), followed by sampling region (R^2 = 0.095, P < 0.001) and sampling season (R^2 = 0.012, P < 0.001). In macroecology, the distance-decay relationship (DDR) is used to explain an increase in community dissimilarity as the geographic distance between samples increases [31]. DDR reflects spatial community turnover and can also be used to infer the underlying ecological processes structuring the metacommunity[32]. Our results revealed a strong DDR for potential PBB communities (slope = -6.805, R^2 = 0.444, P < 0.001, Supplementary Fig. 6). This suggests that the potential PBB communities may be strongly structured by environmental filtering and/or dispersal limitation, which can steepen the DDR slope[33].

It is worth noting that the EMP database contains intrinsic sampling biases (as most samples were collected within the Northern Hemisphere; Supplementary Fig. 1). As such, we acknowledge that caution is warranted in interpreting and extending these results to underrepresented locations within the dataset. To account for this possibility, the biogeographical patterns of potential PBB revealed in our main dataset were cross-validated by an independent global dataset[34], for which the soil samples were collected following a standardized protocol[35]. Using this independent dataset, we also observed a LDG (Pearson r = -0.176, P < 0.001) and DDR (Slope = -1.577, P < 0.001) for potential PBB (Supplementary Fig. 7).

Factors affecting the global distribution of PBB

We used multivariate negative-binomial General Linear Models to examine the role of

environmental variables (climate, soil property, and vegetation) and spatial variables (described by PCNMs: principal coordinate of neighbor matrices [36]) in determining the relative abundance and global distribution of potential PBB, and thus the composition of PBB communities. The relative occurrence of the majority of potential PBB (>50%) was better predicted by environmental variables than spatial variables, and analysis of AIC deviations showed this to be consistent for all potential PBB (environment: 57.55%; space: 40.67%) and its categories biocontrol (environment: 54.75%; space: 43.18%), PGP (environment: 57.57%; space: 40.64%), and stress resistance (environment: 53.54%; space: 44.42%) bacteria (Fig. 3a). Moreover, the potential PBB OTUs strongly influenced by environmental factors accounted for >99% of total potential PBB sequences (Supplementary Fig. 8). This also suggests that potential PBB communities are more consistently structured by environmental filtering than dispersal limitation. We tested the correlations between environmental factors and the relative abundances of dominant potential PBB genera. The results revealed that the relative abundance of all 30 dominant genera was significantly correlated with at least 17 environmental variables (Fig. 3b). These results further confirmed a tight relationship between PBB distribution and local environmental variables.

Based on these relationships, we examined the key environmental factors structuring the distribution of potential PBB using a random forest modeling approach. Seven random forest models, which separately or jointly accounted for different environmental variables, were constructed (Model 1: Climate; Model 2: Soil properties; Model 3: Vegetation; Model 4: Climate & Soil properties; Model 5: Climate & Vegetation; Model 6: Soil properties & Vegetation; Model 7: Climate & Soil properties & Vegetation) and compared for model performance (based on R^2). Then, the model performances (R^2) were compared. All models were statistically significant, with Model

1 (climate factors) consistently performing better than Model 2 (soil properties) and Model 3 (vegetation) in predicting the relative abundance of PBB (Fig. 3c). Besides, adding either soil properties (Model 4) or vegetation (Model 5), or both soil properties and vegetation (Model 7; Fig. 3c) to a climate-only model provided only minor improvement to the explained variation in the relative abundance of potential PBB. Across all variables, climate factors explained 59.1%~60.8% of the variation in the relative abundance of global potential PBB, whereas soil properties and vegetation variables explained 26.5%~31.8% and 8.0%~12.6%, respectively (Fig. 3c). Using an independent global dataset, we also demonstrated that the global distribution of PBB was more affected by bioclimatic variables (Supplementary Fig. 7). While soil pH is considered to be a key factor structuring the distribution of soil bacteria[17, 21], we found soil pH to be only a minor factor in explaining the global distribution of potential PBB when compared to climate factors. Moreover, the effect of vegetation was much lower than that of climate factors on potential PBB communities, possibly because local climates are also primary determinants of global plant distribution[37-39].

PBB abundance under future climate change scenarios

The effect of climate change on the distribution of PBB has remained a major uncertainty. Therefore, we modeled the relative abundance of potential PBB in the year 2100 under four future climate scenarios (SSP126: Sustainability; SSP245: Middle of the road; SSP370: Regional rivalry; SSP585: Fossil-fueled development) using twenty-one different CMIP6 downscaled global change models (GCMs) to minimize the deviations derived from different climate models (Fig. 4a). The multivariate environmental similarity surface (MESS) analysis indicated that predicting global distribution of PBB using our dataset is acceptable (Supplementary Fig. 9), and all projections were cross-validated (Fig. 4b, Supplementary Fig. 10). Our projections showed that the relative abundance of all potential

PBB would potentially decrease by 0.07%, 0.24%, 0.40%, and 0.60% under the scenarios SSP126, SSP245, SSP370, and SSP585, respectively, by the end of this century. This suggests that the relative abundance of PBB would be potentially suppressed under non-sustainable development. Under all four climate scenarios, the PBB increase mainly occurs in equatorial and polar regions. However, in mid-latitude regions, especially in Central Asia, Western Asia, Europe, North Africa, Central North America, and Southern South America, the relative abundance of potential PBB would consistently decline by 2100, especially under the fossil-fuel-dependent future scenarios (Fig. 4a).

For PGP bacteria, their relative abundance is expected to increase by 0.16%, 0.14%, 0.12%, and 0.08% under the scenarios SSP126, SSP245, SSP370, and SSP585, respectively (Fig. 4b; Supplementary Fig. 11). Given the enhanced vegetation productivity by CO_2 fertilization effect in the future[40], more available soil nutrients are expected to be required by plants, which would consequently selectively enrich more PGP bacteria to meet the plant's nutrient requirements[41]. Therefore, the increased PGP bacteria may represent one positive feedback on the CO_2 fertilization effect. In contrast to PGP, the biocontrol and stress resistance categories are expected to consistently decline in the future. The relative abundance of biocontrol is expected to decrease by 0.31%, 0.54%, 0.80%, and 1.03% under SSP126, SSP245, SSP370, and SSP585, respectively; the relative abundance of stress resistance is predicted to decrease by 0.17%, 0.32%, 0.49%, and 0.61% under SSP126, SSP245, SSP370, and SSP585, respectively (Fig. 4b; Supplementary Fig. 11). We initially expected that the proportion of biocontrol agents would increase to antagonize soil phytopathogens, as these represent a group of soilborne taxa predicted to increase by ~1%-2.5% in the future[21]. On the contrary, the expected decline in biocontrol bacteria in our projection implied a potential dysbiosis between phytopathogens and non-phytopathogenic soil taxa, which may

potentially threaten global food production and (agro)ecosystem services.

We also determined the relative area that may be impacted by the decline of PBB relative abundance. This was carried out by calculating the deviated abundance in each grid cell. Our results showed that > 50% of global regions may encounter a decline of potential PBB under all future climate scenarios (Fig. 4c). In particular, > 80% of global regions may encounter a decline of biocontrol bacteria and stress resistance bacteria in the future (Fig. 4c). Moreover, we observed a significant negative correlation between change in stress resistance bacteria and future probability of climate extremes[42] at the global scale (r = -0.294, P < 0.001; Supplementary Fig. 12). This implies that the increased climate extremes may potentially lead to a decrease in stress resistance bacteria in the future. It is important to note that these projections are based on currently available observational data alone, with no mechanistic inference to drive these trends, and that these projections need to be properly experimentally tested.

Limitations, conclusions, and future perspectives

This study provides novel insights into the biogeography, and factors regulating the structure, of communities of plant-beneficial bacteria. However, our PBB database is not yet absolutely comprehensive nor fully resolved to the species level, which may obscure some of the overall PBB abundances. Yet, as more microbes are cultured in the future and their functions are resolved, underrepresented guilds in our database can be further improved. Nevertheless, our characterization of PBB biogeography and the ecological relationships driving their distributions, provides the bases for a robust initial first-order approximation of their underlying ecology and susceptibility to climate change. However, we acknowledge that our predictions of PBB response to climate change do not consider direct or indirect (via host plants) CO_2 fertilization effects, as the

relationship between CO_2 concentration and potential PBB remains largely unknown. Moreover, our projections are founded on a permanent relationship between climate and PBB relative abundance, and the projections may need to be amended if the form of these relationships also changes under future climate scenarios. Our predictions may also deviate from observed changes if major and unforeseen alterations to land use or vegetation types occur in the future.

Applying microbiome management practices, including synthetic microbiomes, to crop production is a promising approach to realizing global food security and sustainable agriculture. By providing a full list of native plant-beneficial microbial candidates in different regions across the globe, our PBB database would contribute to designing the effective synthesis of beneficial native microbiomes. However, we acknowledge that caution is required when directly applying PBB species to fields, since such active management could also potentially raise the risk of introducing unwanted invasive microbes[43-45]. Our study highlights the importance of climate factors in regulating the global biogeography of PBB communities. Moreover, our model projections indicated that non-sustainable development may suppress potential PBB abundance, and more regions would encounter a greater PBB decline under fossil-fuel dependent future scenarios than under sustainable development climate scenarios. These changes are likely to have direct consequences for the productivity and sustainability of managed and natural ecosystems, with direct implications for food production. Our research suggests that sustainable development is highly required to maintain 'stable' PBB abundances, shedding light on optimized sustainable development solutions that promote crop production from plant-beneficial soil microbes.

Methods

Construction of the plant-beneficial bacteria database

Given the wide taxonomic coverage of the Silva database, its taxonomic information (Silva v.138) was extracted at the genus level because most functions are conserved at this level[25-27]. We then classified each genus into one or more plant-beneficial categories based on the current literature to construct the plant-beneficial bacteria database (Fig. 1). A genus was associated with a particular function if all cultured species within the genus have been shown to exhibit that function. The genus must have at least one plant-beneficial trait, either biocontrol potential, plant growth-promoting (PGP), or stress-resistance. The biocontrol potential includes antagonism to either phytopathogenic bacteria, fungi, or nematodes. That is, only those able to directly antagonize phytopathogens were considered in this database. Those that improve disease resistance through indirect pathways, such as activating induced systemic resistance (ISR), were not included. The PGP traits include either nitrogen fixation, phosphorus solubilization, potassium solubilization, zinc solubilization, siderophores production, or production of phytohormones including indole acetic acid (IAA) and acetyl-CoA carboxylase (ACC). Other conditionally plant-beneficial traits, such as nitrification, were discarded because they may be beneficial to only some specific plant species. The stress resistance traits include functions that assist plants to couple with either drought, flood, or changes in soil salinity. Other plant-beneficial traits were not considered in the current PBB database.

It should be noted that all plant-beneficial traits were experimentally tested, either *in situ* or *ex vivo*. That is, we removed plant-beneficial traits if that trait was solely derived from genome prediction or correlation-based analysis. For example, one may find a significant correlation between plant growth and some microbial taxa in field experiments. These taxa would not enter

our database if their PGP traits were not experimentally tested using pure isolates. Therefore, all uncultured taxa were removed from our database. Overall, our PBB database has three plant-beneficial-trait levels (Supplementary Data 1). Level 1 simply defines whether the bacteria are potentially plant-beneficial or plant-harmful (if some species of a genus are also phytopathogens); Level 2 defines if the bacteria are potential phytopathogens, or biocontrol agents, or PGP bacteria, or stress resistance bacteria; Level 3 gives detailed plant-beneficial traits such as nitrogen fixation, phytohormone production, and assist in plant salt tolerance. It should be noted that most stress resistance taxa are halotolerant and anti-drought bacteria in our database due to relatively less experimental evidence on other stresses such as flooding.

As some PBB may be potentially phytopathogenic, we decided to remove these phytopathogenic bacteria after annotating a species-abundance table using the PBB database. We, therefore, also integrated a very comprehensive list of phytopathogenic bacteria (Supplementary Data 2), which covers almost all known phytopathogens recorded by 2022. This list was used to conduct the exclusion procedure. This list of phytopathogenic bacteria contains 57 bacterial genera and 258 species (valid nomenclature under the ICNP), and strain-specific variations within species were ignored. It is worth noting that while the list of phytopathogens may look longer in some other studies, these also include many redundant species that have different nomenclatures[46]. For example, *Clavibacter michiganensis* (valid nomenclature under the ICNP) and *Corynebacterium michiganensis* represent the same species[47]. Besides our full list of phytopathogens, a multiple bacterial pathogen detection pipeline has been recently developed[46], which can detect the phytopathogens based on the 16S rRNA sequences in amplicon sequencing data.

While the risk of false generalizations was minimized via extensive manual investigation of

available literature, we point out that as more microbes are cultured in the future some of these generalizations may turn out to be false. In addition, there may be some missing traits during our manual investigation, and our database can be further extended after adding the missing and newly demonstrated plant-beneficial traits in the future. We note that the use of our PBB database to address further more specific questions may require careful manual refinement of the functional annotations; for example, based on expert knowledge of the system examined. The PBB database is written in a human-readable format that allows for easy modification and extension by any user. Our PBB database is freely available, and we also provide an R script to use with this database (Supplementary Data 3).

Processing the data from field experiments

Field experiment settings and sampling. We conducted multiple field experiments to assess the applicability of the PBB database in positively affecting crop production. Three crops, namely maize (Zea mays L.), rice (Oryza sativa L.) and peanut (Arachis hypogaea L.) were cultivated in different fields. Maize samples were collected from eleven field plots in 2019, and each plot is 20 m length × 5 m width. The fields were located in Ecological Experimental Station of Red Soil at the Chinese Academy of Sciences in Yujiang, Jiangxi province, China (28°13′ N, 116°55′ E). The maize cultivar is Suyu 24. Rice samples were collected from twenty-three field plots in 2022, and each plot is 10 m length × 7 m width. The fields were located in Jiangning, Jiangsu Province, China (32°01′ N, 119°09′ E). The rice cultivar is Nangeng 46. Peanut samples were collected from twelve fields in 2020, and each plot is 6 m length × 4 m width. The fields were located in Comprehensive Experimental Station of red soil in Dongxiang County, Jiangxi Province, China (28°10′ N, 106°35′

E). The peanut cultivar is Ganhua 1.

All crops were harvested at maturity stage. For maize samples, 5 maize plants in each filed plot were randomly selected; for rice samples, 3 rice plants in each plot were randomly selected; for peanut samples, 10 peanut plants in each plot were randomly selected. The selected plants were carefully removed from each plot using a spade, after which the soil attached to the roots was collected and pooled to represent a composite rhizosphere soil sample. The collected rhizosphere soils were stored at −20 °C until DNA extraction and analyses. The maize grains were collected after harvesting maize plants, and the maize yield of each field plot was then calculated after drying the grains; the peanut pods were collected after harvesting peanut plants, and the peanut yield of each plot was calculated after drying the peanut pods. However, we did not collect rice grains after harvesting the rice plants. Instead, only biomass was determined for rice plants. Given the tight correlation between rice plant biomass and rice yield, we used plant biomass to indicate rice production in this study.

Soil DNA extraction. Soil DNA was extracted from 0.5 g of soil (fresh weight) using a Fast®DNA SPIN Kit (MP Biomedicals, CA, USA) and then subsequently purified using a PowerClean® DNA Clean-up Kit (MoBio, CA, USA) according to the manufacturer's instructions. The concentration and quality of the extracted DNA were measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, DE, USA).

Amplicon high-throughput sequencing and data processing. The PCR amplification of the DNA samples was conducted for bacterial community analysis using 519F and 907R primers[48]. We performed high-throughput sequencing using the Illumina MiSeq sequencing platform (Illumina Inc., CA, USA). The raw sequence data were analyzed using the QIIME 2 (version 2021.8)[49]. Raw

sequence data were demultiplexed and quality filtered using the q2-demux plugin followed by denoising with DADA2 (via q2-dada2)[50], and the sequences that were not present in at least 2 samples were filtered out. After quality filtering and the removal of chimaeras, sequences were clustered into ASVs after rarefying sequences to even sequencing depth (based on the sample with the minimum numbers of reads)[51]. Subsequently, taxonomic classification was conducted using plugin *feature-classifier classify-sklearn* by searching against database Silva 138 SSURef NR99 full-length taxonomy[28]. Finally, we annotated the generated ASV tables using our PBB database to extract the PBB ASVs, and then plotted the relative abundance of PBB against crop yield.

Processing of the Earth Microbiome Project data

The microbial abundance table used in the present study was derived from the Silva-based rarefied table generated by the Earth Microbiome Project (EMP)[16]. The EMP employed a unified standard workflow for sample metadata curation, DNA extraction, sequencing, and sequence preprocessing, to avoid known issues in combining multiple amplicons across diverse environments. A total of 4,245 soil samples that have geographical information (latitude and longitude) were extracted from the raw dataset. These 4,245 soil samples were collected across 8 continents and 9 land cover types. Each sample contains 10,000 rarefied high-quality sequences. Based on the taxonomic information, we annotated each OTU (operational taxonomic unit) using our PBB database. Of all 63,094 OTUs in the raw microbial abundance table, 13,979 OTUs were successfully annotated with the PBB database, 6,381 OTUs were classified as biocontrol bacteria, 13,568 OTUs were classified as PGP bacteria, and 4,598 OTUs were classified as stress resistance bacteria. Using these newly generated abundance tables, the biogeographical pattern, driving forces and future changes of PBB

communities were subsequently analyzed.

To infer the global alpha diversity of potential PBB, we also used a 90-bp Deblur BIOM table in EMP, which was generated using the non-reference framework. This table was based on the sequence data from the EMP database after filtering errors and trimming to 90 bp (the length of the shortest sequencing run) using Deblur in Qiime2. This abundance table was filtered to keep tag sequences with at least 25 reads total over all samples. We did not directly use the ready-made rarefied OTU table because the rarefaction procedure was conducted based on all >20,000 samples. Instead, after extracting these 4,245 samples from the raw table (File 'emp_deblur_90bp.qc_filtered.biom' in EMP release), we resampled these samples into the same sequencing depth (6,160 tag sequences per sample). The taxonomic information of 90-bp representative sequences was assigned based on the Silva v.138 database[28].

Climate, vegetation, and soil property data

Nineteen bioclimatic variables for each sample location were extracted from WorldClim2 (https://www.worldclim.org/)[52]. The historical climate data represent the average for the years 1970-2000 and comprise 19 variables, 11 of which are temperature-related, and 8 of which are precipitation-related (for detailed information see Supplementary Table 1; https://www.worldclim.org/data/bioclim.html). The future climate data (2080-2100) are CMIP6 (Coupled Model Intercomparison Project 6, https://esgf-node.llnl.gov/projects/cmip6/) downscaled future climate projections. Monthly values of minimum temperature, maximum temperature, and precipitation were processed for four Shared Socio-economic Pathways (SSP): 126, 245, 370, and 585 (SSP126: sustainability; SSP245: middle of the road; SSP370: regional rivalry; SSP585: fossil-fueled development). The full explanation of different SSP scenarios is available

(https://www.carbonbrief.org/explainer-how-shared-socioeconomic-pathways-explore-future-climate-change). The climate data under different SSP scenarios were separately predicted using twenty-one CMIP6 downscaled global change models (Supplementary Table 2). The vegetation variables are indicated by gross primary production (GPP). The GPP data used in this study were the annual average GPP data during the last four decades derived from satellite near-infrared reflectance data[53]. Soil property data including pH, soil organic carbon (SOC), cation exchange capacity (CEC), soil salinity (indicated by electroconductibility), and base saturation were derived from Harmonized World Soil Database (HWSD v1.2, https://www.fao.org/soils-portal/soil-survey) at a resolution of 250 m.

Statistical analysis

Biogeographical pattern analysis. The richness (defined as the number of observed potential PBB OTUs in this study) was plotted against the absolute latitude to investigate whether the alpha-diversity of PBB followed a latitudinal diversity gradient (LDG). Given the potential unbalanced sampling effect in EMP, we conducted random resampling from densely sampled areas. Briefly, we randomly selected 50 or 100 samples if the samples were >50 or >100 within 5 degrees in latitude., and repeated the resampling for 100 times (Supplementary Fig. 5). Bray–Curtis distances were calculated to quantify taxonomic β-diversity. Three-way PERMANOVA was conducted to compare the effects of continent, land cover type and sampling seasonality on the composition of PBB communities. The Bray–Curtis distances were plotted against the log-transformed geographical distances [log(distance+1)] to determine whether the composition of PBB communities followed a distance-decay relationship (DDR).

Multivariate negative binomial General Linear Models. Multivariate negative binomial General

Linear Models[54] were used to disentangle whether the community composition of PBB communities was more strongly controlled by environmental or spatial factors. We fitted the relative abundance of each potential PBB OTU to environmental variables and spatial variables, respectively. The environment model contains all 19 bioclimatic variables, 5 soil properties, and gross primary production data. Spatial variables were derived from the principal coordinates of neighbor matrices (PCNM) algorithm[36], which was able to deconvolute total spatial variation into a discrete set of explanatory spatial scales[32]. The fit of environment and space models was compared using OTU-specific AIC scores. A model was considered to have support over the other model if the difference in AIC (ΔAIC) was > 2[55, 56].

Random forest model. We applied a machine-learning model, random forest, to quantitatively examine the key environmental variables influencing the relative abundance of potential PBB using the *randomForest* R package[57]. Seven random forest models were constructed. Climatic (indicated by 11 temperature-related and 8 precipitation-related bioclimatic variables), soil property (indicated by pH, SOC, CEC, soil salinity, and base saturation), and vegetation (indicated by gross primary production) variables were separately or jointly considered in these seven random forest models (Model 1: Climate; Model 2: Soil properties; Model 3: Vegetation; Model 4: Climate & Soil properties; Model 5: Climate & Vegetation; Model 6: Soil properties & Vegetation; Model 7: Climate & Soil properties & Vegetation). To reduce collinearity among predictors, we reduced the initial set of 25 environmental variables to 14 variables with a variation inflation factor (VIF) below 10. This final set included eight bioclimatic variables (BIO2: Mean diurnal range; BIO3: Isothermality; BIO8: Mean temperature of the wettest quarter; BIO9: Mean temperature of the driest quarter; BIO14: Precipitation of driest month; BIO15: Precipitation seasonality; BIO18: Precipitation of warmest

quarter; BIO19: Precipitation of coldest quarter), five soil variables (pH, SOC, CEC, electroconductibility, and base saturation), and one vegetation variable (gross primary production). A total of 500 trees were fitted in each model. Each tree was fitted based on a random sample of two-thirds of the observations ("in-bag"), and each tree split was based on a different random subset of one-third of the predictors, while the results were cross-validated against the remaining observations ("out-of-bag"), which is in line with standard protocols[57]. The model performance was assessed based on model R^2 using *rfUtilities* R package with 999 permutations. To express variable importance across all modeled ppSHs, the relative importance of each predictor was calculated as a sum of the predictor relative importance of all Random Forests for potential PBB richness/relative abundance weighted by Random Forest predictive ability (out-of-bag R^2)[18].

Future relative abundance projection. The global pattern of relative abundance of potential PBB under the current climate was estimated using GLMs. A multivariate environmental similarity surface (MESS) analysis was conducted to assess extrapolation reliability of PBB using the variables selected from the GLMs[58]. The GLMs were also cross-validated by common Pearson correlation test using 2/3 samples as a model training dataset and 1/3 as a validation dataset. Using the model constructed based on the current climate data, the global patterns of relative abundance under different future climate scenarios were then estimated based on the model parameters. We predicted the future relative abundance of potential PBB under different climate scenarios using the climate data derived from the above twenty-one different CMIP6 downscaled GCMs, and the relative changes were averaged. The projections were conducted using the formula listed in Supplementary Table 3.

Data availability

All raw data used in the current study including the plant-beneficial bacteria database, sample metadata, climate data and species-abundance dataset are publicly available in Figshare (https://doi.org/10.6084/m9.figshare.22274866). The taxonomy information of bacteria is available in Silva database (https://www.arb-silva.de/). The current and future climate data are available in WorldClim2 (https://www.worldclim.org/). The soil property data are available in Harmonized World Soil Database (https://www.fao.org/soils-portal/soil-survey). Source data are provided with this paper.

Code availability

Most numerical analyses included in this article do not have an associated code. Used codes are available in Figshare (https://doi.org/10.6084/m9.figshare.22274866).

Acknowledgements

This study was supported by the National Key R&D Program of China (2022YFA0912501 to J.J.), Fundamental Research Funds for the Central Universities (KYZZ2023003 to J.J.; KYQN2023027 to P.L.; XUEKEN2022003 to B.W.), National Natural Science Foundation of China (42207349 to P.L.; 41977056 to B.W., and 42107336 to L.L.), Jiangsu Funding Program for Excellent Postdoctoral Talent (2022ZB331 to P.L.), Natural Science Foundation of Jiangsu Province (BK20221005 to P.L.), China Postdoctoral Science Foundation (2022M711653 to P.L.), and grants from D.O.B. Ecology and the Bernina foundation to T.W.C.

Author Contributions Statement

J.J. and B.W. designed the framework. P.L., M.W., Y.J., L.L. and Z.L. contributed the sample collecting. P.L., L.K., T.L. and M.B. performed the data analysis. P.L., L.T., T.W.C., A.J.D., F.D., M.B., L.L., M.S., F.T.V. and J.J. wrote the paper. All authors discussed the results and commented on the manuscript.

Competing Interests Statement

The authors declare no competing interests.

Figure captions

Fig. 1 | Taxonomy of plant-beneficial bacteria in global soils. a, Workflow used to construct the PBB database. Briefly, all taxa in the Silva v.138 database were checked to determine whether they have literature-documented and experimentally confirmed plant-beneficial traits. This yielded a three-level PBB database consisting of 396 bacterial genera. Besides, a comprehensive list of phytopathogens was generated to be used as a reference (PBB also identified as pathogens were removed from the final database). **b,** Taxonomy information of potential PBB. In the left panel, each circle represents a PBB genus, and the circle size is proportional to the mean relative abundance in global soils. The top right panel shows the composition of PBB at different taxonomic levels. The bottom right panel shows the top ten PBB genera, and the rectangle area is proportional to the mean relative abundance in global soils. PGP: Plant growth-promoting.

Fig. 2 | Global biogeographical distribution of plant-beneficial bacteria. a, Average relative abundance of different categories of potential PBB in different continents and land cover types. Purple, orange and blue lines within the bars represent the relative abundance of PGP, biocontrol and stress resistance bacteria, respectively. Data are presented as mean values ± SEM. n

represents the number of samples. **b,** The relationship between absolute latitude and relative abundance of potential PBB. The line shows the second-order polynomial fit based on ordinary least squares regression, and shaded areas represent the 95% confidence intervals. The analysis was based on one-side *F* and two-side *t* tests (model parameters and *P* values are reported as inset panels). n represents the number of samples. **c,** The relationship between absolute latitude and number of observed potential PBB OTUs. The significant negative latitude-richness relationship supports a latitudinal diversity gradient (LDG) of potential PBB. Lines represent the fit of the least squares regressions and shaded areas represent the 95% confidence intervals. The analysis was based on one-side *F* and two-side *t* tests (model parameters and *P* values are reported as inset panels). n represents the number of samples. **d,** Principal coordinates analysis (PCoA) of potential PBB communities based on Bray-Curtis dissimilarity. The effects of land cover type, continent and sampling season on potential PBB communities were assessed by three-way PERMANOVA based on Bray-Curtis dissimilarity. Samples are colored by land cover type (left panel) or continent (right panel). PGP: Plant growth-promoting.

Fig. 3 | Factors affecting the global distribution of plant-beneficial bacteria. a, The effects of environmental and geographic distance on PBB community composition were examined via multivariate negative binomial General Linear Models. The relative abundance of each potential PBB OTU was modeled as a function of either environment (climate, soil property, and vegetation variables) or space (using principal coordinate of neighbor matrices (PCNMs) based on Moran's eigenvector maps). The AIC scores of the space-only and environment-only models of each potential PBB OTU were compared. A lower AIC score represents a superior fit (ΔAIC >2). Pink and blue points represent OTUs that are better explained by environment or space models, respectively.

Green points and the black line represent equal support for the environment-only or space-only models for a given potential PBB OTU based on AIC scores (ΔAIC < 2). Within each plot, the pie chart summarizes the proportion of OTUs that are best supported by either the environment-only or space-only models. **b,** Correlations between environmental variables and the relative abundances of the top 30 dominant potential PBB genera. The correlation coefficient and significance were determined by Spearman test. We applied one-side *F* and two-side *t* tests, and then calculated *P* values. **c,** Random Forest model of key environmental factors structuring the potential PBB communities. The left panel shows the performance (R^2) of different models. Model 1: Climate; Model 2: Soil properties; Model 3: Vegetation; Model 4: Climate & Soil properties; Model 5: Climate & Vegetation; Model 6: Soil properties & Vegetation; Model 7: Climate & Soil properties & Vegetation. The right panel shows the contribution of climatic, soil properties, and vegetation variables to the explained variation based on each PBB category. Each tree was fitted based on a random sample of two-thirds of the observations ("in-bag"), and each tree split was based on a different random subset of one-third of the predictors, while the results were cross-validated against the remaining observations ("out-of-bag"), which is in line with standard protocols. The model performance was assessed based on model R^2 with 999 permutations. Vegetation is indicated by gross primary production. PGP: Plant growth-promoting.

Fig. 4 | Predicted future changes in plant-beneficial bacteria. a, Predicted change in relative abundance of potential PBB under future climate-change scenarios. A relative abundance-climate model was constructed by GLMs using relative abundance of potential PBB and 19 climate variables. This model was used to predict the future relative abundance of potential PBB under four different climate scenarios. All climate variables were derived from WorldClim2 using a 5 min (~10 km)

552 resolution. The future climate data were derived from twenty-one different CMIP6 downscaled
553 global change models (GCMs; See detailed information in Methods). The relative change in the
554 relative abundance of potential PBB under different GCMs compared to current climate conditions
555 was averaged. The right panel shows the latitudinal change in relative abundance of potential PBB
556 under four future climate scenarios. The plot axis labels reflect the shared socioeconomic pathway
557 (SSP), sustainability (SSP126), middle of the road (SSP245), regional rivalry (SSP370), and fossil-
558 fueled development (SSP585) scenarios. **b,** Predicted changes in relative abundance of different
559 categories of potential PBB under future climate-change scenarios. Box plots indicate the median
560 (middle line) with 25th, and 75th percentile (box), and 5th and 95th percentile (whiskers). n = 21 for
561 SSP126; n = 20 for SSP245, and SSP585; n = 19 for SSP370. **c,** The relative area that may be
562 impacted by a decline in relative abundance of potential PBB under different future climate
563 scenarios. We calculated the number of declined grid cells and divided it by the total number of
564 grid cells to determine the relative decline area. PGP: Plant growth-promoting.

565 **References**

566 1. Lam, S.K. et al. Next-generation enhanced-efficiency fertilizers for sustained food security. *Nat.*
567 *Food* **3**, 575-580 (2022).

568 2. Lv, C.D. et al. Selective electrocatalytic synthesis of urea with nitrate and carbon dioxide. *Nat.*
569 *Sustain.* **4**, 868-876 (2021).

570 3. Wu, Y.Y. et al. Policy distortions, farm size, and the overuse of agricultural chemicals in China.
571 *P. Natl. Acad. Sci. USA* **115**, 7010-7015 (2018).

4. Zamioudis, C. & Pieterse, C.M.J. Modulation of host immunity by beneficial microbes. *Mol. Plant Microbe In.* **25**, 139-150 (2012).

5. Busby, P.E. et al. Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol.* **15**, e2001793 (2017).

6. Haney, C.H., Samuel, B.S., Bush, J. & Ausubel, F.M. Associations with rhizosphere bacteria can confer an adaptive advantage to plants. *Nat. Plants* **1**, 1-9 (2015).

7. Hunter, P. Plant microbiomes and sustainable agriculture: Deciphering the plant microbiome and its role in nutrient supply and plant immunity has great potential to reduce the use of fertilizers and biocides in agriculture. *EMBO Rep.* **17**, 1696-1699 (2016).

8. Mendes, R. et al. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* **332**, 1097-1100 (2011).

9. Berendsen, R.L., Pieterse, C.M. & Bakker, P.A. The rhizosphere microbiome and plant health. *Trends Plant Sci.* **17**, 478-486 (2012).

10. Lugtenberg, B. & Kamilova, F. Plant-growth-promoting rhizobacteria. *Annu. Rev. Microbiol.* **63**, 541-556 (2009).

11. French, E., Kaplan, I., Lyer-Pascuzzi, A., Nakatsu, C.H. & Enders, L. Emerging strategies for precision microbiome management in diverse agroecosystems. *Nat. Plants* **7**, 256-267 (2021).

12. Porter, S.S. et al. Beneficial microbes ameliorate abiotic and biotic sources of stress on plants. *Funct. Ecol.* **34**, 2075-2086 (2020).

13. de Vries, F.T., Griffiths, R.I., Knight, C.G., Nicolitch, O. & Williams, A. Harnessing rhizosphere microbiomes for drought-resilient crop production. *Science* **368**, 270-274 (2020).

14. Banerjee, S. & van der Heijden, M.G.A. Soil microbiomes and one health. *Nat. Rev. Microbiol.* **21**, 6-20 (2022).

15. Delgado-Baquerizo, M. et al. A global atlas of the dominant bacteria found in soil. *Science* **359**, 320-325 (2018).

16. Thompson, L.R. et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457-463 (2017).

17. Bahram, M. et al. Structure and function of the global topsoil microbiome. *Nature* **560**, 233-237 (2018).

18. Vetrovsky, T. et al. A meta-analysis of global fungal distribution reveals climate-driven patterns. *Nat. Commun.* **10**, 5142 (2019).

19. Tedersoo, L. et al. Global diversity and geography of soil fungi. *Science* **346**, 1256688 (2014).

20. Oliverio, A.M. et al. The global-scale distributions of soil protists and their contributions to belowground systems. *Sci. Adv.* **6**, eaax8787 (2020).

21. Delgado-Baquerizo, M. et al. The proportion of soil-borne pathogens increases with warming at the global scale. *Nat. Clim. Change* **10**, 550-554 (2020).

22. Nelson, M.B., Martiny, A.C. & Martiny, J.B.H. Global biogeography of microbial nitrogen-cycling

traits in soil. *P. Natl. Acad. Sci. USA* **113**, 8033-8040 (2016).

23. Stehr, N. This changes everything: Capitalism vs the climate. *Nature* **513**, 312-312 (2014).

24. Wang, C.Z. et al. Occurrence of crop pests and diseases has largely increased in China since 1970. *Nat. Food* **3**, 57-65 (2022).

25. Louca, S., Parfrey, L.W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**, 1272-1277 (2016).

26. Martiny, J.B.H., Jones, S.E., Lennon, J.T. & Martiny, A.C. Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**, aac9323 (2015).

27. Voss, M. et al. The marine nitrogen cycle: recent discoveries, uncertainties and the potential relevance of climate change. *Philos. T. R. Soc. B.* **368**, 20130121 (2013).

28. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590-D596 (2013).

29. Nguyen, N.H. et al. FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol.* **20**, 241-248 (2016).

30. Polme, S. et al. FungalTraits: a user-friendly traits database of fungi and fungus-like stramenopiles. *Fungal Divers.* **105**, 1-16 (2020).

31. Clark, D.R., Underwood, G.J.C., McGenity, T.J. & Dumbrell, A.J. What drives study-dependent differences in distance-decay relationships of microbial communities? *Global Ecol. Biogeogr.*

627 **30**, 811-825 (2021).

628 32. Li, P.F. et al. Spatial Variation in Soil Fungal Communities across Paddy Fields in Subtropical
629 China. *mSystems* **5** (2020).

630 33. Martiny, J.B. et al. Microbial biogeography: putting microorganisms on the map. *Nat. Rev.*
631 *Microbiol.* **4**, 102-112 (2006).

632 34. Bahram, M. et al. Structure and function of the soil microbiome underlying N2O emissions
633 from global wetlands. *Nat. Commun.* **13**, 1430 (2022).

634 35. Parn, J. et al. Nitrogen-rich organic soils under warm well-drained conditions are global nitrous
635 oxide emission hotspots. *Nat. Commun.* **9**, 1135 (2018).

636 36. Borcard, D. & Legendre, P. All-scale spatial analysis of ecological data by means of principal
637 coordinates of neighbour matrices. *Ecol. Model.* **153**, 51-68 (2002).

638 37. Root, T.L. et al. Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57-60
639 (2003).

640 38. Pugnaire, F.I. et al. Climate change effects on plant-soil feedbacks and consequences for
641 biodiversity and functioning of terrestrial ecosystems. *Sci. Adv.* **5**, eaaz1834 (2019).

642 39. Antonelli, A. et al. Geological and climatic influences on mountain biodiversity. *Nat. Geosci.* **11**,
643 718-725 (2018).

644 40. Yuan, W. et al. Increased atmospheric vapor pressure deficit reduces global vegetation growth.

Sci. Adv. **5**, eaax1396 (2019).

41. Wang, S.H. et al. Recent global decline of CO2 fertilization effects on vegetation photosynthesis. *Science* **370**, 1295-1300 (2020).

42. Fischer, E.M., Sippel, S. & Knutti, R. Increasing probability of record-shattering climate extremes. *Nat. Clim. Change* **11**, 689-695 (2021).

43. Jack, C.N., Petipas, R.H., Cheeke, T.E., Rowland, J.L. & Friesen, M.L. Microbial inoculants: Silver bullet or microbial jurassic park? *Trends Microbiol.* **29**, 299-308 (2021).

44. Mawarda, P.C., Le Roux, X., Van Elsas, J.D. & Salles, J.F. Deliberate introduction of invisible invaders: A critical appraisal of the impact of microbial inoculants on soil microbial communities. *Soil Biol. Biochem.* **148** (2020).

45. Thomsen, C.N. & Hart, M.M. Using invasion theory to predict the fate of arbuscular mycorrhizal fungal inoculants. *Biol. Invasions* **20**, 2695-2706 (2018).

46. Yang, X. et al. MBPD: A multiple bacterial pathogen detection pipeline for One Health practices. *iMeta* **e82** (2023).

47. Bull, C.T. et al. List of new names of plant pathogenic bacteria (2008-2010). *J. Plant Pathol.* **94**, 21-27 (2012).

48. Biddle, J.F., Fitz-Gibbon, S., Schuster, S.C., Brenchley, J.E. & House, C.H. Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. *P. Natl. Acad. Sci. USA* **105**, 10583-10588 (2008).

664 49. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science
665 using QIIME 2. *Nat. Biotechnol.* **37**, 852-857 (2019).

666 50. Callahan, B.J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat.*
667 *Methods* **13**, 581-583 (2016).

668 51. McKnight, D.T. et al. Methods for normalizing microbiome data: An ecological perspective.
669 *Methods Ecol. Evol.* **10**, 389-400 (2019).

670 52. Fick, S.E. & Hijmans, R.J. WorldClim 2: new 1-km spatial resolution climate surfaces for global
671 land areas. *Int. J. Climatol.* **37**, 4302-4315 (2017).

672 53. Wang, S., Zhang, Y., Ju, W., Qiu, B. & Zhang, Z. Tracking the seasonal and inter-annual variations
673 of global gross primary production during last four decades using satellite near-infrared
674 reflectance data. *Sci. Total Environ.* **755**, 142569 (2021).

675 54. Wang, Y., Naumann, U., Wright, S.T. & Warton, D.I. mvabund- an R package for model-based
676 analysis of multivariate abundance data. *Methods Ecol. Evol.* **3**, 471-474 (2012).

677 55. Alzarhani, A.K. et al. Are drivers of root-associated fungal community structure context specific?
678 *Isme J.* **13**, 1330-1344 (2019).

679 56. Li, P.F. et al. Responses of microbial communities to a gradient of pig manure amendment in
680 red paddy soils. *Sci. Total Environ.* **705** (2020).

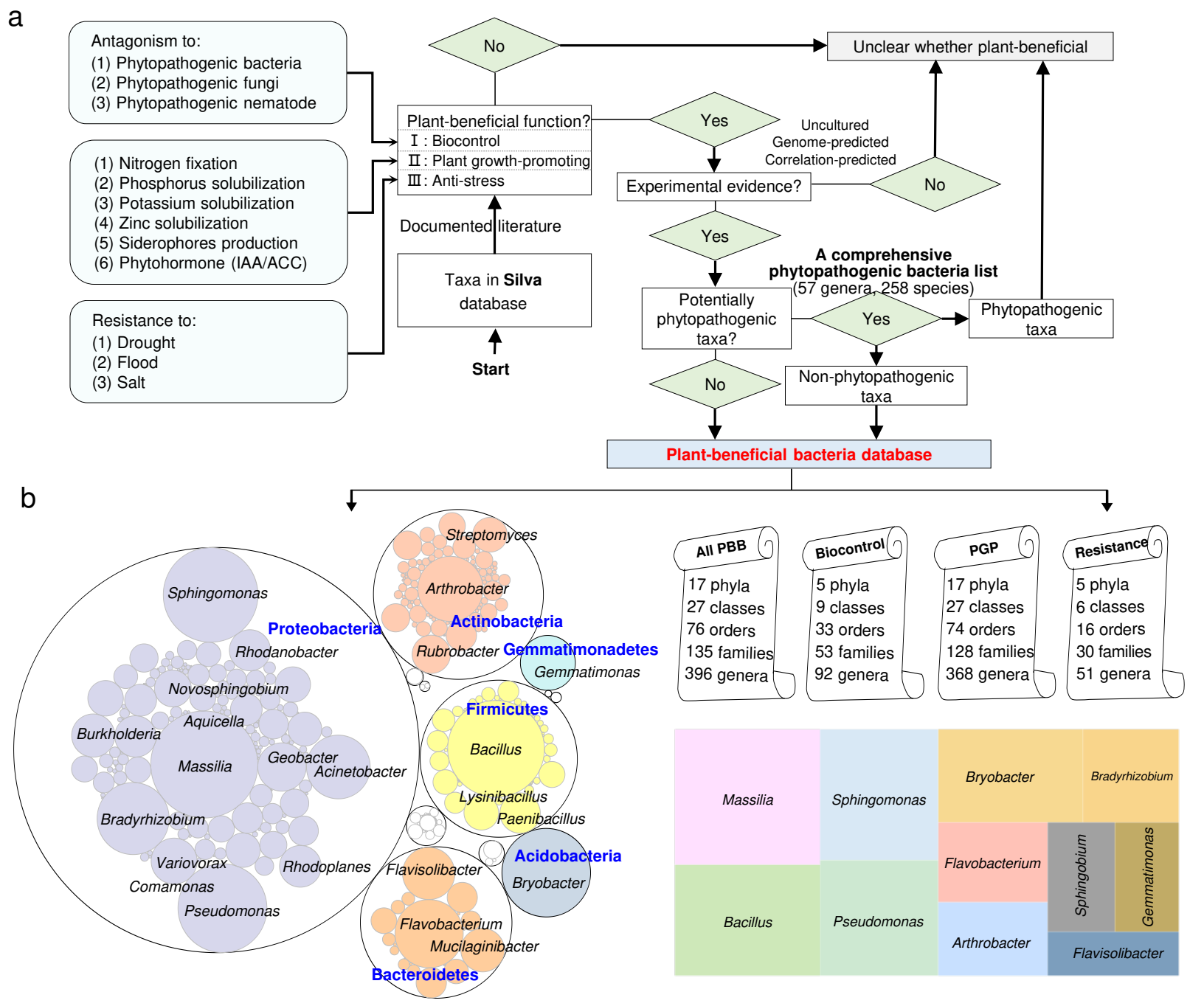
681 57. Sakamoto, T. Early classification method for US corn and soybean by incorporating MODIS-
682 estimated phenological data and historical classification maps in random-forest regression

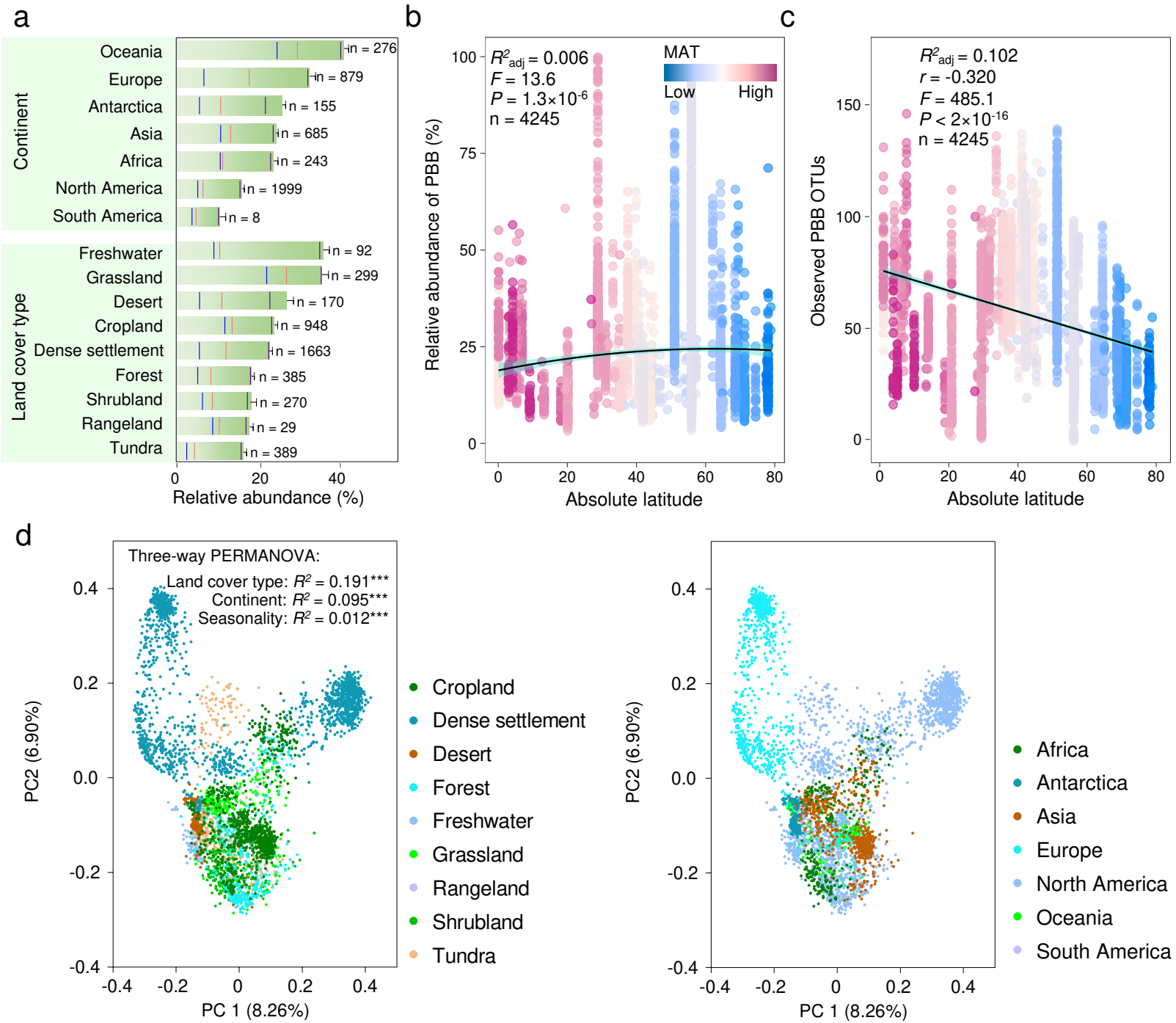
683 algorithm. *Photogramm. Eng. Rem. S.* **87**, 747-758 (2021).

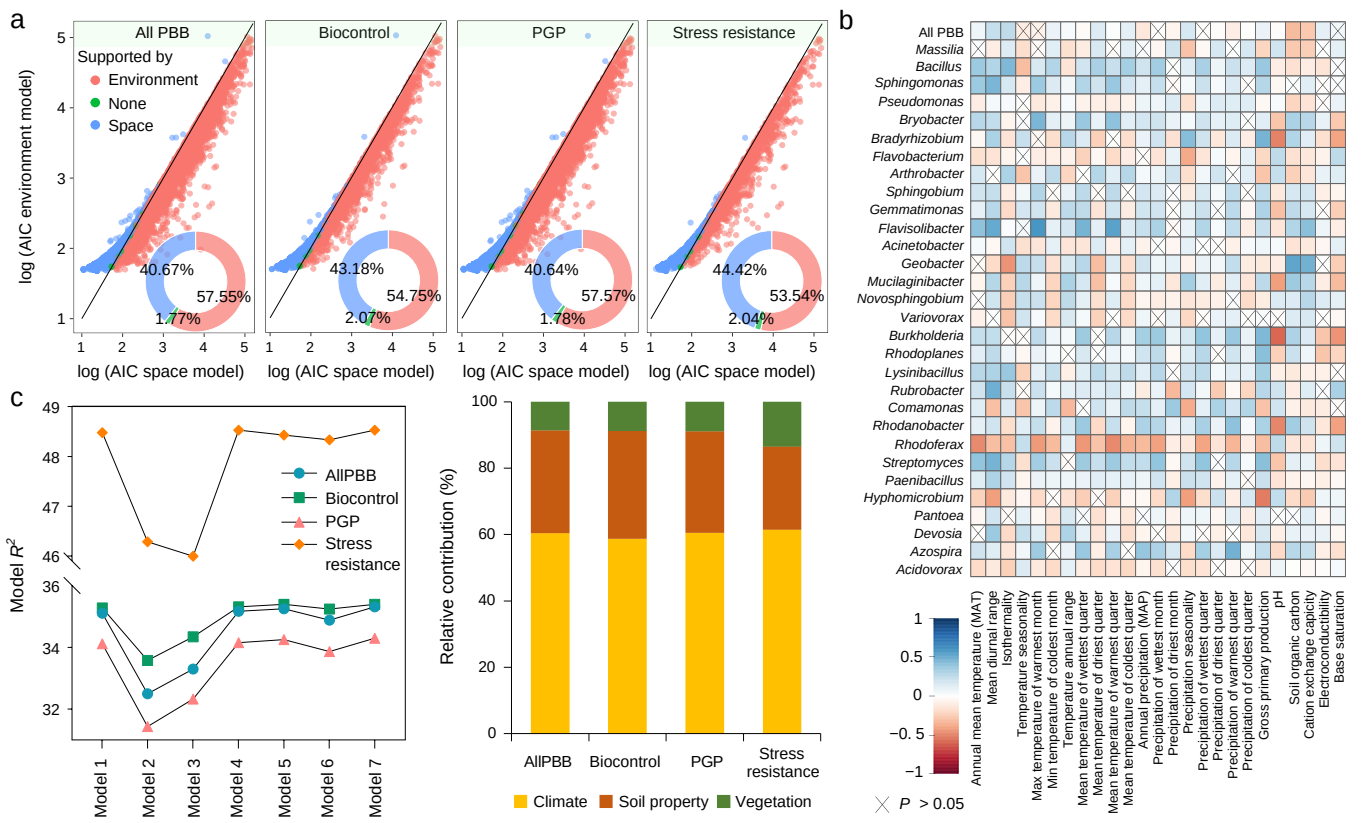
684 58. Elith, J., Kearney, M. & Phillips, S. The art of modelling range-shifting species. *Methods Ecol.*

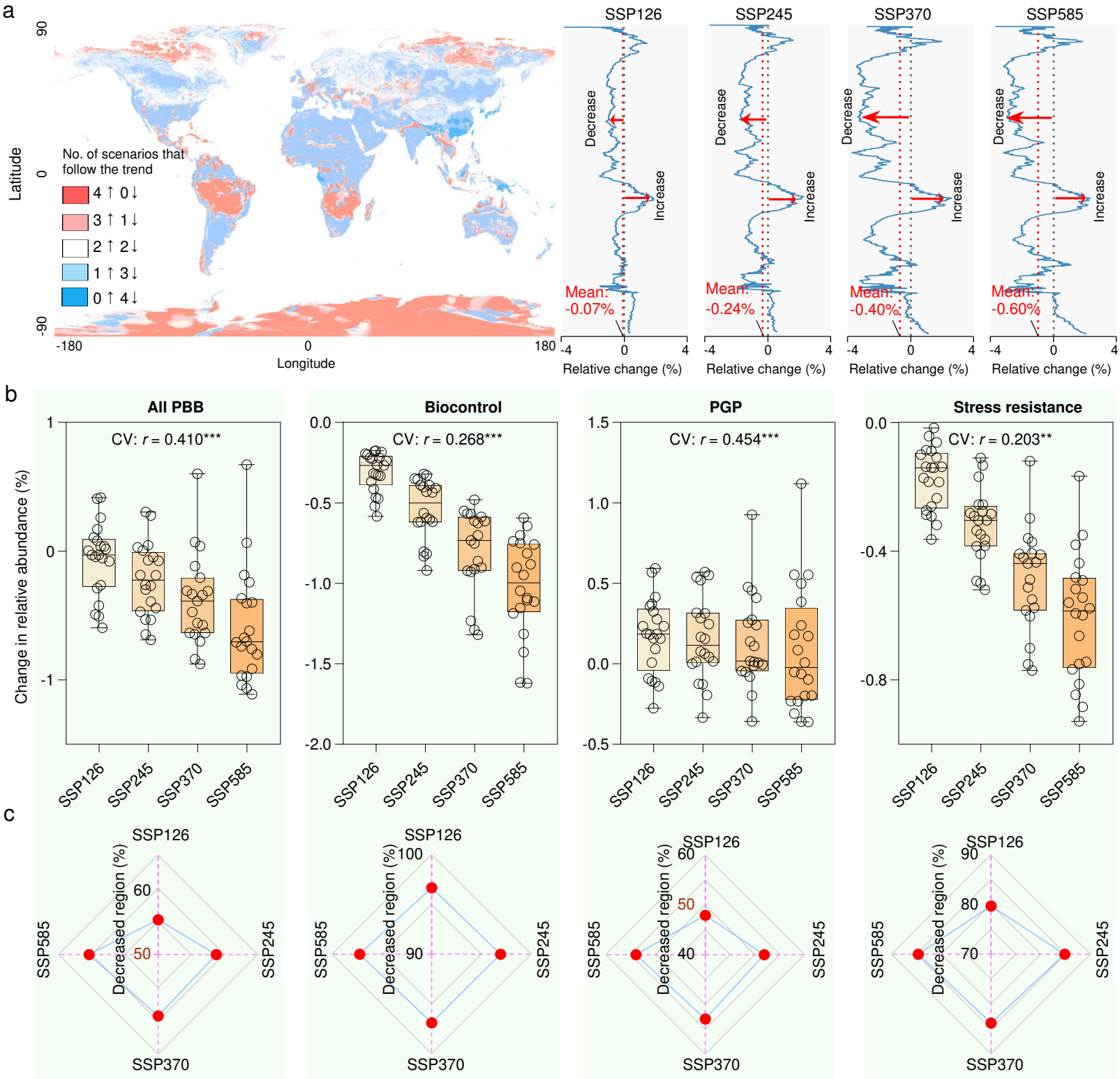
685 *Evol.* **1**, 330-342 (2010).

686


a
Antagonism to:
(1) Phytopathogenic bacteria
(2) Phytopathogenic fungi
(3) Phytopathogenic nematode
(1) Nitrogen fixation
(2) Phosphorus solubilization
(3) Potassium solubilization
(4) Zinc solubilization
(5) Siderophores production
(6) Phytohormone (IAA/ACC)
Resistance to:
(1) Drought
(2) Flood
(3) Salt
No
Unclear whether plant-beneficial
Plant-beneficial function?
Ⅰ: Biocontrol
Ⅱ: Plant growth-promoting
Ⅲ: Anti-stress
Yes
Uncultured
Genome-predicted
Correlation-predicted
Experimental evidence?
No
Yes
Documented literature
Taxa in Silva database
Start
A comprehensive phytopathogenic bacteria list
(57 genera, 258 species)
Potentially phytopathogenic taxa?
Yes
Phytopathogenic taxa
No
Non-phytopathogenic taxa
Plant-beneficial bacteria database
b
Sphingomonas
Proteobacteria
Rhodanobacter
Novosphingobium
Aquicella
Burkholderia
Massilia
Geobacter
Acinetobacter
Bradyrhizobium
Variovorax
Rhodoplanes
Comamonas
Pseudomonas
Streptomyces
Arthrobacter
Actinobacteria
Rubrobacter
Gemmatimonadetes
Gemmatimonas
Firmicutes
Bacillus
Lysinibacillus
Paenibacillus
Acidobacteria
Flavisolibacter
Bryobacter
Flavobacterium
Mucilaginibacter
Bacteroidetes
All PBB
17 phyla
27 classes
76 orders
135 families
396 genera
Biocontrol
5 phyla
9 classes
33 orders
53 families
92 genera
PGP
17 phyla
27 classes
74 orders
128 families
368 genera
Resistance
5 phyla
6 classes
16 orders
30 families
51 genera
Massilia
Sphingomonas
Bryobacter
Bradyrhizobium
Flavobacterium
Sphingobium
Gemmatimonas
Bacillus
Pseudomonas
Arthrobacter
Flavisolibacter


a
Continent
Oceania
Europe
Antarctica
Asia
Africa
North America
South America
n = 276
n = 879
n = 155
n = 685
n = 243
n = 1999
n = 8
Land cover type
Freshwater
Grassland
Desert
Cropland
Dense settlement
Forest
Shrubland
Rangeland
Tundra
n = 92
n = 299
n = 170
n = 948
n = 1663
n = 385
n = 270
n = 29
n = 389
0
20
40
Relative abundance (%)
b
R^2_{adj} = 0.006
F = 13.6
P = 1.3×10-6
n = 4245
MAT
Low
High
Relative abundance of PBB (%)
100
75
50
25
0
0
20
40
60
80
Absolute latitude
c
R^2_{adj} = 0.102
r = -0.320
F = 485.1
P < 2×10-16
n = 4245
Observed PBB OTUs
150
100
50
0
0
20
40
60
80
Absolute latitude
d
Three-way PERMANOVA:
Land cover type: R^2 = 0.191***
Continent: R^2 = 0.095***
Seasonality: R^2 = 0.012***
PC2 (6.90%)
PC 1 (8.26%)
0.4
0.2
0.0
-0.2
-0.4
Cropland
Dense settlement
Desert
Forest
Freshwater
Grassland
Rangeland
Shrubland
Tundra
Africa
Antarctica
Asia
Europe
North America
Oceania
South America


a
All PBB
Biocontrol
PGP
Stress resistance
Supported by
Environment
None
Space
log (AIC environment model)
log (AIC space model)
40.67%
57.55%
1.77%
43.18%
54.75%
2.07%
40.64%
57.57%
1.78%
44.42%
53.54%
2.04%
b
All PBB
Massilia
Bacillus
Sphingomonas
Pseudomonas
Bryobacter
Bradyrhizobium
Flavobacterium
Arthrobacter
Sphingobium
Gemmatimonas
Flavisolibacter
Acinetobacter
Geobacter
Mucilaginibacter
Novosphingobium
Variovorax
Burkholderia
Rhodoplanes
Lysinibacillus
Rubrobacter
Comamonas
Rhodanobacter
Rhodoferax
Streptomyces
Paenibacillus
Hyphomicrobium
Pantoea
Devosia
Azospira
Acidovorax
Annual mean temperature (MAT)
Mean diurnal range
Isothermality
Temperature seasonality
Max temperature of warmest month
Min temperature of coldest month
Temperature annual range
Mean temperature of wettest quarter
Mean temperature of driest quarter
Mean temperature of warmest quarter
Mean temperature of coldest quarter
Annual precipitation (MAP)
Precipitation of wettest month
Precipitation of driest month
Precipitation seasonality
Precipitation of wettest quarter
Precipitation of driest quarter
Precipitation of warmest quarter
Precipitation of coldest quarter
Gross primary production
pH
Soil organic carbon
Cation exchange capicity
Electroconductibility
Base saturation
P > 0.05
c
Model R2
AllPBB
Biocontrol
PGP
Stress resistance
Model 1
Model 2
Model 3
Model 4
Model 5
Model 6
Model 7
Relative contribution (%)
Climate
Soil property
Vegetation


a
Latitude
90
0
-90
-180
0
180
Longitude
No. of scenarios that follow the trend
4 ↑ 0↓
3 ↑ 1↓
2 ↑ 2↓
1 ↑ 3↓
0 ↑ 4↓
SSP126
SSP245
SSP370
SSP585
Decrease
Increase
Mean: -0.07%
Mean: -0.24%
Mean: -0.40%
Mean: -0.60%
-4
0
4
Relative change (%)
b
All PBB
Biocontrol
PGP
Stress resistance
CV: r = 0.410***
CV: r = 0.268***
CV: r = 0.454***
CV: r = 0.203**
Change in relative abundance (%)
c
Decreased region (%)