

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	We assembled a global dataset of 501 long-term agricultural experiments by combining standardized field sampling with a systematic literature synthesis. Field data were collected in April 2024 from 11 long-term fertilization trials across China, using paired unfertilized control and mineral nitrogen fertilization plots with three biological replicates per treatment. Literature data were collected from peer-reviewed studies identified through Web of Science using a search string developed from expert knowledge and refined with litsearchr. The final search was conducted on 14 June 2024 and returned 1,697 records. Studies were included if they were field experiments in croplands, lasted at least 5 years, contained paired unfertilized control and mineral N fertilization treatments, and reported bulk topsoil data from the cultivated layer. Studies involving organic amendments, rhizosphere samples, managed plantations, incubation experiments, soil replacement, liming, reductive soil disinfestation or nitrification inhibitors were excluded. After title, abstract and full-text screening by two authors, 225 primary studies published between 2003 and 2024 were retained. Means, standard deviations and sample sizes were extracted from text, tables or figures using WebPlotDigitizer, and site-level metadata, including coordinates, climate, fertilization duration, cropping regime and fertilizer application rates, were curated and standardized. Annual fertilizer inputs were converted to elemental N, P and K. The field and literature datasets were then harmonized to generate paired observations comparing mineral fertilization with unfertilized controls across 501 long-term experiments.
Data analysis	All statistical analyses were performed in R v4.4.2. Meta-analyses were conducted using the packages metafor v4.8, metagear v0.7, corpcor v1.6.10, orchaRd v2.0, and multcomp v1.4.28. Multilevel meta-analytic models were fitted with metafor using variance-covariance matrices and random effects. Missing standard deviations were imputed using metagear, and variance-covariance matrices were adjusted to ensure positive definiteness using corpcor. Subgroup contrasts were evaluated using multcomp, and orchard plots were generated using orchaRd to visualize effect sizes and confidence intervals. Data processing and general plotting were performed using tidyverse v2.0.0 and tidyplots v0.2.0. Co-occurrence networks of host-associated virulent phage vOTUs were inferred using Spearman correlations implemented in ggClusterNet v2.0, with abundance data organized using phyloseq v1.50. Spatial raster processing and geospatial data handling were

performed using tidygeocoder v1.0.6, terra v1.8.54, and sf v1.0.20. Generalized additive models were implemented using mgcv v1.9.1, including smoothing terms for environmental predictors and a two-dimensional thin-plate spline for geographic coordinates to account for spatial structure. The fitted models were applied to global gridded environmental layers to generate spatial predictions. Predictor multicollinearity was evaluated using variance inflation factors, and predictor importance was assessed using drop-one model comparisons by excluding each covariate and comparing model performance. Partial effects of predictors were visualized from the fitted GAMs. Inverse-variance weighted regressions were used to examine relationships between statistically significant changes in soil physicochemical properties and microbial response variables, with weights equal to the inverse of the sampling variance. All analysis scripts supporting the findings of this study have been deposited in Figshare and are accessible at <https://doi.org/10.6084/m9.figshare.29552153>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Amplicon and metagenomic sequence data generated in this study have been deposited in the NCBI Sequence Read Archive under BioProject accession numbers: PRJNA1273843 (16S rRNA gene amplicon), PRJNA1273925 (ITS amplicon), and PRJNA1273624 (metagenomic sequencing). Source data are provided as a Source Data file.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Not applicable. This study did not involve human participants, human data, or human biological material.
Reporting on race, ethnicity, or other socially relevant groupings	Not applicable. No human participants or socially defined groups were included in this study.
Population characteristics	Not applicable. This study does not include any human research participants.
Recruitment	Not applicable. No recruitment of human participants was performed.
Ethics oversight	Not applicable. No ethics approval for human research was required as this study did not involve human subjects or human materials.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study integrates standardized field surveys and a global synthesis of 501 long-term agricultural experiments (median duration 25 years) to quantify the effects of sustained mineral fertilization on soil physicochemical properties, microbial biomass, diversity, and functional activity.
Research sample	The dataset includes 501 long-term agricultural experiments distributed across the globe, covering upland and paddy systems under varying fertilizer intensities. Field sampling was additionally conducted at 11 long-term experimental sites across China to fill data gaps, yielding 132 composite soil samples analyzed for microbial, enzymatic, and physicochemical variables.
Sampling strategy	Eligible studies were selected using a systematic literature search (Web of Science, last updated June 2024) and strict inclusion criteria: field-based long-term (> 5 years) mineral N fertilization with unfertilized controls, cropland soils only, and exclusion of organic amendments. Field plots followed paired-control designs with triplicate replicates per treatment.

Data collection	Data were obtained from peer-reviewed publications and newly collected field samples. Soil physicochemical parameters, microbial biomass, extracellular enzyme activities, and microbial community composition (amplicon and metagenomic sequencing) were determined using standardized protocols. Literature data were extracted as means $\pm$ SD and sample sizes for control and treatment groups, standardized into effect sizes (lnRR).
Timing and spatial scale	Global dataset of agricultural experiments spanning 5-164 years of continuous mineral fertilization, with worldwide coverage.
Data exclusions	Studies were excluded if they: (i) involved organic amendments such as manure, compost, straw, or biochar; (ii) were incubation or non-field experiments; (iii) had fertilization duration shorter than 5 years; (iv) were conducted in managed plantations rather than croplands; (v) focused on rhizosphere soils or deep layers instead of bulk topsoil; or (vi) included treatments with liming, soil disinfection, or nitrification inhibitors.
Reproducibility	The study ensures reproducibility through full transparency of data and code (publicly archived on Figshare: https://doi.org/10.6084/m9.figshare.29552153), detailed documentation of methods (see Supplementary Information).
Randomization	Randomized block or split-plot designs were used in most long-term field experiments included in the synthesis, following local agronomic standards. Random assignment of treatments was maintained throughout experiment durations.
Blinding	Blinding was not applicable, as the study combined existing field experiments and published datasets where treatment identities were known. Data extraction and verification were independently performed by two authors to minimize bias.
Did the study involve field work?	☒ Yes ☐ No

Field work, collection and transport

Field conditions	Soil samples were collected from paired fertilized and unfertilized plots (0–20 cm) at 11 long-term experimental sites across China in April 2024. Field sites were located within established long-term fertilization trials across major agroecological zones of China, spanning gradients in climate, soil properties and cropping systems. Relevant environmental and management variables included mean annual temperature, mean annual precipitation, agroecosystem type, cropping regime, fertilization duration and annual N, P and K application rates. Site environmental and management variables are provided in the Source Data and Supplementary Information.
Location	Field sampling was conducted at 11 established long-term fertilization trials across China: Fengqiu (35.000°N, 114.400°E; 70 m a.s.l.), Yangling (34.283°N, 108.066°E; 509 m), Luancheng (37.883°N, 114.683°E; 50 m), Changshu (31.583°N, 120.916°E; 6 m), Jinhua (30.016°N, 120.700°E; 12 m), Fukang (44.283°N, 87.933°E; 461 m), Quzhou (36.866°N, 115.016°E; 37 m), Chengdu (30.945°N, 103.733°E; 648 m), Yingtan (28.258°N, 116.925°E; 42 m), Qiqihar (47.500°N, 124.250°E; 151 m) and Guizhou (26.513°N, 106.653°E; 1298 m).
Access & import/export	Soil samples were collected from 11 long-term experimental sites across China in April 2024. Sampling was conducted with permission from the managers or responsible institutions of the participating long-term experimental sites and in accordance with local site-management requirements. Samples from three replicate plots per treatment were homogenized and transported within China for laboratory analyses. No soil samples were imported into or exported from China for this study, and no additional collection or import/export permits were required.
Disturbance	Disturbance was limited to routine collection of small-volume topsoil cores from designated plots within established long-term agricultural experiments. Sampling was restricted to the minimum amount required for laboratory analyses.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	Not applicable.
Novel plant genotypes	Not applicable.
Authentication	Not applicable.