Machine learning predicts ecological risks of nanoparticles to soil microbial communities

Nuohan Xu[1], Jian Kang[1], Yangqing Ye[2], Qi Zhang[1], Mingjing Ke[1], Yufei Wang[3], Zhenyan Zhang[1], Tao Lu[1], W.J.G.M. Peijnenburg[4, 5], Josep Penuelas[6,7], Guanjun Bao[2, *], Haifeng Qian[1, *]

1. College of Environment, Zhejiang University of Technology, Hangzhou, Zhejiang 310032, P. R. of China

2. College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou, Zhejiang 310032, P. R. of China

3. Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430030, P. R. of China

4. Institute of Environmental Sciences (CML), Leiden University, RA Leiden 2300, the Netherlands

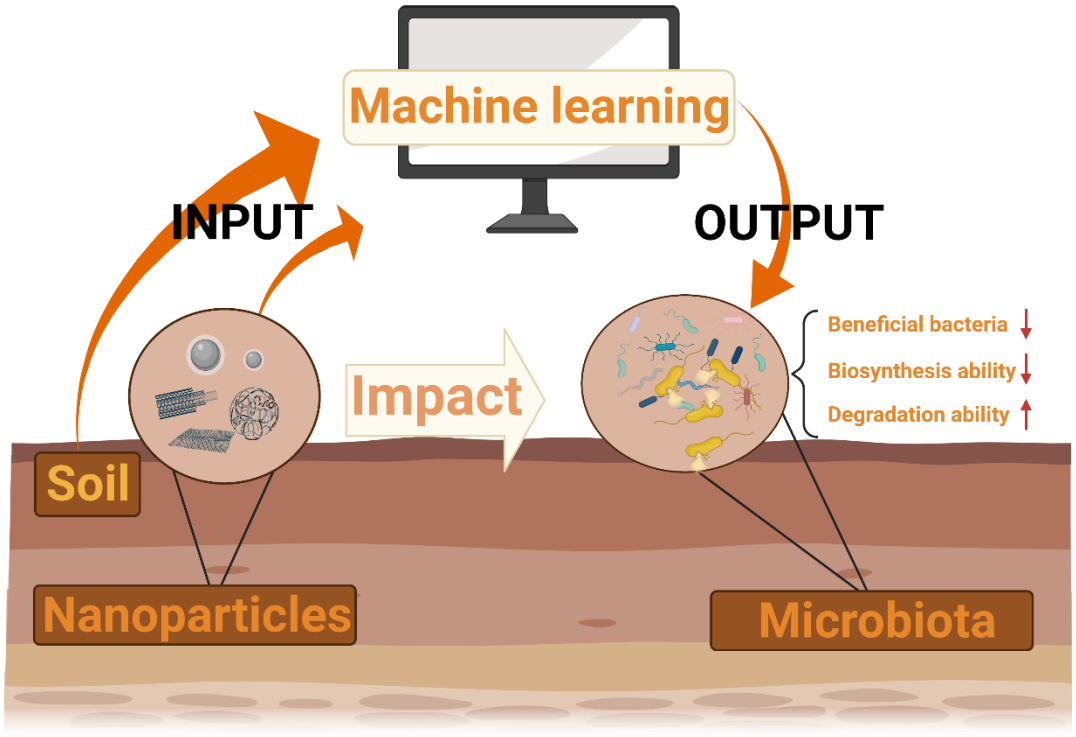
5. National Institute of Public Health and the Environment (RIVM), Center for Safety of Substances and Products, P.O. Box 1, Bilthoven, the Netherlands

6. CSIC, Global Ecology Unit CREAF- CSIC-UAB, Bellaterra, Barcelona 08193, Catalonia, Spain.

7. CREAF, Campus Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Barcelona 08193, Catalonia, Spain.

[*] Correspondence Author: E-mail: gjbao@zjut.edu.cn (Bao G.); E-mail: hfqian@zjut.edu.cn (Qian H.).

22 **TOC ART**



23

24

Abstract

With the rapid development of nanotechnology in agriculture, there is increasing urgency to assess the impacts of nanoparticles (NPs) on the soil environment. This study merged raw high-throughput sequencing (HTS) data sets generated from 365 soil samples to reveal the potential ecological effects of NPs on soil microbial communities by means of metadata analysis and machine learning methods. Metadata analysis showed that treatment with nanoparticles did not have a significant impact on the alpha diversity of the microbial community, but significantly altered the beta diversity. Unfortunately, the abundance of several beneficial bacteria, such as *Dyella*, *Methylophilus*, *Streptomyces*, which promote the growth of soil protozoa and plants, and improve pathogenic resistance, was reduced under the addition of synthetic nanoparticles. Furthermore, metadata demonstrated that nanoparticles treatment weakened the biosynthesis ability of cofactors, carriers, and vitamins, and enhanced the degradation ability of aromatic compounds, amino acids, etc. This is unfavorable for the performance of soil functions. Besides the soil heterogeneity, machine learning uncovered that a) the exposure time of nanoparticles was the most important factor to reshape the soil microbial community, and b) long-term exposure decreased the diversity of microbial communities and the abundance of beneficial bacteria. This study is the first to use a machine learning model and metadata analysis to investigate the relationship between the properties of nanoparticles and the hazards to the soil microbial community from a macro perspective. This guides the rational use of nanoparticles for which the impacts on soil microbiota are minimized.

Keywords: Machine learning; Soil ecosystems; Microbiota; Nanoparticles; Metadata analysis; Ecotoxicity.

1. Introduction

Nanoparticles (NPs) are becoming extensively used additives in modern intensive agriculture. They are widely used in various applications, including the improvement of the quality of saline-alkali soil (Guerriero et al., 2021), the maintenance of vegetables fresh (Fayaz et al., 2009), and the design of new agricultural nanopesticides (You et al., 2018). The unintended release of NPs may though lead to their accumulation in soil and pose potential ecological and human health risks (Peng et al., 2017). The production of NPs in the USA was estimated to be 2.8 to 20 tons per year in 2011 and will reach 2.5 million tons per year by 2025 (Hendren et al., 2011; McShane et al., 2012), and the soil is considered to be the sink of these NPs (Sun et al., 2015). Soil bacteria helps crops grow by improving soil structure, and recycling of soil nutrients (Bahram et al., 2018; Guo et al., 2020; Qu et al., 2020; Ray et al., 2020). Therefore, the stability of soil microbial communities is one of the key factors in the maintenance of the soil ecosystem function (Delgado-Baquerizo et al., 2016). For example, altered microbial communities will interfere with the biosynthesis of nutrients from the soil pool and the degradation of organic matter into inorganic matter (Lehmann et al., 2017).

The negative influence of NPs on soil microbiota is related to the composition, concentration, type, exposure time, and particle size of NPs (Zhang et al., 2020b). NPs have detrimental effects on microbial activity, abundance, and diversity (Simonin and Richaume, 2015) and alter ecosystem functioning (Simonin et al., 2018). The microbiota of soils submitted to NPs treatments have been characterized by using

high-throughput sequencing (HTS) methods that show that microbial communities are redrawing the map of the microbial kingdom by altering its enormous - and largely uncharted - taxonomic and functional diversity (Nilsson et al., 2019). Development of HTS methods help to reveal the effects of NPs on the structure and stability of the soil microbial community (Zhang et al., 2020a). Open-access HTS data from independent studies and the development of analytical workflows provide a new opportunity to evaluate global microbial patterns (Ramirez et al., 2018; Wright et al., 2020).

However, independent studies have occasionally reported contrasting results due to the high heterogeneity of soils from different localities, NPs properties, and variable experimental conditions, which result in a lack of a consensus concerning microbiota responses to NPs treatment (Ben-Moshe et al., 2013; Fan et al., 2018; Zhang et al., 2018). To shed light into these responses, machine learning, can become a key tool for NPs properties identification (Wright et al., 2020; Yuan et al., 2020, Ban et al., 2020;Ahneman et al., 2018; Oh et al., 2016). In fact, machine learning combined with HTS metadata analysis have already been applied to analyze the characteristics of the soil microbiome and predicting the occurrence of *Fusarium* wilt (Yuan et al., 2020).

With the aim of identifying the main factors associated with the ecological impact of NPs overcoming the high heterogeneity of data in independent experiments, we integrated the available raw HTS sequences data for microbiome nanotoxicology and re-analyzed the data using machine learning. Our aims were to: (1) identify the common effects of NP treatment on soil microbial community diversity, composition and function; (2) determine how NPs affect the change in the accumulation of the

94 significant beneficial bacteria and pathogenic bacteria in soils; and (3) build a
95 predictive model to correlate NPs properties with the stability of soil microbial
96 communities. Accurately predicting the contribution of NP characteristics on soil and
97 agricultural ecosystem variation can guide the wide use of eco-friendly NPs.

98

99 **2. Material and methods**

100 *2.1. Data collection and description*

101 The data of microbial high-throughput sequencing (HTS) in NPs treated and untreated
102 soils were collected by searching the keywords “NPs and soil microbial”, “NPs and
103 soil HTS”, “nanomaterials and soil bacterial” in Google Scholar and the National
104 Center for Biotechnology Information (NCBI) SRA database. After filtering the
105 sources of metadata that did not contain basic information, we grouped the data into
106 the categories of nano-treated or -untreated soils. Since the bioproject accession
107 numbers of the HTS data results of most microbial community studies were not
108 uploaded, we collected approximately 63 studies matching the keywords, of which
109 only 18 data sets could be downloaded from public databases (Table S1). We obtained
110 632 microbial HTS samples from 10 countries around the globe (Fig. 1). After
111 removing samples of incomplete, low-abundance sequences, and containing
112 chlorophyll and mitochondria, and then rarefied according to the minimum frequency,
113 there were 365 high quality bacterial samples left. Based on the data preprocessing
114 presented above, we acquired 13 independent studies with 125 controls and 240 NPs-
115 treated samples. The metadata were classified as from “nano-treated” and “nano-

116 untreated" after carefully read the full texts of each paper. We defined nano-treated
117 group as the soil exposed by NPs, and nano-untreated group as the soil without the
118 NPs treatment.
119
120 *2.2. Data Processing*
121 For each independent study, the HTS data were processed using QIIME2 (Version
122 2020.8) core distribution, following the standard workflows (Bolyen et al., 2019).
123 Raw data from each of the studies were imported into QIIME2, and adaptor and
124 primer sequences were then removed from the reads using Cutadapt (Martin, 2011).
125 The latter step was omitted for samples where the sequences had already been
126 removed. Then, we used Dada2 to join paired end reads, denoising sequences as well
127 as resolving amplicon sequence variants (ASVs) under default quality thresholds
128 (Callahan et al., 2016). Dada2 was run using trim lengths determined by the read
129 quality for each study. The samples with low reverse read qualities, which caused too
130 few reads remaining after running Dada2, were only used the forward reads to gain
131 more reads after Dada2. Feature tables and the representative sequence of ASVs were
132 acquired, and then merged by the QIIME2's merge plugin. All feature sequences were
133 annotated with the SILVA v138 database of full-length 16S rRNA gene sequences for
134 subsequent analysis. The workflow of this study is shown in Fig. S1. To eliminate
135 PCR primer deviation and over amplification, we filtered the ASVs table according to
136 the following rules: (1) the samples were rarefied to 2000 reads, and those with <2000
137 reads were removed; (2) sequences annotated to mitochondria and chloroplasts or that

could not be classified at the kingdom level were removed; and (3) ASVs with a maximum number of 20 reads (1% total reads in each sample) were removed. Feature tables for each taxonomic level were then generated, and the data were converted to relative abundance.

2.3. PICRUSt2 for prediction of metagenome functions

Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt2) is a software package suited for predicting functional abundances based on marker gene sequences (Douglas et al., 2020). Function prediction on the merged ASVs table and sequences was applied. The annotation and classification of the functional pathways were acquired through the database MetaCyc (https://metacyc.org) (Caspi et al., 2020).

2.4. Construction of a predictive random forest model

We developed a regression model established by the random forest (RF) algorithm. The established RF model not only took the properties related to NPs as independent variables, but also the properties related to the soil, with a total of 21 input variables (Table. S2). The dependent variables of the RF model were the values of Richness and Shannon of the soil microbial community previously analyzed. To ensure the prediction accuracy and applicability, we utilized the 10-fold cross validation method, R^2 (coefficient) and RMSE (root mean square error) to evaluate the model performance.

The imported dataset was partitioned into 10 folds, in which 70% of the training data

were used to build the regression model, and the remaining 30 % were used for prediction using the trained model. Ten-fold cross-validation with five repeats was used to verify the regression models, and we selected the model with the highest R^2 value for subsequent analysis. To estimate the importance of the different properties, the increase of node purity of each model was calculated by the R package 'RandomForest'. The construction and prediction performance of the random forest regression model (n_estimators=150, max_leaf_nodes=10) and the verification of the 10-fold cross-validation method were both used by the scikit-learn (Version 0.24.0) library of the python language (Python 3.9).

2.5. Statistical methods

The alpha diversity index of Shannon and Richness were calculated using the R vegan and picante package (version 4.0.3). Principal co-ordinates analysis (PCoA) plots were generated from the Bray-Curtis distance created using the R package ggplot2 and vegan (version 4.0.3). Permutational multivariate analysis of variance (PERMANOVA) (ADONIS, transformed data by Bray-Curtis, permutation = 999) was used to determine if beta diversity differed between NPs treated and untreated groups. A hierarchical bubble chart, sunburst chart and world map were drawn by R package ggplot2, ggmap, ggraph and igraph (version 4.0.3). Opportunistic human pathogens were searched from an online database (https://www.bode-science-center.com/center/relevant-pathogens-from-a-z.html). All bar charts were designed by the software Prism 5.0. The graphical abstract was created with BioRender.com. The remaining results are presented as the means ± standard deviations (SDs). Significant

differences ($p < 0.05$) were evaluated by Kruskal-Wallis test using the R function kruskal.test (version 4.0.3).

3. Results

3.1. Diversity and taxonomic difference after NPs treatment

Alpha diversity indices (Richness and Shannon) were not significantly different between the NPs-treated and NPs-free soils (Fig. 2a), while principal coordinates analysis (PCoA) of Bray–Curtis dissimilarity with multivariate permutational analysis of variance (PERMANOVA) revealed that the structure of the microbial communities significantly differed (Fig. 2b, $R^2 = 0.010$, $p < 0.001$; ADONIS). At the phylum level, the soil microbial communities in all samples were mainly composed of *Proteobacteria*, *Actinobacteriota*, *Bacteroidetes*, *Acidobacteriota* etc. (Fig. 2c). Among these phyla, *Proteobacteria* in NPs treatment increased significantly by 29%, while *Acidobacteriota*, *Verrucomicrobiota,* and *Planctomycetota* decreased by 31, 60, and 43%, respectively (Fig. S2). The Venn diagram illustrated that the microbial community composition of the control and NPs-treated groups shared most of the bacterial taxa, with more unique bacterial taxa in the NPs-treated group (219) than in the untreated group (114) (Fig. 2d). In conclusion, our results indicated that NPs treatment altered the microbial community composition and structure.

3.2. The profile of beneficial and pathogenic soil bacteria after NPs treatment

We retrieved 33 plant growth-promoting and soil disease suppression bacteria, as well

as 26 pathogenic bacteria related to human and soil-borne bacteria pathogens. NPs treatment . decreased the abundances of *Methylophilus*, *Dyella* and *Streptomyces*, which are bacteria that promote plant growth, and increased the abundance of symbiotic *Burkholderia-Caballeronia-Paraburkholderia* (Fig. 3a-d). NPs treatment also increased the abundance of one pathogenic bacterium, *Sphingomonas,* (Fig. 3e).

3.3. Functional traits of bacteria after NPs treatment

Most functional pathways were shared between the NPs-treated and untreated groups, but 189 out of a total of 415 functional pathways were different between the two groups (Fig. S3, Table. S3). These 189 significantly different functional pathways belong to four major categories: Biosynthesis, Degradation, Macromolecule modification, Generation of precursor metabolites and energy (Fig. 4), and 91 of them were down-regulated, whereas 98 pathways were up-regulated in the NPs treatment. In the superclass Biosynthesis, although there were a few up-regulated functional pathways, especially in Cofactor, Carrier, and Vitamin Biosynthesis, most functional pathways were down-regulated. On the other hand, most pathways in the superclass Degradation were up-regulated, such as the subclasses Aromatic Compound Degradation, Amino Acid Degradation, Carboxylate Degradation, Amine and Polyamine Degradation (Fig. 4).

3.4. Development of models to connect NPs characteristics and microbial community

traits

The models for the 10-fold cross validation displayed a robust performance (consistently high R^2, mostly > 70%) with values of R^2 mostly ranging between 69 and 91% (Fig. 5a) revealing a proper prediction of the microbial diversity (Fig. S4). Based on the models constructed, we analyzed the importance of each of the characteristics of the models. Regardless of whether it was the model for predicting Richness or the model for predicting the value of the Shannon index, the most important traits were related to heterogeneous soil characteristics, such as longitude, soil sampling depth, and amplification area (Fig. 5b, c). Despite heterogeneities, the exposure time of the NPs had the greatest impact on the soil microbial communities, followed by the extent of aggregation of the NPs (Fig. 5b, c). To explore the specific variation of the most important NPs property, we carried out a two-dimensional scatter diagram analysis of the NPs exposure time and the value of the Richness and Shannon indices of the microbial community. No matter the exposure time, the Richness of the microbial community decreased after NPs treatment (Fig. 5d). We separated exposure time into three phases, with the acute group treated for no more than three days, the subacute group treated in between three to 90 days, and the chronic group treated for more than 90 days. We found that acute NPs treatment increased the Shannon value, but subacute and chronic exposure decreased the Shannon value (Fig. 5d). Three beneficial bacteria were all down-regulated significantly after NPs treatment, including *Dyella*, *Methylophilus*, and *Streptomyces*. (Fig. 5e). The abundance of *Burkholderia-Caballeronia-Paraburkholderia* increased

after long-term exposure (Fig. 5e). However, the abundance of the human pathogen *Sphingomonas* increased under both subacute and chronic exposure conditions (Fig. 5e).

4. Discussion

We merged and normalized the HTS data sets of independent experiments to avoid the deviation of heterogeneous factors in traditional experiments such as sequencing depth, primers, and amplification area, and analyzed the soil microbial characteristics and potential functional features to figure out clearly and uniformly the impact of NPs on soil microbial community diversity, composition, and functions. The results of this analysis show that the treatment of NPs can cause a decrease in biomass and reduce bacterial diversity (Chen et al., 2019; Cocozza et al., 2019). This finding is consistent with previous results (Fig. 2a, 2b and 2c). *Proteobacteria*, *Acidobacteriota*, *Verrucomicrobia*, and *Planctomycetota* were four dominant bacterial phyla in soil (Yang et al., 2019). NPs treatment enriched *Proteobacteria*, including many genera of pathogenic bacteria in this phylum (Shin et al., 2015). Several beneficial soil phyla were down-regulated after NPs treatment, such as *Acidobacteriota*, which is a driver of various ecosystem processes, *Verrucomicrobia*, which is one of the most indispensable phyla that affect soil multifunctional resistance in the soil microbial community, and *Planctomycetota*, which plays an important role in global carbon and nitrogen cycles (Kielak et al., 2010; Wiegand et al., 2018; Yang et al., 2021). The decrease of beneficial bacterial genera of *Dyella*, *Methylophilus*, *Streptomyces* could

decrease resistance of plants to pathogenic microbiota, and weaken exogenous pollutants degradation in soil (Ali et al., 2021; Tapia-García et al., 2020; You et al., 2021). Therefore, these negative effects of NPs contamination on soil-borne beneficial bacteria aggravated potential risks to soil ecosystemic health, and the increase in the abundance of human pathogenic bacteria increased the potential risk for human health (Fang et al., 2018).

The soil microbiota is a crucial driver of global nutrient cycles and of plant growth regulation (Bahram et al., 2018; Finkel et al., 2020). Accurate functional prediction from our merged metadata sets by the PICRUSt2 software showed that degradation pathways in NPs treatment were generally up-regulated, while several biosynthetic pathways were generally down-regulated (Fig. 4). This implies that the ability of the soil microbiota to synthesize antibiotics, siderophores, and hormones related to plant growth, was in all cases weakened (Hibbing et al., 2010). Soil microbiota submitted to NPs treatment increased the degradation of organic nutrients enhancing the tolerance to contamination (Kusi et al., 2020).

A complex set of data was used to obtain meaningful and regular patterns and predict various biological responses through machine learning (Ahneman et al., 2018). The non-linear regression models obtained by applying a Random Forest (RF) algorithm were built to reveal the nano-characteristics that contributed most to variations in the diversity of the soil microbial community. Compared with logistic regression, support vector machines, and other machine learning algorithms, the RF model displayed robust predictive performance for discrete as well as continuous data (Yuan et al.,

292 2020). The RF model provided the possibility to explore the most important
293 properties of NPs that alter the stability of soil microbial community. Our developed
294 RF models had good predictive performance with high R^2 (mostly>70%). From these
295 public data, we determined that the geographic location of the sampling site, soil type,
296 soil acidity and alkalinity, and other heterogeneous characteristics of the soil
297 microbiota, have greater impact on soil microbial diversity than NPs properties
298 themselves. Notably, after removing other soil heterogeneity factors, the exposure
299 time of NPs was the most crucial for the Richness and Shannon models we
300 constructed, implicating that NPs displayed long-term impacts on soil microbiota
301 (Moll et al., 2017). Although in many independent experimental kinds of research, the
302 concentration and type of NPs were considered to be non-negligible traits affecting
303 soil microbial communities, the profile of NPs properties by machine learning
304 provides a baseline for using eco-friendly NPs. The potential adverse effects of NPs to
305 the soil ecosystem can be minimized by adjusting and restricting the exposure time of
306 NPs in soil. With the available heterogeneous data in hand, the RF model can be
307 applied to explore unknown and complex relationships hidden in various quantitative
308 and qualitative factors (Yu et al., 2021). However, the innovative exploration of
309 metadata analysis combined with machine learning still has certain limitations. Due to
310 the lack of standardization workflow of the uploaded HTS data sets, and some low-
311 quality data sets, only 365 samples were available to construct RF models in this
312 study. Although our models displayed good performance, they had certain limitations
313 in explaining the impacts of NPs properties on the diversity and community of the soil

microbiota.

5. Conclusions

Ours results indicate that NPs exposure altered the structure of the soil microbial community in the studies included in this modelling exercise. The decrease of the abundance of bacteria beneficial to soil protozoa and plants increases the risk of diseases of animals and plants in the soil. NPs treatment weakens the biosynthesis ability and strengthens the degradation ability of soil microbiota, thus affecting of soil ecosystemic functioning. Long-term NPs exposure down regulated the diversity of microbial communities and the abundance of beneficial bacteria. This study provided an accurate predictive model to reveal the properties of NPs that determine soil microbial diversity and soil microbial functioning. The model provides guidance for the application of NPs in industrial, agricultural, manufacturing, and other fields to minimize environmental risks.

Author contributions

NX, GB, and HQ conceived the study, JK, MK and ZZ collected the data, NX, YY and TL analyzed the data, NX, QZ, YW, WP, JP and HQ wrote the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (21976161, 21777144, and 41907210).

Data availability

Information for all metadata used was provided in Table S1. Raw data can also be provided by the corresponding author upon request. Codes associated with this study will be publicly available in (https://github.com/Plantxnh/Nanoparticles) upon manuscript acceptance.

Supporting Information

Tables

Table S1. The profile of the data collected.

Table S2. The profile of feature variables of machine learning models.

Table S3. The profile of shared differential functional pathways after NPs treatment.

Figures

Figure S1. Schematic of the bioinformatic processing of the raw High-through sequence data.

Figure S2. There were significant differences across phyla among the top 10. * and *** represent statistically significant differences at $p < 0.05$ and $p < 0.001$ from Kruskal-Wallis test. Error bars represent the means ± SEs.

Figure S3. Venn diagram of shared and unique metabolic pathway numbers observed in nano-untreated and nano-treated soil.

Figure S4. Prediction performance of Richness and Shannon using RF models. The slope of the solid lines was 1, and the dotted lines represented the intercepts of ± RMSE.

References

Ahneman, D.T., Estrada, J.G., Lin, S., Dreher, S.D., Doyle, A.G., 2018. Predicting reaction performance in C–N cross-coupling using machine learning. Science 360, 186–190. doi:10.1126/science.aar5169

Ali, A., Guo, D., Li, Y., Shaheen, S.M., Wahid, F., Antoniadis, V., Abdelrahman, H., Al-Solaimani, S.G., Li, R., Tsang, D.C.W., Rinklebe, J., Zhang, Z., 2021. Streptomyces pactum addition to contaminated mining soils improved soil quality and enhanced metals phytoextraction by wheat in a green remediation trial. Chemosphere 273, 129692. doi:10.1016/j.chemosphere.2021.129692

Ameen, F., Alsamhary, K., Alabdullatif, J.A., ALNadhari, S., 2021. A review on metal-based nanoparticles and their toxicity to beneficial soil bacteria and fungi. Ecotoxicology and Environmental Safety 213, 112027. doi:10.1016/j.ecoenv.2021.112027

Asadishad, B., Chahal, S., Akbari, A., Cianciarelli, V., Azodi, M., Ghoshal, S., Tufenkji, N., 2018. Amendment of Agricultural Soil with Metal Nanoparticles: Effects on Soil Enzyme Activity and Microbial Community Composition. Environmental Science & Technology 52, 1908–1918. doi:10.1021/acs.est.7b05389

Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H., Huerta-Cepas, J., Medema, M.H., Maltz, M.R., Mundra, S., Olsson, P.A., Pent, M., Põlme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., Bork, P., 2018. Structure and function of the global topsoil microbiome. Nature 560, 233–237.

386 doi:10.1038/s41586-018-0386-6
387 Ban, Z., Yuan, P., Yu, F., Peng, T., Zhou, Q., Hu, X., 2020. Machine learning predicts
388 the functional composition of the protein corona and the cellular recognition of
389 nanoparticles. Proceedings of the National Academy of Sciences 117, 10492–
390 10499. doi:10.1073/pnas.1919755117
391 Ban, Z., Zhou, Q., Sun, A., Mu, L., Hu, X., 2018. Screening Priority Factors
392 Determining and Predicting the Reproductive Toxicity of Various Nanoparticles.
393 Environmental Science and Technology 52, 9666–9676. doi:
394 10.1021/acs.est.8b02757
395 Ben-Moshe, T., Frenk, S., Dror, I., Minz, D., Berkowitz, B., 2013. Effects of metal
396 oxide nanoparticles on soil properties. Chemosphere 90, 640–646.
397 doi:10.1016/j.chemosphere.2012.09.018
398 Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A.,
399 Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E.,
400 Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J.,
401 Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C.,
402 Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F.,
403 Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L.,
404 Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H.,
405 Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler,
406 B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I.,
407 Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.-X., Loftfield,

E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37, 852–857. doi:10.1038/s41587-019-0209-9

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13, 581–583. doi:10.1038/nmeth.3869

Cammarota, G., Ianiro, G., Ahern, A., Carbone, C., Temko, A., Claesson, M.J., Gasbarrini, A., Tortora, G., 2020. Gut microbiome, big data and machine learning to promote precision medicine for cancer. Nature Reviews Gastroenterology & Hepatology 17, 635–648. doi:10.1038/s41575-020-0327-3

Caspi, R., Billington, R., Keseler, I.M., Kothari, A., Krummenacker, M., Midford, P.E., Ong, W.K., Paley, S., Subhraveti, P., Karp, P.D., 2020. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. Nucleic Acids Research

48, D445–D453. doi:10.1093/nar/gkz862

Chen, Q.-L., Zhu, D., An, X.-L., Ding, J., Zhu, Y.-G., Cui, L., 2019. Does nano silver promote the selection of antibiotic resistance genes in soil and plant? Environment International 128, 399–406. doi:10.1016/j.envint.2019.04.061

Cocozza, C., Perone, A., Giordano, C., Salvatici, M.C., Pignattelli, S., Raio, A., Schaub, M., Sever, K., Innes, J.L., Tognetti, R., Cherubini, P., 2019. Silver nanoparticles enter the tree stem faster through leaves than through roots. Tree Physiology 39, 1251–1261. doi:10.1093/treephys/tpz046

Delgado-Baquerizo, M., Maestre, F.T., Reich, P.B., Jeffries, T.C., Gaitan, J.J., Encinar, D., Berdugo, M., Campbell, C.D., Singh, B.K., 2016. Microbial diversity drives multifunctionality in terrestrial ecosystems. Nature Communications 7, 10541. doi:10.1038/ncomms10541

Douglas, G.M., Maffei, V.J., Zaneveld, J.R., Yurgel, S.N., Brown, J.R., Taylor, C.M., Huttenhower, C., Langille, M.G.I., 2020. PICRUSt2 for prediction of metagenome functions. Nature Biotechnology 38, 685–688. doi:10.1038/s41587-020-0548-6

Fan, X., Xu, J., Lavoie, M., Peijnenburg, W.J.G.M., Zhu, Y., Lu, T., Fu, Z., Zhu, T., Qian, H., 2018. Multiwall carbon nanotubes modulate paraquat toxicity in Arabidopsis thaliana. Environmental Pollution 233, 633–641. doi:10.1016/j.envpol.2017.10.116

Fang, H., Han, L., Zhang, H., Long, Z., Cai, L., Yu, Y., 2018. Dissemination of antibiotic resistance genes and human pathogenic bacteria from a pig feedlot to the

surrounding stream and agricultural soils. Journal of Hazardous Materials 357, 53–62. doi:10.1016/j.jhazmat.2018.05.066

Fayaz, A.M., Balaji, K., Girilal, M., Kalaichelvan, P.T., Venkatesan, R., 2009. Mycobased Synthesis of Silver Nanoparticles and Their Incorporation into Sodium Alginate Films for Vegetable and Fruit Preservation. Journal of Agricultural and Food Chemistry 57, 6246–6252. doi:10.1021/jf900337h

Finkel, O.M., Salas-González, I., Castrillo, G., Conway, J.M., Law, T.F., Teixeira, P.J.P.L., Wilson, E.D., Fitzpatrick, C.R., Jones, C.D., Dangl, J.L., 2020. A single bacterial genus maintains root growth in a complex microbiome. Nature 587, 103–108. doi:10.1038/s41586-020-2778-7

Guerriero, G., Sutera, F.M., Torabi-Pour, N., Renaut, J., Hausman, J.-F., Berni, R., Pennington, H.C., Welsh, M., Dehsorkhi, A., Zancan, L.R., Saffie-Siebert, S., 2021. Phyto-Courier, a Silicon Particle-Based Nano-biostimulant: Evidence from Cannabis sativa Exposed to Salinity. ACS Nano 15, 3061–3069. doi:10.1021/acsnano.0c09488

Guo, Z., Wan, S., Hua, K., Yin, Y., Chu, H., Wang, D., Guo, X., 2020. Fertilization regime has a greater effect on soil microbial community structure than crop rotation and growth stage in an agroecosystem. Applied Soil Ecology 149, 103510. doi:10.1016/j.apsoil.2020.103510

Hendren, C.O., Mesnard, X., Dröge, J., Wiesner, M.R., 2011. Estimating Production Data for Five Engineered Nanomaterials As a Basis for Exposure Assessment. Environmental Science & Technology 45, 2562–2569. doi:10.1021/es103300g

Hibbing, M.E., Fuqua, C., Parsek, M.R., Peterson, S.B., 2010. Bacterial competition: surviving and thriving in the microbial jungle. Nature Reviews Microbiology 8, 15–25. doi:10.1038/nrmicro2259

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. Science 349, 255–260. doi:10.1126/science.aaa8415Ke, M., Li, Y., Qu, Q., Ye, Y., Peijnenburg, W.J.G.M., Zhang, Z., Xu, N., Lu, T., Sun, L., Qian, H., 2020. Offspring toxicity of silver nanoparticles to Arabidopsis thaliana flowering and floral development. Journal of Hazardous Materials 386, 121975. doi:10.1016/j.jhazmat.2019.121975

Ke, M., Ye, Y., Zhang, Z., Gillings, M., Qu, Q., Xu, N., Xu, L., Lu, T., Wang, J., Qian, H., 2021. Synergistic effects of glyphosate and multiwall carbon nanotubes on Arabidopsis thaliana physiology and metabolism. Science of The Total Environment 769, 145156. doi:10.1016/j.scitotenv.2021.145156

Kusi, J., Scheuerman, P.R., Maier, K.J., 2020. Emerging environmental contaminants (silver nanoparticles) altered the catabolic capability and metabolic fingerprinting of microbial communities. Aquatic Toxicology 228, 105633. doi:10.1016/j.aquatox.2020.105633Kielak, A.M., Veen, J.A. van., Kowalchuk, G.A., 2010. Comparative Analysis of Acidobacterial Genomic Fragments from Terrestrial and Aquatic Metagenomic Libraries, with Emphasis on Acidobacteria Subdivision 6. Applied and Environmental Microbiology 76, 6769–6777. doi:10.1128/AEM.00343-10

Lehmann, A., Zheng, W. Rillig, M.C. Soil biota contributions to soil aggregation.

Nature Ecology and Evolution 1, 1828–1835. doi:10.1038/s41559-017-0344-y

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal. doi:10.14806/ej.17.1.200

McShane, H., Sarrazin, M., Whalen, J.K., Hendershot, W.H., Sunahara, G.I., 2012. Reproductive and behavioral responses of earthworms exposed to nano-sized titanium dioxide in soil. Environmental Toxicology and Chemistry 31, 184–193. doi:10.1002/etc.714

Moll, J., Klingenfuss, F., Widmer, F., Gogos, A., Bucheli, T.D., Hartmann, M., van der Heijden, M.G.A., 2017. Effects of titanium dioxide nanoparticles on soil microbial communities and wheat biomass. Soil Biology and Biochemistry 111, 85–93. doi:10.1016/j.soilbio.2017.03.019

Montes de Oca-Vásquez, G., Solano-Campos, F., Vega-Baudrit, J.R., López-Mondéjar, R., Vera, A., Moreno, J.L., Bastida, F., 2020. Organic amendments exacerbate the effects of silver nanoparticles on microbial biomass and community composition of a semiarid soil. Science of The Total Environment 744, 140919. doi:10.1016/j.scitotenv.2020.140919

Nilsson, R.H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P., Tedersoo, L., 2019. Mycobiome diversity: high-throughput sequencing and identification of fungi. Nature Reviews Microbiology 17, 95–109. doi:10.1038/s41579-018-0116-y

Oh, E., Liu, R., Nel, A., Gemill, K.B., Bilal, M., Cohen, Y., Medintz, I.L., 2016. Meta-analysis of cellular toxicity for cadmium-containing quantum dots. Nature Nanotechnology 11, 479–486. doi:10.1038/nnano.2015.338

518 Peng, C., Xu, C., Liu, Q., Sun, L., Luo, Y., Shi, J., 2017. Fate and Transformation of
519 CuO Nanoparticles in the Soil–Rice System during the Life Cycle of Rice
520 Plants. Environmental Science & Technology 51, 4907–4917.
521 doi:10.1021/acs.est.6b05882
522 Qu, Q., Zhang, Z., Peijnenburg, W.J.G.M., Liu, W., Lu, T., Hu, B., Chen, Jianmeng,
523 Chen, Jun, Lin, Z., Qian, H., 2020. Rhizosphere Microbiome Assembly and Its
524 Impact on Plant Growth. Journal of Agricultural and Food Chemistry 68, 5024–
525 5038. doi:10.1021/acs.jafc.0c00073
526 Rajput, V.D., Minkina, T.M., Behal, A., Sushkova, S.N., Mandzhieva, S., Singh, R.,
527 Gorovtsov, A., Tsitsuashvili, V.S., Purvis, W.O., Ghazaryan, K.A., Movsesyan,
528 H.S., 2018. Effects of zinc-oxide nanoparticles on soil, plants, animals and soil
529 organisms: A review. Environmental Nanotechnology, Monitoring &
530 Management 9, 76–84. doi:10.1016/j.enmm.2017.12.006
531 Ramirez, K.S., Knight, C.G., de Hollander, M., Brearley, F.Q., Constantinides, B.,
532 Cotton, A., Creer, S., Crowther, T.W., Davison, J., Delgado-Baquerizo, M.,
533 Dorrepaal, E., Elliott, D.R., Fox, G., Griffiths, R.I., Hale, C., Hartman, K.,
534 Houlden, A., Jones, D.L., Krab, E.J., Maestre, F.T., McGuire, K.L., Monteux,
535 S., Orr, C.H., van der Putten, W.H., Roberts, I.S., Robinson, D.A., Rocca, J.D.,
536 Rowntree, J., Schlaeppi, K., Shepherd, M., Singh, B.K., Straathof, A.L.,
537 Bhatnagar, J.M., Thion, C., van der Heijden, M.G.A., de Vries, F.T., 2018.
538 Detecting macroecological patterns in bacterial communities across
539 independent studies of global soils. Nature Microbiology 3, 189–196.

doi:10.1038/s41564-017-0062-x

Ranjard, L., Dequiedt, S., Prévost-Bouré, N. C., Thioulouse, J., Saby, N.P.A., Lelievre, M., Maron, P.A., Morin, F.E.R., Bispo, A., Jolivet, C., Arrouays, D., Lemanceau, P., 2013. Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. Nature Communications 4, 1434. doi:10.1038/ncomms2431

Ray, P., Lakshmanan, V., Labbé, J.L., Craven, K.D., 2020. Microbe to Microbiome: A Paradigm Shift in the Application of Microorganisms for Sustainable Agriculture. Frontiers in Microbiology 11. doi:10.3389/fmicb.2020.622926

Shin, N.-R., Whon, T.W., Bae, J.-W., 2015. Proteobacteria: microbial signature of dysbiosis in gut microbiota. Trends in Biotechnology 33, 496–503. doi:10.1016/j.tibtech.2015.06.011

Simonin, M., Cantarel, A.A.M., Crouzet, A., Gervaix, J., Martins, J.M.F., Richaume, A., 2018. Negative effects of copper oxide nanoparticles on carbon and nitrogen cycle microbial activities in contrasting agricultural soils and in presence of plants. Frontiers in Microbiology 9, 3102-3102. doi:10.3389/fmicb.2018.03102

Simonin, M., Richaume, A., 2015. Impact of engineered nanoparticles on the activity, abundance, and diversity of soil microbial communities: a review. Environmental Science and Pollution Research 22, 13710–13723. doi:10.1007/s11356-015-4171-xSun, T., Conroy, G., Donner, E., Hungerbuhler, K., Lombi, E., Nowack, B., 2015. Probabilistic modelling of engineered nanomaterial emissions to the environment: a spatio-temporal approach 2, 340-351. doi:10.1039/c5en00004a

562 Tapia-García, E.Y., Hernández-Trejo, V., Guevara-Luna, J., Rojas-Rojas, F.U., Arroyo-
563 Herrera, I., Meza-Radilla, G., Vásquez-Murrieta, M.S., Estrada-de los Santos,
564 P., 2020. Plant growth-promoting bacteria isolated from wild legume nodules
565 and nodules of Phaseolus vulgaris L. trap plants in central and southern Mexico.
566 Microbiological Research 239, 126522. doi:10.1016/j.micres.2020.126522
567 Tian, L., Shen, J., Sun, G., Wang, B., Ji, R., Zhao, L., 2020. Foliar Application of SiO2
568 Nanoparticles Alters Soil Metabolite Profiles and Microbial Community
569 Composition in the Pakchoi (Brassica chinensis L.) Rhizosphere Grown in
570 Contaminated Mine Soil. Environmental Science & Technology 54, 13137–
571 13146. doi:10.1021/acs.est.0c03767
572 Wang, X., Cui, Y., Zhang, X., Ju, W., Duan, C., Wang, Y., Fang, L., 2020. A novel
573 extracellular enzyme stoichiometry method to evaluate soil heavy metal
574 contamination: Evidence derived from microbial metabolic limitation. Science
575 of The Total Environment 738, 139709. doi:10.1016/j.scitotenv.2020.139709
576 Wiegand, S., Jogler, M., Jogler, C., 2018. On the maverick Planctomycetes. FEMS
577 Microbiology Reviews 42, 739–760. doi:10.1093/femsre/fuy029
578 Wright, R.J., Langille, M.G.I., Walker, T.R., 2021. Food or just a free ride? A meta-
579 analysis reveals the global diversity of the Plastisphere. The ISME Journal 15,
580 789–806. doi:10.1038/s41396-020-00814-9
581 Yang, L., Barnard, R., Kuzyakov, Y., Tian, J., 2021. Bacterial communities drive the
582 resistance of soil multifunctionality to land-use change in karst soils. European
583 Journal of Soil Biology 104, 103313. doi:10.1016/j.ejsobi.2021.103313

584 Yang, Y., Rossel, R.A.V., Li, S., Bissett, A., Lee, J., Shi, Z., Behrens, T., Court, L., 2019.
585 Soil bacterial abundance and diversity better explained and predicted with
586 spectro-transfer functions. Soil Biology and Biochemistry 129, 29–
587 38.doi:10.1016/j.soilbio.2018.11.005
588 You, T., Liu, D., Chen, J., Yang, Z., Dou, R., Gao, X., Wang, L., 2018. Effects of metal
589 oxide nanoparticles on soil enzyme activities and bacterial communities in two
590 different soil types. Journal of Soils and Sediments 18, 211–221.
591 doi:10.1007/s11368-017-1716-2
592 You, X., Suo, F., Yin, S., Wang, X., Zheng, H., Fang, S., Zhang, C., Li, F., Li, Y., 2021.
593 Biochar decreased enantioselective uptake of chiral pesticide metalaxyl by
594 lettuce and shifted bacterial community in agricultural soil. Journal of
595 Hazardous Materials 417, 126047. doi:10.1016/j.jhazmat.2021.126047
596 Yu, F., Wei, C., Deng, P., Peng, T., Hu, X., 2021. Deep exploration of random forest
597 model boosts the interpretability of machine learning studies of complicated
598 immune responses and lung burden of nanoparticles. Science Advances 7,
599 eabf4130. doi:10.1126/sciadv.abf4130
600 Yuan, J., Wen, T., Zhang, H., Zhao, M., Penton, C.R., Thomashow, L.S., Shen, Q., 2020.
601 Predicting disease occurrence with high accuracy based on soil macroecological
602 patterns of Fusarium wilt. The ISME Journal 14, 2936–2950.
603 doi:10.1038/s41396-020-0720-5
604 Zhang, H., Huang, M., Zhang, W., Gardea-Torresdey, J.L., White, J.C., Ji, R., Zhao, L.,
605 2020a. Silver Nanoparticles Alter Soil Microbial Community Compositions and

606 Metabolite Profiles in Unplanted and Cucumber-Planted Soils. Environmental
607 Science & Technology 54, 3334–3342. doi:10.1021/acs.est.9b07562
608 Zhang, P., Guo, Z., Zhang, Z., Fu, H., White, J.C., Lynch, I., 2020b. Nanomaterial
609 Transformation in the Soil–Plant System: Implications for Food Safety and
610 Application in Agriculture. Small 16, doi:10.1002/smll.202000705
611 Zhang, Z., Ke, M., Qu, Q., Peijnenburg, W.J.G.M., Lu, T., Zhang, Q., Ye, Y., Xu, P., Du,
612 B., Sun, L., Qian, H., 2018. Impact of copper nanoparticles and ionic copper
613 exposure on wheat (Triticum aestivum L.) root morphology and antioxidant
614 response. Environmental Pollution 239, 689–697.
615 doi:10.1016/j.envpol.2018.04.066
616 Zhou, Z., Wang, C., Jiang, L., Luo, Y., 2017. Trends in soil microbial communities
617 during secondary succession. Soil Biology and Biochemistry 115, 92–99.
618 doi:10.1016/j.soilbio.2017.08.014
619
620

Figure legends

Figure 1 Overview of sample collection. The geographic region information about sampling sites, and the number of independent studies and samples in this study were displayed. A dot represents a single independent experiment.

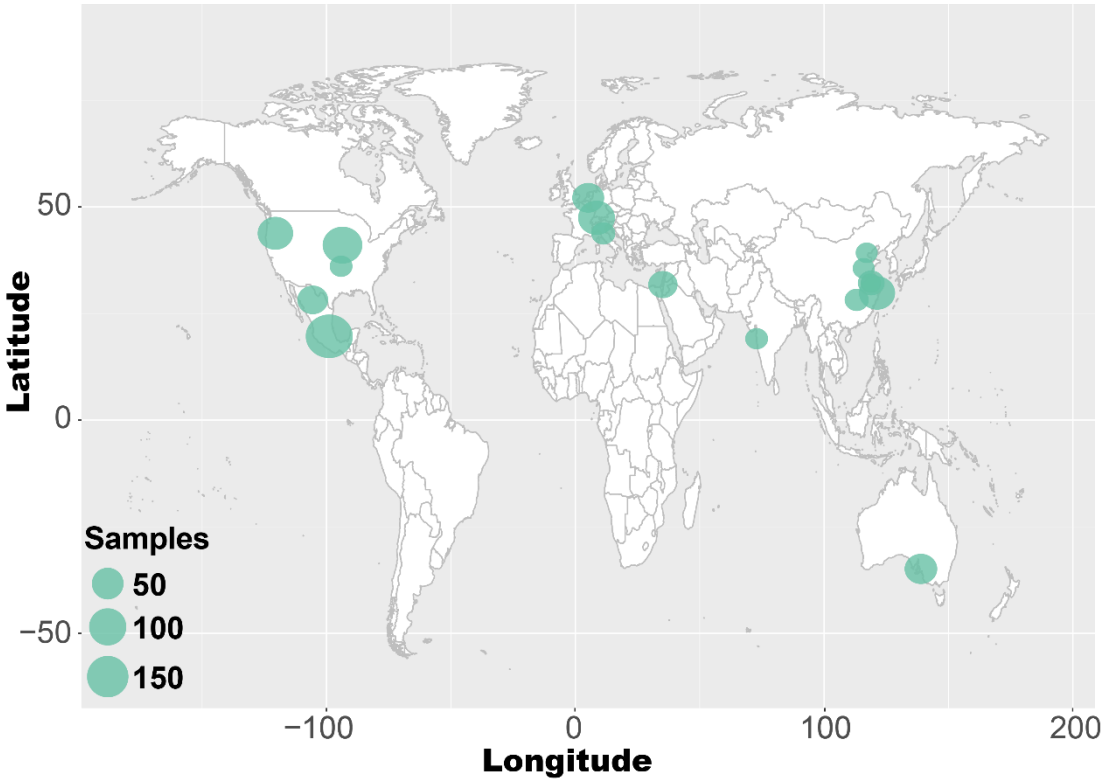
Figure 2 Effects of the NPs on diversity and structure of microbial communities. (a) Alpha diversity of soil bacterial community. The Shannon and Richness indices were calculated with all amplicon sequence variants (ASVs) merged from 365 samples. (b) Principal coordinates analysis (PCoA) with Bray–Curtis dissimilarity performed on the taxonomic (at the genus level) for nano-untreated and nano-treated group. Statstical significance was evaluated via PERMANOVA test. (c) Relative abundance of the 10 most abundant phyla in nano-untreated and nano-treated group. (d) Venn diagram of shared and unique genus numbers observed in nano-untreated and nano-treated soil.

Figure 3 Profile of beneficial and pathogenic soil bacteria following treatment of NPs'. (a), (b), (c) and (d) Differentially beneficial bacteria at genus level after NPs treatment. (e) Differentially pathogenic bacterium at genus level after NPs treatment. The significant difference evaluated by Kruskal-Wallis test ($p < 0.05$).

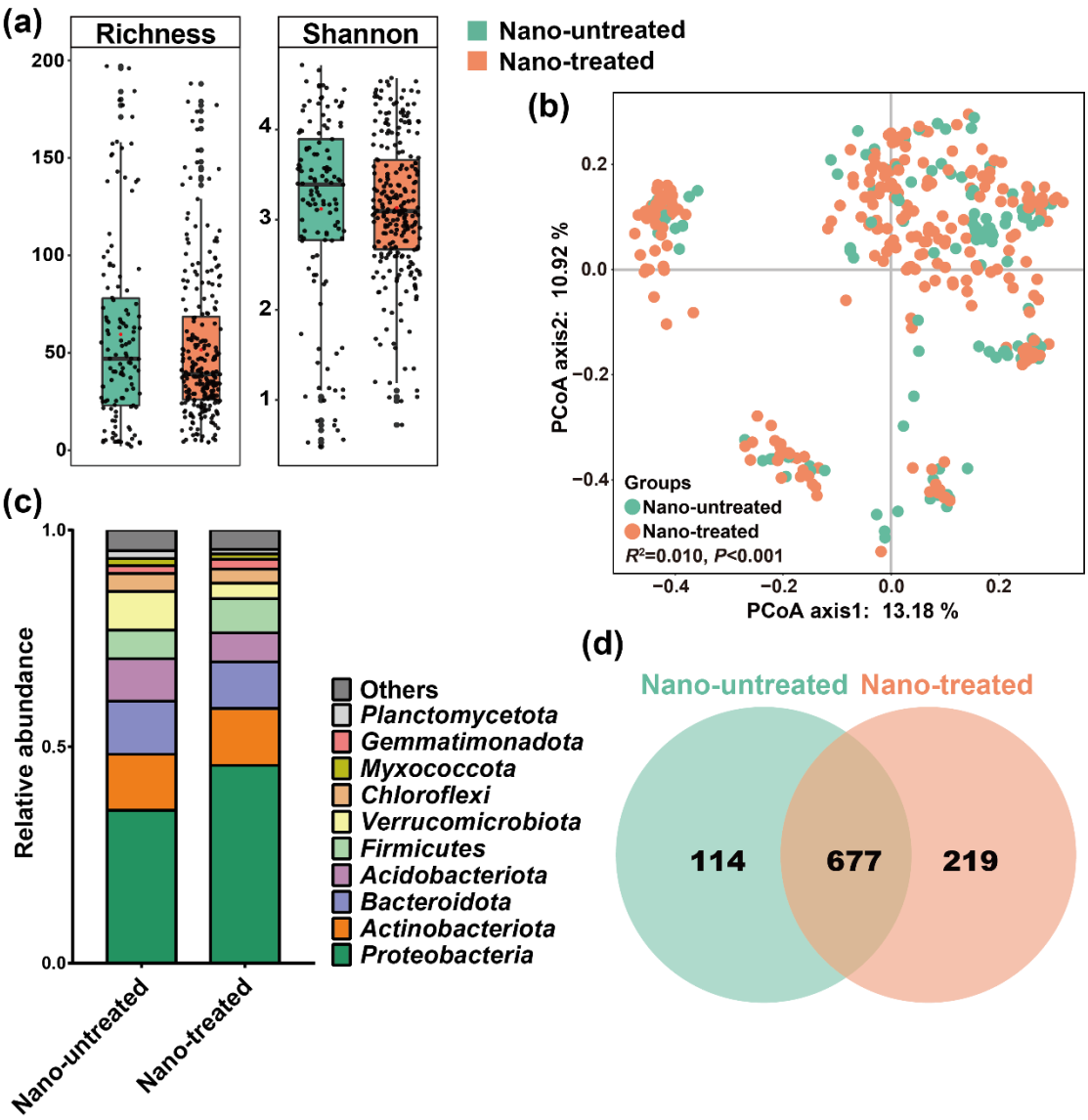
Figure 4 Functional pathways of the microbial communities altered after NPs treatment. The different colors of the outermost layer represented different categories of metabolic pathway functions. The inner circles represented the functional pathways at BioCyc ID level. Pathways showing significant difference ($p < 0.05$) between the nano-untreated and nano-treated group are colored according to direction of change

643 by treatment. The size of circles represented the change-fold of the functional
644 pathway abundance after NPs treatment. The change-fold represented the ratio of
645 nano-treated to nano-untreated functional pathways.
646 **Figure 5** Model building by Random Forest for connection between NPs
647 characteristics and microbial community traits. (a) The values of R^2 in validation
648 model evaluated by tenfold cross-validation. (b) and (c) The contribution of NPs
649 characteristics of Richness and Shannon models on microbial traits. Orange
650 represented the heterogeneity factors, and green represented the NPs factors. The
651 larger the arc of the circle, the more important the factors. The letter abbreviations
652 represent the following: A, Bulk group; B, NPs type; C, NPs category; D, Dispersion
653 medium; E, NPs existence state; F, NPs shape; G, NPs size; H, Zeta potential; I,
654 Dispersion medium pH; J, Exposure time; K, Temperature; L, Extent of
655 agglomeration; M, Crystal structure; N, NPs concentration; O, Soil type; P,
656 Rhizosphere microbiota; Q, Soil pH; R, Soil depth; S, Longitude; T, Latitude; X,
657 Amplicon area. (d) Correlation analysis of the most important NPs factor and
658 microbial diversity. The green dots indicate the changefold of Richness, and the
659 yellow dots mean the changefold of Shannon. (e) Correlation analysis of the most
660 important NPs factor and beneficial and pathogenic bacteria. The green dots indicate
661 the changefold of beneficial bacteria, and the yellow dots mean the changefold of
662 pathogenic bacterium. The shapes of the dots represent different bacteria. The
663 changefold represents the ratio of the value of nano-treated to nano-untreated.

Figure 1



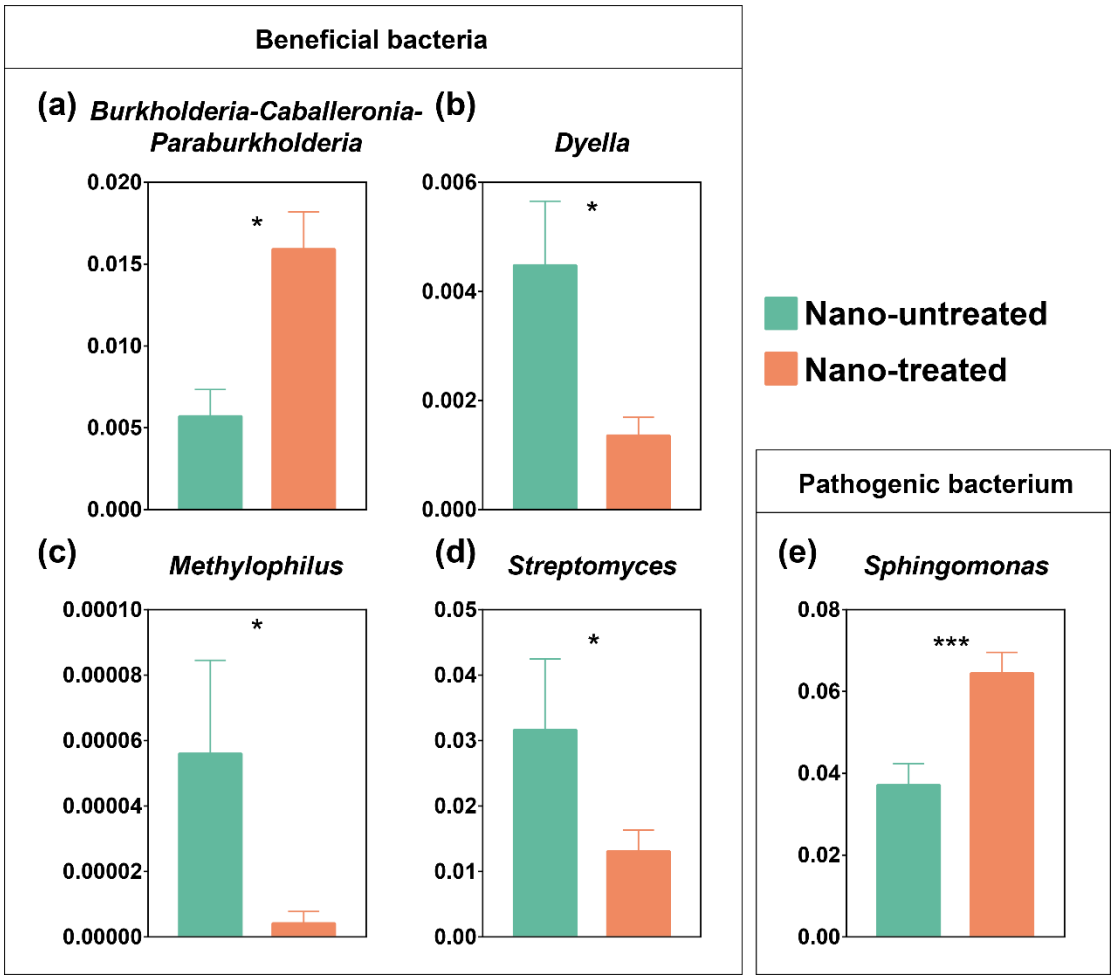
667 **Figure 2**



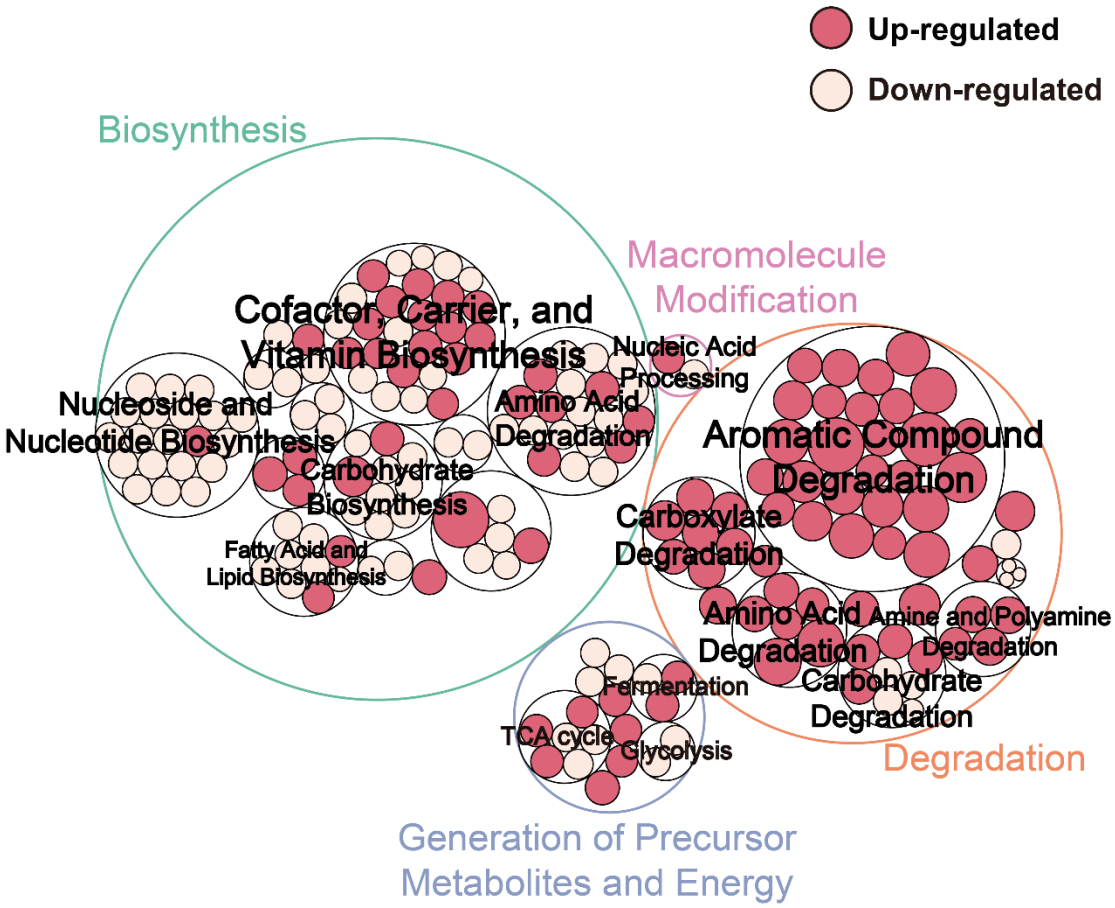
668

669

Figure 3


Beneficial bacteria
(a) Burkholderia-Caballeronia-Paraburkholderia
0.020
0.015
0.010
0.005
0.000
*
(b) Dyella
0.006
0.004
0.002
0.000
*
(c) Methylophilus
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000
*
(d) Streptomyces
0.05
0.04
0.03
0.02
0.01
0.00
*
Nano-untreated
Nano-treated
Pathogenic bacterium
(e) Sphingomonas
0.08
0.06
0.04
0.02
0.00
***

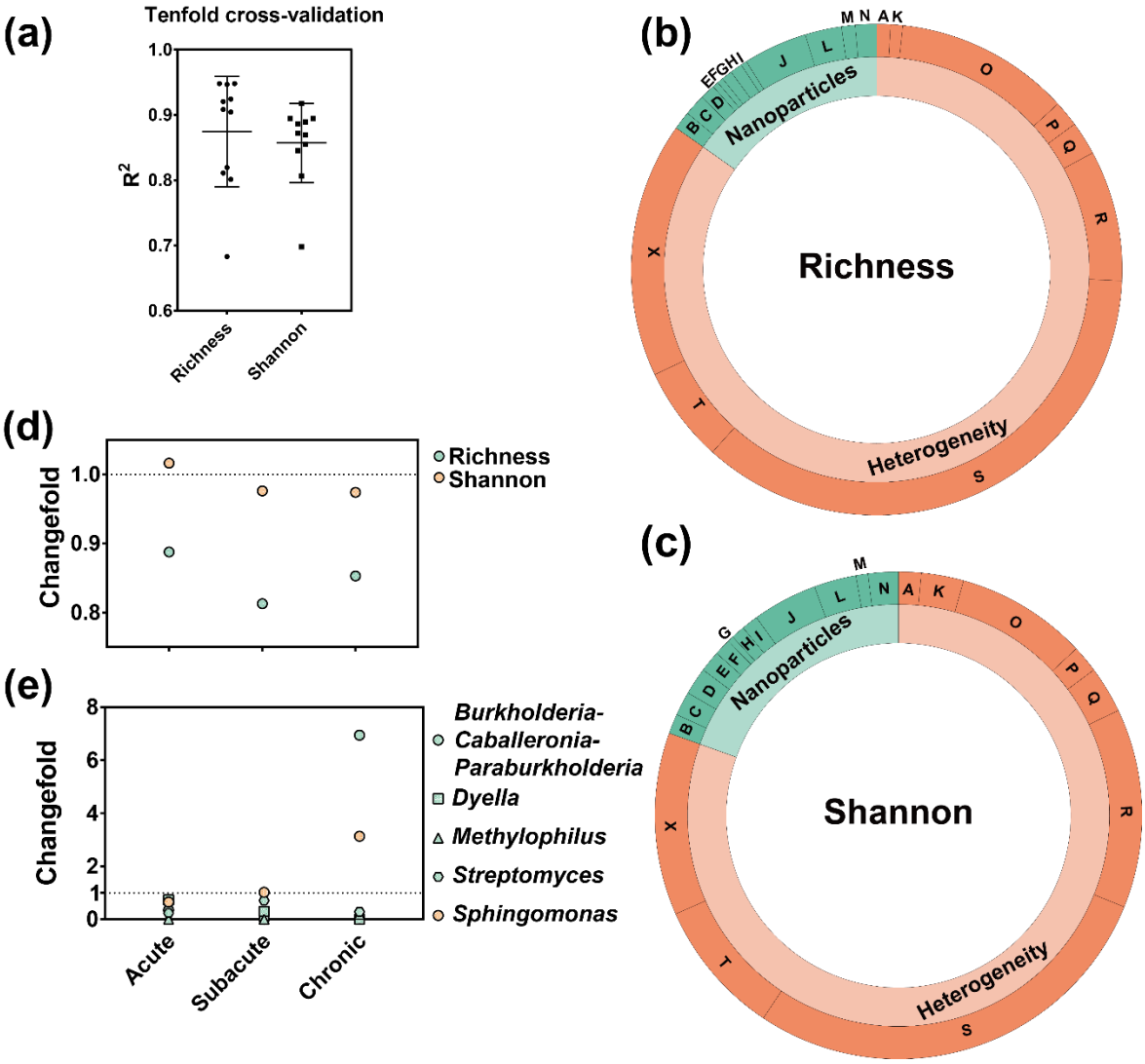
673 **Figure 4**



674

675

676 **Figure 5**



677