Novel insights into the effect of arbuscular mycorrhizal fungi inoculation in soils under long-term biosolids application: emphasis on antibiotic and metal resistance genes, and mobile genetic elements

Tao Sun1,2, Pierre Delaplace2, Guihua Li1, Anina James1, Junting Pan1\*, Jianfeng Zhang1,3\*

1 State Key Laboratory of Efficient Utilization of Arable Land in China, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China

2 Gembloux Agro Bio Tech, University of Liège, 5030, Belgium

3 Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Jilin Province, Changchun 130112, China

**\*Corresponding author**

Prof. Junting Pan ([panjunting@caas.cn](mailto:panjunting@caas.cn))

Prof. Jianfeng Zhang (zhangjianfeng@caas.cn)

# Supplementary materials

**Material and methods**

**DNA extraction and sequencing**

Total DNA was extracted from soil samples using the E.Z.N.A.® Soil DNA Kit (Omega Bio-tek, Norcross, GA, U.S.) according to the manufacturer’s protocols. PCR amplifications targeted different regions: the V4 region for bacteria using primer pair 515F (5'-GTGCCAGCMGCCGCGGTAA-3') / 806R (5'-GGACTACHVGGGTWTCTAAT-3') (Walters et al., 2016), and the ITS region for fungi using primers ITS3F (5'-GCATCGATGAAGAACGCAGC-3') / ITS4R (5'-GCATCGATGAAGAACGCAGC-3') (Op De Beeck et al., 2014) synthesized by Invitrogen. PCR reactions were performed in triplicate 20 μL mixture containing 4 μL of 5 × FastPfu Buffer, 2 μL of 2.5 mM dNTPs, 0.8 μL of each primer (5 μM), 0.4 μL of FastPfu Polymerase, and 10 ng of template DNA. Amplicons were extracted from 2% agarose gels and purified using the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA, U.S.) according to the manufacturer’s instructions. The pooled DNA product was used to construct an Illumina Pair-End library following Illumina’s genomic DNA library preparation procedure. Then the amplicon library was paired-end sequenced (2 × 300) on an NGS platform (Shanghai BIOZERON Biotech. Co., Ltd) according to the standard protocols.

Raw fastq files were first demultiplexed using Trimmomatic (Bolger et al., 2014) and in-house perl scripts according to the barcode sequence information for each sample with the following criteria: (i) The 300bp reads were truncated at any site receiving an average quality score <20 over a 10 bp sliding window, discarding the truncated reads that were shorter than 50bp. (ii) Exact barcode matching, 2 nucleotide mismatch in primer matching, reads containing ambiguous characters were removed. (iii) Only sequences that overlap longer than 10 bp were assembled according to their overlap sequence. Reads which could not be assembled were discarded. The sequencing reads, acquired from paired-end sequencing, were initially sorted into individual samples using unique barcodes. Post sorting, the reads underwent merging and were subsequently subjected to denoising and chimera filtering through DADA2. Taxonomic assignment for bacteria utilized the SILVA v138 database (Green et al., 2022), while the UNITE v8.0 database (Eshaghi et al., 2021) was employed for fungi. For detailed insights into the bioinformatic analyses of the data, refer to the Supporting Materials section, which contains comprehensive information about these procedures.

Metagenomic shotgun sequencing libraries were constructed and sequenced at Shanghai Biozeron Biological Technology Co. Ltd. Briefly, for each sample, the TruSeq DNA Library Preparation kit (catalog no: FC-121-2001, Illumina, USA) was used to construct sequencing libraries and the concentration of all libraries was measured by High Sensitivity Double Stranded DNA kit on a Qubit Fluorometer (Thermo Fisher Scientific). All samples were sequenced on the NGS instrument with pair-end 150bp (PE150) mode. Raw sequence reads underwent quality trimming using Trimmomatic (Bolger et al., 2014) to remove adaptor contaminants and low-quality reads. Reads passing quality control were then mapped against the human genome (NCBI) by the BWA mem algorithm (Li et al., 2015). The reads removing host-genome contaminations and low-quality data were called clean reads and used for further analysis.

Clean sequence reads were generated as a set of contigs for each sample using MegaHit with “--min-contig-len 500” parameters (Li et al., 2015). The open reading frames (ORFs) of assembled contigs were predicted using Prodigal (Hyatt et al., 2010), and all ORFs were generated into a set of unique genes after clustering using CD-HIT (Fu et al., 2012, parameters: -n 9 -c 0.95 -G 0 -M 0 -d 0 -aS 0.9 -r 1). The longest sequence of each cluster was considered the representative sequence of each gene in the unique-gene set. To calculate the gene abundance within total samples, Salmon software (Patro et al., 2017) was applied to get the read numbers for each gene. Finally, the gene abundance was calculated using the following formulas:

Ng, the average number of reads mapped to the g gene; Lg, the number of nucleotides in the g gene; The index j stands for the set of all genes determined in a catalog, and g is an index indicating a particular gene. The unique-gene set was first translated into protein sequences and then searched against the eggNOG database (Huerta-Cepas et al., 2016), NR database, SARG database (Yin et al., 2018), NCycle database (Tu et al., 2019), SCycle database (Yu et al., 2020), PCycle database (Zeng et al., 2022), ASCycle database, PHI database (Urban et al., 2022), VFDB database (Liu et al., 2022), MRG database, BRG database, and Plastic database (Gambarini et al., 2022) using DIAMOND (Buchfink et al., 2015) to identify the gene functions with the following filter parameters (--evalue 0.00001 --id 70). CAZymes were annotated using HMMER (v.3.2.1) to match protein sequences to entries in the hidden Markov model (HMM) libraries of CAZyme families downloaded from the CAZy database (Lombard et al., 2014). KEGG ortholog annotation was performed by KofamScan (<https://www.genome.jp/tools/kofamkoala/>) with the HMMSEARCH package (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmsearch>).

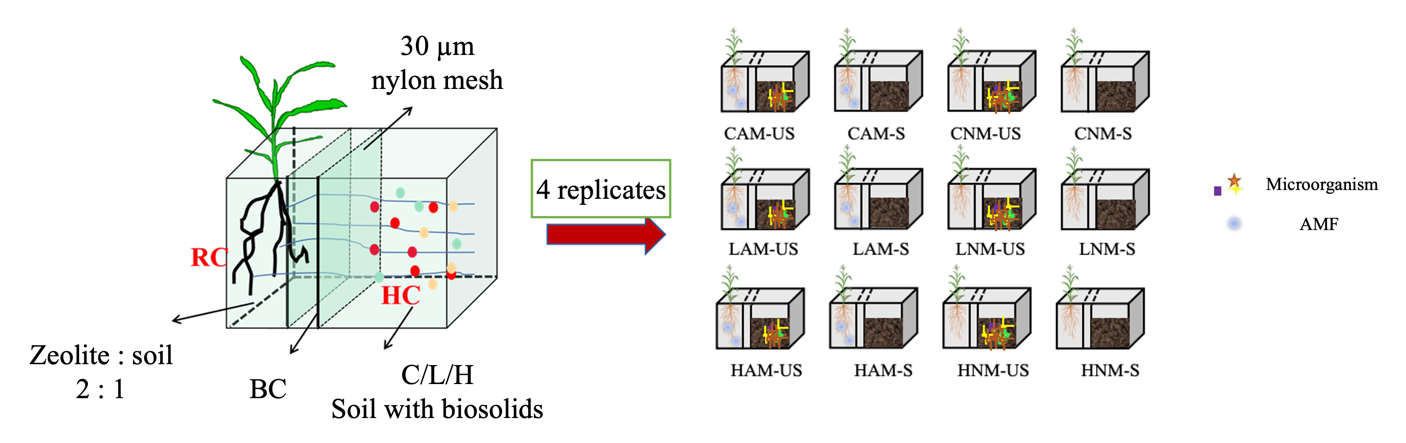
**Supplementary figures**

Fig. S1. Schematic representation of pot experiment setup

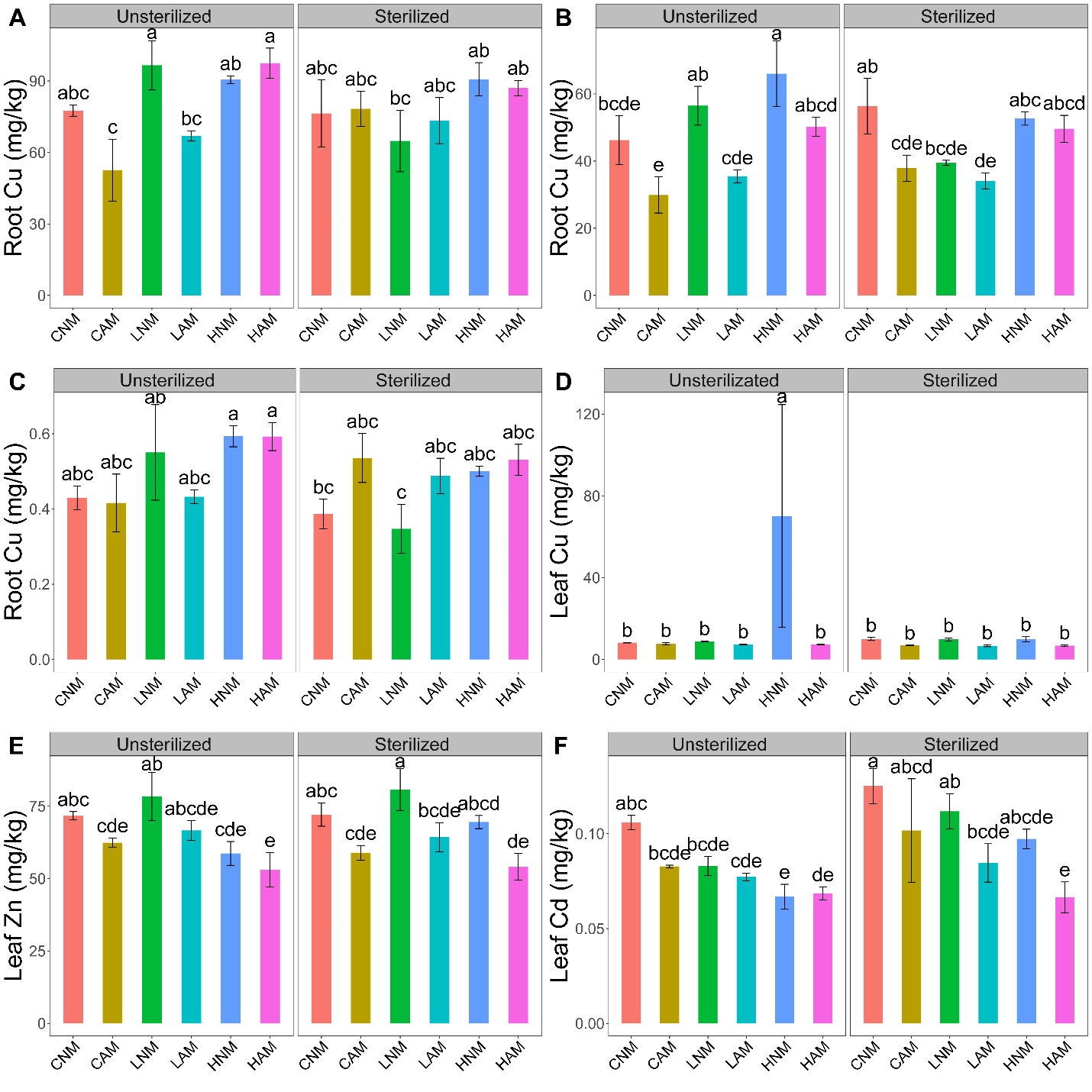


Fig. S2. The effect of AMF inoculation on plant root (A-C) and leaf (D-F) heavy metals contents under different levels of biosolids application.

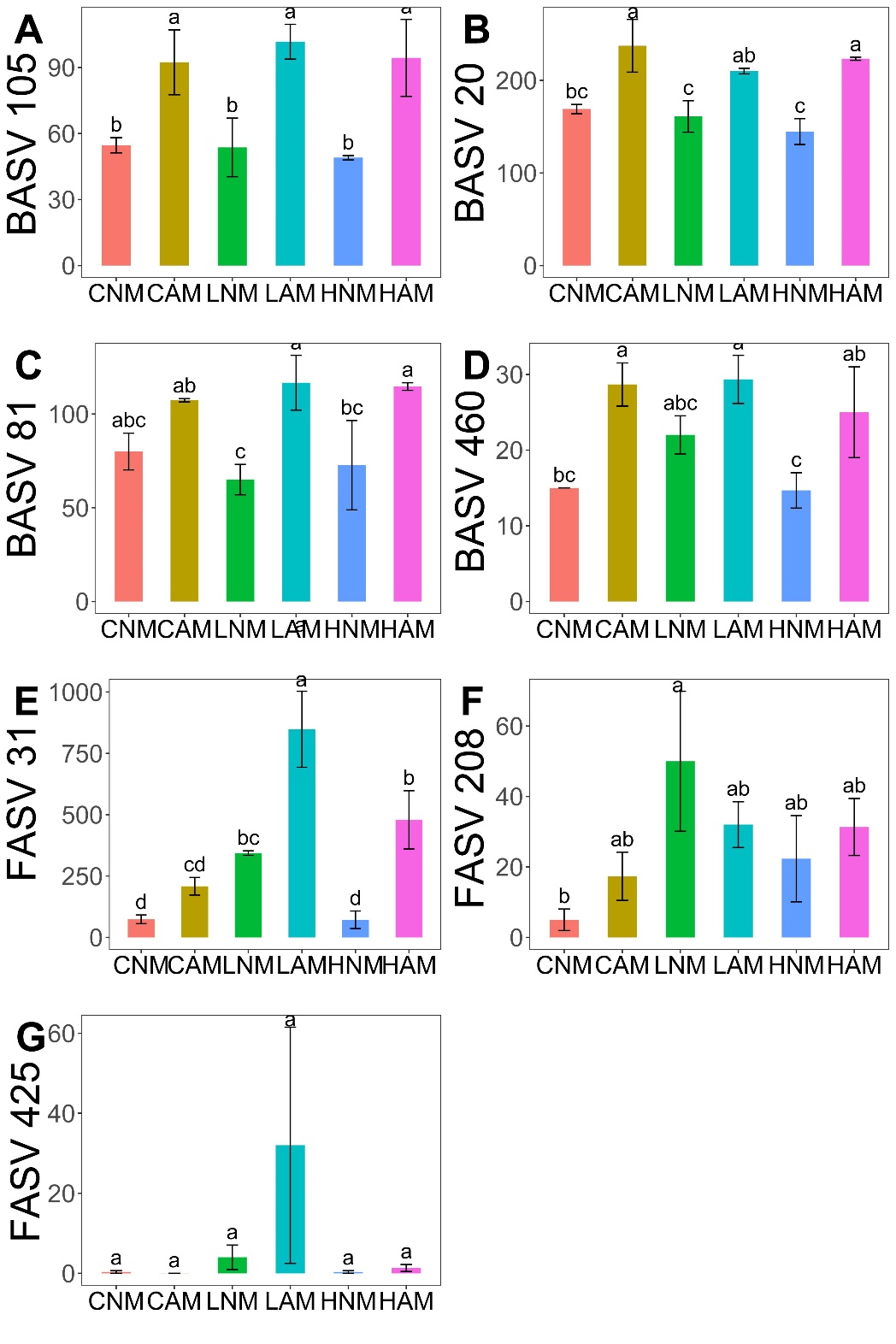


Fig. S3. The effect of AMF inoculation on the abundance of keystone taxa abundance under different levels of biosolids application.

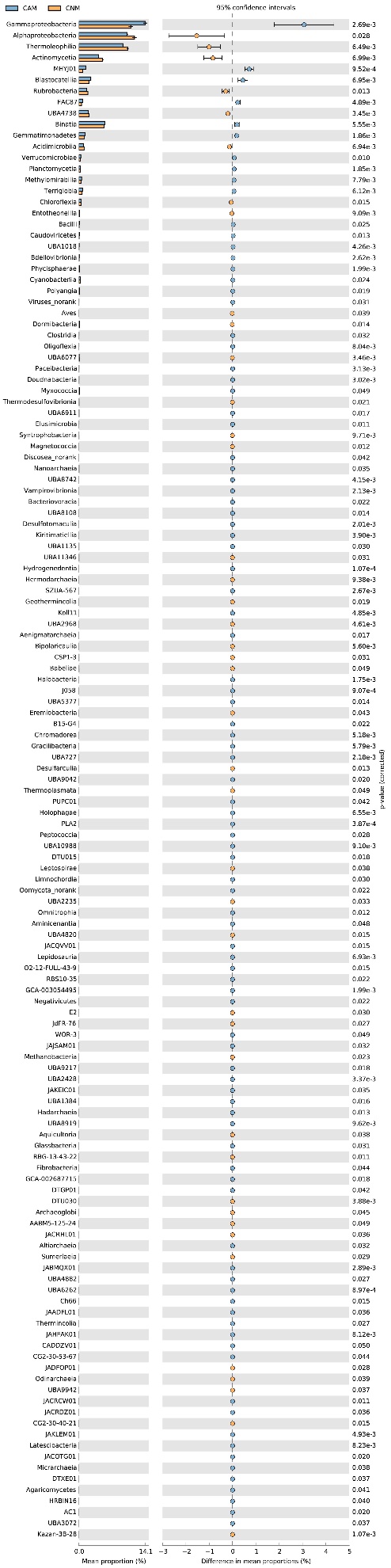


Fig. S4. Comparison of microbial taxa with AMF inoculation and without AMF inoculation in control treatment, visualized using 95 % confidence intervals for the difference in mean proportions.

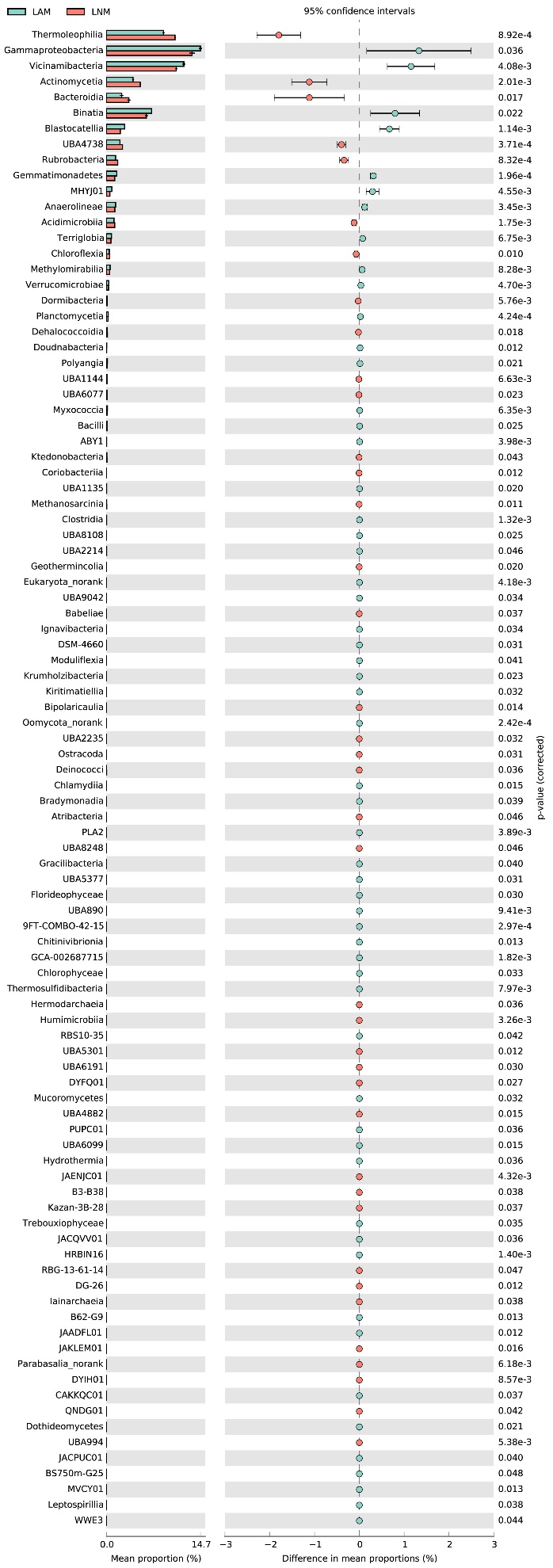


Fig. S5. Comparison of microbial taxa with AMF inoculation and without AMF inoculation in low levels of biosolids application, visualized using 95% confidence intervals for the difference in mean proportions.

**Supplementary tables**

Table S1. Intensity of arbuscular mycorrhizal colonization (M%) and frequency of arbuscules (A%) in maize roots and hyphae length density (HLD) in soil under different treatments.

|  |  |  |  |
| --- | --- | --- | --- |
|  | M | A | HLD |
| CAM-S | 18.79 ± 10.55b | 19.39 ± 10.03b | 1.44 ± 0.07e |
| LAM-S | 41.05 ± 13.88a | 41.05 ± 13.88a | 3.67 ± 0.16a |
| HAM-S | 7.93 ± 1.51b | 8.29 ± 1.46b | 2.37 ± 0.03c |
| CAM-US | 19.72 ± 4.8b | 20.53 ± 5.66b | 2.02 ± 0.06d |
| LAM-US | 41.84 ± 9.27a | 45.01 ± 5.63a | 3.21 ± 0.02b |
| HAM-US | 17.75 ± 8.51b | 19.32 ± 10.79b | 1.07 ± 0.04f |

Table S2. Results of three-way ANOVA investigating the effects of AMF inoculation, soil microorganisms, and biosolids application on plant biomass and heavy metal contents in plants and soil

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | AMF | Biosolids | Microbiome | AMF\*Biosolids | AMF\*Microbiome | Biosolids\*Microbiome | AMF\*Biosolids\*Microbiome | Residuals |
| Plant biomass | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 68.33 | 31.81 | 0.76 | 1.55 | 1.29 | 0.38 | 0.34 |  |
| *p* | < 0.001\*\*\* | < 0.001\*\*\* | 0.39 | 0.23 | 0.26 | 0.69 | 0.71 |  |
| Soil Cu | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 12.67 | 1030.23 | 4.6 | 5.29 | 2.39 | 0.06 | 0.38 |  |
| *p* | 0.002\*\* | < 0.001\*\*\* | 0.04\* | 0.01\* | 0.14 | 0.94 | 0.69 |  |
| Soil Zn | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 53.09 | 8922.53 | 8.48 | 11.88 | 1.08 | 3.51 | 1.45 |  |
| *p* | < 0.001\*\*\* | < 0.001\*\*\* | 0.007\*\* | < 0.001\*\*\* | 0.31 | 0.05\* | 0.25 |  |
| Soil Cd | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 0.09 | 315.51 | 2.65 | 0.91 | 1.02 | 0.91 | 0.26 |  |
| *p* | 0.77 | < 0.001\*\*\* | 0.12 | 0.41 | 0.322 | 0.41 | 0.77 |  |
| Soil Pb | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 0.006 | 68.92 | 0.1 | 1.17 | 0.34 | 2.44 | 0.45 |  |
| *p* | < 0.001\*\*\* | < 0.001\*\*\* | 0.75 | 0.32 | 0.56 | 0.1 | 0.64 |  |
| Soil Hg | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 0.96 | 6.07 | 0.37 | 0.81 | 0.74 | 1.34 | 1.6 |  |
| *p* | 0.33 | 0.007\*\* | 0.54 | 0.45 | 0.39 | 0.27 | 0.22 |  |
| Root Cu | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 1.39 | 4.55 | 0.12 | 0.53 | 2.46 | 1.63 | 1.63 |  |
| *p* | 0.25 | 0.02 \* | 0.74 | 0.59 | 0.13 | 0.22 | 0.22 |  |
| Root Zn | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 14.4 | 5.8 | 0.46 | 0.43 | 1.56 | 2.67 | 0.61 |  |
| *p* | < 0.001\*\*\* | 0.008\*\* | 0.5 | 0.66 | 0.22 | 0.09 | 0.55 |  |
| Root Cd | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 0.66 | 3.43 | 0.97 | 0.22 | 3.87 | 0.98 | 0.74 |  |
| *p* | 0.43 | 0.05\* | 0.34 | 0.8 | 0.06 | 0.39 | 0.49 |  |
| Leaf Cu | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 1.39 | 0.96 | 0.89 | 0.96 | 0.77 | 0.96 | 0.96 |  |
| *p* | 0.25 | 0.4 | 0.35 | 0.4 | 0.39 | 0.4 | 0.4 |  |
| Leaf Zn | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 14.58 | 6.37 | 0.22 | 0.12 | 0.96 | 0.52 | 0.08 |  |
| *p* | < 0.001\*\*\* | 0.006\*\* | 0.64 | 0.89 | 0.34 | 0.6 | 0.92 |  |
| Leaf Cd | Df | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 6 |
| F | 7.32 | 6.31 | 6.55 | 0.16 | 1.81 | 0.05 | 0.5 |  |
| *p* | 0.01\* | 0.006\*\* | 0.2\* | 0.85 | 0.19 | 0.95 | 0.61 |  |

Table S3. Results of two-way ANOVA investigating the effects of AMF inoculation and biosolids application on soil bacterial and fungal Shannon index.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | AMF | Biosolids | AMF\*Biosolids | Residuals |
| Bacterial Shannon index | Df | 1 | 2 | 2 | 12 |
| F | 0.37 | 0.32 | 3.93 |  |
| *p* | 0.56 | 0.73 | 0.05\* |  |
| Fungal Shannon index | Df | 1 | 2 | 2 | 12 |
| F | 0.05 | 0.64 | 2.52 |  |
| *p* | 0.83 | 0.55 | 0.12 |  |

Table S4. PERMANOVA analysis investigating the effects of AMF inoculation and biosolids application on soil bacterial and fungal communities

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Bacterial community | | | Fungal community | | |
|  | F | R2 | *p* | F | R2 | *p* |
| AMF | 1 | 1.24 | 0.22 | 1 | 0.04 | 0.49 |
| Biosolids | 2 | 3.92 | < 0.001\*\*\* | 2 | 3.31 | < 0.001 \*\*\* |
| AMF\*Biosolids | 2 | 1.3 | 0.16 | 2 | 0.11 | 0.53 |
| Residuals | 12 |  |  | 12 |  |  |

Table S5. Keystone taxa identified in soil microbial networks and taxonomic classification

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ASVID | Type | Phylum | Class | Order | Family | Genus |  |
| BASV105 | Connectors | *Proteobacteria* | *Alphaproteobacteria* | *Sphingomonadales* | *Sphingomonadaceae* | *Sphingomonas* | |
| BASV20 | Connectors | *Proteobacteria* | *Alphaproteobacteria* | *Azospirillales* | *Azospirillaceae* | *Skermanella* | |
| BASV460 | Connectors | *Actinobacteriota* | *Actinobacteria* | *Micrococcales* | *Cellulomonadaceae* | *Cellulomonas* | |
| BASV81 | Module hubs | *Acidobacteriota* | *Vicinamibacteria* | *Vicinamibacterales* | *Vicinamibacteraceae* | *Vicinamibacter* | |
| FASV208 | Connectors | *Ascomycota* | *Sordariomycetes* | *Sordariales* | *Lasiosphaeriaceae* | *Apiosordaria* | |
| FASV31 | Connectors | *Ascomycota* | *Sordariomycetes* | *Sordariales* | *Lasiosphaeriaceae* | *Apiosordaria* | |
| FASV425 | Connectors | *Ascomycota* | *Sordariomycetes* | *Myrmecridiales* | *Myrmecridiaceae* | *Myrmecridium* |  |

Table S6. Results of two-way ANOVA investigating the effects of AMF inoculation and biosolids application on soil microbial network parameters

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | AMF | Biosolids | AMF\*Biosolids | Residuals |
| Node number | Df | 1 | 2 | 2 | 12 |
| F | 28.65 | 9.05 | 3.15 |  |
| *p* | < 0.001\*\*\* | 0.004\*\* | 0.08 |  |
| Edge number | Df | 1 | 2 | 2 | 12 |
| F | 14 | 11 | 2.59 |  |
| *p* | 0.002\*\* | 0.002\*\* | 0.12 |  |
| Average degree | Df | 1 | 2 | 2 | 12 |
| F | 47.26 | 6.61 | 1.67 |  |
| *p* | < 0.001\*\*\* | 0.02\* | 0.23 |  |
| Modularity | Df | 1 | 2 | 2 | 12 |
| F | 124,71 | 22.62 | 1.14 |  |
| *p* | < 0.001\*\*\* | < 0.001\*\*\* | 0.35 |  |

**References**

Bolger, A. M., M. Lohse, et al. (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics 30(15): 2114-2120.

Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015 Jan;12(1):59-60. doi: 10.1038/nmeth.3176. Epub 2014 Nov 17.

Fu, L., B. Niu, et al. (2012). "CD-HIT: accelerated for clustering the next-generation sequencing data." Bioinformatics 28(23): 3150-3152.

Gambarini V, Pantos O, Kingsbury JM, Weaver L, Handley KM, Lear G. PlasticDB: a database of microorganisms and proteins linked to plastic biodegradation. Database (Oxford). 2022 Mar 1; 2022: baac008. doi: 10.1093/database/baac008. PMID: 35266524; PMCID: PMC9216477.

Huerta-Cepas, J., D. Szklarczyk, et al. (2016). "eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences." Nucleic Acids Res 44(D1): D286-293.

Hyatt, D., G. L. Chen, et al. (2010). "Prodigal: prokaryotic gene recognition and translation initiation site identification." BMC Bioinformatics 11: 119.

Li, D., C. M. Liu, et al. (2015). "MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph." Bioinformatics 31(10): 1674-1676.

Liu B, Zheng D, Zhou S, Chen L, Yang J. VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic Acids Res. 2022 Jan 7;50(D1): D912-D917. doi: 10.1093/nar/gkab1107. PMID: 34850947; PMCID: PMC8728188.

Lombard, V., H. Golaconda Ramulu, et al. (2014). "The carbohydrate-active enzymes database (CAZy) in 2013." Nucleic Acids Res 42(Database issue): D490-495.

Ningqi Wang, Yarong Li, Shuwen Han, Yaozhong Zhang, Jun Yang, Zheng Yin, Cong Deng, Zijing Liu, Yinhang Wu, Wei Wu, Wei Xue, Tianjie Yang, Yangchun Xu, Qirong Shen, Gaofei Jiang, Zhong Wei. CFViSA: A comprehensive and free platform for visualzation and statistics in omics-data. Computers in Biology and Medicine. 2024.

Patro, R., G. Duggal, et al. (2017). "Salmon provides fast and bias-aware quantification of transcript expression." Nat Methods 14(4): 417-419.

Schloss PD et al (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541.

Tu Q, Lin L, Cheng L, Deng Y, He Z. NCycDB: a curated integrative database for fast and accurate metagenomic profiling of nitrogen cycling genes. Bioinformatics. 2019 Mar 15;35(6):1040-1048. doi: 10.1093/bioinformatics/bty741. PMID: 30165481.

Urban M, Cuzick A, Seager J, Wood V, Rutherford K, Venkatesh SY, Sahu J, Iyer SV, Khamari L, De Silva N, Martinez MC, Pedro H, Yates AD, Hammond-Kosack KE. PHI-base in 2022: a multi-species phenotype database for Pathogen-Host Interactions. Nucleic Acids Res. 2022 Jan 7;50(D1): D837-D847. doi: 10.1093/nar/gkab1037. PMID: 34788826; PMCID: PMC8728202.

Xie F, Jin W, Si H, et al. An integrated gene catalog and over 10,000 metagenome-assembled genomes from the gastrointestinal microbiome of ruminants. Microbiome. 2021 Jun 12;9(1):137. doi: 10.1186/s40168-021-01078-x. Erratum in: Microbiome. 2022..

Yin X, Jiang XT, Chai B, Li L, Yang Y, Cole JR, Tiedje JM, Zhang T. ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. Bioinformatics. 2018 Jul 1;34(13):2263-2270. doi: 10.1093/bioinformatics/bty053. PMID: 29408954.

Yu, X., Zhou, J., Song, W., Xu, M., He, Q., Peng, Y., Tian, Y., Wang, C., Shu, L., Wang, S., Yan, Q., Liu, J., Tu, Q., & He, Z. (2020). SCycDB: A curated functional gene database for metagenomic profiling of sulphur cycling pathways. Molecular Ecology Resources, 21.

Zeng J, Tu Q, Yu X, Qian L, Wang C, Shu L, Liu F, Liu S, Huang Z, He J, Yan Q, He Z. PCycDB: a comprehensive and accurate database for fast analysis of phosphorus cycling genes. Microbiome. 2022 Jul 4;10(1):101. doi: 10.1186/s40168-022-01292-1. PMID: 35787295; PMCID: PMC9252087.