Microbial OMICS, an asset to accelerate sustainability in agricultural and environmental microbiology

Edited by
Adolphe Zeze and Mohamed Hijri

Published in
Frontiers in Genetics
Frontiers in Ecology and Evolution

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public – and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Microbial OMICS, an asset to accelerate sustainability in agricultural and environmental microbiology

Topic editors

Adolphe Zeze — Félix Houphouët-Boigny National Polytechnic Institute,
Côte d'Ivoire

Mohamed Hijri — Université de Montréal, Canada

Table of
contents

Editorial: Microbial OMICS, an asset to accelerate sustainability in agricultural and environmental microbiology

Adolphe Zézé[1]* and Mohamed Hijri[2,3]

[1]Laboratoire de Microbiologie, Biotechnologies et Bioinformatique, Unité Mixte de Recherche et d'Innovation Sciences Agronomiques et Procédés de Transformation, Institut National Polytechnique Félix-Houphouet-Boigny, Yamoussoukro, Côte d'Ivoire, [2]Département de Sciences Biologiques, Institut de Recherche en Biologie Végétale, Université de Montréal, Montréal, QC, Canada, [3]African Genome Center, University Mohammed VI Polytechnic (UM6P), Ben Guerir, Morocco

Editorial on the Research Topic
Microbial OMICS, an asset to accelerate sustainability in agricultural and environmental microbiology

Microorganisms provide numerous ecosystem services to humans, allowing natural systems to benefit from a genetic reservoir essential for their fundamental functioning and sustainability (Rodrigues-Filho et al., 2023). They also play pivotal roles in the functioning of global ecosystems (Li et al., 2023; Liu et al., 2019). The advancement of microbial OMICS has enabled accurate elucidation of microbial functions across diverse ecosystems (Kaur et al., 2023), resulting in the identification and characterization of numerous provisioning services, biological processes, and supporting services (Beale et al., 2022). Moreover, microbial OMICS research has contributed to the advancement of applied biotechnologies and innovations in domains such as food security, agriculture, aquaculture, human health, animal health, and environmental protection (Su et al., 2024; Natnan et al., 2021; Chen et al., 2023; Goossens et al., 2022). The knowledge generated by microbial OMICS technologies, along with the development of related applied biotechnology, represents significant progress in sustainable agriculture and environmental management (Hijri, 2023).

The use of OMICS technologies enables the monitoring of soil health and productivity within agroecosystems. It has been reported that the utilization of specific microbes within agroecosystems, along with the type of soil management, can influence soil microbial community structure and function, consequently impacting soil health and crop productivity (Bertola et al., 2021). In a review, Adedayo and Babalola highlighted the significance of plant genomics in promoting the bioeconomy and the potential offered by advances in plant breeding techniques. Along with persisting challenges of underdevelopment and shifts in average weather conditions, the issue of food scarcity remains unresolved. Consequently, advances in crop production offer potential solutions to tackle these challenges. Furthermore, their review discussed the benefits of beneficial
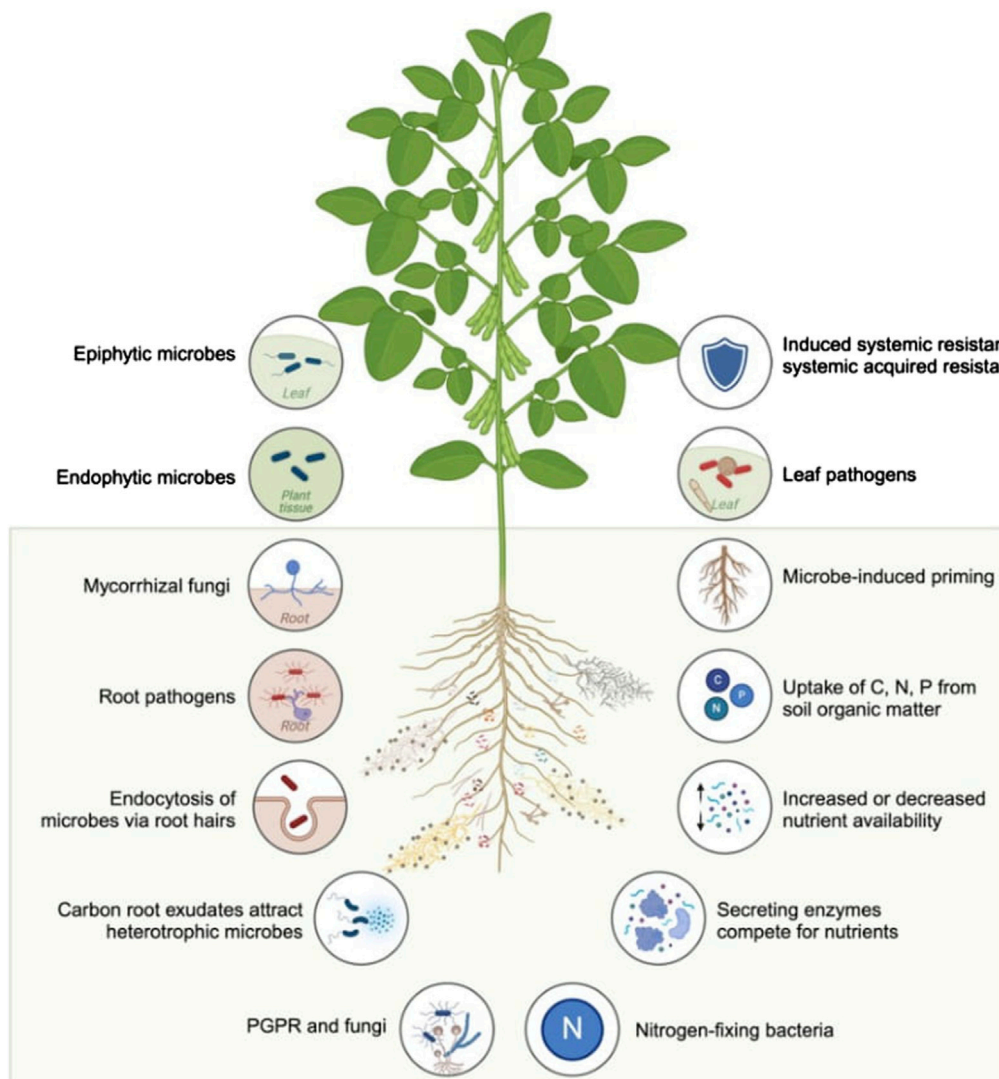
FIGURE 1
An overview of studies investigating plant-microbe interactions in soybean, a model crop, illustrating plant microbiota and microbial functions enabled by OMICS technologies. (Generated using BioRender.com).

microbes in promoting crop growth and outlined the use of OMICS techniques to characterize plant-microbe interactions (Figure 1).

Using 16S rRNA gene metabarcoding targeting bacterial communities, Duan et al. demonstrated that the cultivation of *Morchella esculenta* in fallow paddy fields and drylands can alter the diversity of soil bacteria, affecting soil health and rice productivity. Their findings suggested that *M. esculenta* cultivation may enhance bioavailability of soil phosphorus and potassium in paddy fields, and potassium in dryland soil. *M. esculenta* cultivation had a modest impact on alpha diversity, and it influenced the abundance of certain genera of soil bacteria. Their functional annotation analysis indicated that *M. esculenta* cultivation might reduce methane production potential in paddy field soil and enhance nitrogen cycling in dryland soil. They performed a network analysis and correlation analysis, which revealed that *Gemmatimonas, Bryobacter*, and *Anaeromyxobacter* were key bacterial genera regulating soil chemical properties in paddy field soil under *M. esculenta* cultivation, while *Bryobacter,*

Bacillus, Streptomyces, and *Paenarthrobacter* were key taxa associated with potassium accumulation in dryland soil.

Arunrat et al. compared the variability of soil bacterial communities in maize monoculture and the fallow phase of rotational shifting cultivation fields in Northern Thailand. They selected a continuous 5-year fallow field (CF-5Y) and a continuous 5-year maize cultivation field (M-5Y) with similar microclimate, topography, and duration of field activities. Soil samples were collected from the surface layer at both sites every 3 months for 1 year. Analysis of soil bacterial diversity and composition was conducted using 16S rRNA gene amplicon sequencing. They observed that CF-5Y soil maintained greater stability in bacterial richness and diversity across seasons than M-5Y soil. Notably, fertilization and tillage practices in M-5Y were found to enhance both the diversity and richness of soil bacteria. The study concluded that changes in soil bacterial diversity may result from multifactorial conditions such as land management practices, soil physicochemical properties, weather conditions, and vegetation cover (Arunrat et al.).

Another finding relates to how soil health and crop productivity can be impacted by organic amendment into plant-fungus intercropping systems. Duan et al. assessed the impact of bagasse amendment in a sugarcane - *Dictyophora indusiata* intercropping system on soil health through a field experiment comprising three treatments: bagasse amendment alone, (2) sugarcane amended with bagasse, and the control. Soil chemistry, soil bacterial and fungal diversity using amplicon sequencing, and metabolite composition were analyzed to elucidate the mechanisms underlying the effects of this intercropping system on soil properties. Soil chemistry analyses revealed higher levels of several soil nutrients such as nitrogen and phosphorus in the bagasse application compared to the control. Bacterial diversity was greater in the bagasse application than other treatments, while fungal diversity was lower in the bagasse-amended sugarcane than in other treatments. Soil metabolome analysis revealed significantly lower abundance of carbohydrate metabolites in the bagasse application compared to the control and the bagasse-amended sugarcane. They suggested that the sugarcane amended with bagasse can improve soil health in this intercropping system.

Two studies on microbial genome sequencing revealed new species. A *Bradyrhizobium.* isolated from polluted sediments of a lake in China was identified as a new free-living species *Bradyrhizobium roseum* (Zhang et al.). *B. roseum* displays considerable heterogeneity, exhibiting several functional distinctions from previously described *Bradyrhizobium* genomes. Another study reported a new endophytic fungus, *Cladosporium angulosum*, harboring N uptake-related genes with the potential to enhance plant growth (Yang et al.).

Amon et al. utilized 16S rRNA gene analyses to explore bacterial biogeography and ecosystem services in soils of Côte d'Ivoire. Within bacterial communities comprising 48 phyla, 92 classes, 152 orders, 356 families, and 1,234 genera in Côte d'Ivoire soils, a core bacteriobiome was identified. The distribution of the core genera, along with the 10 major phyla, was influenced by environmental factors including latitude, and soil pH, Al, and K. The predominant distribution pattern observed for the core bacteriobiome was vegetation-independent. Concerning predicted functions, all core bacterial taxa were implicated in assimilatory sulfate reduction, while atmospheric dinitrogen (N_2) reduction was exclusively associated with the genus *Bradyrhizobium*.

Carroll et al. employed phylogenomic and comparative genomics to characterize the population structure and functional potential of 110 *Micrococcus* strains, comprising 104 publicly available genomes and six individuals isolated from South Africa. In terms of functional potential, genes for antimicrobial compounds, including macrolides, beta-lactams, and aminoglycosides (in 81, 61, and 44 of 110 genomes, respectively), were identified across *Micrococcus* genomes. Genome-wide analyses enabled Ma et al. to understand variation in *Mycoplasma gallisepticum*, one of the primary causative agents of chronic respiratory diseases in poultry. It was revealed that the intensity of *M. gallisepticum* biofilm formation strongly correlates with chronic infection, and strains with stronger biofilm-forming abilities exhibit reduced sensitivity to 17 tested antibiotics. Putative key genes associated with biofilm formation, identified through genome-wide analysis of two strains with contrasting biofilm formation, included ManB, oppA, oppD,

PDH, eno, RelA, msbA, deoA, gapA, rpoS, Adhesin P1 precursor, S-adenosine methionine synthetase, and methionyl tRNA synthetase.

Two investigations underscored the significance of microbial OMICS in elucidating natural environmental microbiological processes. de Oliveira et al. utilized shotgun metagenomic sequencing data from water column samples to investigate the resistome and bacterial diversity of two small lakes in the Southern Pantanal region of Brazil. The Abobral lake displayed the highest diversity and abundance of antibiotic resistance genes, antibiotic resistance classes, phyla, and genera. RPOB2 was identified as the most abundant antibiotic resistance gene, and its associated resistance class was the most abundant class. Pseudomonadota emerged as the dominant phylum across all sites, with *Streptomyces* being the most prevalent genus. Ye et al. isolated a thermophilic *Bacillus subtilis* strain from nicotine waste that was resistant to nicotine and possessed high capability to degrade tobacco-derived organics. Whole-genome sequencing revealed that this *B. subtilis* strain also exhibited antibacterial properties, enabling its use in organic fertilizers capable of biological control.

Author contributions

AZ: Conceptualization, Writing–original draft, Writing–review and editing. MH: Writing–review and editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Beale, D. J., Jones, O. A. H., Bose, U., Broadbent, J. A., Walsh, T. K., van de Kamp, J., et al. (2022). Omics-based ecosurveillance for the assessment of ecosystem function, health, and resilience. *Emerg. Top. Life Sci.* 6 (2), 185–199. doi:10.1042/ETLS20210261

Bertola, M., Ferrarini, A., and Visioli, G. (2021). Improvement of soil microbial diversity through sustainable agricultural practices and its evaluation by -omics approaches: a perspective for the environment, food quality and human safety. *Microorganisms* 9 (7), 1400. doi:10.3390/microorganisms9071400

Chen, C., Wang, J., Pan, D., Wang, X., Xu, Y., Yan, J., et al. (2023). Applications of multi-omics analysis in human diseases. *MedComm (2020)* 4 (4):e315. doi:10.1002/mco2.315

Goossens, E., Dehau, T., Ducatelle, R., and Van Immerseel, F. (2022). Omics technologies in poultry health and productivity - part 2: future applications in the poultry industry. *Avian Pathol.* 51 (5), 418–423. doi:10.1080/03079457.2022.2085545

Hijri, M. (2023). Microbial-Based plant Biostimulants. *Microorganisms* 11 (3), 686. doi:10.3390/microorganisms11030686

Kaur, H., Kaur, G., Gupta, T., Mittal, D., and Ali, S. A. (2023). Integrating omics technologies for a Comprehensive understanding of the Microbiome and its impact on Cattle production. *Biol. (Basel)* 12 (9), 1200. doi:10.3390/biology12091200

Li, F., Zi, H., Sonne, C., and Li, X. (2023). Microbiome sustains forest ecosystem functions across hierarchical scales. *Eco Environ. Health* 2 (1), 24–31. doi:10.1016/j.eehl.2023.03.001

Liu, J., Meng, Z., Liu, X., and Zhang, X.-H. (2019). Microbial assembly, interaction, functioning, activity and diversification: a review derived from community compositional data. *Mar. Life Sci. Technol.* 1 (1), 112–128. doi:10.1007/s42995-019-00004-3

Natnan, M. E., Mayalvanan, Y., Jazamuddin, F. M., Aizat, W. M., Low, C. F., Goh, H. H., et al. (2021). Omics Strategies in Current advancements of infectious fish disease management. *Biol. (Basel)* 10 (11), 1086. doi:10.3390/biology10111086

Rodrigues-Filho, J. L., Macedo, R. L., Sarmento, H., Pimenta, V. R. A., Alonso, C., Teixeira, C. R., et al. (2023). From ecological functions to ecosystem services: linking coastal lagoons biodiversity with human well-being. *Hydrobiologia* 850 (12-13), 2611–2653. doi:10.1007/s10750-023-05171-0

Su, G., Yu, C., Liang, S., Wang, W., and Wang, H. (2024). Multi-omics in food safety and authenticity in terms of food components. *Food Chem.* 437 (Pt 2), 137943. doi:10.1016/j.foodchem.2023.137943

Check for updates

Soil chemistry, metabarcoding, and metabolome analyses reveal that a sugarcane—*Dictyophora indusiata* intercropping system can enhance soil health by reducing soil nitrogen loss

Mingzheng Duan[1,2†], Yijie Li[1†], Guanghu Zhu[3], Xiaojian Wu[4], Hairong Huang[1], Jie Qin[1], Shengfeng Long[4], Xiang Li[1], Bin Feng[5], Sunqian Qin[5], Qi-Huai Liu[3], Changning Li[1], Lingqiang Wang[2], Qing Li[2], Tieguang He[4]* and Zeping Wang[1]*

[1]Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences/Sugarcane Research Center, Chinese Academy of Agricultural Science/Key Laboratory of Sugarcane Biotechnology and Genetic Improvement (Guangxi), Ministry of Agriculture, Nanning, China, [2]Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Agriculture, Guangxi University, Nanning, China, [3]Center for Applied Mathematics of Guangxi (GUET), Guilin, China, [4]Guangxi Academy of Agricultural Sciences, Nanning, China, [5]Laibin Academy of Agricultural Sciences, Laibin, China

Introduction: Greater amounts of fertilizer are applied every year to meet the growing demand for food. Sugarcane is one of the important food sources for human beings.

Methods: Here, we evaluated the effects of a sugarcane—*Dictyophora indusiata* (DI) intercropping system on soil health by conducting an experiment with three different treatments: (1) bagasse application (BAS process), (2) bagasse+DI (DIS process), and (3) the control (CK). We then analyzed soil chemistry, the diversity of soil bacteria and fungi, and the composition of metabolites to clarify the mechanism underlying the effects of this intercropping system on soil properties.

Results and discussion: Soil chemistry analyses revealed that the content of several soil nutrients such as nitrogen (N) and phosphorus (P) was higher in the BAS process than in the CK. In the DIS process, a large amount of soil P was consumed by DI. At the same time, the urease activity was inhibited, thus slowing down the loss of soil in the DI process, while the activity of other enzymes such as β-glucosidase and laccase was increased. It was also noticed that the content of lanthanum and calcium was higher in the BAS process than in the other treatments, and DI did not significantly alter the concentrations of these soil metal ions. Bacterial diversity was higher in the BAS process than in the other treatments, and fungal diversity was lower in the DIS process than in the other treatments. The soil metabolome analysis revealed that the abundance of carbohydrate metabolites was significantly lower in the BAS process than in the CK and the DIS process. The abundance of D(+)-talose was correlated with the content of soil nutrients. Path analysis revealed that the content of soil nutrients in the DIS process was mainly affected by fungi, bacteria, the soil metabolome, and soil enzyme activity. Our findings indicate that the sugarcane—DIS intercropping system can enhance soil health.

KEYWORDS

intercropping system, NPK, soil ions, soil enzyme activity, metabarcoding, soil metabolome

1. Introduction

The sustainable production of food has long been an important goal for mankind, and this is being increasingly challenged by the unabated growth of the human population and the intensification of climate change (Zhang et al., 2018; Singh et al., 2020). Large quantities of fertilizer have been applied to accommodate the growing demand for food. However, the excessive application of fertilizer has had deleterious effects both on crop production (e.g., decreased yields) and the environment (e.g., environmental pollution, reductions in soil biodiversity) in both China and several other developing countries (Vitousek et al., 2009; Li et al., 2019; Cui et al., 2020; Xu et al., 2020). There is thus a pressing need to develop approaches that provide some of the beneficial effects of fertilizer applications without compromising agricultural productivity and soil environments (Duan et al., 2022). Sugarcane (*Saccharum* spp.) is the world's largest sugar crop, which widely used in food and industrial fields (Singh et al., 2020). Bagasse is a plant residue that is generated following the extraction of sugar from crops such as sugarcane. It is known that inorganic nutrients (e.g., phosphorus (P) and potassium (K)) required for the growth of plants are abundant in bagasse (Sharma et al., 2015). Bagasse is thus often used as an auxiliary fertilizer in sugarcane fields after harvest and sugar extraction (Seleiman and Kheir, 2018). The application of bagasse is considered an effective strategy for alleviating the deleterious effects of excessive fertilizer application and more generally for achieving the goals of ecological agriculture, which aims to enhance production while promoting ecological sustainability. Bagasse application has been shown to enhance the health of soil by increasing the abundance of various soil nutrients (Sharma et al., 2015; Seleiman and Kheir, 2018; Bento et al., 2019) and pH (Mohamed et al., 2019). Bagasse application also promotes the activity of soil microbes and increases the diversity of soil microbial communities (Bhadha et al., 2020; Ullah et al., 2020).

Dictyophora indusiata (DI), which is in the family Phallaceae, is an economically valuable, edible mushroom that is used in Chinese cuisine (Duan et al., 2023). DI is also widely used as a Chinese medicine material (Wang et al., 2021). Some agricultural waste including bamboo and sugarcane bagasse can be used to grow DI and improve soil nutrients (Tong et al., 2021; Thaisuchat et al., 2022). Consequently, the application of bagasse to sugarcane fields can enhance sugarcane production and promote soil health. However, whether bagasse application can have positive effects on the cultivation of DI when it is intercropped with sugarcane remains unclear. Therefore, additional research is needed to explore approaches that could be employed to enhance the abundance of soil nutrients in sugarcane–DI intercropping systems.

Soil health is affected by various biotic and abiotic factors, and microbes (especially bacteria and fungi) are one of the most important factors affecting soil biogeochemical cycles and the transformation of soil nutrients (Basu et al., 2021). Soil enzymes are mainly derived from soil microbes, and concentrations of soil nutrients are substantially affected by soil microbial activities. One previous study has shown that urease activity can have a major effect on the release of nitrogen (N) in soil (Cantarella et al., 2018). In sugarcane–DI intercropping systems, DI is the main fungus mediating the transformation of soil; this has major implications for soil health because the presence of a single, dominant fungus in the soil can alter the composition of metal ions in soil and thus affect the absorption of metal ions by other organisms (Zotti et al., 2021; Duan et al., 2022). Soil metabolomics can

be used to clarify correlations of metabolic fingerprints with environmental factors, soil nutrients, microbial diversity, and plant phenotypes (Withers et al., 2020). For example, changes in the abundance of soil carbohydrate metabolites can be affected by the content of soil organic matter and microbial activity (Vranova et al., 2013; Gunina and Kuzyakov, 2015). This approach can thus be used to identify key metabolite markers associated with soil health.

Here, we evaluated the effects of a sugarcane–DI intercropping system on soil health through a field experiment. We then conducted analyses of soil chemistry, soil fungi and bacteria diversity, as well as the metabolite composition to clarify the mechanisms by which the sugarcane–DI intercropping system enhanced soil health. Our study suggests a new cultivation model to preserve soil fertility and increase economic value of sugarcane fields.

2. Materials and methods

2.1. Materials

The experiment was conducted in a field located in Laibin City, Guangxi, China (23°16′N, 108°24′E) from February to August 2022. There were three treatments during the experiment: control (CK), in which sugarcane was grown under normal management conditions; BAS process, in which crushed, air-dried bagasse was applied to the soil between rows of sugarcane plants; and DIS process, in which crushed, air-dried BAS, mixed with DI strains obtained via wheat grain fermentation, was applied to the soil between rows of sugarcane plants (Figure 1A). The sugarcane variety GT42 was obtained locally and was used in all experiments. Sugarcane plants were grown with approximatively a 50 cm interrow distance, and the depth of each row was 30 cm of each ridge. When the height of sugarcane plants reached approximately 50 cm after 60 days of cultivation, we demarcated five semi-rectangular areas in which the three treatments were performed. Two duplicates of the CK and BAS process were performed, and the DIS process was only performed in one of these areas, which was located in the middle of the experimental site (Figure 1A). Fruiting bodies of DI matured after approximately 90 d, and high soil moisture levels were maintained during this period. Five soil samples were taken from the three processes (CK, BAS, and DIS), respectively, in the experimental area (red dots in Figure 1A).

After the surface bagasse was cleaned, soil samples were collected from the topsoil layer at a depth of 0–10 cm. The soil was then sieved through a 0.425 mm filter, packed into sterile cryostorage tubes, and frozen in liquid nitrogen.

2.2. Methods

Various analytical methods shown in Figure 1B were used to evaluate the effect of the sugarcane–DI intercropping system on soil properties.

2.2.1. Soil chemical properties
The Kjeldahl method with sulfuric acid–accelerator digestion was used to measure the content of total N and basic nitrogen (BN) in soil (Jiao et al., 2015). NaOH alkali melting and molybdenum–antimony resistance spectrophotometry were used to measure the content of total P in soil (Xiao et al.,1988). NaOH alkali melting and flame
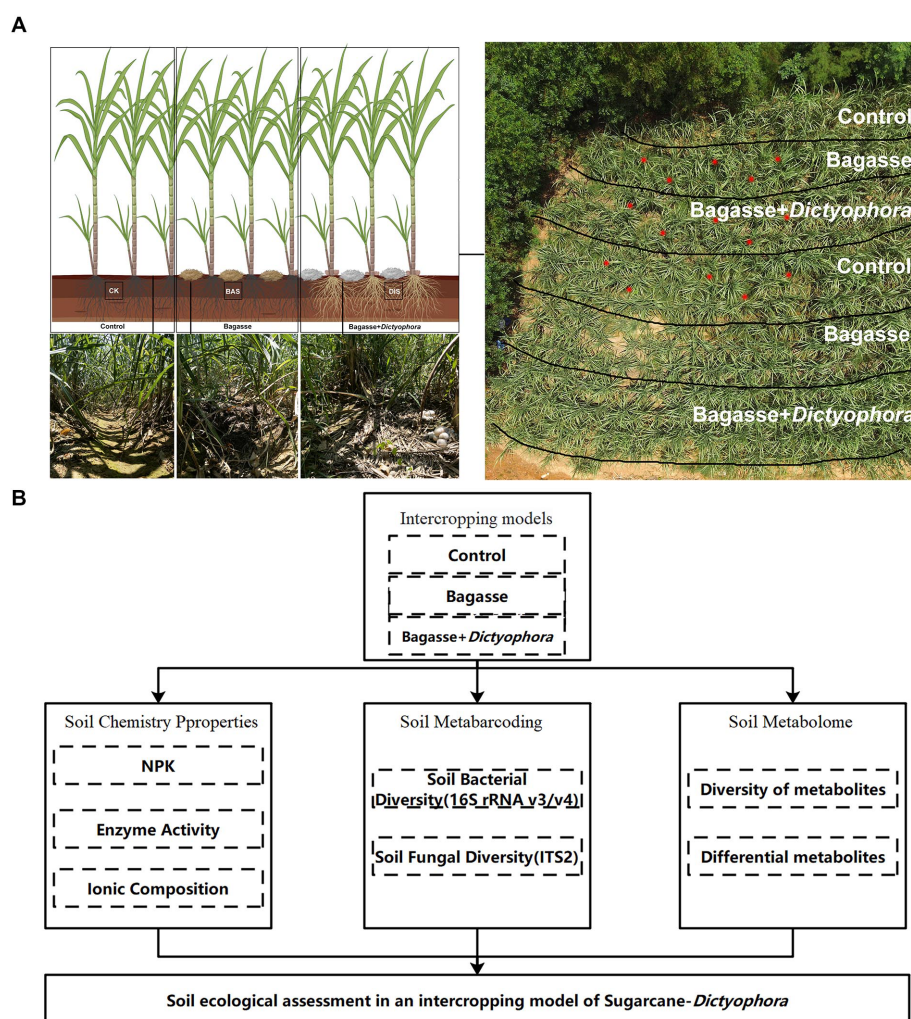
FIGURE 1
The experimental approach used in our study to characterize the effects of the sugarcane−DI intercropping system on soil health. **(A)** Experimental design with three treatments. The three images in the lower left show sugarcane rows corresponding to the different treatments. The image in the upper right shows an aerial view of the experimental field and the division of the treatments within the experimental field; soil samples were taken at each of the red dots shown in this image. **(B)** Schematic diagram of the analytical procedures used to evaluate the effects of each treatment on soil properties. The Figure 1A were build in www.biorender.com.

photometry were used to measure the content of total K in soil (Fu et al., 1988). A carbon and N analyzer was used to measure the content of soil organic carbon (OC) (Dalian Environmental Monitoring Center, 2014). Molybdenum–antimony resistance colorimetry with ammonium fluorine-hydrochloric acid solution and sodium bicarbonate solution was used to measure the content of available P (AP) (Soil Testing-Part7, 2014). Ammonium acetate extraction and flame photometry were used to measure the content of available K (AK) in soil (Du et al., 2004). A Micro Soil Urease Assay Kit (BC0125-100 T/48S, Solarbio, Beijing, China) was used to determine the activity of soil urease. The 3, 5-dinitrosalicylic acid colorimetric method was used to measure the content of sucrase in soil [30; 31]. The p-nitrophenol colorimetric method was used to determine the content of β-glucosidase (Eivazi and Tabatabai, 1988). The 3, 5-dinitrosalicylic acid colorimetric method was used to determine the content of cellulase in soil (Li et al., 2008; Lin, 2010). An Assay Kit (YX-C-B913, Lai Er Bio-Tech, Hefei, China) was used to determine the laccase activity in soil. An Assay Kit (YX-C-B625, Lai Er Bio-Tech, Hefei, China) was used to determine the amylase activity in soil. After

deionized water extraction, inductively coupled plasma optical emission spectroscopy (ICP-OES) was used to measure the content of 27 ions in soil (Bao, 2000). IBM SPSS Statistics 26.0 software was used to compute variances in each of the soil chemical properties.

2.2.2. Metabarcoding analysis

A HiPure Soil DNA Kit (Magen, Guangzhou, China) was used to extract DNA from soil samples (3 g) following the manufacturer's instructions according to a previously described method (Duan et al., 2022). Polymerase chain reaction (PCR) of the 16S rDNA V3–V4 region in the ribosomal RNA gene was conducted using the primers 341F (5′-CCTACGGGNGGCWGCAG-3′) and 806R (5′-GGACTACHVGGGTATCTAAT-3′) for bacteria (Guo et al., 2017) and the primers ITS3_KYO2 (5′-GATGAAGAACGYAGYRAA-3′) and ITS4 (5′-TCCTCCGCTTATTGATATGC-3′) for fungi (Toju et al., 2012). An Illumina NovaSeq 6,000 sequencing platform was used to paired-end sequence the purified amplicons, which were pooled in equimolar ratios following the standard protocol. The clean tags were clustered into operational taxonomic units (OTUs) of ≥97% similarity

using UPARSE (Edgar, 2013) (version 9.2.64) pipeline. The tag sequence with highest abundance was selected as representative sequence within each cluster. A naïve Bayesian model using the Ribosomal Database Project classifier (version 2.2) (Wang et al., 2007) based on SILVA (16S rRNA OTUs) database (Pruesse et al., 2007) (version138.1) and UNITE (ITS OTUs) database (Nilsson et al., 2019) (version 8.3) with the confidence threshold value of 0.8, which is a naïve Bayesian classifier, was used to classify the representative operational taxonomic unit (OTU) sequences. R software was used to create all figures. The VennDiagram package (version 1.6.16) in R software was used to compare OTUs among groups (Chen and Boutros, 2011). Principal component analysis (PCA) and Shannon index analysis were performed to identify the between-group variance. QIIME (version 1.9.1) software was used for Shannon, Chao1, ACE, Simpson and Good's coverage index calculations (Caporaso et al., 2010). For PCA analysis, first, OTUs numbers were normalized into $\log_2$, then vegan package (version 2.5.3; http://CRAN.R-project.org/package=vegan; 2022.11.9) in R software was used to conduct PCA to clarify variation in the composition of OTUs among experimental groups.

2.2.3. Soil metabolome analysis

Soil samples (0.5 g) were mixed with 1 mL of a methanol:isopropanol:water (3:3:2 v/v/v) solution, vortexed for 3 min, and placed on an ultrasound which performed in ice water for 20 min. After centrifugation of the extract at 12,000 r/min at 4°C for 3 min, the supernatant was mixed with 0.020 mL of internal standard (10 μg/mL) in a sample vial, followed by evaporation under N flow. A lyophilizer was used to freeze-dry the sample, and the residue was used for derivatization. The sample was then incubated at 37°C for 2 h after being mixed with 0.1 mL of methoxyamine hydrochloride in pyridine (0.015 g/mL). Next, 0.1 mL of bis-(trimethylsilyl)-trifluoroacetamide with 1% trimethylchlorosilane was added to the mixture and maintained at 37°C for 30 min after vortex mixing. After mixing approximately 0.2 mL of the derivatization solution with n-hexane and diluting to 1 mL, the mixture was filtered through a 0.22 μm organic phase syringe filter and placed in a refrigerator at-20°C; the mixture was then analyzed within 24 h. Gas chromatography–mass spectrometry was conducted on an Agilent 8,890 gas chromatograph with a 5977B mass spectrometer and a DB-5MS column (30 m length × 0.25 mm i.d. × 0.25 μm film thickness, J&W Scientific, United States) to analyze the metabolites in soil samples. The carrier gas used was helium, and the gas flow rate through the column was 1.2 mL/min. Injections of 1 μL were conducted in front inlet mode at a split ratio of 5:1. The oven temperature was maintained at 40°C for 1 min and then increased to 100°C at 20°C/min; it was then increased to 300°C at 15°C/min and held at 300°C for 5 min. The scan mode was used in analyses of all samples. The ion source temperature and transfer line temperature were 230°C and 280°C, respectively. Metware's metabolite database (v2.0; Metware Biotechnology Co., Ltd. Wuhan, China), as well as publicly available databases, including MassBank (http://www.massbank.jp (accessed on 20 October 2022), HMDB (Human Metabolome Database; http://www.hmdb.ca (accessed on 20 October 2022)), and METLIN (http://metlin.scripps.edu/index.php (accessed on 20 October 2022)), were used to qualitatively analyze primary and secondary mass spectrometry data. Unsupervised PCA was conducted using the function prcomp in R (www.r-project.org), and all data were unit variance-scaled prior to the unsupervised PCA.

2.2.4. Correlation analysis

SmartPLS4 software[1] was used to construct a partial least squares path model (PLS–PM). Cronbach alpha index values for each latent variable in the PMs were greater than 0.3, and p-values were evaluated using 20,000 bootstrap replicates. OmicShare Tools (https://www.omicshare.com/tools (accessed on 20 November 2022)) was used to conduct canonical correlation analysis (CCA). All data were $\log_2$-transformed prior to analyses.

3. Results

3.1. Soil chemistry properties

3.1.1. The BAS and DIS process had a significant effect on the content of P, OC, AP, and AK

First, the main soil nutrient indicators, including N, P, K, pH, OC, AP, BN, and AK, in soil from the surface layer among the three experimental groups (CK, BAS process, and DIS process) were compared. The BAS process and DIS process had a significant effect on the content of P, OC, AP, and AK ($F_{2,12} > 100$) (Table 1). Soil P content (CK: 0.66 g/kg; BAS: 0.79 kg) and AP (69.02 mg/kg, 142.46 mg/kg) were higher in the BAS process (0.79 g/kg and 142.46 mg/kg, respectively) than in the CK (0.66 g/kg and 69.02 mg/kg, respectively), while OC content was lower in the BAS process (11.54 mg/kg) than in the CK (15.62 mg/kg). Moreover, P, AP, and AK contents were lower in the DIS process (0.59 g/kg, 65.7 mg/kg, and 92.4 mg/kg, respectively) than in the BAS process (0.79 g/kg, 142.76 mg/kg, and 125.4 mg/kg). It was also noticed that soil pH was higher in the BAS process (4.84) than in the CK (4.26), and soil pH was lower in the DIS process (4.17) than in the CK. Although differences were observed in the content of other nutrients among treatments, such as the content of N, K, and BN, differences in the content of these nutrients among treatments were not significant.

3.1.2. The DIS process restrains soil urease activity

We measured the activity of six soil enzymes, urease, sucrase, β-glucosidase, cellulase, laccase, and amylase, that play a key role in shaping soil chemical properties. The BAS process and the DIS process had significant effects on the activities of all enzymes according to analysis of variance ($F_{2,12} > 100$) (Table 2). The activity of cellulase in soil was higher in the BAS process (1.19 U/g) than in the CK (0.91 U/g); the activity of urease, sucrase, and β-glucosidase was lower in the BAS process (1,228.95 U/g, 3.28 U/g, and 12. 01 U/g, respectively) than in the CK (1,326.66 U/g, 9.87 U/g, and 19.35 U/g, respectively). The activity of sucrase, β-glucosidase, cellulase, laccase, and amylase was higher in the DIS process (9.23 U/g, 31.17 U/g, 1.68 U/g, 9.92 U/g, and 1.45 U/g, respectively) than in the BAS process (3.28 U/g, 12.01 U/g, 1.19 U/g, 1.51 U/g, and 0.87 U/g, respectively). Urease activity in soil was lower in the DIS process (867.52 U/g) than in the BAS process (1,228.95 U/g); the activity of urease was higher than that of all five other enzymes examined. These findings indicate that the BAS process and the DIS process can have significant effects on the activity of soil enzymes.

TABLE 1 Different soil nutrients content of soil samples in the BAS process and DIS process.

ID	CK	BAS	DIS	*F*-value	CK/BAS	BAS/DIS
N(g/kg)	1.45	1.57	1.503	8.449	↑**	ns
P(g/kg)	0.66	0.79	0.59	153.415	↑***	↓***
K(g/kg)	5.22	4.81	4.99	11.164	↓**	ns
pH	4.26	4.84	4.17	37.73	↑***	↓***
OC(mg/kg)	15.62	11.54	15.04	511.063	↓***	↑***
AP(mg/kg)	69.02	142.46	65.7	1781.229	↑***	↓***
BN(mg/kg)	153.42	163.34	142.42	82.226	↑***	↓***
AK(mg/kg)	120.8	125.4	92.4	3,686	↑***	↓***

Numbers are the means of five samples; "/" indicates versus; "↑ and ↓" indicate up-regulated and down-regulated, respectively; OC, organic carbon; AP, available phosphorus; BN, basic nitrogen; AK, available potassium; "*" indicates *p*-value between 0.01 and 0.05; "**" indicates *p* between 0.001 and 0.005; "***" indicates *p*-value less than 0.001; "ns" indicates not significant.

3.1.3. The content of soil ions less affected by both BAS and DIS processes

We measured the content of common ions in soil using ICP–OES to clarify the effects of the BAS process and the DIS process on the composition of ions in soil. The mean content of zinc (Zn), iron (Fe), K, titanium (Ti), magnesium (Mg), calcium (Ca), P, and sodium (Na) in soil was greater than 1 mg/L, indicating that these ions were major components of the soil; the mean content of barium (Ba), manganese (Mn), cerium (Ce), vanadium (V), rubidium (Rb), chromium (Cr), lanthanum (La), strontium (Sr), (lead) Pb, lithium (Li), nickel (Ni), thorium (Th), niobium (Nb), scandium (Sc), copper (Cu), gallium (Ga), cobalt (Co), beryllium (Be), and molybdenum (Mo) was less than 1 mg/L, indicating that these ions were minor components of the soil (Table 3). We characterized variation in the content of 27 ions in soil between treatments, and the average *F*-value for the 27 ions was 32. The content of ions in soil was not greatly affected by the BAS process and the DIS process, as significant differences in the content of most ions between the BAS process and the DIS process were not observed (e.g., Zn content) and variation among treatments was low (e.g., *F*-value in the Zn content was less than 32). However, the *F*-values of six ions of interest (Fe, Ca, P, Cr, La, and Mo) were greater than 32 (33, 300.345, 93.091, 37, 41, and 127, respectively). The content of Ca and La was higher in the BAS process than in the CK; the content of Fe, Cr, and Mo was lower in the BAS process than in the CK. The content of P and Mo was lower in the DIS process than in the CK. Data on the content of the above ions converted into unit soil weight (mg/kg) are shown in Supplementary Table S1.

3.2. The diversity of soil fungi and bacteria inversely impacted by BAS and DIS

The effects of the BAS process and the DIS process on soil microbial diversity were clarified via soil metabarcoding analysis. A total of 3,241,995 effective metabarcoding tags were obtained from 30 samples via sequencing of the 16S rRNA V3–V4 and ITS 1–2 regions.

TABLE 2 Soil enzymes activity in the BAS process and the DIS process.

ID	CK	BAS	DIS	*F*-value	CK/BAS	BAS/DIS
urease	1326.66	1228.95	867.52	125.06	↓**	↓***
sucrase	9.87	3.28	9.23	830.01	↓***	↑***
β-glucosidase	19.35	12.01	31.17	345.86	↓***	↑***
cellulase	0.91	1.19	1.68	112.32	↑***	↑***
laccase	1.18	1.51	9.92	1185.56	ns	↑***
amylase	0.67	0.87	1.45	51.89	ns	↑***

The numbers are means of five soil samples. The units are U/g; "/" indicates versus; "↑ and ↓" indicate up-regulated and down-regulated, respectively; "*" indicates *p*-value between 0.01 and 0.05; "**" indicates *p*-value between 0.001 and 0.005; "***" indicates *p*-value less than 0.001; "ns" indicates not significant.

TABLE 3 Soil metal ions content in the BAS process and DIS process.

ID	CK	BAS	DIS	*F*-value	CK/BAS	BAS/DIS
Zn(mg/L)	189.11	163.59	162.96	7.546	↓**	ns
Fe(mg/L)	102.82	83.31	83.4	33.571	↓***	ns
K(mg/L)	24.32	21.42	20.56	16.176	↓**	ns
Ti(mg/L)	11.68	10.73	10.54	5.33	↓*	ns
Mg(mg/L)	10.04	9.28	9.69	3.118	↓*	ns
Ca(mg/L)	1.46	3.35	1.93	300.345	↑***	↓***
P(mg/L)	1.58	1.66	0.99	93.091	ns	↓***
Na(mg/L)	1.18	1.02	0.88	16.781	↓*	↓*
Ba(ug/L)	376.1	359.86	349.13	6.092	ns	ns
Mn(mg/L)	0.33	0.34	0.34	0.411	ns	ns
Ce(ug/L)	321.71	320.48	337.49	1.055	ns	ns
V(ug/L)	269.6	229.75	224.57	24.751	↓***	ns
Rb(ug/L)	213.79	173.35	186.98	22.624	↓***	↑*
Cr(ug/L)	192.72	160.69	153.46	37.273	↓***	ns
La(ug/L)	127.27	158.31	164.04	40.912	↑***	ns
Sr(ug/L)	117.02	112.44	119.68	1.067	ns	ns
Pb(ug/L)	94.81	79.35	80.01	11.788	↓**	ns
Li(ug/L)	84.66	68.88	68.69	16.266	↓***	ns
Ni(ug/L)	68.14	53.38	55.12	25.241	↓***	ns
Th(ug/L)	53.94	59.43	59.79	12.44	↑**	ns
Nb(ug/L)	55.44	52.41	51.35	2.049	ns	ns
Sc(ug/L)	45.75	48.33	46.46	2.075	ns	ns
Cu(ug/L)	48.6	44.43	42.83	6.703	ns	ns
Ga(ug/L)	44.92	39.27	38.11	21.784	↓***	ns
Co(ug/L)	27.96	22.29	21.94	20.265	↓***	ns
Be(ug/L)	4.53	3.77	3.56	14.214	↓**	ns
Mo(ug/L)	0.99	0.87	0.68	126.819	↓***	↓***

The number indicates the mean value of five soil samples; "/" indicates versus; "↑ and ↓" indicate up-regulated and down-regulated, respectively; "*" indicates *p*-value between 0.01 and 0.05; "**" indicates *p*-value between 0.001 and 0.005; "***" indicates *p*-value less than 0.001; "ns" indicates not significant.

OTU taxonomic classification is shown in Supplementary Tables S3, S4. Alpha diversity indices including Sobs, Simpson, Chao, Ace, and good's coverage are shown in Supplementary Tables S5, S6. OTU

clustering revealed an average of 2,842 bacterial OTUs (16S rDNA) and 637 fungal OTUs (ITS) per sample (Supplementary Table S2). The Venn diagram based on the OTUs revealed that 973 and 416 bacterial and fungal OTUs were shared among all samples (Figures 2A,B). The number of unique bacterial OTUs (Figure 2A) was higher in the BAS process (654) than in the CK (363) and DIS process (368), and the number of unique fungal OTUs (Figure 2B) was significantly lower in the DIS process (152) than in the CK and BAS process (299 and 234, respectively). PCA of the abundance of bacterial and fungal OTUs in the BAS process, DIS process, and CK indicated that the composition of bacterial and fungal OTUs varied among experimental groups (Figures 2C,D). Shannon index was calculated based on the alpha diversity indices to characterize soil microbial diversity. The mean Shannon index of bacterial OTUs for the CK, BAS process, and DIS process was 8.36, 9.12, and 8.93, respectively (Figure 2E). Thus, bacterial diversity was significantly higher in the BAS process than in the CK and DIS process ($p < 0.001$), and no significant difference in bacterial diversity was observed between the DIS process and the CK. The mean Shannon index of fungal OTUs for the CK, BAS process, and DIS process was 5.72, 5.85, and 4.95, respectively (Figure 2F). Thus, fungal diversity was significantly lower in the DIS process than in the BAS process and CK ($p < 0.001$), and there was no significant difference in fungal diversity between the BAS process and the CK. In sum, we found that soil bacterial diversity was higher in the BAS process than in the DIS process and the CK, and soil fungal diversity was higher in the DIS process than in the BAS process and the CK.

3.3. Soil metabolome

3.3.1. BAS and DIS process significantly impacted soil metabolites

A soil metabolomics analysis was conducted to clarify the effects of the BAS process and the DIS process on soil metabolites. A total of 96 metabolites grouped into 12 classes were detected across all soil samples, including acids (19), alcohols (12), amines (3), aromatics (3), carbohydrates (22), esters (4), heterocyclic compounds (3), ketones (1), lipids (17), nitrogen compounds (1), phenols (1), and others (8) (Figure 3; Supplementary Table S7). According to the PCA plot, the metabolites in soil in the BAS process and the DIS process significantly differed from those in the CK and were clustered into two separate groups. This suggests that both BAS and DIS processes had significant effects on the metabolites in soil (Supplementary Figure S1).

3.3.2. Soil metabolites were significantly regulated in both BAS and DIS processes

The abundance of metabolites and different categories of metabolites were estimated according to z-scores (standardized population data). The abundances of carbohydrates were significantly affected by BAS and DIS processes according to a heat map (Figure 3). Specifically, the abundance of carbohydrates was significantly lower in the BAS process than in the CK and significantly higher in the DIS process than in the CK. Moreover, the abundance of lipids was lower in the BAS process than in the CK. Therefore, the BAS process and the DIS process had significant effects on the abundances of soil carbohydrates, and the BAS process mediated the dissolution of soil lipids.

In order to clarify the effects of the BAS process and the DIS process on soil metabolites, the top 20 most abundant metabolites were identified (Table 4). The three most abundant metabolites in the soil samples were (Z)-9-octadecenamide 1, hydroxyurea, and propanedioic acid 2. Significant variation was observed in the content of D(+)-talose, sorbitol 1, rhamnose, fructose 2, D-allose 2, and D-arabinose 2 among the experimental groups ($F_{2,12} > 50$). Specifically, the abundances of these carbohydrates were significantly lower in the BAS process than in the CK and significantly higher in the DIS process than in the CK (Table 4). The abundances of sulfonium, (cyanoamino) diphenyl-, hydroxide, inner salt; sucrose; and glycerin in soil were lower in the BAS process than in the CK, and the abundance of carbonic acid 1 was higher in the BAS process than in the CK. The abundances of hydroxyurea, sucrose, and carbonic acid were lower in the DIS process than in the CK, and the abundances of (Z)-9-octadecenamide 1; sulfonium; and (cyanoamino)diphenyl-, hydroxide, and inner salt were higher in the DIS process than in the CK. These data indicate that the BAS process and the DIS process have significant effects on the abundances of carbohydrates.

3.4. Correlation analyses evidence the regulatory relationship between factors of soil chemistry, microbial diversity, and metabolome

3.4.1. Soil enzyme activity had a highly significant positive effect on soil nutrients

The PLS–PM were built to clarify the effects of several soil factors on soil nutrients in the BAS process and the DIS process. In the BAS process (Figure 4A), the soil metabolome had a significant negative effect on soil nutrients (path coefficient = −0.576, $p = 0.043$), and fungi had a positive but insignificant effect on soil nutrients (path coefficient = 0.312, $p = 0.335$); soil ions (path coefficient = 0.604, $p = 0.665$), bacteria (path coefficient = 0.343, $p = 0.763$), and fungi (path coefficient = 0.426, $p = 0.685$) had positive effects on the soil metabolome, but none of these effects were significant ($p < 0.05$). Bacteria (path coefficient = 0.716, $p = 0.003$) and fungi (path coefficient = −0.464, $p = 0.042$) had significant positive and negative effects on soil ions, respectively. Many of the soil factors examined had significant effects on soil nutrients in the DIS process (Figure 4B). Soil enzyme activity had a highly significant positive effect on soil nutrients (path coefficient = 0.976, $p < 0.001$), which indicates that the activity of soil enzymes has a major effect on soil nutrients in the DIS process. Some soil variables affected soil nutrients through their effects on other variables; for example, fungi had a significant positive effect on bacteria (path coefficient = 0.701, $p < 0.001$); bacteria had a significant positive effect on the soil metabolome (path coefficient = 0.617, $p = 0.007$); and the soil metabolome had a negative effect on soil enzyme activity (path coefficient = −0.473, $p = 0.025$). These relationships between fungi and bacteria, between bacteria and the soil metabolome, and between the soil metabolome and soil enzyme activity drive variation in soil nutrients in the DIS process. Fungi had a negative effect on soil ions, but this effect was not significant (path coefficient = −0.502, $p = 0.122$). Our findings indicate that the BAS process had a major effect on the composition of soil ions and reveal the effects of various soil factors on soil nutrients in the DIS process.
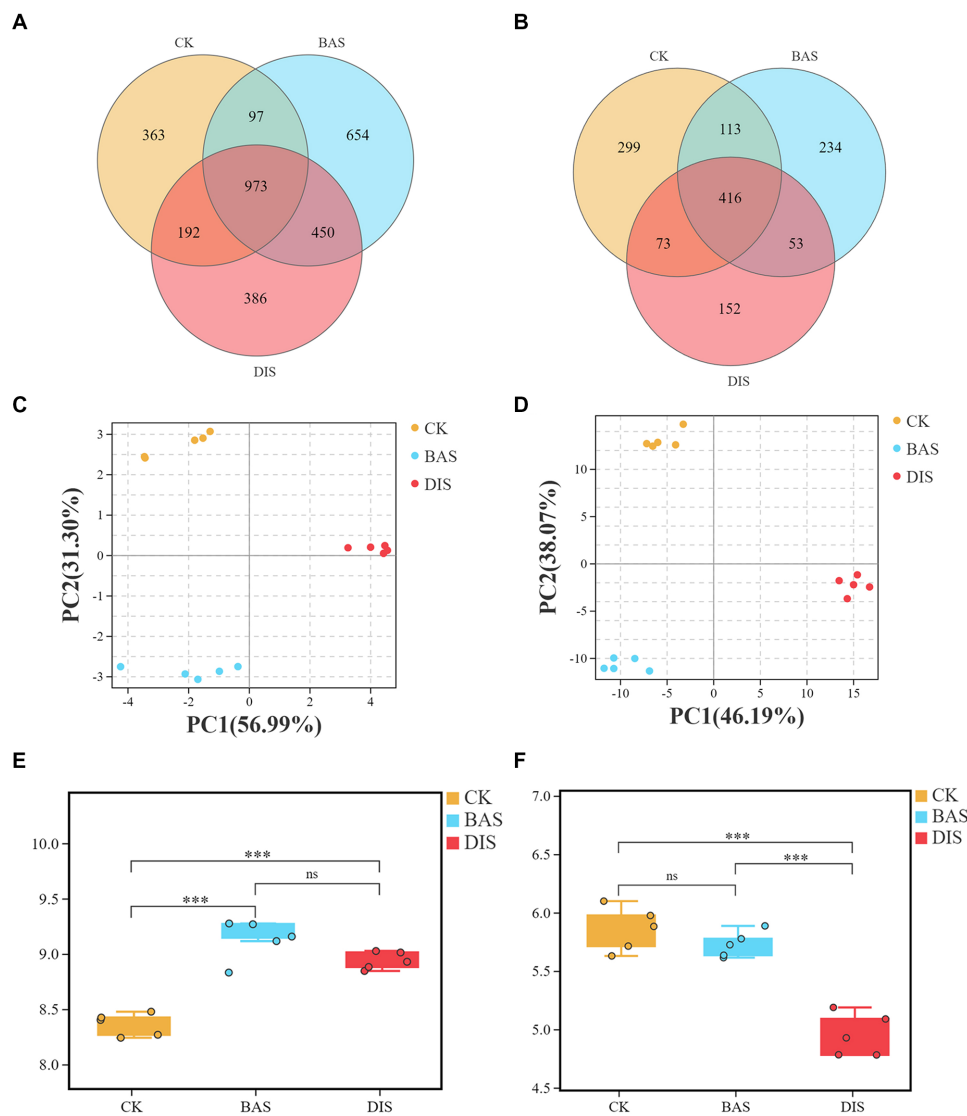
FIGURE 2
Venn analysis, PCA, and alpha diversity (Shannon index) analysis of soil bacterial and fungal diversity in the CK, BAS process, and DIS process (n=5). Venn analysis of **(A)** bacterial and **(B)** fungal OTUs. PCA of **(C)** bacterial and **(D)** fungal OTUs. **(E)** and **(F)** Boxplot showing the alpha diversity (i.e., Shannon index of **(E)** bacteria and **(F)** fungi) in the CK, BAS process, and DIS process. The top and bottom whiskers of the boxes indicate the maximum and minimum values, respectively; the top margin of the box indicates the upper quartile, and the lower margin of the box indicates the lower quartile. The scattered points indicate the distribution of repeated samples within each group. "***" indicates p<0.001, and "ns" indicates not significant.

3.4.2. D(+)-talose was involved in the regulation of bacterial and fungal diversity in DIS process

CCA was conducted to clarify correlations among soil chemistry, microbial diversity, and the soil metabolome in the BAS process and the DIS process. CCA was conducted using bacterial and fungal diversity based on the 50 most abundant bacterial and fungal OTUs, respectively, as response variables and other factors related to soil nutrients, enzyme activity, metal ions, and the 20 most abundant metabolites as explanatory variables (Figure 5). In the BAS process, the content of N, AP, P, and OC and the pH were related to bacterial OTU diversity (Figure 5A); however, these soil nutrient variables were not related to fungal OTU diversity in the BAS process and the DIS process (Figure 5B). The activity of cellulase, amylase, and laccase was related to bacterial OTU diversity (Figure 5C) and fungal OTU diversity (Figure 5D) in the DIS process. The content of Ba, Na, Ti, Zn,

Be, Ki, Li, and Fe was related to bacterial OTU diversity in the BAS process (Figure 5E); the content of Zn, Ti, Li, Fe, Be, Ba, K, and Na was related to fungal OTU diversity in the BAS process (Figure 5F). Sucrose was related to bacterial OTU diversity in the BAS process, and D(+)-talose was related to bacterial OTU diversity (Figure 5G) and fungal OTU diversity (Figure 5H) in the DIS process. The envfit p-values for all the aforementioned correlations were less than 0.01.

4. Discussion

It was found that BAS and DIS processes had significant effects on the content of various nutrients in soil. The application of sugarcane bagasse to the fields enhanced soil nutrient conditions (Seleiman and Kheir, 2018), especially the content of N and P (Table 1). The inverse

FIGURE 3
Composition of soil metabolites in the BAS process and the DIS process. Clustering heat map of all metabolites. Each column corresponds to a sample, and each row corresponds to a metabolite class. A bar with a specific color indicates the abundance of each metabolite. Shades of red and green indicate up-regulated and down-regulated metabolites, respectively.

content of OC in BAS and DIS might be related to increased bacterial activity. Soil transformation may have been mediated later in DIS process. For example, Xiao et al. (2018) noted that fungal diversity can regulate soil OC mineralization, and our findings indicate that fungal diversity was lower in the BAS process than in the CK (Figure 2E), which might affect the soil mineralization process. The fact that the content of OC was higher in the DIS process than in the CK, appeared to be mediated by the cultivation of DI in light of the previous finding that fungal diversity regulates soil OC mineralization. The fact that soil pH was higher in the BAS process than in the CK (Table 1), is consistent with the results of Mohamed (Mohamed et al., 2019) showing that the BAS process can increase the pH of acidic soils. It was found that the

content of P, AP, and AK was significantly higher in the BAS process than in the CK. These nutrient elements are likely derived from ground bagasse, which contains a large amount of the above elements (Sharma et al., 2015). In the DIS process, the cultivation of DI consumes a large amount of soil nutrients, especially P, AP, and AK (Table 1). A previous study of the soil nutrients in soils with DI under continuous cropping over 3 years has shown that the cultivation of DI can increase the content of soil nutrients such as N, P, and K. However, the content of AP and AK decreased gradually over the continuous cropping period (Tong et al., 2021). The content of soil nutrients decreased in the DIS process, which is inconsistent with the results of this previous study. This might be explained by differences in the design of the experiment.

TABLE 4 The 20 most abundant soil metabolites in the BAS process and the DIS process.

Compounds	Class	F-value	CK/BAS	BAS/DIS
(Z)-9-Octadecenamide 1	Amine	4.302	ns	↑*
Hydroxyurea	Nitrogen compounds	8.457	ns	↓**
Propanedioic Acid 2	Acid	1.882	ns	ns
Glycerol 1-palmitate	Ester	1.068	ns	ns
n-Hexadecanoic Acid	Acid	0.124	ns	ns
Octadecanoic acid, 2,3-dihydroxypropyl Ester	Ester	0.212	ns	ns
D(+)-Talose 1	Carbohydrate	245.716	↓**	↑***
Sulfonium, (cyanoamino)diphenyl-, hydroxide, inner salt	Others	37.469	↓***	↑***
Octadecanoic Acid	Acid	1.117	ns	ns
Sucrose	Carbohydrate	19.727	↓**	↓*
Sorbitol 1	Carbohydrate	124.475	↓***	↑***
1-acetyl-2-methyl-Azetidine	Others	1.835	ns	ns
3-Trifluoromethylbenzylamine, N,N-dinonyl	Others	2.764	ns	ns
Rhamnose	Carbohydrate	56.806	↓***	↑***
2,4-Di-tert-butylPhenol 2	Phenol	1.371	ns	ns
Fructose 2	Carbohydrate	90.946	↓***	↑**
carbonic Acid 1	Acid	20.256	↑***	↓***
Glycerin	Alcohol	5.949	↓*	ns
D-Allose 2	Carbohydrate	294.132	↓**	↑***
d-Arabinose 2	Carbohydrate	63.603	↓***	↑***

Numbers are means of five soil samples; "/" indicates versus; "↑ and ↓" indicate up-regulated and down-regulated, respectively; "*" indicates p-value between 0.01 and 0.05; "**" indicates p-value between 0.001 and 0.005; "***" indicates p less than 0.001; "ns" indicates not significant.

For example, in the previous experiment, sugarcane was not planted on the ground as in this experiment. Nevertheless, the findings in this work, as well as the results of this previous study, indicate that DI consumes large amounts of AP and AK in soil. Thus, supplementation of AP and AK is important for enhancing the productivity of the sugarcane–DI intercropping system.

Soil N is generated via the action of urease, but the activity of soil urease is often excessive, and this results in the release of excess N (Hendrickson and Douglass, 1993; Cantarella et al., 2018; Fu et al., 2020). In this study, the content of N and BN in the BAS process and the DIS process did not significantly differ, suggesting that the rate of soil N release in the BAS process and the CK exceeded the sugarcane absorption capacity. This likely caused the urease activity to decrease in the DIS process; however, little variation was observed in the content of N in the DIS process (Tables 1, 2). The above findings suggest that the DIS process inhibits soil urease activity and reduces the rate of soil N loss. Variation in the activity of sucrase and β-glucosidase was similar to variation in OC, suggesting this might be related to changes in the OC content. The significant high activity of β-glucosidase, cellulase, laccase, and amylase in the DIS process, demonstrate that these are common enzymes and are typically synthesized by edible fungi during cultivation (Cai et al., 1998; Sun et al., 2011; Shahryari et al., 2019). These enzymes can therefore promote can promote plant growth through their positive effects on the soil environment (Huang et al., 2010).

In this study, it was found that the BAS process enriched some soil mineral ions meaning that this process might replenish mineral ions in the soil. Among these ions, Ca is essential for plant growth, as it is required for the elongation of plant cells in both the shoots and roots (Burstrom, 1968). Ca can be released from sugarcane bagasse (Bento et al., 2019). The content of La has been shown to accelerate the conversion of soil N (Zhu et al., 2002), and it has no effect on plant growth at low concentrations (von Tucher and Schmidhalter, 2005; Hu et al., 2006). Thus, the BAS process enhances soil health by promoting increases in the content of Ca and La. The content of Fe, Cr, and Mo was lower in the BAS process than in the CK, and the content of P and Mo was lower in the DIS process than in the CK. Fe is an essential micronutrient in nearly all living organisms (Rout and Sahoo, 2015); thus, the loss of soil Fe caused by the BAS process might have a negative effect on sugarcane growth. Cr is a potentially toxic heavy metal that does not have any essential metabolic function in plants (Shahid et al., 2017); therefore, reductions in the content of Cr can enhance soil health. Mo is an essential element that plays a role in various biochemical processes in living organisms (Williams and Fraústo da Silva, 2002); although the content of Mo is low, it was much lower in the BAS process and the DIS process than in the CK. The same factors responsible for the decrease in the content of Mo in the BAS process and the DIS process might also explain the decrease in the P content in the BAS process and the DIS process.

Soil microbes play key roles in soil biogeochemical cycles (Basu et al., 2021). The present study found that the bacterial Shannon index and the number of unique OTUs were higher in the BAS process than in the CK and DIS process, which means that in this study BAS process regulates soil microbial diversity and thus improved soil nutrients. This is consistent with the results of previous studies (Silvia et al., 2014; Bhadha et al., 2020; Ullah et al., 2020). In addition, a previous study of fairy ring soil ecology has shown that increases in the abundance of single dominant fungi in soil decreases the alpha diversity index (Duan and Bau, 2021).
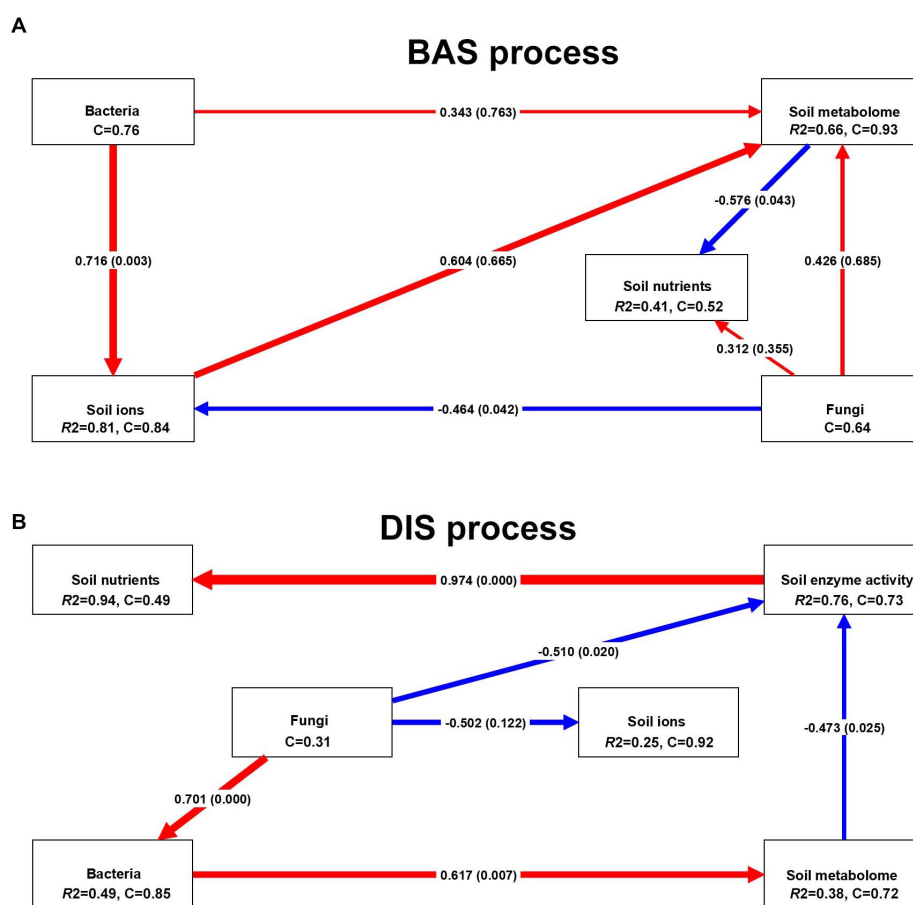
FIGURE 4
PLS–PM analysis of several soil factors, including soil enzyme activity, ions, metabolome, bacteria, and fungi, and soil nutrients in the BAS process and the DIS process. Each box corresponds to a latent variable. C indicates the Cronbach alpha index, and R2 indicates the R-squared index. The red and blue lines indicate positive and negative effects, respectively. The number in the middle of the line indicates the path coefficient, and the p-values are in parentheses; the thickness of the line is related to the absolute value of the path coefficient. **(A)** stands for BIS processing, and **(B)** stands for DIS processing.

Correspondingly, the present findings shown the Shannon index of fungi was lower in the DIS process than in the CK and the BAS process, and this likely stems from the abundance of DI in the soil in the DIS process, which indicated DIS process resulted in a single fungal species, i.e., *Dictyophora indusiata*, dominant in the soil, therefore regulating soil chemistry and metabolism active. Specifically, increases in the bacterial Shannon index in the BAS process were related to increases in soil nutrients, and decreases in the fungal Shannon index in the DIS process were related to decreases in soil nutrients. In light of a previous study suggesting that soil microbes can mediate the transformation of the aforementioned soil nutrients (Basu et al., 2021), we suggest that this finding might stem from the fact that the BAS process increases both bacterial alpha diversity and the absolute abundance of bacteria, whereas the DIS process might inhibit bacterial absolute abundance because of the growth of DI in soil. Additional quantitative microbial surveys are needed to verify this hypothesis.

Soil metabolomics can be used to clarify correlations of metabolic fingerprints with environmental factors, soil nutrients, microbial diversity, and plant phenotypes (Withers et al., 2020). It was found that a large amount of carbohydrate metabolites were consumed in the BAS process,

and the consumption of carbohydrate metabolites was reduced in the DIS process. Carbohydrates are a significant component of the organic matter in all soils and commonly account for 5 to 20% of soil organic matter (Lowe, 1978); carbohydrates are thus often used as indicators of the physical quality of soil (Zaher et al., 2020). A previous study of the effects of fairy ring fungi on soil health has shown that increases in carbohydrates in soil are associated with the enhanced growth of Leymus chinensis (Duan et al., 2022); therefore, our findings indicate that the nutrient status of soils in the DIS process was superior to that in the BAS process. In addition, the content of carbohydrates was negatively correlated with the content of P, AP, and BN and positively correlated with the content of OC. Therefore, the consumption of soil OC in the BAS process might increase the content of P, AP, and BN through the consumption of soil carbohydrates, and this might be achieved via increases in soil bacterial diversity in the BAS process in light of a previous study suggesting that soil microbes play a key role in soil biogeochemical cycling and thus regulate the availability of soil mineral nutrients (Basu et al., 2021).

The identity of the most abundant soil metabolites was consistent with the above results. Carbohydrates were not the only important class of metabolites affected by the BAS process and the DIS process, but they were some of the most abundant metabolites. D(+)-talose 1 was
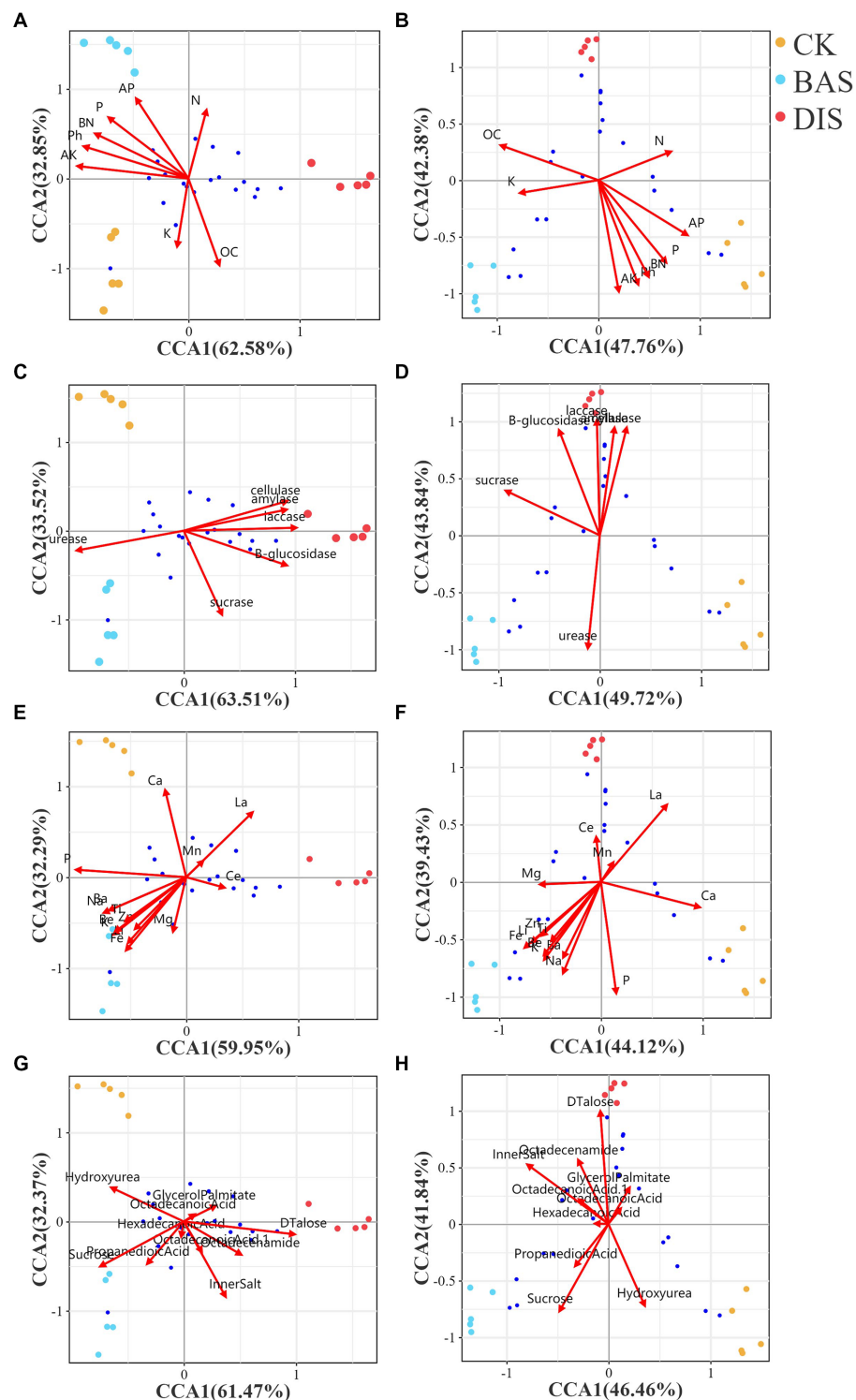
FIGURE 5
CCA with soil bacterial and fungal diversity in the BAS process and the DIS process as response variables and several factors related to soil chemistry, enzyme activity, soil ions, and the metabolome as explanatory variables. Dots of different colors indicate soil samples; the legend is located at the top right. The axes of each graph show the percentage of the total variation in OTU abundance explained. Arrows correspond to environmental factors, and the length of the arrows is positively related to the effect of environmental factors on the distribution of OTUs. Blue dots indicate OTUs; the proximity of the OTUs is positively related to their abundances in samples. CCA with soil bacterial and fungal diversity related to soil chemistry **(A** and **B)**, enzyme activity **(C** and **D)**, soil ions, **(E** and **F)** and the metabolome **(G** and **H)** were shown from top to bottom in the figure.

the most abundant carbohydrate metabolite, and variation in its abundance was consistent with the effects of the BAS process and the DIS process on carbohydrate metabolites (down-regulated in the BAS process and up-regulated in the DIS process). A previous study has shown that the content of D-talose is affected by root rot disease; this suggests that D-talose could be used as an indicator of soil health.

PLS–PM and CCA showed that the inhibition of urease in the DIS process was the main factor affecting soil nutrients, as soil enzyme activity is the main regulatory factor of soil nutrients and then the activity of urease was higher than that of all soil enzymes combined, which suggest cultivation of *Dictyophora indusiata* in sugarcane field effectively could improve soil health. Given that the activity of urease mainly affects the soil N cycle (Fu et al., 2020), other factors might jointly regulate the content of other nutrients, such as P and AK. This might be related to the effect of the DIS process on the abundance of soil bacteria, given that increases in bacterial diversity were associated with increases in soil nutrients. According to PLS–PM (Figure 4B), enzyme activity affected soil nutrients, and this was achieved via the effects of bacteria on the soil metabolome, which were mediated by the effects of fungi on bacteria. D(+)-talose 1 was correlated with bacterial and fungal diversity in the DIS process according to CCA (Figures 5G,H), indicating that it is a key biomarker and was affected by the DIS process. The DIS process inhibits the activity of urease and thus reduces the rate of soil N loss. Our findings thus suggest that D(+)-talose 1 might be related to the regulation of soil urease. Future studies of soil fertility are needed to clarify whether D(+)-talose 1 inhibits urease activity.

5. Conclusion

The present study evaluated the effects of a sugarcane–DI intercropping system and bagasse application on soil health. The findings indicate that the sugarcane–DI system can promote soil health. This study found that this intercropping system could reduce soil N loss, and also explored potential mechanisms underlying these observations, and these findings have implications for the development of novel soil urease inhibitors. In future studies, we plan to characterize the effects of DI on the abundances of soil microbes and the relationship between soil urease activity and the metabolite D(+)-talose 1.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/genbank/, PRJNA924520.

Author contributions

MD, ZW, and TH contributed to the conception and design of the study. MD, YL, GZ, SL, XL, XW, HH, JQ, SQ, BF, Q-HL, CL, QL, and

LW performed investigation. MD and YL performed the statistical analysis. MD wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1193990/full#supplementary-material

References

Bao, S. (2000). *Soil agrochemical analysis. 3rd Edn.* Beijing, China: China Agriculture Press.

Basu, S., Kumar, G., Chhabra, S., and Prasad, R. (2021). "Chapter 13—role of soil microbes in biogeochemical cycle for enhancing soil fertility" in *New and future developments in microbial biotechnology and bioengineering.* eds. J. P. Verma, C. A. Macdonald, V. K. Gupta and A. R. Podile (Elsevier), 149–157. doi: 10.1016/B978-0-444-64325-4.00013-4

Bento, L. R., Castro, A. J. R., Moreira, A. B., Ferreira, O. P., Bisinoti, M. C., and Melo, C. A. (2019). Release of nutrients and organic carbon in different soil types from hydrochar obtained using sugarcane bagasse and vinasse. *Geoderma* 334, 24–32. doi: 10.1016/j.geoderma.2018.07.034

Bhadha, J. H., Xu, N., Khatiwada, R., Swanson, S., and LaBorde, C. (2020). Bagasse: a potential organic soil amendment used in sugarcane production: SL477/SS690, 8/2020. *EDIS* 2020:5. doi: 10.32473/edis-ss690-2020

Burstrom, H. G. (1968). Calcium and plant growth. *Biol. Rev.* 43, 287–316. doi: 10.1111/j.1469-185X.1968.tb00962.x

Cai, Y. J., Buswell, J. A., and Chang, S. T. (1998). β-Glucosidase components of the cellulolytic system of the edible straw mushroom, Volvariella volvacea. *Enzym. Microb. Technol.* 22, 122–129. doi: 10.1016/S0141-0229(97)00151-8

Cantarella, H., Otto, R., Soares, J. R., and Silva, A. G. D. B. (2018). Agronomic efficiency of NBPT as a urease inhibitor: a review. *J. Adv. Res.* 13, 19–27. doi: 10.1016/j.jare.2018.05.008

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Chen, H., and Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinf.* 12:35. doi: 10.1186/1471-2105-12-35

Cui, M., Zeng, L., Qin, W., and Feng, J. (2020). Measures for reducing nitrate leaching in orchards: a review. *Environ. Poll. (Barking, Essex: 1987)* 263:114553. doi: 10.1016/j.envpol.2020.114553

Dalian Environmental Monitoring Center (2014). Soil - Determination of organic carbon - Combustion oxidation nondispersive infrared absorption method. PRC National Standard.

Du, S., Gao, X., Li, H., Lu, R., and Jiang, Q. (2004). *Determination of exchangeable potassium and non-exchangeable potassium content in soil*. Beijing, China:Ministry of Agriculture of the People's Republic of China.

Duan, M., and Bau, T. (2021). Grassland fairy rings of Leucocalocybe mongolica represent the center of a rich soil microbial community. *Braz. J. Microbiol.* 52, 1357–1369. doi: 10.1007/s42770-021-00478-3

Duan, M., Long, S., Wu, X., Feng, B., Qin, S., Li, Y., et al. (2023). Genome, transcriptome, and metabolome analyses provide new insights into the resource development in an edible fungus Dictyophora indusiata. *Front. Microbiol.* 14:1137159. doi: 10.3389/fmicb.2023.1137159

Duan, M., Lu, M., Lu, J., Yang, W., Li, B., Ma, L., et al. (2022). Soil chemical properties, metabolome, and metabarcoding give the new insights into the soil transforming process of fairy ring fungi leucocalocybe mongolica. *J. Fungi* 8:680. doi: 10.3390/jof8070680

Duan, M., Lu, J., Yang, W., Lu, M., Wang, J., Li, S., et al. (2022). Metabarcoding and metabolome analyses reveal mechanisms of Leymus chinensis growth promotion by fairy ring of Leucocalocybe mongolica. *J. Fungi* 8:944. doi: 10.3390/jof8090944

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Eivazi, F., and Tabatabai, M. A. (1988). Glucosidases and galactosidases in soils. *Soil Biol. Biochem.* 20, 601–606. doi: 10.1016/0038-0717(88)90141-1

Fu, Q., Abadie, M., Blaud, A., Carswell, A., Misselbrook, T. H., Clark, I. M., et al. (2020). Effects of urease and nitrification inhibitors on soil N, nitrifier abundance and activity in a sandy loam soil. *Biol. Fertil. Soils* 56, 185–194. doi: 10.1007/s00374-019-01411-5

Fu, S., Wang, Y., Zhang, R., Zhao, M., and Jiang, Y. (1988). *Method for determination of total potassium in soils*. Beijing, China:China Standards Press.

Gunina, A., and Kuzyakov, Y. (2015). Sugars in soil and sweets for microorganisms: review of origin, content, composition and fate. *Soil Biol. Biochem.* 90, 87–100. doi: 10.1016/j.soilbio.2015.07.021

Guo, M., Wu, F., Hao, G., Qi, Q., Li, R., Li, N., et al. (2017). *Bacillus subtilis* improves immunity and disease resistance in rabbits. *Front. Immunol.* 8:354. doi: 10.3389/fimmu.2017.00354

Hendrickson, L. L., and Douglass, E. A. (1993). Metabolism of the urease inhibitor n-(n-butyl)thiophosphoric triamide (nbpt) in soils. *Soil Biol. Biochem.* 25, 1613–1618. doi: 10.1016/0038-0717(93)90017-6

Hu, X., Wang, X.-R., and Wang, C. (2006). Bioaccumulation of lanthanum and its effect on growth of maize seedlings in a red loamy soil. *Pedosphere* 16, 799–805. doi: 10.1016/S1002-0160(06)60116-1

Huang, C., Deng, L., Gao, X., Zhang, S., Luo, T., and Ren, Q. (2010). Effects of fungal residues return on soil enzymatic activities and fertility dynamics in a paddy soil under a rice-wheat rotation in Chengdu plain. *Soil Tillage Res.* 108, 16–23. doi: 10.1016/j.still.2010.03.011

Jiao, R., Dong, Y., and Sun, Q. (2015). *Nitrogen determination methods of forest soils*. Beijing, China:China Standards Press.

Li, S., Lei, Y., Zhang, Y., Liu, J., Shi, X., Jia, H., et al. (2019). Rational trade-offs between yield increase and fertilizer inputs are essential for sustainable intensification: a case study in wheat-maize cropping systems in China. *Sci. Total Environ.* 679, 328–336. doi: 10.1016/j.scitotenv.2019.05.085

Li, Z., Luo, Y., and Teng, Y. (2008). *Soil and environmental microbiology methods*. Beijing, China: Science Press.

Lin, X. (2010). *Principles and methods of soil microbial research*. Beijing, China: Higher Education Press.

Lowe, L. E. (1978). "Chapter 2 carbohydrates in soil" in *Developments in soil science*. eds. M. Schnitzer and S. U. Khan (Elsevier), 65–93. doi: 10.1016/S0166-2481(08)70017-5

Mohamed, I., Ali, M., Ahmed, N., and Chen, F. (2019). Cadmium immobilization and alleviation of its toxicity for soybean grown in a clay loam contaminated soil using sugarcane bagasse-derived biochar. *Environ. Sci. Pollut. Res.* 26, 21849–21857. doi: 10.1007/s11356-019-05501-7

Nilsson, R. H., Larsson, K. H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., et al. (2019). The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucl. Acids Res.* 47, d259–d264. doi: 10.1093/nar/gky1022

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucl. Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864

Rout, G. R., and Sahoo, S. (2015). Role of IRON in plant growth and metabolism. *Rev. Agric. Sci.* 3, 1–24. doi: 10.7831/ras.3.1

Seleiman, M. F., and Kheir, A. M. S. (2018). Saline soil properties, quality and productivity of wheat grown with bagasse ash and thiourea in different climatic zones. *Chemosphere* 193, 538–546. doi: 10.1016/j.chemosphere.2017.11.053

Shahid, M., Shamshad, S., Rafiq, M., Khalid, S., Bibi, I., Niazi, N. K., et al. (2017). Chromium speciation, bioavailability, uptake, toxicity and detoxification in soil-plant system: a review. *Chemosphere* 178, 513–533. doi: 10.1016/j.chemosphere.2017.03.074

Shahryari, Z., Fazaelipoor, M. H., Ghasemi, Y., Lennartsson, P. R., and Taherzadeh, M. J. (2019). Amylase and xylanase from edible fungus Neurospora intermedia: production and characterization. *Molecules* 24:721. doi: 10.3390/molecules24040721

Sharma, S., Thind, H. S., Singh, Y., Singh, V., and Singh, B. (2015). Soil enzyme activities with biomass ashes and phosphorus fertilization to rice–wheat cropping system in the indo-Gangetic plains of India. *Nutr. Cycl. Agroecosyst.* 101, 391–400. doi: 10.1007/s10705-015-9684-7

Silvia, S., Miura, T., Nobuhiro, K., Fujie, K., Hasanuddin, U., Niswati, A., et al. (2014). Soil microbial biomass and diversity amended with bagasse mulch in tillage and no-tillage practices in the sugarcane plantation. *Procedia Environ. Sci.* 20, 410–417. doi: 10.1016/j.proenv.2014.03.052

Singh, R. K., Singh, P., Li, H. B., Song, Q. Q., Guo, D. J., Solanki, M. K., et al. (2020). Diversity of nitrogen-fixing rhizobacteria associated with sugarcane: a comprehensive study of plant-microbe interactions for growth enhancement in Saccharum spp. *BMC Plant Biol.* 20:220. doi: 10.1186/s12870-020-02400-9

Soil Testing-Part7 (2014). *Method for determination of avalible phosphorus in soil*. Beijing, China:Ministry of Agriculture of the People's Republic of China.

Sun, S. J., Liu, J. Z., Hu, K. H., and Zhu, H. X. (2011). The level of secreted laccase activity in the edible fungi and their growing cycles are closely related. *Curr. Microbiol.* 62, 871–875. doi: 10.1007/s00284-010-9794-z

Thaisuchat, H., Karuehanon, W., Boonkorn, P., Meesumlee, J., Malai, S., and Ruttanateerawichien, K. (2022). Bamboo waste recycling using Dictyophora indusiata mycelia cultivation. *Int. J. Recycl. Org. Waste Agriculture*. doi: 10.30486/IJROWA.2022.1955758.1443

Toju, H., Tanabe, A. S., Yamamoto, S., and Sato, H. (2012). High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PLoS One* 7:e40863. doi: 10.1371/journal.pone.0040863

Tong, L., Li, H., Liu, X., Li, B., Chen, L., Chen, G., et al. (2021). Effects of continuous cropping of Dictyophora on soil physical and chemical properties, microbial biomass and enzyme activity. *J. Innovat. Soc. Sci. Res.* 8:6890. doi: 10.53469/jissr.2021.08(07).39

Ullah, Z., Ali, S., Muhammad, N., Khan, N., Rizwan, M., Khan, M. D., et al. (2020). Biochar impact on microbial population and elemental composition of red soil. *Arab. J. Geosci.* 13:757. doi: 10.1007/s12517-020-05671-6

Vitousek, P. M., Naylor, R., Crews, T., David, M. B., Drinkwater, L. E., Holland, E., et al. (2009). Agriculture. Nutrient imbalances in agricultural development. *Science* 324, 1519–1520. doi: 10.1126/science.1170261

von Tucher, S., and Schmidhalter, U. (2005). Lanthanum uptake from soil and nutrient solution and its effects on plant growth. *J. Plant Nutr. Soil Sci.* 168, 574–580. doi: 10.1002/jpln.200520506

Vranova, V., Rejsek, K., and Formanek, P. (2013). Aliphatic, cyclic, and aromatic organic acids, vitamins, and carbohydrates in soil: a review. *Sci. World J.* 2013:524239. doi: 10.1155/2013/524239

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wang, J., Wen, X., Zhang, Y., Zou, P., Cheng, L., Gan, R., et al. (2021). Quantitative proteomic and metabolomic analysis of Dictyophora indusiata fruiting bodies during post-harvest morphological development. *Food Chem.* 339:127884. doi: 10.1016/j.foodchem.2020.127884

Williams, R. J. P., and Fraústo da Silva, J. J. R. (2002). The involvement of molybdenum in life. *Biochem. Biophys. Res. Commun.* 292, 293–299. doi: 10.1006/bbrc.2002.6518

Withers, E., Hill, P. W., Chadwick, D. R., and Jones, D. L. (2020). Use of untargeted metabolomics for assessing soil quality and microbial function. *Soil Biol. Biochem.* 143:107758. doi: 10.1016/j.soilbio.2020.107758

Xiao, D., Huang, Y., Feng, S., Ge, Y., Zhang, W., He, X., et al. (2018). Soil organic carbon mineralization with fresh organic substrate and inorganic carbon additions in a

red soil is controlled by fungal diversity along a pH gradient. *Geoderma* 321, 79–89. doi: 10.1016/j.geoderma.2018.02.003

Xiao, G., Zhang, H., Su, F., and Yang, J. (1988). *Method for determination of soil total phosphorus*. Beijing, China:China Standards Press.

Xu, Q., Ling, N., Chen, H., Duan, Y., Wang, S., Shen, Q., et al. (2020). Long-term chemical-only fertilization induces a diversity decline and deep selection on the soil bacteria. *mSystems* 5:e00337. doi: 10.1128/mSystems.00337-20

Zaher, H., Sabir, M., Benjelloun, H., and Paul-Igor, H. (2020). Effect of forest land use change on carbohydrates, physical soil quality and carbon stocks in Moroccan cedar area. *J. Environ. Manag.* 254:109544. doi: 10.1016/j.jenvman.2019.109544

Zhang, H., Li, Y., and Zhu, J. K. (2018). Developing naturally stress-resistant crops for a sustainable agriculture. *Nat. Plants* 4, 989–996. doi: 10.1038/s41477-018-0309-4

Zhu, J. G., Chu, H. Y., Xie, Z. B., and Yagi, K. (2002). Effects of lanthanum on nitrification and ammonification in three Chinese soils. *Nutr. Cycl. Agroecosyst.* 63, 309–314. doi: 10.1023/A:1021163101692

Zotti, M., Bonanomi, G., Mancinelli, G., Barquero, M., De Filippis, F., Giannino, F., et al. (2021). Riding the wave: response of bacterial and fungal microbiota associated with the spread of the fairy ring fungus Calocybe gambosa. *Appl. Soil Ecol.* 163:103963. doi: 10.1016/j.apsoil.2021.103963

Check for updates

Genus-wide genomic characterization of *Macrococcus*: insights into evolution, population structure, and functional potential

Laura M. Carroll[1,2,3,4], Rian Pierneef[5], Thendo Mafuna[6],
Kudakwashe Magwedere[7] and Itumeleng Matle[8]*

[1]Department of Clinical Microbiology, SciLifeLab, Umeå University, Umeå, Sweden, [2]Laboratory for
Molecular Infection Medicine Sweden (MIMS), Umeå University, Umeå, Sweden, [3]Umeå Centre for
Microbial Research, Umeå University, Umeå, Sweden, [4]Integrated Science Lab, Umeå University, Umeå,
Sweden, [5]Biotechnology Platform, Agricultural Research Council, Onderstepoort Veterinary Research,
Onderstepoort, South Africa, [6]Department of Biochemistry, University of Johannesburg, Auckland Park,
South Africa, [7]Directorate of Veterinary Public Health, Department of Agriculture, Land Reform and
Rural Development, Pretoria, South Africa, [8]Bacteriology Division, Agricultural Research Council,
Onderstepoort Veterinary Research, Onderstepoort, South Africa

Introduction: *Macrococcus* species have been isolated from a range of mammals and mammal-derived food products. While they are largely considered to be animal commensals, *Macrococcus* spp. can be opportunistic pathogens in both veterinary and human clinical settings. This study aimed to provide insight into the evolution, population structure, and functional potential of the *Macrococcus* genus, with an emphasis on antimicrobial resistance (AMR) and virulence potential.

Methods: All high-quality, publicly available *Macrococcus* genomes ($n = 104$, accessed 27 August 2022), plus six South African genomes sequenced here (two strains from bovine clinical mastitis cases and four strains from beef products), underwent taxonomic assignment (using four different approaches), AMR determinant detection (via AMRFinderPlus), and virulence factor detection (using DIAMOND and the core Virulence Factor Database).

Results: Overall, the 110 *Macrococcus* genomes were of animal commensal, veterinary clinical, food-associated (including food spoilage), and environmental origins; five genomes (4.5%) originated from human clinical cases. Notably, none of the taxonomic assignment methods produced identical results, highlighting the potential for *Macrococcus* species misidentifications. The most common predicted antimicrobial classes associated with AMR determinants identified across *Macrococcus* included macrolides, beta-lactams, and aminoglycosides ($n=81$, 61, and 44 of 110 genomes; 73.6, 55.5, and 40.0%, respectively). Genes showing homology to *Staphylococcus aureus* exoenzyme aureolysin were detected across multiple species (using 90% coverage, $n=40$ and 77 genomes harboring aureolysin-like genes at 60 and 40% amino acid [AA] identity, respectively). *S. aureus* Panton-Valentine leucocidin toxin-associated *lukF-PV* and *lukS-PV* homologs were identified in eight *M. canis* genomes ($\geq$40% AA identity, >85% coverage). Using a method that delineates populations using recent gene flow (PopCOGenT), two species (*M. caseolyticus* and *M. armenti*) were composed of multiple within-species populations. Notably, *M. armenti* was partitioned into two populations, which differed in functional potential (e.g., one harbored beta-lactamase family, type II toxin-antitoxin system, and stress response proteins, while the other possessed a Type VII secretion system; PopCOGenT $p<0.05$).

Discussion: Overall, this study leverages all publicly available *Macrococcus* genomes in addition to newly sequenced genomes from South Africa to identify

genomic elements associated with AMR or virulence potential, which can be queried in future experiments.

1. Introduction

Members of the *Macrococcus* genus are Gram-positive, catalase-positive, oxidase-positive, and coagulase-negative cocci (Mazhar et al., 2018; Ramos et al., 2021). The *Macrococcus* genus is a member of the Staphylococcaceae family and was first proposed as a novel genus in 1998, when its four original species (*M. caseolyticus, M. equipercicus, M. bovicus*, and *M. carouselicus*) were differentiated from members of the closely related *Staphylococcus* genus using numerous genetic and phenotypic characteristics (e.g., 16S rDNA sequencing, DNA–DNA hybridization, pulsed field gel electrophoresis, oxidase activity, cell wall composition, plasmid profiles; Kloos et al., 1998; Mazhar et al., 2018). Since the four original *Macrococcus* spp. were described in 1998, eight additional *Macrococcus* spp. have been identified ($n = 12$ total validly published *Macrococcus* spp. per the List of Prokaryotic names with Standing in Nomenclature [LPSN], https://lpsn.dsmz.de/genus/macrococcus; accessed 10 December 2022; Parte et al., 2020): *M. brunensis* (Mannerova et al., 2003)*, M. hajekii* (Mannerova et al., 2003), *M. lamae* (Mannerova et al., 2003), *M. canis* (Gobeli Brawand et al., 2017), *M. bohemicus* (Maslanova et al., 2018), *M. epidermidis* (Maslanova et al., 2018), *M. goetzii* (Maslanova et al., 2018), and *M. armenti* (Keller et al., 2022).

Macrococcus spp. have historically been viewed as animal commensals (Mazhar et al., 2018) and have been isolated from a range of mammals (e.g., the skin of cows, pigs, horses, llamas, dogs) and the products derived from them (e.g., dairy products and meat; Kloos et al., 1998; Mannerova et al., 2003; Cotting et al., 2017; Mazhar et al., 2018; Ramos et al., 2021; Keller et al., 2022). However, the role of *Macrococcus* spp. as opportunistic pathogens has been discussed increasingly in recent years (MacFadyen et al., 2018; Ramos et al., 2021). In veterinary clinical settings, *Macrococcus* spp. have been isolated from infections (e.g., mastitis, otitis, and dermatitis cases, abscesses) in numerous animals, including cattle, sheep, and dogs (Gomez-Sanz et al., 2015; Cotting et al., 2017; Schwendener et al., 2017; Ramos et al., 2021). Notably, in 2018, *Macrococcus* spp. were reportedly isolated from human clinical samples for the first time, when *M. goetzii, M. epidermidis, M. bohemicus*, and *M. caseolyticus* subsp. *hominis* were isolated from infections at several body sites (i.e., wound sites, gynecological cases, and mycoses cases; Maslanova et al., 2018). Since then, *M. canis* has additionally been isolated from a human clinical case (i.e., a skin infection; Jost et al., 2021).

In addition to their pathogenic potential, some *Macrococcus* spp. carry antimicrobial resistance (AMR) genes (Schwendener et al., 2017; MacFadyen et al., 2018; Mazhar et al., 2018; Jost et al., 2021; Ramos et al., 2021). Methicillin resistance in *Macrococcus* spp. is of particular concern, as several mobilizable methicillin resistance determinants (e.g., penicillin-binding protein homologs *mecB, mecD*) have been identified in *Macrococcus* spp. (MacFadyen et al., 2018; Mazhar et al., 2018; Ramos et al., 2021). In this context, methicillin-resistant *Macrococcus* strains become particularly concerning: not only can they

potentially serve as opportunistic human and veterinary pathogens, but they can potentially transfer mobilizable AMR genes to other organisms, including taxa with a higher virulence potential (e.g., pathogenic *Staphylococcus aureus*; MacFadyen et al., 2018; Mazhar et al., 2018; Ramos et al., 2021; Schwendener and Perreten, 2022).

Several studies have employed genomic approaches to gain insight into the evolution and population structure of *Macrococcus*; however, these studies relied on a limited number of genomes (Maslanova et al., 2018; Schwendener and Perreten, 2022) and/or focused on specific taxa within the genus (e.g., *M. caseolyticus*; MacFadyen et al., 2018; Zhang et al., 2022). Furthermore, very few studies—genomic or otherwise—describing *Macrococcus* spp. strains isolated in Africa are available (Tshipamba et al., 2018; Ouoba et al., 2019; Ali et al., 2022). Here, we used whole-genome sequencing (WGS) to characterize six *Macrococcus* spp. strains isolated from bovine-associated sources in South Africa. To gain insight into *Macrococcus* at a genomic scale, we compare our six genomes to all publicly available *Macrococcus* genomes ($n = 110$ total genomes). Overall, our study provides insight into the evolution, population structure, and functional potential of all species—both validly published and putative novel—within the *Macrococcus* genus in its entirety.

2. Materials and methods

2.1. Strain isolation

Macrococcus strains sequenced in this study were isolated from bovine clinical mastitis samples ($n = 2$) and beef products ($n = 4$) and submitted to the Onderstepoort Veterinary Research (OVR) General Bacteriology Laboratory for routine diagnostic services (Supplementary Table S1). From each sample, 10 g (ratio 1:10) were homogenized in buffered peptone water, and then aliquots of 0.1 mL were inoculated onto Baird-Parker agar and Brilliance MRSA 2 agar (both Oxoid, ThermoFisher, Johannesburg) and incubated for 24 h at 37°C. Presumptive macrococci colonies were streaked onto blood agar supplemented with 5% sheep blood (Oxoid, ThermoFisher, Johannesburg), incubated for 24 h at 37°C, and identified by phenotypic characteristics as described by Poyart et al. (2001). Briefly, Gram staining, catalase test, hemolysis, coagulase test, and API 32 ID STAPH (bioMérieux) were used to identify the isolates as macrococci.

2.2. Genomic DNA extraction, whole-genome sequencing, data pre-processing, and quality control

Genomic DNA was prepared from overnight cultures using the QIAGEN® DNeasy blood and tissue kit (Germany) according to the manufacturer's instructions (see section "Strain isolation" above;

Supplementary Table S1). WGS of isolates was performed at the Biotechnology Platform, Agricultural Research Council, Onderstepoort, South Africa. DNA libraries were prepared using TruSeq DNA library preparation kits (Illumina, San Diego, CA, USA), followed by sequencing on a HiSeq 2500 instrument (Illumina, San Diego, CA, USA).

Raw Illumina paired-end reads derived from each of the strains isolated here ($n = 6$) were supplied as input to Trimmomatic v0.38 (Bolger et al., 2014). Trimmomatic was used to remove Illumina adapters (ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:2:keepBothReads), leading and trailing low quality or N bases (i.e., Phred quality <3; LEADING:3 TRAILING:3), and reads <36 bp in length (MINLEN:36). FastQC v0.11.9[1] was used to evaluate the quality of the resulting trimmed paired-end reads (Supplementary Table S2).

The resulting trimmed paired-end reads associated with each strain were assembled into contigs via Shovill v1.1.0,[2] using the following parameters (all other parameters were set to their default values): (i) SKESA v2.4.0 (Souvorov et al., 2018) as the assembler ("--assembler skesa"); (ii) a minimum contig length of 200 ("--minlen 200"); (iii) a minimum contig coverage value of 10 ("--mincov 10"). QUAST v5.0.2 (Gurevich et al., 2013) was used to evaluate the quality of each resulting assembled genome (using a minimum contig length parameter of 1 bp), and the "lineage_wf" workflow in CheckM v1.1.3 (Parks et al., 2015) was used to evaluate genome completeness and contamination. MultiQC v1.12 (Ewels et al., 2016) was used to evaluate the quality of all six *Macrococcus* genomes in aggregate (Supplementary Tables S1, S2).

2.3. Acquisition and quality control of publicly available *Macrococcus* spp. genomes

All publicly available GenBank genomes submitted to the National Center for Biotechnology Information (NCBI) Assembly database as members of *Macrococcus* were downloaded ($n = 102$ genomes; accessed 27 August 2022; Kitts et al., 2016; Schoch et al., 2020). Additionally, all genomes assigned to the *Macrococcus* genus within the Genome Taxonomy Database (GTDB) v207 (Parks et al., 2022), which were not included in the initial set of 102 genomes, were downloaded ($n = 8$ of 88 total GTDB genomes). Together, this search of NCBI and GTDB yielded a preliminary set of 110 publicly available, putative *Macrococcus* genomes.

All 116 putative *Macrococcus* genomes (i.e., 110 publicly available genomes, plus the six genomes sequenced here) were characterized using QUAST and CheckM as described above (see section "Genomic DNA extraction, whole-genome sequencing, data pre-processing, and quality control" above). Six publicly available *Macrococcus* genomes showcased CheckM completeness <95% and/or QUAST N50 < 20 Kbp; these genomes were excluded from further analysis ($n = 104$ publicly available genomes used in subsequent analyses; Supplementary Table S3). One genome (NCBI GenBank

---

1　https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

2　https://github.com/tseemann/shovill

Assembly accession GCA_002119805.1) had >5% CheckM contamination (i.e., 5.11%; Supplementary Table S3). However, because this genome represented the type strain of *M. canis* and was a complete genome, it was used in subsequent steps. Overall, after removing low-quality genomes, the search of NCBI and GTDB, in combination with the six genomes sequenced here, yielded a final set of 110 *Macrococcus* genomes used in subsequent steps (Supplementary Tables S1, S3).

2.4. Taxonomic assignment

The following genomospecies delineation methods were applied to the set of 110 *Macrococcus* genomes (i.e., all 104 high-quality, publicly available *Macrococcus* genomes, plus the six genomes sequenced here; Supplementary Tables S1, S3): (i) the Genome Taxonomy Database Toolkit (GTDB-Tk), a popular genomospecies delineation tool, which relies primarily on a 95 average nucleotide identity (ANI) genomospecies threshold; (ii) bactaxR, which uses pairwise ANI values calculated between a set of genomes to delineate genomospecies *de novo* at any user-specified genomospecies threshold; (iii) the specI taxonomy, a marker gene-based taxonomic assignment approach; (iv) PopCOGenT (Populations as Clusters Of Gene Transfer; Arevalo et al., 2019), a method that relies on a metric of recent gene flow to identify species units. Each of these methods is explained in detail below.

For the GTDB-Tk workflow, all 110 *Macrococcus* genomes were assigned to species using the GTDB-Tk v2.1.0 "classify_wf" workflow (default settings) and version R207_v2 of GTDB (Chaumeil et al., 2019; Parks et al., 2022). GTDB-Tk confirmed that all 110 genomes identified here belonged to the *Macrococcus* genus (i.e., either "g__Macrococcus" or "g__Macrococcus_B," per GTDB's nomenclature; these corresponded to the only two GTDB genus designations, which contained the term "*Macrococcus*"; Supplementary Table S4).

For the bactaxR workflow, pairwise ANI values were calculated between all 110 *Macrococcus* genomes using the command-line implementation of OrthoANI v1.40 (Lee et al., 2016) with default settings (Supplementary Table S5). The resulting pairwise ANI values were supplied as input to the bactaxR v0.2.1 package (Carroll et al., 2020) in R v4.1.2 (R Core Team, 2021); bactaxR was used to construct a dendrogram and graph of all genomes based on pairwise ANI (dis)similarities, using the ANI.dendrogram and ANI.graph functions, respectively. bactaxR's ANI.dendrogram function was further used to construct *de novo* genomospecies clusters using a 95 ANI genomospecies threshold (selected because this genomospecies threshold has been widely adopted by the microbiological community; Supplementary Table S6; Jain et al., 2018). OrthoANI was additionally used to calculate ANI values between all 110 *Macrococcus* genomes identified here (query genomes) relative to all *Macrococcus* spp. type strain genomes available in NCBI (reference genomes, $n = 16$ type strain genomes, accessed 4 October 2022; Supplementary Tables S3, S5).

For the specI workflow, each *Macrococcus* genome was assigned to a marker gene-based species cluster (specI cluster) using classify-genomes (https://github.com/AlessioMilanese/classify-genomes; accessed 3 June 2020; Milanese et al., 2019) and version 3 of the specI taxonomy (Mende et al., 2013). specI clusters reported by classify-genomes were treated as species assignments (Supplementary Table S7).

For the PopCOGenT workflow, the "PopCOGenT" module within PopCOGenT (latest version downloaded 31 August 2022; Arevalo et al., 2019) was used to identify gene flow units among all 110 *Macrococcus* genomes. The resulting "main clusters" reported by PopCOGenT (i.e., gene flow units, which attempt to mimic the classical species definition used for animals and plants) were treated as species assignments (Supplementary Table S8; Arevalo et al., 2019). Two PopCOGenT main clusters (i.e., Main Clusters 0 and 2; Supplementary Table S8) contained >1 subcluster (i.e., within-species populations identified via PopCOGenT, referred to hereafter as "subclusters"); each of these main clusters was additionally queried individually using the "flexible genome sweeps" module in PopCOGenT to identify subcluster-specific orthologues, using an "alpha" (significance) value of 0.05 (Supplementary Tables S9, S10; Arevalo et al., 2019).

2.5. *In silico* multi-locus sequence typing

Each of the 110 *Macrococcus* genomes (Supplementary Tables S1, S3; see section "Acquisition and quality control of publicly available *Macrococcus* spp. genomes" above) was supplied as input to mlst v2.22.0[3] for *in silico* multi-locus sequence typing (MLST). Default settings were used so that mlst could auto-select a MLST scheme from PubMLST (Jolley and Maiden, 2010; Jolley et al., 2018). Of the 110 genomes, 62 and 23 genomes were queried using the *M. caseolyticus* ("mcaseolyticus") and *M. canis* ("mcanis") PubMLST schemes, respectively; for 25 genomes, no scheme could be applied (Supplementary Table S11).

2.6. Genome annotation

Prokka v1.14.6 (Seemann, 2014) was used to annotate each *Macrococcus* genome ($n = 110$, Supplementary Tables S1, S3; see section "Acquisition and quality control of publicly available *Macrococcus* spp. genomes" above), using the "Bacteria" database and default settings. The ".gff" and ".faa" files produced by Prokka, along with the assembled contigs associated with each strain, were supplied as input to AMRFinderPlus v3.10.40 (Feldgarden et al., 2019), which was used to identify antimicrobial resistance (AMR) determinants in each genome, using the "plus" option ("--plus," i.e., to enable a search of the extended AMRFinderPlus database, which includes genes involved in virulence, biocide, heat, metal, and acid resistance) and the Prokka annotation format ("--annotation_format prokka"; Supplementary Table S12).

Amino acid (AA) sequences of virulence factors in the Virulence Factor Database (VFDB) core database (Liu et al., 2019) were downloaded ($n = 4,188$ AA sequences in the VFDB core database; accessed 4 September 2022). CD-HIT v4.8.1 (Li and Godzik, 2006; Fu et al., 2012) was used to cluster all VFDB core database AA sequences using the "cd-hit" command, a sequence identity threshold of 0.4 ("-c 0.4"), and a word length of 2 ("-n 2," the word size recommended for a 0.4 sequence identity threshold; https://github.com/weizhongli/

cdhit/blob/master/doc/cdhit-user-guide.wiki). The "makedb" command in DIAMOND v2.0.15 (Buchfink et al., 2015) was used to construct a DIAMOND database of the VFDB core database in its entirety, and the "diamond blastp" command was used to query AA sequences derived from each *Macrococcus* genome (i.e., ".faa" files produced by Prokka) against the entire VFDB core database, using the following parameters (default values were used for all other parameters): ultra-sensitive mode ("--ultra-sensitive"), one reported maximum target sequence ("--max-target-seqs 1," corresponding to the best match produced by DIAMOND: https://github.com/bbuchfink/diamond/issues/29), a minimum percent AA identity threshold of 60% ("--id 60"), and a minimum subject coverage threshold of 50% ("--subject-cover 50"). Each search was repeated using all combinations of (i) minimum percent AA identity thresholds of 0, 40, and 60%, and (ii) minimum subject coverage thresholds of 50 and 90% (Supplementary Tables S13–S18). Because many VFDB virulence factors are composed of multiple genes, and because some genes in VFDB may be highly similar/redundant, virulence factor presence and absence was considered at the whole virulence factor level, where a gene within a given virulence factor was considered to be "present" if any gene within its CD-HIT cluster could be detected in a given genome using DIAMOND. For example, the *S. aureus* exotoxin Panton-Valentine leukocidin (PVL) is a two-component toxin (Loffler et al., 2010; Shallcross et al., 2013). In the VFDB core database, PVL (VFDB ID VF0018) is composed of two genes: *lukF-PV* and *lukS-PV* (VFDB IDs VFG001276 and VFG001277, respectively). If any gene within the CD-HIT cluster of *lukF-PV* was detected in a *Macrococcus* genome, *lukF-PV* was considered "present"; likewise, if any gene within the CD-HIT cluster of *lukS-PV* was detected, *lukS-PV* was considered "present." If both genes were "present," PVL as a whole was considered to be 100% present. If one gene was "present," PVL was considered to be 50% present. If neither gene was "present," PVL was absent (0% present).

Biosynthetic gene clusters (BGCs) were detected in all 110 *Macrococcus* genomes using the command-line implementations of: (i) antiSMASH v6.1.0, using the "bacteria" taxon option ("--taxon bacteria") and gene finding via Prodigal's metagenomic mode option ("--genefinding-tool prodigal-m"; Blin et al., 2021); (ii) GECCO v0.9.2, using the "gecco run" command and the cluster probability threshold lowered to 0.3 ("-m 0.3"; all other settings were set to their defaults; Carroll et al., 2021). GenBank files (".gbk") for all BGCs identified by antiSMASH and GECCO were supplied as input to BiG-SCAPE v1.1.2 (Navarro-Munoz et al., 2020), which was used to cluster the 309 BGCs identified here, as well as experimentally validated BGCs in the MIBiG v2.1 database ("--mibig") into Gene Cluster Families (GCFs) using default parameter values (Supplementary Table S19; Kautsar et al., 2020).

2.7. Genus-level phylogeny construction

Panaroo v1.2.7 (Tonkin-Hill et al., 2020) was used to identify orthologous gene clusters and construct a core genome alignment ("-a") among the 110 *Macrococcus* genomes (see section "Acquisition and quality control of publicly available *Macrococcus* spp. genomes" above), plus *S. aureus* str. DSM 20231 as an outgroup genome (NCBI RefSeq Assembly accession GCF_001027105.1; $n = 111$ total genomes). The following input/parameters were used (all other parameters were

---

3 https://github.com/tseemann/mlst

set to their default values): (i) each genome's ".gff" file produced by Prokka as input (see section "Genome annotation" above); (ii) MAFFT as the aligner ("--aligner mafft"; Katoh and Standley, 2013); (iii) strict mode ("--clean-mode strict"); (iv) a core genome threshold of 95% ("--core_threshold 0.95"); (v) a protein family sequence identity threshold of 50% ("-f 0.5"). The core gene alignment produced by Panaroo ("core_gene_alignment.aln") was supplied as input to IQ-TREE v1.5.4 (Nguyen et al., 2015), which was used to construct a maximum likelihood (ML) phylogeny, using the General Time-Reversible (GTR) nucleotide substitution model ("-m GTR"; Tavaré, 1986) and 1,000 replicates of the ultrafast bootstrap approximation ("-bb 1,000"; Minh et al., 2013). The resulting ML phylogeny was visualized using the iTOL v6 webserver (https://itol.embl.de/; Letunic and Bork, 2021).

The genus-level phylogeny produced using Panaroo was compared to genus-level trees constructed using other methods, specifically: (i) PEPPAN (Zhou et al., 2020), a pipeline that can construct pan-genomes from genetically diverse bacterial genomes (e.g., spanning the diversity of an entire genus), and (ii) GTDB-Tk, which, in addition to taxonomic assignment, produces a multiple sequence alignment (MSA) of 120 bacterial marker genes detected in all input genomes (Chaumeil et al., 2019). For (i) PEPPAN, ".gff" files produced by Prokka were used as input ($n = 111$ total genomes, including the *S. aureus* outgroup; see section "Genome annotation" above). Default settings were used, except for the "--match_identity" option (the minimal identity of an alignment to be considered during pan-genome construction), which was set to "0.4," and the "--orthology" option (the algorithm for separating paralogous genes from orthologous genes), which was set to "ml" (i.e., the maximum-likelihood algorithm, reportedly the most accurate; Zhou et al., 2020). The PEPPAN_parser command was used to produce a Core Genome Allelic Variation (CGAV) tree (using a core genome threshold of 95%; "-a 95"), a gene presence/absence tree ("--tree"), and pan- and core-genome rarefaction curves ("--curve"; Simonsen et al., 2008; Tettelin et al., 2008; Camacho et al., 2009; Price et al., 2010; Steinegger and Soding, 2017). All aforementioned PEPPAN/PEPPAN_parser steps were repeated three separate times: (a) once as described above, but without the outgroup genome, and (b) using a lower minimal identity threshold (i.e., 20%, "--match_identity 0.2"), with and without the outgroup genome. The resulting trees were annotated using iTOL, and the resulting rarefaction curves were plotted in R using ggplot2 v3.4.0 (Supplementary Figures S1–S6; Wickham, 2016). For the (ii) GTDB-Tk phylogeny, GTDB-Tk was run as described above, with the addition of the outgroup genome (see section "Taxonomic assignment" above). The resulting AA MSA produced by GTDB-Tk was supplied to IQ-TREE, which was used to construct a ML phylogeny as described above, but with the "LG + F + R4" AA substitution model (i.e., the optimal AA substitution model selected using IQ-TREE's implementation of ModelFinder, based on Bayesian Information Criteria [BIC] values; Yang, 1995; Le and Gascuel, 2008; Soubrier et al., 2012; Kalyaanamoorthy et al., 2017). iTOL was used to plot the resulting phylogeny (Supplementary Figure S7).

2.8. Functional enrichment analyses

As mentioned above, two PopCOGenT main clusters (i.e., Main Clusters 0 and 2) contained >1 subcluster (see section "Taxonomic assignment" above; Supplementary Table S8). To gain insight into the functional potential of subcluster-specific genes, which had been acquired post-speciation and differentially swept through subclusters identified via PopCOGenT (i.e., flexible genes identified via PopCOGenT, see section "Taxonomic assignment" above; Supplementary Tables S9, S10), functional enrichment analyses were conducted.

Briefly, for each relevant PopCOGenT main cluster (i.e., Main Cluster 0 and Main Cluster 2; see section "Taxonomic assignment" above), open reading frames (ORFs) produced by PopCOGenT for all members of the given main cluster were supplied as input to the eggNOG-mapper v2.1.9 web server (http://eggnog-mapper.embl.de/; accessed 26 November 2022; Huerta-Cepas et al., 2019; Cantalapiedra et al., 2021). eggNOG-mapper was used to functionally annotate each ORF, using default settings for all parameters except the input data type option (which was set to "CDS," as DNA sequences were used as input) and the "Gene Ontology evidence" option, which was set to "Transfer all annotations (including inferred from electronic annotation)."

For each PopCOGenT subcluster within the given main cluster, enrichment analyses were conducted to identify Gene Ontology (GO) terms (Ashburner et al., 2000; The Gene Ontology Consortium, 2018) assigned via eggNOG-mapper, which were overrepresented among the PopCOGenT flexible genes identified within that particular subcluster: flexible genes identified within the given subcluster were treated as positive instances (PopCOGenT $p < 0.05$; Supplementary Tables S9, S10), and all other genes within the main cluster were treated as negative instances. Only genes with ≥1 assigned GO term were maintained. GO terms enriched within the positive instances (i.e., the subcluster-specific flexible genes identified via PopCOGenT; Supplementary Tables S9, S10) were identified via the "runTest" function in the topGO v2.46.0 R package (Alexa et al., 2006), using a Fisher's exact test (FET) with the "weight01" algorithm. Tests were conducted using each of the Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) ontologies, using a minimum topGO node size of 3 for each ontology (i.e., "nodeSize = 3," where topGO prunes the GO hierarchy from the terms with <3 annotated genes). GO terms were considered to be significantly enriched in the flexible genome of a PopCOGenT subcluster if the resulting FET p-value was <0.05; no additional multiple testing correction was applied, as the "weight01" algorithm accounts for GO graph topology and produces p-values, which can be viewed as inherently corrected or not affected by multiple testing (Alexa et al., 2006). This approach was repeated for each subcluster within PopCOGenT Main Clusters 0 and 2 (Supplementary Tables S20–S24).

2.9. Species-level phylogeny construction

Species-level phylogenies were additionally constructed for the following: (i) GTDB's *M. caseolyticus* genomospecies, as it was composed of multiple PopCOGenT subclusters and contained five of the six South African genomes sequenced in this study ($n = 58$ genomes; Supplementary Tables S4, S8); (ii) bactaxR Cluster 13, corresponding to an unknown GTDB genomospecies, which contained the *M. armenti* type strain, because it, too, was composed of multiple PopCOGenT subclusters ($n = 8$ genomes; Supplementary Tables S6, S8); (iii) bactaxR Cluster 2, as it contained one of the South African genomes sequenced in this study ($n = 4$ genomes; Supplementary Table S6).

For *M. caseolyticus* and bactaxR Cluster 13 (i.e., *M. armenti*), Panaroo was used to construct a core gene alignment as described above (see section "Genus-level phylogeny construction" above), using a protein family sequence identity threshold of 70% ("-f 0.7"), all genomes assigned to the respective species cluster as input, and the following outgroup genomes: (i) a *Macrococcus* spp. genome from bactaxR Cluster 2 for *M. caseolyticus* (NCBI GenBank Assembly accession GCA_019357535.1), and (ii) a *M. canis* genome for bactaxR Cluster 13 (NCBI GenBank Assembly accession GCA_014524485.1; Supplementary Figure S7). Each resulting core gene alignment was supplied as input to IQ-TREE, and ML phylogenies were constructed and annotated as described above (see section "Genus-level phylogeny construction" above).

For *M. caseolyticus*, which was composed of >2 PopCOGenT subclusters, RhierBAPs v1.1.4 (Tonkin-Hill et al., 2018) was additionally employed to cluster the 58 *M. caseolyticus* genomes using two clustering levels. Briefly, Panaroo was used to construct a core gene alignment as described above but with the outgroup genome omitted ($n = 58$ total *M. caseolyticus* genomes). Core SNPs were identified within the resulting core gene alignment using snp-sites v2.5.1 (Page et al., 2016; using the "-c" option), and the resulting core SNP alignment was supplied as input to RhierBAPs.

For bactaxR Cluster 2, all genomes were fairly closely related (>99.2 ANI via OrthoANI); thus, Snippy v4.6.0[4] was used to identify core SNPs among all four genomes within this species cluster, using the closed chromosome of one of the bactaxR Cluster 2 genomes as a reference (NCBI Nucleotide accession NZ_CP079969.1; Li and Durbin, 2009; Li et al., 2009; Quinlan and Hall, 2010; Li, 2011; Cingolani et al., 2012; Garrison and Marth, 2012; Li, 2013; Tan et al., 2015; Page et al., 2016; Li, 2019; Seemann, 2019). For the bactaxR Cluster 2 genome sequenced in this study, trimmed paired-end reads were used as input; for the three publicly available genomes, assembled genomes were used as input. Snippy was run using default settings, and the resulting cleaned alignment was supplied as input to Gubbins v3.1.3 (Croucher et al., 2015) to remove recombination using default settings. The resulting recombination-free alignment produced by Gubbins was queried using snp-sites as described above, and the resulting core SNP alignment was supplied as input to IQ-TREE. IQ-TREE was used to construct a ML phylogeny using an ascertainment bias correction based on the number of constant sites in the Snippy alignment ("-fconst 645,581,381,484,377,636,655,813"), one thousand replicates of the ultrafast bootstrap approximation ("-bb 1,000"), and the optimal nucleotide substitution model selected using ModelFinder ("-m MFP"; the K3Pu model, based on its BIC value; Kimura, 1981; Kalyaanamoorthy et al., 2017). The resulting phylogeny was displayed and annotated using FigTree v1.4.4.[5] The aforementioned steps were repeated, with the genome sequenced in this study omitted, as the remaining three genomes were highly similar on a genomic scale (>99.99 ANI via OrthoANI for the three publicly available bactaxR Cluster 2 genomes; note that Gubbins was not used here, as there were only three genomes available). Pairwise core SNP distances between genomes were calculated in R using the dist.gene function (with "method" set to "pairwise") in ape v5.6.2 (Paradis et al., 2004; Paradis and Schliep, 2019). Snippy was additionally

used to identify SNPs between other closely related genomes identified in the study (i.e., >99.9 ANI via OrthoANI), using default settings.

3. Results

3.1. Multiple GTDB species are represented among bovine-associated South African *Macrococcus* strains

Of the *Macrococcus* strains isolated in South Africa that underwent WGS (i.e., two veterinary isolates from bovine clinical mastitis cases, plus four food isolates from beef products), five were assigned to the *M. caseolyticus* genomospecies using the Genome Taxonomy Database Toolkit (GTDB-Tk; Table 1; Supplementary Table S4). These five genomes each shared 98.0–98.6 average nucleotide identity (ANI) with the closed type strain genome of *M. caseolyticus* (calculated via OrthoANI relative to the *M. caseolyticus* type strain genome with NCBI RefSeq Assembly accession GCF_016028795.1; Supplementary Table S5), which is well above the 95 ANI threshold typically used for prokaryotic species delineation (Jain et al., 2018). When compared to each other, the five *M. caseolyticus* genomes sequenced here shared 97.9–99.4 ANI via OrthoANI. One genome (S135) was assigned to PubMLST *M. caseolyticus* sequence type 2 (ST2), while another (S139) was assigned to ST16; the remaining three *M. caseolyticus* genomes belonged to unknown STs (Supplementary Table S11).

Notably, however, one food isolate (S115) could not be assigned to any known species within GTDB (Table 1; Supplementary Table S4). Strain S115 was isolated in 2015 from beef biltong, a South African spiced intermediate moisture, ready-to-eat (RTE) meat product, which was being sold in a retail outlet in South Africa's Limpopo province (Table 1; Supplementary Table S1). When compared to the five *M. caseolyticus* genomes sequenced here, S115 shared <95 ANI with each (via OrthoANI; Supplementary Table S5). When compared to the type strain genomes of all *Macrococcus* species, S115 was most closely related to *M. caseolyticus* subsp. *hominis* str. CCM 7927 (NCBI RefSeq Assembly accession GCF_002742395.2), sharing 95.3 ANI via OrthoANI (Supplementary Table S5). Comparatively, S115 shared 94.6 ANI with the closed *M. caseolyticus* type strain genome (via OrthoANI, NCBI RefSeq Assembly accession GCF_016028795.1; Supplementary Table S5).

Overall, these results indicate that, among the bovine-associated South African *M. caseolyticus* genomes sequenced here, (i) considerable within-species diversity exists (e.g., multiple STs are represented, novel STs are present, ANI values between strains sequenced in this study are not particularly high); (ii) one or two *Macrococcus* genomospecies are represented, depending on the species delineation method used (i.e., GTDB or ANI-based comparisons to type strain genomes; Table 1).

3.2. Human clinical, veterinary clinical, and food spoilage-associated strains are represented among *Macrococcus* spp. genomes

To compare the bovine-associated South African *Macrococcus* genomes sequenced here to *Macrococcus* genomes collected from other

---

TABLE 1 South African *Macrococcus* spp. genomes sequenced in this study (*n* = 6).

Strain	Year of isolation	Province	Animal	Sample type	Isolation source[a]	Establishment category	GTDB species[b]
S99	1991	Gauteng	Cattle	Veterinary clinical sample	Milk from mastitis case	Farm	*M. caseolyticus*
S125	1992	Gauteng	Cattle	Veterinary clinical sample	Milk from mastitis case	Farm	*M. caseolyticus*
S120	2015	Gauteng	Cattle	Meat sample	RTE beef biltong	Retail outlet	*M. caseolyticus*
S139	2015	Gauteng	Cattle	Meat sample	Minced beef	Butchery	*M. caseolyticus*
S135	2015	Free State	Cattle	Meat sample	Processed beef patties	Retail outlet	*M. caseolyticus*
S115	2015	Limpopo	Cattle	Meat sample	RTE beef biltong	Retail outlet	*M. spp. nov.*

[a]RTE, ready-to-eat.

[b]Species assigned using the Genome Taxonomy Database (GTDB) Toolkit (GTDB-Tk) v2.1.0 and version R207_v2 of GTDB; all six genomes were assigned to GTDB's "Macrococcus_B" genus.

sources in other world regions, the six genomes sequenced here were aggregated with all high-quality, publicly available *Macrococcus* genomes (*n* = 110 total genomes; Figure 1; Supplementary Table S3). Overall, the complete set of 110 *Macrococcus* genomes represented strains collected from at least 10 countries, with most genomes originating from Europe (88 of 110 genomes, 80.0%; Figure 1; Supplementary Tables S1, S3).

A vast majority of the genomes (97 of 110 genomes, 88.2%) originated from animal- and animal product-associated sources, with over half of all strains originating from bovine-associated sources (60 of 110 total genomes, 54.5%; Figure 1; Supplementary Tables S1, S3). Numerous animal-associated strains, including two strains sequenced here, were reportedly clinical in origin (e.g., isolated from bovine mastitis cases, canine ear infection cases, and wound infections in donkeys; Table 1; Supplementary Tables S1, S3). Several animal-associated strains, including four sequenced here, were isolated from food products with the potential for human consumption (i.e., beef and pork meat, cow milk, cheese); one strain was isolated from a food product with a known defect (i.e., "ropy" milk; Table 1; Supplementary Tables S1, S3).

Interestingly, six of the 110 *Macrococcus* genomes (5.5%) were derived from human-associated strains (Figure 1; Supplementary Table S3). At least five of these strains were isolated in conjunction with human clinical cases, including: (i) a hemolytic, methicillin-resistant *M. canis* strain isolated from a 52-year-old immunocompromised patient with cutaneous maculopapular and impetigo lesions (Switzerland, 2019); (ii) a *M. bohemicus* strain from an 80–85 year-old patient with a traumatic knee wound (Czech Republic, 2003); (iii) a *M. goetzii* strain from a foot nail mycosis case in a 30–35 year-old patient (Czech Republic, 2000); (iv) a *M. caseolyticus* subsp. *hominis* strain from an acute vaginitis case in a 40–45 year old patient (Czech Republic, 2003); (v) a *M. epidermidis* strain associated with mycose in a 66–70 year old patient (Czech Republic, 2001; Figure 1; Supplementary Table S3; Maslanova et al., 2018; Jost et al., 2021).

Overall, *Macrococcus* WGS efforts have overwhelming queried animal or animal product-associated strains, although several human clinical strains have undergone WGS (Figure 1).

3.3. *Macrococcus* genomospecies clusters may overlap at a conventional 95 ANI threshold

Overall, using GTDB-Tk, *Macrococcus* encompassed 15 genomospecies: 12 defined genomospecies, plus three undefined/

putative novel genomospecies defined using a conventional 95 ANI threshold (Figure 1; Supplementary Table S4). One of these putative novel GTDB-Tk genomospecies encompassed strain S115 sequenced here (denoted as bactaxR Cluster 2 in Figure 1), plus publicly available genomes submitted to NCBI as *M. caseolyticus* (Supplementary Table S3). All members of this genomospecies shared <95 ANI with the *M. caseolyticus* type strain genome but >95 ANI with the *M. caseolyticus* subsp. *hominis* type strain genome (via OrthoANI, NCBI RefSeq Assembly accessions GCF_016028795.1 and GCF_002742395.2, respectively; Figure 2). The second putative novel GTDB-Tk genomospecies (denoted as bactaxR Cluster 13 in Figure 1) contained the type strain of *M. armenti* (NCBI GenBank Assembly accession GCA_020097135.1); considering *M. armenti* was published as a novel species in 2022, it is likely this genomospecies will be described as such in future versions of GTDB (Keller et al., 2022). The third putative novel GTDB-Tk genomospecies contained a single genome (denoted as bactaxR Cluster 14 in Figure 1), which had been submitted to NCBI as *M. caseolyticus* (strain Ani-LG-066, NCBI GenBank Assembly accession GCA_021366795.1); however, this genome shared <75 ANI with the *M. caseolyticus* and *M. caseolyticus* subsp. *hominis* type strain genomes and shared <81.0 ANI with all other *Macrococcus* spp. genomes (via OrthoANI; Figures 1, 2). Strain Ani-LG-066, which was isolated in 2016 from milk samples taken from a lactating dairy cow in Liege, Belgium, thus likely represents a truly novel *Macrococcus* genomospecies (Figure 1; Supplementary Tables S3, S5).

Importantly, for three of the four genomospecies delineation methods used here (i.e., GTDB-Tk, bactaxR, and PopCOGenT), genomes assigned to separate genomospecies could share >95 ANI with each other (Figure 2; Supplementary Figures S8, S9), indicating that some *Macrococcus* genomospecies defined at a conventional 95 ANI threshold overlap. specI did not yield overlapping genomospecies at 95 ANI (Supplementary Figure S9); however, nearly a third of *Macrococcus* genomes (*n* = 32 of 110 genomes, 29.1%) could not be assigned to a species via specI (Figure 1; Supplementary Figure S9; Supplementary Table S7).

Taken together, these results indicate that (i) three of the four genomospecies delineation methods queried here (i.e., GTDB-Tk, bactaxR with OrthoANI and a 95 ANI threshold, and PopCOGenT) produced similar, albeit not identical, results when applied to *Macrococcus* (Figure 1); (ii) the same three genomospecies delineation methods produced "overlapping genomospecies," in which some genomes could share >95 ANI with members of another genomospecies (Figure 2; Supplementary Figure S9).

FIGURE 1
Maximum likelihood (ML) phylogeny of all 104 high-quality, publicly available *Macrococcus* genomes, plus six bovine-associated South African genomes sequenced here (*n* = 110 total *Macrococcus* genomes). Pink circles denote genomes sequenced in this study ("Study"). Rings surrounding the phylogeny denote the following attributes (from innermost to outermost): **(A)** the continent from which each strain was reportedly isolated ("Continent"); **(B)** the isolation source reported for each strain (because all food-associated isolates were derived from animal products, food-associated isolates were categorized as "Cow" if they were derived from bovine-associated products or "Animal" if they were derived from non-bovine or unspecified animal products); **(C-F)** species assignments obtained using four different taxonomic frameworks: **(C)** Genome Taxonomy Database (GTDB) species, assigned using the Genome Taxonomy Database Toolkit (GTDB-Tk) v2.1.0 and GTDB vR207_v2; (D) PopCOGenT "main clusters" (i.e., gene flow units, which attempt to mirror the classical species definition for animals and plants); **(E)** genomospecies clusters delineated *de novo* using average nucleotide identity (ANI) values calculated via OrthoANI, bactaxR, and a 95 ANI genomospecies threshold (i.e., the threshold largely adopted by the microbiological community); **(F)** marker gene-based species clusters within the specI v3 taxonomy. The ML phylogeny was constructed using an alignment of 649 core genes identified among all 110 *Macrococcus* genomes, plus the genome of *S. aureus* str. DSM 20231 (outgroup genome; NCBI RefSeq Assembly accession GCF_001027105.1), using Panaroo and a 50% protein family sequence identity threshold. The tree was rooted using the outgroup (omitted for readability), and branch lengths are reported in substitutions per site. AF, Africa; AS, Asia; EU, Europe; NA, North America; XX, unknown/unreported geographic location.

3.4. Multiple *Macrococcus* spp. contain genomes that are predicted to be multi-drug resistant

Antimicrobial resistance (AMR) and stress response determinants (detected via AMRFinderPlus; Supplementary Table S12) were variably present throughout *Macrococcus* and were associated with predicted resistance to a variety of antimicrobial classes, heavy metals, and metalloids (Figure 3; Supplementary Figure S10). The most common classes of antimicrobials for which *Macrococcus* was predicted to harbor resistance determinants included macrolides, beta-lactams, and aminoglycosides (*n* = 81, 61, and 44 of 110 genomes with one or more associated AMR determinants, corresponding to 73.6, 55.5, and 40.0% of *Macrococcus* genomes, respectively; Figure 3; Supplementary Table S12). The high proportion of genomes harboring an ATP-binding cassette subfamily F protein (ABC-F)-encoding gene (*abc-f*) contributed to the high proportion of genomes with predicted macrolide resistance (*n* = 74 of 110 *Macrococcus* genomes harbored *abc-f*, 67.3%), although several additional macrolide resistance genes were sporadically present within the genus (Figure 3; Supplementary Figure S10; Supplementary Table S12). The high proportion of genomes showcasing predicted beta-lactam resistance, on the other hand, was largely driven by the presence of *mecD* (*n* = 43 of 110 *Macrococcus* genomes, 39.1%), although *mecB* and *bla* were also present in >10% of genomes (Figure 3; Supplementary Figure S10; Supplementary Table S12). Aminoglycoside resistance genes were sporadically present among *Macrococcus* genomes, the most common being *str* (*n* = 23 of 110 genomes, 20.9%; Figure 3; Supplementary Figure S10; Supplementary Table S12).
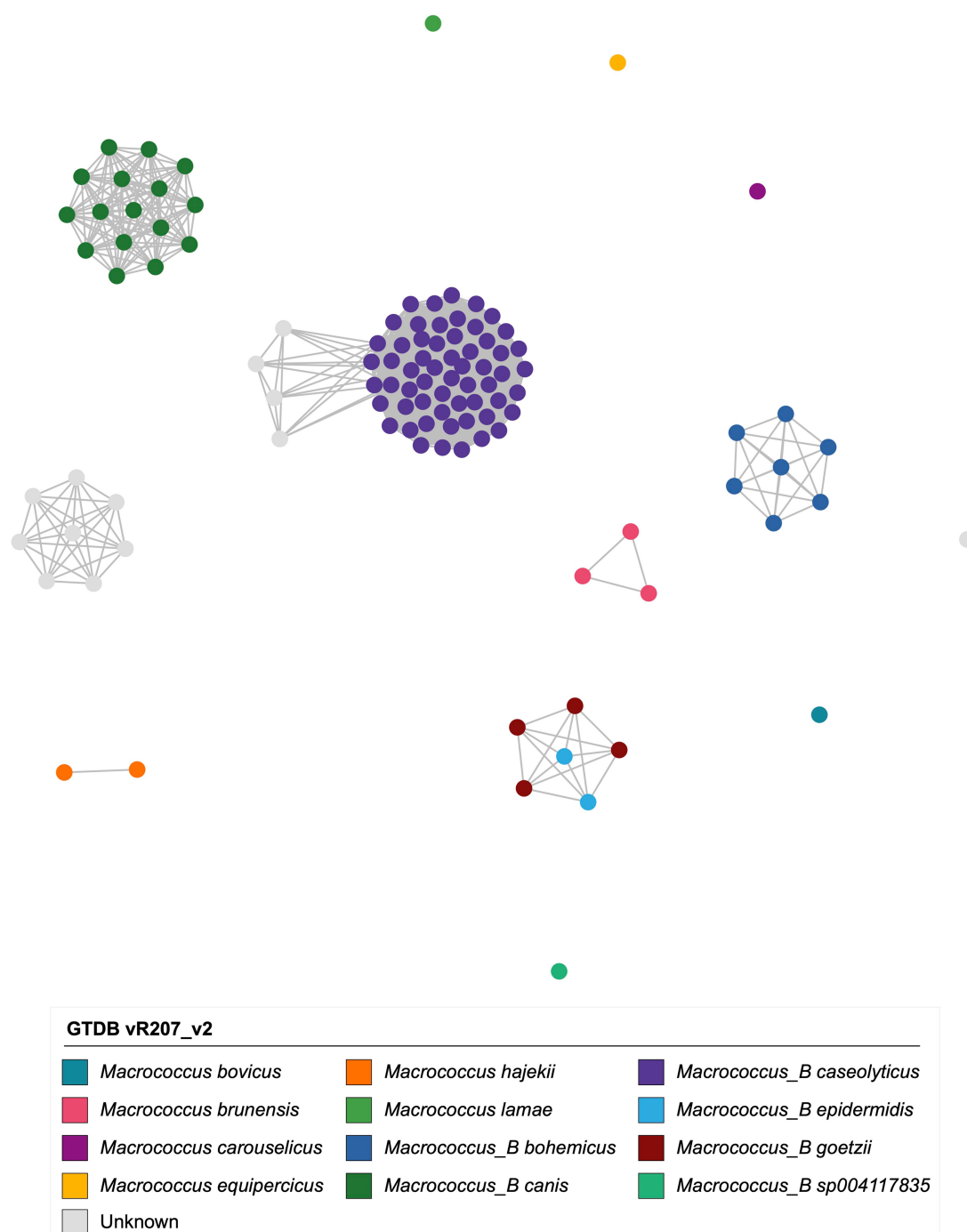
FIGURE 2
Network constructed using pairwise average nucleotide identity (ANI) values calculated between 110 *Macrococcus* genomes. Nodes represent individual genomes, colored by their Genome Taxonomy Database (GTDB) species assigned via the Genome Taxonomy Database Toolkit (GTDB-Tk) v2.1.0 and GTDB vR207_v2. Two nodes (genomes) are connected if they share ≥95 ANI with each other (calculated via OrthoANI). Networks were constructed and displayed using the ANI.graph function in bactaxR (default settings). See Supplementary Figure S9 for an extended version of this figure, which shows results obtained using all four species delineation methods (i.e., GTDB-Tk, bactaxR with a 95 ANI threshold, PopCOGenT, and specI).

The most common AMR profiles among *Macrococcus* genomes harboring one or more AMR determinant were those associated with (i) macrolide and (ii) beta-lactam/macrolide resistance ($n = 18$ and 13 of 110 genomes, corresponding to 16.4 and 11.8% of genomes, respectively; Figure 3; Supplementary Figure S10; Supplementary Table S12). However, numerous predicted multidrug-resistance (MDR) profiles were observed, the most common being (i) aminoglycoside/beta-lactam/macrolide and (ii) aminoglycoside/beta-lactam/macrolide/tetracycline resistance ($n = 11$ and 8 of 110 genomes, corresponding to 10.0 and 7.3% of
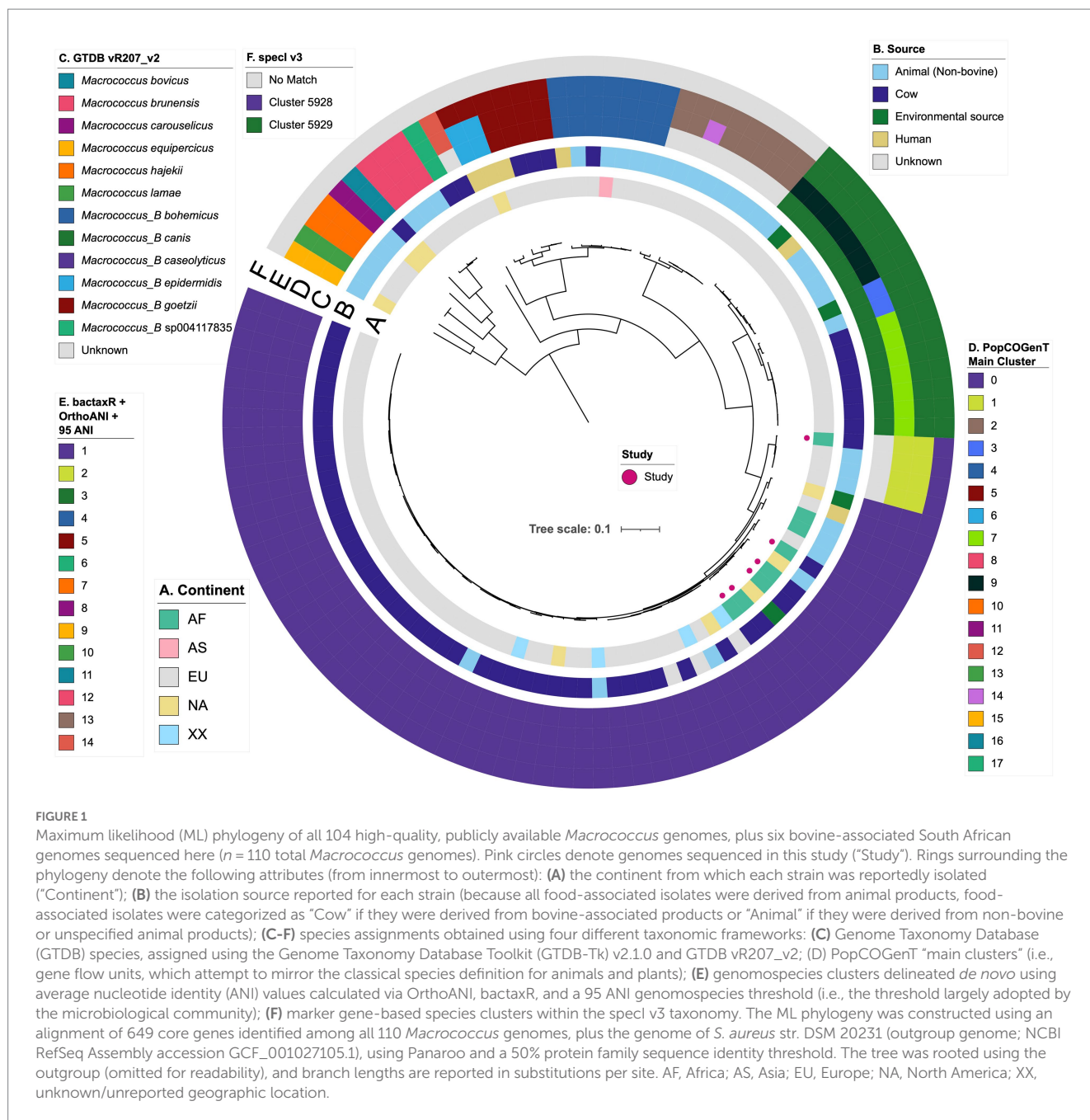
FIGURE 3

Maximum likelihood (ML) phylogeny of all 104 high-quality, publicly available *Macrococcus* genomes, plus six bovine-associated South African genomes sequenced here (*n* = 110 total *Macrococcus* genomes). Color strips and heatmaps to the right of the phylogeny denote (from left to right): (i) genomes of strains isolated and sequenced in this study (denoted by pink circles; "Study"); (ii) genomespecies assignments obtained via the Genome Taxonomy Database Toolkit (GTDB-Tk) v2.1.0 and GTDB vR207_v2 ("GTDB"); (iii) the source from which each strain was reportedly isolated ("Source"); because all food-associated isolates were derived from animal products, food-associated isolates were categorized as "Cow" if they were derived from bovine-associated products or "Animal" if they were derived from non-bovine or unspecified animal products); (iv) the continent from which each strain was reportedly isolated ("Continent"); (v) presence of gene(s) sharing homology to aureolysin at 60% amino acid identity and 50% coverage ("Aureolysin"); (vi) antimicrobial resistance (AMR) and stress response determinants identified in each genome using AMRFinderPlus (default settings; "AMRFinderPlus Determinant"). The ML phylogeny was constructed using an alignment of 649 core genes identified among all 110 *Macrococcus* genomes, plus the genome of *S. aureus* str. DSM 20231 (outgroup genome; NCBI RefSeq Assembly accession GCF_001027105.1), using Panaroo and a 50% protein family sequence identity threshold. The tree was rooted using the outgroup (omitted for readability), and branch lengths are reported in substitutions per site. AF, Africa; AS, Asia; EU, Europe; NA, North America; XX, unknown/unreported geographic location. For an extended version of this figure, see Supplementary Figure S10.

genomes, respectively; Figure 3; Supplementary Figure S10; Supplementary Table S12). The genome displaying predicted AMR to the most antimicrobial classes was the genome of *M. caseolyticus* strain 5813_BC74, which had reportedly been isolated from bovine bulk tank milk in the United Kingdom in 2016 (NCBI GenBank Assembly accession GCA_002834615.1; Supplementary Table S3). This genome displayed predicted aminoglycoside/beta-lactam/ fusidic acid/lincosamide/macrolide/tetracycline resistance (*n* = 6 antimicrobial classes; Figure 3; Supplementary Figure S10; Supplementary Table S12).

Predicted AMR phenotypes observed in <10% of all *Macrococcus* genomes included: (i) fusidic acid resistance (due to the presence of *fusC*; *n* = 10), (ii) lincosamide resistance (per *lnu(A)*, *lnu(G)*; *n* = 7),

(iii) streptothricin resistance (via *sat4*; *n* = 4), (iv) bleomycin resistance (via *bleO*; *n* = 3), (v) trimethoprim (via *dfrE*, *dfrK*) and (vi) fosfomycin resistance (via *fosY*, *n* = 2 genomes each), and (vii) phenicol resistance (via *fexB*, *n* = 1 genome; Figure 3; Supplementary Figure S10; Supplementary Table S12). Interestingly, one genome (South African strain S99, from a bovine mastitis case in Gauteng in 1991) harbored bacitracin resistance genes (i.e., *bcrB* and *bcrC*; Figure 3; Supplementary Figure S10; Supplementary Table S12).

Overall, these results indicate that (i) numerous AMR determinants are variably present within and among *Macrococcus* species; and (ii) *Macrococcus* genomes may harbor AMR determinants predictive of an MDR phenotype (i.e., resistant to three or more antimicrobial classes; Figure 3; Supplementary Figure S10; Supplementary Table S12). However,

these results should be interpreted with caution, as AMR potential was not evaluated phenotypically in this study.

3.5. *Staphylococcus aureus* virulence factor homologues can be detected within some *Macrococcus* genomes at low amino acid identity

To gain insight into the virulence potential of *Macrococcus,* the 110 genomes aggregated here were queried for virulence factors present in the VFDB core database (Figure 3; Supplementary Figure S11; Supplementary Tables S13–S18). Notably, genes encoding the *S. aureus* exoenzyme aureolysin could be detected across multiple *Macrococcus* species (using 90% coverage, $n = 40$ and 77 genomes harboring aureolysin-encoding genes at 60 and 40% amino acid [AA] identity, respectively; Figure 3; Supplementary Figure S11; Supplementary Tables S16, S18).

Further, genes sharing ≥40% AA identity and >85% coverage with *S. aureus* Panton-Valentine leucocidin (PVL) toxin-associated *lukF-PV* and *lukS-PV* were identified in eight *M. canis* genomes (per GTDB-Tk; Supplementary Figure S11; Supplementary Table S15). AA identities of the *Macrococcus lukF-PV* relative to *S. aureus lukF-PV* (VFDB ID VFG001276) ranged from 50.5 to 53.9%, with a mean of 52.6%; for *lukS-PV* (VFDB ID VFG001277), AA identities of 49.5–49.8% were observed, with a mean of 49.6%. For all eight *M. canis* genomes in which they were detected, the *lukF-PV* and *lukS-PV* homologs were located next to each other in the genome (Supplementary Figure S11; Supplementary Table S15).

Overall, these results indicate that proteins homologous to virulence factors present in other species (e.g., *S. aureus*) can be detected in some *Macrococcus* genomes. However, the methods employed here are not adequate to properly evaluate the virulence potential of *Macrococcus* strains that possess these homologs; thus, these results should be interpreted with extreme caution.

3.6. *Macrococcus* species differ in pan-genome composition

Using PEPPAN and a 40% AA identity threshold, a total of 10,300 genes were detected among the 110 *Macrococcus* genomes aggregated here, 1,229 of which were core genes present in all 110 genomes (11.9% of all *Macrococcus* genes; Figure 4; Supplementary Figures S1, S3, S5). Comparatively, at a 20% AA identity threshold, 9,835 total genes were detected, 1,235 of which were core genes present in all 110 genomes (12.6% of all *Macrococcus* genes; Supplementary Figures S2–S4, S6). Based on trees constructed using pan-genome element presence/absence, *Macrococcus* species (per GTDB-Tk) tended to cluster together based on pan-genome composition, although not exclusively (Supplementary Figures S1, S2). Specifically, the topology of the PEPPAN pan-genome tree differed from that of the PEPPAN Core Genome Allelic Variation (CGAV) tree, as some *Macrococcus* species were polyphyletic based on pan-genome element presence/absence (Supplementary Figures S1, S2). Overall, *Macrococcus* species tend to differ via both core genome phylogeny (Figures 1, 3) and pan-genome composition (Figure 4; Supplementary Figures S1–S6).

3.7. *Macrococcus caseolyticus* and *Macrococcus armenti* are composed of multiple within-species subclusters separated by recent gene flow

PopCOGenT identified 18 "main clusters" (species) across *Macrococcus* in its entirety; within two of these main clusters (i.e., PopCOGenT Main Clusters 0 and 2 in Figure 1), PopCOGenT identified multiple "subclusters" separated by recent gene flow (i.e., populations that were still connected by some gene flow, but had significantly more gene flow within the population than between populations; Figure 1). Specifically, (i) within PopCOGenT Main Cluster 0 (corresponding to GTDB-Tk's *M. caseolyticus* genomospecies), five subclusters were identified, and (ii) within PopCOGenT Main Cluster 2 (an unknown species via GTDB-Tk, which contains the *M. armenti* type strain and will thus be referred to as *M. armenti* hereafter), two subclusters were identified. As such, we will discuss these two species individually in detail below (Figure 1).

3.7.1. African and European *Macrococcus caseolyticus* strains largely belong to separate lineages

The 58 *M. caseolyticus* genomes (per GTDB-Tk) were divided into five PopCOGenT subclusters and five RhierBAPS clusters, although the composition of those (sub)clusters differed slightly (Figure 5; Supplementary Figure S12; Supplementary Table S8). Notably, the majority of European *M. caseolyticus* genomes ($n = 33$ of 42 European *M. caseolyticus* genomes, 78.6%) were assigned to a well-supported clade within the species phylogeny (referred to hereafter as the "*M. caseolyticus* major European lineage," which is denoted in Figure 5 as RhierBAPS Cluster 1; ultrafast bootstrap support = 100%). Members of the *M. caseolyticus* major European lineage were overwhelmingly of bovine origin (34 of 36 RhierBAPS Cluster 1 genomes, 94.4%), and nearly all genomes within the lineage were reportedly isolated from European countries: 30 from the United Kingdom (83.3%), and two and one genome(s) from Switzerland and Ireland, respectively (5.6 and 2.8%); the only genome isolated from outside of Europe was reportedly isolated from ropy milk in the United States in 1920 (NCBI GenBank Assembly accession GCA_900453015.1; Figure 5; Supplementary Figure S12; Supplementary Table S3). Interestingly, most genomes within the *M. caseolyticus* major European lineage were predicted to be MDR (Figures 3, 5; Supplementary Figure S12). Specifically, (i) all genomes in the *M. caseolyticus* major European lineage (36 of 36 genomes, 100%) were predicted to be resistant to macrolides, largely due to the presence of *abc-f* (35 of 36 *M. caseolyticus* major European lineage genomes, 97.2%; the only genome in which *abc-f* was not detected possessed *erm(B)* and was thus still predicted to be macrolide-resistant via AMRFinderPlus); (ii) nearly all (33 of 36 genomes, 91.7%) were predicted to be resistant to beta-lactams, largely due to the presence of *mecD* in 28 genomes (77.8% of 36 genomes in the lineage; the remaining five genomes that were predicted to be beta-lactam-resistant harbored *bla* and *mecB*); (iii) a majority (21 of 36 genomes in the lineage, 58.3%) were predicted to be resistant to aminoglycosides, due largely to the presence of *str* and/or *aadD1* (detected in 14 and 9 of 36 genomes, 38.9 and 25.0%, respectively; Figures 3, 5; Supplementary Figure S12). Additionally, three genomes possessed genes conferring resistance to bleomycin; these were the
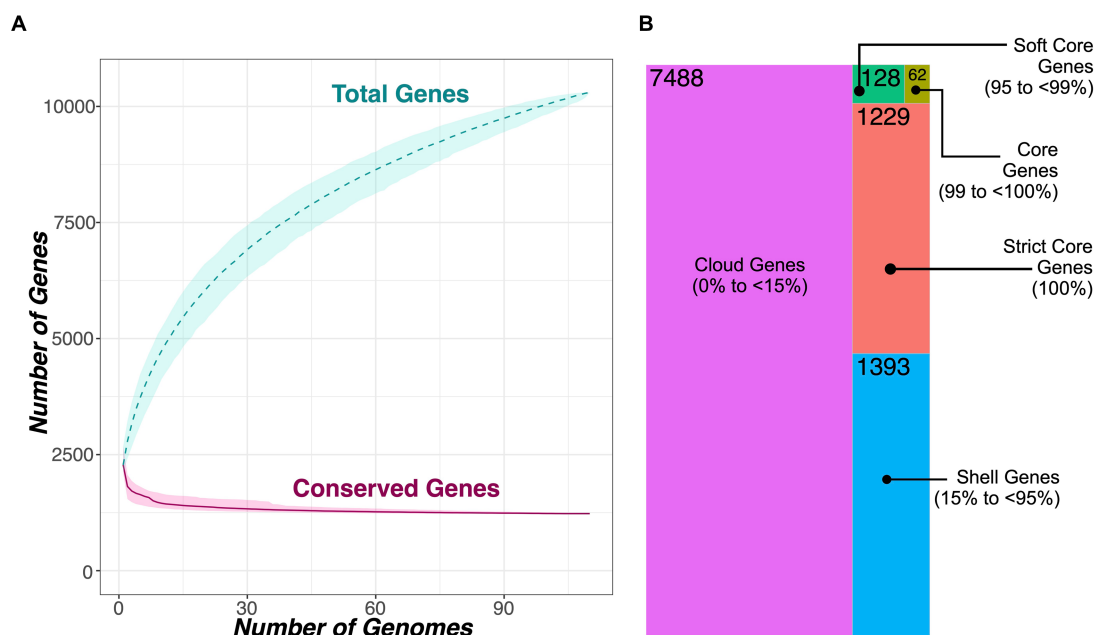
FIGURE 4
(A) Rarefaction curves for the *Macrococcus* pan- and core-genome, constructed using all 104 high-quality, publicly available *Macrococcus* genomes, plus six bovine-associated South African genomes sequenced here (*n* = 110 total *Macrococcus* genomes). Curves showcase the accumulation of pan genes ("Total Genes") and core genes ("Conserved Genes") using 1,000 random permutations. Dashed and solid curved lines denote median values for pan and core genes, respectively, and shading surrounding each line denotes the respective 95% confidence interval. (B) Treemap showcasing the number of genes detected within a given percentage of *Macrococcus* genomes (out of 110 total genomes). Tile sizes are proportional to the number of genes detected within a given percentage of *Macrococcus* genomes; numerical labels within each tile denote the corresponding number of genes. The treemapify v2.5.5 (https://CRAN.R-project.org/package=treemapify) R package was used to construct the plot. For both (A,B), PEPPAN was used to construct the core- and pan-genomes using a 40% amino acid identity threshold and a core genome threshold of 95%.

only genomes within the *Macrococcus* genus, which harbored bleomycin resistance-conferring gene *bleO* (Figures 3, 5; Supplementary Figure S12).

Of the nine European *M. caseolyticus* genomes that were not members of the *M. caseolyticus* major European lineage, seven belonged to a well-supported clade containing 10 genomes (ultrafast bootstrap support = 100%; referred to hereafter as the "*M. caseolyticus* minor European lineage," which is denoted in Figure 5 as RhierBAPS Cluster 2 and PopCOGenT Subcluster 0.1). Aside from two genomes of unknown origin, one genome within the *M. caseolyticus* minor European lineage was reportedly of non-European origin (i.e., strain CCM 3540, reportedly isolated from cow's milk in the Washington, D.C. vicinity of the United States in 1916; NCBI GenBank Assembly accession GCA_003259685.1, Supplementary Table S3; Evans, 1916). Like the *M. caseolyticus* major European lineage, all genomes within the *M. caseolyticus* minor European lineage were predicted to be resistant to macrolides, as all harbored *abc-f* (Figures 3, 5; Supplementary Figure S12). However, a predicted MDR phenotype (i.e., resistant to three or more antimicrobial classes) was less prevalent among genomes within the minor European lineage (*n* = 3 of 10 *M. caseolyticus* minor European lineage genomes, 30%): the MDR genomes were similar on a genomic scale (99.7–99.9 ANI via OrthoANI) and were confined to a single, well-supported clade within the *M. caseolyticus* minor European lineage (ultrafast bootstrap support = 100%; Figure 5; Supplementary Figure S12). Additionally, within the *M. caseolyticus* minor European lineage, PopCOGenT identified six "flexible" genes (i.e., PopCOGenT subcluster-specific

orthologous gene clusters), which were specific to the *M. caseolyticus* minor European lineage (denoted as gene group "C" within the PopCOGenT Flexible Gene heatmap in Figure 5; PopCOGenT *p* < 0.05). All six genes were chromosomal and included (i) large conductance mechanosensitive channel protein MscL, and (ii) genes associated with Y-family DNA polymerases (Figure 5; Supplementary Figure S12; Supplementary Table S9). Compared to all other *M. caseolyticus* genes, numerous biological processes (BPs) and molecular functions (MFs) were enriched in the *M. caseolyticus* minor European lineage flexible genes, including DNA-related BPs/MFs (e.g., DNA biosynthesis, replication, and repair), and those related to ion binding/transport (topGO Fisher's Exact Test [FET] *p* < 0.05; Supplementary Tables S9, S20).

Of the 12 *M. caseolyticus* genomes that were not members of the major and minor European lineages, seven were African in origin, three were North American, and two were European, including the one human-associated *M. caseolyticus* genome (i.e., strain CCM 7927, which was isolated in Pribram, Czech Republic in 2003 from a vaginal swab taken from an acute vaginitis case in a 40–45 year-old patient, NCBI GenBank Assembly accession GCA 002742395.2; Figure 5; Supplementary Figure S12; Supplementary Table S3; Maslanova et al., 2018). Notably, of the five South African *M. caseolyticus* strains isolated and sequenced here, four were assigned to a single PopCOGenT subcluster (i.e., PopCOGenT Subcluster 0.3 in Figure 5). Unlike the major and minor European lineages, members of this subcluster did not possess macrolide resistance genes (Figures 3, 5; Supplementary Figure S12). AMR genes were detected sporadically
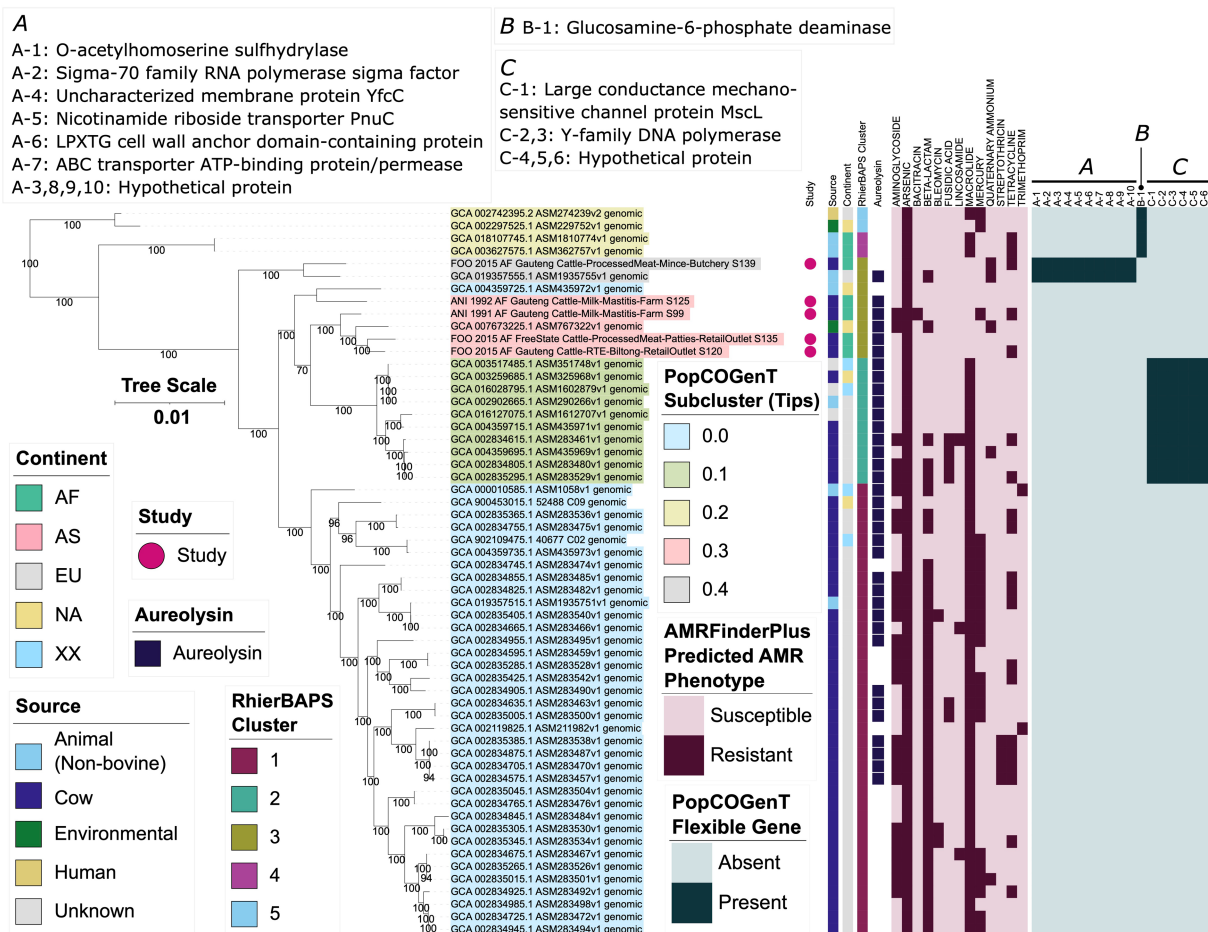
FIGURE 5
Maximum likelihood (ML) phylogeny of 58 genomes assigned to the Genome Taxonomy Database's (GTDB) *M. caseolyticus* genomospecies. Tip label colors correspond to subcluster assignments obtained using PopCOGenT ("PopCOGenT Subcluster"). Pink circles denote genomes sequenced in this study ("Study"). Color strips/heatmaps to the right of the phylogeny denote (from left to right): (i) the source from which each strain was reportedly isolated ("Source"; because all food-associated isolates were derived from animal products, food-associated isolates were categorized as "Cow" if they were derived from bovine-associated products or "Animal" if they were derived from non-bovine or unspecified animal products); (ii) the continent from which each strain was reportedly isolated ("Continent"); (iii) cluster assigned using RhierBAPS ("RhierBAPS Cluster"); (iv) presence of gene(s) sharing homology to aureolysin at 40% amino acid identity and 50% coverage ("Aureolysin"); (v) predicted antimicrobial resistance (AMR) and stress response phenotype, obtained using AMR and stress response determinants identified via AMRFinderPlus ("AMRFinderPlus Predicted AMR Phenotype"); (vi) presence and absence of flexible genes identified via PopCOGenT ("PopCOGenT Flexible Gene"), with corresponding gene annotations displayed in the boxes marked "A," "B," and "C." The ML phylogeny was constructed using an alignment of 1,751 core genes identified among all 58 *M. caseolyticus* genomes, plus an outgroup *Macrococcus* spp. genome from bactaxR Cluster 2 (NCBI GenBank Assembly accession GCA_019357535.1; Figure 1), using Panaroo and a 70% protein family sequence identity threshold. The tree was rooted using the outgroup (omitted for readability), and branch lengths are reported in substitutions per site. Branch labels correspond to branch support percentages obtained using one thousand replicates of the ultrafast bootstrap approximation. AF, Africa; AS, Asia; EU, Europe; NA, North America; XX, unknown/unreported geographic location. For an extended version of this phylogeny, see Supplementary Figure S12.

within these genomes. Specifically, (i) strain S99 possessed genes associated with aminoglycoside (streptomycin), bacitracin, and tetracycline resistance (*str*, *bcrBC*, and *tet(L)*, respectively); (ii) GCA_007673225.1 (an environmental strain isolated in 2018 in Durham, North Carolina, United States) possessed genes associated with aminoglycoside and beta-lactam resistance (i.e., *aph(2″)-IIIa*, *str*, and *mecD*, associated with amikacin/gentamicin/kanamycin/tobramycin, streptomycin, and methicillin resistance, respectively); (iii) S120 possessed tetracycline resistance gene *tet(L)* (Figure 5; Supplementary Figure S12). Despite most genomes being South African in origin, the five *M. caseolyticus* genomes within this subcluster were considerably diverse, sharing 98.6–99.4 ANI with each other (via OrthoANI; Figure 5; Supplementary Figure S12).

The remaining South African genome sequenced in this study (i.e., S139), plus GCA_019357555.1 (isolated from a calf nasal swab in Switzerland in 2019), were assigned to a separate subcluster via PopCOGenT (i.e., PopCOGenT Subcluster 0.4 in Figure 5). Neither genome possessed macrolide resistance genes, although both possessed quaternary ammonium resistance gene *qacH* (Figure 5; Supplementary Figure S12). S139 additionally possessed tetracycline resistance gene *tet(L)*, while GCA_019357555.1 possessed beta-lactam resistance genes *mecB* (methicillin) and *bla* (Figure 5; Supplementary Figure S12). Most notably, however, PopCOGenT identified 10 flexible genes within this subcluster (denoted as gene group "A" within the PopCOGenT Flexible Gene heatmap in Figure 5, PopCOGenT $p < 0.05$; Figure 5; Supplementary Figure S12;

Supplementary Table S9); ATP- and transmembrane-associated BPs/MFs were enriched in this subcluster's flexible genes (topGO FET $p < 0.05$; Supplementary Table S21).

Four additional *M. caseolyticus* genomes were assigned to a single subcluster using PopCOGenT (i.e., PopCOGenT Subcluster 0.2 in Figure 5). Interestingly, like the major and minor European clades, three of the four genomes within this subcluster were predicted to be macrolide resistant, as they possessed *abc-f* and *mef(D)* (Figure 5; Supplementary Figure S12). Two highly similar genomes derived from strains isolated in 2016 from wounded animals in Sudan additionally possessed tetracycline resistance gene *tet(L)* (100.0 ANI and 0 SNPs via OrthoANI and Snippy, respectively, NCBI GenBank Assembly accessions GCA_018107745.1 and GCA_003627575.1; Figure 5; Supplementary Figure S12). Additionally, unlike the other *M. caseolyticus* subclusters described above, none of the genomes within this PopCOGenT subcluster possessed genes sharing homology to aureolysin-encoding genes (Figure 5; Supplementary Figure S12). Further, PopCOGenT identified one flexible gene within this subcluster (denoted as gene group "B" within the PopCOGenT Flexible Gene heatmap in Figure 5, PopCOGenT $p < 0.05$; Supplementary Table S9): glucosamine-6-phosphate deaminase, which was associated with the enrichment of several GO terms, including antibiotic catabolic process, carbohydrate metabolic process, and N-acetylglucosamine-associated processes (topGO FET $p < 0.05$; Figure 5; Supplementary Figure S12; Supplementary Tables S9, S22).

Overall, these results indicate that *M. caseolyticus* genomes from geographic regions outside of Europe, particularly Africa, belong to separate lineages within the species. However, future genomic sequencing efforts are needed to provide further evidence of lineage-geography associations.

3.7.2. Putative virulence factors are differentially associated with *Macrococcus armenti* lineages

Like *M. caseolyticus*, *M. armenti* could be differentiated into subclusters via PopCOGenT (Figure 6A; Supplementary Figure S13; Supplementary Table S8). Specifically, (i) PopCOGenT Subcluster 2.0 contained five genomes from animals in Switzerland (two from strains isolated from the nasal cavities of calves in 2019, and three from the skins of pigs in 2021); and (ii) PopCOGenT Subcluster 2.1 contained two genomes from pigs in Switzerland (one from the nasal cavity of a pig in 2017, and another from the skin of a pig in 2021; Figure 6A; Supplementary Figure S13; Supplementary Tables S3, S8). An additional genome, derived from a pig-associated strain isolated in the United Kingdom in 1963 (NCBI GenBank Assembly accession GCA_022808015.1) was additionally assigned to the *M. armenti* species via ANI-based methods (i.e., using OrthoANI, it shared 96.5–97.7 ANI with all other *M. armenti* genomes; Figure 2; Supplementary Table S5); however, PopCOGenT assigned this genome to a separate main cluster (i.e., "species"), and it was thus not included in the subsequent within-main cluster flexible gene analyses (Figure 6A; Supplementary Figure S13; Supplementary Table S8).

Within PopCOGenT Subcluster 2.0, PopCOGenT identified 43 flexible genes (denoted as gene group "A" within the PopCOGenT Flexible Gene heatmap in Figure 6A; PopCOGenT $p < 0.05$; Supplementary Table S10), which together were associated with the enrichment of eight GO terms (topGO FET $p < 0.05$; Supplementary Table S23). The most highly enriched GO terms were by far "diaminopimelate biosynthetic process" (GO:0019877) and

"lysine biosynthetic process via diaminopimelate" (GO:0009089, topGO FET $p < 1.0 \times 10^{-30}$; Supplementary Table S23), which were assigned to a cluster of three consecutive flexible genes (PopCOGenT $p < 0.05$): (i) 4-hydroxy-tetrahydrodipicolinate reductase/dihydrodipicolinate reductase *dapB* (NCBI Protein accession UBH07557.1); (ii) 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase *dapD* (NCBI Protein accession UBH09720.1); (iii) an amidohydrolase (NCBI Protein accession UBH07558.1; Supplementary Table S10).

Most notably, genes sharing homology to *S. aureus* Type VII secretion system proteins were among the flexible genes within Subcluster 2.0 (PopCOGenT $p < 0.05$), including genes sharing homology to extracellular protein EsxD (VFDB ID VFG049714), chaperone protein EsaE (VFDB ID VFG049701), secreted protein EsxB (VFDB ID VFG002411), secretion substrate EsaC (at 97% query coverage and 38% AA identity; NCBI Protein accession HCD1544785.1), EssB (NCBI Protein accession UBH08107.1), EsaA (NCBI Protein accession UBH08110.1), and secreted protein EsxA (VFDB ID VFG002405; Figure 6A; Supplementary Figure S13; Supplementary Table S10).

Several additional clusters of genes were among the flexible genes within Subcluster 2.0 (PopCOGenT $p < 0.05$; Figure 6A; Supplementary Figure S13; Supplementary Table S10), including: (i) a cluster of genes involved in nitrous oxide reduction, e.g., c-type cytochrome (NCBI Protein accession UBH08788.1), a Sec-dependent nitrous-oxide reductase (NCBI Protein accession UBH08789.1), nitrous oxide reductase family maturation protein NosD (NCBI Protein accession UBH08791.1); (ii) a cluster of genes that included an ImmA/IrrE family metallo-endopeptidase (NCBI Protein accession UBH09010.1), a LacI family DNA-binding transcriptional regulator (NCBI Protein accession UBH09033.1), a sucrose-6-phosphate hydrolase (NCBI Protein accession UBH09034.1), a carbohydrate kinase (NCBI Protein accession UBH09035.1), and sucrose-specific PTS transporter subunit IIBC (NCBI Protein accession UBH09036.1); (iii) a cluster of genes that included a pathogenicity island protein (NCBI Protein accession UBH09209.1; Figure 6A; Supplementary Figure S13; Supplementary Table S10).

Interestingly, a protein most closely resembling immune inhibitor A was also identified by PopCOGenT as a flexible gene (at 98% query coverage and 97.65% AA identity, NCBI Protein accession WP_224185801.1, PopCOGenT $p < 0.05$; Figure 6A; Supplementary Figure S13; Supplementary Table S10). The "immune inhibitor A peptidase M6" protein domain identified in this protein (PFAM ID 05547) has previously been identified in virulence factors secreted by members of the *Bacillus cereus* group (immune inhibitor A; InhA) and *Vibrio cholerae* (secreted metalloprotease PrtV; Vaitkevicius et al., 2008).

Comparatively, within Subcluster 2.1, PopCOGenT identified 45 flexible genes (denoted as gene group "B" within the PopCOGenT Flexible Gene heatmap in Figure 6A; PopCOGenT $p < 0.05$) associated with 22 enriched GO terms (topGO FET $p < 0.05$; Figure 6A; Supplementary Figure S13; Supplementary Tables S10, S24). By far the most highly enriched GO term within this subcluster corresponded to BP "lipoteichoic acid biosynthetic process" (GO:0070395, topGO FET $p = 2.2 \times 10^{-18}$; Supplementary Table S24). Notably, a cluster of five consecutive, chromosomally encoded flexible genes within PopCOGenT Subcluster 2.1 were associated with (lipo teichoic acid synthesis (PopCOGenT $p < 0.05$;
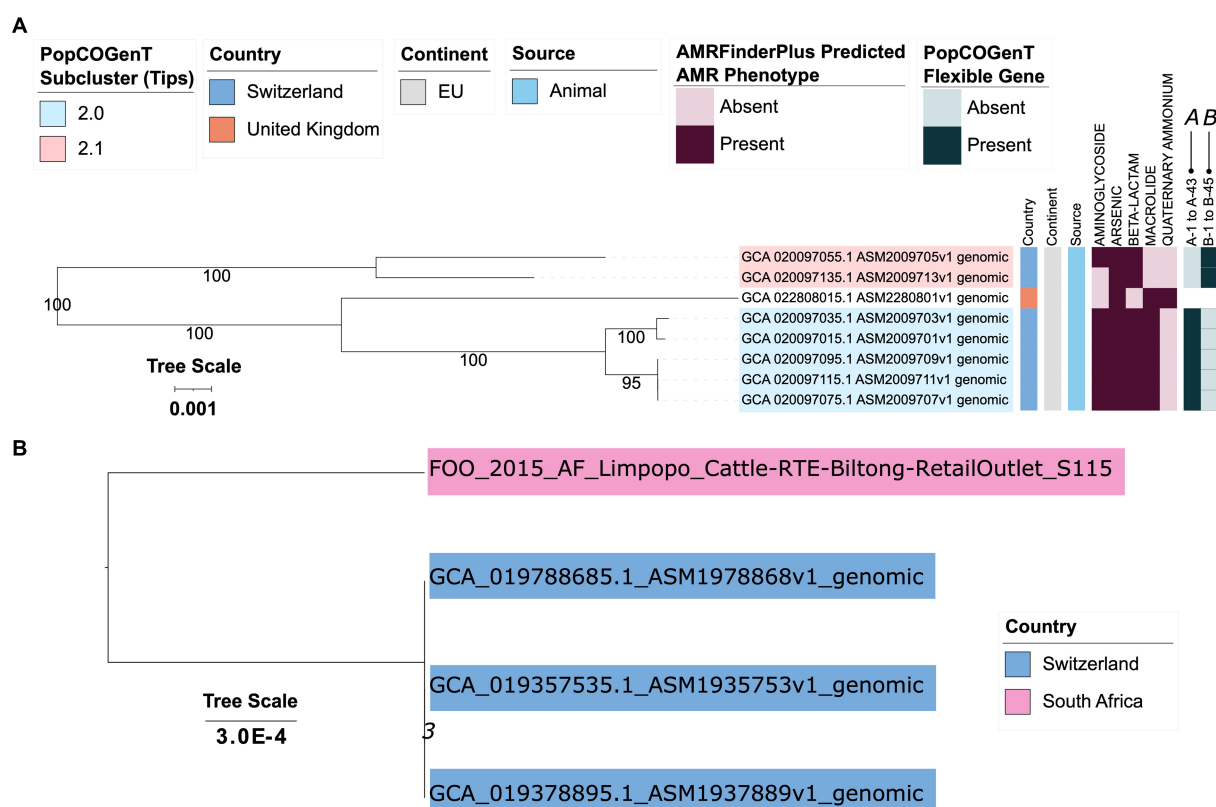
FIGURE 6
(A) Maximum likelihood (ML) phylogeny of eight genomes assigned to bactaxR Cluster 13 (i.e., *M. armenti*, based on average nucleotide identity [ANI]-based comparisons to species type strain genomes; Figure 1). Tip label colors correspond to subcluster assignments obtained using PopCOGenT ("PopCOGenT Subcluster"; one genome was not assigned to the same main cluster via PopCOGenT, and thus is not colored). Color strips/heatmaps to the right of the phylogeny denote (from left to right): (i) the country from which each strain was reportedly isolated ("Country"); (ii) the continent from which each strain was reportedly isolated ("Continent"); (iii) the source from which each strain was reportedly isolated ("Source"); (iv) predicted antimicrobial resistance (AMR) and stress response phenotype, obtained using AMR and stress response determinants identified via AMRFinderPlus ("AMRFinderPlus Predicted AMR Phenotype"); (v) presence and absence of flexible genes identified via PopCOGenT ("PopCOGenT Flexible Gene"; for gene descriptions, see Supplementary Table S10). The ML phylogeny was constructed using an alignment of 1,416 core genes identified among all eight *M. armenti* genomes, plus an outgroup *M. canis* genome (NCBI GenBank Assembly accession GCA_014524485.1; Figure 1), using Panaroo and a 70% protein family sequence identity threshold. The tree was rooted using the outgroup (omitted for readability), and branch lengths are reported in substitutions per site. Branch labels correspond to branch support percentages obtained using one thousand replicates of the ultrafast bootstrap approximation. EU, Europe. For an extended version of this phylogeny, see Supplementary Figure S13. **(B)** ML phylogeny constructed using core SNPs identified among four genomes assigned to bactaxR Cluster 2, a putative novel GTDB genomospecies, which shares >95 ANI with several *M. caseolyticus* genomes but <95 ANI with others (Figures 1, 2). Tip label colors correspond to reported country of isolation. Core SNPs were identified using Snippy, filtered using Gubbins/snp-sites, and the phylogeny was constructed using IQ-TREE. The phylogeny is rooted at the midpoint, and branch lengths are reported in substitutions per site. Branch labels correspond to branch support percentages obtained using one thousand replicates of the ultrafast bootstrap approximation.

Supplementary Table S10): teichoic acid D-Ala incorporation-associated protein DltX (NCBI Protein accession UBH13741.1), D-alanine--poly(phosphoribitol) ligase subunit DltA (NCBI Protein accession UBH13742.1), D-alanyl-lipoteichoic acid biosynthesis protein DltB (NCBI Protein accession UBH13743.1), D-alanine--poly(phosphoribitol) ligase subunit 2 DltC (NCBI Protein accession UBH13744.1), and D-alanyl-lipoteichoic acid biosynthesis protein DltD (NCBI Protein accession UBH13745.1). Interestingly, this cluster of five genes was located several genes downstream of two consecutive, chromosomally encoded beta-lactamase family proteins, which were also both identified as being flexible genes (PopCOGenT $p < 0.05$). Both beta-lactamase family proteins were annotated via eggNOG-mapper as "autolysis and methicillin resistant-related protein PbpX" (NCBI Protein accessions UBH13736.1 and UBH13737.1) and were associated with "response to antibiotic"

(GO:0046677), a BP that was also enriched in PopCOGenT Subcluster 2.1 (topGO FET $p = 2.3 \times 10^{-3}$; Supplementary Tables S10, S24).

Several GO terms associated with transporter activity were also enriched in PopCOGenT Subcluster 2.1 (topGO FET $p < 0.05$), including MF "ABC−type transporter activity" (GO:0140359; Supplementary Table S24). Congruently, four separate clusters of genes containing regions annotated as ABC transporter components were included among PopCOGenT's set of flexible genes (PopCOGenT $p < 0.05$; Supplementary Tables S10, S24).

Interestingly, a protein annotated as immune inhibitor A was also among the flexible genes detected within PopCOGenT Subcluster 2.1 (PopCOGenT $p < 0.05$, NCBI Protein accession UBH13622.1; Figure 6A; Supplementary Figure S13; Supplementary Table S10). Further, genes encoding a type II toxin-antitoxin system were among the flexible genes identified by PopCOGenT within this PopCOGenT

subcluster (PopCOGenT $p < 0.05$; Figure 6A; Supplementary Figure S13; Supplementary Table S10), specifically: (i) a type II toxin-antitoxin system RelE/ParE family toxin, which was immediately upstream of (ii) a type II toxin-antitoxin system Phd/YefM family antitoxin (NCBI Protein accessions UBH12746.1 and UBH12747.1, respectively).

Overall, (i) *M. armenti* boasts two major subclusters, which are largely separated by recent gene flow; and (ii) flexible genes differentially present within these major subclusters (e.g., a type VII secretion system, toxin-antitoxin genes, beta-lactamase family genes) indicate that these two subclusters may differ phenotypically, although future experiments will be necessary to confirm this.

3.8. A novel GTDB genomospecies encompasses *Macrococcus* genomes from Switzerland and South Africa

Of the six *Macrococcus* spp. genomes sequenced in this study, five were assigned to *M. caseolyticus* (per GTDB-Tk; Figure 1; Supplementary Table S4). The genome of S115, however, could not be assigned to a known species via GTDB-Tk (Figure 1; Supplementary Table S4). Using bactaxR and a 95 ANI threshold (i.e., an approach similar to that of GTDB-Tk), three additional, publicly available genomes belonged to this putative novel GTDB genomospecies (i.e., bactaxR Cluster 2, $n = 4$ total genomes; Figures 1, 6B). In addition to (i) S115, a food-associated strain isolated in 2015 from beef biltong sold in South Africa's Limpopo province, this genomospecies included three strains isolated in Switzerland in 2019: (ii) 19Msa1099, isolated from pork meat (NCBI GenBank Assembly accession GCA_019357535.1), plus (iii) 19Msa1047 and (iv) 19Msa0499, each isolated from calf nasal swab samples (NCBI GenBank Assembly accessions GCA_019378895.1 and GCA_019788685.1, respectively; Supplementary Tables S1, S3). Notably, the South African genome was relatively distantly related to the Swiss genomes, sharing 99.2 ANI with each via OrthoANI and differing by 8,614–8,637 SNPs (identified via Snippy relative to each individual Swiss genome).

Comparatively, the three Swiss genomes shared >99.99 ANI with each other via OrthoANI and differed by 1–34 core SNPs (calculated via Snippy with the South African S115 strain excluded): strains 19Msa1047 (from a calf nasal swab) and 19Msa1099 (from pork meat) differed by a single core SNP identified in a gene annotated as a CBS domain-containing protein (NCBI Protein accession WP_219491817.1, corresponding to locus tag KYI07_RS05750 within the *M. caseolyticus* str. 19Msa0499 reference chromosome with NCBI Nucleotide accession NZ_CP079969.1). These two strains differed from strain 19Msa0499 (from a calf nasal swab) by 33 and 34 core SNPs, all of which fell within two regions of the *M. caseolyticus* str. 19Msa0499 reference chromosome: (i) 13 core SNPs within positions 312,553–367,746 bp, and (ii) 20 core SNPs within positions 1,778,236-1,778,444 bp, indicating that genetic differences within these regions may be due to recombination.

4. Discussion

In this study, WGS was used to characterize *Macrococcus* spp. strains isolated from South African cattle (i.e., two strains from bovine clinical mastitis cases) and beef products (i.e., two stains from RTE beef biltong and two from minced/processed beef products). Using these genomes in combination with all publicly available, high quality *Macrococcus* spp. genomes, insight is provided into the evolution, population structure, and functional potential of the *Macrococcus* genus as a whole. Importantly, we observed (i) differences in functional potential (e.g., AMR potential, virulence potential) between and within *Macrococcus* spp., and (ii) that some *Macrococcus* species lack clear boundaries at conventional genomospecies delineation thresholds (i.e., 95 ANI), which may cause taxonomic issues in the future. Below, we discuss these findings in detail, as well as (iii) future opportunities in the *Macrococcus* genomics space.

4.1. Differences in functional potential can be observed between and within *Macrococcus* species

Bacteria can adapt to stressors and stimuli in their respective environments through the acquisition of genomic material in the "flexible" gene pool (Arevalo et al., 2019). Thus, intraspecies differences in genomic content can be observed for many bacterial species (Tonkin-Hill et al., 2020), and differences resulting from recent gene flow (i.e., genomic elements acquired post-speciation) can be used to delineate populations within those species (Arevalo et al., 2019). Here, we queried all *Macrococcus* spp. genomes and identified genomic determinants variably present within species, indicative of within-species differences in functional potential. For example, of the 50 putative AMR and stress response determinants identified across *Macrococcus* in its entirety, nearly half (24 of 50, 48%) were species-specific (based on GTDB-Tk species assignments); of these species-specific AMR and stress response determinants, all (24 of 24, 100%) were variably present within their given species, indicating that AMR potential can vary within *Macrococcus* species. Antimicrobial exposure can select for AMR (Hendriksen et al., 2019; Olesen et al., 2020), and reducing exposure (e.g., limiting antimicrobial use outside of treating human disease, minimizing unnecessary antibiotic use for human illness cases) can reduce the risk of AMR (Murray et al., 2022). Thus, it is not particularly surprising that intraspecies differences in AMR potential exist within *Macrococcus*; the genomes aggregated here were derived from *Macrococcus* strains isolated from a range of sources (e.g., humans, animal hosts, animal products, environmental samples), geographic locations (i.e., four continents), and timeframes (i.e., between the years of 1916 and 2021) and thus have likely been exposed to different selective pressures.

Comparatively, some genomic elements identified here were present across multiple *Macrococcus* spp., indicating shared inter-species functional potential for some phenotypes. Methicillin resistance genes *mecB* and *mecD*, for example, were variably present within multiple *Macrococcus* species (via GTDB-Tk; Figure 3), mirroring previous studies, which have reported *mecB* and/or *mecD* in various *Macrococcus* spp., including *M. caseolyticus* (Schwendener et al., 2017; MacFadyen et al., 2018; Zhang et al., 2022), *M. bohemicus* (Foster and Paterson, 2020), *M. canis* (Chanchaithong et al., 2019), and *M. goetzii* (Maslanova et al., 2018). Outside of the AMR space, we further identified proteins that shared homology with virulence

factors in other species. Perhaps most notably, we detected homologues of aureolysin in multiple *Macrococcus* species (Figure 3). Aureolysin is an extracellular zinc-dependent metalloprotease secreted by *S. aureus*, which plays a crucial role in host immune system evasion (Thammavongsa et al., 2015; Pietrocola et al., 2017). While others have detected aureolysin homologues in *Macrococcus* genomes (Mazhar et al., 2019a,b; Zhang et al., 2022), the roles this protein plays in *Macrococcus* interactions with human or animal hosts (if any) are unknown.

Finally, for *M. caseolyticus* and *M. armenti*, which were composed of multiple populations (subclusters) separated by recent gene flow, some variably present genomic elements were subcluster-specific genes, which had been acquired post-speciation and differentially swept through these subclusters (i.e., flexible genes identified via PopCOGenT). Similar to results observed for *Ruminococcus gnavus* (Arevalo et al., 2019), transporter functions (e.g., ABC-type transporters, genes involved in ion transport) were enriched in subcluster-specific flexible gene sets within both *M. caseolyticus* and *M. armenti*. Notably, within *M. armenti*, we further identified two distinct subclusters with different flexible genes in each, including (i) one subcluster with a type VII secretion system, *S. aureus*-like virulence factors, and a putative pathogenicity island (Subcluster 2.0), and (ii) another with beta-lactamase family proteins and a type II toxin-antitoxin system (Subcluster 2.1). Taken together, these results indicate that there may be differences in the functional potential of these two *M. armenti* subclusters; however future experimental work will be needed to confirm the roles of these subcluster-specific flexible genes in each subcluster, as there are no clear differences in terms of each subcluster's ecological niche (strains in both subclusters were isolated from livestock in Switzerland).

Overall, proteins with potential virulence- and AMR-related functions, which were differentially present within and across *Macrococcus* species were identified. This indicates that there are potential within- and between-species differences in *Macrococcus* virulence and AMR potential. Future experimental efforts will thus be needed to investigate these differences further, as the study conducted here did not consider phenotypic data.

4.2. The lack of clear genomospecies boundaries between some *Macrococcus* species may cause taxonomic issues in the future

The delineation of prokaryotes into species-level taxonomic units is notoriously challenging, as horizontal gene transfer can obscure prokaryotic population boundaries (Jain et al., 2018; Arevalo et al., 2019). With the increasing availability of WGS, taxonomic assignment has largely shifted to *in silico* methods; however, numerous approaches exist for this purpose and may produce conflicting results (e.g., various implementations of ANI-based methods, marker gene-based methods, metrics using recent gene flow, *in silico* DNA–DNA hybridization; Meier-Kolthoff et al., 2013; Mende et al., 2013; Lee et al., 2016; Yoon et al., 2017; Jain et al., 2018; Arevalo et al., 2019; Chaumeil et al., 2019; Meier-Kolthoff et al., 2022; Parks et al., 2022). Here, we applied multiple species-level taxonomic assignment methods

to all publicly available *Macrococcus* genomes, specifically ANI-based approaches (i.e., OrthoANI/bactaxR and GTDB-Tk), an approach that uses a metric based on recent gene flow (i.e., PopCOGenT), and a marker gene-based method (i.e., specI; Figure 1). Overall, we observed similar results for three of four approaches; the marker gene-based approach only recovered two species, likely due to a lack of *Macrococcus* genomes of species other than *M. caseolyticus* and *M. canis* during species cluster database construction (this will likely be remedied in future specI database versions). However, even among the approaches that produced highly similar results, no two methods produced identical results.

Furthermore, at the conventional 95 ANI genomospecies threshold, several *Macrococcus* species were found to overlap (i.e., members of one species shared ≥95 ANI with members of a different species; Figure 2). We have previously observed a similar phenomenon among members of the *Bacillus cereus* group (Carroll et al., 2020), as others have done for *Escherichia/Shigella* spp., *Mycobacterium* spp., and *Neisseria gonorrhoeae/Neisseria meningitidis* (Jain et al., 2018). For *Macrococcus*, ambiguous species boundaries may not seem immediately concerning, as members of the genus are often viewed as animal commensals (Mazhar et al., 2018); thus, taxonomic misidentifications may not be viewed as "high consequence" compared to other organisms plagued by taxonomic issues (e.g., anthrax-causing organisms within the *Bacillus cereus* group, botulinum neurotoxin-producing members of *Clostridium*; Smith et al., 2018; Bower et al., 2022). However, as more *Macrococcus* strains undergo WGS and more is learned about the pathogenic potential of these organisms in animals and humans, there may be a greater need to ensure that species are clearly defined (e.g., in clinical laboratory, diagnostic, or regulatory settings). While there is some evidence that one of the South African genomes sequenced here belongs to a putative novel species (i.e., S115), we do not advocate for any changes to the taxonomy at this time, due to the limited number of genomes available. However, we encourage readers to be aware of ambiguous species boundaries for some *Macrococcus* spp., which may cause taxonomic issues in the future.

4.3. Future genomic sequencing, metadata collection, and phenotypic characterization efforts are needed to gain insight into *Macrococcus* population structure, antimicrobial resistance, and virulence potential

WGS has proven to be revolutionary in the food, veterinary, and human clinical microbiology spaces and is being used for—among other applications—pathogen surveillance, outbreak and cluster detection, source tracking, and diagnostics (Rossen et al., 2018; Brown et al., 2021; Ferdinand et al., 2021; Forde et al., 2023). Massive WGS efforts are being undertaken to query bacterial pathogens such as *Salmonella enterica*, *Escherichia coli*, and *Listeria monocytogenes* (Allard et al., 2016; Stevens et al., 2017; Brown et al., 2019), and large amounts of genomic data and metadata are publicly available for these organisms. As of 5 February 2023, (i) 455,330 genomes had been submitted to NCBI's GenBank Assembly database as *Salmonella*

enterica, (ii) 200,204 as *Escherichia coli*, and (iii) 51,579 as *Listeria monocytogenes*. *S. aureus* is a close relative of *Macrococcus* and boasts a total of 68,631 publicly available, assembled genomes (per NCBI's GenBank Assembly database, accessed 5 February 2023). These numbers dwarf those of *Macrococcus,* with 110 available, high-quality genomes for the entire genus at the time of this study, including the genomes generated here.

While our study provides insight into the evolution, population structure, and functional potential of *Macrococcus*, much more needs to be done to understand the role that *Macrococcus* spp. play as animal commensals, in animal-associated foodstuffs, and as opportunistic pathogens in animals and humans. First and foremost, future WGS efforts are needed to characterize these organisms, as increased availability of genomes will provide further insight into *Macrococcus* evolution (e.g., facilitating the identification of novel species, novel lineages within species). It is equally important that future WGS efforts are complemented with publicly available metadata (e.g., information conveying when and where a given strain was isolated) as this information can be used to identify potential host or geographic associations or potential migration or transmission events (e.g., between hosts or geographic regions). Finally, phenotypic data will be essential to confirm or invalidate the preliminary findings posited here regarding *Macrococcus* functional potential. Genomic AMR prediction, for example, does not necessarily translate to phenotypic AMR (Ransom et al., 2020). Similarly, any genomic determinants identified here based on their homology to known virulence factors (e.g., aureolysin, PVL, immune inhibitor A, the type VII secretion system identified in one *M. armenti* subcluster) must be evaluated experimentally. Thus, we hope that the results provided here can serve as a guide for further studies of the AMR and virulence potential of *Macrococcus* spp.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

References

Alexa, A., Rahnenfuhrer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22, 1600–1607. doi: 10.1093/bioinformatics/btl140

Ali, D. E., Allam, M., Altayb, H. N., Mursi, D., Adalla, M. A., Mohammed, N. O., et al. (2022). A prevalence and molecular characterization of novel pathogenic strains of *Macrococcus caseolyticus* isolated from external wounds of donkeys in Khartoum state -Sudan. *BMC Vet. Res.* 18:197. doi: 10.1186/s12917-022-03297-2

Allard, M. W., Strain, E., Melka, D., Bunning, K., Musser, S. M., Brown, E. W., et al. (2016). Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J. Clin. Microbiol.* 54, 1975–1983. doi: 10.1128/JCM.00081-16

Arevalo, P., Vaninsberghe, D., Elsherbini, J., Gore, J., and Polz, M. F. (2019). A reverse ecology approach based on a biological definition of microbial populations. *Cell* 178:e814, 820–834.e14. doi: 10.1016/j.cell.2019.06.033

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., et al. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, W29–W35. doi: 10.1093/nar/gkab335

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Author contributions

LC performed all computational analyses. IM performed bacterial isolation and identification as well as DNA extraction. RP supervised the sequencing of the isolates. IM and KM sourced the funding for sequencing of the isolates. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1181376/full#supplementary-material

Bower, W. A., Hendricks, K. A., Vieira, A. R., Traxler, R. M., Weiner, Z., Lynfield, R., et al. (2022). What is Anthrax? *Pathogens* 11:690. doi: 10.3390/pathogens11060690

Brown, B., Allard, M., Bazaco, M. C., Blankenship, J., and Minor, T. (2021). An economic evaluation of the whole genome sequencing source tracking program in the U.S. *PLoS One* 16:e0258262. doi: 10.1371/journal.pone.0258262

Brown, E., Dessai, U., Mcgarry, S., and Gerner-Smidt, P. (2019). Use of whole-genome sequencing for food safety and public health in the United States. *Foodborne Pathog. Dis.* 16, 441–450. doi: 10.1089/fpd.2019.2662

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421

Cantalapiedra, C. P., Hernandez-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol. Biol. Evol.* 38, 5825–5829. doi: 10.1093/molbev/msab293

Carroll, L. M., Larralde, M., Fleck, J. S., Ponnudurai, R., Milanese, A., Cappio, E., et al. (2021). Accurate *de novo* identification of biosynthetic gene clusters with GECCO. *bioRxiv*:2021.2005.2003.442509.

Carroll, L. M., Wiedmann, M., and Kovac, J. (2020). Proposal of a taxonomic nomenclature for the *Bacillus cereus* group which reconciles genomic definitions of bacterial species with clinical and industrial phenotypes. *MBio* 11, e00034–e00020. doi: 10.1128/mBio.00034-20

Chanchaithong, P., Perreten, V., and Schwendener, S. (2019). *Macrococcus canis* contains recombinogenic methicillin resistance elements and the *mecB* plasmid found in *Staphylococcus aureus*. *J. Antimicrob. Chemother.* 74, 2531–2536. doi: 10.1093/jac/dkz260

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. doi: 10.1093/bioinformatics/btz848

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80–92. doi: 10.4161/fly.19695

Cotting, K., Strauss, C., Rodriguez-Campos, S., Rostaher, A., Fischer, N. M., Roosje, P. J., et al. (2017). *Macrococcus canis* and *M. caseolyticus* in dogs: occurrence, genetic diversity and antibiotic resistance. *Vet. Dermatol.* 28, 559–e133. doi: 10.1111/vde.12474

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 43:e15. doi: 10.1093/nar/gku1196

Evans, A. C. (1916). The Bacteria of Milk Freshly Drawn From Normal Udders. *J. Infect. Dis.* 18, 437–476. doi: 10.1093/infdis/18.5.437

Ewels, P., Magnusson, M., Lundin, S., and Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354

Feldgarden, M., Brover, V., Haft, D. H., Prasad, A. B., Slotta, D. J., Tolstoy, I., et al. (2019). Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob. Agents Chemother.* 63, e00483–19. doi: 10.1128/AAC.00483-19

Ferdinand, A. S., Kelaher, M., Lane, C. R., Da Silva, A. G., Sherry, N. L., Ballard, S. A., et al. (2021). An implementation science approach to evaluating pathogen whole genome sequencing in public health. *Genome Med.* 13:121. doi: 10.1186/s13073-021-00934-7

Forde, B. M., Bergh, H., Cuddihy, T., Hajkowicz, K., Hurst, T., Playford, E. G., et al. (2023). Clinical implementation of routine whole-genome sequencing for hospital infection control of multi-drug resistant pathogens. *Clin. Infect. Dis.* 76, e1277–e1284. doi: 10.1093/cid/ciac726

Foster, G., and Paterson, G. K. (2020). Methicillin-resistant *Macrococcus bohemicus* encoding a divergent SCC*mecB* element. *Antibiotics* 9:590. doi: 10.3390/antibiotics9090590

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*:1207.3907.

Gobeli Brawand, S., Cotting, K., Gomez-Sanz, E., Collaud, A., Thomann, A., Brodard, I., et al. (2017). *Macrococcus canis* sp. nov., a skin bacterium associated with infections in dogs. *Int. J. Syst. Evol. Microbiol.* 67, 621–626. doi: 10.1099/ijsem.0.001673

Gomez-Sanz, E., Schwendener, S., Thomann, A., Gobeli Brawand, S., and Perreten, V. (2015). First staphylococcal cassette chromosome *mec* containing a *mecB*-carrying gene complex independent of transposon Tn6045 in a *Macrococcus canis* isolate from a canine infection. *Antimicrob. Agents Chemother.* 59, 4577–4583. doi: 10.1128/AAC.05064-14

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Hendriksen, R. S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., et al. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* 10:1124. doi: 10.1038/s41467-019-08853-3

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernandez-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/nar/gky1085

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9

Jolley, K. A., Bray, J. E., and Maiden, M. C. J. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* 3:124. doi: 10.12688/wellcomeopenres.14826.1

Jolley, K. A., and Maiden, M. C. (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinform.* 11:595. doi: 10.1186/1471-2105-11-595

Jost, G., Schwendener, S., Liassine, N., and Perreten, V. (2021). Methicillin-resistant *Macrococcus canis* in a human wound. *Infect. Genet. Evol.* 96:105125. doi: 10.1016/j.meegid.2021.105125

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kautsar, S. A., Blin, K., Shaw, S., Navarro-Munoz, J. C., Terlouw, B. R., Van Der Hooft, J. J. J., et al. (2020). MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 48, D454–D458. doi: 10.1093/nar/gkz882

Keller, J. E., Schwendener, S., Overesch, G., and Perreten, V. (2022). *Macrococcus armenti* sp. nov., a novel bacterium isolated from the skin and nasal cavities of healthy pigs and calves. *Int. J. Syst. Evol. Microbiol.* 72. doi: 10.1099/ijsem.0.005245

Kimura, M. (1981). Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci. U. S. A.* 78, 454–458. doi: 10.1073/pnas.78.1.454

Kitts, P. A., Church, D. M., Thibaud-Nissen, F., Choi, J., Hem, V., Sapojnikov, V., et al. (2016). Assembly: a resource for assembled genomes at NCBI. *Nucleic Acids Res.* 44, D73–D80. doi: 10.1093/nar/gkv1226

Kloos, W.E., Ballard, D.N., George, C.G., Webster, J.A., Hubner, R.J., Ludwig, W., Schleifer, K.H., Fiedler, F., and Schubert, K. (1998). Delimiting the genus *Staphylococcus* through description of *Macrococcus caseolyticus* gen. nov., comb. nov. and *Macrococcus equipercicus* sp. nov., and *Macrococcus bovicus* sp. no. and *Macrococcus carouselicus* sp. nov. *Int. J. Syst. Bacteriol.* 48 Pt, 859–877, doi: 10.1099/00207713-48-3-859

Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067

Lee, I., Ouk Kim, Y., Park, S. C., and Chun, J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* 66, 1100–1103. doi: 10.1099/ijsem.0.000760

Letunic, I., and Bork, P. (2021). Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*:1303.3997.

Li, H. (2019). "Seqtk: a fast and lightweight tool for processing sequences in the FASTA or FASTQ format". 1.2-r102-dirty ed.).

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* 47, D687–D692. doi: 10.1093/nar/gky1080

Loffler, B., Hussain, M., Grundmeier, M., Bruck, M., Holzinger, D., Varga, G., et al. (2010). *Staphylococcus aureus* Panton-Valentine leukocidin is a very potent cytotoxic factor for human neutrophils. *PLoS Pathog.* 6:e1000715. doi: 10.1371/journal.ppat.1000715

Macfadyen, A. C., Fisher, E. A., Costa, B., Cullen, C., and Paterson, G. K. (2018). Genome analysis of methicillin resistance in *Macrococcus caseolyticus* from dairy cattle in England and Wales. *Microb. Genom.* 4:e000191. doi: 10.1099/mgen.0.000191

Mannerova, S., Pantucek, R., Doskar, J., Svec, P., Snauwaert, C., Vancanneyt, M., et al. (2003). *Macrococcus brunensis* sp. nov., *Macrococcus hajekii* sp. nov. and *Macrococcus lamae* sp. nov., from the skin of llamas. *Int. J. Syst. Evol. Microbiol.* 53, 1647–1654. doi: 10.1099/ijs.0.02683-0

Maslanova, I., Wertheimer, Z., Sedlacek, I., Svec, P., Indrakova, A., Kovarovic, V., et al. (2018). Description and comparative genomics of *Macrococcus caseolyticus* subsp. *hominis* subsp. nov., *Macrococcus goetzii* sp. nov., *Macrococcus epidermidis* sp. nov., and *Macrococcus bohemicus* sp. nov., novel macrococci from human clinical material with virulence potential and suspected uptake of foreign DNA by natural transformation. *Front. Microbiol.* 9:1178. doi: 10.3389/fmicb.2018.01178

Mazhar, S., Altermann, E., Hill, C., and Mcauliffe, O. (2019a). Draft genome sequences of *Macrococcus caseolyticus*, *Macrococcus canis*, *Macrococcus bohemicus*, and *Macrococcus goetzii*. *Microbiol. Resour. Announc.* 8, e00343–19. doi: 10.1128/MRA.0343-19

Mazhar, S., Altermann, E., Hill, C., and Mcauliffe, O. (2019b). Draft genome sequences of the type strains of six *Macrococcus* species. *Microbiol. Resour. Announc.* 8, e00343–19. doi: 10.1128/MRA.00344-19

Mazhar, S., Hill, C., and Mcauliffe, O. (2018). The genus *Macrococcus*: an insight into its biology, evolution, and relationship with *Staphylococcus*. *Adv. Appl. Microbiol.* 105, 1–50. doi: 10.1016/bs.aambs.2018.05.002

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform.* 14:60. doi: 10.1186/1471-2105-14-60

Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L., and Goker, M. (2022). TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* 50, D801–D807. doi: 10.1093/nar/gkab902

Mende, D. R., Sunagawa, S., Zeller, G., and Bork, P. (2013). Accurate and universal delineation of prokaryotic species. *Nat. Methods* 10, 881–884. doi: 10.1038/nmeth.2575

Milanese, A., Mende, D. R., Paoli, L., Salazar, G., Ruscheweyh, H. J., Cuenca, M., et al. (2019). Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat. Commun.* 10:1014. doi: 10.1038/s41467-019-08844-4

Minh, B. Q., Nguyen, M. A., and Von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* 30, 1188–1195. doi: 10.1093/molbev/mst024

Murray, C. J. L., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., et al. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 399, 629–655. doi: 10.1016/S0140-6736(21)02724-0

Navarro-Munoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson, E. I., et al. (2020). A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* 16, 60–68. doi: 10.1038/s41589-019-0400-9

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Olesen, S. W., Lipsitch, M., and Grad, Y. H. (2020). The role of "spillover" in antibiotic resistance. *Proc. Natl. Acad. Sci. U. S. A.* 117, 29063–29068. doi: 10.1073/pnas.2013694117

Ouoba, L. I. I., Vouidibio Mbozo, A. B., Anyogu, A., Obioha, P. I., Lingani-Sawadogo, H., Sutherland, J. P., et al. (2019). Environmental heterogeneity of *Staphylococcus* species from alkaline fermented foods and associated toxins and antimicrobial resistance genetic elements. *Int. J. Food Microbiol.* 311:108356. doi: 10.1016/j.ijfoodmicro.2019.108356

Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., et al. (2016). SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom.* 2:e000056. doi: 10.1099/mgen.0.000083

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412

Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi: 10.1093/bioinformatics/bty633

Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P. A., and Hugenholtz, P. (2022). GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* 50, D785–D794. doi: 10.1093/nar/gkab776

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Parte, A. C., Sarda Carbasse, J., Meier-Kolthoff, J. P., Reimer, L. C., and Goker, M. (2020). List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int. J. Syst. Evol. Microbiol.* 70, 5607–5612. doi: 10.1099/ijsem.0.004332

Pietrocola, G., Nobile, G., Rindi, S., and Speziale, P. (2017). *Staphylococcus aureus* manipulates innate immunity through own and host-expressed proteases. *Front. Cell. Infect. Microbiol.* 7:166. doi: 10.3389/fcimb.2017.00166

Poyart, C., Quesne, G., Boumaila, C., and Trieu-Cuot, P. (2001). Rapid and accurate species-level identification of coagulase-negative staphylococci by using the *sodA* gene as a target. *J. Clin. Microbiol.* 39, 4296–4301. doi: 10.1128/JCM.39.12.4296-4301.2001

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

R Core Team (2021). *"R: A language and environment for statistical computing". 4.1.2 ed.* (Vienna, Austria: R Foundation for Statistical Computing).

Ramos, G., Vigoder, H. C., and Nascimento, J. S. (2021). Technological applications of *Macrococcus caseolyticus* and its impact on food safety. *Curr. Microbiol.* 78, 11–16. doi: 10.1007/s00284-020-02281-z

Ransom, E. M., Potter, R. F., Dantas, G., and Burnham, C. D. (2020). Genomic prediction of antimicrobial resistance: ready or not, here it comes! *Clin. Chem.* 66, 1278–1289. doi: 10.1093/clinchem/hvaa172

Rossen, J. W. A., Friedrich, A. W., Moran-Gilad, J., Genomic, E. S. G. F., and Molecular, D. (2018). Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. *Clin. Microbiol. Infect.* 24, 355–360. doi: 10.1016/j.cmi.2017.11.001

Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., et al. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* 2020: baaa062. doi: 10.1093/database/baaa062

Schwendener, S., Cotting, K., and Perreten, V. (2017). Novel methicillin resistance gene *mecD* in clinical *Macrococcus caseolyticus* strains from bovine and canine sources. *Sci. Rep.* 7:43797. doi: 10.1038/srep43797

Schwendener, S., and Perreten, V. (2022). The *bla* and *mec* families of beta-lactam resistance genes in the genera *Macrococcus*, *Mammaliicoccus* and *Staphylococcus*: an in-depth analysis with emphasis on *Macrococcus*. *J. Antimicrob. Chemother.* 77, 1796–1827. doi: 10.1093/jac/dkac107

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Seemann, T. (2019). "samclip: filter SAM file for soft and hard clipped alignments". 0.2 ed.

Shallcross, L. J., Fragaszy, E., Johnson, A. M., and Hayward, A. C. (2013). The role of the Panton-Valentine leucocidin toxin in staphylococcal disease: a systematic review and meta-analysis. *Lancet Infect. Dis.* 13, 43–54. doi: 10.1016/S1473-3099(12)70238-4

Simonsen, M., Mailund, T., and Pedersen, C. N. S. (2008). "Rapid neighbour-joining" in *Algorithms in bioinformatics. Lecture Notes in Computer Science(),* vol 5251. eds. K. A. Crandall and J. Lagergren (Springer Berlin Heidelberg), 113–122.

Smith, T., Williamson Charles, H. D., Hill, K., Sahl, J., and Keim, P. (2018). Botulinum Neurotoxin-Producing Bacteria. Isn't it Time That We Called a Species a Species? *MBio* 9, e01469–e01418. doi: 10.1128/mBio.01469-18

Soubrier, J., Steel, M., Lee, M. S. Y., Der Sarkissian, C., Guindon, S., Ho, S. Y. W., et al. (2012). The Influence of Rate Heterogeneity Among Sites on the Time Dependence of Molecular Rates. *Mol. Biol. Evol.* 29, 3345–3358. doi: 10.1093/molbev/mss140

Souvorov, A., Agarwala, R., and Lipman, D. J. (2018). SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol.* 19:153. doi: 10.1186/s13059-018-1540-z

Steinegger, M., and Soding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* 35, 1026–1028. doi: 10.1038/nbt.3988

Stevens, E. L., Timme, R., Brown, E. W., Allard, M. W., Strain, E., Bunning, K., et al. (2017). The public health impact of a publically available, environmental database of microbial genomes. *Front. Microbiol.* 8:808. doi: 10.3389/fmicb.2017.00808

Tan, A., Abecasis, G. R., and Kang, H. M. (2015). Unified representation of genetic variants. *Bioinformatics* 31, 2202–2204. doi: 10.1093/bioinformatics/btv112

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect. Math. Life Sci.* 17, 57–86.

Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008). Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477. doi: 10.1016/j.mib.2008.09.006

Thammavongsa, V., Kim, H. K., Missiakas, D., and Schneewind, O. (2015). Staphylococcal manipulation of host immune responses. *Nat. Rev. Microbiol.* 13, 529–543. doi: 10.1038/nrmicro3521

The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47, D330–D338. doi: 10.1093/nar/gky1055

Tonkin-Hill, G., Lees, J. A., Bentley, S. D., Frost, S. D. W., and Corander, J. (2018). RhierBAPS: an R implementation of the population clustering algorithm hierBAPS. *Wellcome Open Res.* 3:93. doi: 10.12688/wellcomeopenres.14694.1

Tonkin-Hill, G., Macalasdair, N., Ruis, C., Weimann, A., Horesh, G., Lees, J. A., et al. (2020). Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol.* 21:180. doi: 10.1186/s13059-020-02090-4

Tshipamba, M. E., Lubanza, N., Adetunji, M. C., and Mwanza, M. (2018). Molecular characterization and antibiotic resistance of foodborne pathogens in street-vended ready-to-eat meat sold in South Africa. *J. Food Prot.* 81, 1963–1972. doi: 10.4315/0362-028X.JFP-18-069

Vaitkevicius, K., Rompikuntal, P. K., Lindmark, B., Vaitkevicius, R., Song, T., and Wai, S. N. (2008). The metalloprotease PrtV from *Vibrio cholerae*. *FEBS J.* 275, 3167–3177. doi: 10.1111/j.1742-4658.2008.06470.x

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York.

Yang, Z. (1995). A space-time process model for the evolution of DNA sequences. *Genetics* 139, 993–1005. doi: 10.1093/genetics/139.2.993

Yoon, S. H., Ha, S. M., Lim, J., Kwon, S., and Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110, 1281–1286. doi: 10.1007/s10482-017-0844-4

Zhang, Y., Min, S., Sun, Y., Ye, J., Zhou, Z., and Li, H. (2022). Characteristics of population structure, antimicrobial resistance, virulence factors, and morphology of methicillin-resistant *Macrococcus caseolyticus* in global clades. *BMC Microbiol.* 22:266. doi: 10.1186/s12866-022-02679-8

Zhou, Z., Charlesworth, J., and Achtman, M. (2020). Accurate reconstruction of bacterial pan- and core genomes with PEPPAN. *Genome Res.* 30, 1667–1679. doi: 10.1101/gr.260828.120

Check for updates

Drug sensitivity and genome-wide analysis of two strains of *Mycoplasma gallisepticum* with different biofilm intensity

Xiaoyan Ma[1], Li Wang[1], Fei Yang[1], Jidong Li[1], Lei Guo[2], Yanan Guo[3] and Shenghu He[1]*

[1]Clinical Veterinary Laboratory, Institute of Animal Science and Technology, Ningxia University, Yinchuan, China, [2]Ningxia Xiaoming Agriculture and Animal Husbandry Co., Ltd., Yinchuan, China, [3]Ningxia Academy of Agriculture and Forestry Sciences, Yinchuan, China

Mycoplasma gallisepticum (MG) is one of the major causative agents of chronic respiratory diseases in poultry. The biofilms of MG are highly correlated to its chronic infection. However data on genes involved in biofilm formation ability are still scarce. MG strains with distinct biofilm intensity were screened by crystal violet staining morphotyped and characterized for the drug sensitivity. Two MG strains NX-01 and NX-02 showed contrasted ability to biofilm formation. The biofilm formation ability of NX-01 strain was significantly higher than that of NX-02 strain ($p < 0.01$). The drug sensitivity test showed that the stronger the ability of MG stain to form biofilms, the weaker its sensitivity to 17 antibiotic drugs. Moreover, putative key genes related to biofilm formation were screened by genome-wide analysis. A total of 13 genes and proteins related to biofilm formation, including *ManB, oppA, oppD, PDH, eno, RelA, msbA, deoA, gapA, rpoS,* Adhesin P1 precursor, S-adenosine methionine synthetase, and methionyl tRNA synthetase were identified. There were five major discrepancies between the two isolated MG strains and the five NCBI-published MG strains. These findings provide potential targets for inhibiting the formation of biofilm of MG, and lay a foundation for treating chronic infection.

KEYWORDS

Mycoplasma gallisepticum, biofilm, formation, drug sensitivity, whole-genome analysis

1. Introduction

Mycoplasma gallisepticum (MG) is one of the important pathogens affecting poultry (Ghanem et al., 2017). It is widely distributed in all poultry countries worldwide, causing serious economic losses to the global poultry industry (Roussan et al., 2015; Feberwee et al., 2021; Marouf et al., 2022). MG mainly causes chronic respiratory diseases in chickens and other poultry and infectious sinusitis in turkeys (Sulyok et al., 2019). In addition, the infection of MG can cause low carcass grading and feed conversion efficiency of broiler chickens. With the increase of the weak brood rate, the egg production and egg hatching rate of laying hens tend to decrease (Beaudet et al., 2019).

Many therapeutic drugs have been developed for MG infection, such as macrolides, quinolones, tylomycin, and gentamicin (McAuliffe et al., 2006; Abd El-Hamid et al., 2019). However, the drug resistance of MG is becoming increasingly severe (Beylefeld et al., 2018), and biofilms may be one of the main causes. Bacterial biofilm plays an important role in the bacterial

disease process, allowing bacteria to evade the host's immune defenses, inducing drug resistance and increasing toxin accumulation (Sorci, 2013). The formation of biofilm poses significant clinical challenges and exacerbates the issues of drug resistance and bacterial virulence (Tassew et al., 2017; Feng et al., 2021). Studies have shown that the development of a biofilm by MG could contribute to persistent infections, rendering the eradication of MG challenging (Feng et al., 2020). The complete characterization of MG biofilm formation is key in assessing and combating resistance to infectious poultry diseases, particularly in commercial farming (Taylor-Robinson and Bébéar, 1997). Despite the wealth of information and corresponding treatment modalities available for biofilm formation in bacteria and other mycoplasma (Stewart, 2002; Sharma et al., 2019), there is a lack of research on MG biofilm.

Mycoplasma gallisepticum is a prokaryote intermediate between bacteria and viruses and lacks a cell wall (Ishfaq et al., 2020). A few key genes regulate the formation and attachment of bacterial biofilms to surfaces (Parker et al., 2009; Liu et al., 2022). For example, the formation of *Streptococcus gordonii* biofilm depends on arginine, which regulates the development of *Streptococcus gordonii* biofilm mainly by activating the regulatory pathway of the ArcR gene (Robinson et al., 2018). In addition, the *ssrS* gene of *Haemophilus parasuis* is associated with the antitoxin system, regulating biofilm formation and cell persistence (Jiang et al., 2021). Several genes that have a crucial role in the biofilm formation of various strains have been well-documented. However, there is a significant paucity of research on MG. So far, only a few articles have shown that *ManB*, *VlhA*, *PDH*, ABC transporter penetrase, ABC transporter ATP-binding protein, and other genes may be involved in forming MG biofilm (Chen et al., 2012; Wang et al., 2017). It is very important to study the regulatory genes of biofilm formation for controlling MG chronic infection. Genomic analyses can help elucidate genes functionalities in MG as recently documented (Semashko et al., 2022). Such approach could help to elucidate putative genes involved in biofilm formation ability and drug resistance.

The objective of this study was to screen MG strains with contrasted ability in terms of biofilm formation in Ningxia Hui Autonomous Region in China and study their drug sensitivity. It also aimed to gaining insights in the putative genes involved in biofilm formation ability using a genome-wide analysis on both strains having contrasted ability to biofilm formation.

2. Materials and methods

2.1. Main materials

The improved Frey culture medium was obtained from China Sea Biotechnology Co., LTD., Beijing, China. The bacterial genome DNA extraction kit was purchased from Tiangen Biochemical Technology Co., LTD., Beijing, China. Crystal violet was acquired from Chongming County Yuxi Reagent Factory. Ethanol (95%) and methanol were purchased from Guangnuo Chemical Technology Co., LTD., Shanghai, China. Glutaraldehyde fixative (2.5%, for electron microscopy) was from Proanti Biotechnology Development Co., LTD., Shaanxi, China. Ofloxacin, Ennofloxacin, Norfloxacin, Taymycin, Erythromycin, Talomycin, Temicoxacin, Spiramycin, Tetracycline Hydrochloride, Lincomycin, Gentamicin, Dakanamycin, Kanamycin,

Streptomycin, Doxycycline, and Warnemylin were purchased from Golden Clonix Biotechnology Co., LTD., Beijing, China.

2.2. Bacterial strain isolation

Ethics approval and specific permission were not required for the study, as all samples were collected by the authors during necropsies with the consent of the owners. Air sacs of chickens suspected to be infected with *M. gallisepticum* were collected from poultry farms in Ningxia (Felice et al., 2020; Yadav et al., 2022). They were cut and placed in modified Frey's liquid medium overnight at 4°C under sterile conditions. After filtration, they were inoculated into Frey liquid medium and incubated in a biochemical incubator at 37°C. The MG were purified by Kahya et al. (2010) and Sprygin et al. (2010), the harvested isolates were coated in the modified Frey's solid medium and placed in a 5% CO_2 incubator at 37°C for 3–7 days. Single colonies were selected and cultured in the modified Frey's liquid medium, and the purified isolates were repeated three times to get the purified isolates. The colony morphology on the solid medium was observed with an inverted microscope.

Bacterial genomic DNA was extracted using the kit of Tiangen Biochemical Technology Company (Ahmad et al., 2020; Li L. et al., 2020). After the expanded culture, put 1 mL of fresh bacterial solution into a centrifuge tube and add 20 μL of protease K solution. Add 220 μL anhydrous ethanol and mix thoroughly. Then, the solution and precipitate are added to the adsorption Add 220 μL anhydrous ethanol and mix thoroughly. Then the solution and precipitate are added to the adsorption column, and 500 μL buffer GD is added to the column. Discard the waste liquid after centrifugation. Add 600 μL PW to the adsorption column, and the waste liquid was discarded after centrifugation. Add 50 μL eluent buffer TE in the middle of the adsorbed strain and let it stand at room temperature for 5 min. The solution was collected into the centrifuge tube to obtain the genome extract of MG strains. Meanwhile the isolated MG strain was amplified by PCR based on the OIE 16S rRNA and mgc2 primers to identify the MG strain in the bacterial solution (Table 1), nuclease-free water was used as a negative control in all PCR assays. Finally, the bacterial solution was mixed at a ratio of 1:1 to glycerol and stored at −80°C.

2.3. Determination of biofilm formation capacity

Modified tests using crystal violet staining were used to quantify biofilm formation (Hu et al., 2010). The suspension concentration of the prepared MG strains was adjusted to 1×10^7 CCU/mL, and 200 μL was inoculated on sterile 96-cell plates. Six replicates of the bacterial culture were prepared and subsequently incubated at 37°C for 12, 24, 36, and 48 h. Meanwhile, the crystal violet staining method was employed to measure the OD_{590} value of MG biofilm by enzyme marker, and the absorbance results were recorded. According to ODc value (ODc is equal to the average OD value of blank well), biofilm classification: $OD \leq ODc$, no membrane capacity (−); $ODc < OD \leq 2ODc$, weak biofilm forming ability (+); $2ODc < OD \leq 4ODc$, medium biofilm forming ability (++); and $4ODc < OD$, strong biofilm forming ability (+++).

TABLE 1 List of *M.gallisepticum* strains PCR primer information.

Gene	primer	Sequence (5′-3′)	Stripe/bp	Tm/°C
16S rRNA	MG-F	GAGCTAATCTGTAAAGTTGGTC	183	55°C
	MG-R	GCTTCCTTGCGGTTAGCAAC		
mgc2	MG-F	CGCAATTTGGTCCTNATCCCCAACA	236–302	54°C
	MG-R	TAAACCCRCCTCCAGCTTTATTTCC		

2.4. Determination of matrix growth curves of biofilm models *in vitro*

The growth curves of the biofilm matrix were measured by a semi-quantitative adhesion test (Liang et al., 2019). 200 μL of 1×10^6 CCU/mL bacterial suspension was inoculated on 96-cell plates with six replicates per strain, after which the plates were then incubated at 37°C for 0, 3, 6, 9, 12, 24, 36, 48, and 72 h. In the corresponding period time, crystal violet staining was used to detect the OD_{590} value of two MG biofilms. According to the absorbance results, the biofilm matrix growth curve was drawn.

2.5. Observation of biofilm morphology by scanning electron microscopy

Mycoplasma gallisepticum strains were inoculated into the liquid Frey medium for 12 h, and the suspension concentration was adjusted to 1×10^4 CCU/mL. The sterile six-well plate was used to hold the 14 mm round slip-plate, with 5 mL of bacterial suspension added to each well. The culture was incubated at 37°C for 24 h, and 1 mL 2.5% glutaraldehyde fixed solution (EM Grade) was added to fixed at 4°C overnight. The gradient elution was performed on each well was gradient eluted with ethanol of different concentration (30, 50, 70, 80, 90, and 95%). The obtained sample was sliced and photographed using an ISI-SX-40 scanning electron microscope (Shatila et al., 2021).

2.6. Drug sensitivity test

The drug sensitivity of the isolated stains were determined by Hannan's microdilution method (Hannan, 2000). The suspension concentration of MG strains was adjusted to 1×10^4 CCU/mL. 40 μL 512 μg/mL of antibiotics were added to the first hole of the 96-well plate, and add 20 μL sterile PBS from well 2 to 12. Then 20 μL of liquid was discarded from well 1 to 2 and mixed, and the same operation was performed until well 12. Thus, the antibiotics were continuously diluted at multiple ratios. Finally, 180 μL diluted MG bacterial solution was added to well 1–12. Growth control wells (containing culture medium and bacterial solution) and negative control Wells (containing liquid culture medium) were also set up and each antibiotic were repeated three times, After being inoculated, the 96-well plates were transferred to a constant temperature incubator set at 37°C for a week. The color of the growth control well was yellow, and the negative control well was pink. The highest concentration of the drug with no color change in the test well was considered the lowest inhibitory concentration of the drug (Gautier-Bouchardon, 2018).

2.7. Genome-wide sequencing and quality control

Sangon Bioengineering (Shanghai) Co., LTD. completed the genome-wide sequencing of the MG strains. MG complete genome sequences were obtained using Illumina HiSeq platform (Caporaso et al., 2012) for MG genome sequencing. The original image data file of the MG genome was processed by Base Calling and stored in FASTQ (Brown et al., 2017) file format. Next, the FastQC software was used to evaluate sequencing data quality. Using Trimmomatic (Bolger et al., 2014) was used to clear joint sequences and lower quality sequences to ensure that all the clean reads false discovery rate was <0.01.

2.8. Genome splicing and component analysis and gene function annotation

To splice sequencing data, the SPAdes (Bankevich et al., 2012) was utilized (Table 2). In order to fill gaps in the resulting contigs, the GapFiller (Boetzer and Pirovano, 2012) was employed. The accuracy of the spliced sequences was improved using PrInSeS-G (Massouras et al., 2010) correction tool that can rectify splicing errors during editing and detect small insertion losses. The genetic elements were predicted using Prokka (Seemann, 2014), and repeated sequences in the genome were identified using RepeatMasker (Saha et al., 2008). The CRISPRs were predicted through CRT (Bland et al., 2007) analysis. To obtain functional annotation information, gene sequences were analyzed separately using NCBI NR (Yu and Zhang, 2013), COG (Tatusov et al., 2000), GO (The Gene Ontology, 2019), and KEGG (Kanehisa and Goto, 2000), VFDB (Chen et al., 2016), CARD (McArthur et al., 2013), and CAZy (Lombard et al., 2014), and compared using NCBI Blast + (Altschul et al., 1997).

2.9. Genome-wide comparative analysis

The annotation data of the genome-wide of the locally isolated MG strains were analyzed, including total gene length, GC content, coding gene, etc. Molecular evolutionary tree were constructed using N-J method and MEGA6 software (Tamura et al., 2013). The whole-genome sequences of MG strain were compared with those of the five MG strains available in the NCBI database. Subsequently, the linear analysis of genome-wide sequences of MG strains were conducted using the MUMmer (Delcher et al., 2003; Table 3). Finally, a plot was created to visualize the region of genetic differences.

TABLE 2 List of bioinformatics tools and databases for whole genome sequencing of *M.gallisepticum*.

Name	URL	Edition
Illumina HiSeq platform	https://www.illumina.com.cn/	/
FASTQ	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	v0.11.2
Trimmomatic	http://www.usadellab.org/cms/index.php?page=trimmomatic	v0.36
SPAdes	http://bioinf.spbau.ru/spades	v3.5.0
GapFiller	http://www.baseclear.com/bioinformatics-tools/	v1.11
PrInSeS-G	https://updeplasrv1.epfl.ch/prinses/	v1.0.0
Prokka	http://vicbioinformatics.com/	v1.10
RepeatMasker	www.repeatmasker.org	v4.0.5
NCBI Blast +	https://blast.ncbi.nlm.nih.gov/Blast.cgi	v2.2.28
MEGA6	https://mega6.com.uy/Mega6/inicio	v6.0
CRT	http://www.room220.com/crt	v1.2
MUMmer	https://mummer4.github.io/	v4.0.0
NCBI NR	http://ncbi.nlm.nih.gov/	/
COG	https://www.ncbi.nlm.nih.gov/COG/	/
GO	http://www.geeontology.org	/
KEGG	http://www.kegg.jp	/
VFDB	http://www.mgc.ac.cn/VFs/	/
CARD	https://card.mcmaster.ca/	/
CAZy	http://www.cazy.org/	/
NCBI	http://ncbi.nlm.nih.gov/	/

TABLE 3 Strains of *Mycoplasma gallisepticum* for genome alignment.

Organism	Strain	RefSeq	Source area
Mycoplasma gallisepticum	NX-01	PRJNA972534	China
	NX-02	PRJNA972540	China
	NCTC10115	NZ_LS991952.1	Britain
	KUVMG001	NZ_CP070622.1	South Korea
	6/85	NZ_CP044224.1	United States
	mx-4	NZ_CP044226.1	United States
	f99 lab strain	NZ_CP028146.1	United States
	ts-11	NZ_CP044225.1	United States
	k5234	NZ_CP092251.1	United States
	CA06_2006.052-5-2P	NC_018412.1	United States
	S6	NC_023030.2	Russia
	R(low)	NC_004829.2	United States
	F	NC_017503.1	United States

3. Results

3.1. The bacterial strains showed contrasted ability in biofilm formation

According to strain isolation and identification, five MG strains NX-01, NX-02, NX-03, NX-04, and NX-05 were isolated during this study (Figures 1, 2). The capacity of biofilm formation showed that all 5 MG strains could form biofilms, with the following formation ability: NX-01 > NX-03 > NX-05 > NX-04 > NX-02 (Figure 3A), whereas the *M. gallisepticum* strains NX-01 and NX-03 showed strong biofilm forming ability (+++). Meanwhile, *M. gallisepticum* strain NX-05 showed medium biofilm-forming ability (++), and NX-04 and NX-02 were weak biofilm formers (+). The OD_{590} readings of the strong biofilm-forming strain NX-01 significantly differed from the medium and weak biofilm-forming strains at 12, 24, and 48 h of biofilm culture ($p < 0.01$). Additionally, at 36 h of biofilm culture, the OD_{590} readings of the *M. gallisepticum* strain NX-01 were significantly different from the other four strains ($p < 0.01$; Figure 3B). The results showed that NX-01 strain was the strongest strain and NX-02 strain was the weakest among the five MG strains. Therefore, these two strains were selected as the subsequent test bacteria.

The matrix growth curve test of the biofilm model *in vitro* showed that the biofilm formation of the tested MG was divided into the initial adhesion stage, aggregation stage, and biofilm maturation stage. The two tested strains NX-01 and NX-02 reached the initial adhesion stage after 12 h culture.

The growth rate of the NX-01 strain was the highest at 12–48 h, and the strain gathered to form microcolonies, corresponding to the aggregation stage of the biofilm. There was no significant change in the number of viable cells of stain at 48–72 h ($p > 0.05$), corresponding to the maturation stage of the biofilm, at which the microcolony continued to expand and formed a mature biofilm with a dense structure (Figure 3C).

The cells of the NX-02 strain adhered to each other and gathered to form microcolonies within 12–36 h, after which they rapidly multiplied and began to adhere to microplates. During the logarithmic phase of biofilm growth from 36 to 42 h, the formed strain microcolonies fused with each other and developed into biofilm with a mature three-dimensional structure. At 48–72 h, bacterial metabolic activity in the biofilm decreased (Figure 3D). The results showed that the membrane growth rate of strong membrane-forming MG strain NX-01 was faster than that of weak membrane-forming strain ($p < 0.01$), and the maintenance time of biofilm of NX-01 strain was longer than that of NX-02 strain ($p < 0.05$).

Scanning electron microscopy (SEM) was used to visually detect the morphology of the biofilm adhesion stage of NX-01 and NX-02 strains. Bacterial aggregates of NX-01 strain was detected across the entire visual field at varying magnifications (Figure 4A). They formed the mature biofilm, which is closely adhered together after being wrapped by the extracellular polysaccharide and other substances secreted by pathogenic bacteria. The three-dimensional mushroom-like structure can even be observed at 50,000× magnification (Figure 4E). The distribution of the NX-02 strain was relatively dispersed in the adhesion stage (Figure 4B), and its cell gathered in some areas forming tiny colonies and organized mature biofilms. The results showed that both NX-01 and NX-02 strains could form biofilms. The biofilm formed by NX-01 strain was multilayer and
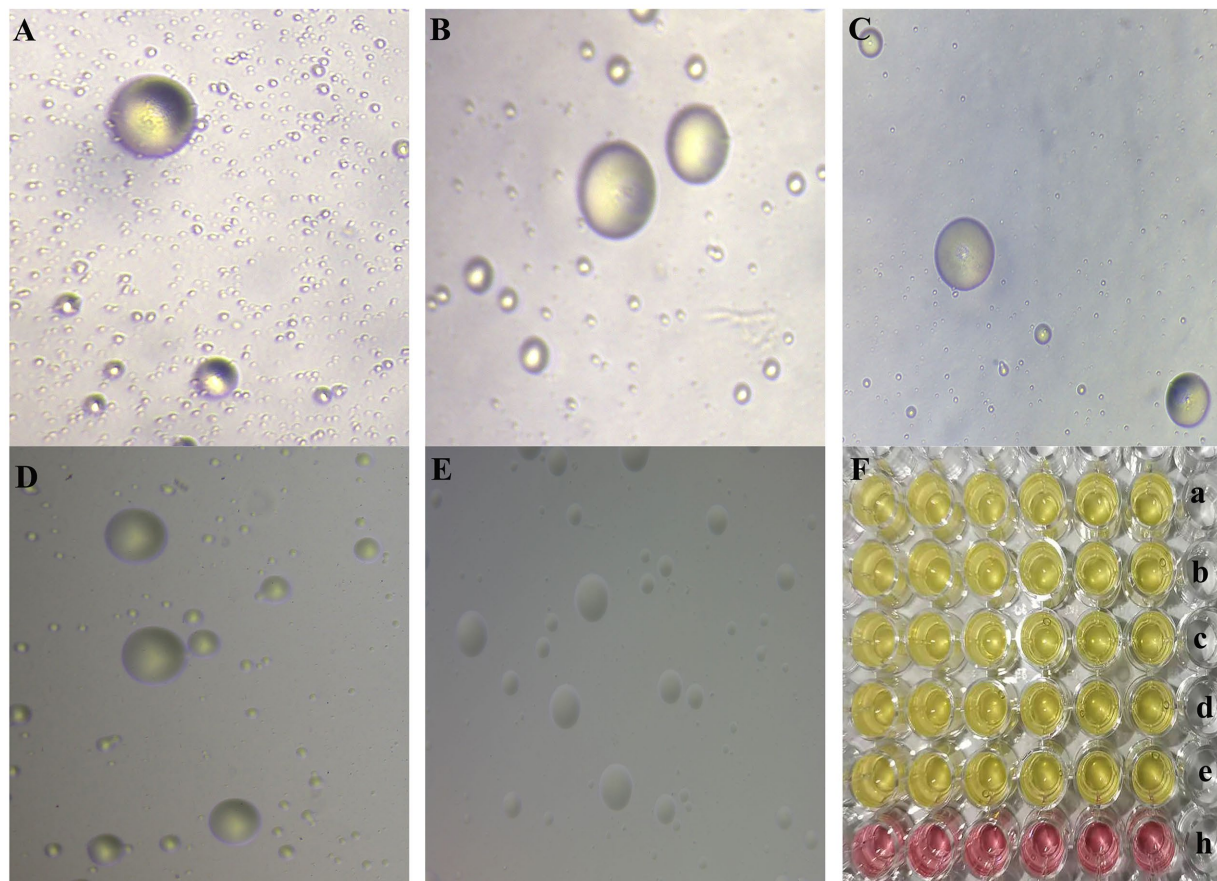
FIGURE 1
(A–E) The morphology of MG NX-01, NX-02, NX-03, NX-04, and NX-05 strains in modified solid medium under inverted microscope (10×), respectively. Consistent with the typical "fried egg" -like appearance of mycoplasma. **(F) a–e** was the color change of MG NX-01, NX-02, NX-03, NX-04, and NX-05 strains cultured in modified Frey liquid medium for 2–3 days, and **h** was the blank control group. Consistent with the characteristics of mycoplasma.

denser than that of NX-02 strain, which was consistent with the results of the biofilm formation ability test (Figures 4C,D).

3.2. Different drug sensitivity between NX-01 and NX-02

The sensitivity results of two MG stains isolated from Ningxia to 17 kinds of antibiotics showed that the order of sensitivity of NX-01 and NX-02 to six antibiotics was similar, which from strong to weak is tetracyclines, quinolones, macrolides, pleulopleuloides, aminoglycosides, and lincomycin (Table 4). The minimum inhibitory concentration of the NX-01 strain was 2 μg/mL, and that of the NX-02 strain was 0.25 μg/mL. The NX-02 strain was more sensitive to antibiotics than the NX-01 strain.

3.3. NX-01 and NX-02 harbors MG genomic charcteristics

DNA of NX-01 and NX-2 stains were, respectively, extracted and tested the integrity. All DNA sample for sequencing were not degrade in 0.7% nucleic acid gel electrophoresis.

The genome-wide of two MG strains was sequenced using the Illumina HiSeq platform, and the total bases count of NX-01 and NX-02 strains were 1,563,147,314 and 1,184,091,000 bp, and the genome coverage was above 1,000× (Table 5). Compare and evaluate the assembly results after the initial assembly using software SPAdes, NX-01, and NX-02 strains were assembled to obtain 40 and 33 nodes, respectively. The average length of assembled sequences of them is 24608.33 and 28679.33 bp. The maximum length of a single sequences of MG strains NX-01 and NX-02 strains is 147,070 and 17,76,641 bp, respectively. The N50 lengths of them are 67,869 and 101,937 bp. NX-01 and NX-02 strains has been deposited in NCBI (Bio project PRJNA972534 and PRJNA972540).

The genome-wide of the two MG strains showed that the genome size of NX-01 and NX-02 strains were 0.98 and 0.94 Mb, respectively (Table 6), and the GC content of their genome were 31.43 and 31.46%. The predicted CDS numbers in the genomes of NX-01 and NX-02 strains were 814 and 1,548 respectively, and predicted rRNAs were 2 and 3, respectively, and both had the tRNAs of 3.

Annotation results of NR database showed that more than 90% of the sequenced genes of the two isolates strains were derived from genome of reported MG, which proved that both of them were MG.
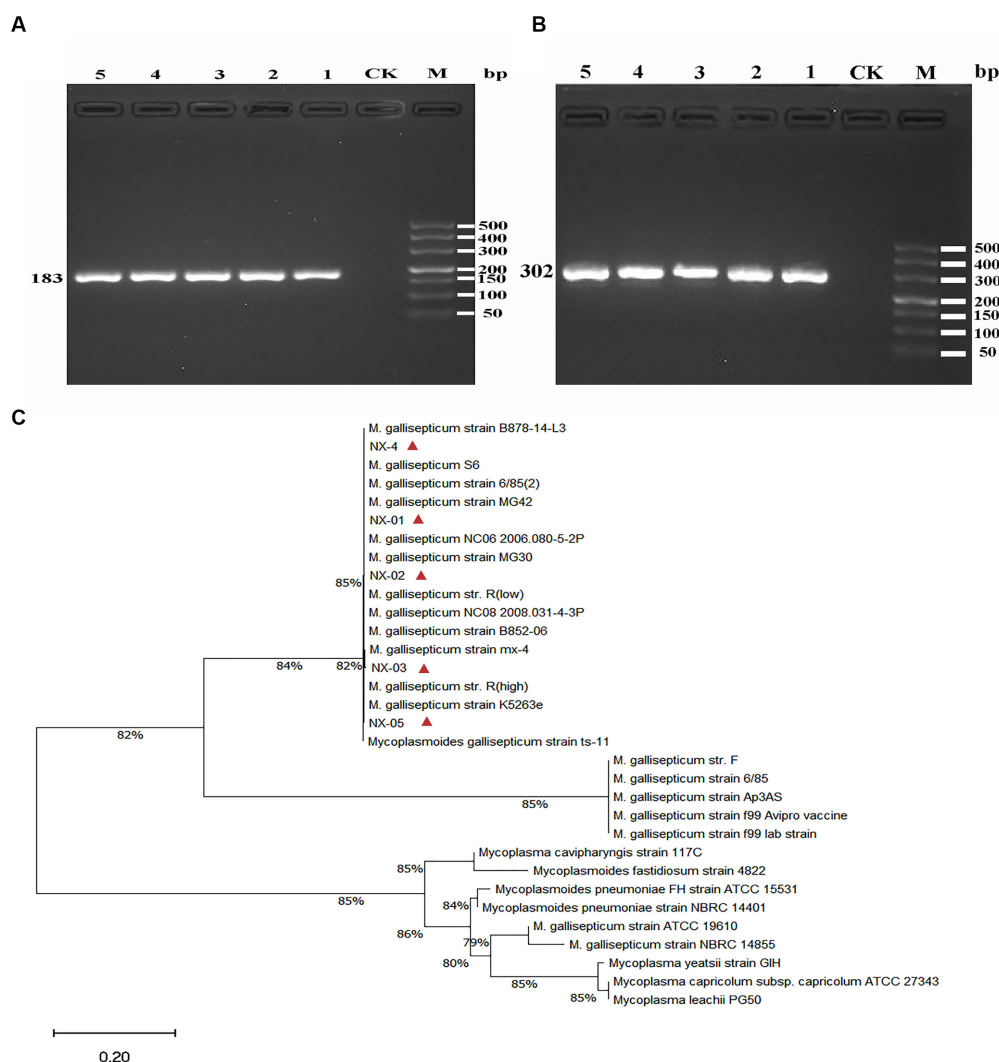
FIGURE 2
(A) The PCR results of MG 16S rRNA isolated from five strains. The fragment size was consistent with the expected size, indicating that the fragment was amplified correctly. **(B)** The PCR results of MG mgc2 isolated from five strains. The fragment size was consistent with the expected size, indicating that the fragment was amplified correctly. **(C)** The 16S rRNA phylogenetic tree of five MG isolates. The results showed that five MG strains were closely related to *Mycoplasma gallisepticum* uploaded by other NCBI. The above results indicated that the isolated strain was *Mycoplasma gallisepticum*.

3.4. Functional genes differ between NX-01 and NX-02 genomes

The functional annotation of COG genes showed (Figures 5A,B) that the number of functional genomes of NX-01 and NX-02 strains were 487 and 648, respectively. The predicted types of functional genes of these two MG strains were identical, whereas the number of certain functional genes differed significantly. The number of the two functional genes, including ribosome structure and biogenesis, and DNA replication and repair, was 40 less in the NX-01 strain than in the NX-02 strain. The number of cell wall or membrane or envelope-related genes of NX-01 strain was higher than that of NX-02 strain. The number of other functional genes of NX-01 was less or equal to that of NX-02. S-adenosylmethionine synthetase and methionyl tRNA synthetase were annotated in the genome-wide of NX-01 and NX-02 strains, both of which are biofilm-related enzymes. A biofilm-related protein, adhesin P1 precursor, was only noted in the genome-wide of the NX-01 strain.

GO database annotation showed (Figures 5C,D) that the functional genes of the two MG strains could be divided into biological process, molecular function, and cell component. NX-01 and NX-02 strains annotated 537 and 947 functional genes, respectively. The same gene may have multiple functions, the total number of genes of NX-01 and NX-02 strains involved in biological processes were 781 and 1,920, involved in molecular function were 1,127 and 1,431, involved in cell components were 1,133 and 1,785, respectively. The most annotated genes of two MG strains,in biological processes category belong to metabolic processes, including 367 genes in NX-01strain and 645 genes in NX-02 strain. The most annotated genes of two MG strains in the category of molecular functions belong to catalytic activity, including 327 genes in NX-01strain and 629 genes in NX-02 strain. The most annotated genes of two MG strains in the category of cell components belong to cells and cell parts, including 343 genes in
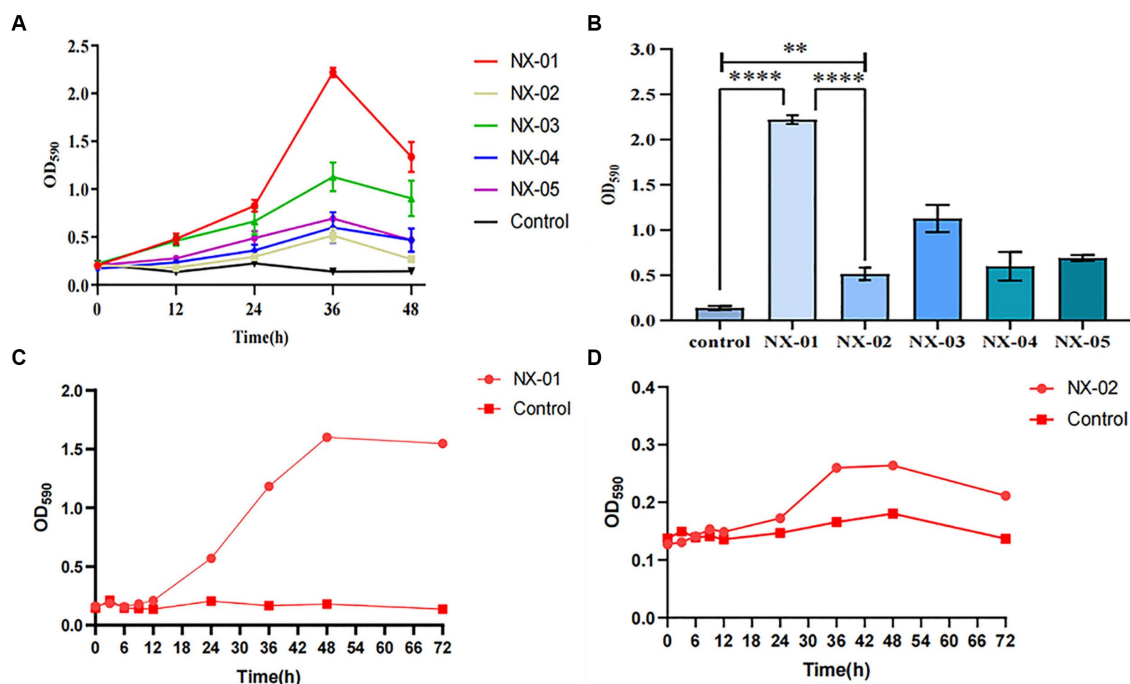
FIGURE 3
(A) Biofilm formation in different strains of MG. The data are means from three independent experiments. The *Mycoplasma gallisepticum* strains NX-01 and NX-03 were determined to be strong biofilm producers. The *M. gallisepticum* strain NX-05 was determined to be a medium biofilm producer. The *M. gallisepticum* strains NX-02 and NX-04 were determined to be weak biofilm producers. **(B)** At 36 h of biofilm culture, the OD$_{590}$ readings of the *M. gallisepticum* strain NX-01 were significantly different from the other four strains. **(C)** Biofilm growth curve of *M. gallisepticum* strain NX-01. **(D)** Biofilm growth curve of *M. gallisepticum* strain NX-02.

NX-01 strain and 554 in NX-02 strain. These results indicated that the number of functional genes of NX-01 was significantly different than that of NX-02.

The results of KEGG database annotation results showed that the genes of two MG strains could be divided into five branches, including cell process, environmental information processing, genetic information processing, metabolism, and organism system (Figures 6A,B). The gene number of the NX-01 strain was 302, and that of the NX-02 strain was 502. Metabolism accounted for the highest proportion of genes in the two strains. Metabolism-related genes of NX-01 strain account for 62.3% of the total genes, including energy metabolism, cofactor, and vitamin metabolism, while which of NX-02 strain lack glycan biosynthesis and metabolism genes. The number of metabolism-related genes in NX-02 strain accounted for 55.9% of the total genes. The genes related to genetic information processing accounted for 26.1% in NX-01 strain compared 27.5% in NX-02 strain. Cell process genes account for the least proportion of total genes, with only 0.3% of genes in NX-01 and 0.1% in NX-02. Five genes, *oppA*, *oppD*, *pdhA*, *eno*, and *msbA* were annotated in the environmental information processing pathway of two MG strains. Seven genes, *manB*, *pdhB*, *pdhC*, *pdhD*, *RelA*, *deoA*, and *gapA*, were annotated in the metabolic pathway. One genes, *ropS*, was annotated in the basic human pathway. All of these genes are biofilm-related gene, and among them, *PDH*, *eno*, and *gapA* participate in multiple pathways at the same time.

The annotation of the VFDB database showed that NX-01 strain has 18 species virulence factors while NX-02 has 22 species (Table 7). NX-01 strain is characterized by Cytadherence organelle, MntA

transporter, and RelA virulence factor. Seven toxic factors, including Clp-type ATPase chaperone protein, type III secretion system ATPase VscN, ATPase EscN, and K1 capsule, were unique to the NX-02 strain. Hemolysin factor was the most annotated MG factor in these two strains, which was cytotoxic to various cell types. From the number of noted virulence factors, it was found that the MG strain NX-02 had more virulence factors than the NX-01 strain, which may lead to stronger virulence of the MG strain NX-02 than NX-01.

The annotation of the CARD database showed seven kinds of drug-resistant genes in the *M. gallisepticum* strain NX-01 and 14 kinds in the MG strain NX-02 (Table 8). Quinolone resistance genes is the most frequent occurring drug resistance gene in two MG strains, which NX-01 strain had 20 quinolone resistance genes, and NX-02 strain had 28. Additional resistance genes were also detected, among which the immunogenic protein *EF-Tu* accounted for the largest proportion which range from 77.7% to 87.5%. Others genes including specific lincomycin resistance factor *lmrD*, macrolide resistance factor *MacB,* aminocoumarin resistance factor *alaS*, coumarin resistance factor *novA*, rifamycin resistance factor *rpoB*, and streptomycin resistance factor *rpsL,* occurs in the NX-02 strain. The results showed that NX-02 strain annotated more drug resistance factors than NX-01 strain, and they may have different degrees of resistance to antibiotics.

The annotation of the CAZy database showed (Figure 6C) that *M. gallisepticum* strain NX-01 was annotated to five carbohydrate enzymes (gene number 14), including glycoside hydrolase (GH), glycosyltransferase (GT), carbohydrate esterase (CE), helper activity enzyme (AA), and carbohydrate-binding module (CBM). NX-02 strain had one less carbohydrate-binding module (CBM) and six
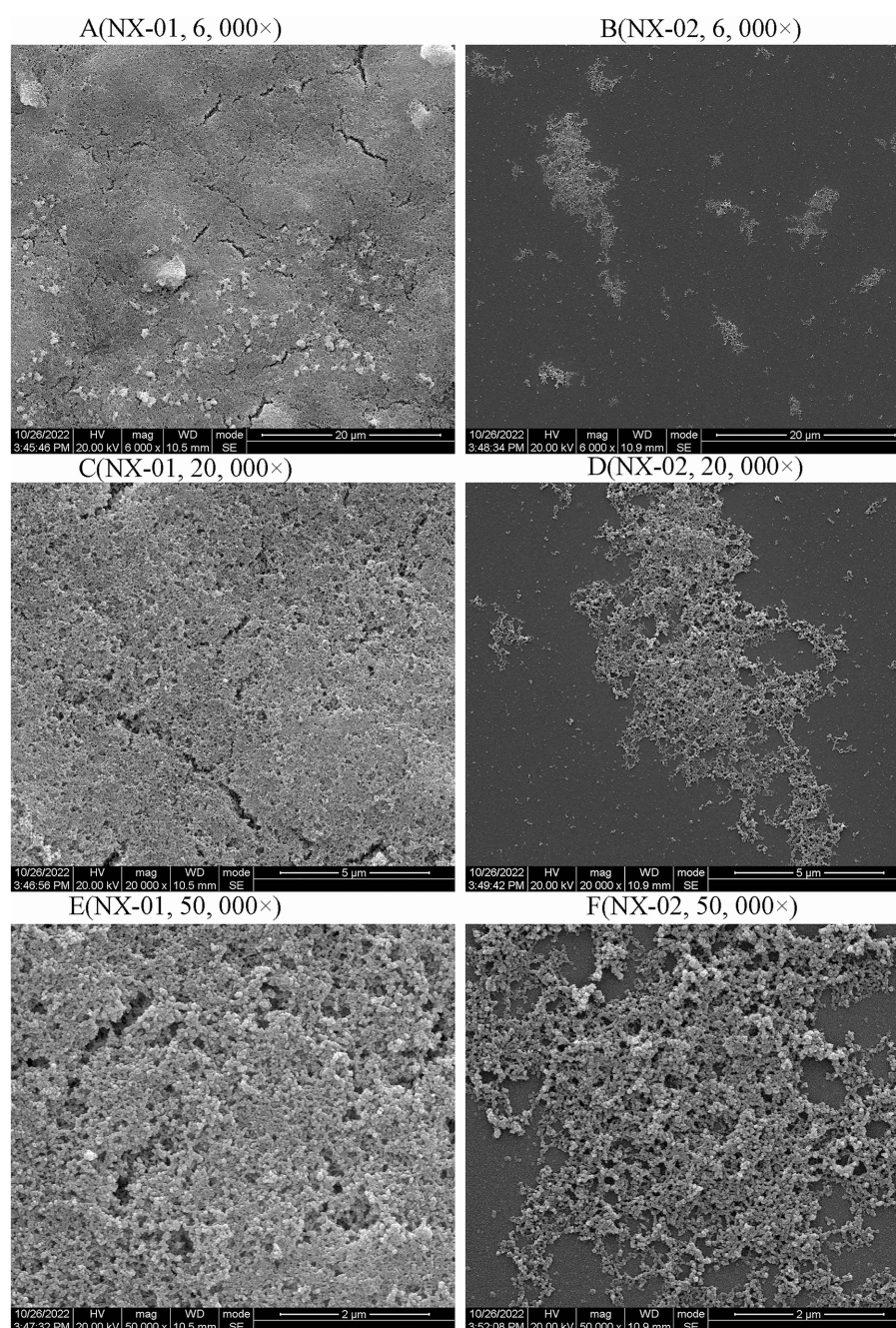
FIGURE 4
Scanning electron microscopy (SEM) analysis of the biofilm 0f NX-01 and NX-02 at gathering period. **(A)** MG NX-01(6,000×), **(B)** NX-02 (6,000×), **(C)** MG NX-01 (20,000×), **(D)** NX-02 (20,000×), **(E)** MG NX-01 (50,000×), and **(F)** NX-02 (50,000×).

genes. The number of related enzymes annotated by NX-01 strain was higher than that of the NX-02 strain, especially the carbohydrate-binding module, which may indicate different metabolic capacities.

3.5. NX-01 and NX-02 belongs to a divergent clade

A phylogenetic tree based on genome-wide construction is shown in Figure 7A. NX-01 and NX-02 strains had the closest homology and were in the same group. They were slightly distant from other MG strains, such as S6, ts-11, 6/85, R(low), NCTC10115, and so on, and were on different evolutionary branches. They were related to *M. gallisepticum* strains mx-4 and CA06_2006-052-5-2p, and furthest related to *M. gallisepticum* strains KUVMG01 and faa lab. The results showed that the 12 MG strains formed three large branches without clear distinctions based on geographic location. The two Ningxia MG isolates were closely related to *M. gallisepticum* Russian strain S6, suggesting that the two Ningxia MG isolates and MG strain S6 may derive from a common ancestor.

TABLE 4 The drug sensitivity of various kinds of drugs on *M.gallisepticum*.

Antibiotic		Drug sensitivity(μg / mL)	
		NX-01	NX-02
Quinolones	Ofloxacin	32	8
	Enrofloxacin	16	4
	Norfloxacin	64	32
Pleuromulin	Tiamulin	32	16
	Valnemulin	32	8
Macrolide	Erythromycin	32	8
	Tylosin	4	0.5
	Tilmicosin	16	2
	Spiramycin	128	64
Tetracycline	Oxytetracycline	2	0.5
	Tetracycline hydrochloride	2	0.25
	Doxycycline	16	2
Lincomycin	Lincomycin	128	128
Aminoglycosides	Gentamicin	4	4
	Spectinomycin	64	32
	Kanamycin	64	64
	Streptomycin	64	64

TABLE 5 Results of 2 MG genome assembly statistics.

Genome assembly statistics	NX-01	NX-02
Total bases count (bp)	1,563,147,314	1,184,091,000
Genome coverage	1,696	1,251
Node	40	33
Average length of assembled sequence (bp)	24,608.33	28,679.33
Maximum length of a single sequence (bp)	147,070	177,641
N50 (bp)	67,869	101,937

The genome-wide circle map of the seven strains of MG (Figure 7B) showed that the genome-wide of NX-01 strain was significantly different from that of the other six strains. *Mycoplasma gallisepticum* strain NX-01 had fewer gene deletions and had the least difference from *M. gallisepticum* Russian strain S6. The MG gene differed less between NX-02 and the other five strains were less than NX-01. Compared with NX-01 strain, NX-02 strain had little genetic difference in MG with five strains from other regions. In addition, NX-02 strain has less genome-wide differences from F and R (low) strains in the United States. Five major differences were found in the comparative genome of seven MG strains, which were 490–500, 525–540, 600–615, 782–792, and 882–892 kb fragment regions. The results indicated that the genetic information of MG isolates from Ningxia had rich regional specificity.

3.6. Evidence of differences of NX-01 and NX-02 with other MG epidemic strains

The two MG were from Ningxia were different from the five endemic *M. gallisepticum* strains [6/85, F, MX-4, R(low), and S6] in five regions of the world, and the two MG isolates were different from *M. gallisepticum* strain F, but little from *M. gallisepticum* strains R(low) and S6.

In the fragment region of 490–500 kb (Figure 8A), the NX-01 and NX-02 strains showed more long fragment deletion than the other five strains. In addition, NX-01 and NX-02 strains showed translocation of MG recombinase Polymerase *recA*, *rpmA*, and *nfo* compared with other five MG strains. The direction of NX-01 and NX-02 strains were consistent with S6, R(low), and 6/85 gene fragments. Compared with MX-4 and F strains, NX-01 and NX-02 strains had similar inversions of larger gene segments.

In the 520–540 kb fragment (Figure 8B), the two MG strains still had longer fragment deletion than the other five MG strains, and the same genes, such as *rsmH*, *scpB*, and *plsC*, were displaced. *Mycoplasma gallisepticum* NX-02 strain-specific *fssZ* can cause host immunogenicity. Two MG strains were in the same direction as *M. gallisepticum* strains S6, R(low), and 6/85, and the whole segment was inverted with *M. gallisepticum* strain MX-4 and F.

The 600–615 kb sequences were compared (Figure 8C). Besides removing longer fragments in two of the MG strains compared to the other five, the NX-01 strain is distinguished by the absence of *pepE*, *geoA*, and *pgcA* genes as opposed to the *M. gallisepticum* strain NX-02. Notably, the *pepE* was associated with the adhesion process of mycoplasma. *Mycoplasma gallisepticum* strain NX-02 had more *deoA*, *pgcA*, and *tpiA* genes than the other six strains. The two MG and *M. gallisepticum* strain F showed an inversion of the entire segment, and the direction of the gene segments of the other four MG strains was the same.

The 782–792 kb sequences were compared (Figure 8D), which revealed found that NX-01 strain had more deletion of MG fragments than the other six strains. NX-01 strain showed more ecfA1 gene than the other five MG isolates from different regions, and the similar genes *cdd* and *gpml* were displaced. The *ecfT*, *pgcA*, and *deoA* genes were unique to the NX-02 strain, and the similar gene *truA* was translocated. The direction of the segments of two MG strains was consistent with that of 6/85, R(low), and S6 strains, and the whole segment was inverted with F strain.

In the 882–892 kb difference region (Figure 8E), the direction of MG gene fragments in two strains was consistent with that in the other five strains. However, fragment deletion was the largest among the five different regions. The same genes *rplK*, *pth*, and *rpoE* of the two MG strains and the other five MG strains were shifted to different degrees. The length of the *tilS* in *M. gallisepticum* strain NX-02 was shorter and more translocated than in the other five MG strains.

4. Discussion

Due to the absence of cell walls and defense mechanisms, mycoplasma typically has a limited lifespan in the external environment (Justice-Allen et al., 2010). The formation of biofilms helps mycoplasma survive in the environment outside of its host (Beko et al., 2022). Thus far, only a limited number of studies have

demonstrated the ability of certain *Mycoplasma gallisepticum* strains to develop biofilms, and subsequent investigations in this area are needed (Perez et al., 2020). The present study measured the

TABLE 6 Basic information of the genome-wide of 2 MG isolates.

Characteristic	NX-01	NX-02
Bp total length	984, 333	946, 418
G + Content(%)	31.43	31.46
Gene number	814	1,548
CD number	870, 411	677, 124
rRNAs number	2	3
tRNAs number	33	33
Repeat region count	0	222
CRISPR	59	31

biofilm-forming ability of five strains of MG isolated from chicken farms in Ningxia, China. Our findings revealed that the MG strains could create biofilms. Yet, proficiency in doing so exhibited notable variations. NX-01 strain was found to have the strongest biofilm formation capacity, while NX-02 strain had the weakest biofilm formation ability.

At the same time, scanning electron microscopy also confirmed that *M. gallisepticum* strain NX-01 could form a more mature and dense multilayer biofilm. On the other hand, the MG strain NX-02 could only form flaky biofilm clusters in some areas.

Biofilms provide a "safe haven" for persistent bacteria to escape antibiotics, dramatically increasing antibiotic resistance (Cepas et al., 2018; Yan and Bassler, 2019). In the drug sensitivity test of two MG strains, the MG strain NX-01, which exhibited robust biofilm formation, displayed greater antibiotic resistance than the NX-02 strain, thus demonstrating weaker biofilm formation. In general, both strains exhibited poor drug sensitivity. Currently, there is no globally

FIGURE 5
Annotation statistics of the genomes of *Mycoplasma gallisepticum* strains NX-01 and NX-02 in databases. **(A)** Functional annotation results of COG database gene of *M. gallisepticum* strains NX-01. **(B)** Functional annotation results of COG database gene of *M. gallisepticum* strain NX-02. **(C)** GO database gene classification annotation results of *M. gallisepticum* strain NX-01. **(D)** GO database gene classification annotation results of *M. gallisepticum* strain NX-02.

FIGURE 6
(A) Annotation results of KEGG database gene function classification of *Mycoplasma gallisepticum* strain NX-01. (B) Annotation results of KEGG database gene function classification of *M. gallisepticum* strain NX-02. (C) The CAZy database gene annotation results of *M. gallisepticum* strains NX-01 and NX-02.

accepted criterion for antimicrobial resistance in MG. Nonetheless, our results showed that the two Ningxia MG isolates had a certain degree of insensitivity to antibiotics commonly used to treat mycoplasma, which was especially evident for the *M. gallisepticum* strain NX-01, probably due to its strong biofilm formation capacity.

Two MG strains were predicted to contain multiple quinolone-resistant genes, such as *Parc*, *ParE*, *gyrA*, and *gyrB*. The drug sensitivity test also verified that the two MG strains were not highly sensitive to norfloxacin, which may be related to the frequent use of quinolone antibiotics in treating MG infection. Two MG-specific *RecA* genes were found in the comparative genome, and they mainly mediated SOS response. The SOS response induces resistance to antibiotic agents by repairing DNA damage caused by antibiotic agents. Quinolones are effective inducers of SOS expression(Pribis et al., 2019). The drug sensitivity test also verified the insensitivity of the two MG strains to quinolones. The NX-02 strain exhibited more specific resistance genes, such as *lmrD*, *MacB*, and *alaS,* among others, than the NX-01 strain. De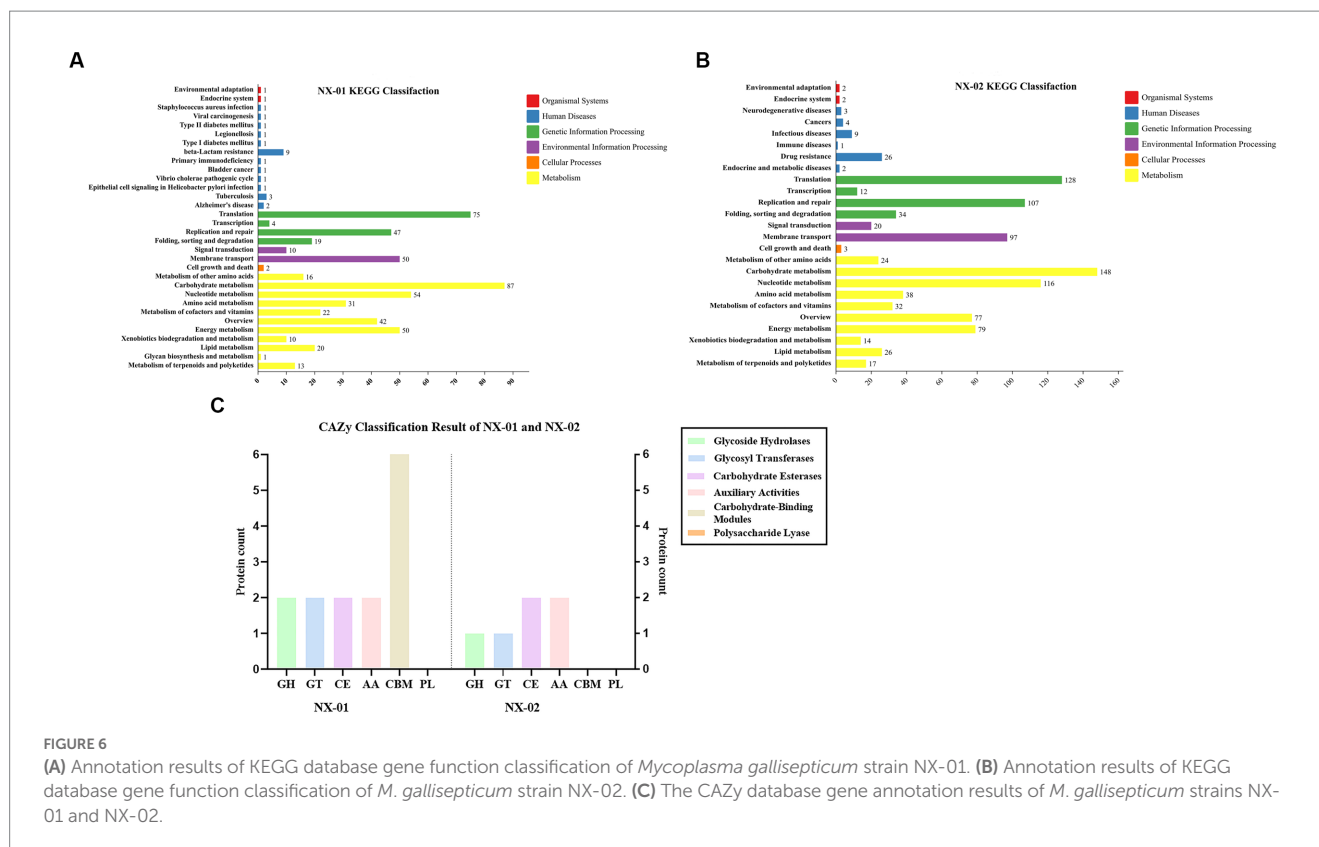spite this, the NX-02 strain demonstrated greater overall sensitivity in the drug sensitivity test than the NX-01 strain. It was observed that both biofilm and drug-resistance genes impacted the drug resistance of the bacteria. However, the mechanism underlying this phenomenon should be further explored.

By genome-wide analysis, 10 genes related to biofilm formation were predicted in both MG strains: *ManB*, *oppA*, *oppD*, *PDH*, *eno*, *RelA*, *msbA*, *deoA*, *gapA,* and *rpoS*. Phosphomanose mutase (*ManB*) is major in regulating extracellular polysaccharide synthesis of *Pseudomonas aeruginosa* biofilms, designed to maintain biofilm structure and antibiotic resistance (Ghafoor et al., 2011; Rachmawati et al., 2022). *oppA* and *oppD* belong to the ATP-dependent transporter family. They exert different roles in bacterial transport and have been

shown to influence biofilm formation (Lee et al., 2004; Li H. et al., 2020). *PdhA*, *pdhB*, *pdhC,* and *pdhD* were all predicted as genes encoding pyruvate dehydrogenase complex (*PDH*). *PDH* can convert pyruvate into acetyl coenzyme A, providing energy for bacterial growth and metabolism (Matic et al., 2003). Monica also found that *PDH* operons affect the formation of *Streptococcus mutans* biofilms (Busuioc et al., 2010). Vania reported that *pdhA* regulates the formation of *Staphylococcus epidermidis* biofilms (Gaio et al., 2021). Enolase encoded by eno is a key enzyme in the glycolysis pathway and a key gene in forming *Staphylococcus aureus* biofilm (Didiasova et al., 2019; Pant et al., 2022). The ribosome-related enzyme *RelA* has been shown to induce metabolic resistance in persistent biofilm cells, used to maintain bacterial biofilms and resist antibiotics (Hall et al., 2020). *MsbA* is a multidrug-resistant protein gene that can affect bacterial adhesion and biofilm formation (Brown et al., 2014; Padayatti et al., 2019). *DeoA* is one of the key enzymes used by deoxyribonucleoside. Previous studies have shown that the microcolony of Mycoplasma enters a state of energy starvation, where deoxyribonucleoside is preferred. It helps them survive under pressure (Fisunov et al., 2022). *GapA* is an essential protein for MG host cell adhesion and virulence and may be involved in biofilm formation (Goh et al., 1998; Ruger et al., 2022). Quorum-sensing phenomenon (QS) is interdependent with biofilm formation, and *rpoS* is considered the major positive modulator of QS response. Previous studies have shown that *rpoS* is closely related to the formation of multiple bacterial biofilms (Mika and Hengge, 2014; Hall et al., 2018; Zhang et al., 2021). However, the mechanism of MG biofilm reaction with QS remains unclear.

In addition, three proteins or enzymes associated with biofilm formation were predicted: Adhesin P1 precursor, S-adenosine methionine synthetase, and methionyl tRNA synthetase. In particular,

TABLE 7 Notes on the VFDB database of MG NX-01 and NX-02.

Virulence factors	Genes number	
	NX-01	NX-02
Capsule	2	2
Chu	0	2
ClpC	1	2
ClpE	1	2
Cya	2	4
Cytadherence organelle	1	0
Cytolysin	2	3
FbpABC	2	3
Hemolysin	10	8
HitABC	1	5
HSI-I	2	3
Hsp60	1	1
K1 capsule	0	1
LplA1	1	2
MgtBC	0	1
MntABC	1	0
MsrAB	1	2
Pyochelin	4	7
RelA	1	0
Shu	0	2
T3SS1	0	1
T6SS	1	2
T6SS-1	0	2
TTSS	0	3
Yersiniabactin	6	8

TABLE 8 Notes on the CARD database of MG NX-01 and NX-02.

Antibiotic type	Gene name	Number of genes	
		NX-01	NX-02
Quinolones	ParC	2	2
	parE	7	10
	gyrA	5	5
	gyrB	6	11
Lincomycin	lmrD	0	1
Macrolides	MacB	0	1
Aminocoumarin	alaS	0	1
Coumarins	novA	0	1
Rifamycins	rpoB	0	2
	rpoC	1	1
Streptomycins	rpsL	0	1
Others	EF-Tu	7	7
	MsbA	1	1
	dfrE	0	1

the Adhesin P1 precursor was a differential protein only predicted in the genome-wide NX-01 strain. Adhesin P1 precursor is related to adhesion protein P1, the most important membrane surface adhesion protein of *Mycoplasma pneumoniae*. It was found that adhesion protein P1 was involved in the biofilm formation of *Mycoplasma pneumoniae* and *Streptococcus mutans* (Kornspan et al., 2011; Tang et al., 2016). Therefore, the Adhesin P1 precursor might be significantly related to the distinct biofilm formation capacity of the two MG strains. However, as the genome-wide study was insufficient, further verification is necessary. In both strains of MG, the genome was analyzed for the presence of S-adenosine methionine synthetase and methionyl tRNA synthetase, where the former has a crucial role in protein methylation and various metabolic pathways in bacteria, making it a significant metabolite. It is also a precursor of QS, producing signal molecules related to inducing biofilm formation(Kleiner et al., 2019; Renard et al., 2021). Methionyl tRNA synthetase is necessary for bacteria to make proteins. Ulrike's research has revealed that this synthetase also has a crucial role in the modulation of virulence and biofilm formation in *Acinetobacter baumannii* (Blaschke et al., 2021).

The primary invasion of host cells by mycoplasma is adhesion, which also seems necessary for forming mycoplasma biofilms (Sachse et al., 1996). Genome-wide analysis also annotated two MG adhesion-related genes. For example, heat shock proteins Hsp60, *Clpc*, and *ClpE* are important virulence factors mediating cell apoptosis, adhesion, invasion, virulence, and reproduction (Garduño et al., 1998). NX-01 strain is characterized by Cytadherence organelle, which is related to Mycoplasma's cellular adherence and invasiveness (Romero-Arroyo et al., 1999). Among the specific virulence genes of the NX-02, TTSS, T3SS, T6SS system, and K1 Capsule had a key role in host bacterial infection (Francis et al., 2001; Schwarz et al., 2010). In addition, a specific non-catalytic carbohydrate-binding module (CBM) of the NX-01 strain was discovered from the CAZy database. CBMs attach to proteins on the surface of bacterial cells and act as carbohydrate-specific adhesins (Higgins et al., 2011). The regulatory mechanisms of MG biofilms and adhesins have not yet been fully explored; however, both were found to be closely associated with severe chronic mycoplasma infection.

The comparative genome showed that NX-01 and NX-02 strains were closely related to the Russian strain S6, which is far away from MG endemic strains in other areas. The results indicated that the two MG Ningxia isolates had regional specificity. The genome-wide of the NX-01 and NX-02 strains differed from each other. It is worth noting that NX-01 and S6 strains slightly differed, just as NX-02 with F and R(low) strains. Previous research has demonstrated that S6 strain exhibits robust biofilm formation capacity, whereas F and R(low) strains exhibit relatively poor biofilm formers. In this study, the genomic association analysis of NX-01 and S6 showed that the genomes of NX-01 and S6 strains were similar in size, 984,333 and 985,433 bp, respectively, and the CG content were also similar, 31.4 and 31.5%, with both strains containing 33 tRNAs. The P1 homologous adhesion gene was discovered in S6 strain (GenBank: U44804.1). The P1 homologous adhesin gene was discovered in strain S6 (GenBank: U44804.1), and the P1 homologous adhesin gene sequence of S6 strain was compared with the P1 adhesin precursor of NX-01 strain at NCBI Blast. The results showed that the Query Cover was 85%, the E-value was 0, and the per identity was 98.28%. This indicates that the gene sequences of the two MG strains have high homology. Only the strong biofilm formation NX-01 strain was found to contain the P1

FIGURE 7

(A) Phylogenetic tree based on different MG genome-wide sequences. NX-01 and NX-02 strains had the closest homology and were in the same group. It was furthest related to *Mycoplasma gallisepticum* strains KUVMG01 and faa lab. **(B)** Genome circle of seven MG strains. On laps 1, 2, and 3 (from inside out), the GC skew is black. Green is skew+ and purple is skew−. In circles 4−10 are the full genomes of seven MG strains, namely *M. gallisepticum* strains NX-01, NX-02, 6/85, F, MX-4, R(LOW), and S6.

FIGURE 8

(A) Differences of 490−500 kb in seven MG genomes. **(B)** Differences of 520−540 kb in seven MG genomes. **(C)** Differences of 600−615 kb in seven MG genomes. **(D)** Differences of 782−792 kb in seven MG genomes. **(E)** Differences of 882−892 kb in seven MG genomes.

mucoadhesive precursor, and it has been shown that the P1 mucoadhesive in *Mycoplasma pneumoniae* is associated with the formation of its biofilm.

Whether these findings are related to the ability of MG biofilm formation needs to be further investigated. We found five big differences between the two MG strains and the five MG strains from other areas, which contained homologous genes and differential genes. Also, some homologous genes were inverted, which may affect the biological characteristics of MG.

5. Conclusion

Mycoplasma gallisepticum predicted key genes that have been shown to regulate biofilm formation in other strains, and these genes may also have a regulatory role in MG biofilm formation. MG biofilm formation capacity has a direct influence on antibiotic drugs. We speculated that Adhesin P1 precursor might significantly differ from MG biofilm formation ability; however, further *in vivo* verification, such as gene knockout or use of transposon-mutated mycoplasma clones, is needed. In summary, our results provide a theoretical basis for the molecular mechanism of MG biofilm formation and a potential target for inhibiting the formation of MG biofilm.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/bioproject/; Two genomic information involved in this paper are PRJNA972534 and PRJNA972540.

Author contributions

SH and FY designed the research and analyzed data. XM and LW cultured *Mycoplasma gallisepticum* biofilm. LG and YG performed genome annotation. XM performed writing—original draft preparation. SH and JL wrote, reviewed, and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

LG was employed by Ningxia Xiaoming Agriculture and Animal Husbandry Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1196747/full#supplementary-material

References

Abd El-Hamid, M. I., Awad, N. F. S., Hashem, Y. M., Abdel-Rahman, M. A., Abdelaziz, A. M., Mohammed, I. A. A., et al. (2019). In vitro evaluation of various antimicrobials against field mycoplasma gallisepticum and *mycoplasma synoviae* isolates in Egypt. *Poult. Sci.* 98, 6281–6288. doi: 10.3382/ps/pez576

Ahmad, T., Ullah, S., Moosa, A., Liu, Y., Chengrong, N., and Shujian, H. (2020). First report of pre-harvest soft rot of peach fruit (*Prunus persica*) caused by *Enterobacter mori* in China. *Plant Dis.* doi: 10.1094/PDIS-06-20-1285-PDN

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Beaudet, J., Tulman, E. R., Pflaum, K., Canter, J. A., Silbart, L. K., and Geary, S. J. (2019). Immunologic pathways in protective versus maladaptive host responses to attenuated and pathogenic strains of *Mycoplasma gallisepticum*. *Infect. Immun.* 87:e00613. doi: 10.1128/IAI.00613-18

Beko, K., Nagy, E. Z., Grozner, D., Kreizinger, Z., and Gyuranecz, M. (2022). Biofilm formation and its impact on environmental survival and antibiotic resistance of Mycoplasma anserisalpingitidis strains. *Acta Vet. Hung.* doi: 10.1556/004.2022.00029

Beylefeld, A., Wambulawaye, P., Bwala, D. G., Gouws, J. J., Lukhele, O. M., Wandrag, D. B. R., et al. (2018). Evidence for multidrug resistance in nonpathogenic mycoplasma species isolated from south African poultry. *Appl. Environ. Microbiol.* 84:e01660-18. doi: 10.1128/AEM.01660-18

Bland, C., Ramsey, T. L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N. C., et al. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* 8:209. doi: 10.1186/1471-2105-8-209

Blaschke, U., Skiebe, E., and Wilharm, G. (2021). Novel genes required for surface-associated motility in *Acinetobacter baumannii*. *Curr. Microbiol.* 78, 1509–1528. doi: 10.1007/s00284-021-02407-x

Boetzer, M., and Pirovano, W. (2012). Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56. doi: 10.1186/gb-2012-13-6-r56

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Brown, N. A., Dos Reis, T. F., Goinski, A. B., Savoldi, M., Menino, J., Almeida, M. T., et al. (2014). The aspergillus nidulans signalling mucin MsbA regulates starvation responses, adhesion and affects cellulase secretion in response to environmental cues. *Mol. Microbiol.* doi: 10.1111/mmi.12820

Brown, J., Pirrung, M., and McCue, L. A. (2017). FQC dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* 33, 3137–3139. doi: 10.1093/bioinformatics/btx373

Busuioc, M., Buttaro, B. A., and Piggot, P. J. (2010). The pdh operon is expressed in a subpopulation of stationary-phase bacteria and is important for survival of sugar-starved *Streptococcus mutans*. *J. Bacteriol.* 192, 4395–4402. doi: 10.1128/JB.00574-10

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624. doi: 10.1038/ismej.2012.8

Cepas, V., Lopez, Y., Munoz, E., Rolo, D., Ardanuy, C., Marti, S., et al. (2018). Relationship between biofilm formation and antimicrobial resistance in gram-negative bacteria. *Microb. Drug Resist.* 25, 72–79. doi: 10.1089/mdr.2018.0027

Chen, H., Yu, S., Hu, M., Han, X., Chen, D., Qiu, X., et al. (2012). Identification of biofilm formation by *Mycoplasma gallisepticum*. *Vet. Microbiol.* 161, 96–103. doi: 10.1016/j.vetmic.2012.07.013

Chen, L., Zheng, D., Liu, B., Yang, J., and Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.*, 44, D694–D697. doi: 10.1093/nar/gkv1239

Delcher, A. L., Salzberg, S. L., and Phillippy, A. M. (2003). Using MUMmer to identify similar regions in large sequence sets. *Curr. Protoc. Bioinformatics* 10:Unit 10.3. doi: 10.1002/0471250953.bi1003s00

Didiasova, M., Schaefer, L., and Wygrecka, M. (2019). When place matters: shuttling of Enolase-1 across cellular compartments. *Front. Cell Dev. Biol.* 7:61. doi: 10.3389/fcell.2019.00061

Feberwee, A., de Wit, S., and Dijkman, R. (2021). Clinical expression, epidemiology, and monitoring of Mycoplasma gallisepticum and *Mycoplasma synoviae*: an update. *Avian Pathol.* 51, 2–18. doi: 10.1080/03079457.2021.1944605

Felice, V., Lupini, C., Mescolini, G., Silveira, F., Guerrini, A., Catelli, E., et al. (2020). Molecular detection and characterization of Mycoplasma gallisepticum and *Mycoplasma synoviae* strains in backyard poultry in Italy. *Poult. Sci.* 99, 719–724. doi: 10.1016/j.psj.2019.12.020

Feng, M., Burgess, A. C., Cuellar, R. R., Schwab, N. R., and Balish, M. F. (2021). Modelling persistent *Mycoplasma pneumoniae* biofilm infections in a submerged BEAS-2B bronchial epithelial tissue culture model. *J. Med. Microbiol.* 70. doi: 10.1099/jmm.0.001266

Feng, M., Schaff, A. C., and Balish, M. F. (2020). *Mycoplasma pneumoniae* biofilms grown in vitro: traits associated with persistence and cytotoxicity. *Microbiology* 166, 629–640. doi: 10.1099/mic.0.000928

Fisunov, G. Y., Pobeguts, O. V., Ladygina, V. G., Zubov, A. I., Galyamina, M. A., Kovalchuk, S. I., et al. (2022). Thymidine utilisation pathway is a novel phenotypic switch of *Mycoplasma hominis*. *J. Med. Microbiol.* 71:001468. doi: 10.1099/jmm.0.001468

Francis, M. S., Lloyd, S. A., and Wolf-Watz, H. (2001). The type III secretion chaperone LcrH co-operates with YopD to establish a negative, regulatory loop for control of Yop synthesis in *Yersinia pseudotuberculosis*. *Mol. Microbiol.* 42, 1075–1093. doi: 10.1046/j.1365-2958.2001.02702.x

Gaio, V., Lopes, N., Cerca, N., and Franca, A. (2021). codY and pdhA expression is induced in *Staphylococcus epidermidis* biofilm and planktonic populations with higher proportions of viable but non-culturable cells. *Front. Cell. Infect. Microbiol.* 11:771666. doi: 10.3389/fcimb.2021.771666

Garduño, R. A., Garduño, E., and Hoffman, P. S. (1998). Surface-associated hsp60 chaperonin of *Legionella pneumophila* mediates invasion in a HeLa cell model. *Infect. Immun.* 66, 4602–4610. doi: 10.1128/IAI.66.10.4602-4610.1998

Gautier-Bouchardon, A. V. (2018). Antimicrobial Resistance in Mycoplasma spp. *Microbiol. Spectr.* 6. doi: 10.1128/microbiolspec.ARBA-0030-2018

Ghafoor, A., Hay, I. D., and Rehm, B. H. (2011). Role of exopolysaccharides in *Pseudomonas aeruginosa* biofilm formation and architecture. *Appl. Environ. Microbiol.* 77, 5238–5246. doi: 10.1128/AEM.00637-11

Ghanem, M., Wang, L., Zhang, Y., Edwards, S., Lu, A., Ley, D., et al. (2017). Core genome multilocus sequence typing: a standardized approach for molecular typing of *Mycoplasma gallisepticum*. *J. Clin. Microbiol.* 56:e01145-17. doi: 10.1128/JCM.01145-17

Goh, M. S., Gorton, T. S., Forsyth, M. H., Troy, K. E., and Geary, S. J. (1998). Molecular and biochemical analysis of a 105 kDa *Mycoplasma gallisepticum* cytadhesin (GapA). *Microbiology* 144, 2971–2978.

Hall, C. W., Hinz, A. J., Gagnon, L. B., Zhang, L., Nadeau, J. P., Copeland, S., et al. (2018). *Pseudomonas aeruginosa* biofilm antibiotic resistance gene ndvB expression requires the RpoS stationary-phase sigma factor. *Appl. Environ. Microbiol.* 84:e02762-17. doi: 10.1128/AEM.02762-17

Hall, D. C. Jr., Król, J. E., Cahill, J. P., Ji, H. F., and Ehrlich, G. D. (2020). The development of a pipeline for the identification and validation of small-molecule RelA inhibitors for use as anti-biofilm drugs. *Microorganisms* 8:1310. doi: 10.3390/microorganisms8091310

Hannan, P. C. (2000). Guidelines and recommendations for antimicrobial minimum inhibitory concentration (MIC) testing against veterinary mycoplasma species. International research programme on comparative Mycoplasmology. *Vet. Res.* 31, 373–395. doi: 10.1051/vetres:2000100

Higgins, M. A., Ficko-Blean, E., Meloncelli, P. J., Lowary, T. L., and Boraston, A. B. (2011). The overall architecture and receptor binding of pneumococcal carbohydrate-antigen-hydrolyzing enzymes. *J. Mol. Biol.* 411, 1017–1036. doi: 10.1016/j.jmb.2011.06.035

Hu, Q., Han, X., Zhou, X., Ding, S., Ding, C., and Yu, S. (2010). Characterization of biofilm formation by *Riemerella anatipestifer*. *Vet. Microbiol.* 144, 429–436. doi: 10.1016/j.vetmic.2010.02.023

Ishfaq, M., Zhang, W., Ali Shah, S. W., Wu, Z., Wang, J., Ding, L., et al. (2020). The effect of *Mycoplasma gallisepticum* infection on energy metabolism in chicken lungs: through oxidative stress and inflammation. *Microb. Pathog.* 138:103848. doi: 10.1016/j.micpath.2019.103848

Jiang, R., Xiang, M., Chen, W., Zhang, P., Wu, X., Zhu, G., et al. (2021). Biofilm characteristics and transcriptomic analysis of *Haemophilus parasuis*. *Vet. Microbiol.* 258:109073. doi: 10.1016/j.vetmic.2021.109073

Justice-Allen, A., Trujillo, J., Corbett, R., Harding, R., Goodell, G., and Wilson, D. (2010). Survival and replication of Mycoplasma species in recycled bedding sand and association with mastitis on dairy farms in Utah. *J. Dairy Sci.* 93, 192–202. doi: 10.3168/jds.2009-2474

Kahya, S., Temelli, S., Eyigor, A., and Carli, K. T. (2010). Real-time PCR culture and serology for the diagnosis of *Mycoplasma gallisepticum* in chicken breeder flocks. *Vet. Microbiol.* 144, 319–324. doi: 10.1016/j.vetmic.2010.01.012

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Kleiner, D., Shmulevich, F., Zarivach, R., Shahar, A., Sharon, M., Ben-Nissan, G., et al. (2019). The interdimeric interface controls function and stability of Ureaplasma urealiticum methionine S-adenosyltransferase. *J. Mol. Biol.* 431, 4796–4816. doi: 10.1016/j.jmb.2019.09.003

Kornspan, J. D., Tarshis, M., and Rottem, S. (2011). Adhesion and biofilm formation of *Mycoplasma pneumoniae* on an abiotic surface. *Arch. Microbiol.* 193, 833–836. doi: 10.1007/s00203-011-0749-y

Lee, E. M., Ahn, S. H., Park, J. H., Lee, J. H., Ahn, S. C., and Kong, I. S. (2004). Identification of oligopeptide permease (Opp) gene cluster in Vibrio fluvialis and characterization of biofilm production by oppA knockout mutation. *FEMS Microbiol. Lett.* 240, 21–30. doi: 10.1016/j.femsle.2004.09.007

Li, L., Li, H., Shi, Y., Chai, A. L., Xie, X., and Li, B. (2020). First report of bacterial leaf spot of *cucurbita pepo* caused by *erwinia persicina* in China. *Plant Dis.* doi: 10.1094/PDIS-06-20-1241-PDN

Li, H., Qiao, Y., Du, D., Wang, J., and Ma, X. (2020). Deletion of the oligopeptide transporter Lmo2193 decreases the virulence of *Listeria monocytogenes*. *J. Vet. Sci.* 21:e88. doi: 10.4142/jvs.2020.21.e88

Liang, Z., Qi, Y., Guo, S., Hao, K., Zhao, M., and Guo, N. (2019). Effect of AgWPA nanoparticles on the inhibition of *Staphylococcus aureus* growth in biofilms. *Food Control* 100, 240–246. doi: 10.1016/j.foodcont.2019.01.030

Liu, Z., Li, H., Li, L., Ma, Q., Fang, Z., Wang, H., et al. (2022). Gene-trait matching analysis reveals putative genes involved in Bifidobacterium spp. biofilm formation. *Gene* 826:146449. doi: 10.1016/j.gene.2022.146449

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178

Marouf, S., Khalf, M. A., Alorabi, M., El-Shehawi, A. M., El-Tahan, A. M., El-Hack, M. E. A., et al. (2022). *Mycoplasma gallisepticum*: a devastating organism for the poultry industry in Egypt. *Poult. Sci.* 101:101658. doi: 10.1016/j.psj.2021.101658

Massouras, A., Hens, K., Gubelmann, C., Uplekar, S., Decouttere, F., Rougemont, J., et al. (2010). Primer-initiated sequence synthesis to detect and assemble structural variants. *Nat. Methods* 7, 485–486. doi: 10.1038/nmeth.f.308

Matic, J. N., Wilton, J. L., Towers, R. J., Scarman, A. L., Minion, F. C., Walker, M. J., et al. (2003). The pyruvate dehydrogenase complex of *Mycoplasma hyopneumoniae* contains a novel lipoyl domain arrangement. *Gene* 319, 99–106. doi: 10.1016/S0378-1119(03)00798-4

McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., et al. (2013). The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* 57, 3348–3357. doi: 10.1128/AAC.00419-13

McAuliffe, L., Ellis, R. J., Miles, K., Ayling, R. D., and Nicholas, R. A. J. (2006). Biofilm formation by mycoplasma species and its role in environmental persistence and survival. *Microbiology* 152, 913–922. doi: 10.1099/mic.0.28604-0

Mika, F., and Hengge, R. (2014). Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia coli*. *RNA Biol.* 11, 494–507. doi: 10.4161/rna.28867

Padayatti, P. S., Lee, S. C., Stanfield, R. L., Wen, P. C., Tajkhorshid, E., Wilson, I. A., et al. (2019). Structural insights into the lipid a transport pathway in MsbA. *Structure* 27, 1114–1123. doi: 10.1016/j.str.2019.04.007

Pant, N., Miranda-Hernandez, S., Rush, C., Warner, J., and Eisen, D. P. (2022). Effect of savirin in the prevention of biofilm-related *Staphylococcus aureus* prosthetic joint infection. *Front. Pharmacol.* 13:989417. doi: 10.3389/fphar.2022.989417

Parker, D., Soong, G., Planet, P., Brower, J., Ratner, A. J., and Prince, A. (2009). The NanA neuraminidase of *Streptococcus pneumoniae* is involved in biofilm formation. *Infect. Immun.* 77, 3722–3730. doi: 10.1128/IAI.00228-09

Perez, K., Mullen, N., Canter, J. A., Ley, D. H., and May, M. (2020). Phenotypic diversity in an emerging mycoplasmal disease. *Microb. Pathog.* 138:103798. doi: 10.1016/j.micpath.2019.103798

Pribis, J. P., Garcia-Villada, L., Zhai, Y., Lewin-Epstein, O., Wang, A. Z., Liu, J., et al. (2019). Gamblers: an antibiotic-induced evolvable cell subpopulation differentiated by reactive-oxygen-induced general stress response. *Mol. Cell* 74:e7. doi: 10.1016/j.molcel.2019.02.037

Rachmawati, D., Fahmi, M. Z., Abdjan, M. I., Wasito, E. B., Siswanto, I., Mazlan, N., et al. (2022). In vitro assessment on designing novel antibiofilms of *pseudomonas aeruginosa* using a computational approach. *Molecules* 27:8935. doi: 10.3390/molecules27248935

Renard, A., Diene, S. M., Courtier-Martinez, L., Gaillard, J. B., Gbaguidi-Haore, H., Mereghetti, L., et al. (2021). 12/111phiA prophage domestication is associated with autoaggregation and increased ability to produce biofilm in *Streptococcus agalactiae*. *Microorganisms* 9:1112. doi: 10.3390/microorganisms9061112

Robinson, J. C., Rostami, N., Casement, J., Vollmer, W., Rickard, A. H., and Jakubovics, N. S. (2018). ArcR modulates biofilm formation in the dental plaque colonizer *Streptococcus gordonii*. *Mol Oral Microbiol* 33, 143–154. doi: 10.1111/omi.12207

Romero-Arroyo, C. E., Jordan, J., Peacock, S. J., Willby, M. J., Farmer, M. A., and Krause, D. C. (1999). *Mycoplasma pneumoniae* protein P30 is required for cytadherence and associated with proper cell development. *J. Bacteriol.* 181, 1079–1087. doi: 10.1128/JB.181.4.1079-1087.1999

Roussan, D. A., Khawaldeh, G., and Shaheen, I. A. (2015). A survey of Mycoplasma gallisepticum and Mycoplasma synovaie with avian influenza H9 subtype in meat-type chicken in Jordan between 2011-2015. *Poult. Sci.* 94, 1499–1503. doi: 10.3382/ps/pev119

Ruger, N., Szostak, M. P., and Rautenschlein, S. (2022). The expression of GapA and CrmA correlates with the *Mycoplasma gallisepticum* in vitro infection process in chicken TOCs. *Vet. Res.* 53:66. doi: 10.1186/s13567-022-01085-2

Sachse, K., Grajetzki, C., Rosengarten, R., Hänel, I., Heller, M., and Pfützner, H. (1996). Mechanisms and factors involved in *Mycoplasma bovis* adhesion to host cells. *Zentralbl Bakteriol* 284, 80–92. doi: 10.1016/S0934-8840(96)80157-5

Saha, S., Bridges, S., Magbanua, Z. V., and Peterson, D. G. (2008). Empirical comparison of ab initio repeat finding programs. *Nucleic Acids Res.* 36, 2284–2294. doi: 10.1093/nar/gkn064

Schwarz, S., West, T. E., Boyer, F., Chiang, W. C., Carl, M. A., Hood, R. D., et al. (2010). Burkholderia type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. *PLoS Pathog.* 6:e1001068. doi: 10.1371/journal.ppat.1001068

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Semashko, T. A., Arzamasov, A. A., Evsyutina, D. V., Garanina, I. A., Matyushkina, D. S., Ladygina, V. G., et al. (2022). Role of DNA modifications in *Mycoplasma gallisepticum*. *PLoS One* 17:e0277819. doi: 10.1371/journal.pone.0277819

Sharma, D., Misba, L., and Khan, A. U. (2019). Antibiotics versus biofilm: an emerging battleground in microbial communities. *Antimicrob. Resist. Infect. Control* 8:76. doi: 10.1186/s13756-019-0533-3

Shatila, F., Yasa, I., and Yalcin, H. T. (2021). Biofilm formation by *Salmonella enterica* strains. *Curr. Microbiol.* 78, 1150–1158. doi: 10.1007/s00284-021-02373-4

Sorci, G. (2013). Immunity, resistance and tolerance in bird-parasite interactions. *Parasite Immunol.* 35, 350–361. doi: 10.1111/pim.12047

Sprygin, A. V., Andreychuk, D. B., Kolotilov, A. N., Volkov, M. S., Runina, I. A., Mudrak, N. S., et al. (2010). Development of a duplex real-time TaqMan PCR assay with an internal control for the detection of *Mycoplasma gallisepticum* and *Mycoplasma synoviae* in clinical samples from commercial and backyard poultry. *Avian Pathol.* 39, 99–109. doi: 10.1080/03079451003604621

Stewart, P. S. (2002). Mechanisms of antibiotic resistance in bacterial biofilms. *Int. J. Med. Microbiol.* 292, 107–113. doi: 10.1078/1438-4221-00196

Sulyok, K. M., Kreizinger, Z., Beko, K., Forro, B., Marton, S., Banyai, K., et al. (2019). Development of molecular methods for rapid differentiation of *mycoplasma gallisepticum* vaccine strains from field isolates. *J. Clin. Microbiol.* 57:e01084-18. doi: 10.1128/JCM.01084-18

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tang, W., Bhatt, A., Smith, A. N., Crowley, P. J., Brady, L. J., and Long, J. R. (2016). Specific binding of a naturally occurring amyloidogenic fragment of *Streptococcus mutans* adhesin P1 to intact P1 on the cell surface characterized by solid state NMR spectroscopy. *J. Biomol. NMR* 64, 153–164. doi: 10.1007/s10858-016-0017-1

Tassew, D. D., Mechesso, A. F., Park, N. H., Song, J. B., Shur, J. W., and Park, S. C. (2017). Biofilm formation and determination of minimum biofilm eradication concentration of antibiotics in *Mycoplasma hyopneumoniae*. *J. Vet. Med. Sci.* 79, 1716–1720. doi: 10.1292/jvms.17-0279

Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33

Taylor-Robinson, D., and Bébéar, C. (1997). Antibiotic susceptibilities of mycoplasmas and treatment of mycoplasmal infections. *J. Antimicrob. Chemother.* 40, 622–630. doi: 10.1093/jac/40.5.622

The Gene Ontology (2019). The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47, D330–D338. doi: 10.1093/nar/gky1055

Wang, Y., Yi, L., Zhang, F., Qiu, X., Tan, L., Yu, S., et al. (2017). Identification of genes involved in *Mycoplasma gallisepticum* biofilm formation using mini-Tn4001-SGM transposon mutagenesis. *Vet. Microbiol.* 198, 17–22. doi: 10.1016/j.vetmic.2016.11.021

Yadav, J. P., Singh, Y., Jindal, N., and Mahajan, N. K. (2022). Rapid and specific detection of Mycoplasma gallisepticum and *Mycoplasma synoviae* infection in poultry using single and duplex PCR assays. *J. Microbiol. Methods* 192:106365. doi: 10.1016/j.mimet.2021.106365

Yan, J., and Bassler, B. L. (2019). Surviving as a community: antibiotic tolerance and persistence in bacterial biofilms. *Cell Host Microbe* 26, 15–21. doi: 10.1016/j.chom.2019.06.002

Yu, K., and Zhang, T. (2013). Construction of customized sub-databases from NCBI-nr database for rapid annotation of huge metagenomic datasets using a combined BLAST and MEGAN approach. *PLoS One* 8:e59831. doi: 10.1371/journal.pone.0083943

Zhang, C., Wang, C., Jatt, A. N., Liu, H., and Liu, Y. (2021). Role of RpoS in stress resistance, biofilm formation and quorum sensing of *Shewanella baltica*. *Lett. Appl. Microbiol.* 72, 307–315. doi: 10.1111/lam.13424

Check for updates

The core bacteriobiome of Côte d'Ivoire soils across three vegetation zones

Chiguié Estelle Raïssa Amon[1]*, Romain Kouakou Fossou[1], Anicet E. T. Ebou[1], Dominiqueua K. Koua[1], Claude Ghislaine Kouadjo[2], Yao Casimir Brou[1], Don Rodrigue Rosin Voko Bi[3], Don A. Cowan[4] and Adolphe Zézé[1]

[1]Laboratoire de Biotechnologies Végétale et Microbienne, UMRI Sciences Agronomiques et Génie rural, Institut National Polytechnique Félix Houphouët-Boigny, Yamoussoukro, Côte d'Ivoire, [2]Laboratoire Central de Biotechnologies, Centre National de la Recherche Agronomique, Abidjan, Côte d'Ivoire, [3]Unité de Formation et de Recherche en Agroforesterie, Université Jean Lorougnon Guédé, Daloa, Côte d'Ivoire, [4]Centre for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa

The growing understanding that soil bacteria play a critical role in ecosystem servicing has led to a number of large-scale biogeographical surveys of soil microbial diversity. However, most of such studies have focused on northern hemisphere regions and little is known of either the detailed structure or function of soil microbiomes of sub-Saharan African countries. In this paper, we report the use of high-throughput amplicon sequencing analyses to investigate the biogeography of soil bacteria in soils of Côte d'Ivoire. 45 surface soil samples were collected from Côte d'Ivoire, representing all major biomes, and bacterial community composition was assessed by targeting the V4-V5 hypervariable region of the 16S ribosomal RNA gene. Causative relationships of both soil physicochemical properties and climatic data on bacterial community structure were infered. 48 phyla, 92 classes, 152 orders, 356 families, and 1,234 genera of bacteria were identified. The core bacteriobiome consisted of 10 genera ranked in the following order of total abundance: *Gp6*, *Gaiella*, *Spartobacteria_genera_incertae_sedis*, *WPS-1_genera_incertae_sedis*, *Gp4*, *Rhodoplanes*, *Pseudorhodoplanes*, *Bradyrhizobium*, *Subdivision3_genera_incertae_sedis*, and *Gp3*. Some of these genera, including *Gp4* and *WPS-1_genera_incertae_sedis*, were unequally distributed between forest and savannah areas while other taxa (*Bradyrhizobium* and *Rhodoplanes)* were consistently found in all biomes. The distribution of the core genera, together with the 10 major phyla, was influenced by several environmental factors, including latitude, pH, Al and K. The main pattern of distribution that was observed for the core bacteriobiome was the vegetation-independent distribution scheme. In terms of predicted functions, all core bacterial taxa were involved in assimilatory sulfate reduction, while atmospheric dinitrogen (N_2) reduction was only associated with the genus *Bradyrhizobium*. This work, which is one of the first such study to be undertaken at this scale in Côte d'Ivoire, provides insights into the distribution of bacterial taxa in Côte d'Ivoire soils, and the findings may serve as biological indicator for land management in Côte d'Ivoire.

1. Introduction

Soils harbour complex microbial communities (Schloss and Handelsman, 2005; Pivato et al., 2015) with the capacity for diverse biogeochemical functions (Lavelle and Spain, 2001; Fenchel et al., 2012). Soil microbiomes provide an interface between soil nutrients and plant nutrient-uptaked mechanisms, and therefore represent key factors in soil health (Sindhu et al., 2019), nutrient cycling and plant health (Fierer, 2017; Jacoby et al., 2017). Soil microbiomes may therefore act as regulators of plant productivity, especially in nutrient-poor ecosystems where plant symbionts are responsible for the uptake of limiting nutrients (Quoreshi, 2008; Sindhu et al., 2019).

A number of studies have investigated microbial community compositions, and the impact of environmental factors on microbiome composition, on a national, continental or global scale (Fierer and Jackson, 2006; Delgado-Baquerizo et al., 2018; Karimi et al., 2018; Gnangui et al., 2021; Cowan et al., 2022). It has been shown that different environmental factors lead to different niches and the specific selection of a core microbiome (Turnbaugh et al., 2007; Liu et al., 2018; Pajares et al., 2018; Koutika et al., 2020). Its influence on soil health, plant growth and agricultural sustainability has been assessed in different ecosystems (Zhao et al., 2019; Castellano-Hinojosa and Strauss, 2021).

Several studies using high-throughput sequencing facilities to assess the functions and drivers of the core microbiome in soils have been conducted at small scale and at local levels (Pershina et al., 2018; Jiao et al., 2022; Kolton et al., 2022). Recent studies of African soil microbiomics have attempted to map both the microbial diversity and functional capacity of soil microbial communities at a local (Maquia et al., 2020; Gnangui et al., 2021) or continental scales (Cowan et al., 2022). However, the authors are unaware of any biogeographical studies focusing on the causative relationships that may exist between environmental factors and the core soil microbiome structure in any sub-Saharan nation.

Côte d'Ivoire can be divided into three main agro-ecological biomes (Ducroquet et al., 2017). The Guinean biome is located in the southern part of the country and represents more than 48% of the national territory (Ducroquet et al., 2017). This is the highest rainfall agro-ecological region and includes the forest zone (FAO, 2009). The climate has a equatorial/subequatorial type with a long rainy season that supports evergreen forest and semi-deciduous humid forest types (Guillaumet and Adjanohoun, 1971; Kassin et al., 2017). The main crop types grown in this area are cash (cocoa, coffee) and food crops (cassava, rice, plantain) (Ducroquet et al., 2017). Here, the pH was often more acidic in the top surface, with a sandy-loam texture (Tondoh et al., 2015; Kassin et al., 2017). At higher latitudes, the soils contain variable and often significant proportions of basic cations (Ca^{2+}, Mg^{2+}, K^+, and Na^+), as well as high amounts of nitrogen, leading to an increase in the rate of organic matters decomposition (Avenard et al., 1971).

The Forest-Savannah transition zone, which corresponds to the Sudano-Guinean biome, is centrally located and represents 19% of the nation's land mass (Ducroquet et al., 2017). This region is characterized by a humid tropical climate within which a transition takes place between the southern forest zone and the north, dominated by the savannah is the transition zone is characterized by a relatively low rainfall and a mixed landscape where the two types of vegetation coexist, termed the forest-savannah mosaic (Gautier, 1990).

The Sudanian biome, located in the north, occupies about 33% of the country's land mass (Ducroquet et al., 2017). This biome is characterized by both a subhumid and semi-arid tropical climate with a longer dry season (Guillaumet and Adjanohoun, 1971; Tano et al., 2020). Nitrogen content have been found to be relatively low (Assémien et al., 2017) with soil pH values varying between 5.7 and 6.4 (Tano et al., 2020).

Despite recent studies on microbiome composition in Côte d'Ivoire soils (Assémien et al., 2017; Kouadio et al., 2017; Séry et al., 2018; Gnangui et al., 2021), nothing is known of the core microbiome distribution across the nation, nor of the putative functions and ecosystem services of these microbiome. The main objectives of this work were therefore: (i) to examine the distribution of bacterial taxa across Côte d'Ivoire soils by high-throughput 16S ribosomal RNA (rRNA) gene sequencing; (ii) to identify the principal environmental factors that may influence the distribution of bacterial communities, at the phylum and genus levels, and (iii) to predict the potential metabolic profiles of the core bacteriobiome across Côte d'Ivoire soils.

2. Methods

2.1. Experimental design

Forty-five (45) surface soil samples were recovered across all the major biomes in Côte d'Ivoire from August to September 2017 (Gnangui et al., 2021; Cowan et al., 2022; Figure 1). Twenty-three (23) soil samples were collected in the forest zone, which represent more than 48% of the national territory, while 17 and 5 samples were

FIGURE 1
Map of Côte d'Ivoire showing the 45 sampled sites. Geographic coordinates of the sampled sites are given in Supplementary Table S1. The map is colored according to the main vegetation types in Côte d'Ivoire (forest, savannah and forest-savannah transition zones).
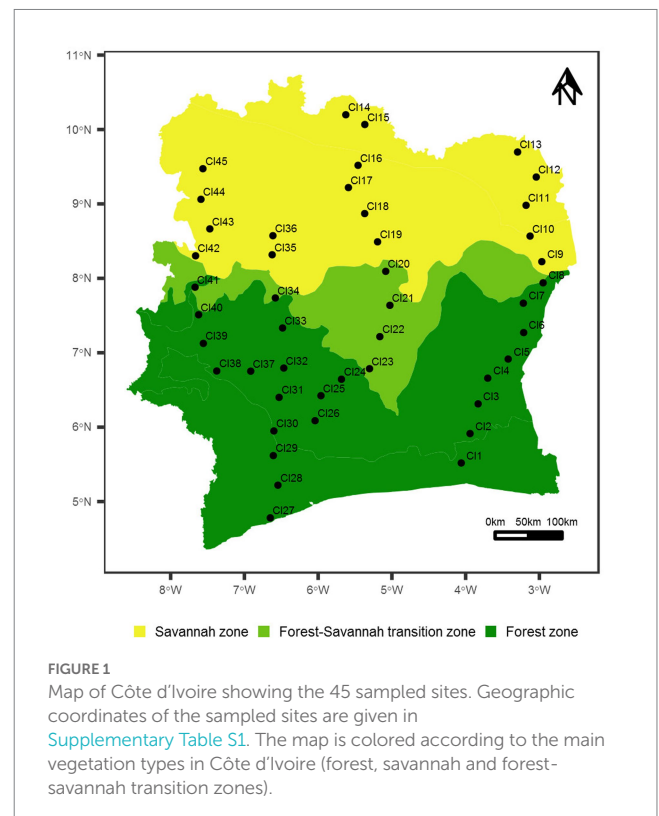
collected in the savannah and the forest-savanna transition zones, respectively. Each sampling point was represented as a virtual rectangle of 100×50 meters, at the corners of which soil subsamples (0–5 cm depth: approx. 25 g each) were recovered into sterile plastic Whirlpak® bags. Geographical coordinates as well as climatic and environmental data of the study locations were recorded (Supplementary Table S1). Soils were sub-divided for chemical analysis (stored at $-4°C$) and eDNA extraction (stored at $-80°C$).

2.2. Soil physico-chemical analyses

Soil physico-chemical characteristics of the soil samples were determined by Bemlab (Strand, Cape Province, South Africa) using standard methods (Gnangui et al., 2021; Cowan et al., 2022). Prior to analyses, the samples were air-dried at room temperature for 4 days, separated from roots and debris, and passed through a 2 mm sieve. The sieved replicate samples of each sampling site were pooled to obtain a composite soil sample. All analyses were performed using the composite soil samples as described elsewhere (Gnangui et al., 2021; Cowan et al., 2022).

2.3. Total DNA extraction, PCR and high throughput sequencing

Total DNA extraction, amplification and Illumina MiSeq amplicon sequencing were carried out as previously reported (Cowan et al., 2022). After DNA extraction at the Centre for Microbial Ecology and Genomics (University of Pretoria, South Africa), DNA sequencing was done at the MRDNA sequencing facility (www.mrdnalab.com, TX, United States). The V4–V5 variable region of the 16S rRNA gene was amplified and sequenced using primers: 515F-Y (5′-GTGYCAGCMGCCGCGGTAA-3′; Parada et al., 2016) and 909–928R (5′-CCCCGYCAATTCMTTTRAGT-3′; Wang and Qian, 2009), with 12 nucleotides unique barcode at 5-end of 515F-Y for each soil sample.

## 2.4. Bioinformatics and statistical analysis for microbial community analyses

The sequences obtained from the Illumina MiSeq platform were analyzed using Qiime 2 v2018.6 (Bolyen et al., 2019). Sequences were initially demultiplexed according to sample barcodes. Sequence processing was done as in Cowan et al., with some modifications (Cowan et al., 2022). A pre-processing step was performed to remove samples barcodes, followed by a filtering performed using the cutadapt v2.10 function under Linux (Martin, 2011) to delete forward primers as well as incomplete sequences. The resulting fastq file was then separated into several files based on the sequences size. A second filtering step was then done on each fastq file via DADA2 algorithm under Qiime 2, allowing the removal of low quality sequences, as well as reverse primers (Callahan et al., 2016; Christensen et al., 2018). The denoised sequences were clustered into Amplicon Sequences Variant (ASVs) with DADA2 (Callahan et al., 2016). The resulting ASVs were classified using the naive Bayesian Ribosomal Database Project (Wang et al., 2007; Cole et al., 2014) with the latest RDP classifier release 2.13

(July 2020[1]). The resulting taxonomic table files were then merged into one file for further analyses. A last filtering has been done on this file to remove 16S sequences affiliated to Archaea and chloroplasts.

All statistical analyses were performed in R v4.0.3 (R Core Team, 2020) with raster v3.5–15 (Hijmans, 2020), vegan v2.5–7 (Oksanen et al., 2020), phyloseq v1.34.0 (McMurdie and Holmes, 2013) and ggplot2 v3.3.6 (Wickham, 2016). Sequences were deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) database under BioProject PRJNA695288.

Soils texture classes were determined according to the standard USDA particle-size classification as in the soil texture R package (Moeys, 2018). Rarefaction curves were produced using the rarecurve function from the vegan package. The Kruskal-Wallis test was used to assess the statistical differences among the three type of vegetation for each of the variables considered in the study, followed by a Wilcoxon rank sum test (with a Benjamini-Hochberg correction) when there were significant differences. Based on their relative abundances, phyla were classified according to a bottom-up hierarchical clustering using Ward's method (Murtagh and Legendre, 2014). Comparisons following Wilcoxon rank sum test among phyla identified those that were abundant (major taxa) with more than 1% of relative abundance or least represented (minor and rare taxa, less than 1% of relative abundance) (Karimi et al., 2018). Raster package was used to map major phyla across Côte d'Ivoire soils, based on phylum abundance. The distribution of the dominant bacterial taxa across the three biomes was determined using the analysis of similarity (ANOSIM) and a variance partitioning performed on physico-chemical, climatic, spatial and land use variables as described elsewhere (Karimi et al., 2018). To assess the contribution of physico-chemical, climatic and spatial variables in the distribution of each core phylum in soils, a redundancy analysis (RDA) was performed as done elsewhere (Shen et al., 2018; Alami et al., 2020). On the total of 16 variables considered, four (i.e., Ca, Mg, N and sand) were removed to avoid collinearity. For each phylum, the most relevant variables were selected by the RDA, with an indication about their adjusted R^2 and p-values. The statistical significance of the results obtained was tested by an analysis of variance (ANOVA) with 999 permutations. Coefficient correlations between each core phylum and significant variables chosen by the RDA were obtained through Hmsic package (Harrell Jr, 2021). The resulting dataset, including adjusted R^2, p-values and correlation coefficients for each core phylum was used to build a bubbleplot with ggplot. The core bacteriobiome of all 45 soils samples was identified at the genus level based on a relative abundance of >0.5 and 95% occurrence in soil samples (Neu et al., 2021). Differences in the distribution of the core bacteriobiome according to the soil samples and the three types of vegetation were evaluated by performing a Kruskal-Wallis test, followed by a Wilcoxon rank sum test with a significant level of $p < 0.05$. RDA analysis was carried out to show variables involved in the distribution of the core genera. Spearman correlations among the abundance of each core genus and environmental variables were done using the Hmsic package (Harrell Jr, 2021). Only correlations greater than 0.5 or less -0.5 with a $p < 0.01$ were considered. A correlogram for

---
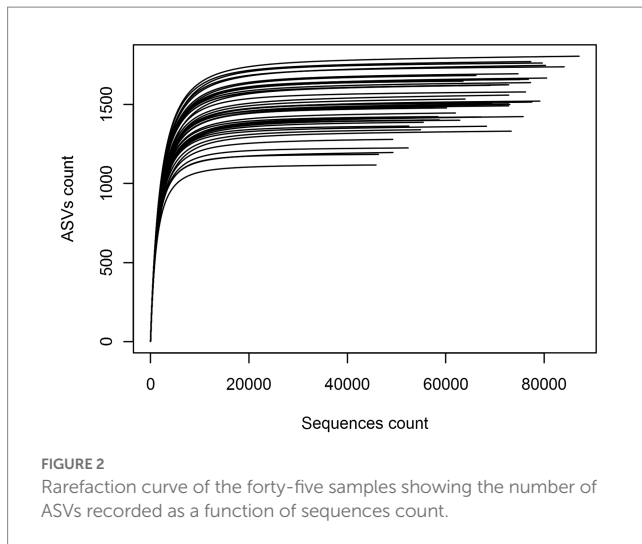
correlation levels among the core genera were built by calculating Pearson correlation.

2.5. Core microbiome function prediction using PICRUSt2

The filtered ASV matrix was used to predict the core bacterial metabolic functions from PICRUSt2 v2.4.2 (Douglas et al., 2020; Djemiel et al., 2022) with implemented tools HMMER[2] (Finn et al., 2011), SEPP (SATe-enabled Phylogenetic Placement) (Mirarab et al., 2012), GAPPA (Genesis Applications for Phylogenetic Placement Analysis) (Czech et al., 2020) and Castor (Louca and Doebeli, 2018). The KEGG Orthology (KO) was used[3] (Kanehisa and Goto, 2000; Kanehisa, 2019; Kanehisa et al., 2021, 2022; Djemiel et al., 2022) and KO numbers obtained from PICRUSt2 analysis was used to map the different functions categories. A manual selection was made among all the metabolic pathways predicted by ASV and only those of interest in agriculture were selected, i.e., nitrogen metabolism, sulfur metabolism, carbon fixation as well as symbiotic pathway (Yan et al., 2015). Heatmaps of the metabolic pathways were generated with the ggplot function. Kruskal-Wallis tests were carried out to determine differences between functions predicted across the samples and the types of vegetation.

3. Results

3.1. Physico-chemical and climate characteristics of Côte d'Ivoire soils

pH values of the 45 soils ranged from 4.9 to 7.8, being consistent with pH range expected in soil of tropical and humid regions, as already reported for Côte d'Ivoire soils (Tondoh et al., 2015; Kassin

---

2 http://hmmer.org/

3 https://www.kegg.jp

et al., 2017; Tano et al., 2020). Approximately 11% of soils were classified as strongly acidic (pH <5.5), 64% as acidic (5.5 to 6.5), 22% neutral (6.5 to 7.5) and only 2% as alkaline (pH over 7.5). Acidic soils were observed across the three vegetation biomes, together with the neutral soils. In contrast, the strongly acidic soils were mostly encountered in the forest zone. The majority of these soils (at least 90%) had a high sand content with a minimum of 53%. Of all the soils, 47% were sandy loam, 13% loam sandy, 31% sandy clay loam and 9% loam.
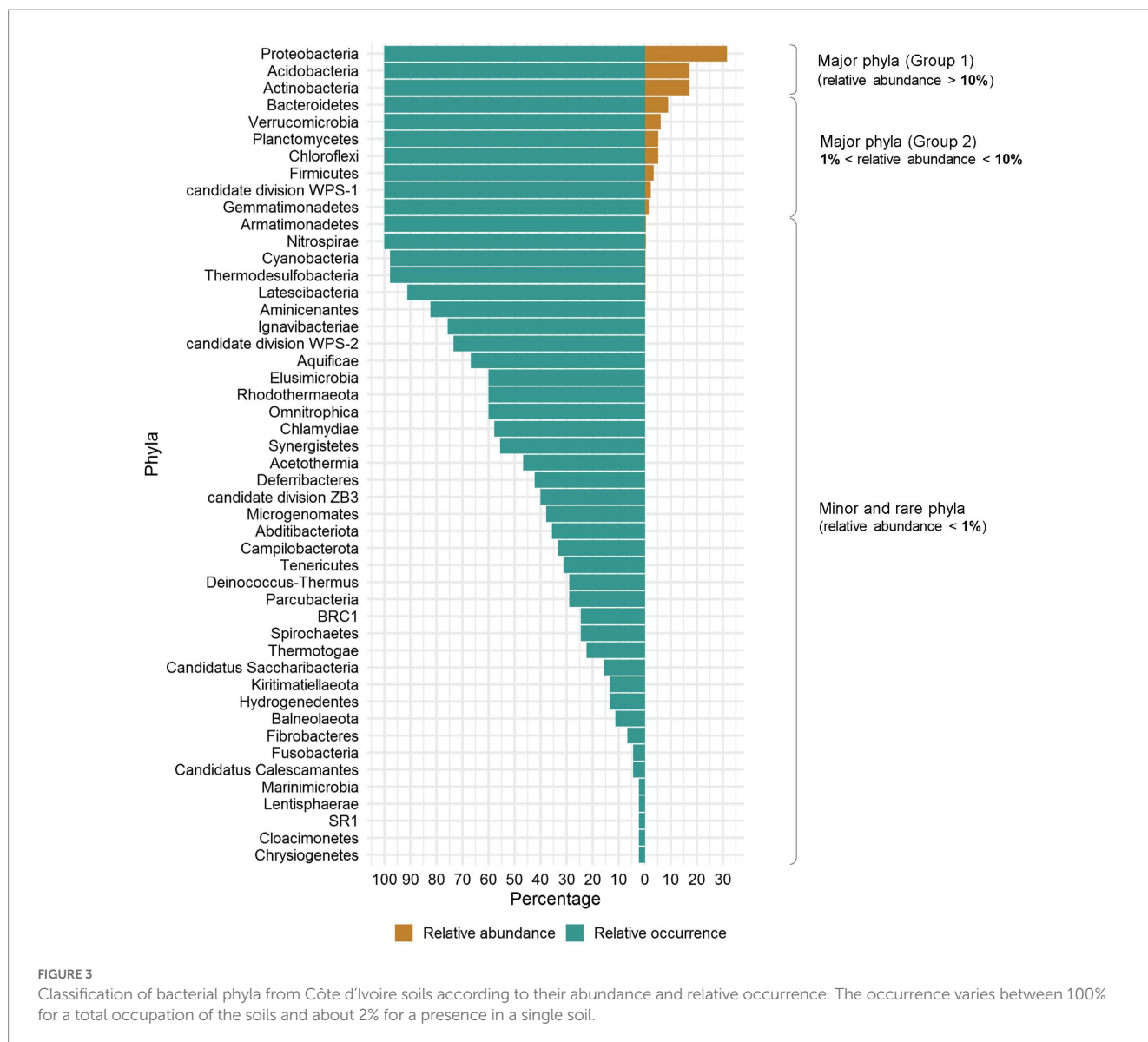
Chemical analyses revealed that organic carbon and total nitrogen contents of soils were statistically different among the three types of vegetation ($p < 0.01$), with a high level in forest soils (1.4 ± 0.46 for C and 0.16 ± 0.06 for N), while savannah and the contact zone had similar C and N contents. pH values of soils, and all the chemical elements tested were not different between the three vegetation zones. Similar results were observed for the proportion of sand, silt and clay. Precipitation ($p = 3.9e\text{-}03$) and temperature ($p = 2e\text{-}03$) data were statistically different among the three vegetation zones. All of these data are presented in Supplementary Tables S1, S2.

3.2. Bacterial communities in Côte d'Ivoire soils are mainly dominated by three ecologically important phyla

To assess sampling quality, rarefaction curves were plotted, showing all the curves reaching asymptotes with less than 30,000 sequences and suggesting that the sequencing effort of each eDNA sample was sufficient (Figure 2). Of the total of 3,645,012 sequences obtained, the 3,175,778 remaining after filtering were clustered and yielded a total of 23,910 bacterial ASVs. Variation in sequences length was shown in Supplementary Table S3. Subsequent analyses allowed their clustering into 48 Phyla, 92 Classes, 152 Orders, 356 Families and 1,234 Genera. When a hierarchical classification of the 48 phyla was performed according to their abundance (Supplementary Figure S1), it was shown that the cumulative sequences of 10 phyla represented more than 98% of the total filtered sequences. Their abundance and occurrence in soils were significantly different ($p < 0.05$) from those of the remaining 37 phyla (Supplementary Table S4). These 10 core phyla included Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Firmicutes, candidate division WPS-1, Gemmatimonadetes Planctomycetes, Proteobacteria, and Verrucomicrobia. According to their relative abundance, they could be divided into two main groups: a first group of phyla with a relative abundance per phylum above 10% (i.e., Acidobacteria, Actinobacteria, and Proteobacteria) and a second group of phyla with a relative abundance per phylum ranging from 1 to 10% (the seven remaining) (Figure 3). The Proteobacteria, Actinobacteria, Bacteroidetes and Chloroflexi communities were differently distributed across the three biomes (ANOSIM: $0.25 < R < 0.5$, $p < 0.05$) while the other phyla (e.g., Acidobacteria, Verrucomicrobia etc.) were equally distributed in the three biomes (ANOSIM: $0.1 < R < 0.25$, $p < 0.05$). The Proteobacteria and Bacteroidetes for example were more spread in the forest zone of Côte d'Ivoire, while Chloroflexi were more abundant in the savannah zone. Supplementary Figure S2 shows the maps of the 10 dominant phyla based on their relative abundance and occurrence in Côte d'Ivoire soils.

FIGURE 3
Classification of bacterial phyla from Côte d'Ivoire soils according to their abundance and relative occurrence. The occurrence varies between 100% for a total occupation of the soils and about 2% for a presence in a single soil.

3.3. Causative relationships between environmental factors and bacterial phyla distribution in Côte d'Ivoire soils

Redundancy analysis was carried out on soil physico-chemical properties, as well as climate and spatial variables, in order to understand their influence on bacterial community distribution. The total variance explained by all these variables in the distribution of the 10 phyla ranged from 6.12 to 19.36%. Their effect on bacterial phylum distribution is summarized by the following relationships: physicochemical properties > spatial data > climate (Figure 4). Among soil nutrients, potassium (K) content had a significant impact on the 10 phyla. For example, the abundance of Proteobacteria ($R^2 = 6.1\%$), Verrucomicrobia ($R^2 = 5.8\%$) and Acidobacteria ($R^2 = 5.3\%$) decreased with increasing K levels, but the opposite effect was observed on Actinobacteria ($R^2 = 5.5\%$) and Bacteroidetes ($R^2 = 6.2\%$). Together with K content, soil pH was one of the factors that significantly influenced all the 10 phyla ($p < 0.05$). Al levels had a higher effect on Verrucomicrobia ($R^2 = 2.3\%$) than on the other phyla. Concerning the

spatial variables, only latitude had a significant effect on all phyla distribution, with a higher influence on Verrucomicrobia ($R^2 = 6.2\%$), Proteobacteria ($R^2 = 6.1\%$) and Acidobacteria ($R^2 = 5.9\%$) (Figure 4).

3.4. The core bacteriobiome of Côte d'Ivoire soils

The core bacteriobiome was determined at the genus level. Among the total of 1,234 genera, only 10 were present with 100% occurrence (Supplementary Table S5) and above 0.5% relative abundance in all samples (Supplementary Table S5; Supplementary Figure S3). The 10 core genera belonged to the phyla of Proteobacteria (30%), Acidobacteria (30%), Verrucomicrobia (20%), Actinobacteria (10%) and candidate division WPS-1 (10%) (Table 1).

The distribution of these genera in the 45 soil samples is represented in Figure 5. *Gp4* and *WPS-1_genera_incertae_sedis* were differently distributed in the soils of the three types of vegetation ($p < 0.01$), while the other genera were homogeneously distributed
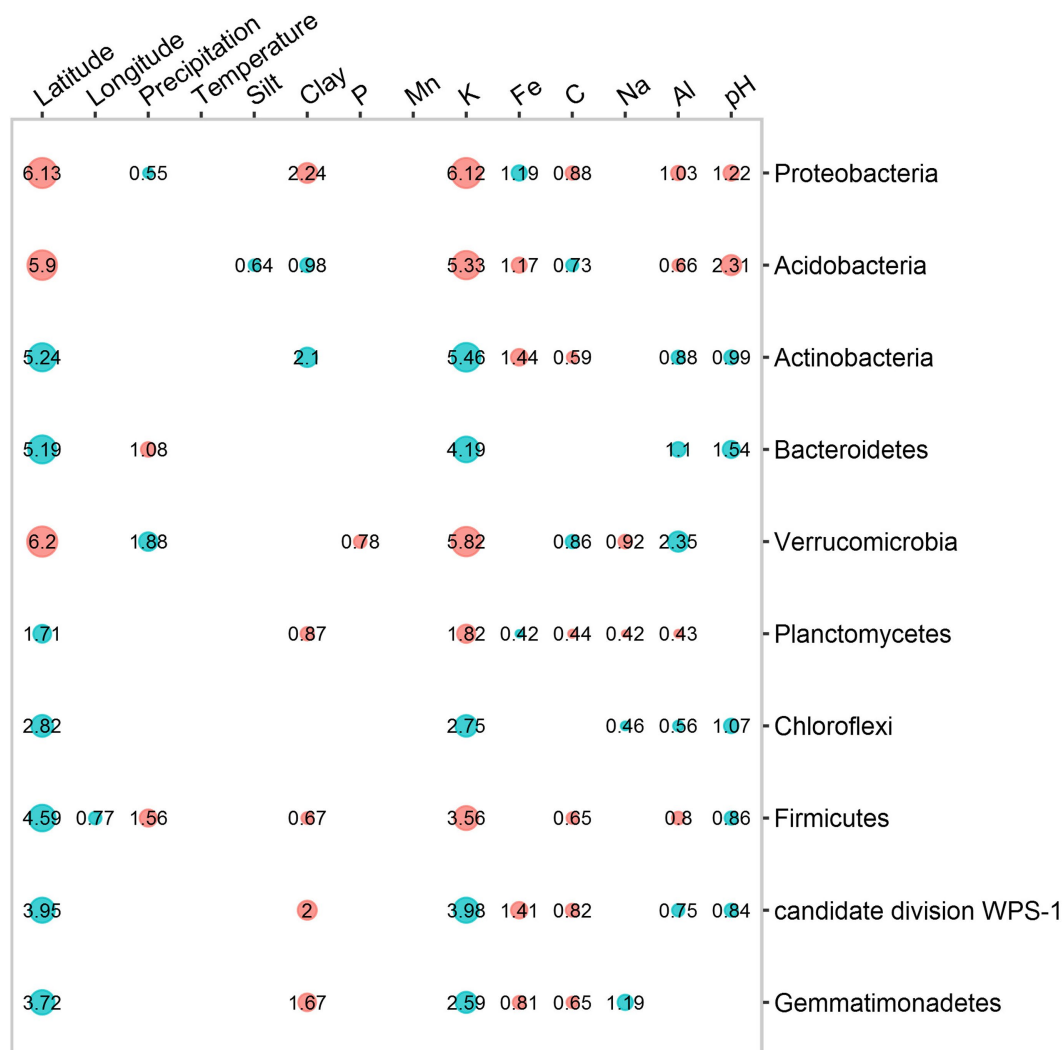
FIGURE 4
Contribution and effects of each variable (geographical location, climate and physico-chemical properties) in the distribution of bacterial taxa. The proportions of variance explained were the significant contributions of these variables ($p < 0.05$). The size of the circles as well as the values written inside provide information on the proportion of variance explained by each variable. The colors blue and red indicate the effect of the regression coefficients based on Spearman's correlation (blue, positive effect, red, negative effect).

($p > 0.05$) (Supplementary Table S6). A hierarchical clustering of the core genera abundance based on the Bray distance revealed three distinct groups of soils (Figure 5). The first group Gr1 (18% of soil samples) had a high abundance of *Gaiella* and low abundance of *Bradyrhizobium*, *Spartobacteria_genera_incertae_sedis* and *Subdivision3_genera_incertae_sedis*, respectively. The second group Gr2 (27%) had high abundance of *Bradyrhizobium*, *Spartobacteria_genera_incertae_sedis*, and *Subdivision3_genera_incertae_sedis*. The third group Gr3 (55%), was dominated by *Gp6* (Figure 5). The Wilcoxon rank sum test used showed that the *Gaiella* genus had different abundances between the three groups while *Bradyrhizobium*, *Gp3*, *Gp6* and *Subdivision3_genera_incertae_sedis* were evenly distributed in both groups Gr1 and Gr3. Moreover, *Rhodoplanes* was equally present in both Gr2 and Gr3 groups. The abundance of *Pseudorhodoplanes* as well as *Gp4* was the same across the Gr1 and Gr2 groups, while the abundance of *WPS-1_genera_incertae_sedis* was found to be different in Gr2 and Gr3 groups (Supplementary Table S7).

3.5. Main variables influencing the core bacteriobiome distribution across Côte d'Ivoire soils

Redundancy analyses were used to identify the main factors that influenced the core bacteriobiome distribution. These factors included the pH ($R^2 = 22.85\%$), K ($R^2 = 9.35\%$), latitude ($R^2 = 5.56\%$) and Al ($R^2 = 2.98\%$) (Figure 6). Moreover, Spearman's correlation coefficient analyses showed that alkaline pH positively affected *Gp4* abundance ($\rho = 0.6$, $p < 1e\text{-}04$) but not *Bradyrhizobium* ($\rho = -0.62$, $p < 1e\text{-}04$) and *Spartobacteria_genera_incertae_sedis* ($\rho = -0.59$, $p < 1e\text{-}04$), that were more abundant in acidic soils. However, K-rich soils were shown to have an increased abundance of *Gp6* genus ($\rho = 0.59$, $p < 1e\text{-}04$). However, soils with high K concentrations had low abundance of *Bradyrhizobium* and *Gp3* (*Bradyrhizobium*: $\rho = -0.6$, $p < 1e\text{-}04$; Gp3: $\rho = -0.65$, $p < 1e\text{-}04$). Lastly, the abundance of *Gp4* ($\rho = 0.52$, $p = 2e\text{-}04$) and

TABLE 1 Supplementary Table summarizing of occurrence and abundance data of the core genera found in Côte d'Ivoire soils.

Phyla	Genera	Occurrence (%)	Min.-Max. soil relative abundance (%)	Median (%)
Actinobacteria	*Gaiella*	100	1.1–10.9	3.6
Acidobacteria	*Gp3*	100	0.5–3.1	1.1
	Gp4	100	0.5–5.4	2.0
	Gp6	100	1.8–12.8	5.4
Proteobacteria	*Pseudorhodoplanes*	100	1.0–5.6	1.8
	Rhodoplanes	100	1.3–4.5	2.1
	Bradyrhizobium	100	0.5–4.0	1.9
Verrucomicrobia	*Spartobacteria_genera_incertae_sedis*	100	0.7–10.7	3.5
	Subdivision3_genera_incertae_sedis	100	0.6–5.5	1.5
candidate division WPS-1	*WPS-1_genera_incertae_sedis*	100	0.8–4.6	2.2

Each genus was related to its phylum in the first column. Relative occurrence values are given for each genus. In addition to median values, minimum and maximum relative abundance from all soils are given to show the range of variation.
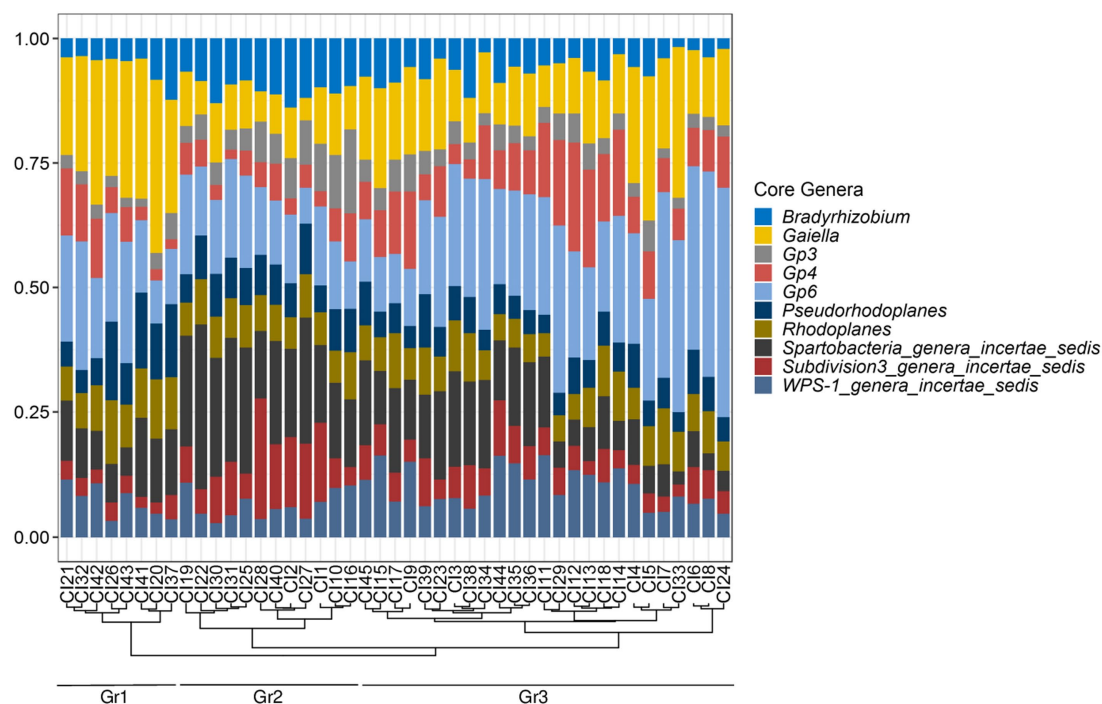
FIGURE 5
The core bacteriobiome of Côte d'Ivoire soils: taxonomic distribution per sample for the most abundant genera with 100% occurrence. Hierarchical clustering of samples based on the relative abundances of the core genera is shown by the dendrogram on the bottom of the plot which revealed three distinct groups: Gr1, Gr2, and Gr3.

WPS-1_genera_incertae_sedis ($\rho = 0.67$, $p < 1e\text{-}04$) were highly correlated with increasing latitude values (Figure 6).

Interactions among genera belonging to the core bacteriobiome were also analyzed (Supplementary Table S8; Supplementary Figure S4). It was shown that the abundance of the genera *Bradyrhizobium*, *Gp3*, *Spartobacteria_genera_incertae_sedis*, *Subdivision3_genera_incertae_sedis* and *Pseudorhodoplanes* were positively correlated ($r \geq 0.5$, $p < 0.001$). Soils with higher abundance of *Gp6* genus shows low abundance of *Bradyrhizobium* ($r = -0.6$, $p < 0.001$). The abundance of *Rhodoplanes* was strongly correlated to *Pseudorhodoplanes* ($r = 0.8$,

$p < 0.001$) and *Gaiella* ($r = 0.5$, $p < 0.001$), whereas *WPS-1_genera_incertae_sedis* abundance was strongly correlated with *Gp4* ($r = 0.7$, $p < 0.001$).

3.6. Predicted functions of the core bacteriobiome in Côte d'Ivoire soils

A total of 44 metabolic capacities were used to predict the functional properties of the core bacteriobiome. The most abundant

FIGURE 6
Redundancy analysis (RDA) ordination diagram (triplot) of core bacteria genera and physicochemical factors across Côte d'Ivoire soils. The plot shows samples in dots, explanatory physicochemical variables (red arrows), and the response variables (orange square with genera names). Axis 1 and axis 2 explained 37.9 and 9.6% of the total variation, respectively. Positively correlated variables are shown as arrows going in the same direction.

predicted metabolic capacities were Carbohydrate Metabolism and Amino Acid metabolism (Figure 7).

   Further predictive functional analyses were performed for all 10 dominant genera, focusing on energy metabolism, particularly the capacity for carbon, nitrogen and sulfur cycling. In general, it was shown that carbon, nitrogen and sulfur cycles were statistically different within each bacterial genus (Figure 8; Supplementary Table S9). In contrast, the analysis revealed that only sulfur metabolism and autotrophic carbon cycle were different according to soil samples ($p < 2.2e$-16) (Supplementary Figure S5). Carbon sequestration was the most abundantly predicted and was significantly different among vegetation type ($p < 2.2e$-16) just like sulfur metabolism ($p = 3.8e$-10). Nitrogen metabolism ($p = 0.13$) and symbiosis ($p = 0.9$) showed no significant difference through the three biomes.

   In terms of occurrence, carbon sequestration was by far the most frequently predicted function, followed by sulfur metabolism, nitrogen metabolism. As for the nitrogen, symbiotic fixation was found to be possible in *Bradyrhizobium* only (Figure 8).

   Interestingly, *Bradyrhizobium* was also the only genus from the core bacteriobiome capable of performing all the energy metabolism functions (Figure 9). All the 10 core genera potentially metabolized sulfate to hydrogen sulfide through assimilatory sulfate reduction, while only *Bradyrhizobium*, *Pseudorhodoplanes* and *Rhodoplanes* (Proteobacteria genera) are predicted to oxidize thiosulfate to sulfate (Figure 9; Supplementary Table S9).

4. Discussion

   This work aimed to assess the landscape-scale distribution of bacteria in Côte d'Ivoire soils and characterize the environmental

factors that may influence the structure of the core bacteriobiome. Studies of the biogeography of soil microorganisms have been performed at large scale worldwide (Griffiths et al., 2011; Malard et al., 2019) including, most recently, Sub-saharian Africa (Cowan et al., 2022). However, nationwide studies, as done elsewhere (Karimi et al., 2018), have not been performed in sub-Saharan Africa and the soil microbiomes remain unexplored in terms of key bacterial taxa abundance, prevalence and functions. It is particularly true for countries like Côte d'Ivoire, where few studies have been conducted on the field of soil bacterial ecology. Despite its large biome diversity, including fallow, gallery-forest, primary-forest, forest-savannah mosaic, shrubland and wooded-savannah landscapes (Tra Bi et al., 2015), soil chemical data suggested that Côte d'Ivoire soils properties were only partially dependent on biome type. For example, acidic soils were observed across all three major biomes. The significant differences observed among the three vegetation types concerning C and N contents and climate data were not surprising as it has been already reported elsewhere (Tondoh et al., 2015; Kassin et al., 2017).

   In order to assess the bacterial distribution through the diverse vegetation types of Côte d'Ivoire soils, HTAS analysis of the 16S rRNA gene have been performed. The analysis revealed that Côte d'Ivoire soils were dominated in terms of abundance and prevalence by 10 bacterial phyla, including Acidobacteria, Actinobacteria, Bacteroidetes, candidate division WPS-1, Chloroflexi, Firmicutes, Gemmatimonadetes, Planctomycetes, Proteobacteria and Verrucomicrobia as reported elsewhere (Cowan et al., 2022). It is commonly accepted that these phyla are the ones most frequently found in the top layers of soils worldwide (Delgado-Baquerizo et al., 2018; Karimi et al., 2018). Here, the distribution of the 10 bacterial phyla was substantially driven by latitude, K content, clay, pH and Al content. pH has repeatedly been reported as one of the major predictors of soil bacterial community composition (Fierer and Jackson, 2006; Wei et al., 2017; Karimi et al., 2018; Liu et al., 2018; Cheng et al., 2020). It was shown that there was an inverse relationship between pH and Actinobacteria abundance where Actinobacteria were less abundant in the forest biome with lower pH while Acidobacteria were higher, as demonstrated elsewhere (Lauber et al., 2009).

   It has been reported that different environmental factors can lead to differences in niches, resulting in the occurrence of a core microbiome (Turnbaugh et al., 2007). Furthermore, it has been argued that, the identification of the core microbiome is important since its persistent presence in particular habitats is likely to be essential for their functioning (Shade and Handelsman, 2012). However, despite its importance, there has been a little real consensus on how a core microbiome should be defined (Neu et al., 2021). Conditions in the determination of a core microbiome vary enormously, depending on the purpose of the study. It is often defined at different taxonomic levels, notably at the genus level (Lupwayi et al., 2021) or at the OTUs/ASVs level as reported in several studies (Gołębiewski et al., 2014; Jiao et al., 2019; Taye et al., 2020; Gschwend et al., 2021). Neu et al. (2021), reported three methods to quantify a core microbiome, for example by the occurrence in samples as in Gschwend et al. (2021), the relative abundance or by the combination of the occurrence and abundance. The latest method has been often used in recent studies (Jiao et al., 2019). Neu et al. (2021) also reported the wide variation in threshold value used to define the core microbiome, varying from 50 to 100%
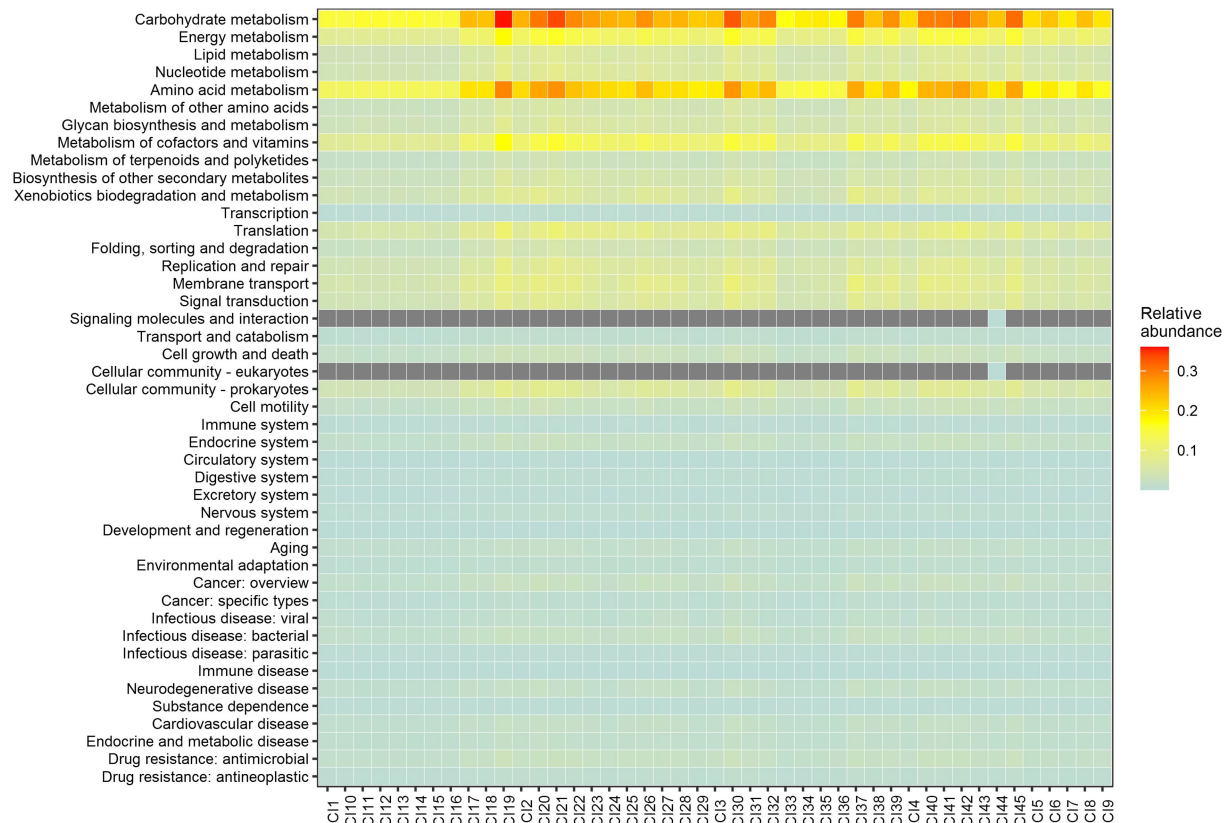
FIGURE 7
PICRUSt2 predicted metabolic capacities of the core genera based on 16S rRNA genes and the KEGG database. Abundance values are ranked from the lowest (light blue) to the highest (red). The grey boxes means no value recorded.
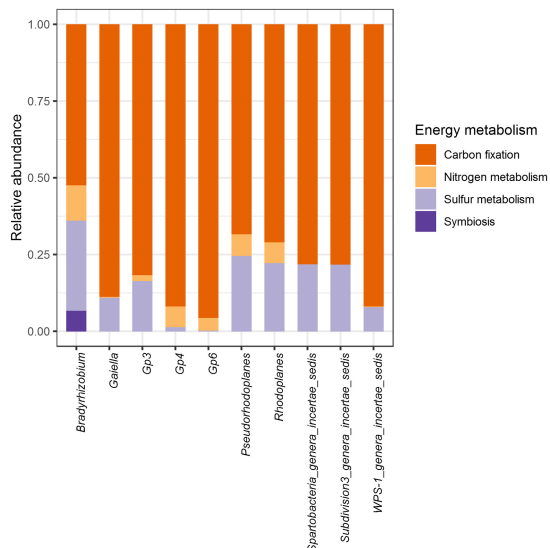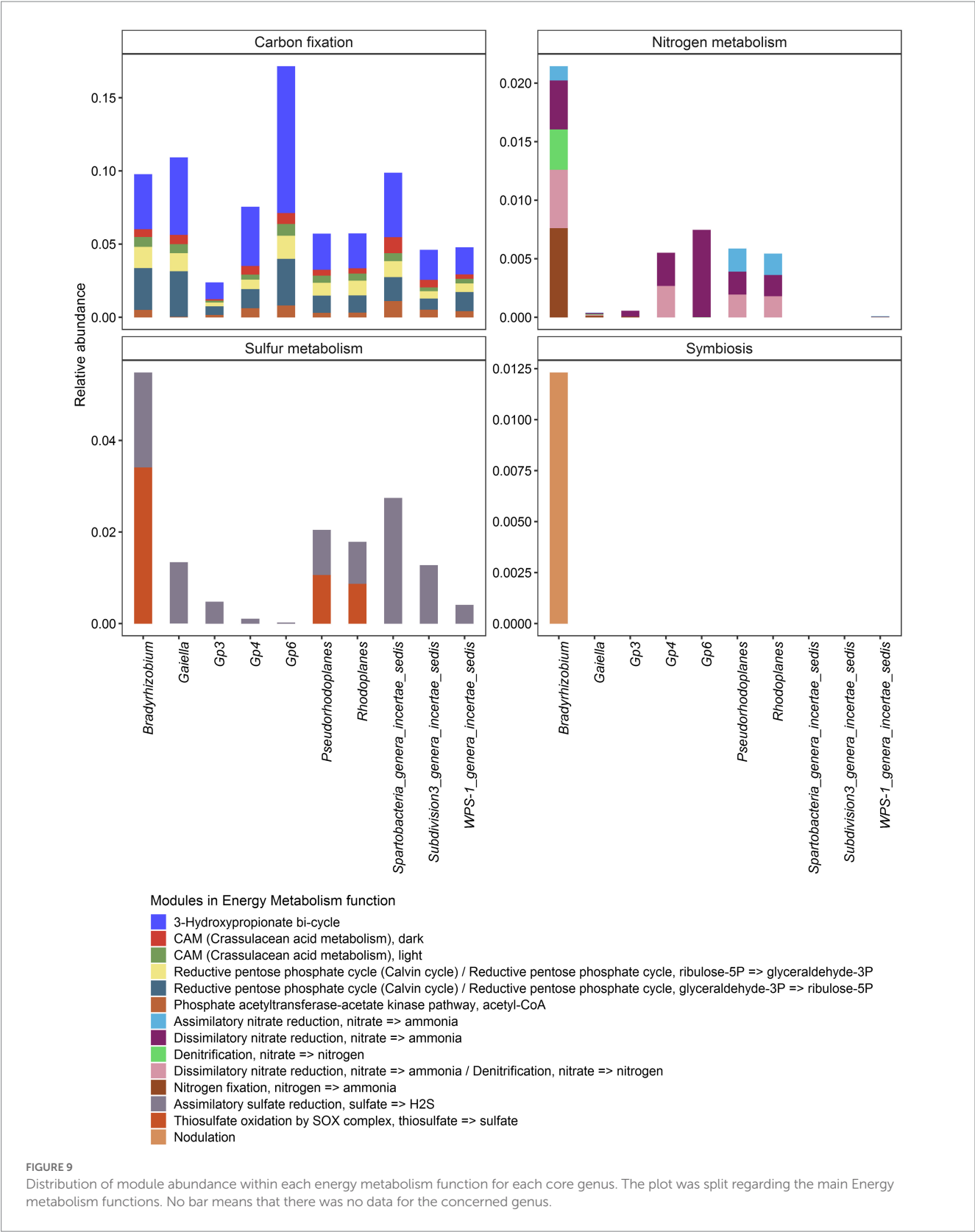

FIGURE 8
Relative abundance of genes associated with agriculturally important energy metabolism functions in the core bacteriobiome of Côte d'Ivoire soils.

for occurrence, and 0.001–4.5% for abundance. Regarding all these different methods, this study choose to focus on a genus level core microbiome based on a combination of abundance and occurrence

conditions. In an analysis of the core bacteriobiome of Côte d'Ivoire soils, a total of 10 genera belonging to the 10 major phyla were found to be prevalent in at least 95% of all soils with an abundance of above 0.5%. These were *Bradyrhizobium*, *Gaiella*, *Gp3*, *Gp4*, *Gp6*, *Pseudorhodoplanes*, *Rhodoplanes*, *Spartobacteria_genera_incertae_sedis*, *Subdivision3_genera_incertae_sedis* and *WPS-1_genera_incertae_sedis*. These taxa are known to be soil-associated (Navarrete et al., 2013; Gołębiewski et al., 2014; Taye et al., 2020; Lupwayi et al., 2021). The ubiquity of these genera may suggest an involvement in soil structure and functioning, or may serve as an indicator of soil quality or be linked to a particular vegetation/crop type (Simonin et al., 2020). For example, members of *Bradyrhizobium* genus, known to be able to establish symbioses with leguminous plants (Woomer et al., 1988; Bouznif et al., 2019), were recently found to dominate in tropical savannah soils of Côte d'Ivoire in a legume cropping area (Gnangui et al., 2021). The *Gaiella* genus may have an important role in agricultural soils as a plant decomposer and inhibitor of *Fusarium oxysporum f.* sp. *Lycopersici* in tomato fields (Zhao et al., 2019). This functional trait in *Gaiella* may be of importance for Côte d'Ivoire, where vegetable-based agriculture are increasing (Soro et al., 2008). Along with *Gp4* and *Gp6*, the predominance of *Gp3* (Navarrete et al., 2013) could potentially be linked to the wide range of agricultural activities (Dedysh and Sinninghe Damsté, 2018), since more than 60% of the soils sampled were recovered from agricultural regions (see BioProject PRJNA695288). The high abundance of genus *Spartobacteria_genera_incertae_sedis* in Côte d'Ivoire soils was consistent with

FIGURE 9
Distribution of module abundance within each energy metabolism function for each core genus. The plot was split regarding the main Energy metabolism functions. No bar means that there was no data for the concerned genus.

previous studies (Brewer et al., 2016) which have demonstrated that this taxon is ubiquitous and abundant in many soils types as grassland. It was evident that in addition to the (weak) influence of vegetation type, there were a combination of different variables that

primarily shaped the distribution of the core bacteriobiome in Côte d'Ivoire soils. For example, *Gp6*, *Gp4* and *WPS-1_genera_incertae_sedis* were abundant in soils with highest values of pH, K, Ca, Mn and lower rainfall (Navarrete et al., 2013). *Gaiella*, *Rhodoplanes* and

Pseudohodoplanes were abundant in soils with a pH level of about 6.1, with an intermediate content of K, Ca, Mn, Fe in biomes with moderately high rainfall level, being consistent with previous reports (Hiraishi and Ueda, 1994; Albuquerque et al., 2011; Supplementary Tables S7, S10). The fact that *Bradyrhizobium* and *Gp3* abundances were also linked to pH levels further confirms the involvement of this parameter in bacterial communities structuration as previously reported (Woomer et al., 1988; Kielak et al., 2016; Dedysh and Sinninghe Damsté, 2018).

Core microbial taxa are thought to be of importance in agroecosystem servicing (Risely, 2020). Most of the predicted functions for *Gp6* were related to carbon metabolism (more than 80%), as for *Gaiella*, *Gp3*, *Gp4* and *WPS- 1_genera_incertae_sedis*. The other genera, on the other hand, shown the ability to vary their source of mineral resources by less carbon fixation (ranging from about 75–50% for *Bradyrhizobium* only) and the use of nitrogen and/or sulfur. The significant difference ($p < 0.05$) in the sulfur metabolism across soils could reveal that depending on the chemical properties of the environment (sulfur availability for example) certain species would be able to perform such oxidation. Unsurprisingly, only members of the Proteobacteria (*Bradyrhizobium*, *Rhodoplanes*, and *Pseudorhodoplanes*) had the putative ability to oxidise thiosulfate, a key sulfur cycle compound, into sulfate (Zhang et al., 2020). Another important predicted function is related to nitrogen metabolism mainly performed by *Bradyrhizobum*. In general important findings have been reported on the benefits of the microbial activities linked to N cycling, including N_2 fixation, mineralization, nitrification and denitrification in Côte d'Ivoire soils. Culture-dependent studies have revealed that the genus *Bradyrhizobium* represents a source of potential bioinoculants capable of fostering the growth of diverse leguminous crops (Fossou et al., 2020). Recently, a pioneering study of soil bacteriobiome carried out in fields of central Côte d'Ivoire has shown that among the key microbial activities linked to N cycling, nitrification plays a crucial role in determining the outcome of agronomic practices (Assémien et al., 2017). Taken together, this exploratory work, as well as previous reports, provided novel insights into understanding the genetic and functional diversity of the key bacterial taxa in Côte d'Ivoire soils, and may serve as indicators for future microbiome explorations and for land-use decision-making in Côte d'Ivoire.

5. Conclusion

This work did a national scale metagenomic study in Côte d'Ivoire in order to investigate the bacterial communities inhabiting the soils sampled in several biomes. As a result, the bacterial population mainly encountered in these soils were unsurprisingly those that were commonly found in soils across the world with a main pH-based distribution. When assessing the core bacteriobiome, the major distribution scheme was found to be vegetation-independent. While the predominance of *Bradyrhizobium* genus in soils is well-known and documented, some other members of the core bacteriobiome such as *Gaiella* lack key information that could help in better understanding their role in Côte d'Ivoire soils. At genus level, *Bradyrhizobium* was the only taxon from the core bacterial microbiome capable of performing all the energy metabolism functions investigated, in addition to its capacity for nodulation through symbiosis. In sum, this work, which

is the one of the pioneering study to ever be undertaken at this scale in Côte d'Ivoire, unraveled the complexity of the Côte d'Ivoire soil bacteriobiome, and opened a new era to investigate in deep its functions and services in the national ecosystem functioning.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/, PRJNA695288.

Author contributions

DC and AZ: study conception, supervision, and project administration. YB, DV, and CA: data acquisition. DC, AZ, CA, and RF: methodology. CA, AE, DK, RF, and CK: software. CA and RF: validation and data curation. CA, AE, and RF: data analysis. CA: writing-original draft preparation. AZ, RF, CA, and AE: writing-review and editing. DC: funding acquisition. All authors read and approved the final manuscript.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1220655/full#supplementary-material

References

Alami, M. M., Xue, J., Ma, Y., Zhu, D., Abbas, A., Gong, Z., et al. (2020). Structure, function, diversity, and composition of fungal communities in Rhizospheric soil of Coptis chinensis Franch under a successive cropping system. *Plants* 9:244. doi: 10.3390/plants9020244

Albuquerque, L., França, L., Rainey, F. A., Schumann, P., Nobre, M. F., and da Costa, M. S. (2011). *Gaiella occulta* gen. Nov., sp. nov., a novel representative of a deep branching phylogenetic lineage within the class *Actinobacteria* and proposal of *Gaiellaceae* fam. Nov. and *Gaiellales* Ord. Nov. *Syst. Appl. Microbiol.* 34, 595–599. doi: 10.1016/j.syapm.2011.07.001

Assémien, F. L., Pommier, T., Gonnety, J. T., Gervaix, J., and Le Roux, X. (2017). Adaptation of soil nitrifiers to very low nitrogen level jeopardizes the efficiency of chemical fertilization in west african moist savannas. *Sci. Rep.* 7:10275. doi: 10.1038/s41598-017-10185-5

Avenard, J.-M., Eldin, M., Girard, G., Sircoulon, J., Touchebeuf, P., Guillaumet, J.-L., et al. (1971). *Le milieu naturel de la Côte d'Ivoire*, Paris, ORSTOM

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Bouznif, B., Guefrachi, I., Rodríguez De La Vega, R. C., Hungria, M., Mars, M., Alunni, B., et al. (2019). Phylogeography of the *Bradyrhizobium* spp. associated with Peanut, *Arachis hypogaea*: fellow travelers or new associations? *Front. Microbiol.* 10:2041. doi: 10.3389/fmicb.2019.02041

Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A., and Fierer, N. (2016). Genome reduction in an abundant and ubiquitous soil bacterium 'Candidatus Udaeobacter copiosus'. *Nat. Microbiol.* 2, 1–7. doi: 10.1038/nmicrobiol.2016.198

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

Castellano-Hinojosa, A., and Strauss, S. L. (2021). Insights into the taxonomic and functional characterization of agricultural crop core rhizobiomes and their potential microbial drivers. *Sci. Rep.* 11, 10068–10011. doi: 10.1038/s41598-021-89569-7

Cheng, J., Zhao, M., Cong, J., Qi, Q., Xiao, Y., Cong, W., et al. (2020). Soil pH exerts stronger impacts than vegetation type and plant diversity on soil bacterial community composition in subtropical broad-leaved forests. *Plant Soil* 450, 273–286. doi: 10.1007/s11104-020-04507-2

Christensen, H., Andersson, A. J., Jørgensen, S. L., and Vogt, J. K. (2018). "16S rRNA amplicon sequencing for metagenomics" in *Introduction to bioinformatics in microbiology* (Springer Nature)

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244

Cowan, D., Lebre, P., Amon, C., Becker, R., Boga, H., Boulangé, A., et al. (2022). Biogeographical survey of soil microbiomes across sub-Saharan Africa: structure, drivers, and predicted climate-driven changes. *Microbiome* 10:131. doi: 10.1186/s40168-022-01297-w

Czech, L., Barbera, P., and Stamatakis, A. (2020). Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* 36, 3263–3265. doi: 10.1093/bioinformatics/btaa070

Dedysh, S. N., and Sinninghe Damsté, J. S. (2018). "Acidobacteria" in *eLS* (Chichester: John Wiley & Sons, Ltd (Wiley))

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., et al. (2018). A global atlas of the dominant bacteria found in soil. *Science* 359, 320–325. doi: 10.1126/science.aap9516

Djemiel, C., Maron, P.-A., Terrat, S., Dequiedt, S., Cottin, A., and Ranjard, L. (2022). Inferring microbiota functions from taxonomic genes: a review. *GigaScience* 11:giab090. doi: 10.1093/gigascience/giab090

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi: 10.1038/s41587-020-0548-6

Ducroquet, H., Tillie, P., Louhich, K., and Gomez-Y-Paloma, S. (2017). L'agriculture de la Côte d'Ivoire à la loupe: Etats des lieux des filières de production végétales et animales et revue des politiques agricoles. Publications Office of the European Union, Luxembourg: Science for Policy report by the Joint Research Centre. EUR 28754 FR, 244. Available at doi:10.2760/126254, JRC107214.

FAO (2009). État des ressources phytogénétiques pour l'alimentation et l'agriculture: Second rapport national. *FAO* 65.

Fenchel, T., King, G. M., and Blackburn, H. (2012). *Bacterial biogeochemistry: The ecophysiology of mineral cycling*. Academic Press, San Diego

Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* 15, 579–590. doi: 10.1038/nrmicro.2017.87

Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci.* 103, 626–631. doi: 10.1073/pnas.0507535103

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367

Fossou, R. K., Pothier, J. F., Zézé, A., and Perret, X. (2020). Bradyrhizobium ivorense sp. nov. as a potential local bioinoculant for *Cajanus cajan* cultures in Côte d'Ivoire. *Int. J. Syst. Evol. Microbiol.* 70, 1421–1430. doi: 10.1099/ijsem.0.003931

Gautier, L. (1990). Contact forêt-savane en Côte-d'Ivoire centrale: évolution du recouvrement ligneux des savanes de la Réserve de Lamto (sud du V-Baoulé). *Acta. Bot. Gall.* 136, 85–92. doi: 10.1080/01811789.1989.10826960

Gnangui, S. L. E., Fossou, R. K., Ebou, A., Amon, C. E. R., Koua, D. K., Kouadjo, C. G. Z., et al. (2021). The Rhizobial microbiome from the tropical Savannah zones in northern Côte d'Ivoire. *Microorganisms* 9:1842. doi: 10.3390/microorganisms9091842

Gołębiewski, M., Deja-Sikora, E., Cichosz, M., Tretyn, A., and Wróbel, B. (2014). 16S rDNA pyrosequencing analysis of bacterial Community in Heavy Metals Polluted Soils. *Microb. Ecol.* 67, 635–647. doi: 10.1007/s00248-013-0344-7

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., and Whiteley, A. S. (2011). The bacterial biogeography of British soils: mapping soil bacteria. *Environ. Microbiol.* 13, 1642–1654. doi: 10.1111/j.1462-2920.2011.02480.x

Gschwend, F., Hartmann, M., Mayerhofer, J., Hug, A.-S., Enkerli, J., Gubler, A., et al. (2021). Site and land-use associations of soil bacteria and fungi define core and indicative taxa. *FEMS Microbiol. Ecol.* 97:fiab165. doi: 10.1093/femsec/fiab165

Guillaumet, J.-L., and Adjanohoun, E. (1971). "La végétation de la Côte d'Ivoire," in *Le milieu naturel de la Côte d'Ivoire* Mémoires ORSTOM.

Hijmans, R. J. (2020). Raster: geographic data analysis and modeling. Available at: https://CRAN.R-project.org/package=raster.

Hiraishi, A., and Ueda, Y. (1994). (1994). Rhodoplanes gen. Nov., a new genus of phototrophic Bacteria including *Rhodopseudomonas rosea* as *Rhodoplanes roseus* comb. nov. and *Rhodoplanes elegans* sp. nov. *Int. J. Syst. Evol. Microbiol.* 44, 665–673. doi: 10.1099/00207713-44-4-665

Jacoby, R., Peukert, M., Succurro, A., Koprivova, A., and Kopriva, S. (2017). The role of soil microorganisms in plant mineral nutrition—current knowledge and future directions. *Front. Plant Sci.* 8:1617. doi: 10.3389/fpls.2017.01617

Jiao, S., Chen, W., and Wei, G. (2022). Core microbiota drive functional stability of soil microbiome in reforestation ecosystems. *Glob. Chang. Biol.* 28, 1038–1047. doi: 10.1111/gcb.16024

Jiao, S., Xu, Y., Zhang, J., Hao, X., and Lu, Y. (2019). Core microbiota in agricultural soils and their potential associations with nutrient cycling. *mSystems* 4:e00313–18. doi: 10.1128/mSystems.00313-18

Jr, Harrell, and Frank, E, (2021). Hmisc: Harrell Miscellaneous R package version, 4.6-0. Available at: https://CRAN.R-project.org/package=Hmisc

Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* 28, 1947–1951. doi: 10.1002/pro.3715

Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., and Tanabe, M. (2021). KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* 49, D545–D551. doi: 10.1093/nar/gkaa970

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Kanehisa, M., Sato, Y., and Kawashima, M. (2022). KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci.* 31, 47–53. doi: 10.1002/pro.4172

Karimi, B., Terrat, S., Dequiedt, S., Saby, N. P. A., Horrigue, W., Lelièvre, M., et al. (2018). Biogeography of soil bacteria and archaea across France. *Sci. Adv.* 4:eaat1808. doi: 10.1126/sciadv.aat1808

Kassin, E., Snoeck, D., Nguessan, J.-C., Yao-Kouamé, A., and Camara, M. (2017). Soil mapping project final report. CNRA, Cirad, Idh, World Cocoa Foundation, Conseil du café cacao, Côte d'Ivoire. Available at: https://www.idhsustainabletrade.com/uploaded/2017/04/CNRA-SOIL-MAPPING-PROJECT-FINAL-REPORT.pdf.

Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A., and Kuramae, E. E. (2016). The ecology of Acidobacteria: moving beyond genes and genomes. *Front. Microbiol.* 7:744. doi: 10.3389/fmicb.2016.00744

Kolton, M., Weston, D. J., Mayali, X., Weber, P. K., McFarlane, K. J., Pett-Ridge, J., et al. (2022). Defining the Sphagnum Core microbiome across the north American continent reveals a central role for Diazotrophic Methanotrophs in the nitrogen and carbon cycles of boreal peatland ecosystems. *MBio* 13, e03714–e03721. doi: 10.1128/mbio.03714-21

Kouadio, A. N. M.-S., Nandjui, J., Krou, S. M., Séry, D. J.-M., Nelson, P. N., and Zézé, A. (2017). A native arbuscular mycorrhizal fungus inoculant outcompetes an exotic commercial species under two contrasting yam field conditions. *Rhizosphere* 4, 112–118. doi: 10.1016/j.rhisph.2017.10.001

Koutika, L.-S., Fiore, A., Tabacchioni, S., Aprea, G., Pereira, A. P. D. A., and Bevivino, A. (2020). Influence of *Acacia mangium* on soil fertility and bacterial Community in Eucalyptus Plantations in the Congolese Coastal Plains. *Sustainability* 12:8763. doi: 10.3390/su12218763

Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* 75, 5111–5120. doi: 10.1128/AEM.00335-09

Lavelle, P., and Spain, A. (2001). *Soil ecology.* Springer Science & Business Media, New York.

Liu, D., Yang, Y., An, S., Wang, H., and Wang, Y. (2018). The biogeographical distribution of soil bacterial communities in the loess plateau as revealed by high-throughput sequencing. *Front. Microbiol.* 9:2456. doi: 10.3389/fmicb.2018.02456

Louca, S., and Doebeli, M. (2018). Efficient comparative phylogenetics on large trees. *Bioinformatics* 34, 1053–1055. doi: 10.1093/bioinformatics/btx701

Lupwayi, N. Z., Janzen, H. H., Bremer, E., Smith, E. G., Kanashiro, D. A., Eastman, A. H., et al. (2021). The core soil bacterial genera and enzyme activities in incubated soils from century-old wheat rotations. *Geoderma* 404:115275. doi: 10.1016/j.geoderma.2021.115275

Malard, L. A., Anwar, M. Z., Jacobsen, C. S., and Pearce, D. A. (2019). Biogeographical patterns in soil bacterial communities across the Arctic region. *FEMS Microbiol. Ecol.* 95. 13. doi: 10.1093/femsec/fiz128

Maquia, I. S., Fareleira, P., Videira E Castro, I., Brito, D. R. A., Soares, R., Chaúque, A., et al. (2020). Mining the microbiome of key species from African savanna woodlands: potential for soil health improvement and plant growth promotion. *Microorganisms* 8:1291. doi: 10.3390/microorganisms8091291

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200

McMurdie, P. J., and Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Mirarab, S., Nguyen, N., and Warnow, T. (2012). SEPP: SATé-enabled phylogenetic placement. *Pac. Symp. Biocomput.*, 247–258. doi: 10.1142/9789814366496_0024

Moeys, J. (2018). Soiltexture: functions for soil texture plot, Classification and Transformation. R package version 1.5.1. Available at: https://CRAN.R-project.org/package=soiltexture

Murtagh, F., and Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J. Classif.* 31, 274–295. doi: 10.1007/s00357-014-9161-z

Navarrete, A. A., Kuramae, E. E., de Hollander, M., Pijl, A. S., van Veen, J. A., and Tsai, S. M. (2013). Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. *FEMS Microbiol. Ecol.* 83, 607–621. doi: 10.1111/1574-6941.12018

Neu, A. T., Allen, E. E., and Roy, K. (2021). Defining and quantifying the core microbiome: challenges and prospects. *PNAS* 118:e2104429118. doi: 10.1073/pnas.2104429118

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2020). Vegan: community ecology package. Available at: https://CRAN.R-project.org/package=vegan.

Pajares, S., Campo, J., Bohannan, B. J. M., and Etchevers, J. D. (2018). Environmental controls on soil microbial communities in a seasonally dry tropical Forest. *Appl. Environ. Microbiol.* 84:17. doi: 10.1128/AEM.00342-18

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Pershina, E. V., Ivanova, E. A., Korvigo, I. O., Chirak, E. L., Sergaliev, N. H., Abakumov, E. V., et al. (2018). Investigation of the core microbiome in main soil types from the east European plain. *Sci. Total Environ.* 631, 1421–1430. doi: 10.1016/j.scitotenv.2018.03.136

Pivato, B., Chemidlin Prévost-Boure, N., and Lemanceau, P. (2015). "Microbiome du sol" in *La métagénomique Développements et futures applications.* ed. S. Faire (France: Versailles Cedex)

Quoreshi, A. (2008). "The use of mycorrhizal biotechnology in restoration of disturbed ecosystem" in *Mycorrhizae: sustainable agriculture and forestry.* ed. Z. A. Siddiqui (New York: Springer)

R Core Team (2020). R: a language and environment for statistical computing. Available at: https://www.R-project.org/

Risely, A. (2020). Applying the core microbiome to understand host–microbe systems. *J. Anim. Ecol.* 89, 1549–1558. doi: 10.1111/1365-2656.13229

Schloss, P. D., and Handelsman, J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol.* 6:229. doi: 10.1186/gb-2005-6-8-229

Séry, D. J.-M., van Tuinen, D., Drain, A., Mounier, A., and Zézé, A. (2018). The genus Rhizophagus dominates arbuscular mycorrhizal fungi communities in contrasted cassava field soils in Côte d'Ivoire. *Rhizosphere* 7, 8–17. doi: 10.1016/j.rhisph.2018.06.007

Shade, A., and Handelsman, J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. *Appl. Environ. Microbiol.* 14, 4–12. doi: 10.1111/j.1462-2920.2011.02585.x

Shen, Y., Ji, Y., Li, C., Luo, P., Wang, W., Zhang, Y., et al. (2018). Effects of phytoremediation treatment on bacterial community structure and diversity in different petroleum-contaminated soils. *Int. J. Environ. Res. Public Health* 15:2168. doi: 10.3390/ijerph15102168

Simonin, M., Dasilva, C., Terzi, V., Ngonkeu, E. L. M., Diouf, D., Kane, A., et al. (2020). Influence of plant genotype and soil on the wheat rhizosphere microbiome: evidences for a core microbiome across eight African and European soils. *FEMS Microbiol. Ecol.* 96:fiaa067. doi: 10.1093/femsec/fiaa067

Sindhu, S., Sharma, R., Sindhu, S., and Sehrawat, A. (2019). "Soil fertility improvement by symbiotic rhizobia for sustainable agriculture" in *Soil Fertility Management for Sustainable Development.* eds. D. G. Panpatte and V. K. Jhala (Singapore: Springer Nature)

Soro, S., Doumbouya, M., and Koné, D. (2008). Potentiel infectieux des sols de cultures de tomate (*Lycopersicon esculentum* Mill.) sous abri et incidence de l'âge de repiquage sur la vigueur des plants vis-à-vis de Pythium sp. à Songon-Dabou en Côte d'Ivoire. *Tropicultura* 26, 173–178.

Tano, B. F., Brou, C. Y., Dossou-Yovo, E. R., Saito, K., Futakuchi, K., Wopereis, M. C. S., et al. (2020). Spatial and temporal variability of soil redox potential, pH and electrical conductivity across a Toposequence in the savanna of West Africa. *Agronomy* 10:1787. doi: 10.3390/agronomy10111787

Taye, Z. M., Helgason, B. L., Bell, J. K., Norris, C. E., Vail, S., Robinson, S. J., et al. (2020). Core and differentially abundant bacterial taxa in the rhizosphere of field grown *Brassica napus* genotypes: implications for canola breeding. *Front. Microbiol.* 10:3007. doi: 10.3389/fmicb.2019.03007

Tondoh, J. E., Kouamé, F. N., Martinez Guéi, A., Sey, B., Wowo Koné, A., and Gnessougou, N. (2015). Ecological changes induced by full-sun cocoa farming in Côte d'Ivoire. *Glob. Ecol. Conserv.* 3, 575–595. doi: 10.1016/j.gecco.2015.02.007

Tra Bi, Z. A., Brou, T., and Mahe, G. (2015). *Analyse par télédétection des conditions bioclimatiques de végétation dans la zone de contact forêt-savane de Côte d'Ivoire: cas du "V" Baoulé.*

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007). The human microbiome project. *Nature* 449, 804–810. doi: 10.1038/nature06244

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wang, Y., and Qian, P.-Y. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* 4:e7401. doi: 10.1371/journal.pone.0007401

Wei, Z., Hu, X., Li, X., Zhang, Y., Jiang, L., Li, J., et al. (2017). The rhizospheric microbial community structure and diversity of deciduous and evergreen forests in Taihu Lake area, China. *PLoS One* 12:e0174411. doi: 10.1371/journal.pone.0174411

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. Available at: https://ggplot2.tidyverse.org

Woomer, P., Singleton, P. W., and Bohlool, B. B. (1988). Ecological indicators of native rhizobia in tropical soils. *Appl. Environ. Microbiol.* 54, 1112–1116. doi: 10.1128/aem.54.5.1112-1116.1988

Yan, N., Marschner, P., Cao, W., Zuo, C., and Qin, W. (2015). Influence of salinity and water content on soil microorganisms. *Int. Soil Water Conserv. Res.* 3, 316–323. doi: 10.1016/j.iswcr.2015.11.003

Zhang, J., Liu, R., Xi, S., Cai, R., Zhang, X., and Sun, C. (2020). A novel bacterial thiosulfate oxidation pathway provides a new clue about the formation of zero-valent sulfur in deep sea. *ISME J.* 14, 2261–2274. doi: 10.1038/s41396-020-0684-5

Zhao, F., Zhang, Y., Dong, W., Zhang, Y., Zhang, G., Sun, Z., et al. (2019). Vermicompost can suppress fusarium oxysporum f. sp. lycopersici via generation of beneficial bacteria in a long-term tomato monoculture soil. *Plant Soil* 440, 491–505. doi: 10.1007/s11104-019-04104

Check for updates

Morchella esculenta cultivation in fallow paddy fields and drylands affects the diversity of soil bacteria and soil chemical properties

Mingzheng Duan[1,2†], Chengcui Yang[1,2†], Liuyuan Bao[1,2], Duo Han[1,2], Huaizheng Wang[1,2], Yongzhi Zhang[1,2], Honggao Liu[1,2]* and Shunqiang Yang[1,2]*

[1]Yunnan Key Laboratory of Gastrodia Elata and Fungal Symbiotic Biology, College of Agronomy and Life Sciences, Zhaotong University, Zhaotong, China, [2]Yunnan Engineering Research Center of Green Planting and Processing of Gastrodia Elata, College of Agronomy and Life Sciences, Zhaotong University, Zhaotong, China

The properties of paddy field (DT) and dry land (HD) soil and food production can be enhanced by the cultivation of *Morchella esculenta* (ME) during the fallow period. However, whether ME cultivation affects the soil health and microbial diversity of paddy fields and drylands during the cultivation period remains unclear, and this has greatly limited the wider use of this cultivation model. Here, we analyzed the soil chemical properties and bacterial diversity (via metabarcoding sequencing) of DT and HD soils following ME cultivation. Our findings indicated that ME cultivation could enhance soil health. The content of soil phosphorus and potassium (K) was increased in DT soil under ME cultivation, and the K content was significantly higher in HD soil than in DT soil under ME cultivation. ME cultivation had a weak effect on alpha diversity, and ME cultivation affected the abundance of some genera of soil bacteria. The cultivation of ME might reduce the methane production capacity of DT soil and enhance the nitrogen cycling process of HD soil based on the results of functional annotation analysis. Network analysis and correlation analysis showed that *Gemmatimonas*, *Bryobacter*, and *Anaeromyxobacter* were the key bacterial genera regulating soil chemical properties in DT soil under ME cultivation, and *Bryobacter*, *Bacillus*, *Streptomyces*, and *Paenarthrobacter* were the key taxa associated with the accumulation of K in HD soil. The results of our study will aid future efforts to further improve this cultivation model.

KEYWORDS

rice, corn, soils, 16S rRNA, metabarcoding, NPK

1 Introduction

The sustainable production of food has long been a major goal for mankind, and this is being increasingly challenged by human population growth and ongoing climate change (Zhang et al., 2018). Thus, much effort has been made to enhance the efficiency of land use and increase food production (Brown et al., 2014; Di Benedetto et al., 2017; Babu et al., 2020).

Morchella esculenta (i.e., Yang-Du-Jun (YDJ); ME) is an economically valuable edible fungus with beneficial medicinal properties (Paul et al., 2018; Sunil and Xu, 2022). ME can be cultivated in both paddy field (DT) and dryland (HD) environments. ME has often been
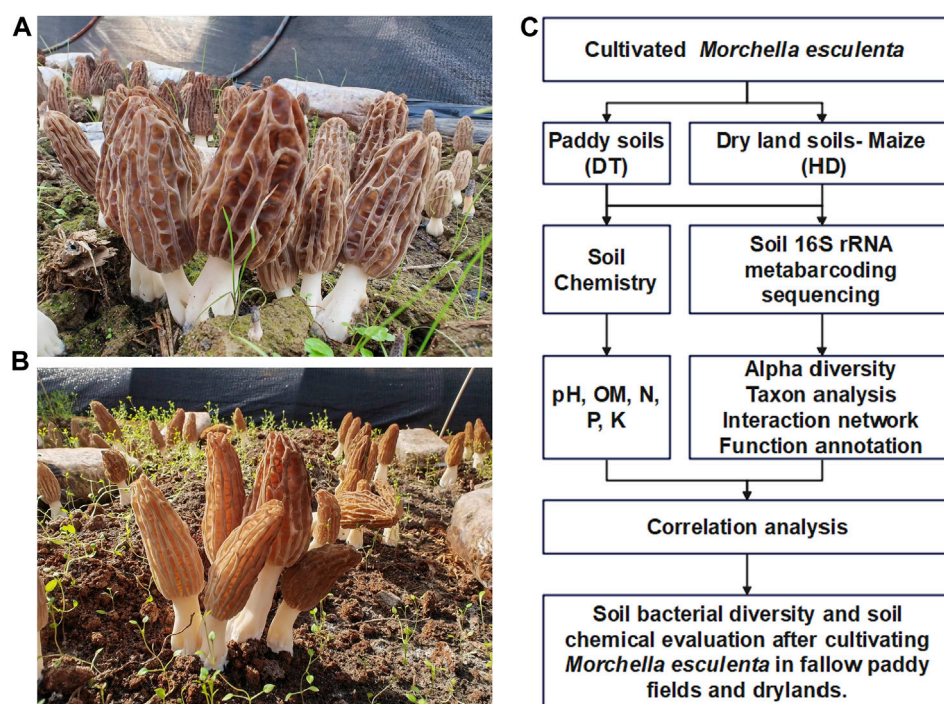
FIGURE 1
Experimental design for the cultivation of ME in paddy fields and dry lands during the fallow period. **(A)** ME cultivated in paddy soil (DT soils). **(B)** ME cultivated in dryland soils (HD soils). **(C)** Technical outline of this study.

cultivated during the fallow period in both DT and HD environments to enhance food production.

Soil health is affected by various biotic and abiotic factors; microbes, especially bacteria and fungi, play a critically important role in soil biogeochemical cycles and the conversion of soil nutrients (Basu et al., 2021). The cultivation of mushrooms is thought to have a substantial effect on soil health. For example, the growth of *Leucocalocybe mongolica* mycelium in soil in a fairy ring ecosystem has been shown to increase soil nutrient abundance (Duan et al., 2022a); this can in turn enhance plant growth, increase the abundance of flavonoid metabolites, and promote hormone synthesis (Duan et al., 2022b). The cultivation of *Dictyophora indusiata* inhibits the activity of soil urease, which reduces the environmental loss of soil N (Duan et al., 2023a). Therefore, the cultivation of macrofungi (mushrooms) has a major effect on soil ecology.

Several previous studies have examined the effects of different *D. indusiata* cultivation techniques on soil bacterial diversity (Orlofsky et al., 2021; Zhang et al., 2023). However, no studies to date have examined the effects of ME cultivation on soil health and the diversity of soil microbes in DT and HD during the fallow period.

The diversity of soil bacteria plays a more important role in mediating the soil-ecological effects of macrofungi than the diversity of soil fungi, and this is related to the fact that the biomass of soil bacteria is much higher than that of fungi (Duan and Bau, 2021). Sequencing technology is one of the most effective approaches for measuring the diversity of soil microbes (Nesme et al., 2016). The diversity of soil microbes in various types of soil has been analyzed in previous studies using metabarcoding sequencing, including grassland soil (Duan et al., 2021), sugarcane field soil

(Duan et al., 2022c; Duan et al., 2023b), and soil under ME cultivation (Orlofsky et al., 2021; Zhang et al., 2023).

Although previous studies have demonstrated that the cultivation of ME during fallow periods in DT and HD can enhance food production, the effects of ME cultivation on soil ecology and the abundances of soil nutrients remain unclear, and this greatly limits the use of this cultivation model. An improved understanding of the effects of ME cultivation on the chemical properties of soil and the bacterial diversity of soil in DT and HD is essential for improving the effectiveness of this cultivation model.

Here, we cultivated ME during the fallow periods in DT and cornfields and collected soil samples at harvest. Soil chemical analysis and metabarcoding sequencing were conducted to analyze the effects of ME cultivation on changes in soil nutrients and bacterial diversity and functions. In addition, we analyzed associations between changes in bacterial populations and the content of soil nutrients under ME cultivation.

2 Materials and methods

2.1 Materials

DT (Figure 1A) and HD (Figure 1B) sampling sites were located in Ludian County, Zhaotong City, Yunnan Province, China. ME (Variety name: Liu-mei, obtained from Longxing Biotechnology Co., LTD, China.) was cultivated in October 2021 and harvested in January 2022. The yield of ME in DT and HD filed was 504 kg/km^2 and 210 kg/km^2.

2.2 Methods

2.2.1 Field sampling method

Soil was collected from under the ME fruiting bodies; samples were obtained from 10 random points and mixed and divide it into 3 portions as biological repetitions for subsequent analysis (including the DT-YDJ and HD-YDJ treatments). Sampling sites in the same area where ME was not cultivated were used as the control groups (including the DT-CK and HD-CK treatments). Soil samples were collected from the top layer at a depth of 0–10 cm. The soil was then sieved through a 0.425 mm filter, packed into sterile cryostorage tubes, and frozen in liquid N for analysis of the metabarcoding sequencing data. The soil samples used in analyses of soil chemical properties were air-dried and stored at ambient temperature.

2.2.2 Analysis of soil chemical properties

Soil pH levels are detected using the potentiometry method (Food and Agriculture Organization, 2007). The content of available N in soil was determined using the Kjeldahl method with sulfuric acid–accelerator digestion (Food and Agriculture Organization, 2015). The content of available P in soil was determined via NaOH alkali melting and molybdenum–antimony resistance spectrophotometry (New south wales, 1988a). The content of available K in soil was determined via NaOH alkali melting and flame photometry (New south wales, 1988b). The K-dichromate oxidation external heating method was used to determine the content of soil organic matter (OM) (Food and Agriculture Organization, 2006).

2.2.3 Metabarcoding analysis

DNA was extracted from soil samples (3 g) using a HiPure Soil DNA Kit (Magen, Guangzhou, China) per the manufacturer's instructions following a previously described method (Duan et al., 2022b). The primers 341F (5′-CCTACGGGNGGCWGCAG-3′) and 806R (5′-GGACTACHVGGGTATCTAAT-3′) were used to amplify the 16S rDNA V3-V4 region in the ribosomal RNA gene for bacteria via polymerase chain reaction (PCR) (Guo et al., 2017). The purified amplicons, which were pooled in equimolar ratios per the standard protocol, were paired-end sequenced using an Illumina Novaseq 6000 platform. The average sequencing depth of each sample is greater than 100000 tags. The UPARSE (Edgar, 2013) (version 9.2.64) pipeline was used to cluster the clean tags into operational taxonomic units (OTUs) with at least 97% similarity. The most abundant tag sequence was selected as the representative sequence within each cluster. The representative OTU sequences were classified using a naïve Bayesian model with The Ribosomal Database Project classifier (version 2.2) (Wang et al., 2007) based on the SILVA (16S rRNA OTUs) database (Pruesse et al., 2007) (version138.1) with a confidence threshold value of 0.8. R software was used to create all figures. Comparisons of OTUs among groups were made using the VennDiagram package (version 1.6.16) in R software (Chen and Boutros, 2011). Nonmetric multidimensional scaling (NMDS) analysis and analysis of the Shannon index were conducted to measure between-group variance. QIIME (version 1.9.1) software was used to calculate the observed species (Sob), Shannon, and Good's

coverage indexes (Caporaso et al., 2010). After the OTU numbers were $\log_2$-transformed, NMDS was conducted in the vegan package (version 2.5.3; http://CRAN.R-project.org/package=vegan; 2022.11.9) in R software to characterize variation in the composition of OTUs among experimental groups. The VennDiagram package (version 1.6.16) in R was used to conduct a Venn analysis that compared the OTUs among the different groups (Chen and Boutros, 2011). Linear discriminant analysis effect size (LEfSe) software was used to conduct LEfSe analysis (Segata et al., 2011), and the value of the linear discriminant analysis (LDA) filtrate score was 4. The ecologically relevant functions of bacteria were predicted using the functional annotations of the prokaryotic taxa (FAPROTAX) database (version 1.0) (Louca et al., 2016) and BugBase database (Ward et al., 2017). OmicShare tools, a dynamic, real-time, and interactive online platform for data analysis (http://www.omicsmart.com) (accessed on 12 April 2023), was used to construct networks based on the correlation coefficients.

2.2.4 Correlation analysis

The OmicShare tools platform was used to analyze the correlation between the metabarcoding data and soil chemical properties data. Data on the soil chemical properties and OTUs were $\log_{10}$ transformed. Correlations were evaluated using the soil chemical properties and the numbers of the 20 most abundant bacterial genera. The correlation heat map tools in OmicShare (accessed on 12 April 2023) were used to make a heat map based on the Pearson correlation coefficients.

3 Results and discussion

3.1 ME cultivation can affect soil chemical properties

We analyzed the pH and abundances of OM, N, P, and K in both DT and HD soils under ME cultivation. In DT soils, no significant differences were observed between the CK and YDJ treatments in the pH (6.4–6.6; Figure 2A), content of OM (45–50.5 g/kg; Figure 2B), and content of N (189.6–201.1 mg/kg; Figure 2C). However, the content of P (15.5–20.1 mg/kg; Figure 2D) and K (170.9–203 mg/kg; Figure 2E) was significantly higher in the YDJ treatment than in the CK treatment (P content: $p = 0.0015$; K content: $p = 0.0193$). Therefore, although no significant differences in pH, OM, and N were observed between the CK and YDJ treatments in DT soils, the content of P and K was markedly higher in the YDJ treatment than in the CK treatment, indicating that ME cultivation increased the availability of these nutrients in the soil. No significant differences in the pH of HD soils (5.1–5.2; Figure 2F) were observed between the CK and YDJ treatments. However, the content of OM (55.5–64.9 g/kg; Figure 2G), N (221.8–295 mg/kg; Figure 2C), and P (144.3–206.6 mg/kg; Figure 2I) was significantly higher in the CK treatment than in the YDJ treatment (OM content: $p = 0.011$; N content: $p = 0.0029$; P content: $p = 0.0017$), and the content of K (123–426.2 mg/kg; Figure 2J) was significantly higher in the YDJ treatment than in the CK treatment ($p = 0.0046$). There were no significant differences in pH in HD soils between the CK and YDJ treatments, and the content of OM, N, and P was markedly higher in the CK treatment
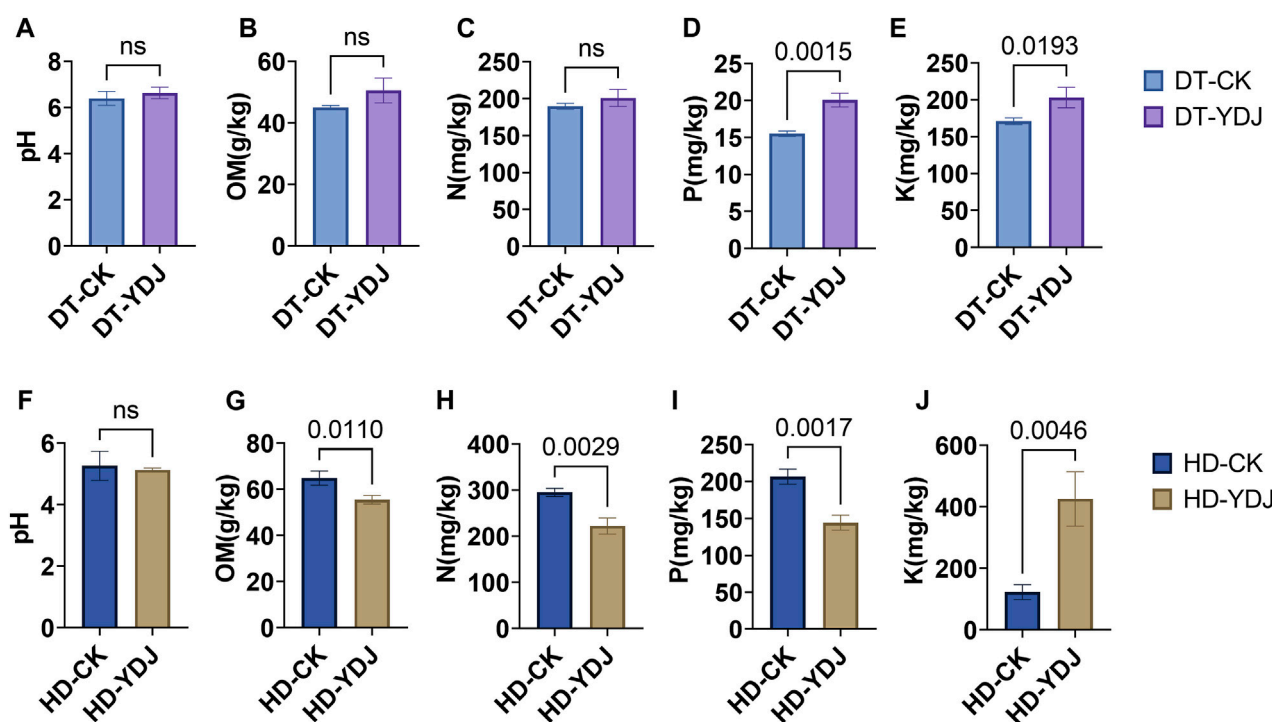
FIGURE 2
Analysis of soil chemical properties following ME cultivation. The abscissa shows the control group (CK) and treatment group (YDJ), and the ordinate shows the chemical properties and their units. DT stands for paddy soil, and HD stands for cornfield soil. **(A, F)** show the pH, **(B, G)** show the content of organic matter (OM), **(C, H)** show the content of hydrolyzable N, **(D, I)** show the content of available P, and **(E, J)** show the content of available K. "ns" stands for not significant, and the number on the bar indicates the *p*-value.

than in the YDJ treatment; the content of K was higher in the YDJ treatment than in the CK treatment, suggesting that the effects of ME cultivation on the availability of nutrients soil vary among soil types. Previous studies have indicated that the cultivation of ME with peach trees and ME fermentation with cultivated maize can increase the content of effective K in the soil (Phanpadith et al., 2020; Song et al., 2021), which is consistent with the results of this study. Similar findings were observed in DT and HD soil, and the effect in HD soil was significant. This indicates that ME cultivation can be used as a replacement for some of the positive effects of K fertilizer. Previous studies have shown that interactions between large fungi and soil can increase soil fertility (Duan et al., 2022a; Duan et al., 2023a), which stems from the fact that interactions of fungi with soil can alter chemical properties, soil enzyme activities, ion concentrations, and other factors; stress resistance can be enhanced via interactions with plants (Duan et al., 2022a; Duan et al., 2022b; Duan et al., 2023a). Overall, the results of our study indicate that the cultivation of ME during the fallow periods in DT and HD not only increases the utilization efficiency of farmland but also possibly improving soil nutrient levels.

3.2 Sequencing statistics and alpha diversity of DT and HD soils under ME cultivation

We conducted a metabarcoding survey to characterize the diversity of soil bacteria under ME cultivation. 16S rRNA

TABLE 1 Bacterial metabarcoding sequencing data.

ID	Clean tags	N90	OTUs
DT-CK-1	124737	441	3603
DT-CK-2	127432	441	3807
DT-CK-3	119486	441	3759
DT-YDJ-1	97653	441	3125
DT-YDJ-2	126552	441	3505
DT-YDJ-3	127208	441	3627
HD-CK-1	115814	441	2565
HD-CK-2	123901	441	2905
HD-CK-3	115817	441	2645
HD-YDJ-1	129155	441	3161
HD-YDJ-2	122425	441	2884
HD-YDJ-3	119843	441	2838

sequencing yielded 1,450,023 effective metabarcoding tags. A total of 3,202 bacterial OTUs were identified per sample on average according to OTU clustering analysis of the soil samples (Table 1). A Venn diagram of the OTUs revealed that 726 (10.56%) of the 6,874 bacterial OTUs were present in all soil samples, and 1,012 (14.72%) and 934 (13.58%) of the OTUs were

FIGURE 3
Venn map, NMDS analysis, and alpha diversity of DT and HD soils under ME cultivation. **(A)** Venn analysis of bacterial OTUs. **(B, C)** NMDS analysis of the bacterial OTUs in the DT-CK/YDJ and HD-CK/YDJ treatments. The colored dots in the figures indicate the different sample groups. **(D, G)** show observed species values, **(E, H)** show Shannon index values, and **(F, I)** show Good's coverage index values. The abscissa shows the CK and YDJ treatments in DT and HD soils, and the ordinate shows the alpha diversity indexes and their units. p-values were calculated by t-tests.

unique to the DT-YDJ and HD-DYJ treatments, respectively (Figure 3A). NMDS plots based on the abundances of OTUs revealed separation in bacterial OTUs between the DT-CK/YDJ (Figure 3B) and HD-CK/YDJ FR (Figure 3C) treatments, indicating that the distribution of bacteria in these treatments significantly differed. We also compared the alpha diversity among samples using a t-test. The mean values of the Sob index of soil samples were 3,723 and 3,419 in the DT-CK and DT-YDJ treatments, respectively (Figure 3D); the mean values of the Sob index were 2,705 and 2,691 in the HD-CK and HD-YDJ treatments, respectively (Figure 3G). The mean values of the Shannon index were 9.8 and 9.58 in the DT-CK and DT-YDJ treatments, respectively (Figure 3E); the mean values of the Shannon index were 8.89 and 9.06 in the HD-CK and HD-YDJ treatments, respectively (Figure 3H). All values of the Good's coverage index were greater than 0.98, which indicates that the sequencing depth of all samples was sufficient. However, the effect

of ME cultivation on the alpha diversity index of soil bacteria was weak in both DT and HD soil, and none of the differences between all comparison groups were significant ($p > 0.05$). A previous metabarcoding study of soil under ME cultivation at different levels of maturity revealed that around 450 of OTUs were identified, and the Shannon diversity index of soil samples was between 4.8 and 5.1 (Orlofsky et al., 2021). The number of OTUs and Shannon index values were much lower in our study than in this previous study. The Shannon index values of soil at different maturity levels ranged from 4 to 6 in another metagenomic sequencing study of soil under ME cultivation at all maturity stages, and this significantly differs from the results of our study (Zhang et al., 2023). This might stem from differences in soil type among studies; it also suggests that OTU richness and alpha diversity might not be the primary factors underlying the interactions between ME and soil. The fact that we did not detect any differences in alpha diversity between soils under

FIGURE 4
Analysis of soil bacterial taxa in DT and HD soils under ME cultivation. The relative abundance of bacterial **(A)** phyla and **(B)** genera. The x-axis indicates the different treatments, the y-axis indicates the relative abundances of bacterial phyla or genera, and the different colors correspond to different phyla or genera. The legend is shown on the right side of the figure. LEfSe analysis of **(C)** DT soils and **(D)** HD soils under ME cultivation. The taxon names represented by specific symbols can be found in Supplementary Figure S1, S2. **(E, F)** indicate the analysis of differentially abundant genera in DT and HD soils following ME cultivation. The vertical axis on the left half of the graph shows differentially abundant species, and the horizontal axis shows mean species abundance. On the right half, the horizontal axis shows the difference in abundance between groups, and the color of the dots indicates the group with higher abundance. The error bars of the dots indicate fluctuations in the 95% confidence interval of the difference. The vertical axis indicates the significance of the difference between corresponding species groups (i.e., the magnitude of the p-value).

ME cultivation and control soil (i.e., YDJ and CK treatments, respectively) provides additional support for this finding.
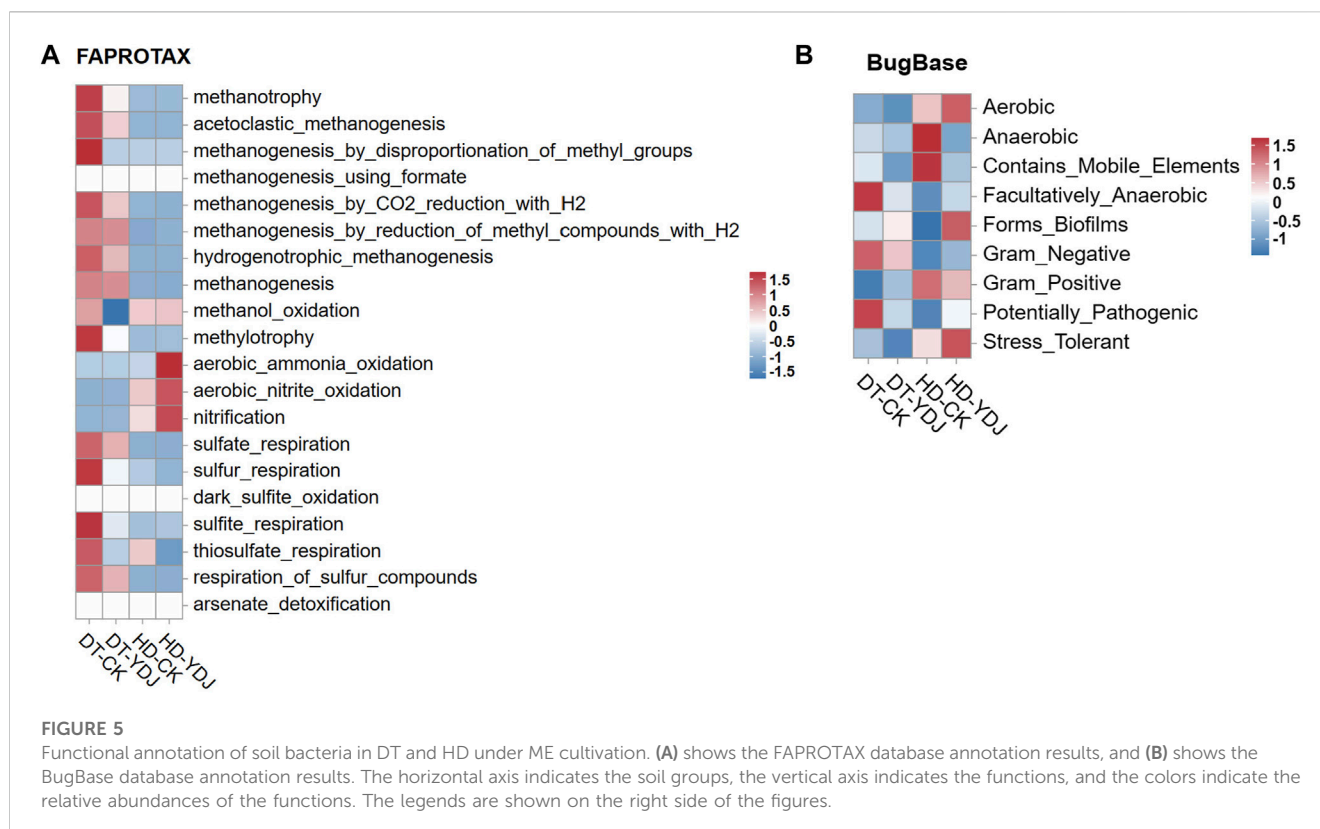
3.3 Composition of soil bacterial taxa

We analyzed differences in the composition of the bacterial community between DT and HD soils under ME cultivation based on the SILVA database. The five most abundant bacterial phyla (Figure 4A) were Proteobacteria (28.99%, 24.68%, 26.97%, and 28.01% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), Actinobacteria (8.07%, 9.07%, 12.96%, and 14.57% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), Acidobacteria (9.48%, 12.62%, 9.88%, and 12.15% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), Chloroflexi (12.87%, 15.44%, 6.39%, and 5.41% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), and Planctomycetes (6.82%, 6.65%, 10.12%, and 9.11% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively); the five most abundant bacterial genera (Figure 4B) were *Sphingomonas* (2.15%, 1.66%, 4.15%, and 6.04% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), *Gemmatimonas* (2.12%, 3.46%, 2.94%, and 3.19% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), *Bryobacter* (1.98%, 2.85%, 1.94%, and 2.56% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), *Flavobacterium* (4.8%, 0.95%, 0.08%, and 0.76% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively), and *Anaeromyxobacter* (1.85%, 2.89%, 0.08%, and 0.76% in the DT-CK, DT-YDJ, HD-CK, and HD-YDJ treatments, respectively). LEfSe analysis revealed that Acidobacteria ($\log_{10}$ LDA score >3, $p < 0.05$) and Gemmatimonadetes ($\log_{10}$ LDA score >3, $p < 0.05$) were the marker phyla in the DT-YDJ vs. DT-CK comparison group (Figure 4C). Cyanobacteria ($\log_{10}$ LDA score >3, $p < 0.05$) and Chlamydiae ($\log_{10}$ LDA score >3, $p < 0.05$) were identified as the marker phyla in the HD-YDJ vs. HD-CK comparison group (Figure 4D). Acidobacteria was enriched in both DT and HD soils under ME cultivation (i.e., DT-YDJ and HD-YDJ treatments); it was also the most abundant bacterial phylum in the DT-YDJ treatment, suggesting that ME cultivation has the most pronounced effect on the abundance of Acidobacteria. Acidobacteria is one of the most common bacteria in soil. It can survive in different soil environments; play a key role in maintaining the structure and function of soil microbial communities, and mediate the conversion of soil nutrients (Lauber et al., 2009; Kielak et al., 2016). Thus, the cultivation of ME in DT and HD soils can enhance soil health by promoting the proliferation of Acidobacteria. Next, we identified bacterial genera that showed significant differences in abundance between DT and HD soils using *t*-tests. In DT soils (Figure 4E), *Bryobacter* (relative abundance of 1.98% and 2.85% in the DT-CK and DT-YDJ treatments, respectively; $p = 0.04$), *Anaeromyxobacter* (1.85% and 2.89% in the DT-CK and DT-YDJ treatments, respectively; $p = 0.01$), *Candidatus_Solibacter* (1.20% and 1.82% in the DT-CK and DT-YDJ treatments, respectively; $p = 0.01$), *Ellin6067* (0.70% and 0.92% in the DT-CK and DT-YDJ treatments,

respectively; $p = 0.01$), and *Bacillus* (0.15% and 0.56% in the DT-CK and DT-YDJ treatments, respectively; $p = 0.001$) were the five most differentially abundant genera; in HD soils (Figure 4F), *Acidibacter* (relative abundance of 0.64% and 0.30% in the HD-CK and HD-YDJ treatments, respectively; $p = 0.04$), *Blastococcus* (0.16% and 0.37% in the HD-CK and HD-YDJ treatments, respectively; $p = 0.01$), and *Mesorhizobium* (0.10% and 0.27% in the HD-CK and HD-YDJ treatments, respectively; $p = 0.04$) were the three most differentially abundant genera. In which, a study has shown that *Bryobacter* and *Candidatus_Solibacter* in rhizosphere soil has the potential to regulate *Cinnamomum migao* metabolism (Li et al., 2023); *Ellin6067* as a kind of nitrosifying bacteria, which may participate in the soil nitrogen cycling process (Lv et al., 2022); *Bacillus* is a well-known plant growth-promoting bacterium (Sheng, 2005; Ali et al., 2021). *Blastococcus* with potential capacity improve the soil nutrient supply (Li et al., 2022); *Mesorhizobium* be found with capacity of improves chickpea growth under chromium stress and alleviates chromium contamination of soil (Naz et al., 2023). Our findings indicate that the composition of the bacterial community differed significantly in DT and HD soils under ME cultivation. The relative abundances of the top bacterial phyla and genera differed between the two soil types, suggesting that the soil type has a major effect on the bacterial community.

3.4 Functional annotation of soil bacteria

To understand the effects of ME cultivation on the ecological functions of soil, we analyzed the functions of soil bacteria using the OTU data. We used the FAPROTAX and Bugbase databases to predict the ecological functions of microbial OTUs identified in the DT-YDJ and HD-YDJ treatments. We found that the abundances of multiple functional modules related to methane synthesis decreased in DT-YDJ soil, including methanotrophy, acetoclastic_methanogenesis, and methanogenesis_by_disproportionation_of_methyl_groups. Rice cultivation produces a large amount of methane, which results in excessive greenhouse gas emissions (Yan et al., 2005); this result indicates that ME cultivation has the potential to reduce the methane emissions of DT soils. The abundances of functional modules related to nitrification were higher in HD-YDJ soil than in HD-CK soil, including aerobic_ammonia_oxidation, aerobic_nitrite_oxidation, and nitrification. Nitrification is the key aerobic process of the biogeochemical N cycle (Koch et al., 2015). These findings indicate that the cultivation of ME in HD soil can promote soil micronutrient cycling. The BugBase database annotations (Figure 5B) were used to predict the biological functions of microorganisms (Ward et al., 2017). The cultivation of ME in DT soils decreased the abundance of the Facultatively_Anaerobic and Potentially_Pathogenic classes. This indicates that the cultivation of ME in DT soil might lower the severity of rice diseases by reducing the abundance of the Potentially_Pathogenic class, and additional studies are needed to confirm this possibility. The cultivation of ME in HD soils increased the abundance of Stress_Tolerant, which suggests that

FIGURE 5
Functional annotation of soil bacteria in DT and HD under ME cultivation. **(A)** shows the FAPROTAX database annotation results, and **(B)** shows the BugBase database annotation results. The horizontal axis indicates the soil groups, the vertical axis indicates the functions, and the colors indicate the relative abundances of the functions. The legends are shown on the right side of the figures.
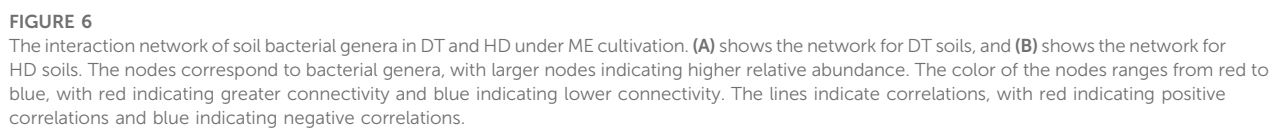
ME cultivation might related to enhance the resistance of corn growth to stress. Overall, ME cultivation had positive effects on the health of both DT and HD soils.

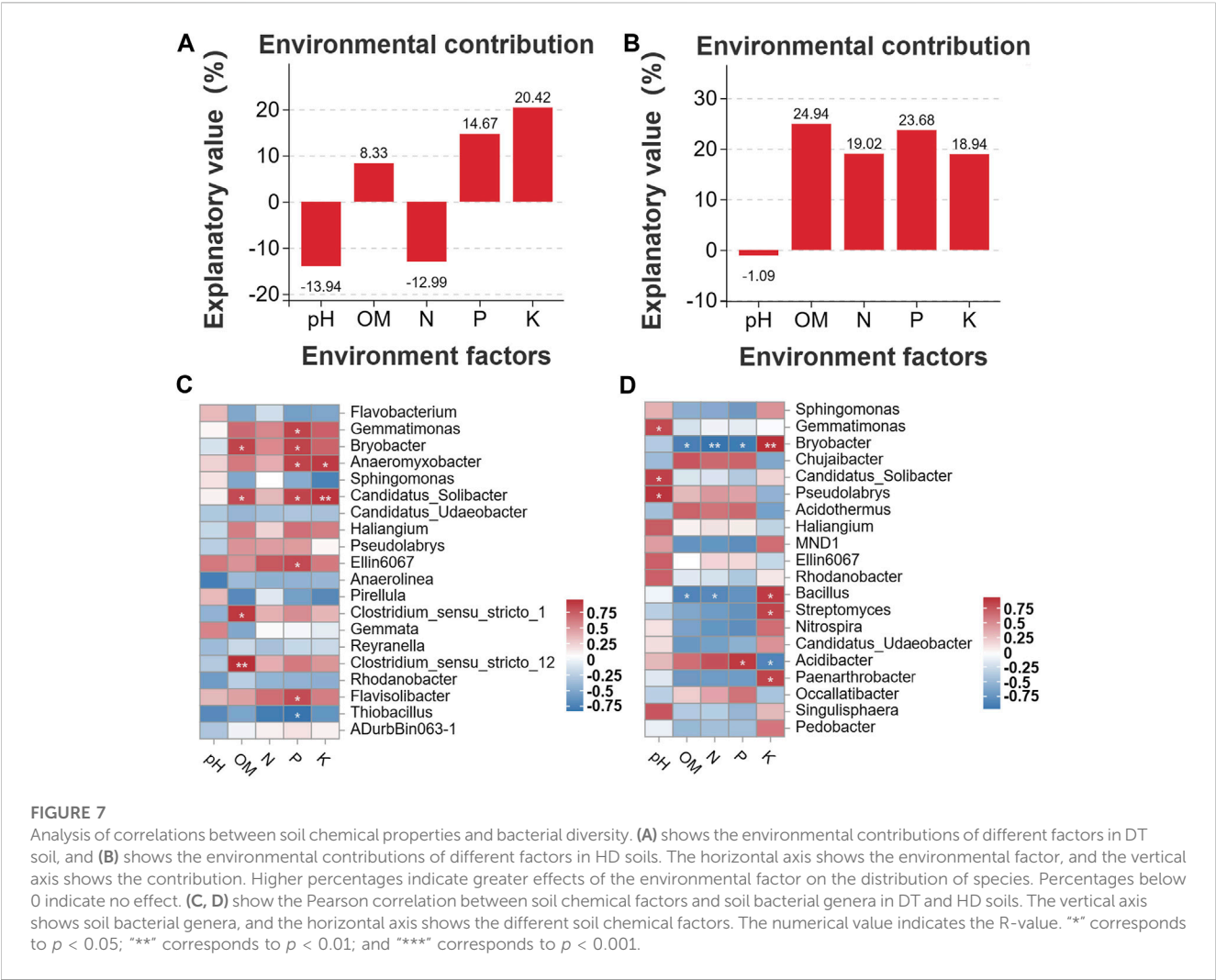## 3.5 Soil bacterial interaction network

We mapped the bacterial interaction network according to the Pearson correlation coefficients to clarify the interactions between bacterial genera in DT and HD soils (Figure 6). In DT soils, the network comprised 96 nodes, and three bacterial genera with high abundance and high connectivity (abundance >2% and connectivity >20) were identified, including *Gemmatimonas*, *Bryobacter*, and *Anaeromyxobacter* (Figure 6A). ME cultivation might affect the abundances of bacterial genera in DT soil, which in turn affects the distribution of soil bacterial diversity. Members of the genus *Gemmatimonas* are capable of anaerobic photosynthesis (Mujakić et al., 2022); however, their effects on soil chemistry remain unclear. The high abundance and connectivity of *Gemmatimonas* in anaerobic rice soil might indirectly affect soil health by affecting other bacteria. The genus *Bryobacter* shows chemoorganotrophic activity and can use sugars, polysaccharides, and organic acids as energy sources; it thus plays a key role in soil metabolism (Pertile et al., 2021). *Anaeromyxobacter* can perform N fixation in soil (Masuda et al., 2020). The network in HD soils consists of 89 nodes (Figure 6B); no bacterial genera with high abundance and high connectivity were detected, and only *Chujaibacter* was found to have a moderate level of connectivity and abundance (abundance >1% and connectivity >10). ME cultivation in DT and HD soils might affect the abundances of these bacteria and thus soil bacterial diversity and health.

3.6 Analysis of the correlations between soil chemical properties and bacterial diversity

Correlation analyses were conducted to clarify the associations between the distribution of microbes and the content of soil chemicals. First, we analyzed the environmental contributions of soil chemical factors in DT and HD soils to soil microbial diversity following ME cultivation. The contribution rates of P and K in DT soil were high (contribution rate >10%) (Figure 7A); the contribution rates of OM, N, P, and K in HD soil were also high (contribution rate >10%) (Figure 7B). In DT soil, the abundances of *Gemmatimonas*, *Bryobacter*, *Anaeromyxobacter*, *Candidatus_Solibacter*, *Ellin6067*, and *Flavisolibacter* were significantly correlated with the content of P, and the abundances of *Anaeromyxobacter* and *Candidatus_Solibacter* were significantly correlated with K (Figure 7C). In HD soil, the abundance of *Acidibacter* was significantly correlated with P, and the abundances of *Bryobacter*, *Bacillus*, *Streptomyces*, and *Paenarthrobacter* were significantly correlated with K (Figure 7D). The abundances of the aforementioned bacterial genera are likely affected by the transformation and accumulation of soil chemical properties. Some of these bacterial genera repeatedly appear in the bacterial interaction network (Figure 6), such as *Gemmatimonas*, *Bryobacter*, and *Anaeromyxobacter*, which emphasizes their importance in regulating soil chemistry processes following ME cultivation. We also identified some bacterial genera that were significantly correlated with the accumulation of K following ME cultivation in HD soil. No studies currently indicate that there is a correlation between the abundance of *Bryobacter* and the accumulation of K in

FIGURE 6
The interaction network of soil bacterial genera in DT and HD under ME cultivation. **(A)** shows the network for DT soils, and **(B)** shows the network for HD soils. The nodes correspond to bacterial genera, with larger nodes indicating higher relative abundance. The color of the nodes ranges from red to blue, with red indicating greater connectivity and blue indicating lower connectivity. The lines indicate correlations, with red indicating positive correlations and blue indicating negative correlations.

soil. Thus, further exploration of the mechanism underlying the effects of this bacterial genus on the accumulation of K in soil is needed to maximize the utilization of this bacterial resource and enhance soil health. *Bacillus* can stimulate the dissolution of soil K and mediate the absorption of K by plants (Sheng, 2005; Ali et al., 2021); ME cultivation in cornfields might increase the content of soil K through the enrichment of *Bacillus*. *Streptomyces* is the most abundant and arguably the most important actinomycete genus; it is also a rich source of bioactive compounds, antibiotics, and extracellular enzymes (Olanrewaju and Babalola, 2019). *Streptomyces* might regulate the content of soil K through its interactions with soil. Members of the genus *Paenarthrobacter*

can play important roles in soil metabolism, as they can degrade various soil pollutants such as herbicides (Zhao et al., 2022). Its metabolic processes might also be related to the accumulation of K, but this possibility needs to be experimentally verified. In general, ME cultivation might affect the abundances of these bacterial genera, and these bacteria might play important roles in regulating soil chemical properties. Additional studies of the interaction between these bacteria and soil are needed. There is a special need for analyses of strains that play a major role in soil K accumulation in *Bryobacter, Bacillus, Streptomyces*, and *Paenarthrobacter*; isolation of these taxa could aid the development of new biofertilizers.

FIGURE 7

Analysis of correlations between soil chemical properties and bacterial diversity. **(A)** shows the environmental contributions of different factors in DT soil, and **(B)** shows the environmental contributions of different factors in HD soils. The horizontal axis shows the environmental factor, and the vertical axis shows the contribution. Higher percentages indicate greater effects of the environmental factor on the distribution of species. Percentages below 0 indicate no effect. **(C, D)** show the Pearson correlation between soil chemical factors and soil bacterial genera in DT and HD soils. The vertical axis shows soil bacterial genera, and the horizontal axis shows the different soil chemical factors. The numerical value indicates the R-value. "*" corresponds to $p < 0.05$; "**" corresponds to $p < 0.01$; and "***" corresponds to $p < 0.001$.

4 Conclusion

The effects of ME cultivation on the soil chemical properties and bacterial diversity in DT and HD soils were revealed in this study. The results of our study indicate that this cultivation model has the potential to be widely applied in light of its positive effects on the soil environment and microbial ecology.

Data availability statement

The original contributions presented in the study are publicly available. The raw amplicon sequencing dataset of metabarcoding is available in the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under BioSample accession number PRJNA1005455.

Author contributions

SY and HL contributed to conception and design of the study. MD, CY, LB, DH, HW, and YZ performed investigation. MD and CY performed the statistical analysis. MD wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1251695/full#supplementary-material

References

Ali, A. M., Awad, M. Y. M., Hegab, S. A., Gawad, A. M. A. E., and Eissa, M. A. (2021). Effect of potassium solubilizing bacteria (Bacillus cereus) on growth and yield of potato. *J. Plant Nutr.* 44, 411–420. doi:10.1080/01904167.2020.1822399

Babu, S., Mohapatra, K. P., Das, A., Yadav, G. S., Tahasildar, M., Singh, R., et al. (2020). Designing energy-efficient, economically sustainable and environmentally safe cropping system for the rainfed maize-fallow land of the Eastern Himalayas. *Sci. total Environ.* 722, 137874. doi:10.1016/j.scitotenv.2020.137874

Basu, S., Kumar, G., Chhabra, S., and Prasad, R. (2021). "Chapter 13 - role of soil microbes in biogeochemical cycle for enhancing soil fertility," in *New and future developments in microbial Biotechnology and bioengineering*. Editors J. P. Verma, C. A. Macdonald, V. K. Gupta, and A. R. Podile (Amsterdam, Netherlands: Elsevier), 149–157.

Brown, C., Murray-Rust, D., van Vliet, J., Alam, S. J., Verburg, P. H., and Rounsevell, M. D. (2014). Experiments in globalisation, food security and land use decision making. *PloS one* 9, e114213. doi:10.1371/journal.pone.0114213

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi:10.1038/nmeth.f.303

Chen, H., and Boutros, P. C. (2011). VennDiagram: A package for the generation of highly-customizable Venn and euler diagrams in R. *BMC Bioinforma.* 12, 35. doi:10.1186/1471-2105-12-35

Di Benedetto, N. A., Corbo, M. R., Campaniello, D., Cataldi, M. P., Bevilacqua, A., Sinigaglia, M., et al. (2017). The role of plant growth promoting bacteria in improving nitrogen use efficiency for sustainable crop production: A focus on wheat. *AIMS Microbiol.* 3, 413–434. doi:10.3934/microbiol.2017.3.413

Duan, M., and Bau, T. (2021). Initial sample processing can influence the soil microbial metabarcoding surveys, revealed by Leucocalocybe mongolica fairy ring ecosystem. *Biotechnol. Biotechnol. Equip.* 35, 1427–1438. doi:10.1080/13102818.2021.1996272

Duan, M., and Bau, T. (2021). Grassland fairy rings of Leucocalocybe mongolica represent the center of a rich soil microbial community. *Braz. J. Microbiol.* 52, 1357–1369. doi:10.1007/s42770-021-00478-3

Duan, M., Li, Y., Zhu, G., Wu, X., Huang, H., Qin, J., et al. (2023a). Soil chemistry, metabarcoding, and metabolome analyses reveal that a sugarcane—Dictyophora indusiata intercropping system can enhance soil health by reducing soil nitrogen loss. *Front. Microbiol.* 14, 1193990. doi:10.3389/fmicb.2023.1193990

Duan, M., Long, Y., Fan, H., Ma, L., Han, S., Li, S., et al. (2022c). Fenlong-ridging promotes microbial activity in sugarcane: A soil and root metabarcoding survey. *Agriculture* 12, 244. doi:10.3390/agriculture12020244

Duan, M., Lu, J., Yang, W., Lu, M., Wang, J., Li, S., et al. (2022b). Metabarcoding and metabolome analyses reveal mechanisms of leymus chinensis growth promotion by fairy ring of Leucocalocybe mongolica. *J. Fungi* 8, 944. doi:10.3390/jof8090944

Duan, M., Lu, M., Lu, J., Yang, W., Li, B., Ma, L., et al. (2022a). Soil chemical properties, metabolome, and metabarcoding give the new insights into the soil transforming process of fairy ring fungi Leucocalocybe mongolica. *J. Fungi* 8, 680. doi:10.3390/jof8070680

Duan, M., Wang, L., Song, X., Zhang, X., Wang, Z., Lei, J., et al. (2023b). Assessment of the rhizosphere fungi and bacteria recruited by sugarcane during smut invasion. *Braz. J. Microbiol.* 54, 385–395. doi:10.1007/s42770-022-00871-6

Edgar, R. C. (2013). Uparse: highly accurate OTU sequences from microbial amplicon reads. *Nat. methods* 10, 996–998. doi:10.1038/nmeth.2604

Food and Agriculture Organization (2007). *Determination of pH in soil*. Philippines: PRC National Standard.

Food and Agriculture Organization (2015). *Nitrogen determination methods of forest soils*. Philippines: PRC National Standard.

Food and Agriculture Organization (2006). *Soil testing. Part 6: Method for determination of soil organic matter*. Philippines: PRC National Standard.

Guo, M., Wu, F., Hao, G., Qi, Q., Li, R., Li, N., et al. (2017). Bacillus subtilis improves immunity and disease resistance in rabbits. *Front. Immunol.* 8, 354. doi:10.3389/fimmu.2017.00354

Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A., and Kuramae, E. E. (2016). The ecology of Acidobacteria: moving beyond genes and genomes. *Front. Microbiol.* 7, 744. doi:10.3389/fmicb.2016.00744

Koch, H., Lücker, S., Albertsen, M., Kitzinger, K., Herbold, C., Spieck, E., et al. (2015). Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11371–11376. doi:10.1073/pnas.1506533112

Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* 75, 5111–5120. doi:10.1128/AEM.00335-09

Li, L., Yang, X., Tong, B., Wang, D., Tian, X., Liu, J., et al. (2023). Rhizobacterial compositions and their relationships with soil properties and medicinal bioactive ingredients in Cinnamomum migao. *Front. Microbiol.* 14, 1078886. doi:10.3389/fmicb.2023.1078886

Li, Z., Zhang, K., Qiu, L., Ding, S., Wang, H., Liu, Z., et al. (2022). Soil microbial Co-occurrence patterns under controlled-release urea and fulvic acid applications. *Microorganisms* 10, 1823. doi:10.3390/microorganisms10091823

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

Lv, H., Ji, C., Zhang, L., Jiang, C., and Cai, H. (2022). Zinc application promotes nitrogen transformation in rice rhizosphere soil by modifying microbial communities and gene expression levels. *Sci. total Environ.* 849, 157858. doi:10.1016/j.scitotenv.2022.157858

Masuda, Y., Yamanaka, H., Xu, Z.-X., Shiratori, Y., Aono, T., Amachi, S., et al. (2020). Diazotrophic anaeromyxobacter isolates from soils. *Appl. Environ. Microbiol.* 86, 6–20. doi:10.1128/AEM.00956-20

Mujakić, I., Piwosz, K., and Koblížek, M. (2022). Phylum gemmatimonadota and its role in the environment. *Microorganisms* 10, 151. doi:10.3390/microorganisms10010151

Naz, H., Sayyed, R. Z., Khan, R. U., Naz, A., Wani, O. A., Maqsood, A., et al. (2023). Mesorhizobium improves chickpea growth under chromium stress and alleviates chromium contamination of soil. *J. Environ. Manage* 338, 117779. doi:10.1016/j.jenvman.2023.117779

Nesme, J., Achouak, W., Agathos, S. N., Bailey, M., Baldrian, P., Brunel, D., et al. (2016). Back to the future of soil metagenomics. *Front. Microbiol.* 7, 73. doi:10.3389/fmicb.2016.00073

New south wales (1988b). *Method for determination of total potassium in soils*. Philippines: PRC National Standard.

New south wales (1988a). *Method for etermination of soil total phosphorus*. Philippines: PRC National Standard.

Olanrewaju, O. S., and Babalola, O. O. (2019). Streptomyces: implications and interactions in plant growth promotion. *Appl. Microbiol. Biotechnol.* 103, 1179–1188. doi:10.1007/s00253-018-09577-y

Orlofsky, E., Zabari, L., Bonito, G., and Masaphy, S. (2021). Changes in soil bacteria functional ecology associated with Morchella rufobrunnea fruiting in a natural habitat. *Environ. Microbiol.* 23, 6651–6662. doi:10.1111/1462-2920.15692

Paul, N., Slathia, P. S., Vaid, A., and Kumar, R. (2018). Traditional knowledge of gucchi, Morchella esculenta (ascomycetes), in doda district, Jammu and Kashmir, India. *Int. J. Med. Mushrooms* 20, 445–450. doi:10.1615/IntJMedMushrooms.2018025995

Pertile, M., Sousa, R. M. S., Mendes, L. W., Antunes, J. E. L., Oliveira, L. M. d. S., de Araujo, F. F., et al. (2021). Response of soil bacterial communities to the application of the herbicides imazethapyr and flumyzin. *Eur. J. Soil Biol.* 102, 103252. doi:10.1016/j.ejsobi.2020.103252

Phanpadith, P., Yu, Z., Yu, D., Phongsavath, S., Shen, K., Zheng, W., et al. (2020). Promotion of maize growth by a yellow morel, Morchella crassipes. *Symbiosis* 80, 33–41. doi:10.1007/s13199-019-00651-1

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., et al. (2007). Silva: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 35, 7188–7196. doi:10.1093/nar/gkm864

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, R60. doi:10.1186/gb-2011-12-6-r60

Sheng, X. F. (2005). Growth promotion and increased potassium uptake of cotton and rape by a potassium releasing strain of Bacillus edaphicus. *Soil Biol. Biochem.* 37, 1918–1922. doi:10.1016/j.soilbio.2005.02.026

Song, H., Chen, D., Sun, S., Li, J., Tu, M., Xu, Z., et al. (2021). Peach-Morchella intercropping mode affects soil properties and fungal composition. *PeerJ* 9, e11705. doi:10.7717/peerj.11705

Sunil, C., and Xu, B. (2022). Mycochemical profile and health-promoting effects of morel mushroom Morchella esculenta (L.) – a review. *Food Res. Int.* 159, 111571. doi:10.1016/j.foodres.2022.111571

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi:10.1128/AEM.00062-07

Ward, T., Larson, J., Meulemans, J., Hillmann, B., Lynch, J., Sidiropoulos, D., et al. (2017). *BugBase predicts organism-level microbiome phenotypes.*133462

Yan, X., Yagi, K., Akiyama, H., and Akimoto, H. (2005). Statistical analysis of the major variables controlling methane emission from rice fields. *Glob. Change Biol.* 11, 1131–1141. doi:10.1111/j.1365-2486.2005.00976.x

Zhang, C., Shi, X., Zhang, J., Zhang, Y., and Wang, W. (2023). Dynamics of soil microbiome throughout the cultivation life cycle of morel (Morchella sextelata). *Front. Microbiol.* 14, 979835. doi:10.3389/fmicb.2023.979835

Zhang, H., Li, Y., and Zhu, J. K. (2018). Developing naturally stress-resistant crops for a sustainable agriculture. *Nat. plants* 4, 989–996. doi:10.1038/s41477-018-0309-4

Zhao, Y., Li, X., Li, Y., Bao, H., Nan, J., and Xu, G. (2022). Rapid biodegradation of atrazine by a novel Paenarthrobacter ureafaciens ZY and its effects on soil native microbial community dynamic. *Front. Microbiol.* 13, 1103168. doi:10.3389/fmicb.2022.1103168

Genomic mechanisms of plant growth-promoting bacteria in the production of leguminous crops

Afeez Adesina Adedayo [ID] and Olubukola Oluranti Babalola [ID] *

Food Security and Safety Focus Area, Faculty of Natural and Agricultural Sciences, North-West University, Mmabatho, South Africa

Legumes are highly nutritious in proteins and are good food for humans and animals because of their nutritional values. Plant growth-promoting bacteria (PGPR) are microbes dwelling in the rhizosphere soil of a plant contributing to the healthy status, growth promotion of crops, and preventing the invasion of diseases. Root exudates produced from the leguminous plants' roots can lure microbes to migrate to the rhizosphere region in other to carry out their potential activities which reveals the symbiotic association of the leguminous plant and the PGPR (rhizobia). To have a better cognition of the PGPR in the rhizosphere of leguminous plants, genomic analyses would be conducted employing various genomic sequences to observe the microbial community and their functions in the soil. Comparative genomic mechanism of plant growth-promoting rhizobacteria (PGPR) was discussed in this review which reveals the activities including plant growth promotion, phosphate solubilization, production of hormones, and plant growth-promoting genes required for plant development. Progress in genomics to improve the collection of genotyping data was revealed in this review. Furthermore, the review also revealed the significance of plant breeding and other analyses involving transcriptomics in bioeconomy promotion. This technological innovation improves abundant yield and nutritional requirements of the crops in unfavorable environmental conditions.

KEYWORDS

abiotic stress, genome editing, genotyping, PGPR interaction, plant breeding

Introduction

The reduction in food crop production nowadays results from the inability to tame wild species, the potential for average condition of weather, and population growth. The production of crops in sustainable agriculture is highly required to feed the population as it increases tremendously in this 21st century (Albahri et al., 2023). Promoting the quantity of crop production is crucial to preserve its nutritive value. Various challenges were encountered in the production of crops due to the application of chemical derivatives, which reveal the prevalence of chemical hazards in the environment (Nguyen et al., 2023). Thus, the a demand for sustainable means of promoting a friendly approach to improving the agricultural system. Integrating lesser crops, also known as underutilized crops, into potential food systems is highly significant. The process of breeding, including molecular and conventional approaches, has been utilized to improve the production of food crops, yet to meet up with food production is far fetch (Jabran et al., 2023).

Implementing genome sequences is significant for comprehending the biochemical and physiological methods of regulating plant stimulus to biotic and abiotic stresses and

plant traits (Raja and Vidya, 2023). The phylogenesis of sequence technologies has led to the production of voluminous plant genome data (Xia et al., 2023). This has produced the possibility of introducing technology in crop promotion via technological advances. In this modern technology of life science, novel observation has been found in clarifying genetic functions and their pathways in line with omic technologies and bioinformatics tools employed to analyze sequence data (Li et al., 2023a). The process has been utilized in genome-wide association studies (GWAS), genomic selection, quantitative linkage locus (QTLs), population genetics, genome editing, marker-assisted breeding, and SNP genotyping detection. The mentioned processes have been modified for the category of crops. Crops, including legumes (Adeleke and Babalola, 2022) and groundnuts (Ajilogba et al., 2022a) have been suggested to assist in the food security sector in sub-Saharan Africa. These crops contain proteins and minerals that make them essential in contributing to the healthy living of people feeding on them, thereby reducing malnutrition and food insecurity. Legumes are majorly cultivated in West African countries (Nigeria, Senegal, Cote d'Ivoire, Togo), and other eastern countries like India and Indonesia (Ajilogba et al., 2022a).

The crops are well known for their potential to fix nitrogen, improving soil fertility (Kebede, 2021). They are also known to undergo crop rotation and their ability to be cultivated without the application of chemical products (Pergner and Lippert, 2023). In rural and other isolated areas, the application of chemical fertilizers is not imminent. So, these crops were relied upon for the fixation of nitrogen into the soil. The richness in nutrients, the diversity of genetic resources, and mineral composition were some of the activities of the crop to improve the soil. Legumes nodulate with the aid of nitrogen-fixing bacteria and further promote nitrogen fixation in the soil (Cheng et al., 2023).

The diversification of crops via the involvement of underutilized crops especially legumes into significant cropping systems will promote food availability (Tan et al., 2020). However, the activities taking place in soil have revealed the immense impact on the production of crops and the effect of various environmental stresses. The rhizosphere soil of these crops contains some microbes that carry out the potential to improve the growth of plants either directly or indirectly. Among these microbes are bacteria that exist as beneficial or pathogenic. The beneficial types are plant growth-promoting microbes (PGPM) that contribute to plant growth and abundant production of crops. There is a symbiotic relationship between these soil microbes, especially rhizobacteria and leguminous plants (Fasusi et al., 2023). These rhizobacteria penetrate the root nodules of the leguminous plants and convert nitrogen from its free state into nitrate that the plant can utilize by nitrogen fixation process. This process assists legumes to flourish especially where nitrogen is deficient in the soil and likewise improves the soil by enhancing the health status of other plants while practicing crop rotation (Hyder et al., 2023). The potential of bacteria species in the rhizosphere consists of the advantages to humans like the significant function they perform in biogeochemical cycles of the major elements like carbon, nitrogen, and sulfur and minor elements including iron,

Nickel, and Mercury (Martínez-Martínez et al., 2023; Noor et al., 2023). The activities improve the engagement of microbes, including plant growth-promoting rhizobacteria (PGPR) and plant growth-promoting fungi (PGPF), in changeable energy embedded in the soil. Microbes manufacture phytohormones like Auxins and other substances like vitamins, enzymes, and other materials contributing to the growth of plants and inhibition of phytopathogens (Solomon et al., 2023). The communities of microbes in the rhizosphere are attributed to the features of soil, plant type, and plant growth stages.

The association of PGPR and leguminous crops has emphasized the potential of genomics. This gives tremendous insight into microbial functions and their genetic makeup. Researchers have determined the specific genes used in the process of nitrogen fixation and likewise the growth-promoting characteristics (Chieb and Gachomo, 2023; Fahde et al., 2023; Kim et al., 2023). Through the study of genomics, scientists identified and characterized major genes that partake in this mutualistic relationship of legumes and rhizobia. This study also presents the discovery of signaling molecules used in communication and recognition between rhizobia and legumes (Khan et al., 2023). A typical example is observed in the production of chemical substances called flavonoids by leguminous plant roots that attract rhizobia to the plant and nod factors, a chemical substance produced by the rhizobia activate the production of nodules in leguminous roots. In addition, genomics also unveil some PGPR which researcher analyze and make comparison of their genetic variation that contribute to plant growth-promotion potential. This knowledge does assist to optimize and choose various microbial strains that contribute to nitrogen fixation (Yao et al., 2023).Biotechnological and novel genomic processes in plants can provide solutions that can reveal the effectiveness of the objectives (Figueiredo et al., 2023). Novel techniques comprising biotechnology and the production of bio-based material can perform the function of improving agricultural sustainability. This is because they are safer for consumer use and for the environment. Biotechnology proffers resolutions for various problems of the modern world by inserting bioenergy, biopharmaceuticals, biofuels, biomaterials, foods, and its products for selling.

This review reveals the importance of plant genomics in bioeconomy promotion and the possibilities brought about by plant breeding techniques. Due to the challenges of underdevelopment and changes in the average weather conditions, the problem of food scarcity has not been resolved. So, the advancement of crop production suggests how to combat the problems. The review further reveals the advantages of the beneficial microbes in crop growth promotion and describes the application of omic techniques to characterize plant-microbe interaction.

Legume-PGPR interaction

Plants preserve the complex environment by stimulating root exudates into the soil environment (Paponov et al., 2023). Plants produce substances that are products of photosynthesis, especially in

the roots embedded in the soil. As a result, the region of soil containing carbon increases the population of microbes, unlike bulk soil which lacks carbon substances. A higher concentration of microbial communities in the soil is a result of interaction between plants and the roots that reveal various associations (Wang et al., 2023b). Some attributes measure the amount and grade of root exudates, plant growth stages (McLaughlin et al., 2023), abiotic stress (high temperature and drought) (Shaposhnikov et al., 2023), soil type (Wang et al., 2023a), phytopathogen invasion (Tortella et al., 2023), and the status of plant feeding (Heuermann et al., 2023). There will be a proposition of promoting beneficial microbes if the production of root exudates is feasible. Amino acids, sugar, phenolics, organic acids, etc., are some of the materials present in the root exudates (Agarwal et al., 2023). These carbon source substances draw the advantageous and nonadvantageous microbes into the rhizosphere soil region (Magnoli and Bever, 2023). The advantageous microbes keep plants safe against phytopathogens via direct and indirect mechanisms (Fahde et al., 2023). Yet, the carbon compounds likewise draw pathogenic microbes, which eventually lead to competition for nutrients, cause diseases in plants, and contribute to rhizobiome structuring (Adedayo and Babalola, 2023). Some plants have shown how microbial communities dwelling in their rhizosphere attract microbes that harbor protection on plants and inhibit pathogen invasions in the rhizosphere (Nwachukwu and Babalola, 2022). The physiochemical and biological features carry out significant functions in the association of plants and microorganisms. Even though phytopathogens have an immense effect on plant health, some beneficial microbes living in the rhizosphere or the tissues of plants contend with phytopathogens for nutrients and space, so end up in biocontrol activities beneficial against pathogenic organisms (Cao et al., 2023).

PGPR are found and multiply in the rhizosphere and tissues of healthy plants and improve plant growth, and abundant crop production, prevent plants against phytopathogens, and inhibit or eradicate biotic and abiotic stresses (Ojuederie et al., 2019; Olanrewaju et al., 2019). A direct form of microbial interaction with plant host is obtained in growth promotion, while the indirect form involves biological control potential against phytopathogens. Direct mechanisms involve phytohormone stimulation that reveals inhibition or promotion of the growth of plant roots, induces systemic resistance, promotes the acquisition of nutrients, and prevents crops from biotic and abiotic stress (Olanrewaju et al., 2017). Due to these properties, PGPR was accepted by researchers as an alternative to chemical fertilizers and other derivatives. Thus, PGPR, including *Bacillus thuringiensis*, contributes to abundant crop production and improves the health of plants, as reported by (Adeniji et al., 2021). Developing plants decrease the availability of Iron (Fe) by assimilating Fe and produce exudates that lure beneficial microbes into the soil to assimilate the Fe (Hyder et al., 2023). Certain PGPR is responsible for the production of siderophores in Fe-stressed environments and is enriched to inhibit phytopathogens via competition for Fe (Sarker et al., 2021). Plants utilize iron-bound siderophores that contribute to their growth (Swarnalatha et al., 2022). This reveals that the microorganisms producing Siderophore can avail the Fe needed by plants because Fe is a component of chlorophyll that is highly required for the photosynthesis process (Molnár et al., 2023).

Another form applied to the reduction of plant diseases is the capability of microbes inhabiting the rhizosphere to inhibit soil or compost from disallowing the invasion of phytopathogen.

Symbiontic association of leguminous plant and rhizosphere microbes

Rhizosphere microbes and leguminous plants have beneficial relationships (Figure 1). Their interaction inhibits or eradicates the invasion of the phytopathogens on the plants. Rhizobia species, including *Azorhizobium*, *Allorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, and *Sinorhizobium*, are common examples inhabiting the rhizosphere of leguminous plants (Babalola et al., 2017). The symbiotic association leads to the acquisition of nitrogen in the soil. Leguminous plants produce nodules and fix nitrogen acquired in the soil during the symbiotic association with *Bradyrhizobium* strains. The production of a nodule in legume plants while interacting with rhizosphere microbes leads to the production of some compounds like flavonoids, betaines, and aldonic acid in the form of root exudates. The exudates signal the PGPR in a mutual association that otherwise promotes the production of the nod gene that stimulates nodulation by associating with the nod proteins present in the PGPR cell wall (Wang et al., 2018; Walker et al., 2020). In the root hair of leguminous plants, the physical effects were revealed when the PGPR (rhizobia) responded to the production of oligosaccharide nod factors. This results in the production of nodules that eventually promote nitrogen fixation. PGPR, especially rhizobia, contributes to plant growth and improves the development of seed by producing nod factors.

Non-symbiotic association of leguminous plants and rhizosphere microbes

Another form of association undergone by leguminous plants is the stimulation of phytohormones by the PGPR and the promotion of plant growth and abundant production of crops. Besides, leguminous plants do produce phenolics that aid the inhibition of phytopathogens, provision of nutrients to plants, and beneficial microbial growth (Egamberdieva et al., 2023). Phenolic compounds produced arbitrate the nod gene elicitors production. Nod gene elicitors control the stimulation of nod factors on the plant roots (Ahmed et al., 2020). After being collected in the rhizosphere, it results in the biological synthesis of flavonoids, which leads to the production of a higher concentration of phytoalexins for plant protection against spoilage organisms, and more reported in Table 1. Riboflavin (Vitamin B2) can likewise be produced by rhizobia and change photochemically (Lopez et al., 2019). Rhizobia has been reported to inhibit the growth of bacterial and fungal spoilage organisms affecting other crops, including soybean (Paravar et al., 2023), okra (Sharma, 2019), tomato (Adedayo et al., 2023a; Adedayo et al., 2023b), sunflower (Adeleke et al., 2021), Bambara groundnut (Ajilogba et al., 2022c), etc. Root exudates, including organic acid anions and phytosiderophores, are significant in the availability of minerals and their presence in agricultural soil (Wang et al., 2023c).

FIGURE 1
Interaction of rhizobia with leguminous plants.
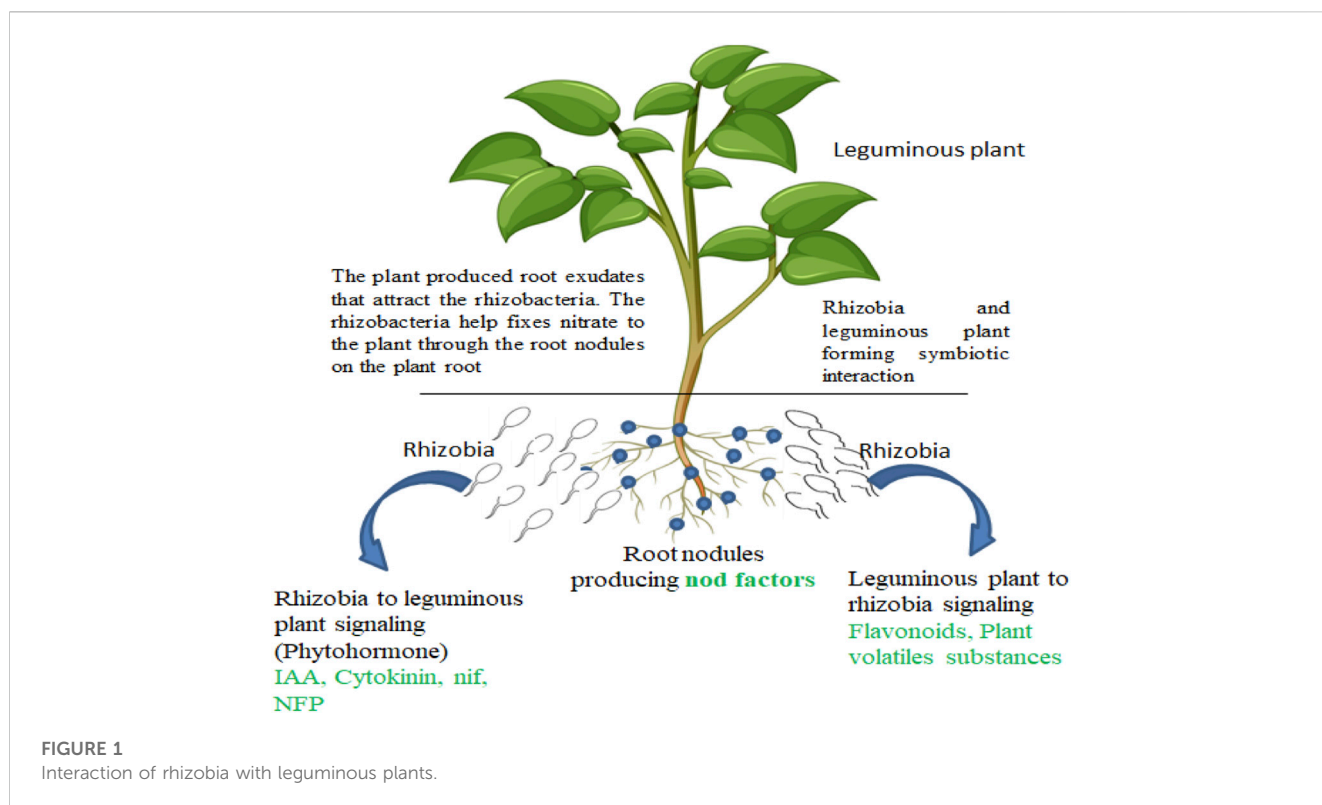
Molecular mechanisms applied in the study of PGPR

Certain biochemical and cultured-based tests have been employed for a long time to study PGPR (Utami et al., 2022). With recent progression in molecular technology, various gene technological data, including genomics, metagenomics, transcriptomics, and proteomics data, are uploaded on global websites (Lata et al., 2023). Genomic sequence analysis of PGPR is grouped into two groups, and they are whole genome sequencing (WGS) and specific genome sequencing (SGS) (Simlat et al., 2023).

WGS reveals all the chromosomes and plasmids present in the PGPR. WGS-employed next-generation sequencing (NGS) shows the appropriate genetic makeup of organisms (Kurotani et al., 2023). Whereas SGS revealed only a certain part of the genome employed for comparison and characterization (Lim et al., 2023). The particular gene in this group is 16S rRNA used to identify bacteria origin to specific levels of organization (Abramova et al., 2023). Another form of identifying the bacteria is an intra-genomic spacer (IGS) 16S-23S, repetitive sequence-based PCR (rep-PCR) that amplify and sequence highly conserved repeats. Restriction fragment length polymorphism (RLFP) analysis has been used to categorize PGPR (Li et al., 2023b).

Some genes, including *atpD*, *dnaK*, *dnaJ*, *gap*, *glnll*, *gltA*, *gyrG*, *pnp*, *recA*, *rpoA*, and *thrC* were employed in strain typing involved in multilocus sequence analysis (Swarnalakshmi et al., 2020). Absolute primers are employed to intensify genomic sequences randomly, known as random amplified polymorphic DNA sequencing. To amplify particular genes together with a secondary mode of sequencing using restriction profiling and denaturing gradient gel electrophoresis (DGGE) to reduce the assembly of mismatches in limited regions or variation in genome GC content (Nakatsu, 2021). The process of fixing nitrogen in the soil by the biological process is a significant method of promoting PGPR with the aid of nitrogenase enzymes producing genes like *nifD*, *nifK*, and *nifH* genes (Özdoğan et al., 2022).

Moreover, PGPR partakes in the solubilization of phosphorus in the production of gluconic acid, which makes use of glucose dehydrogenase. Apart from the mentioned characteristics, PGPR stimulates phytohormones like gibberellin, ethylene, indole acetic acid (IAA), abscisic acid, and cytokinin (Ojuederie et al., 2019; Koza et al., 2022). However, certain genes like *ipdC* and *amlE* are responsible for IAA production, as found in other phytohormones (Adedayo et al., 2023a). For iron utilization, siderophores were produced by PGPR which is regulated by the Siderophore gene. *Pseudomonas* spp. possess membrane receptor that codes for the *pupA* gene, which transports the Siderophore complex in the cell (Sah et al., 2021).

Plant growth-promoting genes contributing to the growth of leguminous plant

Leguminous plants have a special interaction with PGPR especially rhizobia as discussed above. This interaction yields the fixation of atmospheric nitrogen to a plant-utilizable form that contributes to the healthy status of the plant, which is important for plants that have no access to nitrate assimilation (Figure 1). The growth-promoting genes are mainly associated with the production and upkeep of mutualistic association between rhizobia and leguminous plants and the following genes discussed below are responsible for this process.

TABLE 1 Microbial beneficial effects on leguminous crops.

Leguminous crops	Microbes	Activities	References
Vigna subterranean	*Pseudomonas* spp.	Contributes to plant growth and prevents disease invasion	Ajilogba et al. (2022a)
Vigna subterranean and *Arabidopsis thaliana*	*Rhizobium*, *Pseudomonas*, *Azospirillum* and *Bacillus*	Promote the abundant production of legumes and other crops by fixing nitrogen into the soil	Igiehon and Babalola (2018)
Phaseolus vulgaris	*Kosakonia cowanii*, *Pantoea agglomerans*, *Variovorax paradoxus*, *Staphylococcus aureus*, *Enterobacter ludwigii*, and *Bacillus pumilus*	The endophytic microbes help fix plant nutrients, thereby enhancing the growth of plants	Ayilara et al. (2023)
Vigna unguiculata and *Arachis hypogaea*	*Bacillus* spp., *Pseudomonas* spp., *Streptomyces* spp.	The soil microbes are involved in nutrient mineralization, and so promote plant growth	Aremu et al. (2017)
Vigna unguiculata, Phaseolus vulgaris, Arachis hypogaea	PGPR	Aid biological nitrogen fixation that enhances the growth of leguminous crops	Babalola et al. (2017)
	Bacillus spp., *Pseudomonas* spp.		
Vigna subterranean	*Bradyrhizobium japonicum*	The PGPR contributes to the abundant production of groundnut compared to N-fertilizer	Bitire et al. (2022)
Vigna subterranea	*Stenotrophomonas maltophilia*, *Chryseobacterium* spp., *Pseudomonas alcaligenes*, *Pseudomonas plecoglossicida*, and *Pseudomonas hibiscicola*	These microbes have the capacity to produce indole acetic acid (IAA) and the potential to solubilize phosphate	Ayangbenro et al. (2023)
Vigna subterranea	*Rubrobacter*, *Acidobacterium*, and *Skermanella*	The diversity and functions of the microbes attributed to the growth of the crop	Ajilogba et al. (2022b)
Vigna subterranean	*Rhizobia*	The microbes promote the potential of the crops to resist abiotic conditions, prevent disease invasion, and ensure the fixation of nitrogen	Ayilara et al. (2022)
Vigna subterranean	*Bacillus*, *Pseudomonas*, and *Streptomyces*	The microbes improve crop growth and prevent the invasion of diseases on the plant	Olanrewaju and Babalola (2022)
Vigna subterranean	*Bradyrhizobium japonicum*	The PGPR promotes nitrogen fixation and further improves plant growth	Bitire et al. (2023)
Glycine max	*Rhizobium cellulosilyticum* and *Arbuscular mycorrhizal*	The microbes contribute to the growth of the plant in the semi-desert region	Igiehon et al. (2021)
Vigna subterranea	*B. amyloliquefaciens*, *B. thuringiensis* and *Bacillus* sp.	The PGPR contribute to the production of volatile organic compound that acts as antibacterial against phytopathogens	Ajilogba and Babalola (2019)
Vigna subterranean, Glycine max, Vigna unguiculata, Phaseolus vulgaris, Arachis hypogaea	PGPR and PGPR traits	The rhizobacteria enhance plant health by promoting plant growth, preventing the occurrence of phytopathogens through biological control activities, and improving soil health	Hassen et al. (2023)

Nodulation (nod) genes: are the genes used for starting and regulating the production of nodules which is the complex body where the organism (rhizobia) dwells (Malik et al., 2020). The genes examples are *nodA, nodB, nodC,* and *nodD*.

Nitrogen fixation (nif) genes: they are involved in the production and roles of nitrogen-fixing nodules. They also carry out potential action in synthesizing nitrogenase enzyme that converts natural nitrogen gas to ammonia and their example are *nifD, nifH,* and *nifK* (Adedayo et al., 2023a)

Auxin biosynthetic genes: are genes responsible for the production of hormones that contribute to the growth of plants. The stimulation of these hormones is promoted by some rhizobia thereby contributing to the elongation of plant roots and production of the root nodules. The typical examples include indole-3-acetamide hydrolase (*iaah*), an enzyme responsible for the production of auxin, and indole acetic acid (*IAA*) which is a hormone (Adedayo et al., 2023a).

Cytokinin biosynthesis genes: these are responsible for the production of plant hormone Cytokinin. The hormone contributed to the division of cells as well as their differentiation (Šmeringai et al., 2023). Various rhizobia are producing this hormone that contributes to the production of nodules. These genes also contributed to the biological synthesis of cytokines with the aid of the isopentenyl transferase (ipt) enzyme for plant growth.

Plant receptor genes: these are genes constituting particular receptors to identify and associate with the signals produced by rhizobia. Examples include nod factor rhizobia did produces LysM Receptor Kinase (*NFR*) and Nod Factor Perception (*NFP*) genes (Geurts and Huisman, 2023). The genes explained above contribute to structural signaling pathways and their association to produce advantageous and mutualistic associations between leguminous plants and PGPR (rhizobia). They also enhance the development of the plants by supplying them with fixed nitrogen and assimilation of nutrients in the rhizosphere soil.

Plant breeding

The integration of plant and animal production systems to promote environmental quality and natural resources to improve economic viability is defined as sustainable agriculture (Fatima et al., 2023; Liu, 2023). It is also defined as the application of technologies and practices that do not harm the food quality, environment, and production (Gamage et al., 2023). The significant issue of minimizing the application of non-renewable material, e.g., phosphates, crop genotypes, and promoting sustainability of agriculture (Jha et al., 2023). The utilization of modern methods of genetic breeding, manipulation and redesign, and agricultural and ecological management incorporating crop selection is sustainable development (Hemalatha et al., 2023). Agricultural practices nowadays are based on different forms that contain input-intensive systems that make use of energy, fertilizer, oil, water, etc., unlike low-input systems, which bring about low production of crops.

So, scientists' stance should focus on plant breeding which will cut down the rate at which energy is utilized for plant production yet ascertain the production of high-quality food products. Due to climate change and the growth of the population around the globe, plant breeding requires water and, likewise, reduces the rate of water loss in other to promote the potential of sustaining water deprivation and avail food for developing populations. To promote the effectiveness of water use and production of abundant yield, the following features are required, and they include; physiological and molecular modification, osmoregulation, improved plant root system, and improved stomatal conductivity (Woźniak-Gientka and Tyczewska, 2023). Energy in the form of oil and gas is important for modern-day high-input agricultural systems. However, the energy implored for the production of crops during the period of harvest and post-harvest incorporates the application of farm machinery for cultivation, preparation of soil, harvesting, transport of the yield, and post-harvest storage. Although, it has been noted that 15% of utilized energy is demanded by the application of chemical derivatives on farmland and transportation, according to Attfield et al. (2023). The application of chemical fertilizers is regarded as not the best means when studying sustainable agriculture. It is significant to note the breeding process to improve the resistivity and tolerance in agricultural practices to boost sustainability, where the application of chemical pesticides is reduced (Esuola et al., 2016).

Various scientific processes that aid plant feature identification include genetic engineering, hybridization, tissue culture, and plant breeding (Woźniak-Gientka and Tyczewska, 2023). The benefit of genome editing (CRISPR-Cas9) is to supply high target specificity with high precision and speed (Sahel et al., 2023). Furthermore, applying this technique does not yield transgenic plants due to its inability to transfer genes to the host plants. Applying this technique, some features, including biotic and abiotic features, were tolerated to obtain new varieties of crops to improve the abundance of agricultural products (Anguita-Maeso et al., 2023). The process involved in agrobiotechnology is employed to produce Biofertilizers, biopesticides, bioherbicides, etc., in the agricultural field (Woźniak-Gientka and Tyczewska, 2023). These compounds aid in the regulation of soil toxicity and are safe for environmental use. The growth and development of plants mainly depend on the availability of macroelements, including nitrogen (N), phosphorus (P), and potassium (K) (Bitire et al., 2023), and microelements like molybdenum (Mo), iron (Fe), zinc (Zn), manganese (Mn), copper (Cu), boron (B), etc. (Banerjee et al., 2023).

Plant breeding techniques improve the production of crops

The challenge of making available food crops to feed the developing population is the economic, social, and environmental systems of farming that oppose the adaptation of climate, degradation of the agricultural field, and loss invented on biodiversity of the microbial communities. The production of food to support the population is not sufficient as the population increases. Due to this challenge, some procedures were adopted in new breeding techniques to enhance the production of crops through genetically modified organisms (GMOs) (Rahman et al., 2023). However, the usual form of biotechnology in agricultural practices is the cultivation of new crops, also called genetically modified or engineered crops. Some of the genome-edited crops do not shelter foreign DNA (Olanrewaju et al., 2021). CRISPR-Cas9 is a genome-editing tool employed in plant study and its growth globally to improve the features of crop quality, activities against chemical derivatives like pesticides and herbicides, and disease resistance (Marone et al., 2023). Employing a CRISPR technology breeder could promote tolerance to abiotic stresses like extreme temperature and drought and improve nutrient content. Legumes have been monitored upon the application of this technology (Puppala et al., 2023). Climate change has caused extreme heat and drought that affect the production of crops to feed the global populace (Omotoso et al., 2023). CRISPR/Cas9 tools have been employed in crop plantation that requires water for their growth and cognition of significant genes to control drought stress (Banerjee et al., 2023).

Transcriptomics analysis of PGPR dwelling in the rhizosphere of leguminous plants

This involves the genetic expression study pattern of the rhizobia to obtain knowledge about their interaction and their roles in the rhizosphere soil. PGPR are known to be involved in the significant function of improving the growth of plants and likewise promoting the availability of the nutrients required via nitrogen fixation and solubilization of phosphate (Adedayo et al., 2022).

To carry out the analysis of transcriptomes on samples, scientists isolate the specific bacteria species in the soil sample of the leguminous plants. The microbes were lysed to extract the RNA from the bacterial cell and further purified (Saleh-Lakha et al., 2005). The RNA is a precursor for transcribed genes for a specific period and as a starting medium for downstream analysis. Then, the mRNA is converted to complementary strand DNA (cDNA) with the help of an enzyme called reverse transcriptase (Ying, 2003). The cDNA stands for the expressed gene of the microbial isolate.

High throughput sequencing technologies like RNA Seq can be used to collect sequence data of cDNA molecules and further subjected to analyzing the data using bioinformatics tools (Shiau et al., 2023). This constitutes aligned sequences to the reference database or genome of the recognized genes to confirm the genes expressed (Tables 2, 3). Also, the analysis of differential gene expression can be investigated to analyze various conditions of the stages of gene expression, like

TABLE 2 Plant growth-promoting potential of *Paenibacillus polymyxa* contributing to the growth of crop plants like tomatoes (Xie et al., 2016).

Growth parameter	Control[a]	1–18	1–49	1–43	TD94	WLY78
Shoot length (cm)	22.76 (±1.74)	28.16 (±1.13)**	27.57 (±2.37)**	26.34 (±0.49)**	26.13 (±0.62)**	26.43 (±1.07)**
Root length (cm)	12.93 (±1.93)	13.17 (±1.59)	13.85 (±1.18)	13.02 (±0.75)	13.63 (±0.77)	13.03 (±0.75)
Fresh weight (g)	2.68 (±0.71)	4.05 (±0.43)**	3.96 (±0.62)	3.68 (±0.45)**	3.57 (±0.26)**	3.50 (±0.32)**
Dry weight (g)	0.19 (±0.05)	0.31 (±0.03)**	0.30 (±0.05)**	0.29 (±0.02)**	0.29 (±0.02)**	0.27 (±0.02)**

[a]Signifies tomato plant.
**Reveals the level of significance ($p > 0.01$).

TABLE 3 Comparative genomic analysis of *Pantoea agglomerans* strain as plant growth-promoting rhizobacteria according to Shariati et al. (2017).

Strains	Genome size (Mb)	GC (%)	No. of rRNA	No. of tTRNA	Scaffolds	No. of genes	No. of proteins	GenBank Accession no.
P5	5.16726	55.4	7	63	150	4,745	4,674	GCA_002157425.2
Tx10	4.85699	55.1	48	142	22	4,627	4,500	GCA_000475055.1
Eh318	5.03584	54.8	28	73	34	4,791	4,627	GCA_000687245.1
190	5.00257	55.1	24	77	5	4,878	4,778	GCA_000731125.1
MP2	4.73383	55.2	22	71	16	4,352	4,214	GCA_000757415.1
IG1	4.83958	55.0	2	63	18	4,443	4,341	GCA_000241285.2
299 R	4.58148	54.3	27	63	109	4,267	4,157	GCA_000330765.1
RIT273	5.36534	55.1	17	76	26	4,999	4,914	GCA_000627115.1
DAPP-PG734	5.36593	54.7	36	70	195	5,107	5,107	GCA_000710215.1
4	4.8102	55.1	19	72	12	4,370	4,260	GCA_000743785.1
LMAE-2	4.98116	55.1	51	79	155	4,681	4,406	GCA_000814075.1

different rhizobia strains, rhizosphere soil they inhabit, and time range. The differentially identified genes can make available the knowledge of molecular procedures of rhizobacteria's mutualistic association with the leguminous plants (Li et al., 2023c). These genes can contribute to the following mechanisms including; plant growth promotion, nitrogen fixation, stress tolerance, and nutrient cycling, among other functional characteristics of interest. This analysis can likewise be joined together with other high throughput omics technologies like metabolomics and proteomics to get more knowledge of PGPR community structure, functions, and potential influence on abundant production of legumes and improvement of leguminous plant health. On a short note, transcriptomics technology can assist scientists in decrypting the genetic makeup of plant-microbe interactions and make available cogent information for sustainable agriculture. The characteristics of *Paenibacillus polymyxa* (Table 2) and *P. agglomerans* (Table 3) were revealed showing various characteristics of the PGPR ranging from growth features to genomic size of various strains strain isolated. *Pantoea agglomerans* displayed various potential including plant growth promotion, phosphate solubilizing potential, secretion of hormones like IAA and Siderophore, and synthesis of acetoin and 2,3-butanediol according to Shariati et al. (2017).

Conclusion

The sustainability of agricultural soil concerns the agricultural development and application of fertilizers to improve farming activities. Major nutrients required to boost the fertility of the soil and encourage plant growth, including nitrogen, phosphorus, and potassium, were discussed in the context. Nevertheless, PGPR contributes to the health status of plants, thereby contributing to the healthy growth of plants and improving the abundant production of crops and generally regarded as safe for humans, plants, and the environment. Due to these features, they possess, they are alternatives to chemical fertilizers. The diversity of microorganisms in soil contributes to plant growth and improves their various ecosystem functions in the soil. The plant root produces some substances called root exudate that attract soil microbes and reveal the microbial diversity inhabiting the soil. So, high throughput sequencing involving genomics is used to identify unculturable microbes. However, some factors, including soil types and plant genotypes, control the activities of PGPR in the soil. So, genetic variation of advantageous PGPR interaction with plants can be incorporated into plant breeding to boost the production of legumes for agricultural sustainability.

Author contributions

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abramova, A., Berendonk, T. U., and Bengtsson-Palme, J. (2023). A global baseline for qPCR-determined antimicrobial resistance gene prevalence across environments. *Environ. Int.* 178, 108084. doi:10.1016/j.envint.2023.108084

Adedayo, A. A., and Babalola, O. O. (2023). Fungi that promote plant growth in the rhizosphere boost crop growth. *J. Fungi* 9, 239. doi:10.3390/jof9020239

Adedayo, A. A., Babalola, O. O., Prigent-Combaret, C., Cruz, C., Stefan, M., Kutu, F., et al. (2022). The application of plant growth-promoting rhizobacteria in *Solanum lycopersicum* production in the agricultural system: a review. *PeerJ* 10, e13405. doi:10.7717/peerj.13405

Adedayo, A. A., Fadiji, A. E., and Babalola, O. O. (2023a). Unraveling the functional genes present in rhizosphere microbiomes of *Solanum lycopersicum*. *PeerJ* 11, e15432. doi:10.7717/peerj.15432

Adedayo, A. A., Romdhane, S. B., and Babalola, O. O. (2023b). Metagenomic study of the tomato rhizosphere soil microbiome: current state and prospect. *AIMS Geosci.* 9, 330–356. doi:10.3934/geosci.2023018

Adeleke, B. S., Ayangbenro, A. S., and Babalola, O. O. (2021). Bacterial community structure of the sunflower (*Helianthus annuus*) endosphere. *Plant Signal. Behav.* 16, 1974217. doi:10.1080/15592324.2021.1974217

Adeleke, B. S., and Babalola, O. O. (2022). Meta-omics of endophytic microbes in agricultural biotechnology. *Biocatal. Agric. Biotechnol.* 42, 102332. doi:10.1016/j.bcab.2022.102332

Adeniji, A. A., Ayangbenro, A. S., and Babalola, O. O. (2021). Genomic exploration of *Bacillus thuringiensis* MORWBS1.1, candidate biocontrol agent, predicts genes for biosynthesis of zwittermicin, 4,5-DOPA dioxygenase extradiol, and quercetin 2,3-dioxygenase. *Mol. Plant-Microbe Interactions®* 34, 602–605. doi:10.1094/mpmi-10-20-0272-sc

Agarwal, P., Vibhandik, R., Agrahari, R., Daverey, A., and Rani, R. (2023). Role of root exudates on the soil microbial diversity and biogeochemistry of heavy metals. *Appl. Biochem. Biotechnol.* 1–21. doi:10.1007/s12010-023-04465-2

Ahmed, A., Tariq, A., and Habib, S. (2020). "Interactive biology of auxins and phenolics in plant environment," in *Plant phenolics in sustainable agriculture: volume 1*. Editors R. Lone, R. Shuab, and A. N. Kamili (Singapore: Springer), 117–133. doi:10.1007/978-981-15-4890-1_5

Ajilogba, C. F., and Babalola, O. O. (2019). GC–MS analysis of volatile organic compounds from Bambara groundnut rhizobacteria and their antibacterial properties. *World J. Microbiol. Biotechnol.* 35, 83. doi:10.1007/s11274-019-2660-7

Ajilogba, C. F., Olanrewaju, O. S., and Babalola, O. O. (2022a). Improving Bambara groundnut production: insight into the role of omics and beneficial bacteria. *Front. Plant Sci.* 13, 836133. doi:10.3389/fpls.2022.836133

Ajilogba, C. F., Olanrewaju, O. S., and Babalola, O. O. (2022b). Plant growth stage drives the temporal and spatial dynamics of the bacterial microbiome in the rhizosphere of *Vigna subterranea*. *Front. Microbiol.* 13, 825377. doi:10.3389/fmicb.2022.825377

Ajilogba, F. C., Babalola, O. O., Adebola, P., and Adeleke, R. (2022c). Bambara groundnut rhizobacteria antimicrobial and biofertilization potential. *Front. Plant Sci.* 13, 854937. doi:10.3389/fpls.2022.854937

Albahri, G., Alyamani, A. A., Badran, A., Hijazi, A., Nasser, M., Maresca, M., et al. (2023). Enhancing essential grains yield for sustainable food security and bio-safe agriculture through latest innovative approaches. *Agronomy* 13, 1709. doi:10.3390/agronomy13071709

Anguita-Maeso, M., Navas-Cortés, J. A., and Landa, B. B. (2023). Insights into the methodological, biotic and abiotic factors influencing the characterization of xylem-inhabiting microbial communities of olive trees. *Plants* 12, 912. doi:10.3390/plants12040912

Aremu, B. R., Alori, E. T., Kutu, R. F., and Babalola, O. O. (2017). "Potentials of microbial inoculants in soil productivity: an outlook on african legumes," in *Microorganisms for green revolution: volume 1: microbes for sustainable crop production*. Editors D. G. Panpatte, Y. K. Jhala, R. V. Vyas, and H. N. Shelat (Singapore: Springer), 53–75. doi:10.1007/978-981-10-6241-4_3

Attfield, P. V., Bell, P. J. L., and Grobler, A. S. (2023). Reducing carbon intensity of food and fuel production whilst lowering land-use impacts of biofuels. *Fermentation* 9, 633. doi:10.3390/fermentation9070633

Ayangbenro, A. S., Adem, M. R., and Babalola, O. O. (2023). Bambara nut root-nodules bacteria from a semi-arid region of South Africa and their plant growth-promoting traits. *Int. J. Microbiol.* 2023, 8218721. doi:10.1155/2023/8218721

Ayilara, M. S., Abberton, M., Oyatomi, O. A., Odeyemi, O., and Babalola, O. O. (2022). Potentials of underutilized legumes in food security. *Front. Soil Sci.* 2, 1–12. doi:10.3389/fsoil.2022.1020193

Ayilara, M. S., Adeleke, B. S., and Babalola, O. O. (2023). Bioprospecting and challenges of plant microbiome research for sustainable agriculture, a review on soybean endophytic bacteria. *Microb. Ecol.* 85, 1113–1135. doi:10.1007/s00248-022-02136-z

Babalola, O. O., Olanrewaju, O. S., Dias, T., Ajilogba, C. F., Kutu, F. R., and Cruz, C. (2017). "Biological nitrogen fixation: the role of underutilized leguminous plants," in *Microorganisms for green revolution: volume 1: microbes for sustainable crop production*. Editors D. G. Panpatte, Y. K. Jhala, R. V. Vyas, and H. N. Shelat (Singapore: Springer), 431–443. doi:10.1007/978-981-10-6241-4_20

Banerjee, S., Roy, P., Nandi, S., and Roy, S. (2023). Advanced biotechnological strategies towards the development of crops with enhanced micronutrient content. *Plant Growth Regul.* 100, 355–371. doi:10.1007/s10725-023-00968-4

Bitire, T. D., Abberton, M., Oyatomi, O., and Babalola, O. O. (2022). Effect of *Bradyrhizobium japonicum* strains and inorganic nitrogen fertilizer on the growth and yield of Bambara groundnut (*Vigna subterranea (L.) verdc*) accessions. *Front. Sustain. Food Syst.* 6, 1–12. doi:10.3389/fsufs.2022.913239

Bitire, T. D., Abberton, M., Oyatomi, O., and Babalola, O. O. (2023). Yield response of accessions of Bambara groundnut (*Vigna subterranea (L) Verdc*) inoculated with *Bradyrhizobium japonicum* strains. *Front. Sustain. Food Syst.* 7, 1–18. doi:10.3389/fsufs.2023.1142123

Cao, M., Narayanan, M., Shi, X., Chen, X., Li, Z., and Ma, Y. (2023). Optimistic contributions of plant growth-promoting bacteria for sustainable agriculture and climate stress alleviation. *Environ. Res.* 217, 114924. doi:10.1016/j.envres.2022.114924

Cheng, Z., Meng, L., Yin, T., Li, Y., Zhang, Y., and Li, S. (2023). Changes in soil rhizobia diversity and their effects on the symbiotic efficiency of soybean intercropped with maize. *Agronomy* 13, 997. doi:10.3390/agronomy13040997

Chieb, M., and Gachomo, E. W. (2023). The role of plant growth promoting rhizobacteria in plant drought stress responses. *BMC Plant Biol.* 23, 407. doi:10.1186/s12870-023-04403-8

Egamberdieva, D., Eshboev, F., Shukurov, O., Alaylar, B., and Arora, N. K. (2023). Bacterial bioprotectants: biocontrol traits and induced resistance to phytopathogens. *Microbiol. Res.* 14, 689–703. doi:10.3390/microbiolres14020049

Esuola, C. O., Babalola, O. O., Heine, T., Schwabe, R., Schlömann, M., and Tischler, D. (2016). Identification and characterization of a FAD-dependent putrescine N-hydroxylase (GorA) from *Gordonia rubripertincta* CWB2. *J. Mol. Catal. B Enzym.* 134, 378–389. doi:10.1016/j.molcatb.2016.08.003

Fahde, S., Boughribil, S., Sijilmassi, B., and Amri, A. (2023). Rhizobia: a promising source of plant growth-promoting molecules and their non-legume interactions: examining applications and mechanisms. *Agriculture* 13, 1279. doi:10.3390/agriculture13071279

Fasusi, O. A., Babalola, O. O., and Adejumo, T. O. (2023). Harnessing of plant growth-promoting rhizobacteria and arbuscular mycorrhizal fungi in agroecosystem sustainability. *CABI Agric. Biosci.* 4, 26–15. doi:10.1186/s43170-023-00168-0

Fatima, A., Singh, V. K., Babu, S., Singh, R. K., Upadhyay, P. K., Rathore, S. S., et al. (2023). Food production potential and environmental sustainability of different integrated farming system models in northwest India. *Front. Sustain. Food Syst.* 7, 959464. doi:10.1007/s10508-022-02464-8

Figueiredo, M., Dias, A., Neves, J., and Vicente, H. (2023). Assessment of literacy to biotechnological solutions for environmental sustainability in Portugal. *Sustainability* 15, 10056. doi:10.3390/su151310056

Gamage, A., Gangahagedara, R., Gamage, J., Jayasinghe, N., Kodikara, N., Suraweera, P., et al. (2023). Role of organic farming for achieving sustainability in agriculture. *Farming Syst.* 1, 100005. doi:10.1016/j.farsys.2023.100005

Geurts, R., and Huisman, R. (2023). Innovations in two genes kickstarted the evolution of nitrogen-fixing nodules. *Curr. Opin. Plant Biol.* 9, 102446. doi:10.1016/j.pbi.2023.102446

Hassen, A. I., Babalola, O. O., and Carlson, R. (2023). "Rhizobacterial-Mediated interactions for enhanced symbiotic performance of the root nodule rhizobia in legumes," in *Sustainable agrobiology: design and development of microbial consortia*. Editors D. K. Maheshwari and S. Dheeman (Singapore: Springer Nature Singapore), 41–56. doi:10.1007/978-981-19-9570-5_3

Hemalatha, P., Abda, E. M., Shah, S., Venkatesa Prabhu, S., Jayakumar, M., Karmegam, N., et al. (2023). Multi-faceted CRISPR-Cas9 strategy to reduce plant based food loss and waste for sustainable bio-economy – a review. *J. Environ. Manag.* 332, 117382. doi:10.1016/j.jenvman.2023.117382

Heuermann, D., Döll, S., Schweneker, D., Feuerstein, U., Gentsch, N., and Von Wirén, N. (2023). Distinct metabolite classes in root exudates are indicative for field- or hydroponically-grown cover crops. *Front. Plant Sci.* 14, 1122285. doi:10.3389/fpls.2023.1122285

Hyder, S., Rizvi, Z. F., Los Santos-Villalobos, S. D., Santoyo, G., Gondal, A., Khalid, N., et al. (2023). Applications of plant growth-promoting rhizobacteria for increasing crop production and resilience. *J. Plant Nutr.* 46, 2551–2580. doi:10.1080/01904167.2022.2160742

Igiehon, N. O., and Babalola, O. O. (2018). Rhizosphere microbiome modulators: contributions of nitrogen fixing bacteria towards sustainable agriculture. *Int. J. Environ. Res. Public Health* 15, 574. doi:10.3390/ijerph15040574

Igiehon, N. O., Babalola, O. O., Cheseto, X., and Torto, B. (2021). Effects of Rhizobia and Arbuscular mycorrhizal fungi on yield, size distribution and fatty acid of soybean seeds grown under drought stress. *Microbiol. Res.* 242, 126640. doi:10.1016/j.micres.2020.126640

Jabran, M., Ali, M. A., Zahoor, A., Muhae-Ud-Din, G., Liu, T., Chen, W., et al. (2023). Intelligent reprogramming of wheat for enhancement of fungal and nematode disease resistance using advanced molecular techniques. *Front. Plant Sci.* 14, 1132699. doi:10.3389/fpls.2023.1132699

Jha, U. C., Nayyar, H., Parida, S. K., Beena, R., Pang, J., and Siddique, K. H. M. (2023). Breeding and genomics approaches for improving phosphorus-use efficiency in grain legumes. *Environ. Exp. Bot.* 205, 105120. doi:10.1016/j.envexpbot.2022.105120

Kebede, E. (2021). Contribution, utilization, and improvement of legumes-driven biological nitrogen fixation in agricultural systems. *Front. Sustain. Food Syst.* 5, 1–18. doi:10.3389/fsufs.2021.767998

Khan, M., Ali, S., Al Azzawi, T. N., Saqib, S., Ullah, F., Ayaz, A., et al. (2023). The key roles of ROS and RNS as a signaling molecule in plant–microbe interactions. *Antioxidants* 12, 268. doi:10.3390/antiox12020268

Kim, K., Song, I.-G., Yoon, H., and Park, J.-W. (2023). Sub-micron microplastics affect nitrogen cycling by altering microbial abundance and activities in a soil-legume system. *J. Hazard. Mater.* 460, 132504. doi:10.1016/j.jhazmat.2023.132504

Koza, N. A., Adedayo, A. A., Babalola, O. O., and Kappo, A. P. (2022). Microorganisms in plant growth and development: roles in abiotic stress tolerance and secondary metabolites secretion. *Microorganisms* 10, 1528. doi:10.3390/microorganisms10081528

Kurotani, K.-I, Hirakawa, H., Shirasawa, K., Tanizawa, Y., Nakamura, Y., Isobe, S., et al. (2023). Genome sequence and analysis of *Nicotiana benthamiana,* the model plant for interactions between organisms. *Plant Cell Physiology* 64, 248–257. doi:10.1093/pcp/pcac168

Lata, S., Sharma, S., and Kaur, S. (2023). OMICS approaches in mitigating metal toxicity in comparison to conventional techniques used in cadmium bioremediation. *Water, Air, & Soil Pollut.* 234, 148. doi:10.1007/s11270-023-06145-7

Li, H., Tahir Ul Qamar, M., Yang, L., Liang, J., You, J., and Wang, L. (2023a). Current progress, applications and challenges of multi-omics approaches in sesame genetic improvement. *Int. J. Mol. Sci.* 24, 3105. doi:10.3390/ijms24043105

Li, W., Li, W.-B., Xing, L.-J., and Guo, S.-X. (2023b). Effect of arbuscular mycorrhizal fungi (AMF) and plant growth-promoting rhizobacteria (PGPR) on microorganism of phenanthrene and pyrene contaminated soils. *Int. J. Phytoremediation* 25, 240–251. doi:10.1080/15226514.2022.2071832

Li, X., Zhang, X., Zhao, Q., and Liao, H. (2023c). Genetic improvement of legume roots for adaption to acid soils. *Crop J.* 11, 1022–1033. doi:10.1016/j.cj.2023.04.002

Lim, H. J., Choi, M. J., Byun, S. A., Won, E. J., Park, J. H., Choi, Y. J., et al. (2023). Whole-genome sequence analysis of *Candida glabrata* isolates from a patient with persistent fungemia and determination of the molecular mechanisms of multidrug resistance. *J. Fungi* 9, 515. doi:10.3390/jof9050515

Liu, X. (2023). Sustainable intensification: a historical perspective on China's farming system. *Farming Syst.* 1, 100001. doi:10.1016/j.farsys.2023.100001

Lopez, B. R., Palacios, O. A., Bashan, Y., Hernández-Sandoval, F. E., and De-Bashan, L. E. (2019). Riboflavin and lumichrome exuded by the bacterium Azospirillum brasilense promote growth and changes in metabolites in Chlorella sorokiniana under autotrophic conditions. *Algal Res.* 44, 101696. doi:10.1016/j.algal.2019.101696

Magnoli, S. M., and Bever, J. D. (2023). Plant productivity response to inter- and intra-symbiont diversity: mechanisms, manifestations and meta-analyses. *Ecol. Lett.* 26, 1614–1628. doi:10.1111/ele.14274

Malik, G., Singh, N., and Hooda, S. (2020). "Nitrogen stress in plants and the role of phytomicrobiome," in *Phyto-microbiome in stress regulation*. Editors M. Kumar, V. Kumar, and R. Prasad (Singapore: Springer Singapore), 295–322. doi:10.1007/978-981-15-2576-6_15

Marone, D., Mastrangelo, A. M., and Borrelli, G. M. (2023). From transgenesis to genome editing in crop improvement: applications, marketing, and legal issues. *Int. J. Mol. Sci.* 24, 7122. doi:10.3390/ijms24087122

Martínez-Martínez, J. G., Rosales-Loredo, S., Hernández-Morales, A., Arvizu-Gómez, J. L., Carranza-Álvarez, C., Macías-Pérez, J. R., et al. (2023). Bacterial communities associated with the roots of Typha spp. and its relationship in phytoremediation processes. *Microorganisms* 11, 1587. doi:10.3390/microorganisms11061587

Mclaughlin, S., Zhalnina, K., Kosina, S., Northen, T. R., and Sasse, J. (2023). The core metabolome and root exudation dynamics of three phylogenetically distinct plant species. *Nat. Commun.* 14, 1649. doi:10.1038/s41467-023-37164-x

Molnár, Z., Solomon, W., Mutum, L., and Janda, T. (2023). Understanding the mechanisms of Fe deficiency in the rhizosphere to promote plant resilience. *Plants* 12, 1945. doi:10.3390/plants12101945

Nakatsu, C. H. (2021). "4 - microbial genetics," in *Principles and applications of soil Microbiology*. Editors T. J Gentry, J. J Fuhrmann, and D. A Zuberer. Third Edition (Elsevier), 89–109. doi:10.1016/B978-0-12-820202-9.00004-6

Nguyen, M.-K., Lin, C., Nguyen, H.-L., Hung, N. T. Q., La, D. D., Nguyen, X. H., et al. (2023). Occurrence, fate, and potential risk of pharmaceutical pollutants in agriculture: challenges and environmentally friendly solutions. *Sci. Total Environ.* 899, 165323. doi:10.1016/j.scitotenv.2023.165323

Noor, W., Majeed, G., Lone, R., Tyub, S., Kamili, A. N., and Azeez, A. (2023). "Interactive role of phenolics and PGPR in alleviating heavy metal toxicity in wheat," in *Plant phenolics in abiotic stress management*. Editors R. Lone, S. Khan, and A. Mohammed Al-Sadi (Singapore: Springer Nature Singapore), 287–320. doi:10.1007/978-981-19-6426-8_14

Nwachukwu, B. C., and Babalola, O. O. (2022). Metagenomics: a tool for exploring key microbiome with the potentials for improving sustainable agriculture. *Front. Sustain. Food Syst.* 6, 1–15. doi:10.3389/fsufs.2022.886987

Ojuederie, O. B., Olanrewaju, O. S., and Babalola, O. O. (2019). Plant growth promoting rhizobacterial mitigation of drought stress in crop plants: implications for sustainable agriculture. *Agronomy* 9, 712. doi:10.3390/agronomy9110712

Olanrewaju, O. S., Ayangbenro, A. S., Glick, B. R., and Babalola, O. O. (2019). Plant health: feedback effect of root exudates-rhizobiome interactions. *Appl. Microbiol. Biotechnol.* 103, 1155–1166. doi:10.1007/s00253-018-9556-6

Olanrewaju, O. S., Ayilara, M. S., Ayangbenro, A. S., and Babalola, O. O. (2021). Genome mining of three plant growth-promoting Bacillus species from maize rhizosphere. *Appl. Biochem. Biotechnol.* 193, 3949–3969. doi:10.1007/s12010-021-03660-3

Olanrewaju, O. S., and Babalola, O. O. (2022). Plant growth-promoting rhizobacteria for orphan legume production: focus on yield and disease resistance in Bambara groundnut. *Front. Sustain. Food Syst.* 6, 1–8. doi:10.3389/fsufs.2022.922156

Olanrewaju, O. S., Glick, B. R., and Babalola, O. O. (2017). Mechanisms of action of plant growth promoting bacteria. *World J. Microbiol. Biotechnol.* 33, 197. doi:10.1007/s11274-017-2364-9

Omotoso, A. B., Letsoalo, S., Olagunju, K. O., Tshwene, C. S., and Omotayo, A. O. (2023). Climate change and variability in sub-Saharan Africa: a systematic review of trends and impacts on agriculture. *J. Clean. Prod.* 414, 137487. doi:10.1016/j.jclepro.2023.137487

Özdoğan, D. K., Akçelik, N., and Akçelik, M. (2022). Genetic diversity and characterization of plant growth-promoting effects of bacteria isolated from rhizospheric soils. *Curr. Microbiol.* 79, 132. doi:10.1007/s00284-022-02827-3

Paponov, M., Flate, J., Ziegler, J., Lillo, C., and Paponov, I. A. (2023). Heterogeneous nutrient supply modulates root exudation and accumulation of medicinally valuable compounds in *Artemisia annua* and *Hypericum perforatum*. *Front. Plant Sci.* 14, 1174151. doi:10.3389/fpls.2023.1174151

Paravar, A., Piri, R., Balouchi, H., and Ma, Y. (2023). Microbial seed coating: an attractive tool for sustainable agriculture. *Biotechnol. Rep.* 37, e00781. doi:10.1016/j.btre.2023.e00781

Pergner, I., and Lippert, C. (2023). On the effects that motivate pesticide use in perspective of designing a cropping system without pesticides but with mineral fertilizer—a review. *Agron. Sustain. Dev.* 43, 24. doi:10.1007/s13593-023-00877-w

Puppala, N., Nayak, S. N., Sanz-Saez, A., Chen, C., Devi, M. J., Nivedita, N., et al. (2023). Sustaining yield and nutritional quality of peanuts in harsh environments: physiological and molecular basis of drought and heat stress tolerance. *Front. Genet.* 14, 1121462. doi:10.3389/fgene.2023.1121462

Rahman, S. U., Mccoy, E., Raza, G., Ali, Z., Mansoor, S., and Amin, I. (2023). Improvement of soybean; A way forward transition from genetic engineering to new plant breeding technologies. *Mol. Biotechnol.* 65, 162–180. doi:10.1007/s12033-022-00456-6

Raja, B., and Vidya, R. (2023). Application of seaweed extracts to mitigate biotic and abiotic stresses in plants. *Physiology Mol. Biol. Plants* 29, 641–661. doi:10.1007/s12298-023-01313-9

Sah, S., Krishnani, S., and Singh, R. (2021). *Pseudomonas* mediated nutritional and growth promotional activities for sustainable food security. *Curr. Res. Microb. Sci.* 2, 100084. doi:10.1016/j.crmicr.2021.100084

Sahel, D. K., Vora, L. K., Saraswat, A., Sharma, S., Monpara, J., D'souza, A. A., et al. (2023). CRISPR/Cas9 genome editing for tissue-specific *in vivo* targeting: nanomaterials and translational perspective. *Adv. Sci.* 10, 2207512. doi:10.1002/advs.202207512

Saleh-Lakha, S., Miller, M., Campbell, R. G., Schneider, K., Elahimanesh, P., Hart, M. M., et al. (2005). Microbial gene expression in soil: methods, applications and challenges. *J. Microbiol. Methods* 63, 1–19. doi:10.1016/j.mimet.2005.03.007

Sarker, A., Ansary, M. W., Hossain, M. N., and Islam, T. (2021). Prospect and challenges for sustainable management of climate change-associated stresses to soil and plant health by beneficial rhizobacteria. *Stresses* 1, 200–222. doi:10.3390/stresses1040015

Shaposhnikov, A. I., Belimov, A. A., Azarova, T. S., Strunnikova, O. K., Vishnevskaya, N. A., Vorobyov, N. I., et al. (2023). The relationship between the composition of root exudates and the efficiency of interaction of wheat plants with microorganisms. *Appl. Biochem. Microbiol.* 59, 330–343. doi:10.1134/S000368382303016X

Shariati, J. V., Malboobi, M. A., Tabrizi, Z., Tavakol, E., Owlia, P., and Safari, M. (2017). Comprehensive genomic analysis of a plant growth-promoting rhizobacterium Pantoea agglomerans strain P5. *Sci. Rep.* 7, 15610. doi:10.1038/s41598-017-15820-9

Sharma, D. K. (2019). Enumerations on seed-borne and post-harvest microflora associated with okra [*Abelmoschus esculentus* (L.) Moench] and their management. *GSC Biol. Pharm. Sci.* 8, 119–130. doi:10.30574/gscbps.2019.8.2.0149

Shiau, C.-K., Lu, L., Kieser, R., Fukumura, K., Pan, T., Lin, H.-Y., et al. (2023). High throughput single cell long-read sequencing analyses of same-cell genotypes and phenotypes in human tumors. *Nat. Commun.* 14, 4124. doi:10.1038/s41467-023-39813-7

Simlat, M., Ptak, A., Jaglarz, A., Szewczyk, A., Dziurka, M., and Gurgul, A. (2023). Seeds of Stevia rebaudiana bertoni as a source of plant growth-promoting endophytic bacteria with the potential to synthesize rebaudioside A. *Int. J. Mol. Sci.* 24, 2174. doi:10.3390/ijms24032174

Šmeringai, J., Schrumpfová, P. P., and Pernisová, M. (2023). Cytokinins - regulators of *de novo* shoot organogenesis. *Front. Plant Sci.* 14, 1239133. doi:10.3389/fpls.2023.1239133

Solomon, W., Mutum, L., Janda, T., and Molnár, Z. (2023). Potential benefit of microalgae and their interaction with bacteria to sustainable crop production. *Plant Growth Regul.* 101, 53–65. doi:10.1007/s10725-023-01019-8

Swarnalakshmi, K., Yadav, V., Tyagi, D., Dhar, D. W., Kannepalli, A., and Kumar, S. (2020). Significance of plant growth promoting rhizobacteria in grain legumes: growth promotion and crop production. *Plants* 9, 1596. doi:10.3390/plants9111596

Swarnalatha, G. V., Goudar, V., Reddy, E. C. R. G. S., Al Tawaha, A. R. M., and Sayyed, R. Z. (2022). "Siderophores and their applications in sustainable management of plant diseases," in *Secondary metabolites and volatiles of PGPR in plant-growth promotion*. Editors R. Z. Sayyed and V. G. Uarrota (Cham: Springer International Publishing), 289–302. doi:10.1007/978-3-031-07559-9_14

Tan, X. L., Azam-Ali, S., Goh, E. V., Mustafa, M., Chai, H. H., Ho, W. K., et al. (2020). Bambara groundnut: an underutilized leguminous crop for global food security and nutrition. *Front. Nutr.* 7, 601496. doi:10.3389/fnut.2020.601496

Tortella, G., Rubilar, O., Pieretti, J. C., Fincheira, P., De Melo Santana, B., Fernández-Baldo, M. A., et al. (2023). Nanoparticles as a promising strategy to mitigate biotic stress in agriculture. *Antibiotics* 12, 338. doi:10.3390/antibiotics12020338

Utami, D., Meale, S. J., and Young, A. J. (2022). A pan-global study of bacterial leaf spot of chilli caused by *xanthomonas* spp. *Plants* 11, 2291. doi:10.3390/plants11172291

Walker, L., Lagunas, B., and Gifford, M. L. (2020). Determinants of host range specificity in legume-rhizobia symbiosis. *Front. Microbiol.* 11, 585749. doi:10.3389/fmicb.2020.585749

Wang, C., Wang, J., Niu, X., Yang, Y., Malik, K., Jin, J., et al. (2023a). Phosphorus addition modifies the bacterial community structure in rhizosphere of *Achnatherum inebrians* by influencing the soil properties and modulates the *Epichloë gansuensis*-mediated root exudate profiles. *Plant Soil* 1–18. doi:10.1007/s11104-023-06133-0

Wang, Q., Liu, J., and Zhu, H. (2018). Genetic and molecular mechanisms underlying symbiotic specificity in legume-rhizobium interactions. *Front. Plant Sci.* 9, 313–318. doi:10.3389/fpls.2018.00313

Wang, Q., Wang, D., Agathokleous, E., Cheng, C., Shang, B., and Feng, Z. (2023b). Soil microbial community involved in nitrogen cycling in rice fields treated with antiozonant under ambient ozone. *Appl. Environ. Microbiol.* 89, e0018023. doi:10.1128/aem.00180-23

Wang, Q., Zhang, P., Zhao, W., Li, Y., Jiang, Y., Rui, Y., et al. (2023c). Interplay of metal-based nanoparticles with plant rhizosphere microenvironment: implications for nanosafety and nano-enabled sustainable agriculture. *Environ. Sci. Nano* 10, 372–392. doi:10.1039/d2en00803c

Woźniak-Gientka, E., and Tyczewska, A. (2023). Genome editing in plants as a key technology in sustainable bioeconomy. *EFB Bioeconomy J.* 3, 100057. doi:10.1016/j.bioeco.2023.100057

Xia, H., Deng, H., Li, M., Xie, Y., Lin, L., Zhang, H., et al. (2023). Chromosome-scale genome assembly of a natural diploid kiwifruit (*Actinidia chinensis var. deliciosa*). *Sci. Data* 10, 92. doi:10.1038/s41597-023-02006-4

Xie, J., Shi, H., Du, Z., Wang, T., Liu, X., and Chen, S. (2016). Comparative genomic and functional analysis reveal conservation of plant growth promoting traits in *Paenibacillus polymyxa* and its closely related species. *Sci. Rep.* 6, 21329. doi:10.1038/srep21329

Yao, K., Wang, Y., Li, X., and Ji, H. (2023). Genome-wide identification of the soybean LysM-RLK family genes and its nitrogen response. *Int. J. Mol. Sci.* 24, 13621. doi:10.3390/ijms241713621

Ying, S. Y. (2003). "Complementary DNA libraries," in *Generation of cDNA libraries: methods and protocols*. Editor S. Y. Ying (Totowa, NJ: Humana Press), 1–12. doi:10.1385/1-59259-359-3:1

Stability of soil bacteria in undisturbed soil and continuous maize cultivation in Northern Thailand

Noppol Arunrat[1]*, Chakriya Sansupa[2], Sukanya Sereenonchai[1] and Ryusuke Hatano[3]

[1]Faculty of Environment and Resource Studies, Mahidol University, Nakhon Pathom, Thailand, [2]Department of Biology, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand, [3]Laboratory of Soil Science, Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan

Rotational shifting cultivation (RSC) in Northern Thailand serves the dual purpose of ensuring food security and meeting economic goals through maize cultivation. However, the research question remains: Does the dynamics of soil bacterial communities differ between maize monoculture and RSC fields with continuous fallow throughout the season? Therefore, the objective of this study was to investigate and compare the variation of soil bacterial communities in maize monoculture and fallow RSC fields. A continuous 5-year fallow field (undisturbed soil; CF-5Y) and a continuous 5-year maize cultivation field (M-5Y) in Mae Chaem District, Chiang Mai Province, Northern Thailand, were selected due to their similarities in microclimate, topography, and the 5-year duration of different field activities. Over the span of a year, we collected soil samples from the surface layer (0−2 cm depth) at both sites. These collections occurred at 3-month intervals, starting from March 2022 (summer season) and followed by June (rainy season), September (rainy season), December (winter season), and March 2023 (summer season). Soil bacterial diversity and composition were analyzed using 16S rRNA gene-based metagenomic analysis. The results found that undisturbed soil over a 5-year period exhibited more stability in the richness and diversity of bacteria across seasons compared with M-5Y. Notably, fertilizer application and tillage practices in M-5Y can enhance both the diversity and richness of soil bacteria. In terms of bacterial abundance, Proteobacteria prevailed in CF-5Y, while Actinobacteria dominated in M-5Y. At the genus level, Candidatus *Udaeobacter* dominated during the summer and winter seasons in both CF-5Y and M-5Y sites. Interestingly, during the rainy season, the dominant genus shifted to *Bacillus* in both CF-5Y and M-5Y fields. The soil bacterial community in M-5Y was strongly influenced by organic matter (OM) and organic carbon (OC). In contrast, in CF-5Y, there was no correlation between soil properties and the soil bacterial community, likely due to the lower variation in soil properties across seasons. β-Glucosidase was the dominant enzyme in both CF-5Y and M-5Y sites, and it showed a positive correlation with OM and OC. Further studies should continue to investigate soil bacteria dynamics, considering the changes in land management practices.

1. Introduction

Soil microorganisms play a crucial role in the dynamic transformation of soil nutrients (Schulz et al., 2013), organic matter transformation, the enhancement of plant productivity, and the control of soil-borne diseases (Pieterse et al., 2016). However, soil microorganisms are highly sensitive to soil disturbance and management practices, which can lead to changes in their functionality (Cordovez et al., 2019). Changes in land use, specifically alterations in vegetation cover, can significantly impact soil physicochemical properties and drive shifts in microbial community composition (Szoboszlay et al., 2017; Meng et al., 2019). The intensification of land use, such as the conversion of natural vegetation to arable agriculture, has been associated with a substantial loss of soil organic carbon (SOC) by up to 44.9% (Arunrat et al., 2022a), primarily due to the loss of vegetation cover, litter, and root biomass (Whitbread et al., 1999; Shahzad et al., 2018). These changes ultimately lead to a decrease in microbial populations and a reduction in soil microbial diversity (Williams et al., 2017; Newbold et al., 2018), resulting in degraded soil ecosystems and long-term decline in soil fertility (Gouda et al., 2018). Therefore, the identification and quantification of soil microorganisms are essential for assessing soil quality status, which can contribute to maintaining soil nutrient levels and improving crop productivity (Schloter et al., 2003).

Shifting cultivation, also known as swidden farming or slash-and-burn farming, is a widely practiced agricultural method in the highlands of Southeast Asia (Rerkasem and Rerkasem, 1995). It is deeply rooted in the cultural heritage of indigenous communities. The practice involves a cyclic rotation of farming areas, encompassing conversion, cultivation, harvesting, burning, and fallow periods lasting for 5–15 years (Kleinman et al., 1995). In Northern Thailand, rotational shifting cultivation (RSC) is a traditional farming practice among the inhabitants of mountainous regions (Arunrat et al., 2022b). RSC involves managing fallow cycles where one area is temporarily cultivated and subsequently left fallow to allow vegetation and soil fertility to recover, while the farmers move on to another area (Arunrat et al., 2023a). Fire is utilized as a land preparation tool during the conversion process, altering soil characteristics depending on the severity and intensity of the burn (Caon et al., 2014). A study conducted by Wapongnungsang et al. (2021) examined soil properties in shifting agricultural areas over 10 and 15 years. The findings indicated that pH, as well as the availability of phosphorus (P) and nitrogen (N), increased in the burned regions compared with the pre-burned areas. Moreover, previous research suggests that shifting agriculture not only affects soil physical structure and biochemical characteristics but also alters soil microbial communities due to the heat generated during combustion (Certini, 2005; Barreiro and Díaz-Raviña, 2021), potentially influencing the survival and growth of bacteria. *Bacillus* sp. have been commonly found in burned swidden areas (Smith et al., 2008; Miah et al., 2010). However, when comparing bacterial communities in shifting agriculture areas with a 13-year-old reforested area, distinct differences were observed. Cultivated areas showed the presence of *Pseudomonas diminuta* and *Shigella*, whereas forested areas contained *Bacillus firmus* and *Xanthomonas*, with variations in bacterial strains linked to soil properties (Miah et al., 2014). During the fallow period, the field accumulates organic matter (OM) inputs from litter and roots (Sarkar et al., 2015), which significantly impacts soil microbial communities (Kim et al., 2020).

Maize (*Zea mays* Lin.) is a vital food and industrial crop, ranking among the most widely cultivated crops worldwide. In 2022, Thailand exported approximately 977.15 tons of maize and imported approximately 1.4 million tons, indicating a high demand for maize in the country. Maize is predominantly used in the animal feed industry, which contributes to its economic significance (Ekasingh et al., 2004). Thai farmers are increasingly shifting toward maize production for animal feed due to its easier management, shorter growth cycle, and lower water requirements compared with rice cultivation. However, continuous maize cultivation in certain areas has led to challenges such as soil degradation (Moebius-Clune et al., 2011) and a decline in biodiversity due to the prolonged cultivation of a single plant variety. Additionally, soil fertility limitations often impact maize development and yield (Essel et al., 2020). The growth of maize is influenced by various factors, including soil microorganisms, organic fertilizer treatments, chemical fertilizers, and soil nutrient availability. Previous studies have compared the response of bacteria in maize-growing areas under organic fertilizer application versus no organic fertilizer application, revealing that the use of organic fertilizer can positively impact maize growth (Wu et al., 2021). Furthermore, *Mycobacterium phlei* was identified as an efficient bacterial strain for enhancing maize growth (Egamberdiyeva, 2007). In Northeast China, Yang et al. (2022) demonstrated that maize cropping and continuous cropping of alfalfa increased soil bacterial alpha diversity compared with meadow cropping. This can be attributed to soil management practices, including tillage operations, which promote increased decomposition and mineralization of available nutrients, thereby influencing the structure of soil bacterial communities.

However, the extensive cultivation of maize in Northern Thailand, driven by high price incentives, has resulted in a reduction in the number of RSC fields. As a consequence, there is a lack of direct evidence regarding the dynamics of soil bacterial communities under maize monoculture compared with RSC fields with continuous fallow. Therefore, the objective of this study was to investigate and compare the variation of soil bacterial communities in areas of maize monoculture and RSC fields with a continuous 5-year fallow period. Our hypothesis posited that 5 years of undisturbed soil would lead to a more stable community composition compared with monoculture.

2. Materials and methods

2.1. Study areas

The study sites were situated in Ban Mae Pok, Ban Thab Subdistrict, Mae Chaem District, Chiang Mai Province, Northern Thailand (Figure 1). Average weather data, including maximum, minimum, and mean temperatures, precipitation, and the number of rainy days, for the study period (March 2022 to March 2023) were obtained from the Thai Meteorological Department, sourced from Doi Ang Khang and Mueang Chiang Mai stations (see Supplementary Table S1; Supplementary Figure S1). A continuous 5-year fallow field (CF-5Y) and a continuous 5-year maize cultivation field (M-5Y) were selected due to their similarity in microclimate, topography, and 5-year duration of different field activities. The CF-5Y field (18°23′3″N, 98°11′49″E; elevation 659 m a.s.l) was used for cultivating upland rice, and after the harvest, it has been left abandoned for 5 years. The M-5Y field (18°23′30″N, 98°10′52″E;
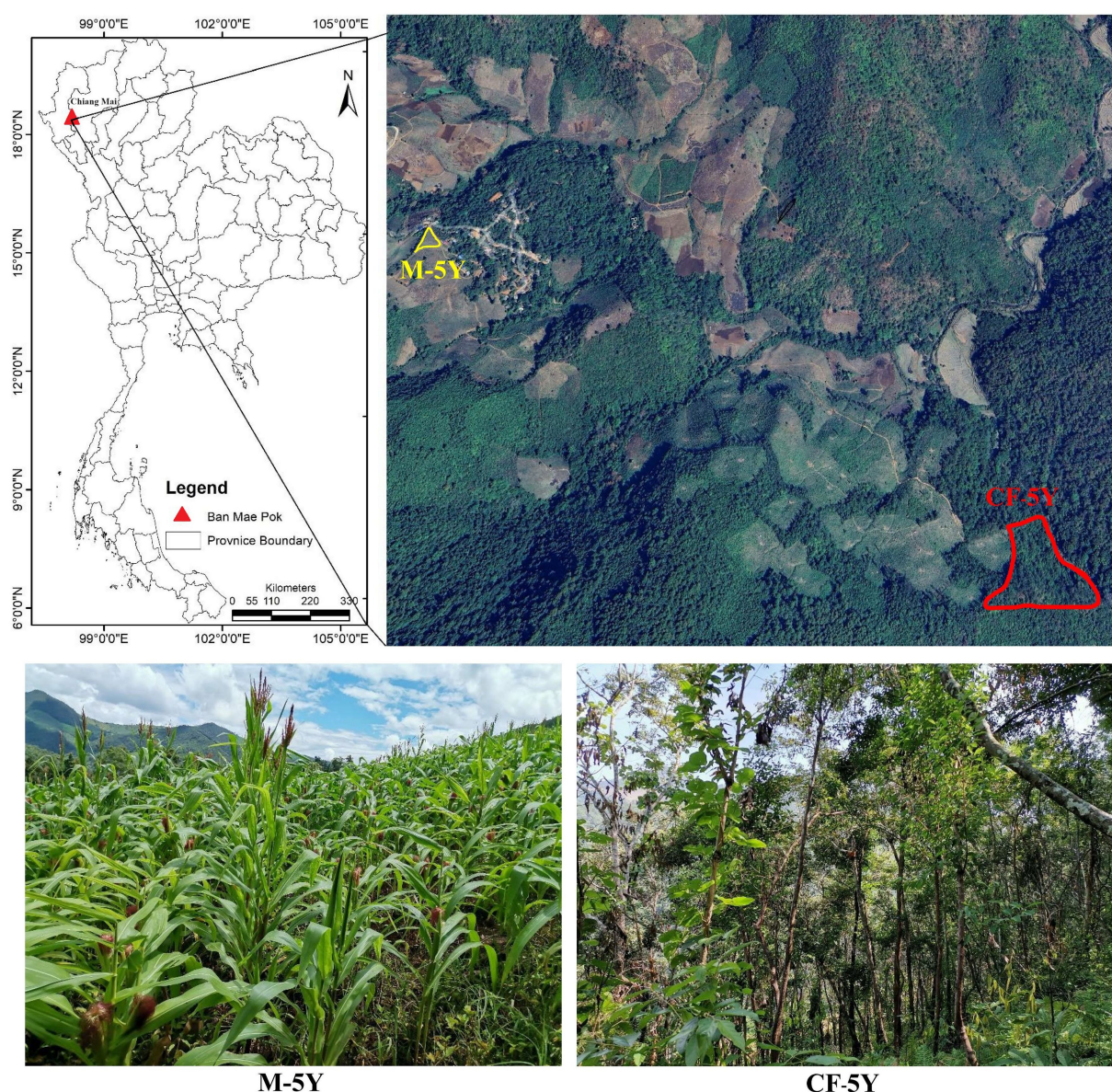
FIGURE 1
Study areas. The aerial image was captured from Google Earth on 29 August 2023. The photos were taken by Noppol Arunrat.

elevation 823 m a.s.l) was previously abandoned land from upland rice cultivation, and it has been converted into a maize field for 5 years. Tillage is typically performed in approximately May, which marks the early rainy season. This practice involves plowing to a depth of 20–30 cm. Subsequently, maize seeds were sown in June, and the harvest was carried out in November. Chemical fertilizer was applied only once at the flowering stage during August (40–50 days after sowing). Nitrogen fertilizer was applied at a rate of approximately 25.5 to 30.2 kg N ha^{-1} year^{-1}, while phosphorus fertilizer was used at a rate of approximately 5.8–7.5 kg P ha^{-1} year^{-1}. During maize growth, herbicides were not used, but pesticides were needed to control pest infestations. After the maize harvest, residues were left in the field without burning. While the field was left fallow until the next rainy season, cows were brought to the maize field to consume maize residues and grasses.

2.2. Experimental design and soil sampling

The CF-5Y (~180 m × 270 m) and M-5Y fields (~50 m × 78 m) were divided into three transects, starting from the uppermost slope to the lowermost slope. Along each transect in each field, five plots were designated at equal distances, and soil samples at a depth of 0–2 cm (surface layer) were collected. Previous studies (Higashida and Takao, 1985; Mehring et al., 2011; Liu C. et al., 2021; Barbour et al., 2022) have also reported that soil bacteria in the surface soil layer (0–2 cm) are more susceptible to changes in environmental conditions. Following this, the individual soil samples were combined to create a single composite sample for each transect. This approach was adopted due to the potential variation in soil nutrients caused by erosion processes along the slope (Arunrat et al., 2023b), which could impact the diversity of soil bacteria due to slope-related differences.

Approximately 1 kg of soil was collected and placed in a plastic bag for analysis of the physical and chemical properties of the soil. The steel soil core (5.0 cm width × 5.5 cm length) was used to collect soil samples from each plot at a depth of 5 cm for bulk density analysis. Additionally, approximately 100 g of soil was collected, placed in zip-lock plastic bags, and kept at 10°C in the field and at −20°C in the laboratory for DNA extraction. Soil samplings were collected for a period of 1 year, at 3-month intervals, starting from March 2022 (summer season) and followed by June (rainy season), September (rainy season), December (winter season), and March 2023 (summer season). For each soil sampling plot during each time period, soil temperature and soil moisture were recorded at a depth of 2 cm using a Thermocouple Type K and a soil moisture meter, respectively. The results of soil temperature and soil moisture measurements are presented in Supplementary Table S2.

2.3. Soil physicochemical properties analysis

The soil bulk density was determined by weighing the dry soil sample in a steel soil core after it was dried at 105°C for 24 h. The soil texture was determined using the hydrometer method. Soil pH was measured using a pH meter with a 1:1 ratio of solid to water (National Soil Survey Center, 1996). Electrical conductivity (ECe) was measured by determining the saturation of paste extracts using an EC meter (USDA, 1954). The cation exchange capacity (CEC) was determined using the NH_4OAc pH 7.0 method. Total nitrogen (TN) was analyzed using the micro-Kjeldahl method. Ammonium nitrogen (NH_4-N) and nitrate-nitrogen (NO_3-N) were measured using the KCl extraction method. The exchangeable calcium (exch.Ca), magnesium (exch.Mg), and potassium (exch.K) values were determined using atomic absorption spectrometry with NH_4OAc pH 7.0 extraction. Available phosphorus (avail.P) was measured using the molybdate blue method (Bray II extraction) (Bray and Kurtz, 1945). Organic carbon (OC) was determined using potassium dichromate ($K_2Cr_2O_7$) in sulfuric acid (Walkley and Black, 1934).

2.4. DNA extraction, bacterial 16 s amplification, and sequencing

Soil DNA was extracted using a DNeasy PowerSoil Pro DNA Kit (Qiagen). Subsequently, the DNA was amplified, targeting the V3–V4 region of the 16S rRNA gene, using primers 341F (5′-CCTAYGG-GDBGCWSCAG) and 805R (5′-GGACTAC-NVGGGTHTCTAAT-3′) (Klindworth et al., 2013). The amplicon was then subjected to a sequencing step using the paired-end Illumina Miseq platform at the Omics Sciences and Bioinformatics Center of Chulalongkorn University (Bangkok, Thailand).

2.5. Bioinformatics analysis

QIIME2 was used for bioinformatics analysis (Estaki et al., 2020). Forward and reverse primers were removed by Cutadapt (Martin, 2011). The sequences were then quality-filtered, merged, and chimera removed using the DADA2 plugin (Callahan et al., 2016). After this

analysis, similar sequences were grouped together as amplicon sequence variants (ASVs). Then, the ASVs that contained less than two sequences were removed. Bacterial taxonomies were identified using the Silva v.138 database (Quast et al., 2013). Then, the ASVs that were identified as mitochondria or chloroplasts were removed. The remaining ASVs were then normalized to the smallest number of sequences from each sample using the rarefy plugin. The rarefied data were subjected to PICRUSt2 to predict enzyme abundance that was potentially performed by the bacteria found in this dataset (Douglas et al., 2020). Specifically, we highlight 13 enzymes involved in the soil system, such as those in the C, N, and P cycles (Das and Varma, 2011). These included acid phosphatase, alkaline phosphatase, alpha-N-acetylglucosaminidase, amidase, β-glucosidase, cellulase, chitinase, endo-1,4-β-xylanase, pectin lyase, rease, Xylan 1,4-β-xylosidase, nitrogenase, and nitrate reductase.

2.6. Statistical analysis

Statistical analyses were performed on PAST (Hammer et al., 2001) and R program (R Development Core, 2019). Beta diversity, which presents bacterial community composition, was visualized by Non-Metric Multidimensional Scaling (NMDS). The affected site and sampling times to bacterial community composition in both study sites were tested using two-way permutational multivariate analyses of variance (PERMANOVA). The correlation between soil properties and soil bacterial community compositions was analyzed using the Mantel test. Soil properties with significant correlation implied that these properties significantly shape the bacterial community composition in both study sites. The alpha diversity, presenting bacterial richness (observed richness and abundance-based coverage estimator (ACE) indexes) and bacterial diversity (Simpson and Shannon indexes), was computed using the microeco package (Liu X. et al., 2021). Differences between each alpha diversity index, relative abundance of top 10 genera, and abundance of soil enzymes were tested using ANOVA with repeated-measure. Spearman's rank correlation was tested to determine the effect of soil properties on the abundance of bacterial genera and predictive enzymes.

3. Results

3.1. The effects of site and sampling time on bacterial community composition

NMDS ordination of bacterial community composition in CF-5Y and M-5Y across all time points is presented in Figure 2. The bacterial community in CF-5Y was separated from M-5Y and the bacterial community at different sampling times. This indicated that the community composition of bacteria in CF-5Y was different from that in M-5Y. Two-way PERMANOVA analysis indicated that site, sampling time, and the interaction between these two factors significantly shape the bacterial community structure ($p < 0.05$; Figure 2). It indicated that CF-5Y and M-5Y fields, as well as the timing of sampling, were the vital factors, influencing the bacterial community structure. Soil physical and chemical properties of CF-5Y and M-5Y are presented in Tables 1, 2. There were relatively small variations in soil properties across seasons for both sites. This suggests

FIGURE 2
Bacterial community composition across all samples.
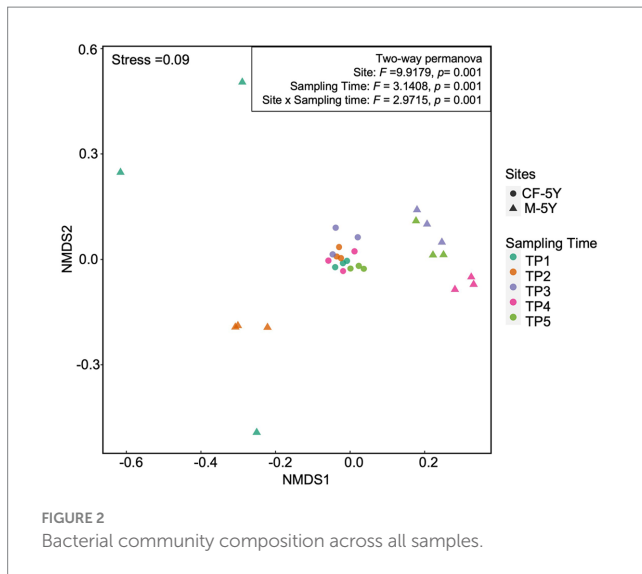
that the soil properties remained relatively stable throughout the year. However, it is worth noting that OM and OC contents showed a declining trend over the seasons. TN, Avail. P, and Exch. K levels were highest during the rainy season (September), while Exch. Ca, Exch. Mg, NH$_4$-N, and NO$_3$-N peaked after the rainy season in December. The Mantel test revealed significant correlations between the bacterial community and soil properties, including OM, OC, Avail.P, Exch.Mg, TN, BD, sand, silt, and clay in M-5Y. The strongest correlations were observed in OM and OC, with correlation coefficients exceeding 0.5 (Table 3). However, no significant correlation was detected between soil properties and the bacterial community in CF-5Y (Table 3), likely due to the lower variation in soil properties across seasons.

3.2. Bacterial richness and diversity

The alpha diversity index was used to assess the richness of bacteria, which included observed richness and ACE indexes. Moreover, the diversity of bacteria was evaluated using the Shannon and Simpson indexes. The richness and diversity indexes varied in different sampling times and study sites. In CF-5Y, all alpha diversity indexes dropped in TP2 and TP3 (rainy season) and increased during dry season (TP4—winter, TP5—summer). Significant differences were found in observed richness and ACE between TP3 and TP4, which indicate the transition from the rainy season to the winter season. The richness indexes significantly increased in this period (T4; Figures 3A–D). After 1 year of first sampling, all diversity indexes did not significantly change.

On the other hand, in M-5Y, all alpha diversity indexes were slightly increased in TP2 (rainy season); however, only a significant difference was found in the Shannon index. After that, observed richness, ACE, and Shannon indexes were significantly increased from TP3 to TP4, the transition between rainy and winter seasons. Then, these indexes did not change between TP4 and TP5, which are winter and summer seasons (Figures 3E–H). After 1 year of first sampling, all diversity indexes significantly increased.

Spearman's rank correlation showed that the variation of alpha diversity indexes in both study sites was correlated with three soil

properties, including BD, sand, and clay. Observed richness, ACE, Shannon, and Simpson indexes negatively correlated with BD and sand, whereas Shannon and Simpson indexes positively correlated with clay (Figure 3I).

3.3. Bacterial taxonomic distribution

A total of 37 phyla, 93 classes, 190 orders, 258 families, and 506 genera were detected in this study. The most abundant phyla across all samples were Actinobacteria, followed by Proteobacteria, Planctomycetes, and Chloroflexi.

3.3.1. Bacterial taxonomic distribution in CF-5Y

The most abundant taxa in the first sampling time (TP1: March 2022—summer) were Proteobacteria (22.94%), followed by Actinobacteria (20.93%) and Planctomycetes (16.23%). Although there are some changes in the relative abundance in each taxon across all time points, the changes were small and did not affect the abundant group in each time point. For example, the abundance of proteobacteria increased to 25.91% (TP2) and 29.70% (TP3) in the rainy season and decreased to 24.90% (TP4) in winter and 18.46% (TP5) in summer. This phylum was still the most abundant phylum across all sampling points (Figure 4A).

At the genus level, the most abundant genus in TP1 was Candidatus *Udaeobacter* (5.59%), followed by *Acidothermus* (3.21%) and *Conexibacter* (3.04%). In the rainy season, *Bacillus* became the most abundant genus, accounting for 4.39% of TP2 and 3.54% of TP3. Furthermore, the abundance of *Conexibacter* (from 3.04% in TP1 to 1.23% in TP3) and *Gemmata* (from 1.59% in TP1 to 0.80% in TP3) significantly decreased in this season. The most abundant genus in TP4 (winter) was Candidatus *Udaeobacter* (5.36%), followed by *Bacillus* (3.83%). In TP5 (summer- 1 year after the first sampling), the proportion of bacterial genera was changed compared with the first sampling point. Although Candidatus *Udaeobacter* (5.71%) was still the most abundant genus, the second most abundant genera were changed from *Acidothermus* to *Bacillus* (4.07%). Furthermore, the abundance of *Bradyrhizobium* significantly decreased after 1 year (Figure 5A).

Spearman's rank correlation showed the correlation between soil properties and the relative abundance of bacterial genera. Here, we found that the significant genera, including *Bradyrhizobium, Acidothermus,* and *Conexibacter,* were positively correlated with OM and OC while negatively correlated with Avail.P, Exch.K, and Exch. Ca. On the other hand, *Gemmata* were positively correlated with CEC (Figure 5B).

3.3.2. Bacterial taxonomic distribution in M-5Y

The relative abundance of bacterial phyla in M-5Y were notably varied at different time points. The most abundant taxa in the first sampling time (TP1: March 2022—summer) were Actinobacteria (31.18%) and Chloroflexi (17.71%). In TP2 (rainy season), the proportion of bacterial phyla notably changed. Chloroflexi became the most abundant phyla, accounting for 31.60%, followed by Proteobacteria (16.84%) and Acidobacteria (14.49%). After 3 months (TP3 rainy season), the most abundant phyla belonged to Proteobacteria (21.40%) and Actinobacteria (19.07%). In winter (TP4), the most abundant phyla belonged to Planctomycetes (24.71%)

TABLE 1 Variation in soil properties: bulk density (BD) (Mg m^{-3}), electrical conductivity (EC$_e$) (dS m^{-1}), organic matter (OM) (%), organic carbon (OC) (%), total nitrogen (TN) (%), and proportion of sand, silt, and clay (%) with sites and time points.

Site	Season	Time point	pH (1:1)		EC$_e$		BD		OM		OC		TN		%Sand		%Silt		%Clay	
			Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
CF-5Y	Summer	March 2022	5.07	0.01	0.17	0.01	1.33[a]	0.01	8.80[a]	0.13	5.10[a]	0.08	0.33[a]	0.02	32.00	0.17	51.73[a]	1.45	16.28	1.28
	Rainy	June 2022	5.01	0.01	0.17	0.02	1.33[a]	0.01	8.65[a]	0.18	5.02[a]	0.10	0.35[a]	0.02	31.22	0.26	50.63	1.01	18.15	0.95
		September 2022	4.95	0.07	0.14	0.03	1.31[a]	0.01	7.69[a]	0.16	4.46[a]	0.09	0.37[a]	0.02	32.15	0.26	48.60[a]	0.97	19.25	0.91
	Winter	December 2022	4.76	0.11	0.15	0.01	1.32[a]	0.02	7.79[a]	0.06	4.52[a]	0.03	0.33[a]	0.02	33.75	0.28	45.19	0.90	21.06[a]	0.84
	Summer	March 2023	4.59	0.02	0.14	0.02	1.31[a]	0.01	7.55[a]	0.07	4.38[a]	0.04	0.34[a]	0.03	35.66	0.76	45.45	0.69	18.88[a]	0.57
M-5Y	Summer	March 2022	4.75	0.17	0.15	0.01	1.53[b]	0.01	2.35[b]	0.09	1.36[b]	0.05	0.17[b]	0.02	40.95	1.32	40.72[b]	1.45	18.33	0.34
	Rainy	June 2022	5.03	0.12	0.16	0.01	1.52[b]	0.02	2.35[b]	0.03	1.37[b]	0.01	0.19[b]	0.02	39.99	1.58	40.53	2.26	19.47	0.75
		September 2022	4.98	0.11	0.16	0.02	1.48[b]	0.02	2.14[b]	0.04	1.24[b]	0.02	0.21[b]	0.02	37.11	2.20	35.49[b]	3.61	27.40	1.99
	Winter	December 2022	4.61	0.04	0.17	0.02	1.45[b]	0.03	2.12[b]	0.02	1.23[b]	0.01	0.16[b]	0.01	26.98	1.76	41.88	2.11	31.13[b]	1.45
	Summer	March 2023	4.96	0.27	0.16	0.01	1.48[b]	0.01	2.08[b]	0.03	1.21[b]	0.02	0.18[b]	0.02	30.40	1.02	38.47	2.01	31.13[b]	1.45

[a,b]Significant statistical differences ($p < 0.05$) between CF-5Y and M-5Y fields.

TABLE 2 Variation in soil properties: cation exchange capacity (CEC) (meq 100 g^{-1}), available P (mg kg^{-1}), exchangeable K, Ca, and Mg (mg kg^{-1}), NH$_4$-N, and NO$_3$-N (mg kg^{-1}) contents with sites and time points.

Site	Season	Time point	CEC		Avail. P		Exch. K		Exch. Ca		Exch. Mg		NH$_4$-N		NO$_3$-N	
			Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
CF-5Y	Summer	March 2022	18.48	0.45	6.19[a]	0.12	155.81	3.24	204.32[a]	0.56	97.90	1.29	17.77[a]	5.03	7.11[a]	0.00
	Rainy	June 2022	15.97	0.67	10.60[a]	0.31	177.15	4.19	214.73[a]	3.92	100.77[a]	2.04	14.21[a]	0.00	7.11[a]	0.00
		September 2022	14.80	0.64	15.12[a]	0.33	194.86	4.52	231.90[a]	4.23	108.82	2.20	14.21[a]	0.00	7.11[a]	0.00
	Winter	December 2022	14.97	0.15	5.82[a]	0.37	164.03	4.98	255.09	4.65	119.70	2.42	21.32	0.00	14.21	0.00
	Summer	March 2023	16.25	0.33	8.33[a]	0.39	172.30	8.38	255.09	4.65	117.24	5.71	21.32	0.00	14.21	0.00
M-5Y	Summer	March 2022	11.36	1.16	40.66[b]	1.89	171.79	6.56	306.66[b]	18.21	119.13	6.31	35.53[b]	0.00	28.42[b]	0.00
	Rainy	June 2022	14.40	0.00	41.89[b]	2.28	195.25	6.00	385.99[b]	27.98	203.33[b]	11.37	35.53[b]	0.00	35.53[b]	0.00
		September 2022	14.40	0.00	53.57[b]	1.70	232.39	19.76	348.57[b]	31.06	161.65	3.15	35.53[b]	0.00	35.53[b]	0.00
	Winter	December 2022	12.13	0.21	19.48[b]	0.76	183.87	3.67	289.73	24.46	148.46	1.90	28.42	0.00	14.21	0.00
	Summer	March 2023	11.20	0.57	37.82[b]	2.47	177.95	2.49	270.79	17.66	135.11	3.75	21.32	0.00	14.21	0.00

[a,b]Significant statistical differences ($p < 0.05$) between CF-5Y and M-5Y fields.

TABLE 3 The correlation coefficient and significant value of bacterial community and soil properties determined by the Mantel test.

Soil properties	CF-5Y		M-5Y	
	Correlation coefficient	Value of p	Correlation coefficient	Value of p
pH	0.1348	0.129	0.0739	0.205
OM	0.1212	0.136	0.6281	0.002*
OC	0.1308	0.098	0.6396	0.001*
ECe	0.1856	0.099	−0.0842	0.803
CEC	0.0255	0.417	0.1693	0.057
NH$_4$-N	0.2252	0.008	0.0693	0.254
NO$_3$-N	0.1883	0.045	0.1452	0.053
Avail.P	0.1282	0.122	0.2478	0.024*
Exch.K	0.1548	0.114	−0.0483	0.55
Exch.Ca	0.1670	0.045	0.0008	0.427
Exch.Mg	0.0681	0.209	0.4396	0.011*
Total.N	0.0185	0.390	0.3632	0.008*
BD	0.1551	0.117	0.4500	0.001*
Sand	0.1373	0.179	0.3514	0.004*
Silt	0.0668	0.247	−0.1987	0.934
Clay	−0.0551	0.609	0.5863	0.001*

followed by Actinobacteria (19.19%). At the last sampling time which is 1 year after the first sampling, the relative abundance of bacterial phyla changed from the previous year. The most abundant phyla in TP5 were Actinobacteria (25.32%), followed by Planctomycetes (20.70%) and Chloroflexi (14.18%) (Figure 4B).

At the genus level, *Bacillus* (5.89%), *Acidothermus* (3.84%), and *Conexibacter* (2.16%) dominated the bacterial community in TP1. *Arcobacter* were found only at this time point. In TP2 (rainy season), the abundance of HSB OF53-F07 (9.25%) and Candidatus *Udaeobacter* (6.43%) increased by approximately 8 and 6%, becoming the most abundant genera at this time point. After that, these two genera significantly decreased after 3 months (TP3). The most abundant genera in TP3 were *Bacillus* (5.89%) and *Gemmata* (2.29%). The latter significantly increased compared with TP1 and TP2. The most abundant genera in TP4 (winter) were Candidatus *Udaeobacter* (4.37%) and *Gemmata* (2.50%). In the last sampling time, *Gemmata* (3.61%) became the most abundant genus, followed by Candidatus *Udaeobacter* (2.47%) and *Bacillus* (2.19%) (Figure 6A). It is observed that the proportion of bacterial genera and the most abundant genera changed after 1 year.

Spearman's rank correlation showed that the abundance of several genera were significantly correlated with soil properties. For example, Candidatus *Udaeobacter* was positively correlated with Exch. Mg. HSB OF53-F07, and *Acidodermus* was positively correlated with OM, OC, BD, and sand, while *Gemmata* was negatively correlated with those properties (Figure 6B).

3.4. Predictive function

In this study, we used PICRUSt2 as a tool to predict bacterial-associated function and showed enzyme activity. Here, we found that

in CF-5Y, urease, chitinase, cellulase, and β-glucosidase significantly decreased in TP3 (rainy season). After that, urease continuously decreased until TP5 (summer—1 year after the first sampling), becoming the least abundant across all time points. On the other hand, chitinase, cellulase, and β-glucosidase significantly increased in TP4 (winter) and TP5 (summer). The abundance of amidase did not change during the rainy season (TP3–TP4), but it significantly decreased in TP4 (winter) and TP5 (summer—1 year after the first sampling). The abundance of acid- and alkaline-phosphatase is quite stable, but it significantly decreases after 1 year of the first sampling (Figure 7A). Spearman's rank correlation showed that chitinase, β-glucosidase, and urease positively correlated with OM and OC, while they negatively correlated with Avail. P, Exch. K, Exch. *Ca*. Cellulase positively correlated with NH$_4$-N and NO$_3$-N, whereas urease negatively correlated with these properties (Figure 7C).

In M-5Y, the abundance of chitinase, urease, and nitrogenase significantly decreased in the rainy season (TP2 or TP3). Chitinase slightly increased in TP5 (summer—1 year after the first sampling), but the urease and nitrogenase did not change. Their abundance is still low in TP4–TP5. On the other hand, xylan 1,4-β-xylosidase, and acid phosphatase significantly increased in the rainy season (TP2 and TP3), decreased in winter (TP4), then bounced back to a similar number of TP3 in summer (TP5). Overall, the abundance of these two enzymes in TP5 did not change from TP1 (Figure 7B). Spearman's rank correlation showed that chitinase, usease, and nitrogenase positively correlated with OM, OC, NH$_4$-N, NO$_3$-N, total N, and BD, whereas xylan 1,4-β-xylosidase, and acid phosphatase positively correlated with pH (Figure 7D).

4. Discussion

4.1. Stability of soil bacteria in undisturbed soil and continuous maize cultivation

The soil bacterial community plays several important roles in soil functioning. Our study found that there was more variation in the richness and diversity of bacteria in M-5Y, whereas more stability was observed in CF-5Y (Figure 3).

The undisturbed soil over a 5-year period demonstrated more stability across seasons compared with M-5Y, as evidenced by the consistent values of diversity indexes between the TP1 and TP5 sampling times (Figure 3). Proteobacteria displayed a higher abundance in CF-5Y (Figure 4A), consistent with the findings of Wan and He (2020). Their study revealed that Proteobacteria were more prevalent in the soil of a natural forest in Northwest China, which was attributed to the nutrient-rich soil resulting from the rapid turnover of OM. Proteobacteria are generally classified as copiotrophic organisms, thriving and reproducing in nutrient-rich environments (Li et al., 2020; Zheng et al., 2020). It is evident that the availability of nutrients in CF-5Y primarily originated from the decomposition of OM, ranging from 7.55 to 8.80% across the seasons (Table 1).

Agricultural land management practices directly and indirectly affect soil environments, resulting in the alteration of activity and diversity of soil bacterial communities (Bending et al., 2002; Bulluck and Ristaino, 2002; Steenwerth et al., 2003). In cultivated land, higher soil microbial diversity can enhance soil fertility and maintain soil nutrient balance, thereby influencing crop growth
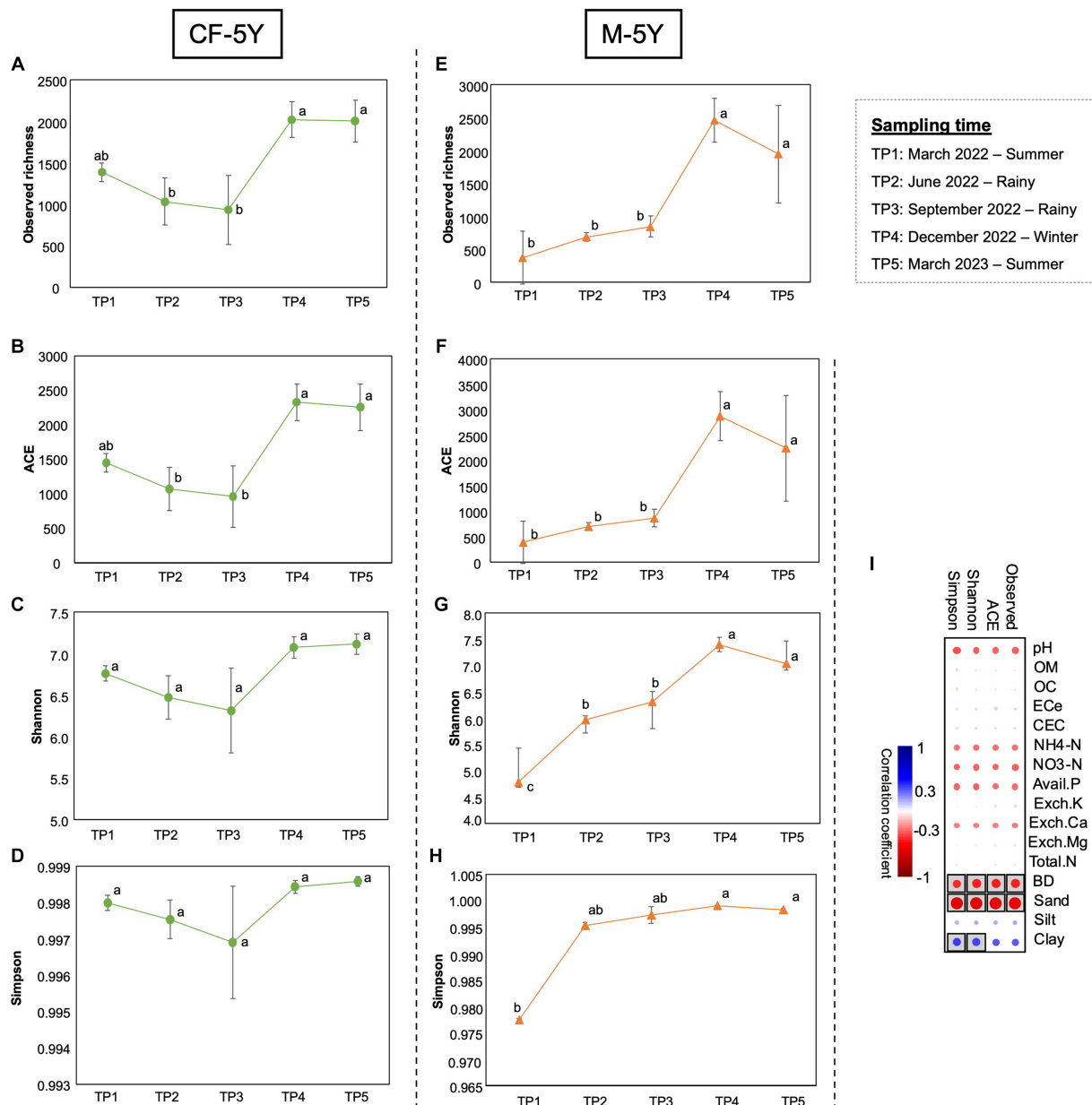
FIGURE 3
Alpha diversity indexes in CF-5Y **(A–D)** and M-5Y **(E-H)** across all time points. **(I)** Spearman's rank correlation between the alpha diversity indexes and soil properties. Circle with box is the significant correlation ($p < 0.05$).

(Turner et al., 2013). Zhang et al. (2021) reported that application fertilization, particularly nitrogen fertilizer, has a positive effect on maintaining the soil nutrients and increasing maize yield. Furthermore, tillage can enhance soil aeration and oxygen diffusion rates, potentially resulting in increased degradation of OM (Khan, 1996). This, in turn, could lead to elevated bacterial diversity and richness, along with enhanced soil enzyme activities (Zhou et al., 2019). Our study demonstrated that tillage practices in M-5Y during May had a notable impact, leading to a significant increase in the Shannon index from TP1 to TP2 (Figure 3G). This suggests that tillage practices could enhance both the number of bacteria species and the evenness of individual distribution among species, leading to increased diversity. The application of fertilizer

in August significantly influenced the increase in both bacteria richness and diversity, as evidenced by a significant rise in observed richness and ACE indexes (Figures 3E,F), as well as the Shannon index from TP3 to TP4 (Figure 3G).

At 1 year of the first sampling (TP5), the richness and diversity significantly increased and higher than TP1 (Figure 3), assuming that the remaining synthetic nutrients support the increase in the abundance of soil bacteria. This is in line with the finding in Figure 4B that Actinobacteria were the most abundant in M-5Y. Similar to Proteobacteria, Actinobacteria are classified as copiotrophic with faster growth rates and rapidly increased under nitrogen-rich conditions in cropland soils (Fierer et al., 2012; Dai et al., 2018). Calderoli et al. (2017) reported that Actinobacteria are
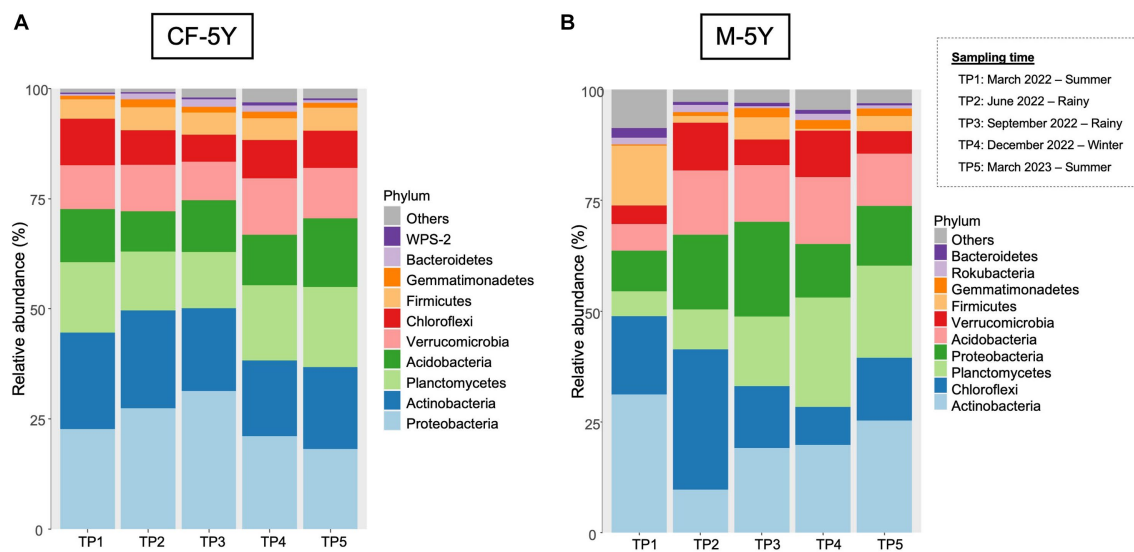
FIGURE 4
The most abundant bacterial phyla in **(A)** CF-5Y and **(B)** M-5Y.

involved in nitrogen fixation through nitrogenase (*nifH*) gene expression.

Soil texture plays a crucial role in shaping soil bacterial taxonomic diversity and composition (Hemkemeyer et al., 2018). As shown in Figure 3I, clay content displayed a positive correlation with diversity indexes, while sand content and soil bulk density exhibited negative correlations with richness indexes. This observation aligns with the findings of Seaton et al. (2020) who indicated that clay and silt contents have a greater influence on soil bacterial composition compared with sand particles. Notably, clay content can provide microhabitats that facilitate organic matter stabilization, water retention, and the proliferation of soil bacteria biomass (Sparling, 1992; Aponte et al., 2014). In contrast, higher soil bulk density reduces soil porosity and limits water availability and aeration, which, in turn, may lead to a decrease in the abundance of soil bacteria (Jensen et al., 1996; Li et al., 2002).

Soil enzymes play a vital role in the decomposition of OM and the cycling of nutrients within ecosystems (Uwituze et al., 2022). β-Glucosidase was the most dominant enzyme in both CF-5Y and M-5Y sites (Figure 7). According to Wick et al. (2002), β-glucosidase is involved in the mineralization and cycling of carbohydrates in the soil. It catalyzes the hydrolysis of glycosidic bonds in cellobiose, resulting in the production of glucose (Chen et al., 2021). Eivazi and Tabatabai (1990) found positive correlations of β-glucosidase with OC concentration, which is in line with our study (Figures 7C,D). Amidase was considered the second dominant soil enzyme in both CF-5Y and M-5Y sites (Figure 7). Amidase enzyme is important in the nitrogen cycle, as it catalyzes the hydrolysis of amides to carboxylates and ammonia (Frankenberger and Tabatabai, 1980; Fournand and Arnaud, 2001). Our study showed lower amidase enzymes in M-5Y compared with CF-5Y sites (Figures 7A,B). We hypothesized that the utilization of synthetic fertilizers and pesticides in M-5Y could potentially inhibit certain soil bacterial communities, including those that play a role in amidase production. Moreover, M-5Y contained lower OM accumulation.

According to Frankenberger and Dick (1983), alkaline phosphatase, amidase, and catalase have been identified as good indicators of soil microbial activity and biomass. Our study is consistent with this finding, as both alkaline phosphatase and amidase enzymes dominated in both CF-5Y and M-5Y fields (Figures 7A,B). This suggests that the differences between the 5-year maize monoculture and 5-year undisturbed soil had no negative impact on soil bacterial communities and soil enzyme activity.

4.2. Seasonal influence on soil bacterial richness, diversity, and genera

Seasonal fluctuations can impact microbial diversity and richness (Samaritani et al., 2017; Keet et al., 2019) as changes in precipitation and temperature influence microbial activity and nutrient availability (Wu et al., 2020).

In the CF-5Y site, there was no significant change in bacterial diversity throughout the seasons (Figures 3C,D). However, a significant increase in bacterial richness was observed from the rainy season (TP3) to the winter season (TP4) (Figures 3A,B). On the other hand, a notable increase in the Shannon diversity index was observed from the summer season to the rainy season in M-5Y (Figure 3G). This increase could be attributed to the rise in soil moisture, which likely stimulated the growth of various bacterial genera. Moreover, the application of chemical fertilizer in August, coupled with the subsequent increase in soil moisture, could have facilitated the enhancement of bacterial richness (Figures 3E,F) and Shannon diversity indexes (Figure 3G). The increase in soil moisture is linked to an accelerated decomposition rate and the subsequent release of nutrients, thereby enhancing nutrient availability (Brockett et al., 2012; Madegwa and Uchida, 2021).

Tillage practices and reduced soil covering expose the soil surface to variations in weather conditions, leading to fluctuations in soil bacteria throughout seasonal changes. This study focused on a
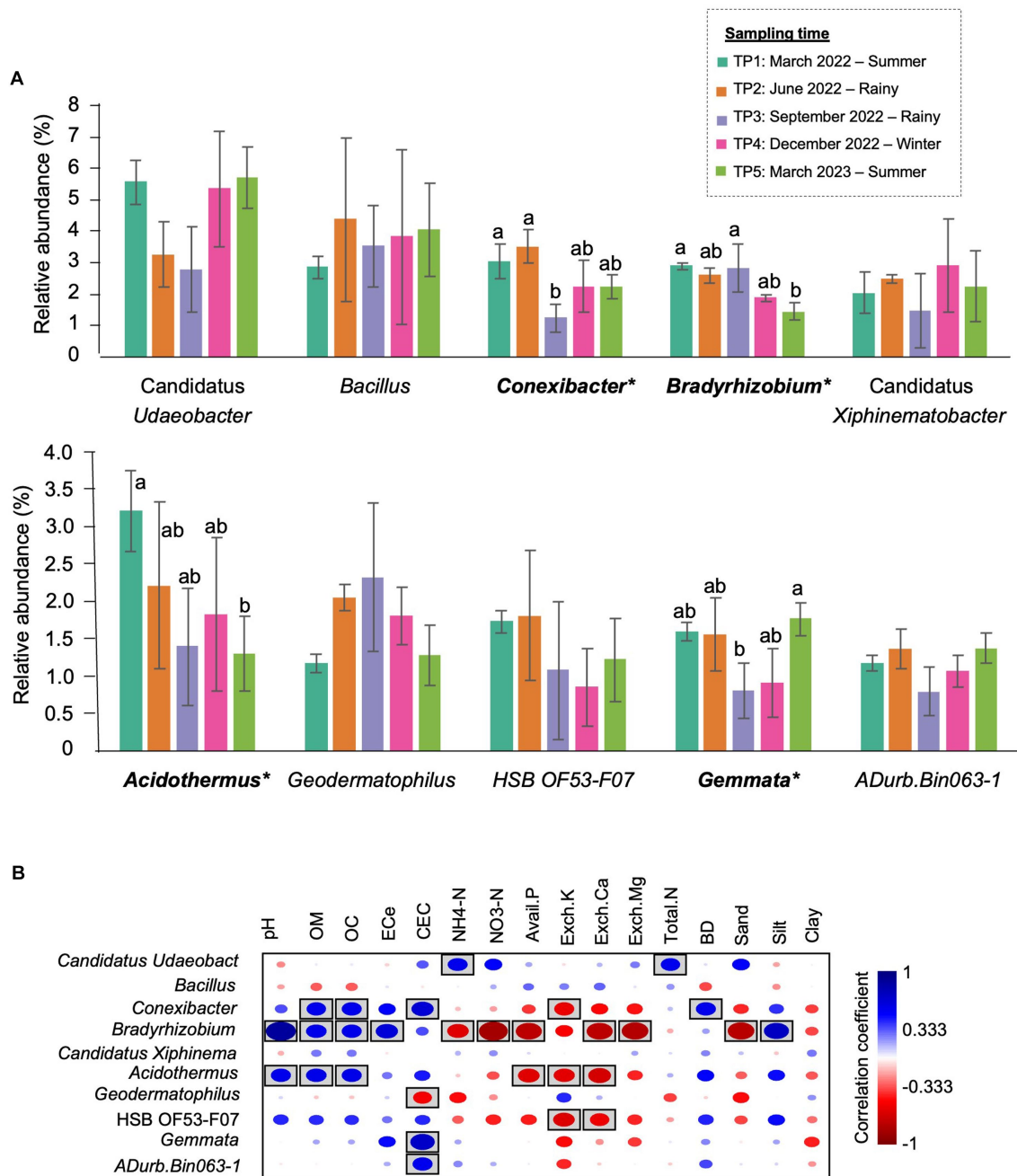
FIGURE 5
The most prevalent genera in CF-5Y and their correlation with soil properties. **(A)** The top 10 most abundant genera in CF-5Y. Taxa with asterisk (*) were statistically different between the sampling points ($p < 0.05$). **(B)** Spearman's rank correlation between the abundant genera in CF-5Y and soil properties. Circle with box is the significant correlation ($p < 0.05$).

mountainous area, where the relationship between bacterial community structure and temperature fluctuations is significant. However, it is important to note that the effects of temperature on different taxonomic groups vary. For example, Firmicutes had a relatively high abundance in the summer but decreased in the winter season, while Acidobacteria showed the opposite trend (Figure 4B). The lower abundance of Firmicutes in winter may be attributed to reduced activity. The most prevalent genera in CF-5Y remained more stable throughout the seasons compared with M-5Y (Figures 5A, 6A).

Meanwhile, some genera in M-5Y disappeared in certain seasons. Particularly, *Saccharopolyspora* and *Arcobacter* appeared only during the summer season (Figure 6A).

At the genus level, Candidatus *Udaeobacter* was identified as the most dominant during the summer season (TP1 and TP5), as well as in the winter season (TP4) in both CF-5Y and M-5Y sites. However, during the rainy season (TP2 and TP3), the dominant genus shifted to *Bacillus* in both CF-5Y and M-5Y sites (Figure 5A). Candidatus *Udaeobacter* belongs to the Verrucomicrobia phylum, which is
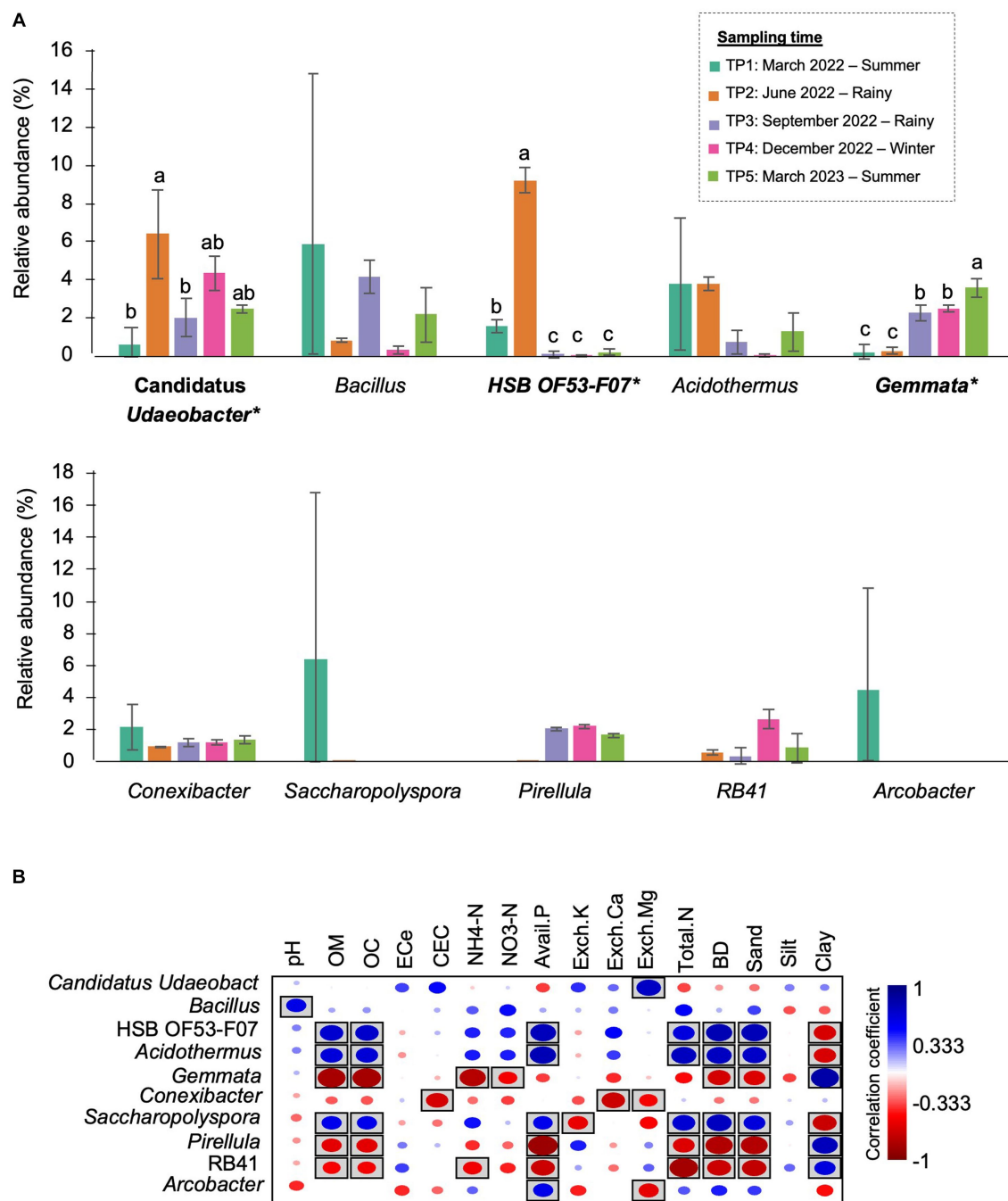
FIGURE 6
The most prevalent genera in M-5Y and their correlation with soil properties. **(A)** The top 10 most abundant genera in M-5Y. Taxa with asterisk (*) were statistically different between the sampling points (p < 0.05). **(B)** Spearman's rank correlation between the abundant genera in M-5Y and soil properties. Circle with box is the significant correlation (p < 0.05).

composed of aerobic heterotrophs (Kadnikov et al., 2021). Candidatus *Udaeobacter* plays a role as amino acid and vitamin transporters, which involves sacrificing metabolic versatility in favor of storing surplus carbon as glycogen or starch. This alternative strategy contributes to its dominance in the soil environment even under resource-limiting conditions and high competition for labile carbon (Brewer et al., 2016). Huang et al. (2012) observed a decrease in the abundance of Verrucomicrobia, following an increase in available N,

P, and K resulting from cotton straw application. Similarly, Lima et al. (2014) and Navarrete et al. (2015) found that higher verrucomicrobial community abundance is related to soils containing lower P and K contents. This may explain the abundance of Candidatus *Udaeobacter* during the summer and winter seasons, as these periods coincide with low available P and exchangeable K compared with the rainy season (Table 2). On the other hand, *Bacillus* sp. are the most dominant plant growth-promoting rhizobacteria (Tsotetsi et al., 2022), which colonize
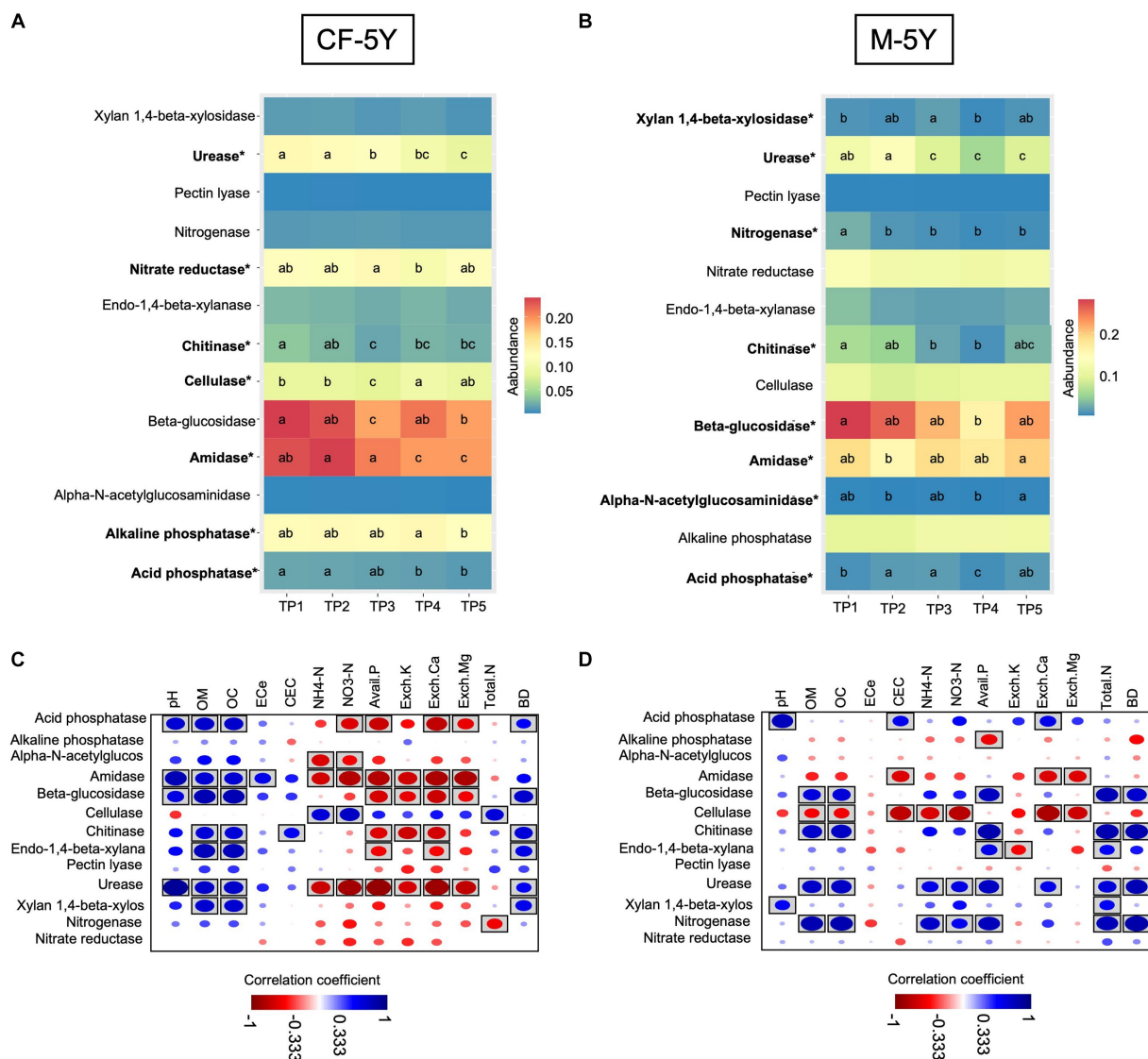
FIGURE 7
The abundance of soil enzyme predicted by PiCRUSt2 in **(A)** CF-5Y and **(B)** M-5Y and Spearman's rank correlation between soil enzymes and soil properties in **(C)** CF-5Y and **(D)** M-5Y. Enzymes' names with an asterisk (*) were statistically different between the sampling points ($p < 0.05$). Circle with box is the significant correlation ($p < 0.05$).

the area around plant roots and play a role in nutrient uptake (nitrogen fixation, solubilization, and mineralization of phosphorus and other nutrients) (Goswami et al., 2016). These abilities indicate that *Bacillus* can rapidly grow in nutrient-rich conditions during the rainy season. This is consistent with the observation of higher TN, available P, and exchangeable K during the rainy season compared with the summer and winter seasons (Tables 1, 2).

Seasonal variation in soil enzymes is driven by changes in soil moisture and temperature dynamics (Wallenstein et al., 2009; Machmuller et al., 2016). In our study, we observed a significant decrease in various soil enzymes during the rainy season in both CF-5Y (urease, chitinase, cellulase, and β-glucosidase) (Figure 7A) and M-5Y (chitinase, urease, and nitrogenase) sites (Figure 7B). This could be attributed to the effects of leaching and dilution. Heavy rainfall has the potential to leach soluble organic compounds and nutrients from

the soil, causing these enzymes to be washed away from the active decomposition zone. Additionally, heavy rainfall can dilute the concentration of OM and nutrients in the soil, resulting in a reduced supply of substrates for the enzymes to act upon. Ladwig et al. (2015) reported that the rainy season could lead to a decrease in photosynthetic availability, potentially resulting in a reduced rate of expression of new enzymes compared with the loss of enzymes through decomposition processes. However, certain enzymes (chitinase, cellulase, and β-glucosidase) exhibited an increase during the summer seasons (Figures 7A,B), potentially attributed to the favorable combination of optimum temperature and soil moisture conditions. This is consistent with the study by Tomar and Baishya (2020), who reported that soil enzymes, with the exception of β-glucosidase activity, exhibited higher levels during the monsoon season in a semi-arid forest of Delhi, India, due to the optimum moisture and temperature conditions.

It is important to note that, unlike in a laboratory experiment, shifts in soil bacterial diversity can be driven by a combination of various physical factors, including land management practices, soil physicochemical properties, weather conditions, and vegetation cover. The presence and recovery of soil bacteria might have played important roles in the soil decomposition process, warranting further studies. The CF-5Y field, with its high soil moisture content (Table 1), could mitigate the fluctuations in daily and seasonal soil temperatures and potentially create a localized microclimate for soil bacteria. This, in turn, may contribute to the stability of the soil bacterial community throughout the changing seasons.

Our findings provide insights into the relative stability of soil bacteria in undisturbed soil compared with a maize cultivation field. Considering the context of climate change, with expectations of rising temperatures and shifting seasons in Northern Thailand, it is noteworthy that the dominant taxa in the bacterial communities of undisturbed soil and maize fields were found to be relatively similar. Therefore, it may be suggested that changing climate conditions may not lead to a significant shift in the soil bacterial community.

5. Conclusion

Our study found that undisturbed soil over a 5-year period exhibited more stability in the richness and diversity of bacteria across seasons compared with M-5Y. Fertilizer application and tillage practices in M-5Y influenced the diversity and richness of soil bacteria. Proteobacteria were the most abundant in CF-5Y, while Actinobacteria dominated in M-5Y. At the genus level, Candidatus *Udaeobacter* was the most dominant during the summer and winter seasons in both CF-5Y and M-5Y sites. However, during the rainy season, the dominant genus shifted to *Bacillus* in both CF-5Y and M-5Y fields. The soil bacterial community in M-5Y was strongly influenced by OM and OC, whereas in CF-5Y, there was no correlation between soil properties and the soil bacterial community. β-Glucosidase was the dominant enzyme in both CF-5Y and M-5Y sites, and it showed a positive correlation with OM and OC. Tillage practices and reduced soil covering in the M-5Y field exposed the soil surface to variations in weather conditions, leading to variations in soil bacteria throughout seasonal changes such as *Saccharopolyspora* and *Arcobacter,* which appeared only during the summer season. This study highlights that the shifts in soil bacterial diversity can be driven by a combination of land management practices, soil physicochemical properties, weather conditions, and vegetation cover. Longer time periods should be continuously investigated, as our study only covered a 1-year pattern of soil bacterial dynamics in these two environments.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: National Center for Biotechnology Information (NCBI) under the BioProject accession number: PRJNA1003674, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1003674.

Author contributions

NA: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing. CS: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. SS: Conceptualization, Investigation, Writing – original draft. RH: Conceptualization, Methodology, Supervision, Writing – original draft.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1285445/full#supplementary-material

References

Aponte, C., Matías, L., González-Rodríguez, V., Castro, J., García, L. V., Villar, R., et al. (2014). Soil nutrients and microbial biomass in three contrasting Mediterranean forests. *Plant Soil* 2014, 57–72. doi: 10.1007/s11104-014-2061-5

Arunrat, N., Sereenonchai, S., and Hatano, R. (2022b). Effects of fire on soil organic carbon, soil total nitrogen, and soil properties under rotational shifting cultivation in Northern Thailand. *J. Environ. Manag.* 302:113978. doi: 10.1016/j.jenvman.2021.113978

Arunrat, N., Sereenonchai, S., Kongsurakan, P., and Hatano, R. (2022a). Soil organic carbon and soil erodibility response to various land-use changes in Northern Thailand. *Catena* 219:106595. doi: 10.1016/j.catena.2022.106595

Arunrat, N., Sereenonchai, S., Kongsurakan, P., Iwai, C. B., Yuttitham, M., and Hatano, R. (2023a). Post-fire recovery of soil organic carbon, soil total nitrogen, soil nutrients, and soil erodibility in rotational shifting cultivation in Northern Thailand. *Front. Environ. Sci.* 11:1117427. doi: 10.3389/fenvs.2023.1117427

Arunrat, N., Sereenonchai, S., Kongsurakan, P., Yuttitham, M., and Hatano, R. (2023b). Variations of soil properties and soil surface loss after fire in rotational shifting cultivation in Northern Thailand. *Front. Environ. Sci.* 11:1213181. doi: 10.3389/fenvs.2023.1213181

Barbour, K. M., Weihe, C., Allison, S. D., and Martiny, J. B. H. (2022). Bacterial community response to environmental change varies with depth in the surface soil. *Soil Biol. Biochem.* 172:108761. doi: 10.1016/j.soilbio.2022.108761

Barreiro, A., and Díaz-Raviña, M. (2021). Fire impacts on soil microorganisms: mass, activity, and diversity. *Curr. Opin. Environ. Sci. Health.* 22:100264. doi: 10.1016/j.coesh.2021.100264

Bending, G. D., Turner, M. K., and Jones, J. E. (2002). Interactions between crop residue and soil organic matter quality and the functional diversity of soil microbial communities. *Soil Biol. Biochem.* 34, 1073–1082. doi: 10.1016/S0038-0717(02)00040-8

Bray, R. A., and Kurtz, L. T. (1945). Determination of total organic and available form of phosphorus in soil. *Soil Sci.* 59, 39–45. doi: 10.1097/00010694-194501000-00006

Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A., and Fierer, N. (2016). Genome reduction in an abundant and ubiquitous soil bacterium '*Candidatus* Udaeobacter copiosus'. *Nat. Microbiol.* 2:16198. doi: 10.1038/nmicrobiol.2016.198

Brockett, B. F. T., Prescott, C. E., and Grayston, S. J. (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biol. Biochem.* 44, 9–20. doi: 10.1016/j.soilbio.2011.09.003

Bulluck, L. R., and Ristaino, J. B. (2002). Effect of synthetic and organic soil fertility amendments on southern blight soil microbial communities and yield of processing tomatoes. *Phytopathology* 92, 181–189. doi: 10.1094/PHYTO.2002.92.2.181

Calderoli, P. A., Collavino, M. M., Behrends Kraemer, F., Morrás, H. J. M., and Aguilar, O. M. (2017). Analysis of *nif*H-RNA reveals phylotypes related to *Geobacteria* and cyanobacteria as important functional components of the N₂-fixing community depending on depth and agricultural use of soil. *MicrobiologyOpen* 6:e502. doi: 10.1002/mbo3.502

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

Caon, L., Vallejo, V. R., Ritsema, C. J., and Geissen, V. (2014). Effects of wildfire on soil nutrients in Mediterranean ecosystems. *Earth Sci. Rev.* 139, 47–58. doi: 10.1016/j.earscirev.2014.09.001

Certini, G. (2005). Effects of fire on properties of forest soils: a review. *Oecologia* 143, 1–10. doi: 10.1007/s00442-004-1788-8

Chen, A., Wang, D., Ji, R., Li, J., Gu, S., Tang, R., et al. (2021). Structural and catalytic characterization of TsBGL, a β-glucosidase from *Thermofilum* sp. ex4484_79. *Front. Microbiol.* 12:723678. doi: 10.3389/fmicb.2021.781826

Cordovez, V., Dini-Andreote, F., Carrión, V. J., and Raaijmakers, J. M. (2019). Ecology and evolution of plant microbiomes. *Annu. Rev. Microbiol.* 73, 69–88. doi: 10.1146/annurev-micro-090817-062524

Dai, Z., Su, W., Chen, H., Barberán, A., Zhao, H., Yu, M., et al. (2018). Long-term nitrogen fertilization decreases bacterial diversity and favors the growth of actinobacteria and proteobacteria in agro-ecosystems across the globe. *Glob. Chang. Biol.* 24, 3452–3461. doi: 10.1111/gcb.14163

Das, S. K., and Varma, A. (2011). "Role of enzymes in maintaining soil health" in *Soil Enzymology*. eds. G. Shukla and A. Varma, vol. *22* (Berlin/Heidelberg, Germany: Springer), 25–42.

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi: 10.1038/s41587-020-0548-6

Egamberdiyeva, D. (2007). The effect of plant growth promoting bacteria on growth and nutrient uptake of maize in two different soils. *Appl. Soil Ecol.* 36, 184–189. doi: 10.1016/j.apsoil.2007.02.005

Eivazi, F., and Tabatabai, M. A. (1990). Factors affecting glucosidase and galactosidase activities in soils. *Soil Biol. Biochem.* 22, 891–897. doi: 10.1016/0038-0717(90)90126-K

Ekasingh, B., Gypmantasiri, P., Thong-Ngam, K., and Grudloyma, P. (2004). *Maize in Thailand: Production Systems, Constraints, and Research Priorities*. Mexico, DF: CIMMYT: International Maize and Wheat Improvement Center, Maize Production Systems Papers.

Essel, B., Abaidoo, R. C., Opoku, A., and Ewusi-Mensah, N. (2020). Economically optimal rate for nutrient application to maize in the semi-deciduous Forest zone of Ghana. *J. Soil Sci. Plant Nutr.* 20, 1703–1713. doi: 10.1007/s42729-020-00240-y

Estaki, M., Jiang, L., Bokulich, N. A., McDonald, D., González, A., Kosciolek, T., et al. (2020). QIIME 2 enables comprehensive end-to-end analysis of diverse microbiome data and comparative studies with publicly available data. *Curr. Protoc. Bioinformatics* 70:e100. doi: 10.1002/cpbi.100

Fierer, N., Lauber, C. L., Ramirez, K. S., Zaneveld, J., Bradford, M. A., and Knight, R. (2012). Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J.* 6, 1007–1017. doi: 10.1038/ismej.2011.159

Fournand, D., and Arnaud, A. (2001). Aliphatic and enantioselective amidases: from hydrolysisto acyl transfer activity. *J. Appl. Microbiol.* 91, 381–393. doi: 10.1046/j.1365-2672.2001.01378.x

Frankenberger, W. T. Jr., and Dick, W. A. (1983). Relationships between enzyme activities and microbial growth and activity indices in soil. *Soil Sci. Soc. Am. J.* 47, 945–951. doi: 10.2136/sssaj1983.03615995004700050021x

Frankenberger, W. T. Jr., and Tabatabai, M. A. (1980). Amidase activity in soils: I. Method of assay. *Soil Sci. Soc. Am. J.* 44, 282–287. doi: 10.2136/sssaj1980.03615995004400020016x

Goswami, D., Thakker, J. N., and Dhandhukia, P. C. (2016). Portraying mechanics of plant growth promoting rhizobacteria (PGPR): a review. *Cogent Food Agric* 2:1127500. doi: 10.1080/23311932.2015.1127500

Gouda, S., Kerry, R. G., Das, G., Paramithiotis, S., Shin, H.-S., and Patra, J. K. (2018). Revitalization of plant growth promoting rhizobacteria for sustainable development in agriculture. *Microbiol. Res.* 206, 131–140. doi: 10.1016/j.micres.2017.08.016

Hammer, O., Harper, D., and Ryan, P. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4, 1–9.

Hemkemeyer, M., Dohrmann, A. B., Christensen, B. T., and Tebbe, C. C. (2018). Bacterial preferences for specific soil particle size fractions revealed by community analyses. *Front. Microbiol.* 9:149. doi: 10.3389/fmicb.2018.00149

Higashida, S., and Takao, K. (1985). Seasonal fluctuation patterns of microbial numbers in the surface soil of a grassland. *Soil Sci. Plant Nutr.* 31, 113–121. doi: 10.1080/17470765.1985.10555222

Huang, W., Bai, Z., Hoefel, D., Hu, Q., Lv, Q., Zhuang, G., et al. (2012). Effects of cotton straw amendment on soil fertility and microbial communities. *Front. Environ. Sci. Eng.* 6, 336–349. doi: 10.1007/s11783-011-0337-z

Jensen, L. S., McQueen, D. J., and Shepherd, T. G. (1996). Effects of soil compaction on N mineralization and microbial-C and -NI field measurements. *Soil Tillage Res.* 38, 175–188. doi: 10.1016/S0167-1987(96)01033-1

Kadnikov, V. V., Mardanov, A. V., Beletsky, A. V., Grigoriev, M. A., Karnachuk, O. V., and Ravin, N. V. (2021). Thermophilic Chloroflexi dominate in the microbial community associated with coal-fire gas vents in the Kuznetsk Coal Basin, Russia. *Microorganisms*. 9:948. doi: 10.3390/microorganisms9050948

Keet, J. H., Ellis, A. G., Hui, C., and Le Roux, J. J. (2019). Strong spatial and temporal turnover of soil bacterial communities in South Africa's hyperdiverse fynbos biome. *Soil Biol. Biochem.* 136:107541. doi: 10.1016/j.soilbio.2019.107541

Khan, A. R. (1996). Influence of tillage on soil aeration. *J. Agron. Crop Sci.* 177, 253–259. doi: 10.1111/j.1439-037X.1996.tb00243.x

Kim, N., Zabaloy, M. C., Guan, K., and Villamil, M. B. (2020). Do cover crops benefit soil microbiome? A meta-analysis of current research. *Soil Bio Biochem* 142:107701. doi: 10.1016/j.soilbio.2019.107701

Kleinman, P. J. A., Pimentel, D., and Bryant, R. B. (1995). The ecological sustainability of slash-and-burn agriculture. *Agric. Ecosyst. Environ.* 52, 235–249.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41:e1. doi: 10.1093/nar/gks808

Ladwig, L. M., Sinsabaugh, R. L., Collins, S. L., and Thomey, M. L. (2015). Soil enzyme responses to varying rainfall regimes in Chihuahuan Desert soils. *Ecosphere.* 6:40. doi: 10.1890/ES14-00258.1

Li, S. F., Huang, X. B., Lang, X. D., Shen, J., and Su, J. (2020). Cumulative effects of multiple biodiversity attributes and abiotic factors on ecosystem multifunctionality in the Jinsha river valley of southwestern China. *For. Ecol. Manag.* 472:118281. doi: 10.1016/j.foreco.2020.118281

Li, C. H., Ma, B. L., and Zhang, T. Q. (2002). Soil bulk density effects on soil microbial populations and enzyme activities during the growth of maize (*Zea mays* L.) planted in large pots under field exposure. *Can. J. Plant Sci.* 82, 147–154. doi: 10.4141/S01-026

Lima, A. B., Cannavan, F. S., Navarrete, A. A., Teixeira, W. G., Kuramae, E. E., and Tsai, S. M. (2014). Amazonian dark earth and plant species from the Amazon region

contribute to shape rhizosphere bacterial communities. *Microb. Ecol.* 69, 855–866. doi: 10.1007/s00248-014-0472-8

Liu, C., Cui, Y., Li, X., and Yao, M. (2021). Microeco: an R package for data mining in microbial community ecology. *FEMS Microbiol. Ecol.* 97:fiaa255. doi: 10.1093/femsec/fiaa255

Liu, X., Liu, Y., Zhang, L., Yin, R., and Wu, G.-L. (2021). Bacterial contributions of bio-crusts and litter crusts to nutrient cycling in the mu us Sandy land. *Catena* 199:105090. doi: 10.1016/j.catena.2020.105090

Machmuller, M. B., Mohan, J. E., Minucci, J. M., Phillips, C. A., and Wurzburger, N. (2016). Season, but not experimental warming, affects the activity and temperature sensitivity of extracellular enzymes. *Biogeochemistry* 131, 255–265. doi: 10.1007/s10533-016-0277-6

Madegwa, Y. M., and Uchida, Y. (2021). Land use and season drive changes in soil microbial Communities and related functions in agricultural soils. *Environmental DNA* 3, 1214–1228. doi: 10.1002/edn3.244

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. doi: 10.14806/ej.17.1.200

Mehring, M., Glaser, B., de Camargo, P. B., and Zech, W. (2011). Impact of forest organic farming change on soil microbial C turnover using 13C of phospholipid fatty acids. *Agronomy Sust. Developm.* 31, 719–731. doi: 10.1007/s13593-011-0013-5

Meng, M., Lin, J., Guo, X., Liu, X., Wu, J., Zhao, Y., et al. (2019). Impacts of forest conversion on soil bacterial community composition and diversity in subtropical forests. *Catena* 175, 167–173. doi: 10.1016/j.catena.2018.12.017

Miah, S., Dey, S., and Sirajul Haque, S. M. (2010). Shifting cultivation effects on soil fungi and bacterial population in Chittagong Hill tracts, Bangladesh. *J. For. Res.* 21, 311–318. doi: 10.1007/s11676-010-0076-1

Miah, S., Haque, S. M. S., Sumi, W., and Hossain, M. M. (2014). Effects of shifting cultivation on biological and biochemical characteristics of soil microorganisms in Khagrachari hill district, Bangladesh. *J. For. Res.* 25, 689–694. doi: 10.1007/s11676-014-0453-2

Moebius-Clune, B. N., Van Es, H. M., Idowu, O. J., Schindelbeck, R. R., Kimetu, J. M., Ngoze, S., et al. (2011). Long-term soil quality degradation along a cultivation chronosequence in western Kenya. *Agric. Ecosyst. Environ.* 141, 86–99. doi: 10.1016/j.agee.2011.02.018

National Soil Survey Center. (1996). *Soil Survey Laboratory Methods Manual. Soil Survey Investigations Report No. 42, Version 3.0, USDA*. National Soil Survey Center Washington, DC, USA.

Navarrete, A. A., Diniz, T. R., Braga, L. P. P., Silva, G. G. Z., Franchini, J. C., Rossetto, R., et al. (2015). Multi-analytical approach reveals potential microbial indicators in soil for sugarcane model systems. *PLoS One* 10:e0129765. doi: 10.1371/journal.pone.0129765

Newbold, T., Hudson, L. N., Contu, S., Hill, S. L. L., Beck, J., Liu, Y., et al. (2018). Widespread winners and narrow-ranged losers: land use homogenizes biodiversity in local assemblages worldwide. *PLoS Biol.* 16:e2006841. doi: 10.1371/journal.pbio.2006841

Pieterse, C. M. J., de Jonge, R., and Berendsen, R. L. (2016). The soil-borne supremacy. *Trends Plant Sci.* 21, 171–173. doi: 10.1016/j.tplants.2016.01.018

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R Development Core. (2019). *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 1.

Rerkasem, K., and Rerkasem, B. (1995). Montane mainland South-East Asia: agroecosystems in transition. *Glob. Environ. Change.* 5, 313–322. doi: 10.1016/0959-3780(95)00065-V

Samaritani, E., Mitchell, E. A. D., Rich, J., Shrestha, J., Fournier, B., and Frey, B. (2017). Soil bacterial communities and ecosystem functioning change more strongly with season than habitat in a restored floodplain. *Appl. Soil Ecol.* 112, 71–78. doi: 10.1016/j.apsoil.2016.12.010

Sarkar, D., Meitei, C. B., Baishya, L. K., Das, A., Ghosh, S., Chongloi, K. L., et al. (2015). Potential of fallow chronosequence in shifting cultivation to conserve soil organic carbon in Northeast India. *Catena* 135, 321–327. doi: 10.1016/j.catena.2015.08.012

Schloter, M., Dilly, O., and Munch, J. C. (2003). Indicators for evaluating soil quality. *Agric. Ecosyst. Environ.* 98, 255–262. doi: 10.1016/S0167-8809(03)00085-9

Schulz, S., Brankatschk, R., Dümig, A., Kögel-Knabner, I., Schloter, M., and Zeyer, J. (2013). The role of microorganisms at different stages of ecosystem development for soil formation. *Biogeosciences* 10, 3983–3996. doi: 10.5194/bg-10-3983-2013

Seaton, F. M., George, P. B., Lebron, I., Jones, D. L., Creer, S., and Robinson, D. A. (2020). Soil textural heterogeneity impacts bacterial but not fungal diversity. *Soil Biol. Biochem.* 144:107766. doi: 10.1016/j.soilbio.2020.107766

Shahzad, T., Rashid, M. I., Maire, V., Barot, S., Perveen, N., Alvarez, G., et al. (2018). Root penetration in deep soil layers stimulates mineralization of millennia-old organic carbon. *Soil Biol. Biochem.* 124, 150–160. doi: 10.1016/j.soilbio.2018.06.010

Smith, N. R., Kishchuk, B. E., and Mohn, W. W. (2008). Effects of wildfire and harvest disturbances on Forest soil bacterial communities. *Appl. Environ. Microbiol.* 74, 216–224. doi: 10.1128/AEM.01355-07

Sparling, G. (1992). Ratio of microbial biomass carbon to soil organic carbon as a sensitive indicator of changes in soil organic matter. *Soil Res.* 30, 195–207. doi: 10.1071/SR9920195

Steenwerth, K. L., Jackson, L. E., Calderón, F. J., Stromberg, M. R., and Scow, K. M. (2003). Soil microbial community composition and land use history in cultivated and grassland ecosystems of coastal California. *Soil Biol. Biochem.* 35, 489–500. doi: 10.1016/S0038-0717(03)00028-2

Szoboszlay, M., Dohrmann, A. B., Poeplau, C., Don, A., and Tebbe, C. C. (2017). Impact of land-use change and soil organic carbon quality on microbial diversity in soils across Europe. *FEMS Microbiol. Ecol.* 93:fix146. doi: 10.1093/femsec/fix146

Tomar, U., and Baishya, R. (2020). Seasonality and moisture regime control soil respiration, enzyme activities, and soil microbial biomass carbon in a semi-arid forest of Delhi, India. *Ecol. Proces.* 9:50. doi: 10.1186/s13717-020-00252-7

Tsotetsi, T., Nephali, L., Malebe, M., and Tugizimana, F. (2022). Bacillus for plant growth promotion and stress resilience: what have we learned? *Plan. Theory* 11:2482. doi: 10.3390/plants11192482

Turner, T. R., James, E. K., and Poole, P. S. (2013). The plant microbiome. *Genome Biol* 14:209. doi: 10.1186/gb-2013-14-6-209

USDA. (1954). *Diagnosis and Improvement of Saline and Alkali Soils, Agriculture. Handbook No. 60, U.S. Salinity Laboratory*. Government Printing Office, Washington, DC.

Uwituze, Y., Nyiraneza, J., Fraser, T. D., Dessureaut-Rompré, J., Ziadi, N., and Lafond, J. (2022). Carbon, nitrogen, phosphorus, and extracellular soil enzyme responses to different land use. *Front. Soil Sci.* 2:814554. doi: 10.3389/fsoil.2022.814554

Walkley, A., and Black, J. A. (1934). An examination of the dichromate method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Sci.* 37, 29–32. doi: 10.1097/00010694-193401000-00003

Wallenstein, M. D., McMahon, S. K., and Schimel, J. P. (2009). Seasonal variation in enzyme activities and temperature sensitivities in Arctic tundra soils. *Glob. Chang. Biol.* 15, 1631–1639. doi: 10.1111/j.1365-2486.2008.01819.x

Wan, P., and He, R. (2020). Soil microbial community characteristics under different vegetation types at the national nature reserve of Xiaolongshan Mountains, Northwest China. *Eco. Inform.* 55:101020. doi: 10.1016/j.ecoinf.2019.101020

Wapongnungsang, O. E., Upadhyay, K. K., and Tripathi, S. K. (2021). Soil fertility and rice productivity in shifting cultivation: impact of fallow lengths and soil amendments in Lengpui, Mizoram Northeast India. *Heliyon.* 7:e06834. doi: 10.1016/j.heliyon.2021.e06834

Whitbread, A., Blair, G., Naklang, K., Lefroy, R., Wonprasaid, S., Konboon, Y., et al. (1999). The management of rice straw, fertilisers and leaf litters in rice cropping systems in Northeast Thailand: 2. Rice yields and nutrient balances. *Plant Soil* 209, 21–28.

Wick, B., Kühne, R. F., Vielhauer, K., and Vlek, P. L. (2002). Temporal variability of selected soil microbiological and biochemical indicators under different soil quality conditions in South-Western Nigeria. *Biol Fert Soils.* 35, 155–167. doi: 10.1007/s00374-002-0455-7

Williams, D. R., Alvarado, F., Green, R. E., Manica, A., Phalan, B., and Balmford, A. (2017). Land-use strategies to balance livestock production, biodiversity conservation and carbon storage in Yucatan, Mexico. *Glob. Chang. Biol.* 23, 5260–5272. doi: 10.1111/gcb.13791

Wu, J., Sha, C., Wang, M., Ye, C., Li, P., and Huang, S. (2021). Effect of organic fertilizer on soil bacteria in maize fields. *Land.* 10:328. doi: 10.3390/land10030328

Wu, K., Xu, W., and Yang, W. (2020). Effects of precipitation changes on soil bacterial community composition and diversity in the Junggar desert of Xinjiang China. *PeerJ.* 8:e8433. doi: 10.7717/peerj.10403

Yang, Z., Xu, Y., Li, H., Li, S., Wang, X., and Chai, H. (2022). Difference of bacterial community structure in the meadow, maize, and continuous cropped alfalfa in Northeast China. *Front. Microbiol.* 13:794848. doi: 10.3389/fmicb.2022.1082025

Zhang, M., Zhang, X., Zhang, L., Zeng, L., Liu, Y., Wang, X., et al. (2021). The stronger impact of inorganic nitrogen fertilization on soil bacterial community than organic fertilization in short-term condition. *Geoderma* 382:114752. doi: 10.1016/j.geoderma.2020.114752

Zheng, X., Lin, C., Guo, B., Yu, J., Ding, H., Peng, S., et al. (2020). Effects of re-vegetation restoration on soil bacterial community structure in degraded land in subtropical China. *Eur. J. Soil Biol.* 98:103184. doi: 10.1016/j.ejsobi.2020.103184

Zhou, Y., Zhou, B., Xu, F., Muhammad, T., and Li, Y. (2019). Appropriate dissolved oxygen concentration and application stage of micro-nano bubble water oxygation in greenhouse crop plantation Agric. *Water Manag.* 223:105713. doi: 10.1016/j.agwat.2019.105713

Check for updates

Isolation of starch and protein degrading strain *Bacillus subtilis* FYZ1-3 from tobacco waste and genomic analysis of its tolerance to nicotine and inhibition of fungal growth

Changwen Ye[1,2†], Dandan Liu[2*†], Kuo Huang[2], Dong Li[2], Xinxin Ma[3], Yiying Jin[3] and Hanguo Xiong[1]

[1]College of Food Science and Technology, Huazhong Agricultural University, Wuhan, China, [2]China Tobacco Standardization Research Center, Zhengzhou Tobacco Research Institute of CNTC, Zhengzhou, China, [3]School of Environment, Tsinghua University, Beijing, China

Aerobic fermentation is an effective technique for the large-scale processing of tobacco waste. However, the specificity of the structure and composition of tobacco-derived organic matter and the toxic alkaloids in the material make it currently difficult to directly use microbial agents. In this study, a functional strain FYZ1-3 was isolated and screened from thermophilic phase samples of tobacco waste composting. This strain could withstand temperatures as high as 80°C and grow normally at 0.6% nicotine content. Furthermore, it had a strong decomposition capacity of tobacco-derived starch and protein, with amylase activity of 122.3 U/mL and protease activity and 52.3 U/mL, respectively. To further understand the mechanism of the metabolic transformation of the target, whole genome sequencing was used and the secondary metabolite gene cluster was predicted. The inhibitory effect of the strain on common tobacco fungi was verified using the plate confrontation and agar column methods. The results showed that the strain FYZ1-3 was *Bacillus subtilis*, with a genome size of 4.17 Mb and GC content of 43.68%; 4,338 coding genes were predicted. The genome was annotated and analyzed using multiple databases to determine its ability to efficiently degrade starch proteins at the molecular level. Moreover, 14 functional genes related to nicotine metabolism were identified, primarily located on the distinct genomic island of FYZ1-3, giving a speculation for its nicotine tolerance capability on the molecular mechanism. By mining the secondary metabolite gene cluster prediction, we found potential synthetic bacteriocin, antimicrobial peptide, and other gene clusters on its chromosome, which may have certain antibacterial properties. Further experiments confirmed that the FYZ1-3 strain was a potent growth inhibitor of *Penicillium chrysogenum*, *Aspergillus sydowii*, *A. fumigatus*, and *Talaromyces funiculosus*. The creation and industrial use of the functional strains obtained in this study provide a theoretical basis for its industrial use, where it would be of great significance to improve the utilization rate of tobacco waste.

1 Introduction

China is the largest producer and consumer of tobacco in the world (Zhang G. et al., 2022). The amount of tobacco waste generated in 2020 has exceeded 2.6 million tons (Zhang Y. et al., 2022). Tobacco waste is a valuable biomass resource (Wang J. et al., 2022), while traditional treatment means often involve random piling or centralized incineration, which not only leads to serious secondary pollution but also to resource waste (Dobaradaran et al., 2020; Koutela et al., 2020; Lee et al., 2021). Studies have confirmed that the aerobic fermentation of tobacco waste is an effective method for the large-scale utilization of waste resources, and exogenous microbial agents are usually added for high-temperature aerobic fermentation to fasten the process of composting and production of organic fertilizer that has added value (Kayikçioglu and Okur, 2011; Awasthi et al., 2017). Among them, *Bacillus subtilis* has become the most widely used microbial strain owing to its rapid growth rate, high tolerance, and strong degradation ability (Nayak, 2020). The screening and identification of *Bacillus* spp. to make complex microbial agents can significantly shorten the composting cycle and improve the degradation rate of sludge, garden waste, food waste, etc. (Fang et al., 2019; Liu et al., 2022a,b; Wang M. et al., 2022).

However, tobacco waste usually contains toxic alkaloids, among which, nicotine accounts for >95% (Zhou, 2008). The U.S. Environmental Protection Agency had added nicotine to the list of environmentally restricted release inventories in 1994, and the European Union also introduced laws stipulating that tobacco waste with a mass fraction exceeding 0.05% should be classified as hazardous waste (Ashengroph, 2015; Weishaar et al., 2016). Tobacco waste usually contains 0.1%–4.0% (mass fraction) of nicotine (Civilini et al., 1997), and some studies have shown that nicotine has a certain inhibitory effect on microbial metabolic pathways, gene expression, cell reproduction, enzyme activity, and other biological functions (Yildiz, 2004; Fitzpatrick, 2018; Nölke et al., 2021). Additionally, there are significant differences in the structure and background impurities between tobacco-derived starch and proteins and ordinary starch and proteins according to current studies (Masson and Rossignol, 1995; Song et al., 2010; Wang et al., 2012). Although there are some studies and applications on efficient organic matter-degrading bacterial agents, there are no relevant reports on bacterial agents that can degrade tobacco-derived organics and withstand nicotine at high temperatures. Therefore, it is necessary to identify a functional strain that can simultaneously withstand high temperatures and nicotine content, as well as have a high degradation ability of tobacco source organic matter.

In this study, tobacco waste was selected as the sampling object, and tobacco starch and tobacco proteins were used as the sole carbon sources as the enrichment medium to isolate the strains that could degrade tobacco-derived organic matter at 80°C. The strain with the maximum amylase and protease activities was selected as the functional strain. This study utilized tobacco waste composting samples during the high-temperature period to screen bacteria, using tobacco starch and tobacco protein as the only carbon sources. Enrichment took place at 80°C, followed by preliminary screening using a transparent hydrolysis ring and enzyme activity determination to obtain a strain capable of withstanding high temperatures, degrading tobacco organic matter, and displaying nicotine tolerance. The functional strain with the greatest comprehensive amylase and

protease activity was chosen as the target for further analysis. Moreover, the tolerance of functional strains to nicotine levels was further determined. The classification status of functional strain was determined using phylogenetic analysis based on 16S rRNA gene sequencing and average nucleotide identity (ANI) analysis.

Studying the basic genomic information of functional strains can help researchers further understand the mechanism of metabolic transformation of functional strains toward target substances, and also facilitate the in-depth understanding of the relationship between genes and proteins, metabolic function, and individual behavior. Therefore, the whole genome sequence of the functional strain was determined to determine its ability to efficiently degrade starch and protein at the molecular level, and to further explore whether there is a secondary metabolite gene cluster on its chromosome that synthesizes bacteriocins and antimicrobial peptides. The growth-inhibition ability of common tobacco fungi was verified using the confrontation plate culture and agar column methods.

The thermophilic strain obtained in this study was resistant to nicotine and had high degradation ability to tobacco source organics. Using whole genome sequencing, it was found that it also has a certain antibacterial effect, producing organic fertilizers capable of biological control, which providing a theoretical basis for the creation and industrial application of functional microbials in the later stage. Thus, this study could be of significance for improving the utilization rate of tobacco waste resources.

2 Materials and methods

2.1 Materials

2.1.1 Source of samples

Compost samples were obtained from tobacco waste piles in their natural state from Luliang County, Qujing City, Yunnan Province of China.

2.1.2 Culture media

Bacterial culture medium: LB medium (tryptone 10 g/L, yeast extract g/L, NaCl 10 g/L, pH 7.2–7.4); fungal medium: potato dextrose agar (PDA) medium (potato 200 g/L, glucose 20 g/L, natural pH). Agar powder (20 g/L) was added to the solid medium and sterilized at 121°C for 20 min.

Tobacco starch-screening medium: tobacco starch 1 g/L, NaCl 5 g/L, yeast powder 2 g/L, agar powder 15 g/L. Tobacco-derived protein-screening medium: tobacco protein 2 g/L, sodium chloride 5 g/L, agar powder 15 g/L.

Nicotine-fermentation medium: 100% nicotine was added to LB liquid medium according to the volume ratio to prepare nicotine media (0.3, 0.6, 0.9, 1.2, and 1.5%).

2.2 Methods

2.2.1 Isolation and screening of the functional strain

Sampling and enrichment: 5 g of the pile was weighed using the cross-sampling method. Next, 45 mL of sterilized phosphate buffer solution was added and the sample was homogenized for 10 min. The

mixture was heated at 80°C in a water bath for 15 min. Then the heated liquid was centrifuged at 3000 rpm/min for 10 min and 1 mL of the supernatant was added to a liquid medium containing 50 mL of tobacco starch and that of tobacco protein, respectively. The two media were incubated at 37°C and 180 r/min for 48 h.

Screening of bacterial strains: after heating the above bacterial solution in a water bath at 80°C for 15 min, gradient dilution was performed. Next, 100 μL of bacterial suspensions with five dilutions of 10^{-4} to 10^{-8} were drawn and inoculated on tobacco starch or tobacco protein screening medium and cultured at 37°C for 48 h in a constant temperature and humidity box. The tobacco starch culture medium was dripped with iodine solution, and single-colony strains with larger transparent circles were visually selected. The diameter of the transparent circles and colony diameters were measured using a digital vernier caliper and the ratio of the two (HC value) was calculated.

Determination of enzyme activity: using the iodine colorimetric method, the amylase secreted by the bacterial strain interacts with starch under certain conditions. The reaction solution is detected by a spectrophotometer at a specific time to color with iodine, and the results are obtained by referring to the alpha-amylase activity table. Specific operating procedures were based on China National Standard (GB1886.174-2016, 2016). Protease activity is often measured using casein as a substrate to produce amino acids under specific hydrolysis conditions. Its activity is measured by the amount of tyrosine produced, which is evaluated by the Folin-phenol reagent method. Specific operating procedures were based on China National Standard (GB/T 23527-2009, 2009).

Determination of nicotine tolerance: 100 μL of strains in the logarithmic phase were added to different concentrations of the nicotine fermentation medium and cultured in a shaker at 37°C and 220 rpm for 16 h, followed by measurement of the OD_{600}. An equal volume of sterile water was used as the blank control group.

2.2.2 Functional strain identification and whole genome sequencing

The genomic DNA of FYZ1-3 was extracted using the FastPure Gel DNA Extraction Mini Kit (Vazyme, Nanjing, China) according to the manufacturer's instructions. The pair of bacterial universal primer (upstream primer 27F: 5′-AGAGTTTGATCCTGGCTCAG-3′; downstream primer 1492R: 5′-GGYCCCTTGTTACGATT-3′) was used as the 16S rDNA amplification primer. Polymerase chain reaction (PCR) amplification reaction system (50 μL): 2 μL upstream and downstream primers (10 pmol/μL), 2 μL 10× PCR buffer (containing Mg^{2+}), 0.6 μL Taq DNA polymerase, 4 μL dNTP mix. PCR amplification reaction conditions: 94°C, pre-denaturation for 5 min; denaturation at 94°C for 30 s, annealing at 55°C for 30 s (renaturation), extension at 72°C for 1 min, 30 cycles; 72°C for 10 min. PCR products were sent to Shenzhen Huada Gene Co., Ltd. for sequencing and the results were compared with the sequence in the NCBI database for BLAST homology.

Genomic DNA extraction and database construction: the genomic DNA of the screened functional strain, FYZ1-3, was extracted using the STE method, which was comprised of a mixture of sodium chloride-Tris-HCl-EDTA to create a protective environment for the DNA (Pan et al., 2006; Shui et al., 2014). In addition, the purity and integrity of the DNA were evaluated through agarose gel electrophoresis and quantified utilizing Qubit.

The SMRT Bell library was constructed using the SMRT Bell™ template kit. The constructed library was quantified by Qubit concentration, and the size of the inserted fragment was detected using Agilent 2100. The PacBio platform was utilized for sequencing purposes whereby the PacBio RS II featured single molecule real-time sequencing technology. Upon the complex creation with DNA and polymerase and its containment within the ZMW (zero mode waveguide), four diverse fluorescent labeled dNTPs entered the detection area randomly through Brownian motion and bind with the polymerase. Chemical bonds required more time to form with the template-matching base in comparison to the other bases. The duration of fluorescence information's existence could discern between matched and free bases. Analyzing the relationship between time and four fluorescence information types enabled sequencing of DNA template sequences.

2.2.3 Validation of the fungal-inhibition ability of the functional strain

Confrontation plate culture method: four typical tobacco pathogens, *Penicillium chrysogenum*, *Aspergillus sydowii*, *A. fumigatus* and *Talaromyces funiculosus*, were selected as indicators. The active bacterial solution was coated on PDA plates and incubated at 28°C for 72 h. After a single colony was established, the suitable bacteria were selected by means of a sterilized puncher and shaped into a 5 mm bacterial cake. The cake was first connected to a fresh PDA plate and the functional strain was spotted at a distance of 2.5 cm from the cake; a double-distilled water (ddH₂O)-spotted plate was used as a control. Fungi were cultured in a constant-temperature incubator at 28°C for 72 h, and fungal growth was recorded.

Agar column method: the fermentation supernatant of the functional strain was added to the PDA medium in a 1:10 (v/v) ratio, and the activated indicator bacterial cake was punched with a 5 mm sterilization punch and placed in the center of the PDA medium. A plate with the same amount of ddH₂O was used as a control. The cells were cultured in a constant temperature incubator at 28°C for 48 h. The cross-crossing method was used to estimate the scattered area of bacteria and spores. The antibacterial rate was calculated as follows:

$$\text{Inhibition rate}\,(\%) = \frac{\begin{array}{c}\text{Bacterial cake area in the blank group} - \\ \text{Bacterial cake area in the experimental group}\end{array}}{\text{Bacterial cake area in the blank group}} \times 100 \quad (1)$$

2.2.4 Functional annotation of genes

The default algorithm of GeneMarkS (Version 4.17) was used to predict the encoding genes of the sequenced genome. Based on sequence composition, the IslandPath-DIOMB software (Version 0.2) was used to predict gene islands, which determine gene islands and potential horizontal gene transfer by detecting nucleotide bias and mobility genes (such as transposases or integrases) in the sequence. With a library based on the specified type of HMM algorithm built, the antiSMASH 4.0.2 program was used to predict and identify all known secondary metabolic clusters in the genome. The annotated universal functional databases used in this study include GO, KEGG, and CAZy. The basic steps of functional annotation are as follows: (1) perform diamond alignment of the predicted gene protein sequence with various functional

databases(evaluation $\leq 1e^{-5}$); (2) filter alignment results: for each sequence's alignment results, select the alignment result with the highest score (default identity ≥40%, coverage ≥40%) for annotation. Lastly, whole genome sequencing data of the measured functional strains were submitted to the NCBI database, and the GenBank accession number is SRR24954540.

3 Results

3.1 Screening and characterization of functional strain FYZ1-3

A total of 31 culturable single colonies were obtained through enrichment culture of tobacco waste compost samples at high-temperature period. Next, 10 strains with clear degradation abilities on starch and protein medium were then selected based on the hydrolysis transparent circle diameter to colony diameter ratio (H/C) on selective medium, as shown in Table 1. Finally, the amylase activity and protease activity of these 10 strains were determined, and the rescreening results were shown in Figure 1A. It could be seen that FYZ1-3 exhibit the highest amylase activity of 122.3 U/mL among all ten strains, which was considerably higher than that of the other strains (20.0–48.4 U/mL). The protease activity of 10 bacterial strains ranged from 20.7 to 90.2 U/mL. Despite not displaying the highest protease activity, strain FYZ1-3 demonstrated a value of 52.3 U/mL, which was not significantly different from the highest recorded value. Furthermore, its overall enzyme activity was the highest among all strains evaluated. Therefore, FYZ1-3 was chosen as the functional strain for the subsequent experiments. The FYZ1-3 strain was inoculated in a nicotine fermentation medium with varying concentration gradients, and the OD_{600} was measured after 16 h. The results indicated that the FYZ1-3 strain could tolerate a nicotine concentration of 0.6%, as shown in Figure 1B. Compared with the same type of high temperature and nicotine resistant *B. subtilis* B5221 (Wang et al., 2012), the amylase activity of FYZ1-3 was 122.3 U/mL, which was much higher than the highest enzyme activity of 39.9 U/mL after optimization of culture conditions. Although the protease activity of FYZ1-3 was only 52.3 U/mL, it still exceeded that of the H010 strain screened by Li et al. (2012), which measured 39.2 U/mL. The FYZ1-3 strain was demonstrated extensive functionality, thriving in an environment with a temperature of 80°C and a nicotine content of 0.6%. This strain also exhibited a superior capacity to degrade tobacco starch and protein. As shown in Figures 1C–E, the analysis of single bacterial characteristics, including colony morphology, microscopic cell structure, and Gram staining,

identified strain FYZ1-3 as *Bacillus* sp. The colony appeared opaque and milky white with visible surface folds. Rod-shaped cells with visible spore structures were observed under the microscope, and Gram staining was positive.

3.2 Molecular biology-based identification and whole genome sequencing of functional strains

### 3.2.1 Phylogenetic tree and ANI analysis based On 16S rRNA gene sequencing

16S rRNA gene sequencing results of the selected strain FYZ1-3 were uploaded to the NCBI database for BLAST comparison. The similarity of the FYZ1-3 strain with the model strains *B. subtilis* DSM 10, *B. subtilis* NBRC 13719, and *B. subtilis* BCRC 10255 was 100%. The phylogenetic tree was constructed using the neighbor-joining method and MEGA software (version 11) as shown in Figure 2A. *B. subtilis* FYZ1-3 was clustered with other *B. subtilis* spp. and had the closest relationship with *B. subtilis* DSM 10 strain. Among them, based on the information in the GeneBank database, *B. subtilis* DSM 10 strain (GeneBank No.: CP120681.1) was isolated from fermentation materials in western Africa. This high-temperature fermentation environment is similar to the high-temperature fermentation environment of tobacco waste, which may be a reason why DSM 10 and FYZ1-3 strains are closely related. *B. subtilis* DSM10 has been used as a model strain and is widely used as a reference strain in various studies (Lilge et al., 2021).

ANI is mainly used to evaluate the genetic relationship between species at the genome-wide level (Arahal, 2014). Based on the results from 16S rRNA gene sequencing, *B. subtilis* DSM 10 having the highest similarity with the 16S rRNA gene sequence of FYZ1-3 was selected. The complete genome sequence was downloaded from the Assembly database of NCBI, and the ANI value between the FYZ1-3 and *B. subtilis* DSM 10 strains was calculated using JSpecies. Figure 2B shows the homology between FYZ1-3 and *B. subtilis* DSM 10 to be 98.56%, which is higher than the classification threshold of 95%. It was determined that FYZ1-3 and *B. subtilis* DSM 10 were the same species; therefore, it was named *B. subtilis* FYZ1-3.

3.2.2 Genome overview and gene island prediction of the FYZ1-3 strain

PacBio Sequel is a third-generation sequencing platform based on nanopore single-molecule real-time technology, which has the advantages of long sequencing reads, high-throughput analysis, and high accuracy (Mclaughlin et al., 2021). In this study, the whole

TABLE 1 Results of initial screening of strains from selected medium.

Item		k8-1	k11-1	k12-2	k13-1	k13-3	k13-4	k14-2	k15-1	FX19-4	FYZ1-3
Starch selective medium	H/C	2.82	2.22	2.57	2.48	2.31	3.09	2.51	2.24	2.06	2.74
	Image										
Protein selective medium	H/C	1.43	1.17	1.48	1.32	1.15	1.30	1.20	1.17	3.40	1.15
	Image										

FIGURE 1
Screening process and characteristics of the FYZ1-3 strain: **(A)** starch and protein enzyme activity; **(B)** nicotine tolerance; **(C)** colony morphology; **(D)** microscope images; **(E)** Gram staining.

FIGURE 2
Analysis and identification of strain FYZ1-3: **(A)** 16S rRNA phylogenetic tree; **(B)** ANI value analysis diagram of DSM 10 with the closest genetic relationship.
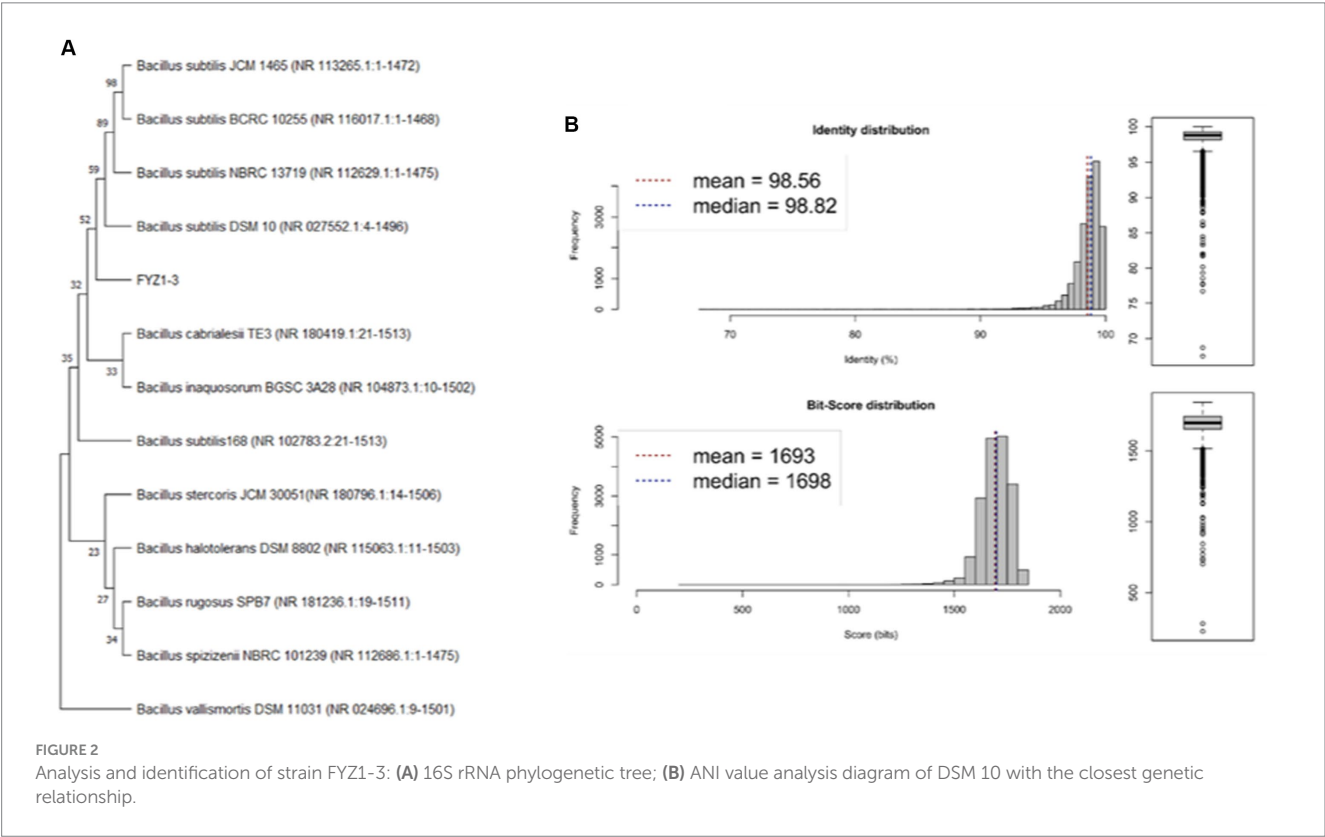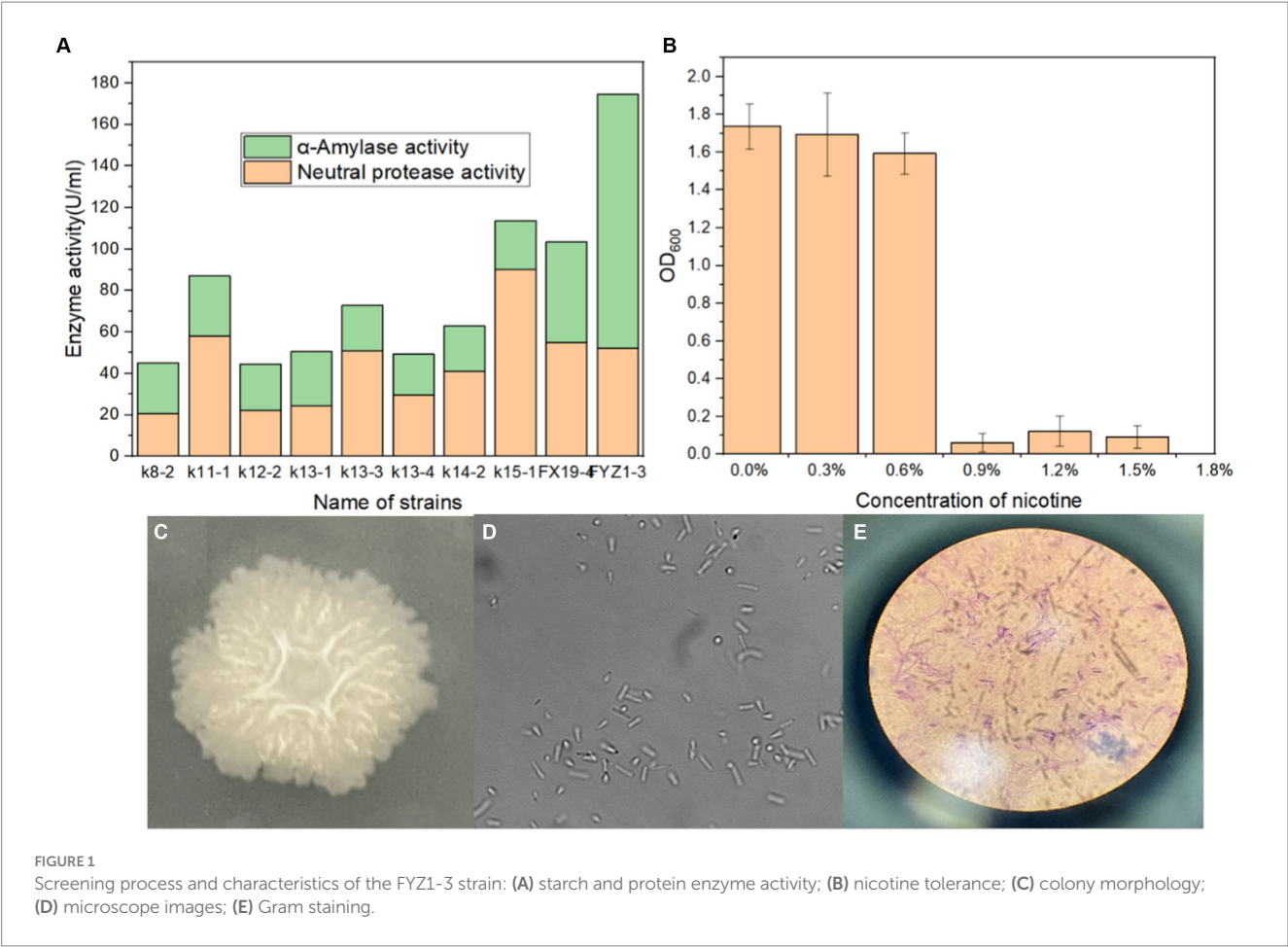
genome of strain FYZ1-3 was sequenced using the PacBio Sequel sequencing platform, and the genome assembly of the sequence was performed using SMRT Link assembly software to obtain 1 contig. After the data were corrected by Illumina Nova Seq PE150 sequencing results, the chromosome sequence of strain FYZ1-3 was assembled into a circular genome. The whole genome spectrum is shown in Figure 3A. The genome size of the FYZ1-3 strain was 4,167,567 bp, the average GC content was 43.68%, and there were no plasmids. In addition, the genome also contains 86 tRNA genes and 30 rRNA genes.

Genomic island is a gene cluster that can be horizontally transferred in the microbial genome. The transfer of genomic islands is a way of microbial gene exchange, which can improve the diversity of microorganisms and their adaptability to the environment. It is of great significance in the evolutionary analysis of bacteria and the study of special functions that may be obtained during evolution (Dhillon et al., 2015). As shown in Table 2, 15 genomic islands were predicted on the chromosome of strain FYZ1-3. Among the 15 genomic islands, GIs011 was the longest genomic island with a length of 47,019 bp. GIs013 was the shortest genomic island with a length of 5,517 bp. Comparing the average GC content of the genome with 43.68%, it was found that the GC content of the other genomic islands except GIs 006 was lower than the average GC content of the genome. These 15 genomic islands carry a large number of additional functional genes, and the encoded proteins are mainly related to core metabolism, metal ion metabolism, motility chemotaxis, and quorum sensing.

3.2.3 Genome difference analysis and collinearity analysis of FYZ1-3 and DSM 10

The chromosome of strain FYZ1-3 was circular with a length of 4,167,567 bp and average GC content of 43.68%. Ori-Finder 2022 was used to annotate the replication origin of the chromosome of the FYZ1-3 strain. The results showed that there were two replication origins in the chromosome. The length of replication origin 1 was 993 nt, the AT content was 0.62, and the location was 4,166,656–4,167,648 bp, which was composed of 15 Dna A boxes. The length of replication origin 2 was 188 nt, the AT content was 0.63, and the location was 1,342–1,529 nt, which was composed of 8 Dna A boxes. From Figure 3B, it can be seen that the chromosome of strain FYZ1-3 has high similarity with most regions of the chromosome of the B. subtilis DSM 10 strain and has high homology. However, the chromosome of the FYZ1-3 strain also has some differences; regions 1–4 are unique. After comparison with the predicted genomic islands, regions 1–4 correspond to GIs 004, GIs 008, GIs 011, and GIs 012, respectively.

In general, if the relationship between two species is close, the sequence and order of genes will be very close and is called collinearity (Magain et al., 2017). Collinearity mainly reflects the structural variation between genomes, reflecting the coding sequence and structural homology between genomes. Based on the phylogenetic tree analysis, the model strain B. subtilis DSM 10 with the closest homology was selected for Mauve collinearity alignment with the genome of the screening strain FYZ1-3, and the large fragment sequence rearrangement between the genomes was quickly analyzed.
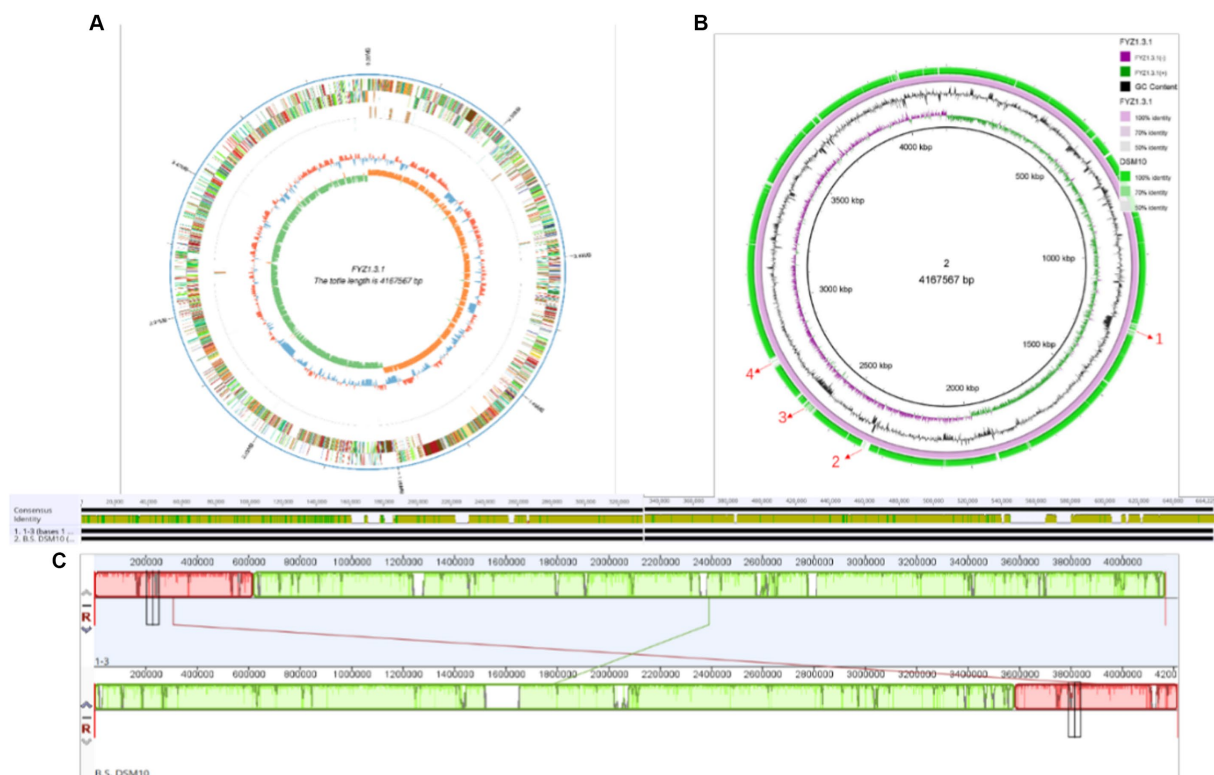
Genome analysis of the strain FYZ1-3: **(A)** whole genome map of strain FYZ1-3; **(B)** comparison of genomic differences between the FYZ1-3 and B. subtilis DSM 10 strains; **(C)** analysis of genome collinearity between the FYZ1-3 and B. subtilis DSM 10 strains.

It can be seen from Figure 3C that collinearity between strains FYZ1-3 and *B. subtilis* DSM 10 was poor, and there were obvious changes in gene rearrangement time during deletion, inversion, and translocation, indicating that FYZ1-3 strain may evolve a set of phenomena to adapt to the new environmental genome during the high-temperature fermentation of nicotine-containing tobacco waste.

TABLE 2 Prediction results of genomic islands contained in chromosome 1 of FYZ1-3.

Genomic island number	Start position	End position	GI_length (bp)	GC%
GIs 001	115,497	150,760	35,264	42.33
GIs 002	526,592	534,784	8,193	34.16
GIs 003	632,327	642,064	9,738	36.97
GIs 004	1,235,447	1,278,734	43,288	33.98
GIs 005	1,449,628	1,458,581	8,954	38.95
GIs 006	1,674,943	1,683,536	8,594	43.79
GIs 007	1,792,145	1,801,079	8,935	37.59
GIs 008	1,902,091	1,923,017	20,927	38.93
GIs 009	2,081,031	2,096,390	15,360	35.64
GIs 010	2,352,227	2,382,751	30,525	40.45
GIs 011	2,571,091	2,618,109	47,019	36.13
GIs 012	2,793,519	2,808,954	15,436	42.15
GIs 013	3,671,848	3,677,364	5,517	35.49
GIs 014	3,993,864	4,004,575	10,712	33.56
GIs 015	4,132,293	4,142,551	10,259	35.41

3.3 Functional annotation of degrading genes in the functional strain

3.3.1 GO database annotation

The predicted genes were compared with the GO database using BLAST to obtain the classification annotation information of genes, including cellar components (CC), molecular functions (MF), and biological processes (BP). The statistical results of strain FYZ1-3 in the GO database are shown in Figure 4A. There were 2,926 gene annotations in this classification and a total of 49 functional classifications. According to the MF annotation, 347 genes were involved in FYZ1-3 in the GO-CC classification. The first three functions were cell, cell part, and cell membrane. According to the MF annotation, 1,130 genes were involved in strain FYZ1-3 in the GO-MF classification. The first three functions were catalytic activity, binding, and transport activity, and the gene with amylase activity was *GM000743*. According to the classification of BP, there were 650 GO-BP classifications of genes, and the three most annotated pathways were metabolic processes, cellular processes, and single-cell processes.

3.3.2 KEGG database annotation

KEGG is a database that systematically analyzes the metabolic pathways of gene products in cells and the functions of these gene products. KEGG can be used to further study the complex behavior of genes in the field of biology (Aoki-Kinoshita and Kanehisa, 2007). The genome sequence of strain FYZ1-3 was compared with the KEGG database using BLAST, and the functional annotation results were obtained. A total of 2,502 genes in the KEGG database were functionally annotated on six major functional pathways, namely, cell processes, environmental information processing, genetic information processing, human disease, metabolism, and organism system. The

FIGURE 4
Gene function annotation of the FYZ1-3 strain: **(A)** GO database functional classification; **(B)** KEGG database function classification; **(C)** CAZy database functional classification.
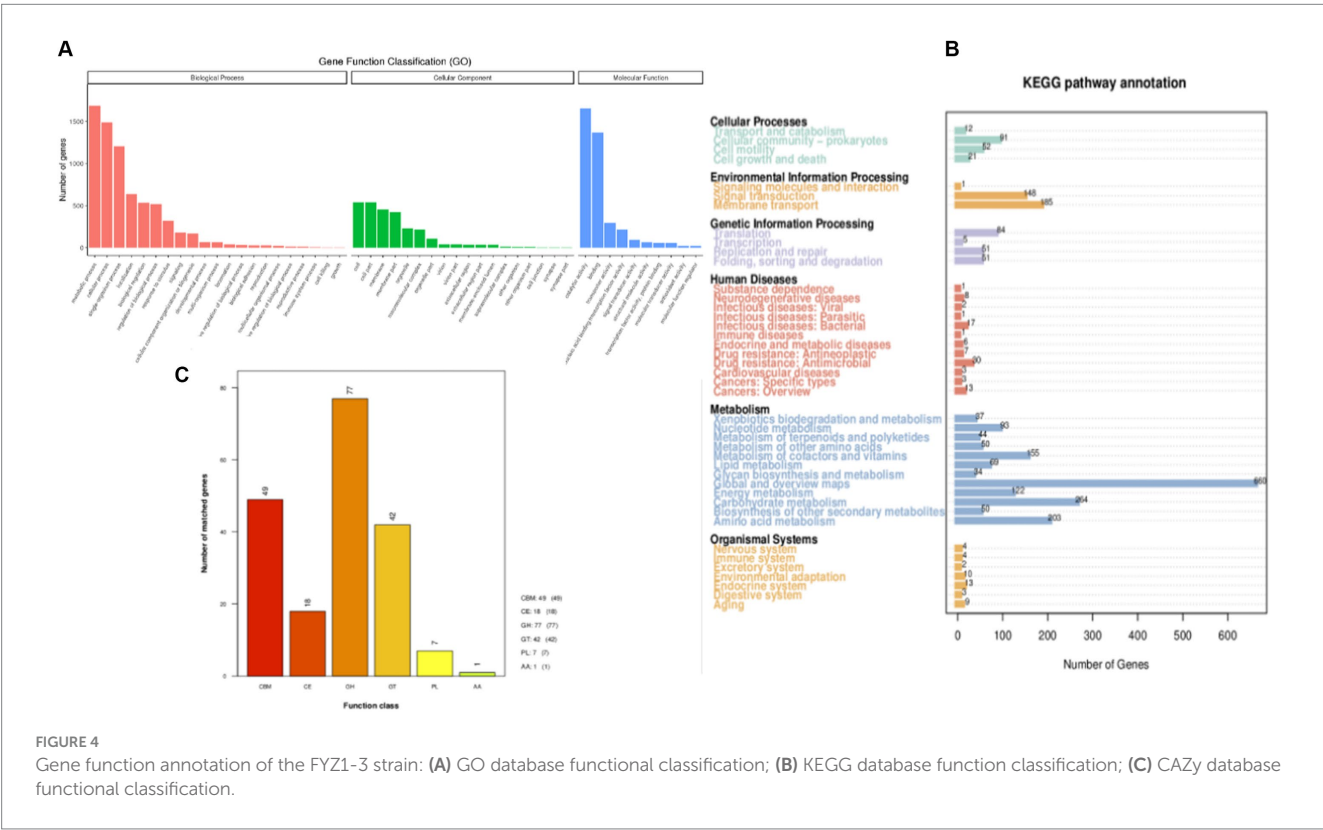
results are shown in Figure 4B. Among them, 1781 genes were annotated in the metabolic pathway. Among the 12 metabolic pathways, 264 genes were related to carbohydrate metabolism, accounting for 14.82% of the annotated genes in the metabolic pathway. Table 3 shows the pathway information of genes related to carbohydrate and protein metabolism. It can be seen that the genome of strain FYZ1-3 contains multiple metabolic pathways and involves a large number of genes, which is one of the reasons why it is efficient in degrading starch and proteins. In addition, 134 genes were annotated at the level of environmental information processing, including 54 genes related to membrane transport and 54 genes related to the two-component signal transduction system. A total of 334 genes were annotated at the level of environmental information processing, including 185 genes related to membrane transport and 148 genes related to the signal transduction system.

Furthermore, the KEGG database identified 14 functional genes associated with nicotine metabolism in Table 4. Compared with the genome island information in Table 2, it could be seen that the ppnK gene of FYZ1-3 strain was located on GIs004. Additionally, this strain contained gene clusters namely nadE, nadD, nadA, nadC, and nadB which were related to nicotinamide metabolism, with nadD, nadA, nadC, and nadB located on GIs011 and the pncB gene found on GIs012. From Figure 3B, it was evident that the FYZ1-3 strain possessed three distinctive genomic islands that were absent in the DSM 10 strain. These genes might be responsible for the molecular mechanism of nicotine tolerance in the functional strain.

3.3.3 CAZy database annotation

The CAZy database includes carbohydrate-related enzyme families that catalyze carbohydrate degradation, modification, and biosynthesis. The enzyme family is divided into the following six categories: glycoside hydrolase (GH), polysaccharide lyase, carbohydrate esterase, glycosyltransferase, auxiliary activity (AA), and carbohydrate-binding module (CBM). The annotation results are shown in Figure 4C. There are 194 CAZy enzyme gene families encoded by the FYZ1-3 strain, of which the GH family accounts for the highest (39.7%) of the whole gene family, followed by the CBM family, accounting for approximately 25.3%. GH, as a key component of carbohydrate metabolism, can catalyze the hydrolysis of glycosidic bonds in various glycosyl compounds. CBM is a noncatalytic protein domain that binds to carbohydrates. It is a polysaccharide-binding module that does not have catalytic activity but participates in carbohydrate degradation. It can specifically bind to polysaccharides and effectively improve the catalytic efficiency of carbohydrate-degrading enzymes. The main components of tobacco waste are carbohydrates and proteins, and the FYZ1-3 strain contains genes encoding endoglucanase (lichenase/endo-beta-1,3-1,4-glucanase [EC 3.2.1.73], β-glucosidase [β-glucosidase, EC 3.2.1.21], α-amylase [amyE, EC 3.2.1.1], and other enzymes). From the perspective of gene analysis, the degradation mechanism of the FYZ1-3 strain on tobacco waste was further understood.

3.4 Biological control effect of the functional strain

3.4.1 Prediction of the secondary metabolic gene cluster of the FYZ1-3 strain

AntiSMASH was used to predict and analyze the secondary metabolite synthesis gene cluster of FYZ1-3. There were 10 secondary

TABLE 3 Carbohydrate and protein metabolic degradation pathways and related genes of the FYZ1-3 strain genome.

Item	Metabolic pathway No.	Function description	Gene number
Carbo-hydrate	ko00500	Starch and sucrose metabolism	47
	ko00010	Glycolysis/ gluconeogenesis	39
	ko00520	Amino sugar and nucleotide sugar metabolism	43
	ko04973	Carbohydrate digestion and absorption	1
Protein	ko00460	Cyanoamino acid metabolism	5
	ko01230	Biosynthesis of amino acids	122
	ko00250	Alanine, aspartate and glutamate metabolism	33
	ko00471	D-Glutamine and D-glutamate metabolism	6

TABLE 4 Information of functional genes related to nicotine metabolism based on KEGG of the FYZ1-3 strain genome.

Item	Gene_ ID	Ko_ name	Ko_ EC	Identity%	GIs_ id
Nicotine	GM000325	nadE	6.3.1.5	100	
	GM000414	gabD	1.2.1.16 1.2.1.79 1.2.1.20	99.6	
	GM000824	yfkN	3.1.4.16 3.1.3.6 3.1.3.5	98.8	
	GM001226	ppnK	2.7.1.23	100	GIs004
	GM001855	pncC	3.5.1.42	99.8	
	GM002193	deoD	2.4.2.1	100	
	GM002298	cca	2.7.7.72 3.1.3.– 3.1.4.	100	
	GM002414	punA	2.4.2.1	99.6	
	GM002666	nadD	2.7.7.18	99	
	GM002885	nadA	2.5.1.72	99.7	GIs011
	GM002886	nadC	2.4.2.19	99.7	
	GM002887	nadB	1.4.3.16	99.2	
	GM003111	ppnK	2.7.1.23	100	GIs012
	GM003348	pncB	6.3.4.21	99.8	

metabolites in total. The specific prediction analysis results are shown in Table 5. After comparing all gene clusters of strain FYZ1-3 with the known secondary metabolite gene clusters using BLAST, 8 functional

TABLE 5 Prediction and analysis of the secondary metabolite gene cluster of FYZ1-3.

Gene cluster number	Gene dosage	Gene cluster type	The most similar gene cluster	Similarity/%
Cluster 1	26	NRPS	Surfactin	82
Cluster 2	20	Terpene	—	—
Cluster 3	55	NRPS, PKS-like, T3PKS, transAT-PKS	Bacillaene	100
Cluster 4	37	NRPS, betalactone	Fengycin	100
Cluster 5	18	Terpene	—	—
Cluster 6	44	T3PKS	1-carbapen-2-em-3-carboxylic-acid	16
Cluster 7	44	NRP-metallophore, NRPS	Bacillibactin	100
Cluster 8	26	Lanthipeptide-class-i	Subtilin	100
Cluster 9	18	Sactipeptide	Subtilosin A	100
Cluster 10	39	Other	Bacilysin	100

synthetic gene clusters and 2 unknown synthetic gene clusters were found, which indicated the likelihood of new active substance synthesis gene clusters in the gene sequence of FYZ1-3. Lignosulfan is a heat-stable antimicrobial peptide on the cell membranes of Gram-positive bacteria, which can be used as a substitute for antibacterial drugs (Willey and Donk, 2007). Similarity between the Willey J., Donk wool sulfur antimicrobial peptide gene cluster 8 and lanthipeptides: Bateq7PJ16_RS18530 in the genome of *B. subtilis* FYZ1-3 reached 100%, and the similarity between gene cluster 9 and sactipeptides: BSU_37350 reached 100%. Thus, it may be used as a development strain for the synthesis of wool sulfur antibiotics (Figure 5). The strain contains 44 wool sulfur antimicrobial peptide gene clusters, which are related to the antibacterial properties of bacilli. The annotated sactipeptide is a class of polypeptides with an intramolecular thioether bond connected by a cysteine sulfhydryl group and an α-carbon atom of the receptor amino group. It is formed by free radical S-adenosylmethionine enzyme-dependent catalysis of the propeptide. This unique bond is completely different from the bond that connects sulfur-containing amino acid residues to β-carbons (Flühe and Marahiel, 2013).

Bacteriocin is a polypeptide with antibacterial activity synthesized by bacteria during the metabolic process. It is a naturally found microbial preservative that is safe and nontoxic with strong antibacterial activity and mild action conditions (Holo et al., 1991). According to the results from previous studies, the antibacterial substances produced by *Bacillus* are mainly lipopeptide antibiotics and polyketide antibiotics, which have antagonistic activity against several pathogenic bacteria. The application of these substances to soil can enhance plants' own resistance (Ongena and Jacques, 2007). Based on

the prediction of the secondary metabolite gene cluster, it was found that the genome of strain FYZ1-3 contained multiple genes related to bacteriocins, and it was speculated that it had certain antibacterial properties.

3.4.2 Analysis of antibacterial effect gene of the FYZ1-3 strain

Fengycin and bacilysin are the most common antibacterial substances in *Bacillus*, which can be synthesized by non-ribosomal pathway with a molecular weight ranging from 300 to 3,000 μ. Existing studies have demonstrated the effectiveness of lipopeptide antimicrobial compounds, including surfactin, iturin, and fengycin, which are produced by the lipopeptide synthase gene cluster, against various plant pathogenic bacteria, fungi, and oomycetes (Meena and Kanwar, 2015; Abdalah et al., 2018). Among them, fengycin synthase genes had higher antibacterial activity (Li et al., 2012; Chen et al., 2019). According to the KEGG database annotation, fengycin and bacilysin synthase regulatory genes could be detected in FYZ1-3 strain, including 5 genes of fengycin and 7 genes of bacilysin, as shown in Table 6. The molecular structure of fengycin contains a circular part composed of 10 amino acids and a long-chain fatty acid branch chain. The synthetic gene is composed of five genes: *fenC, fenD, fenE, fenA* and *fenB*, as shown in Figure 6. FYZ1-3 strain had complete Fengycin synthase regulatory genes, indicating a strong inhibitory effect on filamentous fungi. Based on the prediction of secondary metabolite gene clusters and KEGG antibacterial genes, it was found that the genome of strain FYZ1-3 contained multiple bacteriocin-related gene clusters, indicating that it could possess antibacterial qualities.

3.4.3 Verification of the fungal-inhibition ability of the FYZ1-3 strain

Four common pathogenic fungi in tobacco waste, *P. chrysogenum*, *A. sydowii*, *A. fumigatus*, and *T. funiculosus* were selected for plate confrontation experiments to determine the inhibitory ability of FYZ1-3 to fungi. The growth of the strain after 72 h of culture at 28°C is shown in Figure 7A. No filamentous fungal growth was observed around strain FYZ1-3, and the colony surface in the *P. chrysogenum* group could not form lemon-yellow droplets by normal exudate accumulation. The spores of *A. sydowii* in the experimental group could not be dispersed in the whole culture dish similar to that in the control group. Although multiple colonies were formed in the experimental group of *T. funiculosus*, no sterile colonies were growing around strain FYZ1-3. The *A. fumigatus* colony in the experimental group was small, and exhibited obviously shrunken mycelia. In summary, strain FYZ1-3 had an obvious growth-inhibition effect on the four fungi.

Since there were two microorganisms grown on a culture medium in the antagonistic experiment, there was a competition for limited space and nutrients. The observed inhibitory effect might be due to the low growth competition rate of the fungi rather than the secretion of growth inhibitors by the strain FYZ1-3, as the functional bacteria had a growth rate significantly exceeding the fungi. To eliminate this confounding factor, the agar column method was employed, which inoculated a single colony of filamentous fungi onto PDA culture medium containing sterile fermentation supernatant from functional bacteria FYZ1-3. The growth state of fungi was then observed in the absence of other microorganisms. In this way, the growth of fungi was

FIGURE 5
Linear map of the wool thiopeptide antibiotic synthesis gene cluster of strain FYZ1-3: **(A)** cluster 8 lanthipeptide-class-i; **(B)** cluster 9 sactipeptide.

dependent solely on the metabolic products produced by the added functional microorganisms.

The growth of fungal strains was recorded after 48 h of culture at 28°C (Figure 7B). The mycelium of *P. chrysogenum* and *A. sydowii* in the control group had spread throughout the culture dish, whereas fungal growth in the culture dish containing the fermentation supernatant of FYZ1-3 strain was strongly inhibited by up to 75.8%. Observation of the colonies of *T. funiculosus* and *A. fumigatus* revealed only a small amount of diffusion at the edge of the inoculated fungal colony in the experimental group, and a small amount of mycelium was found growing in other parts of the culture dish. The color and morphology were significantly different from those of the normally growing mycelium in the control group. There was some *Penicillium* growth in the Petri dish to which the sterile fermentation supernatant of the FYZ1-3 strain was added. The inhibition rates of the four groups were calculated to be 75.8, 63.4, 48.6, and 44.1%, respectively. Observation of the colony morphology revealed that the experimental group had larger fungal colonies with more loose hyphae and unclear boundaries. The control group had smaller colonies, dense hyphae, and clear boundaries.
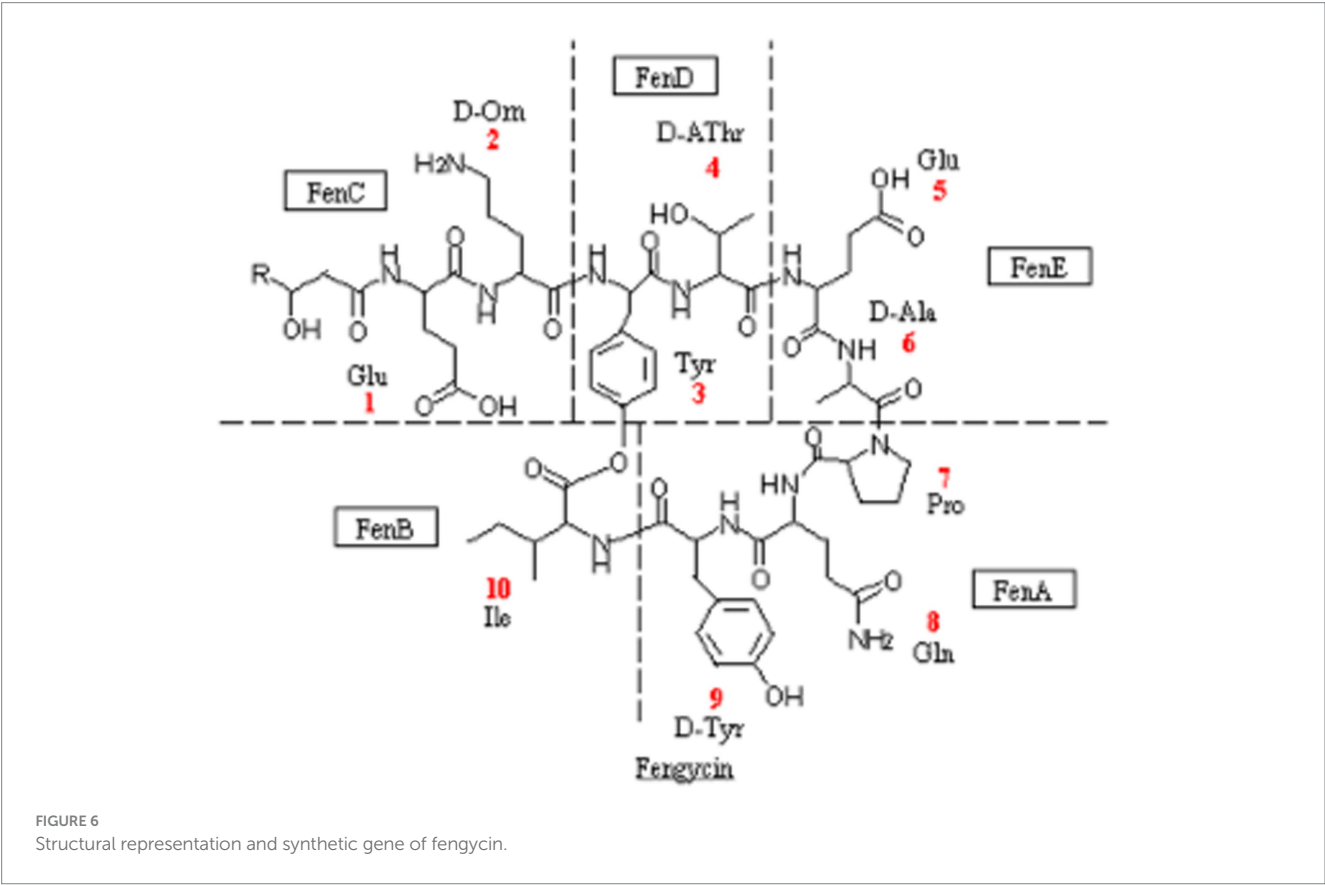
Based on these experiments, it could be inferred that the fermentation supernatant of the FYZ1-3 strain contains substances that can inhibit the growth and development of fungi, inhibit the germination of spores and bacterial growth, and change the normal morphology of mycelium. Therefore, strain FYZ1-3 is the target strain that can inhibit fungal growth.

4 Discussion

Currently, *Bacillus* spp., as the most commonly selected representative, has become the core genus of effective microorganisms (EM) and other commercial bacterial agents. *Bacillus* spp. can produce several extracellular enzymes such as amylase, protease, and cellulase, while its thicker cell wall and spores can help the bacterium survive in harsh high-temperature and high-salt environments (Sonenshein et al., 1993). There are many studies on the degradation mechanism and applications of *B. subtilis*; however, owing to the unique structure and composition of organic substances in tobacco sources, as well as the presence of the biological toxin nicotine in tobacco waste, bacterial agents that have been identified and demonstrated to show potential still cannot be directly used for the aerobic degradation of tobacco waste. Therefore, the search for functional strains that are resistant to nicotine and have a high degradation ability for tobacco source organic matter has important scientific significance and practical value in improving the resource-utilization rate of tobacco waste.

Understanding the basic genomic information of functional strains allows researchers to further understand the mechanism of metabolic transformation of functional strains and facilitates understanding the relationship between genes and proteins, metabolic functions, and individual behaviors. Therefore, genomics plays a crucial role in understanding the physiological performance and ecological significance of functional bacteria (Seminago et al., 2022). In this study, the genome of *B. subtilis* FYZ1-3, a strain isolated from

TABLE 6 Information of bacteriostasis genes based on KEGG of FYZ1-3.

Type	Gene _ID	Identity%	Ko_id	Ko_name	Function
Fengycin	GM002040	99.1	K15668	ppsE, fenB	Fengycin family lipopeptide synthetase E
	GM002041	97.7	K15667	ppsD, fenA	Fengycin family lipopeptide synthetase D
	GM002042	98.5	K15666	ppsC, fenE	Fengycin family lipopeptide synthetase C
	GM002043	98.9	K15665	ppsB, fenD	Fengycin family lipopeptide synthetase B
	GM002044	98.5	K15664	ppsA, fenC	Fengycin family lipopeptide synthetase A
Bacilysin	GM003982	99.6	K19550	bacG	Bacilysin biosynthesis oxidoreductase BacG
	GM003983	99.7	K19549	bacF	Bacilysin biosynthesis transaminase BacF
	GM003984	100	K19552	bacE	MFS transporter, DHA3 family, bacilysin exporter BacE
	GM003985	99.8	K13037	bacD	L-alanine-L-anticapsin ligase
	GM003986	98.8	K19548	bacC	Dihydroanticapsin dehydrogenase
	GM003987	99.6	K19547	bacB	3-[(4R)-4-hydroxycyclohexa-1,5-dien-1-yl]-2-oxopropanoate isomerase
	GM003988	100	K19546	bacA	Prephenate decarboxylase

FIGURE 6
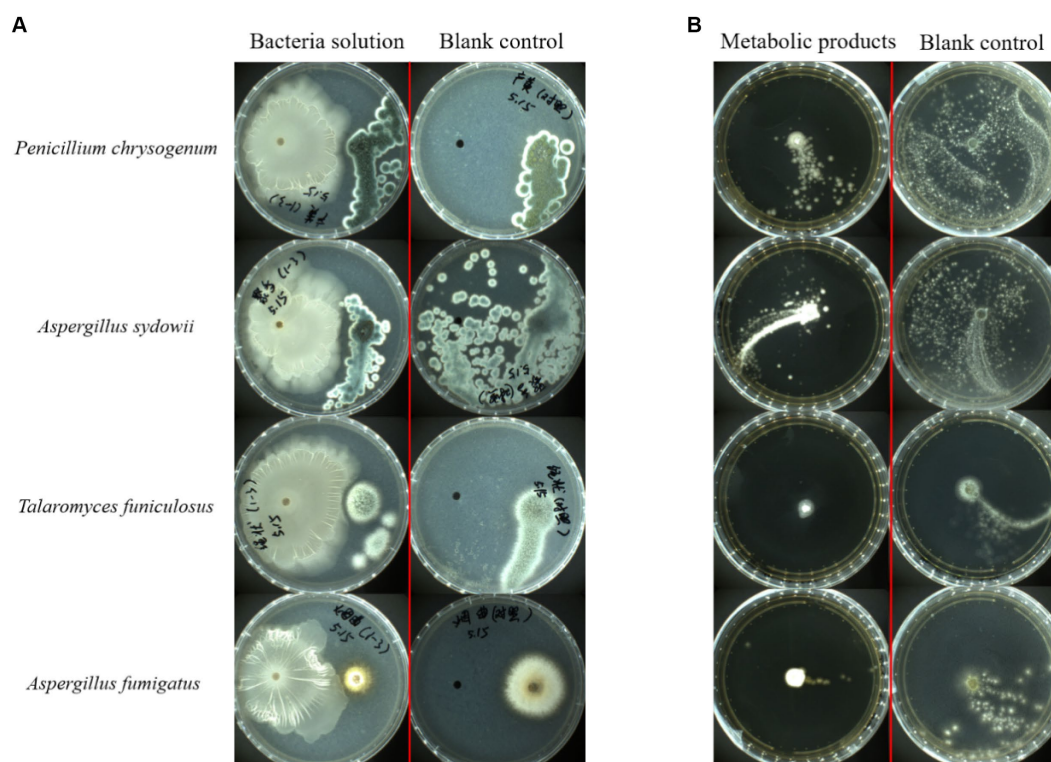Structural representation and synthetic gene of fengycin.

FIGURE 7
Bacterial inhibition by strain FYZ1-3: **(A)** growth results using FYZ1-3 bacterial solution in the plate confrontation method for 72 h; **(B)** 48 h growth results of FYZ1-3 in the metabolite agar column method.

tobacco waste compost, was analyzed. Based on phylogenetic analysis and metabolic pathway analysis of its genome, and comparative analysis with the genome of the standard strain DSM 10, the evolutionary development status and metabolic potential of the functional strain FYZ1-3 in *B. subtilis* were revealed at the molecular level. Moreover, the mobile genome islands carrying a large number of functional genes were identified, providing an important basis for gene function analysis and environmental niche research.

Due to the wide antibacterial spectrum and strong adaptability, *B. subtilis* has been increasingly used in the treatment of diseases in humans, animals, and plants, as well as biosurfactants, food preservatives, feed additives, and molecular biology research tools. Its antibacterial mechanism was mainly manifested in competition, antagonism, and induction of biological resistance. This study aimed to improve the efficiency of aerobic fermentation degradation of tobacco waste and produce high-quality organic fertilizers. A strain of *B. subtilis* FYZ1-3 was screened, and its biological control ability against common pathogenic fungi was speculated through genome analysis and laboratory experimentation. Therefore, this functional strain FYZ1-3 presented itself as a crucial contender for microbial organic fertilizers, providing a harmonious blend of organic fertilizers and advantageous microorganisms. Organic fertilizers could create an optimal environment for microorganisms, extending their survival and enhancing their ability to control soil pathogens. The implementation of microbial organic fertilizers could increase the treatment of tobacco waste, minimize secondary pollution, and promote green circular development. Moreover, it could curtail the

incidence of soil-borne diseases by suppressing the activity of pathogenic microorganisms in the soil due to the antibacterial impact of FYZ1-3.

5 Conclusion

This study isolated a functional strain, FYZ1-3, from aerobic fermentation samples of tobacco waste, which could tolerate high temperatures up to 80°C and nicotine concentrations as high as 0.6%. FYZ1-3 exhibited high degradation ability towards tobacco starch and tobacco protein, with amylase and protease activities of 122.3 U/mL and 52.3 U/mL, respectively. Through morphology and molecular biology analysis, FYZ1-3 was identified as *B. subtilis*. Using whole genome sequencing, the reasons for high carbohydrate and protein degradation rates and nicotine tolerance were analyzed at the molecular level, and the ability of FYZ1-3 strain to inhibit fungal growth was identified and validated. This study marks the initial step in exploring and studying the functional strain FYZ1-3 for tobacco waste treatment. In the future, we will continue to explore and verify the functional genes through molecular biology, and conduct material and genetic analysis of metabolites. To summarize, FYZ1-3 exhibits promising potential as a microbial preparation candidate for utilizing tobacco waste resources, based on our research findings. This study enhances microbial resources and establishes a theoretical basis for enhancing aerobic fermentation efficiency and product utilization for tobacco waste.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

CY: Conceptualization, Methodology, Formal analysis, Funding acquisition. DaL: Investigation, Validation, Methodology, Formal analysis, Data curation. KH: Methodology, Visualization. DoL: Resources, Methodology, Project administration. XM: Data curation, Visualization. YJ: Investigation, Methodology. HX: Methodology.

Funding

Conflict of interest

DaL, CY, KH, and DoL were employed by Zhengzhou Tobacco Research Institute of CNTC.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdalah, D., Tounsi, S., Gharallah, H., Hammami, A., and Gargouri, O. (2018). Lipopeptides from *Bacillus amyloliquefaciens* strain 32a as promising biocontrol compounds against the plant pathogen *Agrobacterium tumefaciens*. *Environ. Sci. Pollut. Res. Int.* 25, 36518–36529. doi: 10.1007/s11356-018-3570-1

Aoki-Kinoshita, K., and Kanehisa, M. (2007). Gene annotation and pathway mapping in KEGG. *Methods Mol. Biol.* 396, 71–91. doi: 10.1007/978-1-59745-515-2_6

Arahal, D. (2014). Whole-genome analyses: average nucleotide identity. *Methods Microbiol.* 41, 103–122. doi: 10.1016/bs.mim.2014.07.002

Ashengroph, M. (2015). Isolation and screening of a native *Citrobacter* sp. with high nicotine-tolerant and its application as a biocatalyst for biodegradation of nicotine. *Biol. J. Microorg.* 4, 69–82.

Awasthi, M., Zhang, Z., Wang, Q., Shen, F., Li, R., Li, D., et al. (2017). New insight with the effects of biochar amendment on bacterial diversity as indicators of biomarkers support the thermophilic phase during sewage sludge composting. *Bioresour. Technol.* 238, 589–601. doi: 10.1016/j.biortech.2017.04.100

Chen, M., Wang, J., Zhu, Y., Liu, B., Yang, W., and Rang, C. (2019). Antibacterial activity against *Ralstonia solanacearum* of the lipopeptides secreted from the *Bacillus amyloliquefaciens* strain FJAT-2349. *J. Appl. Microbiol.* 126, 1519–1529. doi: 10.1111/jam.14213

Civilini, M., Domenis, C., Sebastinanutto, N., and Bertoldi, M. (1997). Nicotine decontamination of tobacco agro-industrial waste and its degradation by micro-organisms. *Waste Manag. Res.* 15:4, 349–358. doi: 10.1177/0734242X9701500403

Dhillon, B., Laird, M., Shay, J., Winsor, G., Lo, R., Nizam, F., et al. (2015). IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Res.* 43, 104–108. doi: 10.1093/nar/gkv401

Dobaradaran, S., Schmidt, T., Lorenzo-Parodi, N., Kaziur-Cegla, W., Jochmann, M., Nabipour, I., et al. (2020). Polycyclic aromatic hydrocarbons (PAHs) leachates from cigarette butts into water. *Environ. Pollut.* 259:113916. doi: 10.1016/j.envpol.2020.113916

Fang, Y., Jia, X., Chen, L., Lin, C., Zhang, H., and Chen, J. (2019). Effect of thermotolerant bacterial inoculation on the microbial community during sludge composting. *Can. J. Microbiol.* 65, 750–761. doi: 10.1139/cjm-2019-0107

Fitzpatrick, P. (2018). The enzymes of microbial nicotine metabolism. *Beilstein J. Org. Chem.* 14, 2295–2307. doi: 10.3762/bjoc.14.204

Flühe, L., and Marahiel, M. (2013). Biosynthesis of Sactipeptides: characterization of the thioether bond forming radical SAM enzymes AlbA und SkfB. *Curr. Opin. Chem. Biol.* 17, 605–612. doi: 10.1016/j.cbpa.2013.06.031

GB/T 23527-2009. (2009) *Proteinase Preparations of China*. Standardization Administration of the People's Republic of China.

GB1886.174-2016. (2016) *National Food Safety Standards Food Additives Enzyme Preparations for the Food Industry of China*. National Health Commission of the People's Republic of China.

Holo, H., Nilssen, O., and Nes, I. (1991). Lactococcin a, a new bacteriocin from *Lactococcus lactis* subsp. cremoris: isolation and characterization of the protein and its gene. *J. Bacteriol.* 173, 3879–3887. doi: 10.1128/jb.173.12.3879-3887.1991

Kayikçioglu, H., and Okur, N. (2011). Evolution of enzyme activities during composting of tobacco waste. *Waste Manag. Res.* 29, 1124–1133. doi: 10.1177/0734242X10392813

Koutela, N., Fernández, E., Saru, M., and Psillakis, E. (2020). A comprehensive study on the leaching of metals from heated tobacco sticks and cigarettes in water and natural waters. *Sci. Total Environ.* 714:136700. doi: 10.1016/j.scitotenv.2020.136700

Lee, T., Jung, S., Lin, K., Tsang, Y., and Kwon, E. (2021). Mitigation of harmful chemical formation from pyrolysis of tobacco waste using CO_2. *J. Hazard. Mater.* 2020, 401–123416. doi: 10.1016/j.jhazmat.2020.123416

Li, L., Ma, M., Huang, R., Qu, Q., Li, G., Zhou, J., et al. (2012). Induction of chlamydospore formation in fusariumby cyclic lipopeptide antibiotics from *Bacillus subtilis* C2. *J. Chem. Ecol.* 38, 966–974. doi: 10.1007/s10886-012-0171-1

Lilge, L., Hertel, R., Heravi, K., Henkel, M., Commichau, F., and Hausmanna, R. (2021). Draft genome sequence of the type strain *Bacillus subtilis* subsp. subtilis DSM10. *Microbiolol. Resour. Announc.* 10:e00158. doi: 10.1128/MRA.00158-21

Liu, D., Ma, X., Huang, J., Shu, Z., Chu, X., Li, Y., et al. (2022a). Enhancing the fermentation of acidified food waste using a self-formulated thermophilic and acid-resistant bacterial agent. *J. Environ. Chem. Eng.* 10:107350. doi: 10.1016/j.jece.2022.107350

Liu, D., Ma, X., Huang, J., Shu, Z., Chu, X., Li, Y., et al. (2022b). Study on personalized microbial formulation during high-temperature aerobic fermentation of different types of food wastes. *Sci. Total Environ.* 814:152561. doi: 10.1016/j.scitotenv.2021.152561

Magain, N., Miadlikowska, J., Mueller, O., Gajdeczka, M., Truong, C., Salamov, A., et al. (2017). Conserved genomic collinearity as a source of broadly applicable, fast evolving, markers to resolve species complexes: a case study using the lichen-forming genus peltigera section polydactylon. *Mol. Phylogenet. Evol.* 117, 10–29. doi: 10.1016/j.ympev.2017.08.013

Masson, F., and Rossignol, M. (1995). Basic plasticity of protein expression in tobacco leaf plasma membrane. *Plant J.* 8, 77–85. doi: 10.1046/j.1365-313X.1995.08010077.x

Mclaughlin, I., Harting, J., and Van Wageningen, S. (2021). High-resolution HLA typing with NGSgo-MX6-1by HiFi NGS data of a PacBio sequel II platform. *GenDX* 2021:97.

Meena, K., and Kanwar, S. (2015). Lipopeptides as the antifungal and antibacterial agents: applications in food safety and therapeutics. *Biomed. Res. Int.* 2015:473050. doi: 10.1155/2015/473050

Nayak, S. (2020). Multifaceted applications of probiotic Bacillus species in aquaculture with special reference to *Bacillus subtilis*. *Rev. Aquac.* 13, 862–906. doi: 10.1111/raq.12503

Nölke, G., Chudobova, I., Houdelet, M., Volke, D., Lusso, M., Frederick, J., et al. (2021). Impact of nicotine pathway downregulation on polyamine biosynthesis and leaf ripening in tobacco. *Plant Direct* 5:e00329. doi: 10.1002/pld3.329

Ongena, M., and Jacques, P. (2007). Bacillus lipopeptides: versatile weapons for plant disease biocontrol. *Trends Microbiol.* 16, 115–125. doi: 10.1016/j.tim.2007.12.009

Pan, H., Yang, C., Wei, Z., and Jiang, J. (2006). DNA extraction of birch leaves by improved CTAB method and optimization of its ISSR system. *J. For. Res.* 4, 298–300. doi: 10.1007/s11676-006-0068-3

Seminago, R., Castano, E., and Amador, A., (2022). *Genomics: more than genes*. Dandbook of instrumental techniques from CCiTUB.

Shui, P., Zheng, X., Lin, J., and Guo, L. (2014). A rapid and efficient method for isolating high quality total RNA from edible fungi. *Acta Edulis Fugi* 36:2008. doi: 10.16488/j.cnki.1005-9873.2008.01.008

Sonenshein, A. L., Hoch, J. A., and Losick, R., (1993). *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics. *Amer. Soc. Microbiol.* doi: 10.1128/9781555818388

Song, C., Sun, F., Xu, Z., Xu, X., Wng, S., and Gong, C. (2010). Research advance in fine structure of tobacco starch during curing. *Chin. Tob. Sci.* 1:78.

Wang, T., He, F., Tian, B., Song, C., Shi, L., Xu, C., et al. (2012). Changes of granule structure and enzyme hydrolysability of upper flue-cured tobacco leaves starch during bulk flue-curing. *Sci. Agric. Sin.* 13, 2704–2710. doi: 10.3864/j.issn.0578-1752.2012.13.014

Wang, J., Jiang, B., Liu, L., Cao, L., Yuan, Q., Tian, J., et al. (2022). Tobacco waste biomass for electrochemical energy storage application. *J. Phys. Conf. Ser.*, 2022 2160:012052. doi: 10.1088/1742-6596/2160/1/012052

Wang, M., Wu, Y., Wang, X., Wu, Y., Wang, X., Zhao, J., et al. (2022). Effects of thermophiles inoculation on the efficiency and maturity of rice straw composting. *Bioresour. Technol.* 354:127195. doi: 10.1016/j.biortech.2022.127195

Weishaar, H., Collin, J., and Amos, A. (2016). Tobacco control and health advocacy in the European Union: understanding effective coalition-building. *Nicotine Tob. Res.* 18, 122–129. doi: 10.1093/ntr/ntv016

Willey, J., and Donk, W. (2007). Lantibiotics: peptides of diverse structure and function. *Annu. Rev. Microbiol.* 61, 477–501. doi: 10.1146/annurev.micro.61.080706.093501

Yildiz, D. (2004). Nicotine, its metabolism and an overview of its biological effects. *Toxicon* 43, 619–632. doi: 10.1016/j.toxicon.2004.01.017

Zhang, Y., Li, R., Shang, G., Zhu, H., Wang, H., Pandiselvam, R., et al. (2022). Effects of multiscale-mechanical fragmentation on techno-functional properties of industrial tobacco waste. *Powder Technol.* 2022:402. doi: 10.1016/j.powtec.2022.117327

Zhang, G., Zhan, J., and Fu, H. (2022). Trends in smoking prevalence and intensity between 2010 and 2018: implications for tobacco control in China. *Int. J. Environ. Res. Public Health* 19:670. doi: 10.3390/ijerph19020670

Zhou, Q. (2008). Research progress in the nicotine content in tobacco leaf. *J. Anhui Agric. Sci.* 36, 2359–2361. doi: 10.1016/S1872-2075(08)60071-0

*CORRESPONDENCE
Long Jin
✉ isacckim@alumni.kaist.ac.kr
Hyung-Gwan Lee
✉ trustin@kribb.re.kr

†These authors have contributed equally to this
work

Genetic diversity into a novel free-living species of *Bradyrhizobium* from contaminated freshwater sediment

Naxue Zhang[1†], Chun-Zhi Jin[2†], Ye Zhuo[2,3], Taihua Li[1], Feng-Jie Jin[1], Hyung-Gwan Lee[2]* and Long Jin[1]*

[1]College of Ecology and Environment, Nanjing Forestry University, Nanjing, China, [2]Cell Factory Research Centre, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea, [3]Department of Environmental Biotechnology, KRIBB School of Biotechnology, Korea University of Science and Technology (UST), Daejeon, Republic of Korea

A free-living *Bradyrhizobium* strain isolated from a contaminated sediment sample collected at a water depth of 4 m from the Hongze Lake in China was characterized. Phylogenetic investigation of the 16S rRNA gene, concatenated housekeeping gene sequences, and phylogenomic analysis placed this strain in a lineage distinct from all previously described *Bradyrhizobium* species. The sequence similarities of the concatenated housekeeping genes support its distinctiveness with the type strains of the named species. The complete genome of strain S12-14-2 consists of a single chromosome of size 7.3M. The strain lacks both a symbiosis island and important nodulation genes. Based on the data presented here, the strain represents a new species, for which the name *Bradyrhizobium roseus* sp. nov. is proposed for the type strain S12-14-2^T. Several functional differences between the isolate and other published genomes indicate that the genus *Bradyrhizobium* is extremely heterogeneous and has functions within the community, such as non-symbiotic nitrogen fixation. Functional denitrification and nitrogen fixation genes were identified on the genomes of strain S12-14-2^T. Genes encoding proteins for sulfur oxidation, sulfonate transport, phosphonate degradation, and phosphonate production were also identified. Lastly, the *B. roseus* genome contained genes encoding ribulose 1,5-bisphosphate carboxylase/oxygenase, a trait that presumably enables autotrophic flexibility under varying environmental conditions. This study provides insights into the dynamics of a genome that could enhance our understanding of the metabolism and evolutionary characteristics of the genus *Bradyrhizobium* and a new genetic framework for future research.

Introduction

The genus *Bradyrhizobium* is a relatively well-studied rhizobial bacteria and well known for symbiosis and endophytic interaction with leguminous plants, where they exchange nutrients for survival. The genus includes a diverse collection of globally distributed bacteria that can nodulate a wide range of legumes (Salminen and Streeter, 1987; Van Rhijn and Vanderleyden, 1995; Okubo et al., 2012; Parker, 2015; Remigi et al., 2016; Sprent et al., 2017; Claassens et al., 2023; Lafay et al., 2023). It was widely recognized that the *Bradyrhizobium* species were capable of nitrogen fixation and the formation of symbiotic nodules in the roots of leguminous plants. The symbiotic species are crucial for soil health, plant growth, and food security, whereas nitrogen-fixing species are critical for the global nitrogen cycle (Jordan, 1982; Ladha and So, 1994; Stacey et al., 1995; Remigi et al., 2016; Gogoi et al., 2018). Members of *Bradyrhizobium* are characterized extensively for their symbiotic lifestyle with leguminous plants to form symbiotic nodules. *Bradyrhizobium* has biologically significant functions in soils, including photosynthesis, nitrogen fixation, denitrification, and degradation of aromatic compounds and others. Especially in the global nitrogen cycle, *Bradyrhizobium* removes nitrogen through heterotrophic denitrification and also has multiple roles in agriculture (Kaneko et al., 2002; Okubo et al., 2012). Nodulating *Bradyrhizobium* species typically contain *nod* and *nif* genes associated with nodulation and nitrogen fixation, respectively, the so-called symbiotic properties. Due to these characteristics, *Bradyrhizobium* has been regarded as a model of legume-rhizobia symbiosis and an ecologically significant microbe. However, nodulation-deficient isolates have been discovered and characterized, even though they possess genes for nitrogen fixation, and some *Bradyrhizobium* isolates recovered from forest soils have been found to lack both nodulation and nitrogen-fixing functions (VanInsberghe et al., 2015; Jones et al., 2016). A culture-independent investigation has shown that the population of *Bradyrhizobium* in soil habitats is different from the rhizosphere of leguminous plants (VanInsberghe et al., 2015), suggesting that non-symbiotic *Bradyrhizobia* inhabit physically or functionally distinct niches.

Understanding the mechanisms of bradyrhizobial adaptation to independent existence in diverse environments is becoming increasingly important and urgent and may reveal the genetic potential of this globally significant genus. In the present study, we characterized a novel species of *Bradyrhizobium* isolated from freshwater sediment of the Hongze Lake in China. To reveal the genetic diversity and physiological features, here, we report the isolation, and the sequencing, assembly, taxonomic classification, and annotation of the genome of a new species of the genus *Bradyrhizobium*. These investigations will be extremely beneficial for elucidating its putative metabolism, contributing to a greater understanding of free-living *Bradyrhizobium* isolates, enhancing knowledge of the genes and enzymes involved in metabolic pathways, and possibly identifying genomic markers that can be utilized in future ecological studies.

Materials and methods

Sampling, isolation, and growth condition

Strain S12-14-2 was recovered from a contaminated sediment sample (33°14′26″N, 118°35′40″E) collected at a water depth of 4 m from Hongze Lake, the fourth largest freshwater lake in China. The Hongze Lake is located in the Jiangsu Province and surrounded by Suqian and Huai'an cities; it has a surface area of 1,960 km^2, a mean water depth of 1.77 m, a maximum depth of 4.37 m, and a volume of 28×108 m^3 (Gao et al., 2009; Li et al., 2011). The ecological risk of the Hongze Lake sediments has been assessed. Heavy metals including copper (Cu), zinc (Zn), lead (Pb), cadmium (Cd), chromium (Cr), mercury (Hg), arsenic (As), iron (Fe), aluminum (Al), and manganese (Mn) have been identified in the lake sediments of the Hongze Lake, and antibiotics like atrazine and ofloxacin were also detected (Yao et al., 2016; Wu et al., 2023). One gram of a sediment sample was used to screen for bacterial strains with the serial dilution method. The strain was grown on modified R2A medium at 30°C under dark conditions (Jin et al., 2019).

Morphological and physiological characteristics

The mobility and morphology of the colonies were examined using a phase-contrast microscope (Nikon Optiphot, 1,000× magnification) and transmission electron microscopy (Philips CM-20). The growth temperature range, pH range, and salt tolerance were determined following the methods described by Jin et al. (2021). Duplicated antibiotic-susceptibility of strain S12-14-2 was checked on R2A agar medium using the filter-paper disk method (Li et al., 2020), where the disks contained the following: nalidixic acid (30 μg ml^{-1}), tetracycline (30 μg ml^{-1}), amikacin (30 μg ml^{-1}), ampicillin/sulbactam (20 μg ml^{-1}, 1:1), kanamycin (30 μg ml^{-1}), vancomycin (30 μg ml^{-1}), chloramphenicol (30 μg ml^{-1}), teicoplanin (30 μg ml^{-1}), spectinomycin (25 μg ml^{-1}), gentamicin (30 μg ml^{-1}), streptomycin (25 μg ml^{-1}), rifampicin (30 μg ml^{-1}), lincomycin (15 μg ml^{-1}), and erythromycin (30 μg ml^{-1}). Positive results were observed for halo diameters greater than 10 mm after 5 days of incubation at 30°C. For the comparative cellular fatty acids analysis, strain S12-14-2 together with the reference strains *Bradyrhizobium erythrophlei* LMG 28425^T, *Bradyrhizobium jicamae* LMG 24556^T, *Bradyrhizobium lablabi* LMG 25572^T, *Bradyrhizobium mercantei* LMG 30031^T, *Bradyrhizobium elkanii* KACC 10647^T, and *Bradyrhizobium japonicum* KACC 10645^T were cultured on R2A agar for 2 days at 30°C. Standardized cell harvesting and extraction were performed according to Jin et al. (2022a), and the identification was performed using the TSBA 6 (Trypticase Soy Broth Agar) database with the Sherlock software 6.1. Carbon source utilization, enzyme activities and other physiological and biochemical activities were observed using the API 20NE, API ID 32GN, and API ZYM kits (bioMérieux, l'Etoile, France) according to the manufacturer's instructions.

DNA extraction, PCR amplification, and genomic sequencing

Genomic DNA was extracted using the FastDNA^{TM} SPIN kit following the manufacturer's instructions. The concentration and purity of the DNA were then checked on a ND-2000 spectrophotometer (Thermo Fisher Scientific). The 16S rRNA genes of strain S12-14-2 were amplified by PCR method using the universal primer sets 27F/1492R (Weisburg et al., 1991) and then identified using the EzBioCloud web service (Yoon et al., 2017). The use of multilocus sequence analysis (MLSA) has been proven useful for the classification and identification of several groups of rhizobia. In the genus *Bradyrhizobium*, the 16S rDNA sequences provide minimal taxonomic information due to the high degree of conservation among species and the relatively high sequence similarity. Consequently, MLSA studies of *glnII*, *recA*, *dnaK*, *rpoB*, and ITS genes were performed to assess the taxonomic position of the free-living strain among identified species (Martens et al., 2008; Menna et al., 2009; Rivas et al., 2009; Pérez-Yépez et al., 2014; Gnat et al., 2016). The partial sequences of the housekeeping genes, *glnII*, *recA*, *dnaK*, *rpoB*, and ITS genes, were amplified by PCR method. The primer sets and PCR amplification conditions were described in a previous study (Menna et al., 2009; Delamuta et al., 2013). Whole genome sequencing was done at Novogene Biotechnology, Beijing, China, using the Pacific Biosciences (PacBio) RS II single molecule real-time (SMRT) platform together with the Illumina HiSeq technology. The quality of the genome assemblies was assessed using SMRT Link version 5.0.1.

Phylogenetic analyses and genome annotation

For phylogenetic analysis of strain S12-14-2, the 16S rRNA sequences of *Bradyrhizobium* type strains were downloaded from the NCBI database. The sequences were edited and aligned with BIOEDIT and CLUSTAL X software, respectively (Thompson et al., 1997; Hall, 1999), and the phylogenetic tree was reconstructed using the MEGA7 software (Kumar et al., 2016). For genomic annotations, the draft genome sequence was applied to the RAST pipeline, and comparisons of the genome were performed in the SEED Viewer (Aziz et al., 2008, 2012). The protein coding sequences (CDSs) were submitted to the COG (Clusters of Orthologous Groups) database[1] for functional classification and summary statistics (Tatusov et al., 1997, 2003). For the pangenome analysis, we followed Delmont and Eren's description of the Anvi'o workflow (Delmont and Eren, 2018), a community-driven, open-source analysis and visualization platform for microbial-omics, which is available at http://merenlab.org/software/anvio/. The structures of protein monomeric units were predicted using the SWISS-MODEL (Bienert et al., 2017; Waterhouse et al., 2018). Biosynthetic gene clusters (BGC) were identified and analyzed using the secondary metabolite databases antiSMASH database.[2] The average nucleotide identity (ANI) was determined using the OrthoANI tool in the EZBioCloud server, and

the digital DNA–DNA hybridization (dDDH) value was calculated with the Genome-to-Genome Distance Calculator (GGDC 2.1) (Meier-Kolthoff et al., 2013; Yoon et al., 2017). The average amino acid identity (AAI) was calculated with the AAI web-based calculator developed by Konstantinidis Lab.[3]

Results and discussion

Physiological features

The strain displayed growth on R2A and NA agar media after 5 days of incubation at 28°C but not on TSA, LB, or MA media. Colonies were pink in color and less than 1 mm in diameter after 5 days of growth at 28°C on R2A agar medium. Microscopically, the strain represented a Gram-negative and rod-shaped bacterium that ranged in size from 2.0-2.3 × 0.7-0.8 µm (Supplementary Figure 1). Strain S12-14-2^{T} showed resistance to the following antibiotics tested: nalidixic acid (30 µg ml^{-1}), vancomycin (30 µg ml^{-1}), chloramphenicol (30 µg ml^{-1}), teicoplanin (30 µg ml^{-1}), streptomycin (25 µg ml^{-1}), rifampicin (30 µg ml^{-1}), and lincomycin (15 µg ml^{-1}); and susceptible to tetracycline (30 µg ml^{-1}), amikacin (30 µg ml^{-1}), ampicillin/sulbactam (20 µg ml^{-1}, 1:1), kanamycin (30 µg ml^{-1}), spectinomycin (25 µg ml^{-1}), gentamicin (30 µg ml^{-1}), and erythromycin (30 µg ml^{-1}). The major fatty acids (>10%) of the strain were sorted into two groups C$_{16:0}$ and C$_{18:1}$ ω7c and/or C$_{18:1}$ ω6c, which was consistent with species from the genus *Bradyrhizobium* (Supplementary Table 1). The carbon source assimilation, enzyme activity, and some other physiological features are summarized in Supplementary Table 2.

Genomic analysis: the taxonomic status

The phylogenetic analysis based on 16S rRNA gene sequences revealed that strain S12-14-2 should taxonomically be classified to the genus *Bradyrhizobium*. The strain shared 99.36–99.69% pairwise similarity with *B. lablabi* CCBAU 23086^{T}, *B. jicamae* PAC68^{T}, *B. erythrophlei* CCBAU 53325^{T}, and *B. algeriense* RST89^{T} and less than 99.3% with the other species within the genus *Bradyrhizobium* (Table 1). The 16S rRNA gene sequence neighbor-joining tree revealed that strain S12-14-2 unambiguously clustered with members of *Bradyrhizobium*. Due to its high level of conservation within the genus *Bradyrhizobium*, the 16S rRNA gene is an inadequate molecular marker for distinguishing species (Germano et al., 2006; Rivas et al., 2009). Therefore, a genome-based phylogenetic analysis was used. The topology of the phylogenetic and phylogenomic trees is consistent with the sequence similarities of the 16S rRNA gene between the novel strain and the type strains of the *Bradyrhizobium* species (Figure 1 and Supplementary Figure 2).

Within the genus *Bradyrhizobium*, phylogenetic analysis based on MLSA (multilocus sequence analysis) of housekeeping genes

---

[1] http://www.ncbi.nlm.nih.gov/COG/

[2] https://antismash.secondarymetabolites.org

[3] http://enve-omics.ce.gatech.edu/aai/

TABLE 1   16S rRNA gene, ANI, AAI similarities (%), and digital DNA–DNA relatedness (%) between novel strain S12-14-2^T and related type strains of *Bradyrhizobium*.

| No.	Type strains	16S rDNA	ANI	AAI	dDDH
1	*B. sediminis* S2-20-1^T (CP076134)	99.3	81.9	80.0	24.9
2	*B. algeriense* RST89^T (PYCM00000000)	99.7	86.6	86.5	32.5
3	*B. daqingense* CGMCC 1.10947^T (VLKL00000000)	97.7	79.2	75.0	22.8
4	*B. elkanii* USDA 76^T (ARAG00000000)	99.0	81.2	76.6	24.7
5	*B. embrapense* SEMIA 6208^T (LFIP00000000)	99.3	81.0	76.8	24.1
6	*B. erythrophlei* GAS138 (LT670817)	99.6	79.9	74.6	23.7
7	*B. icense* LMTR 13^T (CP016428)	99.3	85.1	85.1	29.5
8	*B. japonicum* USDA 6^T (AP012206)	98.2	79.7	74.6	23.0
9	*B. jicamae* PAC68^T (LLXZ00000000)	99.4	85.5	85.4	30.0
10	*B. lablabi* CCBAU 23086^T (LLYB01000065)	99.4	85.5	71.6	29.8
11	*B. mercantei* SEMIA 6399^T (MKFI00000000)	99.3	81.1	76.5	24.5
12	*B. neotropicale* BR 10247^T (LSEF00000000)	98.9	79.3	74.4	22.9
13	*B. pachyrhizi* PAC 48^T (LFIQ01000000)	99.0	81.1	76.5	24.5
14	*B. retamae* Ro19^T (LLYA00000000)	99.0	85.0	84.8	29.2
15	*B. tropiciagri* SEMIA 6148^T (LFLZ00000000)	99.0	80.9	75.4	24.5
16	*B. uaiense* UFLA03-164^T (VKHP00000000)	99.8	80.8	76.3	24.5
17	*B. viridifuturi* SEMIA 690^T (LGTB00000000)	99.3	80.9	77.0	24.3
18	*B. yuanmingense* CGMCC 1.3531^T (VLKS00000000)	98.1	79.3	74.9	22.9
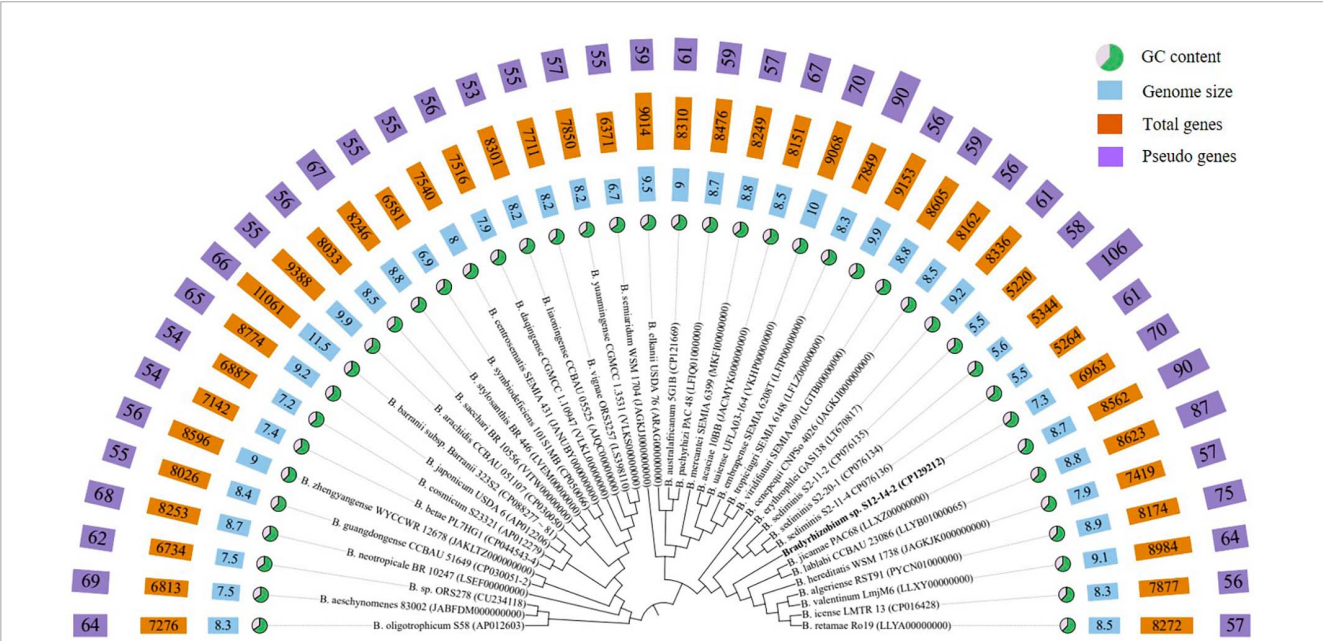
FIGURE 1

Phylogenomic tree based on genome sequences of strain S12-14-2^T and its closely related type strains of *Bradyrhizobium* in the TYGS (https://tygs.dsmz.de/) and iTOL.

is a powerful tool for discriminating species (Rivas et al., 2009; Tang et al., 2012; Duran et al., 2014). The phylogenetic tree of five concatenated housekeeping gene sequences (*ITS*, *dnaK*, *glnII*, *recA*, and *rpoB*) located the novel strain in a clearly distinct lineage from the named *Bradyrhizobium* species. The phylogenetic tree of the

concatenated *ITS-dnaK-glnII-recA-rpoB* gene sequences confirmed the placement of the novel strain in a lineage distinct from the named *Bradyrhizobium* species (Supplementary Figure 3). ANI, AAI, and dDDH analyses based on whole-genome sequences were implemented for additional characterization. We estimated

the values for the complete genome sequence of S12-14-2 by conducting pair-wise comparisons with genome sequences of the type strains of *Bradyrhizobium* species obtained from public databases. The ANI values of the named species ranged from 79.2% (*Bradyrhizobium daqingense*) to 86.6% (*B. algeriense*), which is significantly lower than the 95–96% threshold value for circumscription of bacterial species (Konstantinidis and Tiedje, 2005; Rivas et al., 2009; Chun et al., 2018; Ciufo et al., 2018). The AAI (dDDH) values for all pairwise comparisons between the novel strain and related *Bradyrhizobium* species ranged from 71.6 to 86.5% (22.8–32.5%) (Table 1), which is also under the borderline of 95% (ANI) and 70% (dDDH) species delineation (Goris et al., 2007; Richter and Rossello-Mora, 2009; Luo C. et al., 2014; Konstantinidis et al., 2017).

General genome properties

The genome of isolate S12-14-2 has a size of 7,319,411 base pairs, a GC content of 63.3%, three rRNA genes, 62 tRNA genes, and 6893 CDS, of which 5,838 genes were allocated to COG (Supplementary Table 3). *Bradyrhizobium* spp. have relatively large genomes compared to other groups, ranging between 5 and 11M, whereas the smallest genome size found to date is approximately 5M in the free-living strain S2-11-2 (Ormeño-Orrillo and Martínez-Romero, 2019; Jin et al., 2022b; Klepa et al., 2022; Zhang et al., 2023). The GC content is consistent with other *Bradyrhizobium* type strains (Figure 1 and Supplementary Table 2). The abundant genes were associated with amino acid transport and metabolism (COG category E), inorganic ion transport and metabolism (COG category P), energy production and conversion (COG category C), transcription (COG category K), signal transduction mechanisms (COG category T), and cell wall/membrane/envelope biogenesis (COG category M), and lipid transport and metabolism (COG category I), and approximately 32.5% of the genes were assigned to the unknown function COG category (Supplementary Table 4).

Pangenomics and secondary metabolites

To investigate the genomic diversity of *Bradyrhizobium* sp. S12-14-2 and related *Bradyrhizobium* spp., we compiled a total of 44 complete whole genomes (including 7 free-living and 34 symbiotics) downloaded from the NCBI. This dataset contains genomes representing 41 species validly published. Across the entire dataset, the genome size (GS) varied between 5.5 and 11.5M (mean, 8.5M); the GC content varied between 61.4 and 65.1% (mean, 64.9), and the number of predicted genes ranged from 5,220 to 11,061 (mean, 8,111).

Following the workflow, pangenomic analysis of 37 *Bradyrhizobium* strains was performed using Anvi'o (version 7.1). Single copy gene clusters (SCGs) are shown in purple color on the outside, which are drawn by 1,235 gene clusters from 43,916 splits and 44,526 objects. *Bradyrhizobium* sp. S12-14-2^T exhibited a similar pattern of blue-colored single copy gene clusters as the other three free-living microorganisms *Bradyrhizobium sediminis* S2-20-1^T, *B. sediminis* S2-11-2, and

B. sediminis S2-11-4. SCGs were also the same with the core genes, and others not included in the SCGs were accessory genes. *B. erythrophlei* GAS138^T (LT670817) has the largest number of singleton gene clusters which is almost equal to 2,046 (Figure 2). BGC from 44 *Bradyrhizobium* genomes from the free-living and symbiotic groups were identified using antiSMASH, and a comprehensive network analysis was performed in BiG-SCAPE using all similar BGC sequences from the MIBiG databases. The average number of BGCs from the 44 genomes is 11, and the network analysis distributed 557 BGCs in different types, including 99 terpenes, 84 NRPS (non-ribosomal peptides), 93 hserlactones (homoserine lactones), 44 redox-cofactors (redox-cofactors such as PQQ), 47 RiPP-like (other unspecified ribosomally synthesized and post-translationally modified peptide products), 35 NRPS-like (NRPS-like fragmenta), and 37 TPKS (type I, II, and III PKSs), and generally, the number of secondary metabolites from free living bacteria are less than symbiotic bacteria (Supplementary Figure 4). We identified common BGCFs shared by all strains, such as terpenes and redox-cofactors and nine BGCFs not found in free-living strains, such as *Bradyrhizobium* sp. S12-14-2 (Supplementary Figure 4). Additionally, *Bradyrhizobium* sp. S12-14-2 possesses a phosphonate-producing gene cluster, confirmed by phosphorous metabolism-related genetic analysis.

Genome size of *Bradyrhizobium*

Genome size variation is a biological trait, although its evolutionary sources and effects are debated. GS's neutrality, selective pressures, and strength are unknown. The genetic segments responsible for this diversity directly affect evolutionary outcomes and are targeted by diverse evolutionary pressures. *Bradyrhizobium* isolates have rather big genomes ranging from 5 to 11M, with an average size of 8.6M (Ormeño-Orrillo and Martínez-Romero, 2019; Avontuur et al., 2022; Jin et al., 2022b; Claassens et al., 2023; Lafay et al., 2023; Zhang et al., 2023). Nevertheless, the range of GSs within the genus remains unknown. The smallest genome was 5.5M in length and related to strain S12-14-2 recovered from freshwater sediment. The largest full assembly (11.7M) was discovered from the soil isolate GAS478 (FSRD01000000). Notably, strains S12-14-2 (7.3M), S2-20-1^T (5.6M), S2-11-2 (5.5M), S2-11-4 (5.5M), and GAS478 (11.7M) lacked symbiosis genes, suggesting that GS in *Bradyrhizobium* is unrelated to its capacity to interact with legumes. In the *Nitrobacteraceae* family, *Bradyrhizobium* has the largest genome, followed by *Tardiphaga* and *Bosea*. *Bradyrhizobium*, *Tardiphaga*, and *Bosea*, three taxa with large genomes, interact with plants (De Meyer et al., 2012; Ormeño-Orrillo and Martínez-Romero, 2019). And taxa with small genomes, including *Nitrobacter*, *Oligotropha*, and *Variibacter*, were discovered to be metabolically constrained, indicating that the *Nitrobacteraceae* GS appears to be associated with lifestyle. Understanding the GS variation and its causes and effects requires trustworthy methodologies; however, there are still obstacles. From genome annotation and content analysis to laboratory- and sequence-based GS measurement accuracy, there are problems and recommended practices. Even among big consortia building high-quality genome
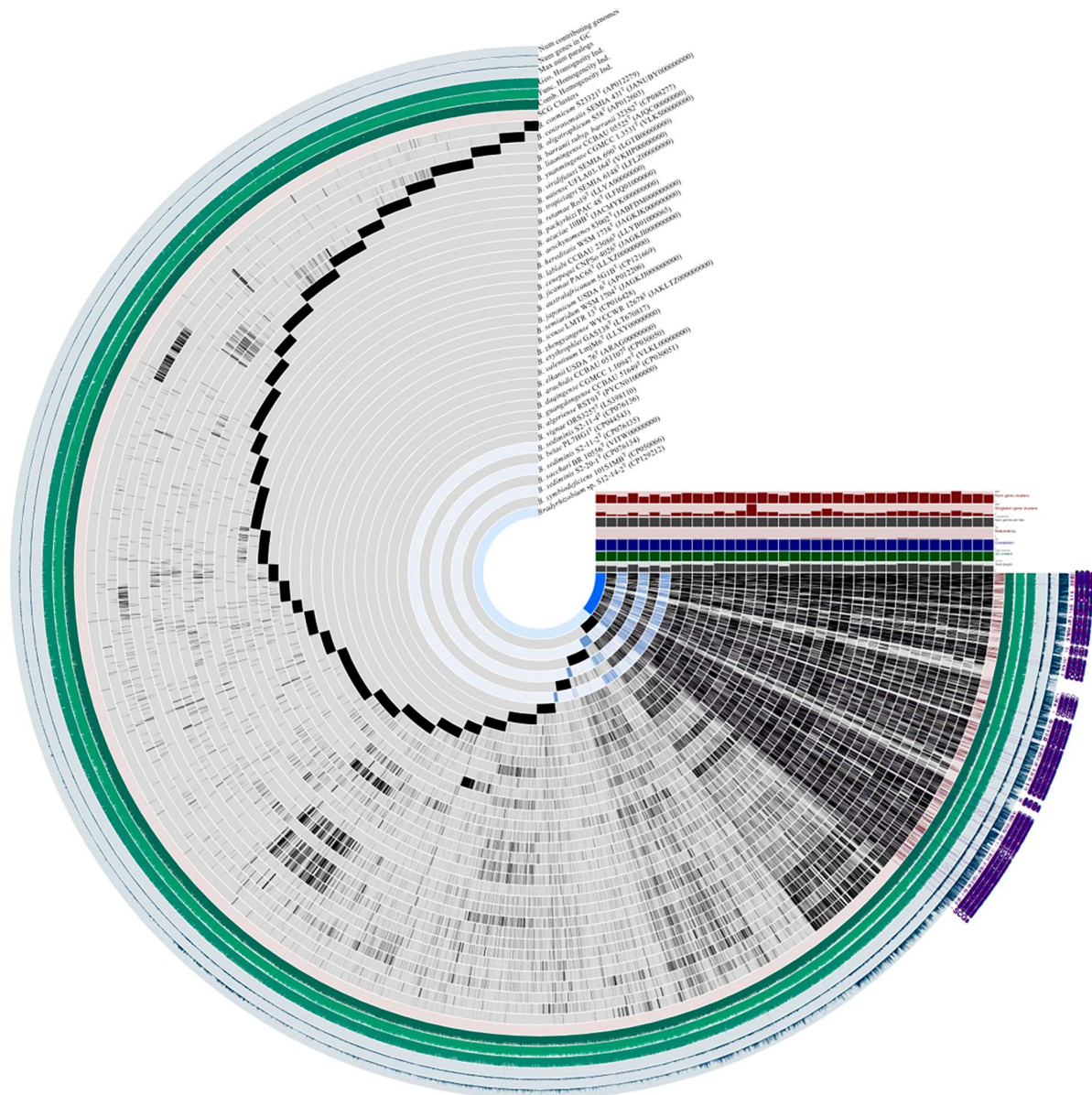
FIGURE 2
Pangenome analysis of *Bradyrhizobium* species including *Bradyrhizobium* sp. S12-14-2^T and closely related strains calculated with Anvi'o version 7.1. Dark regions in purple color (SCGs) represent genes (black) found in each genome at least one time.

assemblies, gold standards and best practices are continually changing.

Motility

The genome of strain S12-14-2 contains 192 genes for motility via flagella and genes for chemotaxis (data not shown). At least 132 genes related to flagellar motility (components of the flagellar apparatus or specific regulators) are grouped in two clusters in the genome separated by unknown genes. Notably, the cultured S12-14-2 cells lack any discernible flagella or pili (**Supplementary Figure 1**). The reason could be that twitching motility may be a more efficient mode of locomotion in environments devoid of

conditions that require or prohibit pilus-facilitated or flagellum-powered movement.

Genes involved in carbon metabolism and photosynthesis

Free-living *Bradyrhizobium* organisms also perform vital roles in soil ecology, nutrient cycling, and carbon metabolism. As free-living microbes, they must acquire carbon from other environmental sources. *Bradyrhizobium* can metabolize a vast array of carbon sources, including simple carbohydrates, organic acids, and amino acids, via diverse biochemical pathways (VanInsberghe et al., 2015; Tao et al., 2021). The TCA (tricarboxylic acid) cycle,

also known as the Krebs cycle, is one of the primary pathways involved in carbon metabolism in free-living *Bradyrhizobium*. This pathway is essential for the growth and survival of microorganisms as it is involved in producing energy from carbon sources. The complete gene sets of the TCA cycle, Entner–Doudoroff pathway, and glycolysis and gluconeogenesis were observed in the genome of strain S12-14-2. Furthermore, genes for encoding phosphoenolpyruvate (PEP) carboxylase (EC 4.1.1.31) and pentose phosphate pathway were also observed. The genome of S12-14-2 contained genes for encoding RuBisCO (ribulose 1,5-bisphosphate carboxylase/oxygenase) or the light-harvesting complex (**Figure 3A**), which means that strain S12-14-2 may fix CO_2 and potentially have photosynthetic function.

The gene cluster responsible for photosynthesis in *Bradyrhizobium* encompasses several genes that have a crucial role in the production of pigments and proteins necessary for the process of photosynthesis. The genes in question encode various proteins, such as bacteriochlorophyll, carotenoids, and reaction center proteins (Gregor and Klug, 1999; Igarashi et al., 2001; Hata et al., 2023). The presence of genes associated with photosynthesis in the genome of S12-14-2 has been verified. The genetic components in this set are as follows: bacteriochlorophyll genes (*bchCXYZGFPF*), genes for light-independent protochlorophyllide reductase (*chlIDNBL*), carotenoid genes (*crtCBICDEF*), genes for light harvesting polypeptides (*pucCAB*), and genes for reaction center subunits (*puhBA/pufLM*) (**Figure 3B**). Several species of *Bradyrhizobium* have the ability to utilize carbon dioxide through the Calvin cycle, also known as the Calvin–Benson–Bassham cycle. In this metabolic pathway, the enzyme ribulose-1, 5-bisphosphate carboxylase oxygenase (RuBisCo) has a crucial role. The strain S12-14-2 was found to possess all gene sets associated with the Calvin cycle except for a gene responsible for encoding sedoheptulose-1,7-bisphosphatase (SBP) (**Figure 3A**). The Calvin cycle involves the execution of distinct tasks by RuBisCo (RbcL), phosphoribulokinase (PRK), and SBP. The combined utilization of RuBisCo and PRK can be used as reliable indicators for the presence of the cycle within a particular organism. However, the enzymatic activity of SBP in bacteria is commonly facilitated by fructose-1,6-bisphosphatase (FBP) with a dual sugar (Jiang et al., 2012).

Genes involved in nitrogen metabolism and nodulation

Nitrogen fixation is extremely important in environmental sustainability, particularly in nitrogen-deficient soil (Lindström et al., 2010; Hungria et al., 2015). In legume-rhizobium symbiosis, this activity is key for soil health and plant growth (Jensen and Hauggaard-Nielsen, 2003; Meena et al., 2018). Genes encoding NifKDH (nitrogenase α-, ß-subunits and reductase, respectively); and NifENB (nitrogenase assembly proteins) are considered the minimal gene set required for a diazotrophic phenotype; those genes encode a FeMo-cofactor-containing nitrogenase (Kaneko et al., 2002; Dos Santos et al., 2012; Sprent et al., 2017). All genes encoding these nitrogen-fixation like proteins in S12-14-2 (genes and proteins abbreviated *nif* and Nif, respectively, and numbered 1–2), *nifABDEHKNOQSTUVWXZ*, were clustered (**Figure 4A**).

The genes encoding $NifH_{1,2}DK$ are located upstream of the cluster, while the genes encoding NifNE are adjacent to *nifKDH*. In contrast to symbiotic strains, free-living *Bradyrhizobium* bacteria do not have *nif* or *nod* genes, and some nod-independent strains contain just *nif* genes (Ferrieres et al., 2004; Bobik et al., 2006; Meilhoc et al., 2010; VanInsberghe et al., 2015). No nodulation genes were identified in the genome, indicating that strain S12-14-2 could fix atmospheric nitrogen; as for nodulation, it is uncertain whether strain S12-14-2 forms nodules at this time because of nod-factor-independence of symbiosis.

Denitrification is a prominent contributor to nitrous oxide (N_2O) emission in soil, serving as a substantial constituent of the worldwide nitrogen cycle. This process takes place in both terrestrial and marine ecosystems. Certain bacteria can adjust their metabolic processes in response to oxygen-depleted settings, wherein the concentration of dissolved and accessible oxygen is diminished. In such conditions, these bacteria use nitrate as a respiratory substrate for the process of denitrification (Coates et al., 2001; Luo F. et al., 2014; Zhang et al., 2020). The strain S12-14-2 contains a periplasmic nitrate reductase (Nar), a copper (Cu)-containing nitrite reductase (Nir), and a Cu-dependent nitrous oxide reductase (Nos) encoded by the *narGHJI, nirKV*, and *nosRZDFYLX* genes, respectively (**Figure 4B**). Nitrite produced by NarGHJI will be released to the periplasm through a membrane transporter and reduced to nitric oxide (NO) by Cu-containing nitrite reductase (NirK) (Richardson, 2000; Simon and Klotz, 2013). Due of its high cytotoxicity, NO will immediately reduce to N_2O by Nor (nitric oxide reductase). Typically, nor enzymes are usually periplasmic cytochrome *c* (cNor) or quinones (qNor) (Nordling et al., 1990; Pearson et al., 2003; Rinaldo and Cutruzzola, 2007; Simon and Klotz, 2013). Notably, the genome does not contain NO reductase genes for the reduction stage. The apparent absence of Nor-encoding genes in strain S12-14-2 is of special interest, but it is speculated that some cNor or qNor-like proteins are encoded by some other unknown genes.

Genes involved in sulfur metabolism

Sulfur, an amino acid component and enzyme cofactor, is essential to all organisms. Sulfide is formed by the sulfate assimilation pathway in many bacteria and integrated into sulfur-containing organic compounds. Under sulfur-limiting conditions, bacteria must receive sulfur from the environment because inorganic sulfate is not prevalent in nature. Sulfonates and sulfonate esters are natural or xenobiotic compounds (Kertesz, 2000; Van Der Ploeg et al., 2001b). Genes encoding a sulfur-oxidizing (Sox) function were discovered, and the *sox* gene cluster consists of 16 genes. *soxR* encodes a repressor protein, and *soxSW* encodes a periplasmic thioredoxin, which is essential for full expression of the *sox* genes. The subsequent *soxXYZABCD* genes encode four periplasmic proteins that reconstitute the Sox enzyme system: SoxXA, SoxYZ, SoxB, and SoxCD, which can convert a variety of reduced sulfur compounds to sulfates (**Figure 5**). Strain S12-14-2 also possess the *soxF* gene, which encodes flavoprotein SoxF with a sulfide dehydrogenase activity (Quentmeier et al., 2004). These sulfur oxidation Sox systems suggest that strain S12-14-2 might be capable of oxidizing thiosulfate to sulfate.
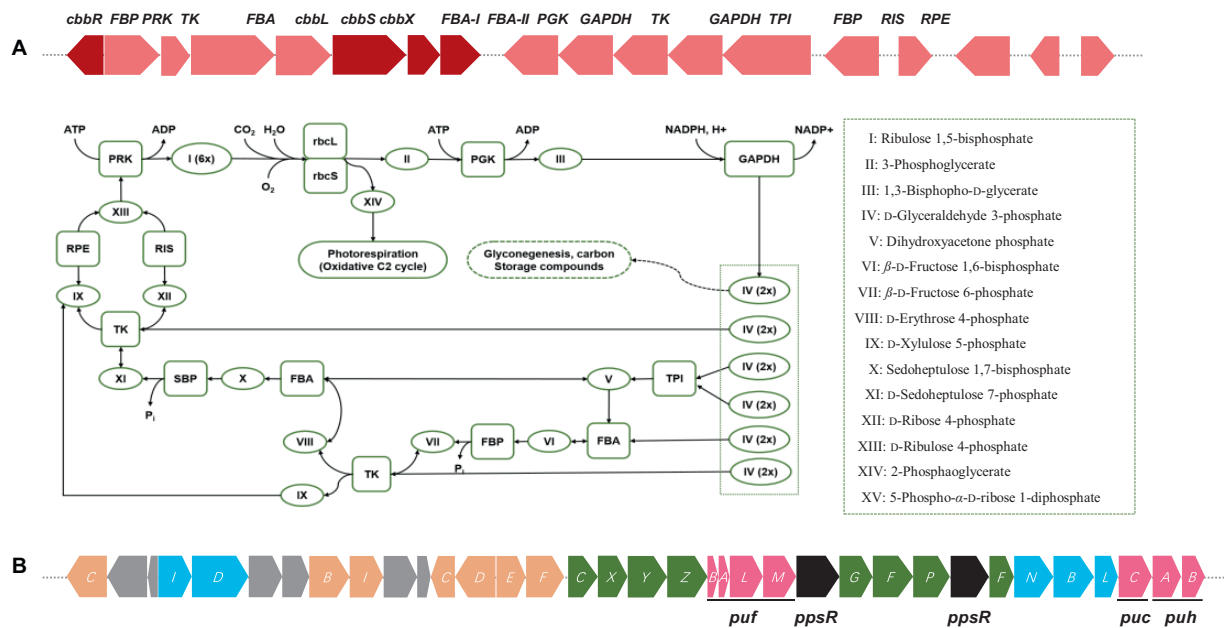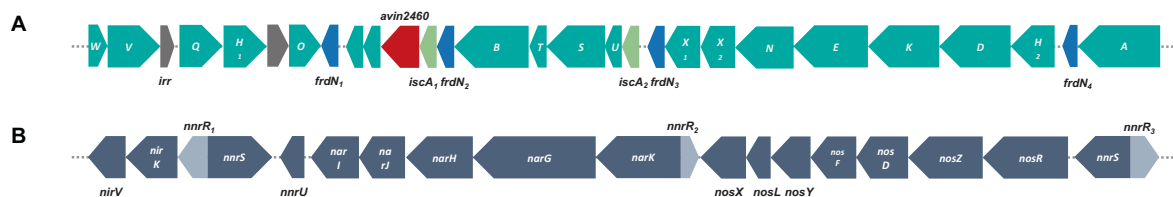
FIGURE 3

(A) Diagram of Calvin cycle and RuBisCO gene cluster arrangements including RuBisCO (RbcL), phosphoribulokinase (PRK), fructose-1,6-bisphosphatase (FBP), etc. Color keys: red, RuBisCO subunits and RuBisCo activation protein coding genes: rose, other carbon fixation related genes. **(B)** Photosynthetic gene cluster of *Bradyrhizobium* sp. S12-14-2^T. Color keys: green, *bch* genes (bacteriochlorophyll synthesis); orange, *crt* genes (carotenoid synthesis); blue, *chl* gene (light-independent protochlorophyllide reductase); pink, *puh* (photosynthetic reaction centers), *puf* and *puc* (light-harvesting complexes); black, regulatory proteins; gray, uncharacterized proteins.

FIGURE 4

Genetic organization of nitrogen-fixation and denitrification related genes and surrounding genes. **(A)** Nitrogen-fixation gene clusters located on genome of strain S12-14-2^T. **(B)** Denitrification gene clusters; color keys: teal, *nif* genes; gray, *irr* genes (iron-responsive regulator); blue, *frdN* genes (4Fe-4S ferredoxin, nitrogenase-associated); red, *avin2460* [LRV (FeS)4 cluster domain protein clustered with nitrogenase cofactor synthesis]; light green, *iscA* genes (probable iron binding protein in *nif* operon); gray, denitrification related genes: *nor*, nitric oxide reductase; *nar*, nitrate reductase; *nir*, nitrite reductase; *nnr* genes [*nnrS* (NO sensing protein); *nnrU* (denitrification regulatory protein)].

Under sulfur-limiting conditions, several bacteria express proteins to use organosulfonates or sulfonate esters as sulfur sources. Sulfate starvation-induced proteins ingest sulfonate, acquire sulfur from organic molecules, and defend against reactive oxygen species (Hummerjohann et al., 1998; Van Der Ploeg et al., 2001a; Ellis, 2011). A sulfonate transport system has been detected in the genome of *Bradyrhizobium* sp. S12-14-2, which includes sulfonate-binding protein (SsuA), ATP-binding protein (SsuB), and sulfonate permease protein (SsuC). The genes *tauD* and *ssuD,* which encode a taurine dioxygenase and an alkanesulfonate monooxygenase, respectively, were also found to be responsible for sulfonate and taurine metabolism. The *tau* operon encodes an ABC-type transporter and taurine dioxygenase (TauD), which oxidize taurine to hydroxytaurine. The unstable hydroxytaurine breaks down to yield the aminoacetaldehyde and sulfite products (Eichhorn et al., 1997). The SsuD monooxygenase enzyme catalyzes

the desulfonation of a wide variety of organosulfonated products, producing the corresponding aldehyde and sulfonate (Eichhorn et al., 1999).

The structures of TauD and SsuD monomeric units were predicted using the SWISS-MODEL. The SsuD monomer consists of a total of 359 amino acid residues and belongs to the structural family of TIM-barrel proteins. SsuD is classified as a flavin-dependent oxidoreductase belonging to the LLM (luciferase-like monooxygenase) class. On the other hand, the TauD monomer consists of 297 amino acid residues and functions as an α-ketoglutarate-dependent dioxygenase (**Figure 5**). TauD prefers taurine but may also desulfonate alkanesulfonates. Although alkanesulfonate monooxygenase enzymes can use various organosulfate substrates, they have no enzymatic activity against taurine (Ellis, 2011). These TauD and SsuD enzymes possibly work together to ensure that the bacterial cell has sufficient
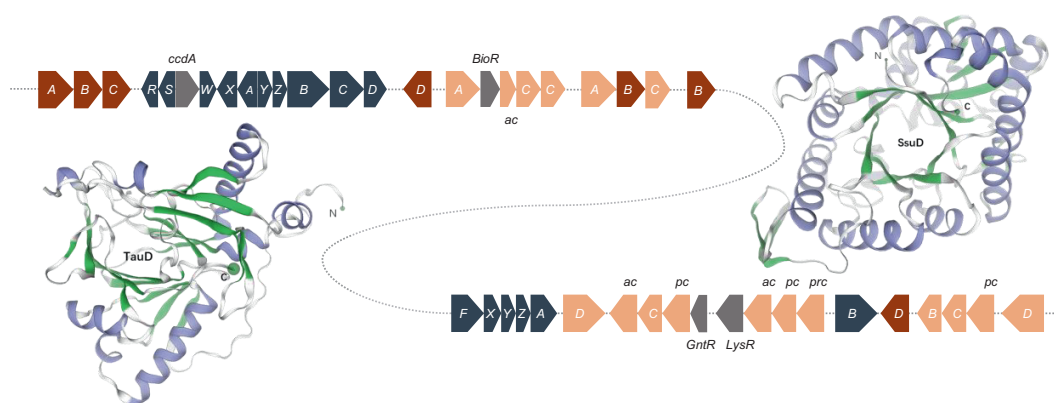
FIGURE 5

Sulfur oxidation gene clusters and of structure of the monomeric unit of TauD and SsuD. The TauD and SsuD enzyme exist in the form of a TIM-barrel containing 297 and 359 amino acid residues, respectively. Color key: blue gray, *sox* genes; red, *tau* genes; orange, *ssu* genes (*ac*, ABC type nitrate/sulfonate/bicarbonate transport system, ATPase component; *pc*, ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components; *prc*, ABC-type nitrate/sulfonate/bicarbonate transport system, permease component.); gray, regulator genes.
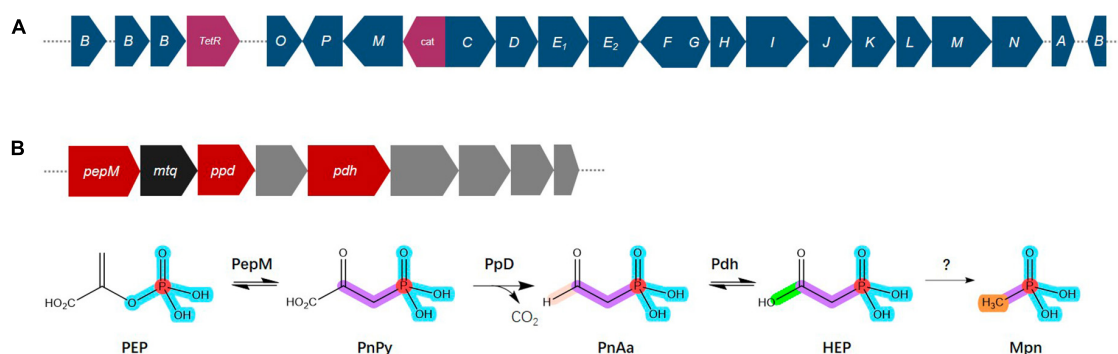
FIGURE 6

Overview of the C–P lyase gene cluster and pathway. **(A)** Overview of *phn* gene organization. **(B)** Pathway for phosphonate production and gene organization. Phosphoenolpyruvate (PEP) mutase (PepM) isomerizes PEP to phosphonopyruvate (PnPy), which is the precursor to all known phosphonate molecules. Due to this functional conservation, the *pepM* gene is a reliable genomic marker for the potential to produce phosphonates. Phosphonopyruvate decarboxylase (Ppd) converts PnPy to phosphonoacetaldehyde (PnAa), phosphonoacetaldehyde dehydrogenase (Pdh) dehydrogenates of PnAa to 2-hydroxyethylphosphonic acid (HEP), and finally to methylphosphonate (Mpn). Color key: dark blue, *phn* genes; purple, *TelR* (transcriptional regulator), *cat* (chloramphenicol acetyltransferase); red, phosphonate production related genes; black, *mtq* (O-methyltransferase); gray, transporters.

sulfur for biosynthetic processes. While most of these enzymes' kinetic properties are known, much remains to be learned about their detailed mechanistic properties.

Genes involved in phosphate metabolism

Phosphorus is an essential nutrient for all organisms and has crucial roles in respiration, energy production and storage, the biosynthesis of nucleic acids, ATP, and phospholipids and other physiological and biochemical processes in specific environments (Balemi and Negisho, 2012; Nath et al., 2017). Phosphorus availability and plant assimilation are governed by microorganisms, and phosphorus-solubilizing microorganisms are capable of dissolving insoluble phosphorus in soil and converting it into soluble phosphorus for plant absorption and use (Cao, 2014; Moro et al., 2021). Phosphonate compounds, which contain a direct C–P bond instead C–O–P ester linkage, are prevalent in the environment, where they are a significant source of organic phosphorus and a common pollutant (Horsman and Zechel, 2017; Rott et al., 2018). The degradation of phosphonates is especially advantageous for microorganisms, where the bioavailability of phosphorus is frequently a growth-limiting factor (Sosa et al., 2019). The genome of *Bradyrhizobium* sp. S12-14-2 contains all the predicted C–P lyase genes necessary for phosphonate degradation and the additional putative auxiliary functions (**Figure 6A**). These genes included a *TetR* family transcriptional regulator, phosphonate ABC transporters *phnCDE$_1$E$_2$*, purine ribonucleoside triphosphate phosphorylase components *phnGHIJKL*, a chloramphenicol acetyltransferase, a triphosphoribosyl 1-phosphonate diphosphohydrolase *phnM*, and a phosphoribosyl cyclic phosphodiesterase *phnP*.

Phosphonates are a type of phosphorus-based metabolites distinguished by a highly stable C–P bond. They comprise

TABLE 2 Descriptions of *Bradyrhizobium roseus* sp. nov.

Genus name	*Bradyrhizobium*
Species name	*Bradyrhizobium roseus*
Species epithet	*roseus* sp. nov.
Species etymology	ro′se.us. L. masc. adj. *roseus*, rose colored, pink
Description of the new taxon and diagnostic traits	Cells are Gram-stain-negative and rods. Colonies grown on R2A agar are pink. Growth occurs on R2A at temperatures from 15 to 30°C (optimum temperature 25–30°C) but not 10 or 37°C. The pH range for growth is from pH 6.0 to 9.0 and optimum pH 7–8 but not at pH 5.0 and 10.0. Cells are found to be positive for nitrate reduction and indole production, but negative for glucose acidification, urease, esculin hydrolysis, gelatin hydrolysis, and β-galactosidase. Cells utilize L-proline and L-serine but not acetate, N-acetyl glucosamine, adipate, L-alanine, L-arabinose, caprate, citrate, L-fucose, gluconate, D-glucose, glycogen, histidine, 3-hydroxybenzoate, 4-hydroxybenzoate, 3-hydroxybutyrate, inositol, itaconate, 2-ketogluconate, 5-ketogluconate, lactate, malate, malonate, D-maltose, D-mannitol, D-mannose, D-melibiose, phenylacetate, propionate, L-rhamnose, D-ribose, salicin, D-sorbitol, suberate, D-sucrose, or valerate. Cells are found to be positive for the following enzyme activities: alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase and naphthol-AS-BI-phosphohydrolase (weakly). Cells are found to be negative for the following enzyme activities: *N*-acetyl-β-glucosaminidase, acid phosphatase, α-chymotrypsin, cystine arylamidase, α-fucosidase, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, β-glucuronidase, lipase (C14), α-mannosidase, trypsin, and valine arylamidase. The major fatty acids are grouped into two categories $C_{18:1}$ ω7*c* and/or $C_{18:1}$ ω6*c* and $C_{16:0}$.
Region of origin and sampling date (mm/yyyy)	Hongze Lake, China (10/2018)
Latitude (xx°xx′xx″N/S, xx°xx′xx″E/W)	33°14′26″N, 118°35′40″E
The GenBank accession number for 16S rRNA gene and genome	MZ315001, CP129212 (complete genome)
Genome size (GC content)	7,319,411 bp (63.3%)
Designation of the type strain and isolation source	S12-14-2[T] isolated from freshwater sediment
Strain collection numbers	JCM 34606[T]; CGMCC 1.19422[T]

a significant portion of the dissolved organic phosphorus reservoir in the oceans and are widely distributed among more primitive life forms (Villarreal-Chiu et al., 2012; Acker et al., 2022). Almost all known pathways of phosphonate biosynthesis involve a central C–P bond-forming reaction in which the enzyme PEP phosphomutase (Ppm) intramolecularly rearranges the intermediary metabolite phosphoenolpyruvate (PEP) to form phosphonopyruvate (PnPy) (Metcalf and Van Der Donk, 2009; Acker et al., 2022). The genome of *Bradyrhizobium* sp. S12-14-2 contains the genes *pepM* which encodes Ppm, *ppd* encoding phosphonopyruvate decarboxylase (Ppd) which subsequently converts PnPy to phosphonoacetaldehyde (PnAa) in next reaction, and *pdh* encoding phosphonoacetaldehyde dehydrogenase (Pdh) which dehydrogenates of PnAa to 2-hydroxyethylphosphonic acid (HEP) (Figure 6B).

Conclusion

To summarize our investigations, the strain S12-14-2[T] from a contaminated freshwater sediment belonging to the genus *Bradyrhizobium* was studied through genome analysis and a polyphasic approach. The phylogenetic analysis indicated that the strain S12-14-2[T] was affiliated closely with the species of *Bradyrhizobium*, based on the phylogenetic, genomic, and physiological differences; thus, we propose strain S12-14-2[T] as a novel species, *Bradyrhizobium roseus* sp. nov., in the family *Nitrobacteraceae* (Table 2). The *B. roseus* pan-genome is considered to be in an open state, genetically diverse, and with a flexible gene repertoire associated with multiple functions. *B. roseus* contained

genes encoding all enzymes of the Calvin–Benson cycle and phototrophic systems with carotenoids synthesizing genes and possesses complete genes for nitrogen fixing and denitrification. *B. roseus* S12-14-2[T] also contains genes for sulfur oxidation and the essential genes encoding for the sulfonate transport system. Moreover, strain S12-14-2[T] contains all the necessary genes for phosphonate degradation and phosphate biosynthesis. The physiological characteristics and comparative genome analyses of *B. roseus* S12-14-2[T] help us to understand the metabolism and evolutionary features of the genus *Bradyrhizobium* and provide understanding of this microorganism and a new genetic framework for future studies.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary material**.

Author contributions

NZ: Conceptualization, Data curation, Writing – original draft. C-ZJ: Conceptualization, Data curation, Writing – original draft. YZ: Data curation, Validation, Visualization, Writing – original draft. TL: Data curation, Validation, Visualization, Writing – original draft. F-JJ: Data curation, Validation, Visualization,

Writing – original draft. H-GL: Resources, Writing – original draft, Writing – review and editing. LJ: Resources, Supervision, Writing – original draft, Writing – review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1295854/full#supplementary-material

References

Acker, M., Hogle, S. L., Berube, P. M., Hackl, T., Coe, A., Stepanauskas, R., et al. (2022). Phosphonate production by marine microbes: Exploring new sources and potential function. *Proc. Natl. Acad. Sci. U. S. A.* 119:e2113386119. doi: 10.1073/pnas.2113386119

Avontuur, J. R., Palmer, M., Beukes, C. W., Chan, W. Y., Tasiya, T., Van Zyl, E., et al. (2022). *Bradyrhizobium altum* sp. nov., *Bradyrhizobium oropedii* sp. nov. and *Bradyrhizobium acaciae* sp. nov. from South Africa show locally restricted and pantropical nodA phylogeographic patterns. *Mol. Phylogenet. Evol.* 167:107338. doi: 10.1016/j.ympev.2021.107338

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Evol. Biol.* 9:75. doi: 10.1186/1471-2164-9-75

Aziz, R. K., Devoid, S., Disz, T., Edwards, R. A., Henry, C. S., Olsen, G. J., et al. (2012). SEED servers: High-performance access to the SEED genomes, annotations, and metabolic models. *PLoS One* 7:e48053. doi: 10.1371/journal.pone.0048053

Balemi, T., and Negisho, K. (2012). Management of soil phosphorus and plant adaptation mechanisms to phosphorus stress for sustainable crop production: A review. *J. Soil Sci. Plant. Nutr.* 12, 547–561. doi: 10.4067/s0718-95162012005000015

Bienert, S., Waterhouse, A., de Beer, T. A. P., Tauriello, G., Studer, G., Bordoli, L., et al. (2017). The SWISS-MODEL Repository - new features and functionality. *Nucleic Acids. Res.* 45, D313–D319. doi: 10.1093/nar/gkw1132

Bobik, C., Meilhoc, E., and Batut, J. (2006). FixJ: A major regulator of the oxygen limitation response and late symbiotic functions of *Sinorhizobium meliloti*. *J. Bacteriol.* 188, 4890–4902. doi: 10.1128/jb.00251-06

Cao, N. D. (2014). Phosphate and potassium solubilizing bacteria from weathered materials of denatured rock Mountain, Ha Tien, Kien Giang Province, Vietnam. *Am. J. Life. Sci.* 1, 88–92. doi: 10.11648/j.ajls.20130103.12

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., et al. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466. doi: 10.1099/ijsem.0.002516

Ciufo, S., Kannan, S., Sharma, S., Badretdin, A., Clark, K., Turner, S., et al. (2018). Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int. J. Syst. Evol. Microbiol.* 68, 2386–2392. doi: 10.1099/ijsem.0.002809

Claassens, R., Venter, S. N., Beukes, C. W., Stępkowski, T., Chan, W. Y., and Steenkamp, E. T. (2023). *Bradyrhizobium xenonodulans* sp. nov. isolated from nodules of Australian Acacia species invasive to South Africa. *Syst. Appl. Microbiol.* 46:126452. doi: 10.1016/j.syapm.2023.126452

Coates, J. D., Chakraborty, R., Lack, J. G., O'Connor, S. M., Cole, K. A., Bender, K. S., et al. (2001). Anaerobic benzene oxidation coupled to nitrate reduction in pure culture by two strains of *Dechloromonas*. *Nature* 411, 1039–1043. doi: 10.1038/35082545

De Meyer, S. E., Coorevits, A., and Willems, A. (2012). *Tardiphaga robiniae* gen. nov., sp. nov., a new genus in the family *Bradyrhizobiaceae* isolated from *Robinia pseudoacacia* in Flanders (Belgium). *Syst. Appl. Microbiol.* 35, 205–214. doi: 10.1016/j.syapm.2012.02.002

Delamuta, J. R. M., Ribeiro, R. A., Ormeño-Orrillo, E., Melo, I. S., Martínez-Romero, E., and Hungria, M. (2013). Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov. *Int. J. Syst. Evol. Microbiol.* 63, 3342–3351. doi: 10.1099/ijs.0.049130-0

Delmont, T. O., and Eren, A. M. (2018). Linking pangenomes and metagenomes: The *Prochlorococcus metapangenome*. *PeerJ* 6:e4320. doi: 10.7717/peerj.4320

Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., and Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* 13:162. doi: 10.1186/1471-2164-13-162

Duran, D., Rey, L., Mayo, J., Zuniga-Davila, D., Imperial, J., Ruiz-Argüeso, T., et al. (2014). *Bradyrhizobium paxllaeri* sp. nov. and *Bradyrhizobium icense* sp. nov., nitrogen-fixing rhizobial symbionts of Lima bean (*Phaseolus lunatus* L.) in Peru. *Int. J. Syst. Evol. Microbiol.* 64, 2072–2078. doi: 10.1099/ijs.0.060426-0

Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997). Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J. Biol. Chem.* 272, 23031–23036. doi: 10.1074/jbc.272.37.23031

Eichhorn, E., van der Ploeg, J. R., and Leisinger, T. (1999). Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. *J. Biol. Chem.* 274, 26639–26646. doi: 10.1074/jbc.274.38.26639

Ellis, H. R. (2011). Mechanism for sulfur acquisition by the alkanesulfonate monooxygenase system. *Bioorg. Chem.* 39, 178–184. doi: 10.1016/j.bioorg.2011.08.001

Ferrieres, L., Francez-Charlot, A., Gouzy, J., Rouille, S., and Kahn, D. (2004). Fix-regulated genes evolved through promoter duplication in *Sinorhizobium meliloti*. *Microbiology* 150, 2335–2345. doi: 10.1099/mic.0.27081-0

Gao, Y., Lesven, L., Gillan, D., Sabbe, K., Billon, G., Galan, S. D., et al. (2009). Geochemical behavior of trace elements in sub-tidal marine sediment of the Belgian coast. *Mar. Chem* 117, 88–96. doi: 10.1016/j.marchem.2009.05.002

Germano, M. G., Menna, P., Mostasso, F. L., and Hungria, M. (2006). RFLP analysis of the rRNA operon of a Brazilian collection of bradyrhizobial strains from 33 legume species. *Int. J. Syst. Evol. Microbiol.* 56, 217–229. doi: 10.1099/ijs.0.02917-0

Gnat, S., Małek, W., Ole´nska, E., Wdowiak-Wróbel, S., Kalita, M., Rogalski, J., et al. (2016). Multilocus sequence analysis supports the taxonomic position of *Astragalus glycyphyllos* symbionts based on DNA–DNA hybridization. *Int. J. Syst. Evol. Microbiol.* 66, 1906–1912. doi: 10.1099/ijsem.0.000862

Gogoi, N., Baruah, K. K., Meena, R. S., Das, A., Yadav, G. S., and Lal, R. (eds) (2018). "Grain legumes: Impact on soil health and agroecosystem," in *Legumes for Soil Health and Sustainable Management*, ed. R. S. Meena (Singapore: Springer Nature Singapore Pte Ltd), 511–539. doi: 10.1007/978-981-13-0253-4_16

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0

Gregor, J., and Klug, G. (1999). Regulation of bacterial photosynthesis genes by oxygen and light. *FEMS. Microbiol. Lett.* 179, 1–9. doi: 10.1111/j.1574-6968.1999.tb08700.x

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95–98. doi: 10.1021/bk-1999-0734.ch008

Hata, S., Kojima, S., Tsuda, R., Kawajiri, N., Kouchi, H., Suzuki, T., et al. (2023). Characterization of photosynthetic *Bradyrhizobium* sp. strain SSBR45 isolated from the root nodules of *Aeschynomene indica*. *Plant Signal. Behav.* 18:2184907. doi: 10.1080/15592324.2023.2184907

Horsman, G. P., and Zechel, D. L. (2017). Phosphonate biochemistry. *Chem. Rev.* 117, 5704–5783. doi: 10.1021/acs.chemrev.6b00536

Hummerjohann, J., Küttel, E., Quadroni, M., Ragaller, J., Leisinger, T., and Kertes, M. A. (1998). Regulation of the sulfate starvation response in *Pseudomonas aeruginosa*: Role of cysteine biosynthetic intermediates. *Microbiology* 144, 1375–1386. doi: 10.1099/00221287-144-5-1375

Hungria, M., Menna, P., and Delamuta, J. R. M. (2015). "*Bradyrhizobium*, the ancestor of all rhizobia: Phylogeny of housekeeping and nitrogen-fixation genes," in *Biological Nitrogen Fixation*, ed. F. Bruijn (New Jersey, NJ: Wiley, Sons, Inc), 191–202. doi: 10.1002/9781119053095.ch18

Igarashi, N., Harada, J., Nagashima, S., Matsuura, K., Shimada, K., and Nagashima, K. V. (2001). Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple bacteria. *J. Mol. Evol.* 52, 333–341. doi: 10.1007/s002390010163

Jensen, E. S., and Hauggaard-Nielsen, H. (2003). How can increased use of biological N_2 fixation in agriculture benefit the environment? *Plant Soil* 252, 177–186. doi: 10.1023/A:1024189029226

Jiang, Y. H., Wang, D. Y., and Wen, J. F. (2012). The independent prokaryotic origins of eukaryotic fructose-1, 6-bisphosphatase and sedoheptulose-1, 7-bisphosphatase and the implications of their origins for the evolution of eukaryotic Calvin cycle. *BMC Evol. Biol.* 12:208. doi: 10.1186/1471-2148-12-208

Jin, C. Z., Jin, L., Liu, M. J., Lee, J. M., Park, D. J., and Kim, C. J. (2022a). *Solihabitans fulvus* gen. nov., sp. nov., a member of the family *Pseudonocardiaceae* isolated from soil. *Int. J. Syst. Evol. Microbiol.* 72:005110. doi: 10.1099/ijsem.0.005110

Jin, C. Z., Wu, X. W., Zhuo, Y., Yang, Y., Li, T. H., and Jin, F. J. (2022b). Genomic insights into a free-living, nitrogen-fixing but non nodulating novel species of *Bradyrhizobium sediminis* from freshwater sediment: Three isolates with the smallest genome within the genus *Bradyrhizobium*. *Syst. Appl. Microbiol.* 45:126353. doi: 10.1016/j.syapm.2022.126353

Jin, L., Jin, C. Z., Lee, H. G., and Lee, C. S. (2021). Genomic insights into denitrifying methane-oxidizing bacteria *Gemmobacter fulva* sp. nov., isolated from an *Anabaena* culture. *Microorganisms.*9:2423. doi: 10.3390/microorganisms9122423

Jin, L., Ko, S. R., Jin, C. Z., Jin, F. J., Li, T. H., and Ahn, C. Y. (2019). Description of novel members of the family *Sphingomonadaceae*: *Aquisediminimonas profunda* gen. nov., sp. nov., and *Aquisediminimonas sediminicola* sp. nov., isolated from freshwater sediment. *Int. J. Syst. Evol. Microbiol.* 69, 2179–2186. doi: 10.1099/ijsem.0.003347

Jones, F. P., Clark, I. M., King, R., Shaw, L. J., Woodward, M. J., and Hirsch, P. R. (2016). Novel European free-living, non-diazotrophic *Bradyrhizobium* isolates from contrasting soils that lack nodulation and nitrogen fixation genes - a genome comparison. *Sci. Rep.* 6:25858. doi: 10.1038/srep25858

Jordan, D. C. (1982). NOTES: Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing, root nodule bacteria from leguminous plants. *Int. J. Syst. Bacteriol.* 32, 136–139. doi: 10.1099/00207713-32-1-136

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., and Sasamoto, S. (2002). Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res.* 9, 189–197. doi: 10.1093/dnares/9.6.189

Kertesz, M. A. (2000). Riding the sulfur cycle–metabolism of sulfonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiol. Rev.* 24, 135–175. doi: 10.1016/S0168-6445(99)00033-9

Klepa, M. S., Helene, L. C. F., O´ Hara, G., and Hungria, M. (2022). *Bradyrhizobium cenepequi* sp. nov., *Bradyrhizobium semiaridum* sp. nov., *Bradyrhizobium hereditatis* sp. nov. and *Bradyrhizobium australafricanum* sp. nov., symbionts of different leguminous plants of Western Australia and South Africa and definition of three novel symbiovars. *Int. J. Syst. Evol. Microbiol.* 72:005446. doi: 10.1099/ijsem.0.005446

Konstantinidis, K. T., Rosselló-Móra, R., and Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J.* 11, 2399–2406. doi: 10.1038/ismej.2017.113

Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 102, 2567–2572. doi: 10.1073/pnas.0409727102

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Ladha, J. K., and So, R. B. (1994). Numerical taxonomy of *photosynthetic rhizobia* nodulating *Aeschynomene* species. *Int. J. Syst. Evol. Microbiol.* 44, 62–73. doi: 10.1099/00207713-44-1-62

Lafay, B., Coquery, E., and Oger, P. M. (2023). *Bradyrhizobium commune* sp. nov., isolated from nodules of a wide range of native legumes across the Australian continent. *Int. J. Syst. Evol. Microbiol.* 73:005971. doi: 10.1099/ijsem.0.005971

Li, S. H., Guo, W., and Mitchell, B. (2011). Evaluation of water quality and management of Hongze Lake and Gaoyou Lake along the Grand Canal in Eastern China. *Environ. Monit. Assess.* 176, 373–384. doi: 10.1007/s10661-010-1590-5

Li, T. H., Zhuo, Y., Jin, C. Z., Wu, X. W., Ko, S. R., Jin, F. J., et al. (2020). Genomic insights into a novel species *Rhodoferax aquaticus* sp. nov., isolated from freshwater. *Int. J. Syst. Evol. Microbiol.* 70, 4653–4660. doi: 10.1099/ijsem.0.004325

Lindström, K., Murwira, M., Willems, A., and Altier, N. (2010). The biodiversity of beneficial microbe-host mutualism: The case of rhizobia. *Res. Microbiol.* 161, 453–463. doi: 10.1016/j.resmic.2010.05.005

Luo, C., Rodriguez-R, L. M., and Konstantinidis, K. T. (2014). MyTaxa: An advanced taxonomic classifier for genomic and metagenomic sequences. *Nucleic Acids Res.* 42:e73. doi: 10.1093/nar/gku169

Luo, F., Gitiafroz, R., Devine, C. E., Gong, Y., Hug, L. A., Raskin, L., et al. (2014). Metatranscriptome of an anaerobic benzene-degrading, nitrate reducing enrichment culture reveals involvement of carboxylation in benzene ring activation. *Appl. Environ. Microb.* 80, 4095–4107. doi: 10.1128/AEM.00717-14

Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., and Willems, A. (2008). Advantages of multilocus sequence analysis for taxonomic studies: A case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *Int. J. Syst. Evol. Microbiol.* 58, 200–214. doi: 10.1099/ijs.0.65392-0

Meena, R. S., Das, A., Yadav, G. S., and Lal, R. (eds) (2018). *Legumes for Soil Health and Sustainable Management*. Singapore: Springer. doi: 10.1007/978-981-13-0253-4_13

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. doi: 10.1186/1471-2105-14-60

Meilhoc, E., Cam, Y., Skapski, A., and Bruand, C. (2010). The response to nitric oxide of the nitrogen-fixing symbiont *Sinorhizobium meliloti*. *Mol. Plant. Microbe Interact.* 23, 748–759. doi: 10.1094/MPMI-23-6-0748

Menna, P., Barcellos, F. G., and Hungria, M. (2009). Phylogeny and taxonomy of a diverse collection of *Bradyrhizobium* strains based on multilocus sequence analysis of the 16S rRNA gene, ITS region and *glnII*, *recA*, *atpD* and *dnaK* genes. *Int. J. Syst. Evol. Microbiol.* 59, 2934–2950. doi: 10.1099/ijs.0.009779-0

Metcalf, W. W., and Van Der Donk, W. A. (2009). Biosynthesis of phosphonic and phosphinic acid natural products. *Annu. Rev. Biochem.* 78, 65–94. doi: 10.1146/annurev.biochem.78.091707.100215

Moro, H., Park, H. D., and Kunito, T. (2021). Organic phosphorus substantially contributes to crop plant nutrition in soils with low phosphorus availability. *Agronomy* 11:903. doi: 10.3390/agronomy11050903

Nath, D., Maurya, B. R., and Meena, V. S. (2017). Documentation of five potassium- and phosphorus-solubilizing bacteria for their K and P-solubilization ability from various minerals. *Biocatal. Agric. Biotechnol.* 10, 174–181. doi: 10.1016/j.bcab.2017.03.007

Nordling, M., Young, S., Karlsson, B. G., and Lundberg, L. G. (1990). The structural gene for cytochrome c551 from *Pseudomonas aeruginosa*. The nucleotide sequence shows a location downstream of the nitrite reductase gene. *FEBS Lett.* 259, 230–232. doi: 10.1016/0014-5793(90)80015-B

Okubo, T., Tsukui, T., Maita, H., Okamoto, S., Oshima, K., Fujisawa, T., et al. (2012). Complete genome sequence of *Bradyrhizobium* sp. S23321: Insights into symbiosis evolution in soil oligotrophs. *Microbes Environ.* 27, 306–315. doi: 10.1264/jsme2.ME1132

Ormeño-Orrillo, E., and Martínez-Romero, E. (2019). A Genomotaxonomy view of the *Bradyrhizobium* Genus. *Front. Microbiol.* 10:1334. doi: 10.3389/fmicb.2019.01334

Parker, M. A. (2015). The spread of *Bradyrhizobium lineages* across host legume clades: From *Abarema* to *Zygia*. *Microb. Ecol.* 69, 630–640. doi: 10.1007/s00248-014-0503-5

Pearson, I. V., Page, M. D., van Spanning, R. J., and Ferguson, S. J. (2003). A mutant of *Paracoccus denitrificans* with disrupted genes coding for cytochrome c_{550} and pseudoazurin establishes these two proteins as the in vivo electron donors to cytochrome cd_1 nitrite reductase. *J. Bacteriol.* 185, 6308–6315. doi: 10.1128/JB.185.21.6308-6315.2003

Pérez-Yépez, J., Armas-Capote, N., Velázquez, E., Pérez-Galdona, R., Rivas, R., and León-Barrios, M. (2014). Evaluation of seven housekeeping genes for multilocus sequence analysis of the genus *Mesorhizobium*: Resolving the taxonomic affiliation of the *Cicer canariense* rhizobia. *Syst. Appl. Microbiol.* 37, 553–559. doi: 10.1016/j.syapm.2014.10.003

Quentmeier, A., Hellwig, P., Bardischewsky, F., Wichmann, R., and Friedrich, C. G. (2004). Sulfide dehydrogenase activity of the monomeric flavoprotein SoxF of *Paracoccus pantotrophust*. *Biochemistry* 43, 14696–14703. doi: 10.1021/bi048568y

Remigi, P., Zhu, J., Young, J. P. W., and Masson-Boivin, C. (2016). Symbiosis within symbiosis: Evolving nitrogen-fixing legume symbionts. *Trends Microbiol.* 24, 63–75. doi: 10.1016/j.tim.2015.10.007

Richardson, D. J. (2000). Bacterial respiration: A flexible process for a changing environment. *Microbiology* 146, 551–571. doi: 10.1099/00221287-146-3-551

Richter, M., and Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106

Rinaldo, S., and Cutruzzola, F. (2007). Nitrite reductases in denitrification. *Biol. Nitrogen Cycle* 37, 37–55. doi: 10.1016/B978-044452857-5.50004-7

Rivas, R., Martens, M., de Lajudie, P., and Willems, A. (2009). Multilocus sequence analysis of the genus *Bradyrhizobium*. *Syst. Appl. Microbiol.* 32, 101–110. doi: 10.1016/j.syapm.2008.12.005

Rott, E., Steinmetz, H., and Metzger, J. W. (2018). Organophosphonates: A review on environmental relevance, biodegradability and removal in wastewater treatment plants. *Sci. Total Environ.* 615, 1176–1191. doi: 10.1016/j.scitotenv.2017.09.223

Salminen, S. O., and Streeter, J. G. (1987). Uptake and metabolism of carbohydrates by *Bradyrhizobium japonicum* bacteroids. *Plant Physiol.* 83, 535–540. doi: 10.1104/pp.83.3.535

Simon, J., and Klotz, M. G. (2013). Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochim. Biophys. Acta* 1827, 114–135. doi: 10.1016/j.bbabio.2012.07.005

Sosa, O. A., Repeta, D. J., DeLong, E. F., Ashkezari, M. D., and Karl, D. M. (2019). Phosphate-limited ocean regions select for bacterial populations enriched in the carbon-phosphorus lyase pathway for phosphonate degradation. *Environ. Microbiol.* 21, 2402–2414. doi: 10.1111/1462-2920.14628

Sprent, J. I., Ardley, J., and James, E. K. (2017). Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol.* 215, 40–56. doi: 10.1111/nph.14474

Stacey, G., Sanjuan, J., Luka, S., Dockendorff, T., and Carlson, R. W. (1995). Signal exchange in the *Bradyrhizobium*–soybean symbiosis. *Soil Biol. Biochem.* 27, 473–483. doi: 10.1016/0038-0717(95)98622-u

Tang, J., Bromfield, E. S. P., Rodrigue, N., Cloutier, S., and Tambong, J. T. (2012). Microevolution of symbiotic *Bradyrhizobium* populations associated with soybeans in east North America. *Ecol. Evol.* 2, 2943–2961. doi: 10.1002/ece3.404

Tao, J., Wang, S., Liao, T., and Luo, H. (2021). Evolutionary origin and ecological implication of a unique nif island in free-living *Bradyrhizobium* lineages. *ISME J.* 15, 3195–3206. doi: 10.1038/s41396-021-01002-z

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., et al. (2003). The COG database: An updated version includes eukaryotes. *BMC Bioinformatics* 4:41. doi: 10.1186/1471-2105-4-41

Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997). A genomic perspective on protein families. *Science* 278, 631–637. doi: 10.1126/science.278.5338.631

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The Clustal X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 24, 4876–4882. doi: 10.1093/nar/25.24.4876

Van Der Ploeg, J. R., Barone, M., and Leisinger, T. (2001a). Expression of the *Bacillus subtilis* sulphonate–sulphur utilization genes is regulated at the levels of transcription initiation and termination. *Mol. Microbiol.* 39, 1356–1365. doi: 10.1046/j.1365-2958.2001.02327.x

Van Der Ploeg, J. R., Eichhorn, E., and Leisinger, T. (2001b). Sulfonate-sulfur metabolism and its regulation in *Escherichia coli*. *Arch. Microbiol.* 176, 1–8. doi: 10.1007/s002030100298

Van Rhijn, P., and Vanderleyden, J. (1995). The Rhizobium-plant symbiosis. *Microbiol. Rev.* 59, 124–142. doi: 10.1128/mr.59.1.124-142.1995

VanInsberghe, D., Maas, K. R., Cardenas, E., Strachan, C. R., Hallam, S. J., and Mohn, W. W. (2015). Non-symbiotic *Bradyrhizobium* ecotypes dominate North American forest soils. *ISME J.* 9, 2435–2441. doi: 10.1038/ismej.2015.54

Villarreal-Chiu, J. F., Quinn, J. P., and McGrath, J. W. (2012). The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. *Front. Microbiol.* 3:19. doi: 10.3389/fmicb.2012.00019

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427

Weisburg, W. G., Barns, S. M., Pelletier, D. A., and Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* 173, 697–703. doi: 10.1128/jb.173.2.697-703.1991

Wu, Y. S., Huang, T. Y., Zhang, J. G., Tian, Y. J., Pang, Y., and Xu, Q. J. (2023). Distribution characteristics and risk assessment of PPCPs in surface water and sediments of lakes in the lower reaches of the Huaihe River. *Huan Jing Ke Xue* 44, 3217–3227. doi: 10.13227/j.hjkx.202207169

Yao, Y., Wang, P., Wang, C., Hou, J., Miao, L., Yuan, Y., et al. (2016). Assessment of mobilization of labile phosphorus and iron across sediment-water interface in a shallow lake (Hongze) based on in situ high-resolution measurement. *Environ. Pollut.* 219, 873–882. doi: 10.1016/j.envpol.2016.08.054

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., et al. (2017). Introducing EzBio-Cloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* 67, 1613–1617. doi: 10.1099/ijsem.0.001755

Zhang, J., Wang, N., Li, S., Peng, S., Andrews, M., Zhang, X., et al. (2023). *Bradyrhizobium zhengyangense* sp. nov., isype strains of the most closely related species of isolated from effective nodules of *Arachis hypogaea* L. in central China. *Int. J. Syst. Evol. Microbiol* 73:005723. doi: 10.1099/ijsem.0.005723

Zhang, Z., Guo, H., Sun, J., and Wang, H. (2020). Investigation of anaerobic phenanthrene biodegradation by a highly enriched co-culture, PheN9, with nitrate as an electron acceptor. *J. Hazard Mater.* 383:121191. doi: 10.1016/j.jhazmat.2019.121191

A novel endophytic fungus strain of *Cladosporium*: its identification, genomic analysis, and effects on plant growth

Nan Yang[1†], Wenbin Zhang[1†], Dan Wang[1], Dingding Cao[1], Yanyu Cao[1], Weihong He[1], Ziting Lin[1], Xiaofeng Chen[1], Guiping Ye[1], Zhiming Chen[2], Jianjun Chen[3] and Xiangying Wei[1*]

[1]Fujian Key Laboratory on Conservation and Sustainable Utilization of Marine Biodiversity, Fuzhou Institute of Oceanography, College of Geography and Oceanography, Minjiang University, Fuzhou, Fujian Province, China, [2]Institute of Cytology and Genetics, Hengyang Medical School, University of South China, Hengyang, Hunan, China, [3]Mid-Florida Research and Education Center, Department of Environmental Horticulture, Institute of Food and Agricultural Sciences, University of Florida, Apopka, FL, United States

Introduction: Endophytic microorganisms are bacteria or fungi that inhabit plant internal tissues contributing to various biological processes of plants. Some endophytic microbes can promote plant growth, which are known as plant growth-promoting endophytes (PGPEs). There has been an increasing interest in isolation and identification of PGPEs for sustainable production of crops. This study was undertaken to isolate PGPEs from roots of a halophytic species *Sesuvium portulacastrum* L. and elucidate potential mechanisms underlying the plant growth promoting effect.

Methods: Surface-disinfected seeds of *S. portulacastrum* were germinated on an *in vitro* culture medium, and roots of some germinated seedlings were contaminated by bacteria and fungi. From the contamination, an endophytic fungus called BF-F (a fungal strain isolated from bacterial and fungal contamination) was isolated and identified. The genome of BF-F strain was sequenced, its genome structure and function were analyzed using various bioinformatics software. Additionally, the effect of BF-F on plant growth promotion were investigated by gene cluster analyses.

Results: Based on the sequence homology (99%) and phylogenetic analysis, BF-F is likely a new *Cladosporium angulosum* strain or possibly a new *Cladosporium* species that is most homologous to *C. angulosum*. The BF-F significantly promoted the growth of dicot *S. portulacastrum* and Arabidopsis as well as monocot rice. Whole genome analysis revealed that the BF-F genome has 29,444,740 bp in size with 6,426 annotated genes, including gene clusters associated with the tryptophan synthesis and metabolism pathway, sterol synthesis pathway, and nitrogen metabolism pathway. BF-F produced indole-3-acetic acid (IAA) and also induced the expression of plant N uptake related genes.

Discussion: Our results suggest that BF-F is a novel strain of *Cladosporium* and has potential to be a microbial fertilizer for sustainable production of crop plants. The resulting genomic information will facilitate further investigation of its genetic evolution and its function, particularly mechanisms underlying plant growth promotion.

1 Introduction

Many microorganisms, particularly fungi and bacteria that can colonize intercellular spaces, epidermis, endodermis, and vascular bundles of host plant tissues or inside the cells for their entire or partial life cycle without causing damage to the host plant. These microbes are defined as endophytes (Saikkonen et al., 2003; Perotto et al., 2012). Endophytes are considered as a symbiotic association with host plants. Plants from nearly all the genera of the Kingdom Plantae harbor fungal endophytes within their internal tissues. Plants with endophytic fungi have been shown to be more resistant to pathogens, better utilize resources, and exhibit increased productivity (Baron and Rigobelo, 2021). Endophytic fungi can protect plants from pathogen infections by either direct inhibition through competition, antibiosis or mycoparasitism or indirect inhibition through enhancing resistance (Latz et al., 2018; Silva et al., 2018; Ebrahimi et al., 2022; Wei et al., 2022). Many protective secondary metabolites have been identified from endophytic fungi, such as flavonoids, terpenoids, alkaloids, phenols, volatile organic compounds (VOC). Some bioactive compounds identified from endophytic fungi can be used for drug discovery and agricultural applications for controlling pathogens and insect pest (Zhang et al., 2006; Meenavalli et al., 2011; Gouda et al., 2016; Toghueo et al., 2017; Strobel, 2018). In addition, some endophytic fungi inside host plants are capable of degrading a portion of plant lignin and cellulose which helps the host plant's defense against invasive pathogens (Marques et al., 2018; Li et al., 2021) or secrete extracellular chitinase to decompose chitin and damage the cell wall structure of phytopathogenic fungi (Hartl et al., 2012; Kumar et al., 2018).

Endophytes also contribute to multiple bioprocesses of plant growth and development, including nitrogen fixation, phosphate solubilization, and phytohormones production (Santoyo et al., 2016; Priyadharsini and Muthukumar, 2017; Zhang et al., 2018). Many endophytic fungi possess phosphate solubilizing activity, such as *Pestalotiopsis*, *Trichoderma, Penicillium,* and *Aspergillus* species (Zhang et al., 2016; Hossain et al., 2017; Adhikari and Pandey, 2019; Erman et al., 2022). The solubilization of inorganic phosphate by endophytes has been shown to contribute significantly to the bioavailability of phosphorus (Adhikari and Pandey, 2019). Phytohormones are important plant signaling compounds regulating plant growth and development as well as adaptation to stressful conditions. Many endophytic microbes can synthesize indole-3-acetic acid (IAA), gibberellins (GAs), or some cytokinin compounds (Hamayun et al., 2009; Aly et al., 2010; Sasirekha et al., 2012; You et al., 2013; Khan et al., 2014, 2015; Wei et al., 2020). In the plant growth-promoting endophytic fungi group, some can secrete GAs or IAA to regulate host plant growth and developmental (Hamayun et al., 2009; Shah et al., 2018). Since endophytic fungi not only act as biocontrol agents but as biostimulants and biofertilizers, a considerable effort has been placed on the isolation of endophytic fungi from different plants and identification of their role in improving plant growth and/or adaptation to abiotic and biotic stresses.

Cladosporium fungi belong to the order Capnodiales in the class Dothideomycete and a member of the dematiaceous hyphomycetes (Dugan et al., 2004; Bensch et al., 2018). *Cladosporium* is commonly found in plant, fungal, and other organic debris, can be isolated from soil, air, food, paint, textiles, and diverse types of organic materials (Farr et al., 1989; Flannigan, 2001; Mullins, 2001; Sandoval-Denis et al., 2015). In plants, *Cladosporium* species are commonly known as endophytes (El-Morsy, 2000; Hamayun et al., 2009; Venkateswarulu et al., 2018) and phylloplane fungi (Levetin and Dorseys, 2006). However, the effects of *Cladosporium* on plants are diverse. Some *Cladosporium* species are pathogens that cause plant disease (Heuchert et al., 2005; Watson and Napier, 2009; Kryczyński and Weber, 2011; Ogórek et al., 2012), and some other species are beneficial to plants by promoting plant growth (Hamayun et al., 2009; Paul and Park, 2013; Sandoval-Denis et al., 2015) or enhancing plant tolerance to biotic and abiotic stresses (Chaibub et al., 2016; Torres et al., 2017).

The beneficial effects of *Cladosporium* on plants include the release of some secondary metabolites for improving plant growth (Hamayun et al., 2009; Paul and Park, 2013). Thus, the analysis of the metabolic pathways of *Cladosporium* species will help define species functions within *Cladosporium*. Advances in sequencing technology have accelerated fungal genomic studies, and new third-generation sequencing techniques allow researchers to describe the molecular diversity of more closely related strains (Xu et al., 2021). Clusters of genes for secondary metabolic pathways are an emerging theme in plant biology and microbiology, which provides some provocative insights into genome plasticity and evolution (Osbourn, 2010). The potential functions of *Cladosporium* species in plants can be quickly predicted based on the gene clusters of secondary metabolic pathways obtained from the genome.

Halophytic *Sesuvium portulacastrum* L. is a pioneer plant species used for sand-dune fixation, desalination, and phytoremediation along coastal regions of Asia, Arabian Peninsula, and South China due to its tolerance to salinity. It is also used as a vegetable, ornamental plant, and fodder for domestic animals (Lokhande et al., 2013). The extracts of *S. portulacastrum* are widely used as essential oils and folk medicine (Lokhande et al., 2013; Patel et al., 2019). There is a speculation that its salt tolerance and medicinal properties may be related to endophytic and rhizosphere microorganisms. A halotolerant PGPR isolated from *S. portulacastrum* was found to improve salt tolerance of *Vigna mungo* L. (John et al., 2023). Endophytic and rhizosphere fungi isolated from *S. portulacastrum* increased cucumber plants tolerance to damping off caused by *Pythium aphanidermatum* (Karunasinghe et al., 2020). Therefore, application of endogenous microorganisms isolated from halophytes, such as *S. portulacastrum* could be a strategy to improve crop plants to tolerate salinity and increase crop productivity.

We recently isolated an endophytic fungus from roots of *in vitro* cultured *S. portulacastrum* seedlings, which was able to substantially enhance plant growth. The objectives of this study were to identify and characterize this isolate, evaluate its effects on plant growth, perform

genomic sequencing, analyze hormones in mycelium, and determine likely mechanisms underlying the plant growth promoting effects.

2 Materials and methods

2.1 Isolation of an endophytic fungus from *Sesuvium portulacastrum*

Seeds of *S. portulacastrum* were disinfected with 75% (v/v) ethanol for 1 min, and then in 3% NaClO (v/v) for 6 min. After rinse with sterile distilled water three times, the seeds were placed on Murashige and Skoog (MS) medium (Murashige and Skoog, 1962) supplemented with 3% (w/v) sucrose and 0.6% (w/v) agar at 25°C for germination. The MS has been the most widely used plant tissue culture medium since 1962.

After germination, we found that roots of few seedlings appeared to be contaminated with fungi and bacteria, and the contaminated seedlings were taller and stronger than those uncontaminated. Using the methods described by Khan et al. (2015) and Wei et al. (2016), a fungus was isolated. Briefly, roots of contaminated seedlings were disinfected as mentioned above and cut into about 0.5 cm pieces and placed on a water agar (WA) medium and incubated in the dark at 25°C until fungal growth started. The fungus was then cultured on malt extract agar (MEA) solid medium (1 g/L tryptone, 20 g/L malt extract, 20 g/L glucose, and 20.0 g/L agar) for morphological identification. The colony characteristics of this fungus were determined after growing on solid potato dextrose agar (PDA) medium (6.0 g/L potato powder +20.0 g/L glucose +20.0 g/L agar; pH = 5.6–5.8) for 7 days in the dark at 25°C (described below). This isolate was stored in the Chinese Center for Microbial Species Preservation and Management (CGMCC NO: 23053) and referred to as BF-F (a fungus isolated from bacterial and fungal contaminated culture) in the library.

2.2 Morphological identification of *Cladosporium* BF-F

Morphological characteristics of BF-F, including colony diameter, color, pigments, and reverse side of colony color, were examined after culturing on PDA for 7 days in the dark at 25°C. Conidia and conidiophores were induced using 20% potato dextrose broth, a small volume (about 50 μL) of 20% potato dextrose broth was dropped into each of two holes of a glass slide placed in a 9 cm plastic Petri dish, hyphae of isolate were added, and then incubated in the dark at 25°C for 7–21 days. Conidia were observed under a light microscope. Hypha was able to separate and branch to 2.0–5.5 μm in width. The fungus was putatively identified with taxonomic keys in Barron (1962) and Domsch et al. (1980).

2.3 Molecular identification of *Cladosporium* BF-F

The isolate of BF-F was cultured in a liquid PDB medium with shaking (120 r/min) at 25°C for 14 days. The culture was collected in sterile 50 mL centrifuge tube by centrifuging at 10,000 × g for 10 min

and then frozen in liquid nitrogen. After grinding the fungi with a mortar and pestle. Genomic DNA was extracted using the SDS method (Lim et al., 2016). The isolated DNA was confirmed by agarose gel electrophoresis and quantified by Qubit® 2.0 Fluorometer. The internal transcribed spacer (ITS) region was amplified using the ITS1 (5' TCCGTAGGTGAACCTGCGG 3') and ITS4 (5' TCCTCCGCTTATTGATATGC 3') primers. The amplifications were performed in a 50 μL reaction volume containing 50 ng of genomic DNA, 50 pmol of each primer, 100 μM of dNTP, 1 U of Taq polymerase, and 5 μL of 10 × PCR buffer. The tubes were incubated at 95°C for 2 min and then subjected to 35 cycles as follows: 94°C for 30 s; 60°C for 30 s, and 72°C for 45 s; a final incubation was carried out for another 5 min at 72°C. The PCR products were cloned with the PMD19-T Easy Vector System and analyzed by an ABI 3730XI automatic DNA Sequencer. The molecular identification was carried out using the ITS region and *EF-1α* sequence (Supplementary Table S1). The isolate was subjected to DNA sequencing and ITS-based phylogenetic analysis as previously described with slight modification (Yew et al., 2014). The BLAST search was conducted for the maximum percentage of sequence homology and query coverage as well as the lowest E values. The rDNA ITS sequences of species in *Cladosporium* were searched for in the UNITE database (Nilsson et al., 2019), and the search parameters used were as follows: "Threshold" was "1.5″; "include" was "All SH-s." The complete ITS1-5.8S-ITS4 and *EF-1α* sequences for 12 *Cladosporium* species were obtained from GenBank for phylogenetic tree construction with BF-F. Neighbor-joining tree analysis was performed using Molecular Evolutionary Genetics Analysis (MEGA) software with 1,000 bootstrap replications.

2.4 Effects of *Cladosporium* BF-F on plant growth

The effect of BF-F on plant growth was determine on *S. portulacastrum, Arabidopsis thaliana*, and rice (Tp309), respectively. Seeds of the three plants were disinfected in 75% (v/v) ethanol for 1 min and then in 3% NaClO (v/v) for 6 min. After rinse with sterile distilled water three times, the *S. portulacastrum* seeds were placed on MS medium at 25°C for germination. The germination took place in a growth room under white-fluorescent light with 16 h light and 8 h dark for 14–16 days. Seedlings were inoculated with a diameter 0.5 cm BF-F fungus colony as described by Wei et al. (2016). Plant height, root number, and leaf number of the *S. portulacastrum* seedlings were recorded after 1–4 weeks of co-cultivation. The same procedure was used for germination of *A. thaliana* and rice seeds. BF-F effects on *Arabidopsis* and rice seedling growths were evaluated after 14 and 28 days of inoculation, respectively.

2.5 Genome sequencing and assembly

The genome of BF-F strain was sequenced and assembled by Beijing Novogene Technology Co., Ltd. (Beijing, China) using the Illumina Novaseq 6,000 and PacBio Sequel I high-throughput sequencing platform (Beijing Nuohe Zhiyuan Technology Co., Ltd., Beijing, China). The obtained raw data were filtered to obtain high-quality DNA sequence data and then processed in accordance with the platform gene assembly process. The third-generation sequencing

data generated by PacBio Sequel I Sequencing, SMRT Link v5.0.1 software was used to assemble the reads to obtain preliminary assembly results that reflected the basic situation of the sample genome, and the WTDBG software assembly results were selected for subsequent processing. PacBio Sequel I Arrow software (Chaisson and Tesler, 2012) was then used to perform self-single base correction. With the use of the variant Caller module of SMRT Link software, the arrow algorithm was used to correct and count the variant sites in the preliminary assembly results, and Pilon software (Walker et al., 2014) was used to correct the sequence assembled by the third-generation sequencing and to finally obtain the corrected genome. The sequence was submitted to the NCBI database under the accession number JAMRMS000000000.

2.6 Functional genome annotation

The Gene Ontology Resource was used for Gene Ontology (GO),[1] analysis (Ashburner et al., 2000). The Kyoto Encyclopedia of Genes and Genomes (KEGG) database[2] was employed for the construction of metabolic pathways (Kanehisa et al., 2012). The NCBI Clusters of Orthologous Groups (COG) server[3] was used to assign the obtained COG annotations for Eukaryotic Orthologous Groups (KOG) to functional categories (Koonin et al., 2004). Genes involved in plant growth promotion were excavated from the predicted annotated genes of the Cladosporium sp. BF-F genome.

2.7 Analysis of tryptophan and tryptophan metabolites in mycelium of BF-F

Tryptophan (Trp) and tryptophan metabolites were analyzed using mass spectrometry based on the method described by Krause et al. (2015). A diameter 0.5 cm BF-F colony was cultured in the flask containing potato dextrose broth medium on a rotating shaker at 150 rpm, 25°C in total darkness for 5 days. Mycelium was collected by filtration through Whatman filter paper No. 1 and frozen in liquid N (Wei et al., 2020). For analysis of tryptophan and tryptophan metabolites, 1 g mycelium was extracted by acetonitrile or methanol and diluted by diethyl ether to different concentration gradients. Tryptophan and tryptophan metabolites including indole-3-acetic acid, indole-3-butyric acid were analyzed using the Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer by Sci-Tech Inovation, Inc. (Qindao, China). CSH C18 RP column (150×2.1 mm, particle size of 1.7 μm) was used, and the analysis conditions were set as follows: cone/desolvation gas flow at 150/1000 L h^{-1}; capillary voltage at 2.1 kV for ESI(+) and 1.5 kV for ESI(−); source/desolvation temperature at 125°C/600°C; collision energy at 12 to 30 eV; cone voltage at 10 to 40 V; and collision argon gas flow at 0.21 mL min^{-1}.

2.8 Qrt-PCR analysis of genes related to auxin response and N absorption

Expression of key genes related to auxin biosynthesis and N uptake were investigated in *A. thaliana* seedlings with and without BF-F inoculated. Seedlings of *A. thaliana* were harvested after 2 weeks inoculation, and total RNAs were extracted using E.Z.N.A Plant RNA Kit (Omega Bio-tek, Norcross, GA, United States). The first stand cDNA was synthesized using the PrimeScript first cDNA Synthesis Kit (Takara, Dalian, China). qRT-PCR was used to analyze the expression level of auxin response genes, including *PRE1, IAA19, SAUR-AC1, GH3.3, ARF9*, and *PIN7* as well as nitrate transporter genes *NRT1.1, NRT1.7*, and *NRT2.7*. *PP2A* was used as an internal control. For a given gene, the relative expression level was performed as mean ± SD with three replicates. All primers used are listed in Supplementary Table S2.

3 Results

3.1 Morphological characteristics of BF-F strain

The morphology of the BF-F colony exhibited grayish-green, radial wrinkles, which was velvety in texture (Figure 1A), indigo on the back, and produced no soluble pigment (Figure 1B) when cultured on MEA medium in the dark at 25°C for 7 days. The hyphae of BF-F were loose separation, concave performance, and dark brown in color (Figure 1C). Conidia were subglobose to cylindrical with light brown color under the microscope when cultured on PDA medium (Figure 1C). The color became darker with aging; conidial peduncles occurred in aerial hyphae. Their stems were divided, upright or slightly curved. Conidia were oval, $2.5–5.5 \times 2–4$ μm in size. Branch conidia occasionally formed, long column or oval with a size of 6–12 μm in length and 2.5–4.5 μm in width.

3.2 Molecular characterization

Phylogenetic analysis of BF-F based on the ITS rDNA and *EF-1α* sequences is presented in Supplementary Table S1. The *EF-1α* and ITS rDNA sequences of BF-F were compared to the available sequences obtained by BLAST from GenBank database. Neighbor-joining trees were constructed using 13 taxa (12 references and one clone) with 1,000 bootstrap replications using MEGA software, respectively (Figure 2 and Supplementary Figure S1). The selection of these strains was based on the BLAST search as they showed the maximum percentage of sequence homology and query coverage as well as the lowest E values. In the dendrogram of the ITS rDNA sequence, BF-F fell into *Cladosporium* sp. group with strong support (Supplementary Figure S1). However, BF-F formed a monoclade with *C. angulosum* CBS140692 (type strain), which had 88% bootstrap support in the phylogenetic tree of the *EF-1α* sequence (Figure 2). Based on the sequence homology (99%) and phylogenetic analysis, BF-F may be delimited as a new strain of *C. angulosum* or a *Cladosporium* sp. homologous to *C. angulosum*.
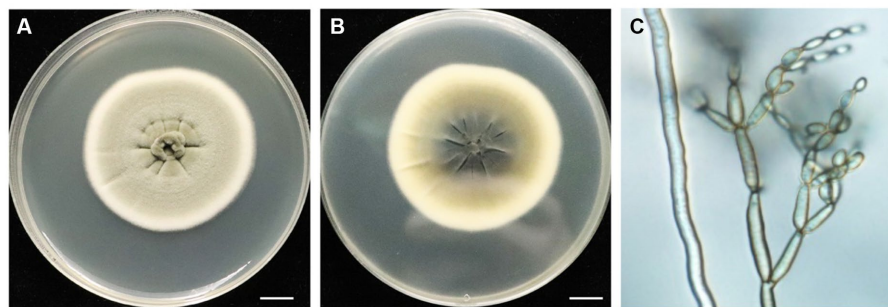
FIGURE 1
Morphological characteristics of endophytic fungi BF-F. **(A)** and **(B)** Colonial morphology of BF-F grown on PDA for 7 days at 25°C. A, front; B, back. **(C)** Light micrograph showing conidia morphology of BF-F grown on PDA for 7 days at 25°C.
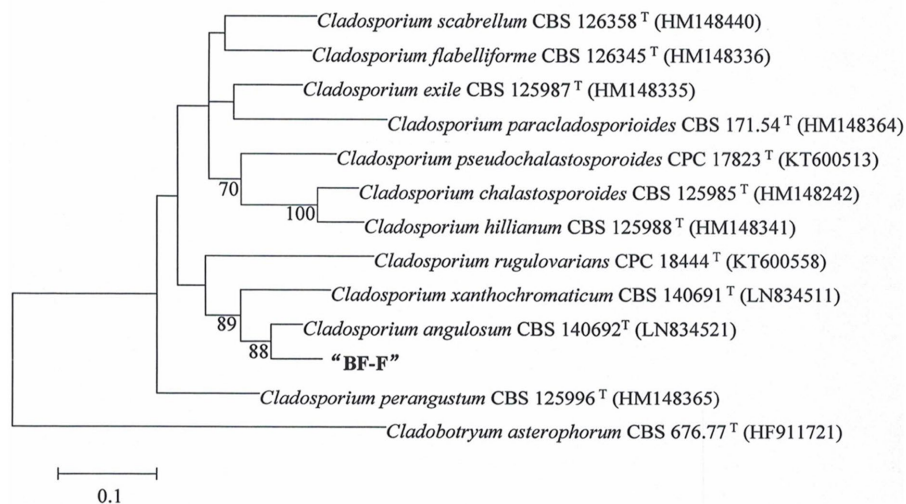

FIGURE 2
Phylogenetic tree analysis of BF-F based on *EF-1α* sequences. Tree topology of BF-F strain based on *EF-1α* sequences using the Neighbour-Joining (N-J) method. The 12 reference sequences were obtained from GenBank. Bootstrap values (calculated from 1000 repetitions) ≥ 70% are shown at their relevant nodes and the superscript "T" represented the model strains.

3.3 Effects of *Cladosporium* BF-F on plant growth

Although some *Cladosporium* sp. are known to promote plant growth (El-Morsy, 2000; Hamayun et al., 2009; Paul and Park, 2013) and enhance plant stress tolerance (Chaibub et al., 2016; Torres et al., 2017), little information is available on *C. angulosum* effects on plant growth. After the inoculation with BF-F, the growth of *S. portulacastrum* seedlings were significantly larger (Figure 3A) and roots were more abundant (Figure 3B) than control. Plant height, root numbers, and leaf numbers were substantially greater than those of plants without BF-F inoculation, and the growth-promoting effects became increasingly different with plant growth time (Figures 3C–E). It was unknown if there was a dose dependent effect on plant growth because BF-F inoculant used in this study was a piece of BF-F fungus colony (0.5 cm). Further study regarding this is warranted.

In order to be certain that the growth promotion was not due to specific plant species, BF-F was first inoculated in seedlings of dicot model plant *A. thaliana*. It was found that BF-F also enhanced the growth of *A. thaliana* seedlings (Figure 4A). *A. thaliana* seedling sizes significantly increased (Figure 4B), and the fresh weight of *A. thaliana* seedlings with BF-F inoculation was more than four times greater than that of the control group (Figure 4C). Root elongation appeared to be inhibited by BF-F, but root hairs were extremely abundant with a bushy root system (Figure 4A). Growth promoting effects of BF-F were further evaluated using monocot rice. Results showed that rice seedlings inoculated with BF-F grew more vigorously (Figure 4D) and had more abundant roots than the control (Figure 4E).

3.4 General features of the BF-F genome

In order to understand the molecular mechanism behind the plant growth promoting effects, the genome of BF-F (GenBank accession number JAMRMS000000000) was sequenced and assembled using Illumina Novaseq 6,000 and PacBio Sequel I. The complete
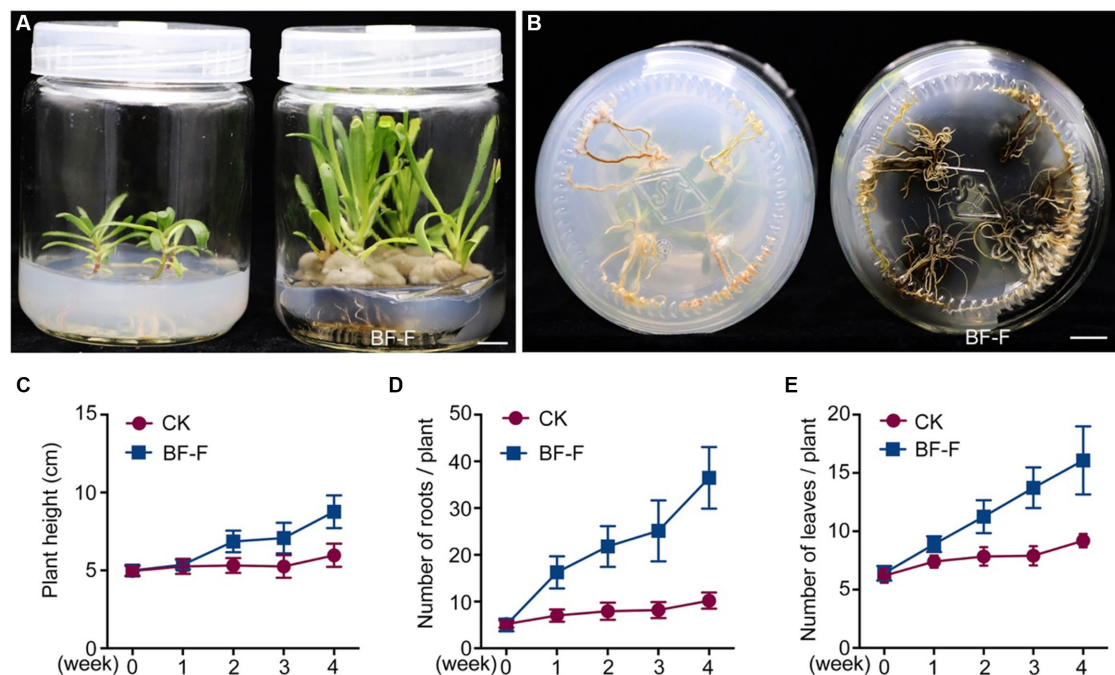
FIGURE 3
Growth promoting effects of BF-F on *Sesuvium portulacastrum* seedlings. **(A)** Shoot phenotypes of *S. portulacastrum* seedlings after 4 weeks of growth in MS medium uninoculated (left) and inoculated (right) with BF-F. **(B)** Root phenotypes of *S. portulacastrum* seedlings after 4 weeks of growth in MS medium uninoculated (left) and inoculated with BF-F. **(C−E)** Plant height, leaf and root numbers of *S. portulacastrum* seedlings grown in MS medium uninoculated and inoculated with BF-F for 4 weeks. Data represent means ± SD with three biological replicates, each had at least 20 seedlings.

genome of BF-F consisted of a total of 29.4 Mb with a GC content of 52.57%, including 18 contigs (N50 size was 1.75 Mb) (Figure 5). The general genome features are shown in Table 1, of which 6,426 genes were identified as protein-coding genes, with an average length of 1,314 bp. In addition to the predicted genes, a total of 214 tRNA, two sRNA, 18 snoRNA, and 59 rRNA loci (5S, 18S, and 28S) were predicted in the BF-F genome sequence (Table 1).

3.5 Gene function annotation

The 6,426 predicted genes in the BF-F genome were mapped to the KOG, KEGG, and GO databases (Figure 6). A total of 1,871 predicted proteins were annotated redundantly based on KOG function classification (Figure 6A). Among the highest annotated groups, post-translational modifications of proteins appeared to be significant in cellular regulation, development, and adaptation to stress. Annotation identified 35 putative genes encoding chaperones and 52 genes encoding ubiquitination degradation-related proteins and proteasomes in group O (post-translational modification, protein turnover, and chaperones), which are associated with the regulation of misfolded proteins and metabolic enzymes degraded via the ubiquitin-proteasome system. In group E (amino acid transport and metabolism), tryptophan syntheses and metabolism appeared to be significant. In group G (carbohydrate transport and metabolism), the highest number of genes was related to putative glycolysis/gluconeogenesis.

The KEGG pathway analysis annotated a total of 5,5,572 genes in five metabolic pathways, including amino acid metabolism (348), carbohydrate metabolism (339), overview (335), energy metabolism (195), and metabolism of cofactors and vitamins (195) (Figure 6B). It is not surprising that BF-F contains many genes involved in amino acid metabolism because some amino acids are precursors of plant hormones. For example, IAA is synthesized using tryptophan as the precursor. IAA is a natural auxin and can facilitate the interactions between fungi and other organisms and also promote plant growth (Fu et al., 2015).

According to the GO classification, 4,321 predicted genes received a GO assignment (Figure 6C). A total of 2,267 genes were assigned to the metabolic process category, which may indicate that *Cladosporium* BF-F metabolites are closely associated with plant growth. Therefore, the gene dataset obtained from this study not only enriches the gene database resources but also provides a basis for further investigation of the gene functions of *Cladosporium*.

3.6 Genes associated with plant growth-promoting effects

Genome annotation of BF-F revealed several genes contributing directly or indirectly to the promotion of plant growth (Table 2). In the genome annotation, different gene clusters related to IAA biosynthesis were identified. These included tryptophan cluster genes with strong involvement in the biosynthesis (*trpA*, *trpB*, *trpD*, *trpE*, *TRP1*, and *TRP3*) and metabolism (*TDC*, *MAO*, and *ALDH*)
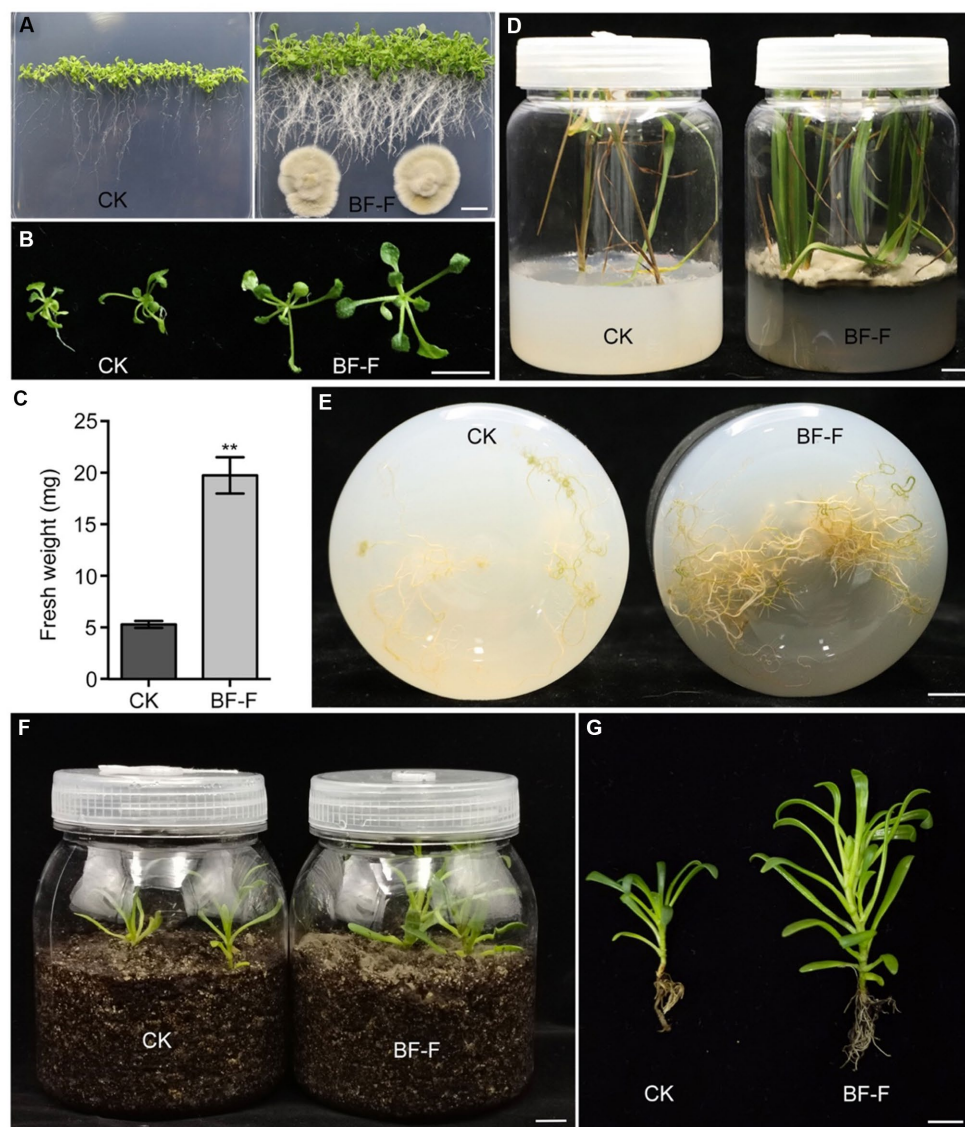
FIGURE 4
Growth promoting effects of BF-F on different plants. **(A)** Overall growth phenotypes and **(B)** shoot phenotypes of Arabidopsis seedlings grown on MS medium without and with BF-F for 2 weeks. **(C)** Shoot fresh weight of Arabidopsis seedlings after 2 weeks of growth without and with BF-F. Data represent means $\pm$ SD of at least 20 seedlings with three biological replicates. **(D)** Growth phenotypes of rice seedlings after uninoculated and inoculated with BF-F for 4 weeks. **(E)** Root phenotypes of rice seedlings after uninoculated and inoculated with BF-F for 4 weeks. **(F–G)** Growth phenotypes of *S. portulacastrum* seedlings growing in soil without and with BF-F for 5 weeks. Bar = 1 cm.

of tryptophan. As mentioned above, tryptophan is the precursor of IAA (Table 2). In addition, genes involved in the production of steroids were clustered, and they are implicated in plant growth promotion (Schena et al., 1991; Feldmann, 2006). For example, genes related to plant hormone brassinosteroid (BR) biosynthesis process (Bhar, 2021) were also identified (Table 2). Furthermore, genes involved in nitrogen metabolism and regulation were found in the BF-F genome, including the nitrogen regulation protein cluster (*NRT*, *glnA*, *gtl1*, and *gdhA*) and the nitrite reductase (*NIT-6*) and nitrate reductase (*NR*) clusters (Table 2). Together, these results may imply that phytohormone production and enhanced nutrient absorption in BF-F may contribute to plant growth promoting effects.

3.7 Hormones and precursors of IAA produced by BF-F

In order to determine whether or not BF-F could produce IAA. Mycelium of BF-F was analyzed for precursors involved in IAA biosynthesis. Results showed that Trp was extremely higher with a mean concentration of 100.12 mg kg^{-1} (Figure 7A). Other compounds were detected with IBA at 8.73 μg kg^{-1} and IAA at 6.87 μg kg^{-1} (Figure 7A). qRT-PCR analysis showed that *PRE1, IAA19, SAUR-AC, ARF9, PIN7,* and *GH3.3*, which are responsible for IAA, were highly upregulated in *A. thaliana* seedlings inoculated with BF-F compared to the uninoculated control plants (Figure 7B), and the upregulation corresponded to enhanced growth of seedlings (Figures 4A,B). These
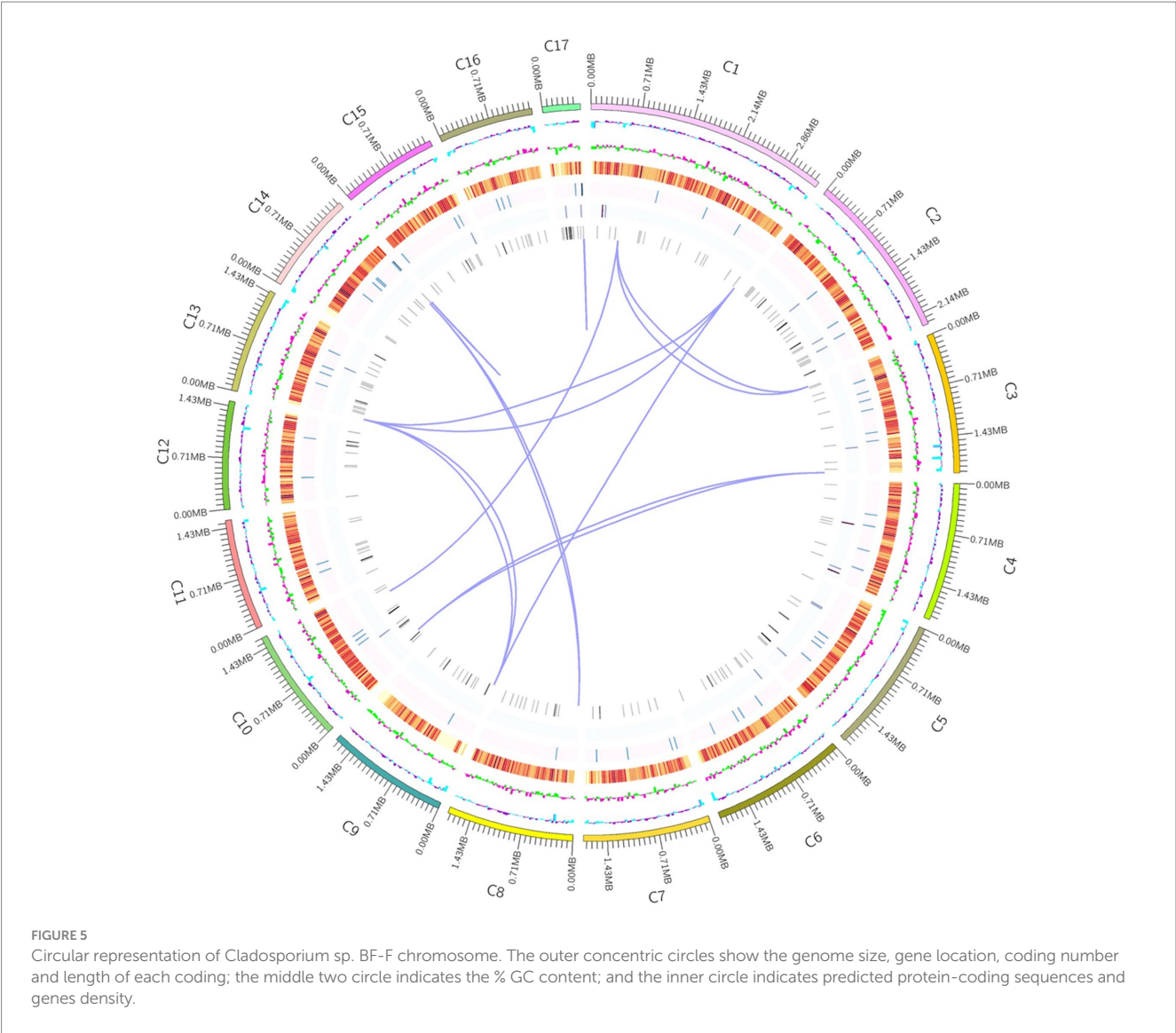
FIGURE 5

Circular representation of Cladosporium sp. BF-F chromosome. The outer concentric circles show the genome size, gene location, coding number and length of each coding; the middle two circle indicates the % GC content; and the inner circle indicates predicted protein-coding sequences and genes density.

TABLE 1 General genome features of BF-F.

Feature	Chromosome
Genome size (bp)	29,444,740
Number of Genes	6,425
Gene length (bp)	8,439,339
% GC content	52.57
% of Genome (genes)	28.66
Gene average length (bp)	1,314
% of Genome(internal)	71.34
rRNAs	52,3,4 (5S,18S,28S)
tRNAs	214
sRNA	2
snRNA	18

results indicated that the production of IAA in BF-F could be an important factor contributing to plant growth promotion.

3.8 BF-F promotes N uptake related genes expression in *Arabidopsis* seedlings

Some key genes involved in N uptake, including *NRT1.1*, *NRT1.7*, and *NRT2.7* were analyzed in plants, and results showed that their expression in *A. thaliana* seedlings inoculated with BF-F was significantly higher than those uninoculated with BF-F (Figure 7C). This result indicated that the endophytic BF-F also enhanced plant uptake of N.

4 Discussion

This study isolated a new *Cladosporium* fungus strain, referred to as BF-F, from roots of *S. portulacastrum* seedling. It exhibits beneficial effects on plant growth by increasing root and leaf numbers and plant height. Based on the genomic sequence data, BF-F has gene clusters that are implicated in tryptophan biosynthesis and metabolism, which lead to the production IAA and IBA (Figure 7A). It has gene clusters involved in the sterol biosynthesis pathway, which would result in the
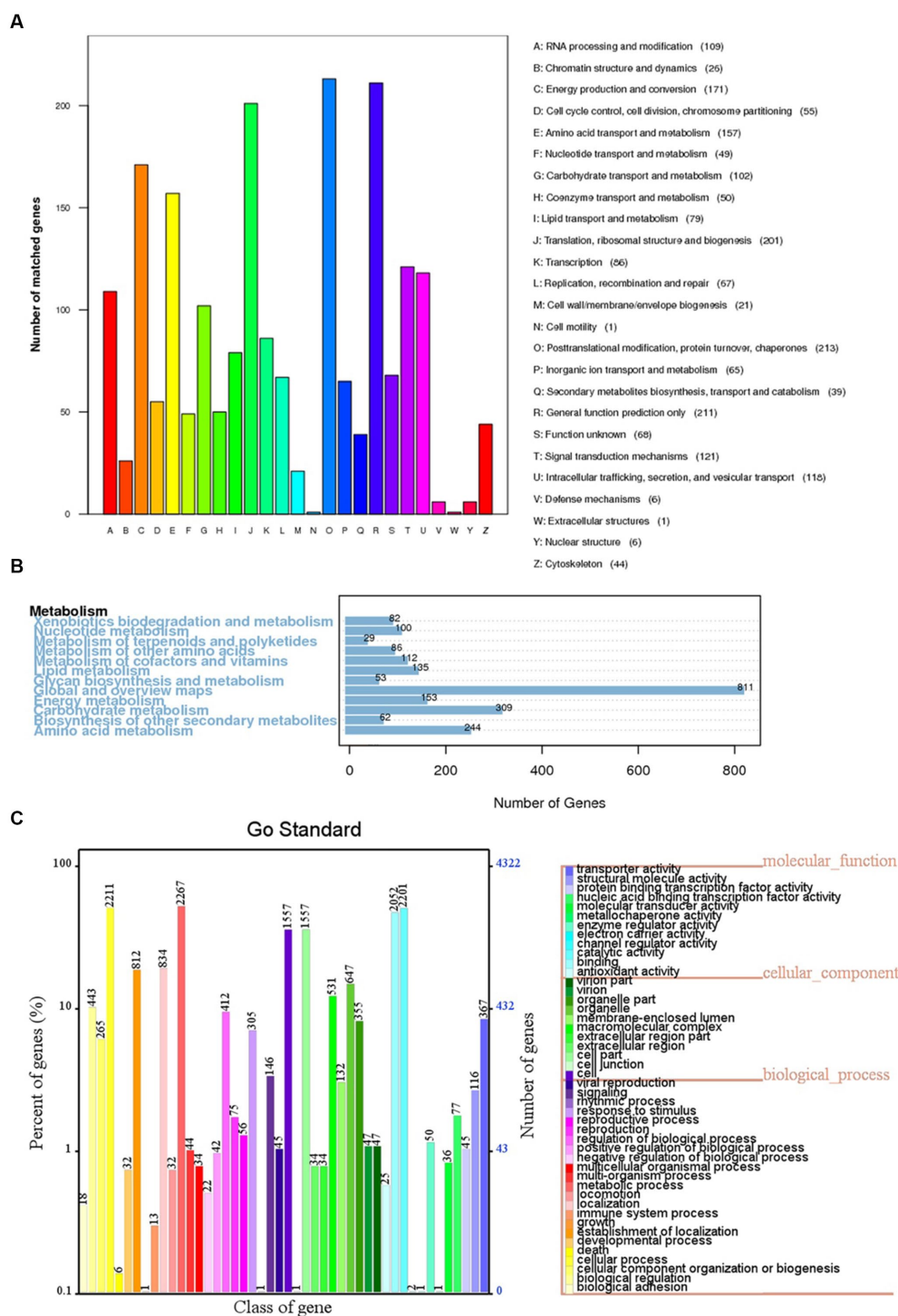
FIGURE 6

KOG, KEGG and GO classifications of predicted genes in BF-F. **(A)** Distribution of KOG classes for predicted proteins in BF-F. **(B)** KEGG classifications of predicted genes in metabolic pathway. **(C)** Gene Ontology of predicted genes in BF-F. The gene numbers were noted at the right of each bar; The length of each bar represents the percentage of genes in each GO class (the percentage of total GO-annotated genes was set as 100%); Each GO class with number of genes ≤10 was not shown.

biosynthesis of phytohormone BR. In addition to phytohormones produced by BF-F that directly contribute to the plant growth promotion, BF-F also support plant growth indirectly by increased expression of genes related to N uptake, such as *NRT1.1*, *NRT1.7*, and

NRT2.1 (Figure 7B). As far as it is known, the genome of *Cladosporium* has not yet been investigated. This study provides the first genomic sequence information about *Cladosporium* species and presents a mechanistic explanation as to how BF-F as a *Cladosporium* strain

TABLE 2 Genes potentially associated with PGP traits in the genome of BF-F.

Function/PGP trait	Gene	Description	Location/locus tag
Tryptophan biosynthesis	aroQ	3- dehydroquinate dehydratase II	Contig10:487946:488395
	nagZ	beta-N-acetylhexosaminidase	Contig10:502674:503914
	QUIB	qa-3 quinate dehydrogenase	Contig10:1023683:1024675
	TRP	tryptophan synthase	Contig10:1441293:1442593
	TRP1	prephenate dehydrogenase (NADP+)	Contig10:1496599:1497988
	trpD	anthranilate phosphoribosyltransferase	Contig13:1209593:1210753
	aroF	4- deoxy-7-phosphoheptulonate synthase	Contig14:239569:240808
	aroB	3- dehydroquinate synthase	Contig14:540431:542150
	aroC	chorismate synthase	Contig16:405385:406659
	TRP1	anthranilate synthase / indole-3-glycerol phosphate synthase	Contig1:2203175:2205490
	trpE	anthranilate synthase component I	Contig1:2786166:2787913
	aroF	3-deoxy-7-phosphoheptulonate synthase	Contig1:321179:322343
	ARO1	pentafunctional AROM polypeptide	Contig7:769064:773896
	GOT1	aspartate aminotransferase, cytoplasmic	Contig7:1236542:1237861
	ARO8	aromatic amino acid aminotransferase I / 2-aminoadipate transaminase	Contig8:394272:396702
Tryptophan metabolism	DDC, TDC	aromatic-L-amino-acid/L-tryptophan decarboxylase	Contig3:1013651:1015306
	ALDH	aldehyde dehydrogenase	Contig16:1124610:1136277
	MAO	monoamine oxidase	Contig1:3334092:3335552
Steroid biosynthesis	FDFT1	farnesyl-diphosphate farnesyltransferase	Contig2:34905:36471
	SQL	squalene monooxygenase	Contig9:1052550:1054004
	SMT1	sterol 24-C-methyltransferase	Contig10:183335:184477
	CYP51	sterol 14alpha-demethylase	Contig14:1153188:1154868
	TM7S	Delta14-sterol reductase	Contig1:1361362:1362948
	EBP	cholestenol Delta-isomerase	Contig5:1697694:1698497
	SC5DL	Delta7-sterol 5-desaturase	Contig4:1047053:1048255:-
	ERG3	Delta7-sterol 5-desaturase	Contig4:1047053:1048255
Nitrogen metabolism	NRT	MFS transporter, NNP family, nitrate/nitrite transporter	Contig10:801362:803187
	glnA,	glutamine synthetase	Contig3:216124:217758
	cyns	cyanaten lyase	Contig3:5281684:529873
	GTL1	glutamate synthase (NADH)	Contig5:744675:751205
	gdha	glutamate dehydrogenase	Contig16:221788:223217
Nitrogen reduction	NR	nitrate reductase (NAD(P)H)	Contig5:71331:73058
	NIT-6	nitrite reductase (NAD(P)H)	Contig8:321642:325025

promotes plant growth. To better understand the growth promoting effects of BF-F, further studies based on the genomic information are warranted.

Cladosporium is one of the most common fungal genera distributed in diverse substrates with various lifestyles. The identification of Cladosporium species based on morphology has been a challenge. Although conidiophore and conidia size and shape are important characteristics of different Cladosporium species, dimensions usually overlap among species in the genus. Therefore, molecular analysis has been used to identify Cladosporium species. The ITS rDNA, ACT, and EF-1α regions of the genome are often used

to explain the diversity and evolutionary trends in the Cladosporium genus (Bensh et al., 2010, 2012). However, ITS rDNA alone does not provide acceptable species resolution (Zalar et al., 2007). In contrast, ACT and EF-1α demonstrate a high degree of divergence among species (Voigt and Wöstemeyer, 2001; Bensh et al., 2010). In this study an integrated approach based on both molecular and morphological characteristics were used to identify the isolated BF-F. Molecular analysis showed that EF-1α sequence of BF-F had only 88% similarity to C. angulosum (Figure 2). However, morphologically, BF-F produces no soluble pigment when cultured in PDA medium (Figures 1A,B) and thus differs from C. angulosum as it releases sulfur-yellow pigment
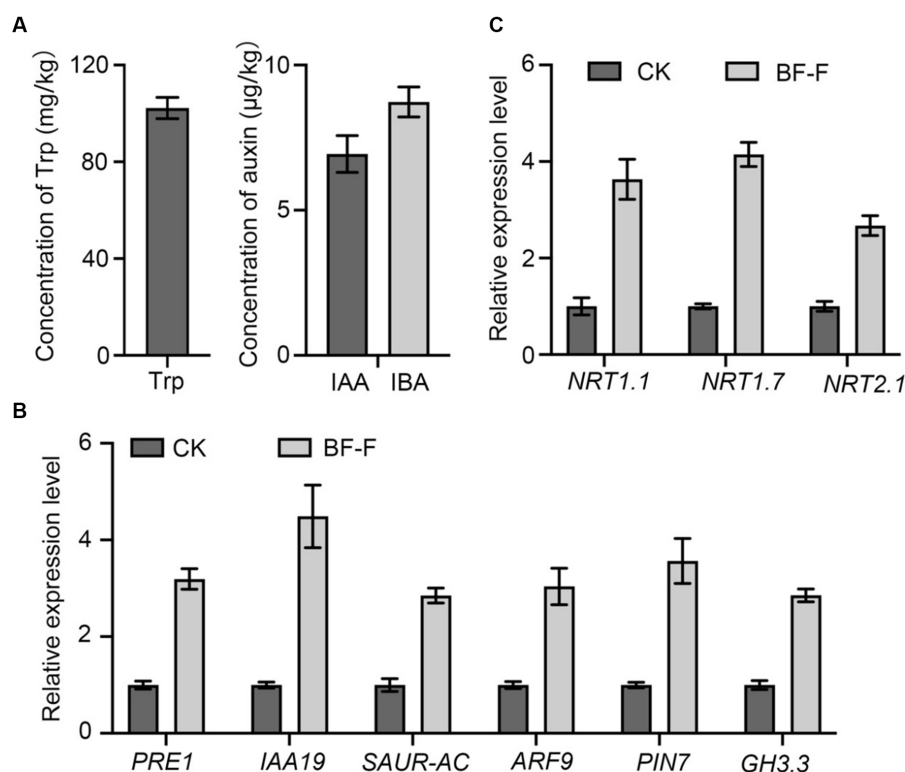
into PDA (Bensch et al., 2018). On the basis of sequence homology (99%) and phylogenetic analysis, BF-F is likely a new *C. angulosum* strain or possibly a new *Cladosporium* species that is most homologous to *C. angulosum*. Further research is needed to determine its identity.

Endophytic fungi, as an important class of PGPF species have attracted considerable interest due to their multiple beneficial effects on plant growth and improved plant tolerance to abiotic and biotic stresses (Zhang et al., 2006; Meenavalli et al., 2011; Perotto et al., 2012; Murali et al., 2013; Gouda et al., 2016; Toghueo et al., 2017; Strobel, 2018; Wei et al., 2022). Their secondary metabolites also served as an excellent source of bioactive compounds for potential use as antimicrobial, anti-insect, and anticancer materials (Gouda et al., 2016). Plant growth promotion executed by PGPF is a complex process that is often attributed to a variety of direct and indirect factors, including the solubilization of minerals, biosynthesis of growth-stimulating hormones, production of volatile organic compounds (VOCs), increased uptake of nutrients, and improved tolerance to biotic and abiotic stresses (Santoyo et al., 2016; Hossain et al., 2017; Priyadharsini and Muthukumar, 2017; El-Maraghy et al., 2020; Hossain and Sultana, 2020). Many species in the genus *Cladosporium* function as endophytic PGPF through different strategies for plant growth regulation, including gibberellic acid (GA) biosynthesis, the production of volatile substances, or antagonism against plant pathogens (Hamayun et al., 2009; Paul and Park, 2013; Torres et al., 2017). In this study, the isolated BF-F can substantially promote plant growth (Figures 3, 4). The promotion is not

only limited to its host plant *S. portulacastrum* but also other dicot plants, such as *Arabidopsis* and monocot rice (Figure 4). Metabolic pathway gene cluster analysis of the BF-F genome showed that the tryptophan synthesis and metabolism pathway happened in BF-F (Figures 5, 7A,B and Table 2), tryptophan metabolite IBA and IAA occurred in BF-F mycelium (Figure 7A), and auxin response genes correspondingly upregulated in *A. thaliana* seedlings inoculated with BF-F. All these results suggest that the production of IAA by BF-F is one of main reasons facilitating plant growth promotion, which has not been reported in *Cladosporium* species thus far. Moreover, sterol synthesis pathways were also found in BF-F genome, suggesting that BF-F has the potential to produce sterols (Figure 5 and Table 2). Although sterols may have limited effects on plants, sterols are the precursors of BR, which promotes plant growth. In addition, the annotated BF-F genome also revealed the presence of several gene clusters involved in nitrogen metabolism and reduction (Table 2). Plant N uptake related genes were upregulated in *A. thaliana* seedlings when inoculated with BF-F (Figure 7C), which may be another mechanism underpinning plant growth promotion executed by BF-F. These new metabolites and metabolic pathways in BF-F may not only explain the molecular mechanisms of plant growth promotion but also probably represent important characteristics for distinguishing BF-F from other *Cladosporium* species.

The investigation of clusters of genes involved in metabolic pathways is an emerging area in plant biology, and research has

provided some provocative insights into plant genome plasticity and evolution (Osbourn, 2010; Hautbergue et al., 2018). Clusters of functionally related genes has been investigated in different PGPF and plant growth-promoting rhizobacteria (PGPR). Genes clusters of phytohormone pathways are often obtained in plant growth-promoting microbes. Gene clusters associated with IAA production are most common in PGPF species (Liu et al., 2016; Abdullahi et al., 2021). *Cladosporium* is a diverse family of fungi with various species and functions. Further analyses of this genus based on a large number of isolates collected from different geographical regions are necessary to redefine species borders and growth-promoting potential within *Cladosporium* (Wirsel et al., 2002). However, the genomic information of *Cladosporium* species, especially PGPF, is still lacking. Although some *Cladosporium* species that can produce GA or volatile compounds have been reported, associated gene clusters remain unclear due to the lack of genomic information (Hamayun et al., 2009; Paul and Park, 2013). In the present study, gene clusters associated with IAA production by tryptophan-dependent pathways and sterol synthesis were identified in the BF-F genome (Table 2). These findings indicate that *Cladosporium* also possesses the potential to produce IAA and/or sterols for improving plant growth. Further analysis of *Cladosporium* species genomes will help uncover new metabolites and new metabolic pathways associated with plant growth promotion.

5 Conclusion

Cladosporium BF-F strain is a broad-spectrum PGPF that can promote growth of both monocot and dicot plants. The genome of BF-F has been sequenced, which should provide the first glimpse into the genetic basis of the plant growth promotion traits of *Cladosporium*. Based on the genomic information, BF-F also has the ability to produce IAA, similar to other PGPF species. The production of IAA along with indirect effects on N uptake may be the basis of BF-F in enhanced plant growth.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

NY: Data curation, Writing – original draft. WZ: Writing – original draft, Writing – review & editing. DW: Formal analysis, Writing – original draft. DC: Formal analysis, Writing – original draft. YC: Formal analysis, Writing – original draft. WH: Data curation, Writing – original draft. ZL: Data curation, Writing – original draft. XC: Writing – review & editing. GY: Data curation, Formal analysis, Writing – original draft. ZC: Writing – review & editing. JC: Writing – review & editing. XW: Writing – review & editing, Funding acquisition.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1287582/full#supplementary-material

References

Abdullahi, S., Haris, H., Zarkasi, K. Z., and Amir, H. G. (2021). Complete genome sequence of plant growth-promoting and heavy metal-tolerant Enterobacter tabaci 4M9(CCB-MBL 5004). *J. Basic. Microbiol.* 61, 293–304. doi: 10.1002/jobm.202000695

Adhikari, P., and Pandey, A. (2019). Phosphate solubilization potential of endophytic fungi isolated from Taxus wallichiana Zucc. Roots. *Rhizosphere* 9, 2–9. doi: 10.1016/j.rhisph.2018.11.002

Aly, A., Debbab, A., Kjer, J., and Proksch, P. (2010). Fungal endophytes from higher plants: a prolific source of phytochemicals and other bioactive natural products. *Fungal Divers.* 41, 1–16. doi: 10.1007/s13225-010-0034-4

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Michael Cherry, J., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Baron, N. C., and Rigobelo, E. C. (2021). Endophytic fungi: a tool for plant growth promotion and sustainable agriculture. *Mycology* 13, 39–55. doi: 10.1080/2150 1203.2021.1945699

Barron, G. L. (1962). New species and new records of Oidiodendron. *Can. J. Bot.* 40, 589–607. doi: 10.1139/b62-055

Bensch, K., Groenewald, J. Z., Meijer, M., Dijksterhuis, J., Jurjevic, Ž., Houbraken, J., et al. (2018). Cladosporium species in indoor environments. *Stud. Mycol.* 89, 177–301. doi: 10.1016/j.simyco.2018.03.002

Bensh, K., Braun, U., Groenewald, J. Z., and Crous, P. W. (2012). The genus Cladosporium. *Stud. Mycol.* 72, 1–401. doi: 10.3114/sim0003

Bensh, K., Groenewald, J. Z., Starink-Willemse, M., Andersen, B., Sumerell, B. A., Shin, H. D., et al. (2010). Species and ecological diversity within the Cladosporium cladosporioides complex (Davidiellaceae, Capnodiales). *Stud. Mycol.* 67, 1–94. doi: 10.3114/sim.2010.67.01

Bhar, A. (2021). Brassinosteroid: the potential steroid hormone in crop improvements. *Biot. Res. Today* 3, 1062–1064.

Chaibub, A. A., de Carvalho, J. C. B., de Sousa Silva, C., Collevatti, R. G., Gonvalves, F. J., de Carvalho Barros Côrtes, M. V., et al. (2016). Defence responses in

rice plants in prior and simultaneous applications of Cladosporium sp. during leaf blast suppression. *Environ. Sci. Pollut. R.* 23, 21554–21564. doi: 10.1007/s11356-016-7379-5

Chaisson, M. J., and Tesler, G. (2012). Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* 13:238. doi: 10.1186/1471-2105-13-238

Domsch, K. H., Gams, W., and Anderson, T. H. (1980). Compendium of Soil Fungi. *EurUR. J. SoilOIL. SciCI.* 59:1007. doi: 10.1111/j.1365-2389.2008.01052_1.x

Dugan, F. M., Schubert, K., and Braun, U. (2004). Check-list of Cladosporium names. *Schlechtendalia* 11, 1–103.

Ebrahimi, E., Rda, S. H., and Etebarian, H. R. (2022). Apple endophytic fungi and their antagonism against apple scab disease. *Front. Microbiol.* 13:10240001. doi: 10.3389/fmicb.2022.1024001

El-Maraghy, S. S., Tohamy, T. A., and Hussein, K. A. (2020). Role of plant-growth promoting Fungi (PGPF) in defensive genes expression of *Triticum aestivum* against wilt disease. *Rhizosphere* 15:100223. doi: 10.1016/j.rhisph.2020.100223

El-Morsy, E. M. (2000). Fungi isolated from the endorhizosphere of halophytic plantsfrom the Red Sea coast of Egypt. *Fungal Divers.* 5, 43–54.

Erman, M., Yurnaliza, A. L., and Adrian, H. (2022). Isolation and characterization of phosphate solubilizing activity of endophytic Fungi from Zingiberaceous species. *OnLine J. Biol. Sci.* 22, 149–156. doi: 10.3844/ojbsci.2022.149.156

Farr, D. F., Bills, G. F., Chamuris, G. P., and Rossman, A. Y. (1989). *Fungi on plants plan products in the United States*. ed. G. M. Plunkett (Minnesota: The American Phytopathological Society Press), pp. 340.

Feldmann, K. A. (2006). Steroid regulation improves crop yield. *Nat. Biotechnol.* 24, 46–47. doi: 10.1038/nbt0106-46

Flannigan, B. (2001). "Microorganisms in indoor air" in *Microorganisms in home and indoor work environments: Diversity, health impacts, investigation and control*. eds. B. Flannigan, R. A. Samson and J. D. Miller (London: Taylor & Francis), 17–31.

Fu, S. F., Wei, J. Y., Chen, H. W., Liu, Y. Y., Lu, H. Y., and Chou, J. Y. (2015). Indole-3-acetic acid: widespread physiological code in interactions of fungi with other organisms. *Plant Signal. Behav.* 10:e1048052. doi: 10.1080/15592324.2015.1048052

Gouda, S., Das, G., Sen, S. K., Shin, H. S., and Patra, J. K. (2016). Endophytes: a treasure house of bioactive compounds of medicinal importance. *Front. Microbiol.* 7:1538. doi: 10.3389/fmicb.2016.01538

Hamayun, M., Khan, S. A., Ahmad, N., Tang, D. S., Kang, S. M., Na, C. I., et al. (2009). Cladosporium sphaerospermum as a new plant growth-promoting endophyte from the roots of *Glycine max* (L.) Merr. *World J. Microbiol. Biotechnol.* 25, 627–632. doi: 10.1007/s11274-009-9982-9

Hartl, L., Zach, S., and Seidl-Seiboth, V. (2012). Fungal Chitinases: diversity, mechanistic properties, and biotechnological applications. *Appl. Microbiol. Biotechnol.* 93, 533–543. doi: 10.1007/s00253-011-3723-3

Hautbergue, T., Jamin, E. L., Debrauwer, L., Puel, O., and Oswald, I. P. (2018). From genomics to metabolomics, moving toward an integrated strategy for the discovery of fungal secondary metabolites. *Nat. Prod. Rep.* 35, 147–173. doi: 10.1039/C7NP00032D

Heuchert, B., Braun, U., and Schubert, K. (2005). Morphotaxonomic revision of fungi colous Cladosporium species (hyphomycetes). *Schlechtendalia* 13, 1–78.

Hossain, M. M., and Sultana, F. (2020). Application and mechanisms of plant growth promoting Fungi (PGPF) for Phytostimulation. *IntechOpen Organ. Agric.* 2, 1–31. doi: 10.5772/intechopen.92338

Hossain, M. M., Sultana, F., and Hyakumachi, M. (2017). Role of ethylene signaling in growth and systemic resistance induction by the plant growth promoting fungus Penicillium viridicatum in Arabidopsis. *J. Phytopathol.* 165, 432–441. doi: 10.1111/jph.12577

John, J. E., Maheswari, M., Kalaiselvi, T., Prasanthrajan, M., Poornachandhra, C., Rakesh, S. S., et al. (2023). Biomining *Sesuvium portulacastrum* for halotolerant PGPR and endophytes for promotion of salt tolerance in *Vigna mungo* L. *Front. Microbiol.* 14:1085787. doi: 10.3389/fmicb.2023.1085787

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40, D109–D114. doi: 10.1093/nar/gkr988

Karunasinghe, T. G., Maharachchikumbura, S. S. N., Velazhahan, R., and Al-Sadi, A. M. (2020). Antagonistic activity of endophytic and rhizosphere Fungi isolated from sea purslane (*Sesuvium portulacastrum*) against Pythium damping off of cucumber. *Plant Dis.* 104, 2158–2167. doi: 10.1094/PDIS-01-20-0003-RE

Khan, A. R., Ullah, I., Waqas, M., Shahzad, R., Hong, S. J., Park, G. S., et al. (2015). Plant growth-promoting potential of endophytic fungi isolated from *Solanum nigrum* leaves. *World J. Microbiol. Biotechnol.* 31, 1461–1466. doi: 10.1007/s11274-015-1888-0

Khan, A. L., Waqas, M., Hussain, J., Al-Harrasi, A., Al-Rawahi, A., Al-Hosni, K., et al. (2014). Endophytes aspergillus caespitosus LK12 and Phoma sp. LK13 of Moringa peregrina produce gibberellins and improve rice plant growth. *J. Plant Interact.* 9, 731–737. doi: 10.1080/17429145.2014.917384

Koonin, E. V., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Krylov, D. M., Makarova, K. S., et al. (2004). A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5:R7. doi: 10.1186/gb-2004-5-2-r7

Krause, K., Henke, C., Asiimwe, T., Ulbricht, A., Klemmer, S., Schachtschabel, D., et al. (2015). Biosynthesis and secretion of indole-3-acetic acid and its morphological effects on Tricholoma vaccinum-spruce ectomycorrhiza. *Appl. Environ. Microbiol.* 81, 7003–7011. doi: 10.1128/AEM.01991-15

Kryczyński, S., and Weber, Z. (2011). Fitopatologia. T. 2. Choroby roślin uprawnych [Diseases of culti-vated plants]. PWRiL, Poznań, 350–352.

Kumar, M., Brar, A., Yadav, M., Chawade, A., Vivekanand, V., and Pareek, N. (2018). Chitinases-potential candidates for enhanced plant resistance towards fungal pathogens. *Agriculture* 8:88. doi: 10.3390/agriculture8070088

Latz, M. A. C., Jensen, B., Collinge, D. B., and Jørgensen, H. J. L. (2018). Endophytic fungi as biocontrol agents: elucidating mechanisms in disease suppression. *Plant Ecol. Divers.* 11, 555–567. doi: 10.1080/17550874.2018.1534146

Levetin, E., and Dorsey, K. (2006). Contribution of leaf surface fungi to the air spora. *Aerobiologia.* 22, 3–12. doi: 10.1007/s10453-005-9012-9

Li, H., Dou, M., Wang, X., Guo, N., Kou, P., Jiao, J., et al. (2021). Optimization of cellulase production by a novel endophytic fungus Penicillium oxalicum R4 isolated from *Taxus cuspidata*. *Sustainability* 13:6006. doi: 10.3390/su13116006

Lim, H. J., Lee, E. H., Yoon, Y., Chua, B., and Son, A. (2016). Portable lysis apparatus for rapid single-step DNA extraction of *Bacillus subtilis*. *J. Appl. Microbiol.* 120, 379–387. doi: 10.1111/jam.13011

Liu, W., Wang, Q., Hou, J., Tu, C., Luo, Y., and Christie, P. (2016). Whole genome analysis of halotolerant and alkalotolerant plant growth-promoting rhizobacterium Klrbsiella sp. D5A. *Sci. Rep.* 6:26710. doi: 10.1038/srep26710

Lokhande, V. K., Gor, B. K., Desai, N. S., Nikam, T. D., and Suprasanna, P. (2013). *Sesuvium portulacastrum*, a plant for drought, salt stress, sand fixation, food and phytoremediation. *Agron. Sustain. Dev.* 33, 329–348. doi: 10.1007/s13593-012-0113-x

Marques, N. P., de Cassia Pereira, J., Gomes, E., da Silva, R., Araújo, A. R., Ferreira, H., et al. (2018). Cellulases and xylanases production by endophytic fungi by solid state fermentation using lignocellulosic substrates and enzymatic sacchari cation of pretreated sugarcane bagasse. *Ind. Crop Prod.* 122, 66–75. doi: 10.1016/j.indcrop.2018.05.022

Meenavalli, B., Rajulu, G., Irunavukkarasu, N., Suryanarayanan, T. S., Ravishankar, J. P., Gueddari, N. E. E., et al. (2011). Chitinolytic enzymes from endophytic fungi. *Fungal Divers.* 47, 43–53. doi: 10.1007/s13225-010-0071-z

Mullins, J. (2001). "Microorganisms in outdoor air" in *Microorganisms in home and indoor work environments: Diversity, health impacts, investigation and control*. eds. B. Flannigan, R. A. Samson and J. D. Miller (London: Taylor & Francis), 3–16.

Murali, M., Sudisha, J., Amruthesh, K. N., Ito, S. I., and Shetty, H. S. (2013). Rhizosphere fungus Penicillium chrysogenum promotes growth and induces defence related genes and downy mildew disease resistance in pearl millet. *Plant Biol.* 15, 111–118. doi: 10.1111/j.1438-8677.2012.00617.x

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Plant Physiol.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Nilsson, R. H., Larsson, K. H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., et al. (2019). The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* 47, 259–264. doi: 10.1093/nar/gky1022

Ogórek, R., Lejman, A., Pusz, W., Miłuch, A., and Miodyńska, P. (2012). Characteristics and taxonomy of Cladosporium fungi. *Mikologia Lekarska* 19, 80–85.

Osbourn, A. (2010). Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. *Plant Physiol.* 154, 531–535. doi: 10.1104/pp.110.161315

Patel, M. K., Pandey, S., Brahmbhatt, H. R., Mishra, A., and Jha, B. (2019). Lipid content and fatty acid profile of selected halophytic plants reveal a promising source of renewable energy. *Biomass Bioenergy.* 124, 25–32. doi: 10.1016/j.biombioe.2019.03.007

Paul, D., and Park, K. S. (2013). Identification of volatiles produced by Cladosporium cladosporioides CL-1, a fungal biocontrol agent that promotes plant growth. *Sensors* 13, 13969–13977. doi: 10.3390/s131013969

Perotto, S., Martino, E., Abba, S., and Vallino, M. (2012). "Genetic diversity and functional aspects of ericoid mycorrhizal fungi" in *The Mycota IX*. ed. B. Hock. *2nd* ed (Berlin: Springer-Verlag), Fungal Associations), 255–285.

Priyadharsini, P., and Muthukumar, T. (2017). The root endophytic fungus Curvularia geniculate from *Parthenium hysterophorus* roots improves plant growth through phosphate solubilization and phytohormone production. *Fungal Ecol.* 27, 69–77. doi: 10.1016/j.funeco.2017.02.007

Saikkonen, K., Faeth, S. H., Helander, M., and Sullivan, T. J. (2003). FUNGAL ENDOPHYTES: a continuum of interactions with host plants. *Annu. Rev. Ecol. Syst.* 29, 319–343. doi: 10.1146/annurev.ecolsys.29.1.319

Sandoval-Denis, M., Sutton, D. A., Martin-Vicente, A., Cano-Lira, J. F., Wiederhold, N., Guarro, J., et al. (2015). Cladosporium species recovered from clinical samples in the United States. *J. Clin. Microbiol.* 53, 2990–3000. doi: 10.1128/JCM.01482-15

Santoyo, G., Moreno-Hagelsieb, G., Orozco-Mosqueda, M. C., and Glick, B. R. (2016). Plant growth-promoting bacterial endophytes. *Microbiol. Res.* 183, 92–99. doi: 10.1016/j.micres.2015.11.008

Sasirekha, B., Shivakumar, S., and Sullia, S. B. (2012). Statistical optimization for improved indole-3-acetic acid (IAA) production by Pseudomonas aeruginosa and demonstration of enhanced plant growth promotion. *J. Soil Sci. Plant Nutr.* 12, 863–873. doi: 10.4067/S0718-95162012005000038

Schena, M., Lioyd, A. M., and Davis, R. W. (1991). A steroid-inducible gene expression system for plant cells. *PNAS* 88, 10421–10425. doi: 10.1073/pnas.88.23.10421

Shah, S., Shrestha, R., Maharjan, S., Selosse, M. A., and Pant, B. (2018). Isolation and characterization of plant growth-promoting endophytic Fungi from the roots of Dendrobium moniliforme. *Plan. Theory* 8:5. doi: 10.3390/plants8010005

Silva, N. I. D. E., Brooks, S., Lumyong, S., and Hyde, K. D. (2018). Use of endophytes as biocontrol agents. *Fungal Biol. Rev.* 33, 133–148. doi: 10.1016/j.fbr.2018.10.001

Strobel, G. (2018). The emergence of endophytic microbes and their biological promise. *J. Fungi* 4:57. doi: 10.3390/jof4020057

Toghueo, R. M. K., Zabalgogeazcoa, I., Vazquez de Aldana, B. R., and Boyom, F. F. (2017). Enzymatic activity of endophytic fungi from the medicinal plants *Terminalia catappa*, Terminalia mantaly. *Cananga odorataS. Afr. J. Bot.* 109, 146–153. doi: 10.1016/j.sajb.2016.12.021

Torres, D. E., Rojas-Martinez, R. I., Zavaleta-Mejia, E., Guevara-Fefer, P., Marquez-Guzman, G. J., and Perez-Martinez, C. (2017). Cladosporium cladosporioides and Cladosporium pseudocladosporioides as potential new fungal antagonists of Puccinia horiana Henn., the causal agent of chrysanthemum white rust. *PLoS One* 12:e0170782. doi: 10.1371/journal.pone.0170782

Venkateswarulu, N., Shameer, S., Bramhachari, P. V., Thaslim Basha, S. K., Nagaraju, C., and Vijaya, T. (2018). Isolation and characterization of plumbagin (5- hydroxyl-2-methylnaptalene-1, 4-dione) producing endophytic fungi Cladosporium delicatulum from endemic medicinal plants. *Biotechnol. Rep.* 20:e00282. doi: 10.1016/j.btre.2018.e00282

Voigt, K., and Wöstemeyer, J. (2001). Phylogeny and origin of 82 zygomycetes from all 54 genera of the Mucorales and Mortierellales based on combined analysis of actin and translation elongation factor EF-1 genes. *Gene* 270, 113–120. doi: 10.1016/S0378-1119(01)00464-4

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:112963. doi: 10.1371/journal.pone.0112963

Watson, A., and Napier, T. (2009). Diseases of cucurbit vegetables. *Prime Fact* 832, 2–3.

Wei, X., Chen, J., Zhang, C., Liu, H., Zheng, X., and Mu, J. (2020). Ericoid mycorrhizal fungus enhances microcutting rooting of Rhododendron fortunei and subsequent growth. *Hortic. Res.* 7:140. doi: 10.1038/s41438-020-00361-6

Wei, X., Chen, J., Zhang, C., and Pan, D. (2016). A New Oidiodendron maius Strain Isolated from Rhododendron fortune and its Effects on Nitrogen Uptake and Plant Growth. *Front Microbiol.* 7: 1327. doi: 10.3389/fmicb.2016.01327

Wei, X., Zhang, W., Zulfiqar, F., Zhang, C., and Chen, J. (2022). Ericoid mycorrhizal fungi as biostimulants for improving propagation and production of ericaceous plants. *Front. Plant Sci.* 13:1027390. doi: 10.3389/fpls.2022.1027390

Wirsel, S. G., Runge-Froböse, C., Ahren, D. G., Kemen, E., Oliver, R. P., and Mendgen, K. W. (2002). Four or more species of Cladosporium sympatrically colonize *Phragmites Australis*. *Fungal Genet. Biol.* 35, 99–113. doi: 10.1006/fgbi.2001.1314

Xu, P., Wang, H., Qin, C., Li, Z., Lin, C., Liu, W., et al. (2021). Analysis of the taxonomy and pathogenic factors of Pectobacterium aroidearum L6 using whole-genome sequencing and comparative genomics. *Front. Microbiol.* 12:679102. doi: 10.3389/fmicb.2021.679102

Yew, S. M., Chan, C. L., Lee, K. W., Na, S. L., Tan, R., Hoh, C. C., et al. (2014). A five-year survey of dematiaceous fungi in a tropical hospital reveals potential opportunistic species. *PLoS One.* 9: e104352. doi: 10.1371/journal.pone.0104352

You, Y. H., Yoon, H., Kang, S. M., Woo, J. R., Choo, Y. S., Lee, I. J., et al. (2013). Cadophora malorum Cs-8-1 as a new fungal strain producing gibberellins isolated from *Calystegia soldanella*. *J. Basic Microbiol.* 53, 630–634. doi: 10.1002/jobm.201200002

Zalar, P., De Hoog, G. S., Schroers, H. J., Crous, P. W., Groenewald, J. Z., and Gunde-Cimerman, N. (2007). Phylogeny and ecology of the ubiquitous saprobe Cladosporium sphaerospermum, with descriptions of seven new species from hypersaline environments. *Stud. Mycol.* 58, 157–183. doi: 10.3114/sim.2007.58.06

Zhang, Y., Chen, F. S., Wu, X. Q., Luan, F. G., Zhang, L. P., Fang, X. M., et al. (2018). Isolation and characterization of two phosphate-solubilizing fungi from rhizosphere soil of moso bamboo and their functional capacities when exposed to different phosphorus sources and pH environments. *PLoS One* 13:e0199625. doi: 10.1371/journal.pone.0199625

Zhang, S. W., Gan, Y. T., and Xu, B. L. (2016). Application of plant-growth-promoting fungi Trichoderma longibrachiatum T6 enhances tolerance of wheat to salt stress through improvement of antioxidative defense system and gene expression. *Front. Plant Sci.* 7:1405. doi: 10.3389/fpls.2016.01405

Zhang, H. W., Song, Y. C., and Tan, R. X. (2006). Biology and chemistry of endophytes. *Nat. Prod. Rep.* 23, 753–771. doi: 10.1039/b609472b

A comparative genomics study of the microbiome and freshwater resistome in Southern Pantanal

André R. de Oliveira[1], Bárbara de Toledo Rós[2], Rodrigo Jardim[1], Nelson Kotowski[1], Adriana de Barros[3], Ricardo H. G. Pereira[3], Nalvo Franco Almeida[2] and Alberto M. R. Dávila[1]*

[1]Laboratório de Biologia Computacional e Sistemas, Instituto Oswaldo Cruz, Rio de Janeiro, Brazil, [2]Universidade Federal do Mato Grosso do Sul, Campo Grande, Brazil, [3]Universidade Federal do Mato Grosso do Sul, Aquidauana, Brazil

This study explores the resistome and bacterial diversity of two small lakes in the Southern Pantanal, one in Aquidauana sub-region, close to a farm, and one in Abobral sub-region, an environmentally preserved area. *Shotgun* metagenomic sequencing data from water column samples collected near and far from the floating macrophyte *Eichhornia crassipes* were used. The Abobral small lake exhibited the highest diversity and abundance of antibiotic resistance genes (ARGs), antibiotic resistance classes (ARGCs), phylum, and genus. RPOB2 and its resistance class, multidrug resistance, were the most abundant ARG and ARGC, respectively. Pseudomonadota was the dominant phylum across all sites, and *Streptomyces* was the most abundant genus considering all sites.

KEYWORDS

resistome, metagenomics, microbiome, pantanal, freshwater

1 Introduction

Antibiotic resistance is emerging as a significant global public health issue due to the swift rise of resistant bacteria and the concurrent decline in new drugs entering the market (Serwecińska, 2020). While resistance is a natural phenomenon in microbial communities, where it serves as a form of competition (Frost et al., 2018), its effects can be amplified in environments with high antibiotic concentrations. These environments include livestock farms (Qian et al., 2018), aquacultures (Preena et al., 2020), hospital effluents (Hassoun-Kheir et al., 2020), and wastewater treatment plants (Raza et al., 2022). Bacteria possess horizontal gene transfer mechanisms (integrons, MGE, and plastids) that facilitate the spread of antibiotic resistance genes (ARGs) within the community (Sun et al., 2019). Consequently, a pathogenic species can develop resistance to a specific antibiotic without direct exposure to it.

Although antibiotic resistance in pathogenic bacteria is well-studied, these bacteria represent only a small fraction of the microbial community (Doron and Gorbach, 2008). Therefore, our understanding of the antibiotic-resistance genes in non-pathogenic bacteria, particularly those inhabiting rivers, lakes, soils, and oceans, remains limited. This gap in knowledge is due to the difficulty in cultivating these bacteria using current protocols. Metagenomics, which allows for the analysis of a microbial community through sequencing of environmental genetic material without the need for cultivation (Doron and Gorbach, 2008), has emerged as a promising solution.
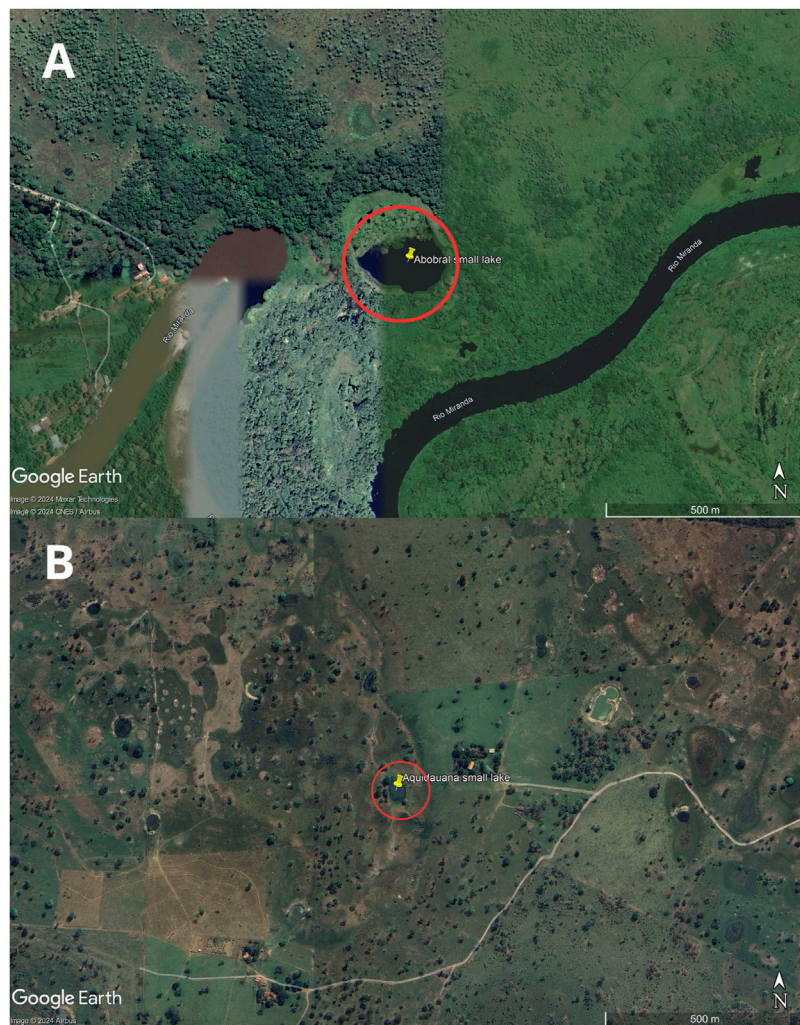
FIGURE 1
Map of the sample collection location. The small lakes are highlighted by a red circle. Image **(A)** is from Abobral and **(B)** is from Aquidauana.

Given the central role these microorganisms play in biogeochemical processes, studies on this topic have increased. Understanding the resistome of non-culturable bacterial communities is crucial for identifying potential gene reservoirs that could contribute to the evolution and spread of antibiotic resistance (Fresia et al., 2019).

The significance of biotic antibiotic removal mechanisms, often carried out by microorganisms, is highlighted in this context. These mechanisms are intertwined with various plant-based processes, a phenomenon known as phytoremediation. Certain plants have demonstrated the ability to eliminate and tolerate high levels of antibiotics without experiencing toxic effects (Kurade et al., 2021; Polińska et al., 2021). As a result, the microbial community present is shaped by the presence of plants and their impact on the rhizosphere, which in turn affects antibiotic resistance.

Our study is centered on the Pantanal Sul Matogrossense, one of the world's largest wetland ecosystems (Assine, 2015). This biome covers parts of Brazil (78%), Bolivia (18%), and Paraguay (4%). Despite its recognition, it has been significantly impacted by livestock activities, which often involve the use of large quantities of antibiotics (Ferrante and Fearnside, 2022). This region, abundant in water, serves as an efficient medium for the spread of Mobile Genetic Elements (MGE). The vast water systems in the Pantanal enhance the potential for ARGs to disseminate widely (Aminov and Mackie, 2007; Baquero et al., 2008). Hence, the examination of this region's resistome is of paramount importance from both a public health and academic perspective.

This study is set to explore the resistome and taxonomy of filtered water samples from two small lakes in Pantanal. One small lake is in a farm area, the other in a reserve, where samples were taken both near and far from the *Eichhornia crassipes* (Mart.) Solms macrophyte. We'll analyze sequencing data, looking into the identity and diversity of the Antibiotic Resistance Genes (ARGs), and also classify the samples taxonomically. This is the first time a metagenomic approach is being used to study the bacterial resistome of this area.
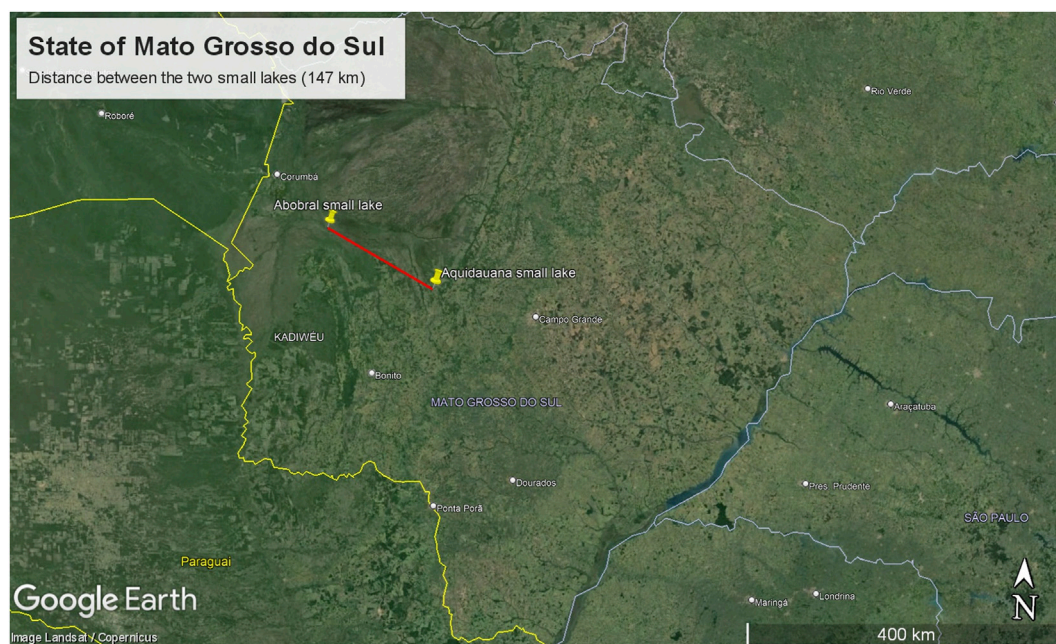
FIGURE 2
Map displaying the distance between the small lakes.

2 Methods

2.1 Sampling

The sampling site comprised two small lakes located in the Pantanal Sul Matogrossense: one within an environmental reserve unit in the Pantanal de Abobral subregion (Figure 1A), municipality of Corumbá (19°34′35″S 57°00′46″W), and the other in a farm region in the Pantanal de Aquidauana subregion (Figure 1B), municipality of Aquidauana (20°12′30″S 55°46′29″W), 147 km away from each other (Figure 2). In each small lake, water column samples were collected near the floating macrophyte (*Eichhornia crassipes*) and at a distance of 10 m from it. Access to genetic samples was properly registered at Brazilian SisGen under the code: AB3AE33. The sequencing files were submitted to the NCBI Bioproject under the code PRJNA1078255.

For clarity and objectivity, the sample from Abobral near the macrophyte will be referred to as "Site 1", Abobral distant from the macrophyte as "Site 2", Aquidauana near the macrophyte as "Site 3", and Aquidauana distant from the macrophyte as "Site 4". Each sample consisted of 10 L of water, collected in autoclaved bottles and subsequently stored at 4°C.

In the laboratory, the water samples were filtered through 1.2μm, 0.8μm, and 0.45 μm membranes, but only the data from the last one were used for this work. After the DNA extraction with QIAGEN DNEasy PowerWater Kit, the genetic material extracted from each of the 10 L bottled water was sent to the Sequencing Platform of the Histocompatibility and Cryopreservation Laboratory of the State University of Rio de Janeiro (UERJ) for a *shotgun* sequencing on the Illumina HiSeq-

2500 platform. The sequenced data were then stored on the servers of the Laboratory of Computational Biology and Systems at Oswaldo Cruz Institute/Fiocruz and submitted to NCBI SRA (PRJNA1078255).

2.2 Data analysis

The sequencing data were processed using two sequence analysis tools: the Metawrap pipeline (version 1.3) (Uritskiy et al., 2018) and DeepArg (version 2.0) (Arango-Argoty et al., 2018). The Metawrap pipeline was utilized for the following steps: i) Quality verification of the sequences using FastQC, ii) Sequence cleaning with Trimmomatic, iii) Taxonomic inference using Kraken, and iv) Taxonomy visualization with Krona. Following the cleaning process conducted by Trimmomatic, the data were further analyzed by DeepArg. This allowed for the prediction of Antibiotic Resistance Genes (ARGs) and their corresponding Antibiotic Resistance Classes (ARGCs), following the classification scheme provided by the tool.

All statistical analyses were performed using the R programming language. The descriptive statistics, including all diversity indices, were computed using the Vegan package (version 2.15–1)(Dixon, 2003) The graphics were generated with the ggplot2 package (version 3.4.4).

Due to the large volume of data, a selection criterion was established for the ARGs, ARGCs, phylum, and genus to be included in the graphical analysis. The criteria for inclusion were a relative abundance of 5% or higher for ARGs, ARGCs, and phylum in at least one sample site. For genus, the threshold was set at a relative abundance of 1% or higher. This adjustment was necessary

Information	Site 1	Site 2	Site 3	Site 4
ARGs	74	55	48	55
ARGCs	19	18	14	15
reads	1,685	1,399	874	974
phylum	26	24	22	26
genus	468	485	384	442

3 Results

3.1 Diversity of ARGs and ARGCs

A total of 232 ARGs and 66 ARGCs were found adding up all sites. Among them, we were able to identify a total of 103 unique ARGs and 21 Antibiotic Resistance Gene Classes (ARGCs). Comparing all sites, Site one stood out with the most ARGs (74), ARGCs (19), and the highest number of reads (1,685). On the other hand, Site 3 had the least with 48 ARGs, 14 ARGCs, and 874 reads (Table 1).

A comparative analysis between the small lakes revealed that Abobral's small lake (Sites 1 and 2) had a greater diversity of unique ARGs (84) and ARGCs (21) than Aquidauana's small lake (Sites 3 and 4), which had 69 ARGs and 17 ARGCs.

When considering the presence (Sites 1 and 3) and absence (Site 2 and 4) of the macrophyte, its presence seemed to slightly increase the number of different ARGs (82 vs. 79) but decrease the number of ARGCs (19 vs. 21). The co-occurrence of ARGs and ARGCs, whether together or isolated, is illustrated in Supplementary Figures S1 and S2. At least 24 ARGs and 11 known ARGCs were found to occur together in all sites.

The most abundant ARG was RPOB2, accounting for more than half of the total reads in all samples, with Site 3 having the highest percentage (73.8%). Other ARGs that had high percentages compared to the others were BACA and UGD in Sites 1 and 2, both accounting for approximately 10% of the total reads in each site (Figure 3).

Given that RPOB2 was the most abundant ARG across all samples, its corresponding antibiotic resistance class, Multidrug Resistance, would consequently be the most prevalent. Bacitracin and Peptides come in sequence also reflecting the proportions of the

as only two genera had a relative abundance higher than 5%. This approach ensured a manageable and representative subset of data for graphical analysis.

The inferential statistics involved testing for normality using the Shapiro-Wilk test, Lilliefors test, and QQ plot. Based on these normality tests, only non-parametric tests were appropriate for our data. Therefore, group comparisons were made using the Kruskal–Wallis test, followed by a *post hoc* Dunn's test with Bonferroni correction.

These comparisons were performed on alpha and beta diversity indices for ARGs, ARGCs, and the diversity of phyla and genera. Given that the Abobral and Aquidauana small lakes do not have direct contact, and therefore, their communities are isolated from each other, beta diversity was analyzed only among locations that establish a habitat gradient (WHITTAKER, 1972), that is, in each isolated small lake, taking into account only the presence and absence of macrophyte. The bootstrap resampling process (R = 100) was used to enhance the reliability of the results. All tests were considered significant at $p < 0.05$.
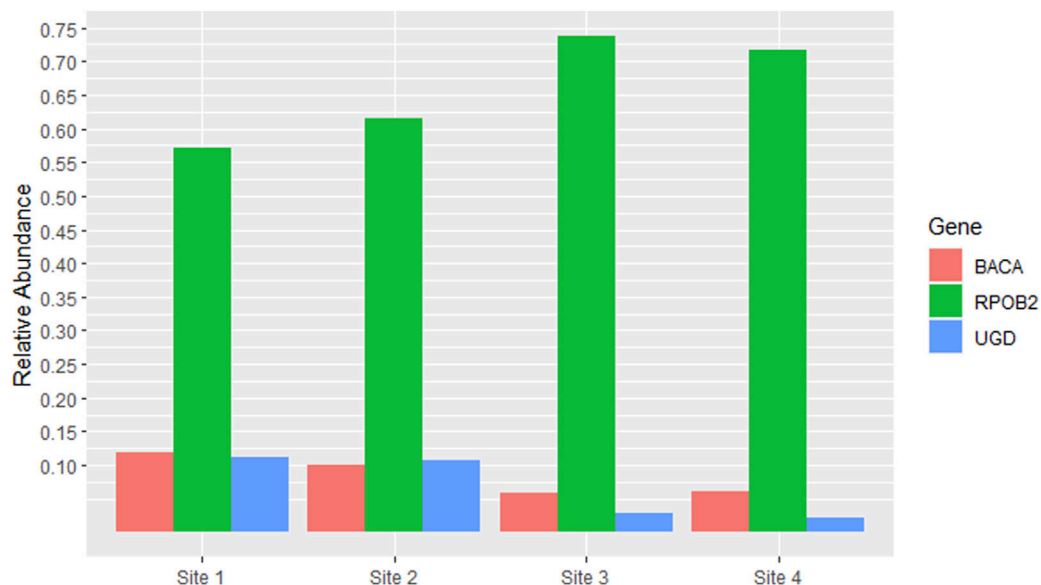

FIGURE 3
Antibiotic Resistance Genes (ARGs) which had relative abundance equal to or greater than 5% in at least one site. The y-axis represents the percentage of relative abundance (0.75 = 75%).
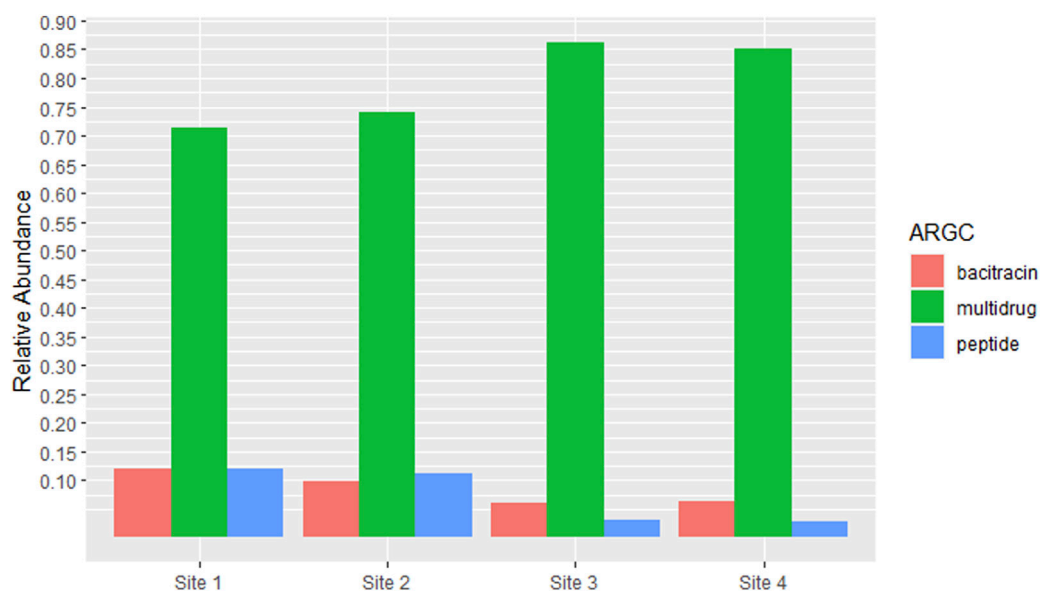
FIGURE 4
Antibiotic Resistance Classes (ARGCs) which had relative abundance equal to or greater than 5% in at least one site. The *y*-axis represents the percentage of relative abundance.
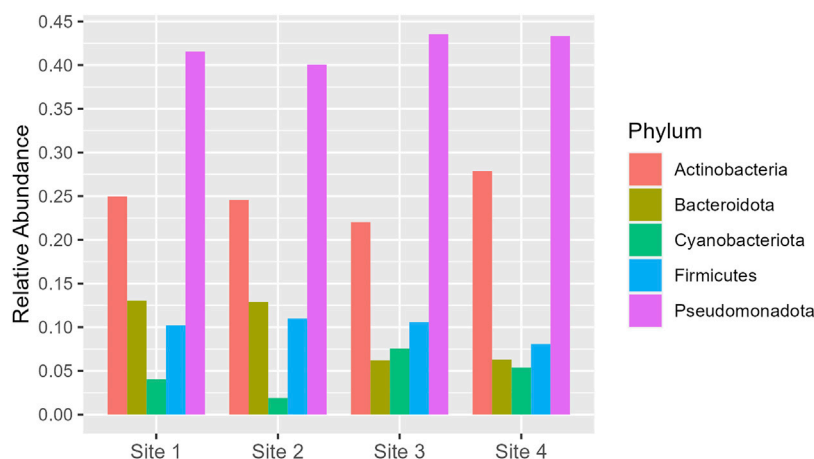

FIGURE 5
Phyla which had relative abundance equal to or greater than 5% in at least one site. The *y*-axis represents the percentage of relative abundance.

genes BACA and UGD (Figure 4). The richness and abundance of all ARGs and ARGC present in each site can be found in Supplementary Figures S2 and S4.

A total of 98 phyla and 1779 genera were found adding up all sites. Among them, we were able to identify a total of 31 unique phyla and 917 genera. Comparing all sites, Site 1 and 4 stood out with the most unique phyla (26) and Site 2 with the most genera (485). On the other hand, Site 3 had the least with 22 phyla and 384 genera (Table 1).

Similarly to the ARGs and ARGCs comparative analysis, Abobral's small lake (Sites 1 and 2) had a greater diversity of unique phyla (31) and genera (687) than Aquidauana's small lake (Sites 3 and 4), which had 27 phyla and 576 genera. When considering the presence (Sites 1 and 3) and absence (Site 2 and 4) of the macrophyte, its presence seemed to also slightly increase the number of different phyla (32 vs. 29) but decrease the number of genera (653 vs. 700).

Pseudomonadota is the dominant phylum across all sites, with relative abundance varying from 0.40 to 0.43. Actinobacteria follows as the second most abundant, with a slight increase in proportion at Site 4 (0.28) compared to the range of 0.22–0.25 at the other sites. Bacteroidota and Firmicutes exhibit similar distributions, but Bacteroidota shows a notable decrease at Site 3 and Site 4 (around 0.06) compared to Site one and Site 2 (around 0.13). Cyanobacteriota, although the least abundant, show a significant increase at Site 3 (0.08) compared to Site 2
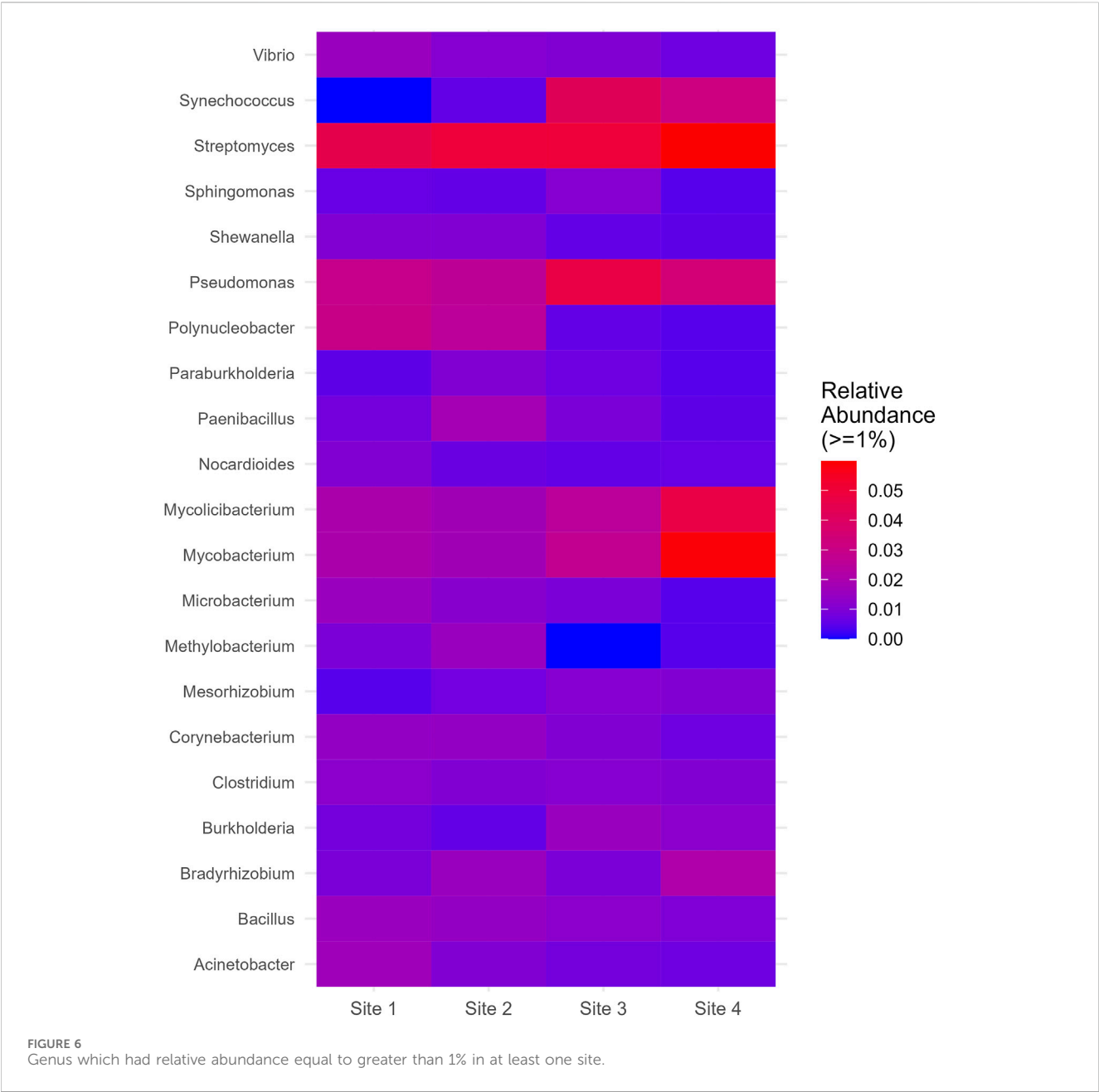
FIGURE 6
Genus which had relative abundance equal to greater than 1% in at least one site.

TABLE 2 Simpson Indexes for ARGs, ARGCs, phylum and genus.

Simpson Index	Site 1	Site 2	Site 3	Site 4
ARG	0.64	0.60	0.44	0.47
ARGC	0.46	0.42	0.25	0.26
phylum	0.74	0.74	0.74	0.71
genus	0.99	0.99	0.99	0.99

TABLE 3 Shannon Indexes for ARGs, ARGCs, phylum and genus.

Shannon Index	Site 1	Site 2	Site 3	Site 4
ARG	1.85	1.67	1.37	1.48
ARGC	0.99	0.94	0.63	0.69
phylum	1.67	1.74	1.75	1.68
genus	5.51	5.57	5.32	5.28

(0.02) (Figure 5). More information about all the phyla found in all sites can be seen in Supplementary Material (Supplementary Figure S5).

On the genus level, *Streptomyces* and *Pseudomonas* were relatively abundant across all sites and displayed a slight increase

at Site 4 and Site 3, respectively. *Mycobacterium* and *Mycolicibacterium*, while exhibiting similar distributions, showed a significant increase at Site 4. Synechococcus, which was absent at Site 1, manifested a substantial increase at Site 3 and Site 4. *Polynucleobacter*, on the other hand, showed a significant
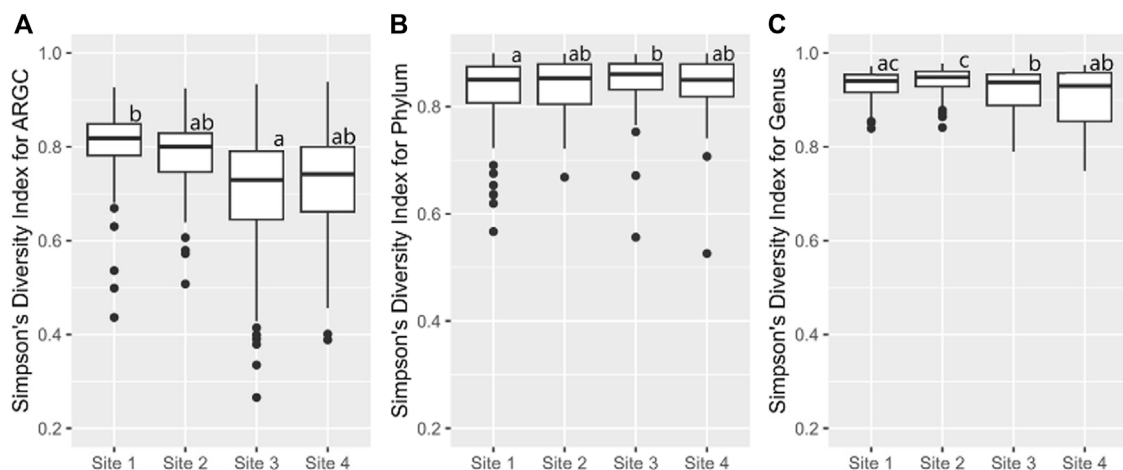
FIGURE 7
From left to right, bootstrap boxplots from the Simpson's Index of ARGC **(A)**, phylum **(B)**, and genus **(C)** diversity of each site. Letters inside the graph represent Dunn's test results.
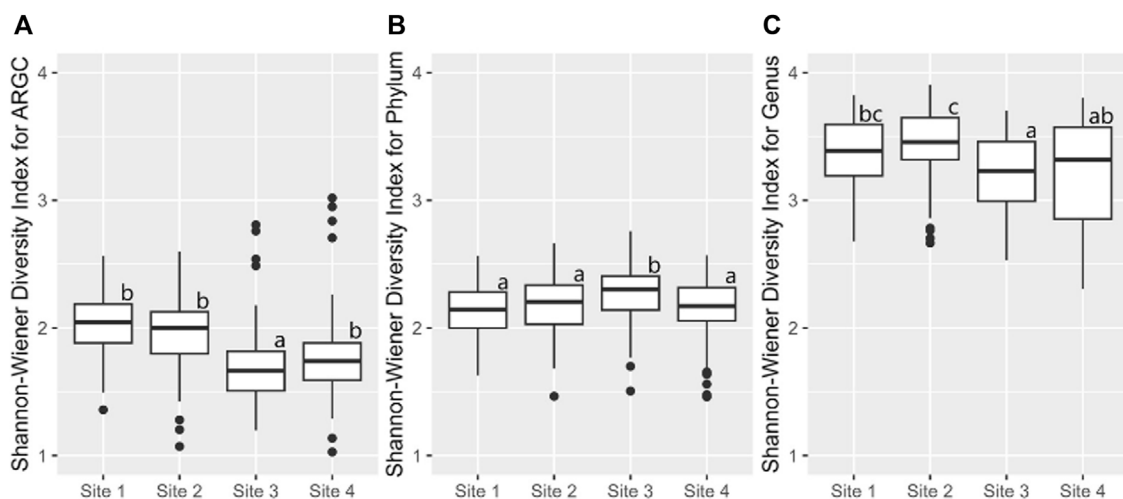

FIGURE 8
From left to right, bootstrap boxplots from the Shannon-Wiener Index of ARGC **(A)**, phylum **(B)**, and genus **(C)** diversity of each site. Letters inside the graph represent Dunn's test results.

decrease at Site 3 and Site 4. *Methylobacterium*, despite its low relative abundance at Site one and Site 2, was completely absent at Site 3 and appeared in lower proportions at Site 4 (Figure 6). More information about all the genera found in all sites can be seen in Supplementary Material (Supplementary Figure S6).

3.2 Diversity indexes

Both Simpson and Shannon indices were employed to assess the diversity for the following categories: ARGs, ARGCs, phylum, and genus. The results are listed in Tables 2, 3.

Further statistical analysis revealed distinct differences in diversity across various sites. The Simpson index revealed key differences in ARGC diversity between Sites one and 3

(Figure 7A), phylum diversity between Sites 1 and 3 (Figure 7B), and genus diversity between Sites 2 and 3, and Sites 2 and 4 (Figure 7C).

Building on this, the Shannon index identified additional disparities. In ARGC diversity, it showed differences between Sites 1, 2, and 3 (Figure 8A). For phylum diversity, it highlighted differences between Sites 2, 3, and 4 (Figure 8B). For genus diversity, it confirmed differences between Sites 1, 2, and 4 (Figure 8C). No significant differences were observed for ARGs in either index.

The beta diversity analysis, utilizing Bray Curtis and Sørensen dissimilarity indices, is listed in Table 4:

Further decomposition of the Sørensen index into turnover and nestedness components revealed noteworthy turnover. Sites 1 and 2 had the highest turnover for genus at 0.4295 and the lowest in phylum at 0.2083. Sites 3 and 4 had pronounced turnover, especially

TABLE 4 Beta diversity indexes.

Sites	ARGs		ARGCs		genus		phylum	
	Bray-Curtis Index	Sørensen Index	Bray-Curtis Index	Sørensen Index	Bray-Curtis Index	Sørensen Index	Bray-Curtis Index	Sørensen Index
Site 1 vs. Site 2	0.096	0.448	0.047	0.135	0.343	0.440	0.055	0.240
Site 3 vs. Site 4	0.072	0.340	0.022	0.172	0.335	0.391	0.075	0.125

FIGURE 9
Sørensen Dissimilarity decomposition into Turnover and Nestedness for ARGs, ARGCs, genus and phylum between sites close and far from the macrophyte for each small lake (Site 1 vs. Site 2 and Site 3 vs. Site 4).

FIGURE 10
Bray-Curtis Dissimilarity for ARGs, ARGCs, genus, and phylum between sites close and far from the macrophyte for each small lake. **(A)** Site 1 vs Site 2, **(B)** Site 3 vs Site 4.

in ARGs at 0.2917 and phylum at 0.0455. Nestedness was observed to a lesser extent, emphasizing the distinct ecological compositions between the small lake sites (Figure 9).

The inferential statistics were conducted to evaluate whether the diversity difference caused by the presence of macrophytes was significantly greater for one group compared to the others. In Abobral, only ARGC and phylum were not different from one another (Figure 10A). A similar pattern was noted in Aquidauana, where only ARGs and phylum showed no significant difference (Figure 10B).

4 Discussion

4.1 ARGs, ARGC and read numbers

Antibiotics and antibiotic resistance genes (ARGs) in freshwater environments are influenced by a myriad of factors, including soil type, macrophyte species, type of macrophyte, antibiotic concentration, water flow, microbiota, community dynamics, nutrients, pH, oxygenation, temperature, and the method of antibiotic introduction (Overton et al., 2023). While this study provides valuable insights into the resistome and taxonomy of the sampled sites, it is important to acknowledge that these biotic and abiotic parameters were not measured during the sample collection and could potentially explain the observed results.

In the present study a higher concentration of ARGs, ARGCs, and reads in a conserved area with no human activity compared to a farm region were found, suggesting that human activity may not always lead to an increase in ARGs in nearby bacterial communities. This could be potentially explained by the well-established fact that conserved areas have higher biodiversity in contrast to anthropized areas (McDonald et al., 2020; Glidden et al., 2021).

The significant presence of the RPOB2 gene in both Abobral and Aquidauana suggests the existence of a natural reservoir for this gene within the Pantanal region. The geographical extent of this reservoir warrants further investigation. This finding holds considerable interest for both public health and academic research.

The observed shift in the proportions of the ARGs RPOB2, BACA, UGD and their respective ARGCs within the Aquidauana may be attributed to a complex interplay of factors. While the small lake's proximity to an agricultural region suggests that farming practices could have influenced these shifts, it is important to question this assumption. The use of antibiotics, heavy metals, and other agrochemicals, which can infiltrate local water bodies through various means, could have contributed to the loss of genetic diversity in this region (Holt, 2000; Schmitt et al., 2015).

However, it is crucial to conduct further studies to determine whether the observed decrease in biodiversity is a permanent or transient phenomenon, possibly due to a recent change in environmental conditions or farming practices. A study demonstrated that the introduction of an antibiotic to the microbiological community was initially harmful. However, from the second to the fifth week, the levels of bacterial activity returned to levels similar to those found in communities not exposed to the antibiotic (Weber et al., 2011). This observation further underscores the need for ongoing monitoring and research to fully understand the dynamics at play.

There were unique ARGs and ARGCs found exclusively in certain locations, and these did not appear in isolation but were always found together, suggesting a possible relationship or co-dependence (Supplementary Figure S1). Moreover, the presence of the macrophyte influenced the occurrence of certain ARGs and ARGCs, demonstrating that their presence was more influenced by the macrophyte rather than the small lake environment.

One specific group of clinically important genes found in all samples were the mcr-N genes. These genes are significant because they offer resistance to colistin, an antibiotic used as a last resort against super-resistant bacteria. Although they have already been found in remote places, such as Antarctica, any occurrence of this group of genes is important to be reported (Cuadrat et al., 2020).

4.2 Diversity indexes

In this study, we used two different indices to measure alfa diversity: the Shannon index and the Simpson index. Overall the Shannon index identified Abobral small lake with the higher biodiversity compared to Aquidauana. However, for the Simpson index, it was the opposite. This might seem contradictory at first, however, this is due to the different aspects of biodiversity these indices measure.

The Shannon index is sensitive to species richness, meaning it increases with the number of different species present. Conversely, the Simpson index emphasizes species evenness, meaning it increases when a few species are significantly more prevalent than others (Guiaşu and Guiaşu, 2003).

A study on the effects of simulated nitrogen deposition on soil microbial community diversity in a coastal wetland found that with increasing levels of nitrogen deposition, alpha diversity (Shannon and Simpson indices) decreased significantly (Lu et al., 2021). This decrease in diversity may be due to soil acidification resulting from long-term nitrogen deposition in the supersaturated state. Long-term nitrogen deposition may also decrease the available organic matter of soil microorganisms, reducing microbial activity and diversity. Another study found that the presence of certain pollutants, such as chromium, can also reduce alpha diversity (Wei et al., 2023). Although we do not have data regarding nutrient and contaminant concentrations, bodies of water near farms often have high concentrations of these nutrients and heavy metals due to agricultural practices.

On the contrary, studies conducted in a wastewater treatment plant and different types of soil found that alpha diversity was higher in areas with more pollution and human intervention (Ndlovu et al., 2016; Geng et al., 2020). The addition of Fe^{2+} was also found to increase microbial diversity (Song et al., 2016). These findings suggest that while certain conditions and substances can decrease alpha diversity, others can lead to an increase, highlighting the complex interplay of factors that influence microbial diversity.

The results of beta diversity indexes indicate that the influence of macrophytes on the diversity of the small lake ecosystem varies across different categories. The genus category showed the highest dissimilarity values across both indices and all sites, indicating that the genus-level diversity is most affected by the proximity to the macrophyte. This is expected given that genus was the category with the highest diversity of all, allowing for a greater variation.

In contrast, the ARGCs and phylum categories showed relatively lower dissimilarity values. This suggests that these categories are less influenced by the proximity to the macrophyte compared to the genus and ARGs categories.

The dissimilarity values were generally higher for the Abobral small lake (Site 1 vs. Site 2) compared to Aquidauana's (Site 3 vs. Site 4) across all categories. This is also expected since the former exhibited greater diversity, allowing for more variation.

The decomposition of the Sørensen index into turnover and nestedness components provided a more nuanced understanding of the differences in microbial diversity between the sites. The results suggest that turnover is the dominant component of dissimilarity for all categories at both sites, indicating that the differences in diversity between the sites are primarily due to the replacement of species, rather than the presence of a subset of species at one site. In contrast, the nestedness component was relatively low for all categories, suggesting that the sites do not contain many species that are subsets of the species at other sites.

These findings emphasize the importance of considering multiple indexes when assessing biodiversity. Each index provides a different lens through which to view diversity, and together they offer a more comprehensive picture of the ecological structure of the sites.

4.3 Taxonomy

The abundance results found for the diversity of phyla and genera are consistent with the genetic diversity of the small lakes, with the Abobral small lake exhibiting greater diversity than Aquidauana.

The Pseudomonadota phylum (previous Proteobacteria), known for its high abundance in the majority of microbial communities, was, indeed, the most prevalent across all samples. This group plays a significant role in removing various pollutants, making it a crucial component of the microbial community. Its dominance in most systems can be attributed to its diverse metabolic capabilities and adaptability to different environmental conditions, which allow it to thrive in both polluted and unpolluted freshwater environments (Holt et al., 1984; Balows et al., 1992; Bergey, 1994; Topley, 2005).

In addition to its prevalence, this phylum of microorganisms plays a significant role in various environmental processes. For instance, the removal of antibiotics is largely attributed to the majority of functional microorganisms within this phylum (Alexandrino et al., 2017; Huang et al., 2017; Li et al., 2019; Shan et al., 2020). Additionally, the class Deltaproteobacteria, which is part of the Pseudomonadota phylum, contains most of the sulfate-reducing bacteria essential for heavy metal removal (Chen et al., 2021a; 2021b). The phylum is also involved in the removal of phosphorus, as indicated by research from (Si et al., 2018; 2019; Huang et al., 2020). Lastly, the role of the Pseudomonadota phylum in nitrogen removal from various wastewaters is well-documented, with genera such as Nitrosomonas, Nitrobacter, and Nitrosospira being associated with nitrification (Aguilar et al., 2019; Ajibade et al., 2021).

The Pseudomonas genus, a member of the Pseudomonadota phylum, is renowned for its adaptability and metabolic flexibility

(He et al., 2004; Høiby et al., 2010). This genus is found in various environments (Özen and Ussery, 2012), demonstrating its resilience and ability to thrive even in polluted areas. Its metabolic versatility allows it to play a crucial role in the removal of different pollutants. Pseudomonas exhibits a remarkable capability for environmental remediation. It effectively absorbs phosphorus from wastewater, storing it internally as polyphosphate (Tian et al., 2017; Huang et al., 2020). This process not only aids in the purification of wastewater but also contributes to the recycling of this essential nutrient.

In addition to phosphorus absorption, Pseudomonas shows resistance to heavy metals, aiding in their extraction from the environment. This is achieved through the synthesis of extracellular substances that bind to these metals, thereby preventing their spread within the biofilm and offering protection to the cells from stress (Teitzel and Parsek, 2003; Giovanella et al., 2017).

Pseudomonas also can metabolize glucose and mitigate sulfonamides through the co-metabolism of organic matter and sulfamethoxazole, contributing to antibiotic removal (Zheng et al., 2021). This unique metabolic capability further underscores the importance of Pseudomonas in environmental remediation and pollutant removal.

A slight increase in the abundance of Pseudomonas was observed in Site 3, which has plants nearby. Given that most pathogenic members of this genus are related to plants (Özen and Ussery, 2012), this relationship could be a pivotal factor in its distribution. The presence of plants and the specific environmental conditions at this site might have provided a competitive advantage for Pseudomonas, leading to its increased abundance.

On the other hand, the genus Polynucleobacter showed high abundance in the Abobral small lake and low abundance in the Aquidauana small lake, suggesting a higher susceptibility to pollution from the farm. Our results differ from a study that found a high abundance of Polynucleobacter in a river influenced by effluents from backyard aquacultures (Nakayama et al., 2017). It has also been found in polluted rivers and is known to live as chemoorganotrophs by utilizing low-molecular-weight substrates derived from the photooxidation of humic substances (Hahn et al., 2012; Ma et al., 2016).

Contrary to expectations, there was an increase in Bacteroidetes in the Aquidauana small lake. Although this phylum is more known for adapting well or even preferring polluted environments (Da Silva et al., 2015; Tai et al., 2020), pollution-sensitive Bacteroidetes have been observed, to function as bioindicators (Wolińska et al., 2017). This suggests that our environment might have conditions that are not unfavorable for this group.

The Actinobacteria phylum, which does not seem to be significantly influenced by the environmental conditions of the two regions studied, was found to be consistently present. Its genus, Streptomyces, was notably abundant across all four sites as well. This prevalence could be attributed to the fact that Streptomyces bacteria are the source of most antibiotics used in medicine, veterinary practice, and agriculture (Chater et al., 2010), making them efficient competitors in natural environments.

Site 3 showed a minor increase in Streptomyces abundance. This could be due to their close relationship with plants, as they are common in the rhizosphere and are frequent endophytes. Their ecological function in the natural decomposition of plant and fungi

cell walls, which are globally abundant, could also contribute to this increase. *Streptomyces* are known for their significant role in breaking down plant, fungi, and insect cell walls or surface components (Chater et al., 2010).

On the other hand, the genera *Mycobacterium* and *Mycolicibacterium* of this phylum, demonstrated a different pattern. While their abundance remained relatively stable and low in Sites 1, 2, and 3, a significant peak was observed in Site 4. This pattern can be attributed to the natural resistance of *Mycobacterium* species to most antimicrobial agents currently available (Nguyen and Pieters, 2009).

The absence of plants in Site 4, which could help eliminate antibiotics, might have exposed this site to more antibiotics, favoring the growth of these genera. In particular, the genus *Mycolicibacterium*, which was recently separated from *Mycobacterium*, indicating its close evolutionary proximity, may share similar characteristics, such as a higher resistance to antibiotics (Gupta et al., 2019).

Cyanobacteria, a crucial component of many aquatic ecosystems, were found in significant amounts in all samples, demonstrating their resilience and adaptability in diverse environments. Despite being common in aquatic environments, they were found in greater relative abundance in Aquidauana compared to Abobral, likely due to the higher availability of nutrients such as phosphorus and nitrogen.

Synechococcus, a genus of cyanobacteria, exhibited a unique distribution pattern. Its abundance was low in the Abobral small lake, with no instances at Site 1. Conversely, in the Aquidauana small lake, there was a marked increase in its abundance. According to our data, this genus seems to be more sensitive to nutrient availability variations, as indicated by its differing abundance in Aquidauana and Abobral.

The presence of *Synechococcus* has been linked to total nitrogen, dissolved nitrogen, dissolved organic carbon, and dissolved phosphorus (Le et al., 2022). Furthermore, a study by Pishbin et al. (2021) found that under mixotrophic conditions, *Synechococcus elongatus* could remove up to 85.1% of phosphorus and 87.4% of nitrogen. This nutrient removal efficiency, coupled with the likely high levels of phosphorus and nitrogen in the Aquidauana small lake due to its location in a farming region, could account for the observed distribution pattern of *Synechococcus*, and cyanobacteria in general.

5 Conclusion

In this study, we collected samples from two different small lakes. The first small lake is located in Abobral sub-region, which is in a protected reserve area, and the second one is in Aquidauana sub-region, characterized by farming activities. For each of these locations, we collected samples from areas close to and far from the floating macrophyte *Eichornia crassipes*, giving us a total of four samples.

In our successful endeavor to unravel these two areas' resistome, we were able to identify the primary antibiotic resistance genes and taxa, and how these vary depending on the level of pollution in the area and the presence or absence of plants nearby. We also identified a potential natural reservoir of the RPOB2 gene, as it occurred in high abundance in both areas. This finding is of academic, economic, and public health interest as it could influence decisions regarding the use of antibiotics in the area.

Our analysis of the collected data revealed a significant loss of both genetic and taxonomic biodiversity in the sample from the farm in Aquidauana sub-region when compared to the sample from the reserve in Abobral sub-region. This finding supports the widely accepted view that human activities can lead to a decrease in biodiversity.

However, our study also brought to light an interesting observation that contradicts a well-studied phenomenon. We found that human activity does not always result in an increase in the number of antibiotic resistance genes in the nearby bacterial community. This is one of the first studies to report such a finding, highlighting the need for further research in this area.

While the impact of human activity on the loss of genetic and taxonomic biodiversity is well-documented, our understanding of its effect on the diversity of resistance genes is still limited. More research is needed to investigate if human activity is causing a loss in diversity of resistance genes, and if so, whether certain genes are being favored over others. This will help us gain a more comprehensive understanding of the complex interactions between human activity and microbial communities.

Furthermore, we may have identified a potential natural reservoir for the RPOB2 gene, given its significant presence in both Aquidauana and Abobral. This discovery could facilitate informed decision-making regarding the use of antibiotics and public health.

Data availability statement

The raw sequences have been deposited in the NCBI Sequence Read Archive under the code PRJNA1078255.

Author contributions

AO: Investigation, Methodology, Writing–original draft, Writing–review and editing. BR: Investigation, Methodology, Writing–original draft. RJ: Investigation, Methodology, Writing–original draft. NP: Investigation, Methodology, Writing–original draft. AB: Investigation, Methodology, Writing–original draft. RP: Investigation, Methodology, Writing–original draft. NA: Writing–original draft. AD: Conceptualization, Methodology, Supervision, Writing–original draft, Writing–review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1352801/full#supplementary-material

References

Aguilar, L., Gallegos, Á., Arias, C. A., Ferrera, I., Sánchez, O., Rubio, R., et al. (2019). Microbial nitrate removal efficiency in groundwater polluted from agricultural activities with hybrid cork treatment wetlands. *Sci. total Environ.* 653, 723–734. doi:10.1016/j.scitotenv.2018.10.426

Ajibade, F. O., Wang, H.-C., Guadie, A., Ajibade, T. F., Fang, Y.-K., Sharif, H. M. A., et al. (2021). Total nitrogen removal in biochar amended non-aerated vertical flow constructed wetlands for secondary wastewater effluent with low C/N ratio: microbial community structure and dissolved organic carbon release conditions. *Bioresour. Technol.* 322, 124430. doi:10.1016/j.biortech.2020.124430

Alexandrino, D. A. M., Mucha, A. P., Almeida, C. M. R., Gao, W., Jia, Z., and Carvalho, M. F. (2017). Biodegradation of the veterinary antibiotics enrofloxacin and ceftiofur and associated microbial community dynamics. *Sci. Total Environ.* 581, 359–368. doi:10.1016/j.scitotenv.2016.12.141

Aminov, R. I., and Mackie, R. I. (2007). Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol. Lett.* 271, 147–161. doi:10.1111/j.1574-6968.2007.00757.x

Arango-Argoty, G., Garner, E., Pruden, A., Heath, L. S., Vikesland, P., and Zhang, L. (2018). DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome* 6, 23–15. doi:10.1186/s40168-018-0401-z

Assine, M. L. (2015). "Brazilian Pantanal: a large pristine tropical wetland," in *Landscapes and landforms of Brazil*, 135–146.

Balows, A., Trüper, H. G., Dworkin, M., Harder, W., and Schleifer, K.-H. (1992). *The prokaryotes: a handbook on the biology of bacteria: ecophysiology, isolation, identification, applications*. Cham: Springer.

Baquero, F., Martínez, J.-L., and Cantón, R. (2008). Antibiotics and antibiotic resistance in water environments. *Curr. Opin. Biotechnol.* 19, 260–265. doi:10.1016/j.copbio.2008.05.006

Bergey, D. H. (1994). *Bergey's manual of determinative bacteriology*. Pennsylvania, United States: Lippincott Williams & Wilkins.

Chater, K. F., Biró, S., Lee, K. J., Palmer, T., and Schrempf, H. (2010). The complex extracellular biology of Streptomyces. *FEMS Microbiol. Rev.* 34, 171–198. doi:10.1111/j.1574-6976.2009.00206.x

Chen, J., Deng, S., Jia, W., Li, X., and Chang, J. (2021a). Removal of multiple heavy metals from mining-impacted water by biochar-filled constructed wetlands: adsorption and biotic removal routes. *Bioresour. Technol.* 331, 125061. doi:10.1016/j.biortech.2021.125061

Chen, J., Li, X., Jia, W., Shen, S., Deng, S., Ji, B., et al. (2021b). Promotion of bioremediation performance in constructed wetland microcosms for acid mine drainage treatment by using organic substrates and supplementing domestic wastewater and plant litter broth. *J. Hazard Mater* 404, 124125. doi:10.1016/j.jhazmat.2020.124125

Cuadrat, R. R. C., Sorokina, M., Andrade, B. G., Goris, T., and Davila, A. M. R. (2020). Global ocean resistome revealed: exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples. *Gigascience* 9, giaa046. doi:10.1093/gigascience/giaa046

Da Silva, M. L. B., Cantão, M. E., Mezzari, M. P., Ma, J., and Nossa, C. W. (2015). Assessment of bacterial and archaeal community structure in swine wastewater treatment processes. *Microb. Ecol.* 70, 77–87. doi:10.1007/s00248-014-0537-8

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927–930. doi:10.1111/j.1654-1103.2003.tb02228.x

Doron, S., and Gorbach, S. L. (2008). Bacterial infections: overview. *Int. Encycl. Public Health* 273, 273–282. doi:10.1016/b978-012373960-5.00596-7

Ferrante, L., and Fearnside, P. M. (2022). Brazil's Pantanal threatened by livestock. *Science* 377, 720–721. doi:10.1126/science.ade0656

Fresia, P., Antelo, V., Salazar, C., Giménez, M., D'Alessandro, B., Afshinnekoo, E., et al. (2019). Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. *Microbiome* 7, 35–39. doi:10.1186/s40168-019-0648-z

Frost, I., Smith, W. P. J., Mitri, S., Millan, A. S., Davit, Y., Osborne, J. M., et al. (2018). Cooperation, competition and antibiotic resistance in bacterial colonies. *ISME J.* 12, 1582–1593. doi:10.1038/s41396-018-0090-4

Geng, X. D., Zhou, Y., Wang, C. Z., Yu, M. H., and Qian, J. L. (2020). BACTERIAL COMMUNITY STRUCTURE AND DIVERSITY IN THE SOIL OF THREE DIFFERENT LAND USE TYPES IN A COASTAL WETLAND. *Appl. Ecol. Environ. Res.* 18, 8131–8144. doi:10.15666/aeer/1806_81318144

Giovanella, P., Cabral, L., Costa, A. P., de Oliveira Camargo, F. A., Gianello, C., and Bento, F. M. (2017). Metal resistance mechanisms in Gram-negative bacteria and their potential to remove Hg in the presence of other metals. *Ecotoxicol. Environ. Saf.* 140, 162–169. doi:10.1016/j.ecoenv.2017.02.010

Glidden, C. K., Nova, N., Kain, M. P., Lagerstrom, K. M., Skinner, E. B., Mandle, L., et al. (2021). Human-mediated impacts on biodiversity and the consequences for zoonotic disease spillover. *Curr. Biol.* 31, R1342–R1361. doi:10.1016/j.cub.2021.08.070

Guiaşu, R. C., and Guiaşu, S. (2003). Conditional and weighted measures of ecological diversity. *Int. J. Uncertain. Fuzziness Knowledge-Based Syst.* 11, 283–300. doi:10.1142/s0218488503002089

Gupta, R. S., Lo, B., and Son, J. (2019). Corrigendum: phylogenomics and comparative genomic studies robustly support division of the genus Mycobacterium into an emended genus Mycobacterium and four novel genera. *Front. Microbiol.* 10, 714. doi:10.3389/fmicb.2019.00714

Hahn, M. W., Scheuerl, T., Jezberová, J., Koll, U., Jezbera, J., Šimek, K., et al. (2012). The passive yet successful way of planktonic life: genomic and experimental analysis of the ecology of a free-living Polynucleobacter population. *PLoS One* 7, e32772. doi:10.1371/journal.pone.0032772

Hassoun-Kheir, N., Stabholz, Y., Kreft, J.-U., De La Cruz, R., Romalde, J. L., Nesme, J., et al. (2020). Comparison of antibiotic-resistant bacteria and antibiotic resistance genes abundance in hospital and community wastewater: a systematic review. *Sci. Total Environ.* 743, 140804. doi:10.1016/j.scitotenv.2020.140804

He, J., Baldini, R. L., Déziel, E., Saucier, M., Zhang, Q., Liberati, N. T., et al. (2004). The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc. Natl. Acad. Sci.* 101, 2530–2535. doi:10.1073/pnas.0304622101

Høiby, N., Ciofu, O., and Bjarnsholt, T. (2010). *Pseudomonas aeruginosa* biofilms in cystic fibrosis. *Future Microbiol.* 5, 1663–1674. doi:10.2217/fmb.10.125

Holt, J. G., Krieg, N. R., Sneath, P. H. A., Staley, J. T., and Williams, S. T. (1984). *Bergey's manual of systematic bacteriology, vol. 1*. Baltimore: The Williams and Wilkins Co, 1–1388.

Holt, M. S. (2000). Sources of chemical contaminants and routes into the freshwater environment. *Food Chem. Toxicol.* 38, S21–S27. doi:10.1016/s0278-6915(99)00136-2

Huang, J., Xiao, J., Guo, Y., Guan, W., Cao, C., Yan, C., et al. (2020). Long-term effects of silver nanoparticles on performance of phosphorus removal in a laboratory-scale vertical flow constructed wetland. *J. Environ. Sci.* 87, 319–330. doi:10.1016/j.jes.2019.07.012

Huang, X., Zheng, J., Liu, C., Liu, L., Liu, Y., Fan, H., et al. (2017). Performance and bacterial community dynamics of vertical flow constructed wetlands during the treatment of antibiotics-enriched swine wastewater. *Chem. Eng. J.* 316, 727–735. doi:10.1016/j.cej.2017.02.029

Kurade, M. B., Ha, Y.-H., Xiong, J.-Q., Govindwar, S. P., Jang, M., and Jeon, B.-H. (2021). Phytoremediation as a green biotechnology tool for emerging environmental pollution: a step forward towards sustainable rehabilitation of the environment. *Chem. Eng. J.* 415, 129040. doi:10.1016/j.cej.2021.129040

Le, K. T. N., Maldonado, J. F. G., Goitom, E., Trigui, H., Terrat, Y., Nguyen, T. L., et al. (2022). Shotgun metagenomic sequencing to assess cyanobacterial community composition following coagulation of cyanobacterial blooms. *Toxins (Basel)* 14, 688. doi:10.3390/toxins14100688

Li, H., Liu, F., Luo, P., Chen, X., Chen, J., Huang, Z., et al. (2019). Stimulation of optimized influent C: N ratios on nitrogen removal in surface flow constructed wetlands: performance and microbial mechanisms. *Sci. Total Environ.* 694, 133575. doi:10.1016/j.scitotenv.2019.07.381

Lu, G., Xie, B., Cagle, G. A., Wang, X., Han, G., Wang, X., et al. (2021). Effects of simulated nitrogen deposition on soil microbial community diversity in coastal wetland of the Yellow River Delta. *Sci. Total Environ.* 757, 143825. doi:10.1016/j.scitotenv.2020.143825

Ma, L., Mao, G., Liu, J., Gao, G., Zou, C., Bartlam, M. G., et al. (2016). Spatial-temporal changes of bacterioplankton community along an exhorheic river. *Front. Microbiol.* 7, 250. doi:10.3389/fmicb.2016.00250

McDonald, R. I., Mansur, A. V., Ascensão, F., Colbert, M., Crossman, K., Elmqvist, T., et al. (2020). Research gaps in knowledge of the impact of urban growth on biodiversity. *Nat. Sustain* 3, 16–24. doi:10.1038/s41893-019-0436-6

Nakayama, T., Hoa, T. T. T., Harada, K., Warisaya, M., Asayama, M., Hinenoya, A., et al. (2017). Water metagenomic analysis reveals low bacterial diversity and the presence of antimicrobial residues and resistance genes in a river containing wastewater from backyard aquacultures in the Mekong Delta, Vietnam. *Environ. Pollut.* 222, 294–306. doi:10.1016/j.envpol.2016.12.041

Ndlovu, T., Khan, S., and Khan, W. (2016). Distribution and diversity of biosurfactant-producing bacteria in a wastewater treatment plant. *Environ. Sci. Pollut. Res.* 23, 9993–10004. doi:10.1007/s11356-016-6249-5

Nguyen, L., and Pieters, J. (2009). Mycobacterial subversion of chemotherapeutic reagents and host defense tactics: challenges in tuberculosis drug development. *Annu. Rev. Pharmacol. Toxicol.* 49, 427–453. doi:10.1146/annurev-pharmtox-061008-103123

Overton, O. C., Olson, L. H., Majumder, S. D., Shwiyyat, H., Foltz, M. E., and Nairn, R. W. (2023). Wetland removal mechanisms for emerging contaminants. *Land (Basel)* 12, 472. doi:10.3390/land12020472

Özen, A. I., and Ussery, D. W. (2012). Defining the Pseudomonas Genus: where do we draw the line with Azotobacter? *Microb. Ecol.* 63, 239–248. doi:10.1007/s00248-011-9914-8

Polińska, W., Kotowska, U., Kiejza, D., and Karpińska, J. (2021). Insights into the use of phytoremediation processes for the removal of organic micropollutants from water and wastewater; a review. *Water (Basel)* 13, 2065. doi:10.3390/w13152065

Preena, P. G., Swaminathan, T. R., Kumar, V. J. R., and Singh, I. S. B. (2020). Antimicrobial resistance in aquaculture: a crisis for concern. *Biol. Bratisl.* 75, 1497–1517. doi:10.2478/s11756-020-00456-4

Qian, X., Gu, J., Sun, W., Wang, X.-J., Su, J.-Q., and Stedfeld, R. (2018). Diversity, abundance, and persistence of antibiotic resistance genes in various types of animal manure following industrial composting. *J. Hazard Mater* 344, 716–722. doi:10.1016/j.jhazmat.2017.11.020

Raza, S., Shin, H., Hur, H.-G., and Unno, T. (2022). Higher abundance of core antimicrobial resistant genes in effluent from wastewater treatment plants. *Water Res.* 208, 117882. doi:10.1016/j.watres.2021.117882

Schmitt, N., Wanko, A., Laurent, J., Bois, P., Molle, P., and Mosé, R. (2015). Constructed wetlands treating stormwater from separate sewer networks in a residential Strasbourg urban catchment area: micropollutant removal and fate. *J. Environ. Chem. Eng.* 3, 2816–2824. doi:10.1016/j.jece.2015.10.008

Serwecińska, L. (2020). Antimicrobials and antibiotic-resistant bacteria: a risk to the environment and to public health. *Water (Basel)* 12, 3313. doi:10.3390/w12123313

Shan, A., Wang, W., Kang, K. J., Hou, D., Luo, J., Wang, G., et al. (2020). The removal of antibiotics in relation to a microbial community in an integrated constructed wetland for tail water decontamination. *Wetlands* 40, 993–1004. doi:10.1007/s13157-019-01262-8

Si, Z., Song, X., Wang, Y., Cao, X., Zhao, Y., Wang, B., et al. (2018). Intensified heterotrophic denitrification in constructed wetlands using four solid carbon sources: denitrification efficiency and bacterial community structure. *Bioresour. Technol.* 267, 416–425. doi:10.1016/j.biortech.2018.07.029

Si, Z., Wang, Y., Song, X., Cao, X., Zhang, X., and Sand, W. (2019). Mechanism and performance of trace metal removal by continuous-flow constructed wetlands coupled with a micro-electric field. *Water Res.* 164, 114937. doi:10.1016/j.watres.2019.114937

Song, X., Wang, S., Wang, Y., Zhao, Z., and Yan, D. (2016). Addition of Fe²⁺ increase nitrate removal in vertical subsurface flow constructed wetlands. *Ecol. Eng.* 91, 487–494. doi:10.1016/j.ecoleng.2016.03.013

Sun, D., Jeannot, K., Xiao, Y., and Knapp, C. W. (2019). Editorial: horizontal gene transfer mediated bacterial antibiotic resistance. *Front. Microbiol.* 10, 1933. doi:10.3389/fmicb.2019.01933

Tai, X., Li, R., Zhang, B., Yu, H., Kong, X., Bai, Z., et al. (2020). Pollution gradients altered the bacterial community composition and stochastic process of rural polluted ponds. *Microorganisms* 8, 311. doi:10.3390/microorganisms8020311

Teitzel, G. M., and Parsek, M. R. (2003). Heavy metal resistance of biofilm and planktonic *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* 69, 2313–2320. doi:10.1128/aem.69.4.2313-2320.2003

Tian, J., Yu, C., Liu, J., Ye, C., Zhou, X., and Chen, L. (2017). Performance of an ultraviolet Mutagenetic polyphosphate-accumulating bacterium PZ2 and its application for wastewater treatment in a newly designed constructed wetland. *Appl. Biochem. Biotechnol.* 181, 735–747. doi:10.1007/s12010-016-2245-y

Topley, W. W. C. (2005). *Topley and Wilson's microbiology and microbial infections.* London: Hodder Arnold.

Uritskiy, G. V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6, 1–13. doi:10.1186/s40168-018-0541-1

Weber, K. P., Mitzel, M. R., Slawson, R. M., and Legge, R. L. (2011). Effect of ciprofloxacin on microbiological development in wetland mesocosms. *Water Res.* 45, 3185–3196. doi:10.1016/j.watres.2011.03.042

Wei, Z., Sixi, Z., Baojing, G., Xiuqing, Y., Guodong, X., and Baichun, W. (2023). Effects of Cr stress on bacterial community structure and composition in rhizosphere soil of Iris tectorum under different cultivation modes. *Microbiol. Res. (Pavia)* 14, 243–261. doi:10.3390/microbiolres14010020

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon* 21, 213–251. doi:10.2307/1218190

Wolińska, A., Kuźniar, A., Zielenkiewicz, U., Izak, D., Szafranek-Nakonieczna, A., Banach, A., et al. (2017). Bacteroidetes as a sensitive biological indicator of agricultural soil usage revealed by a culture-independent approach. *Appl. Soil Ecol.* 119, 128–137. doi:10.1016/j.apsoil.2017.06.009

Zheng, Y., Liu, Y., Qu, M., Hao, M., Yang, D., Yang, Q., et al. (2021). Fate of an antibiotic and its effects on nitrogen transformation functional bacteria in integrated vertical flow constructed wetlands. *Chem. Eng. J.* 417, 129272. doi:10.1016/j.cej.2021.129272

Frontiers in
Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

frontiers

Frontiers in
Genetics

frontiers | Research Topics