

Nitrogen-rate effects on rhizosphere gene-enzyme coupling and maize yield in drip-irrigated maize

Kanghui Liu

Baoguo Yang

lkanghui0626@163.com

Ning Li

Zhi Chen

Fei Liang

Qinzheng Xu

Wenxin Ma

Baoqiang Li

Dongmei Zhao

Tongxu Liu

Hassan Waseem

Muqarrab Ali

Research Article

Keywords: Maize, Nitrogen fertilization, Functional genes, Soil enzymes, Gene-enzyme coupling, Nitrogen use efficiency

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10992392/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Nitrogen-rate effects on rhizosphere gene-enzyme coupling and maize yield in drip-irrigated maize

Kanghui Liu ^{a,b}, Baoguo Yang ^{a,b*}, Zhi Chen ^c, Ning Li ^d, Fei Liang ^a, Qinzheng Xu ^e, Wenxin Ma ^e,
Baoqiang Li ^e, Dongmei Zhao ^e, Tongxu Liu ^f, Waseem Hassan ^g, Muqarrab Ali ^h, Shengguang
Zhang ^a

^a Key Laboratory of Pollutant Chemistry and Environmental Treatment, School of Resources and Environment, Yili Normal University, Yining 835000, China

^b Xinjiang Key Laboratory of Clean Conversion and High Value Utilization of Biomass Resources, Yili Normal University, Yining 835000, China

^c Xinjiang Agricultural Vocational and Technical University, Changji 831100, China

^d College of Resources and Environment, Xinjiang Agricultural University, Urumqi 830052, China

^e Xinjiang Huier Agriculture Group Co., Ltd., Changji 831100, China

^f National-Regional Joint Engineering Research Center for Soil Pollution Control and Remediation in South China, Guangdong Key Laboratory of Integrated Agro-Environmental Pollution Control and Management, Institute of Eco-Environmental and Soil Sciences, Guangdong Academy of Sciences, Guangzhou 510650, China

^g Department of Soil and Environmental Sciences Muhammad Nawaz Shareef University of Agriculture, Multan, Pakistan

^h Department of Climate Change Muhammad Nawaz Shareef University of Agriculture, Multan, Pakistan

*Corresponding author: bgyang@cug.edu.cn (B. Yang)

Abstract

Background and aims Functional-gene abundance reflects microbial functional potential, whereas measured soil enzyme activity reflects realized biochemical activity under assay conditions; these components may respond asynchronously to nitrogen (N) supply. We examined how rhizosphere gene-enzyme matching varied across N rates and whether it was associated with maize (*Zea mays* L.) yield.

Methods A drip-irrigated field experiment comprised five N rates (0, 150, 225, 300, and 375 kg N ha⁻¹; three replicate plots per treatment). We measured maize yield, rhizosphere soil properties, bacterial community composition, functional-gene abundances involved in carbon (C), nitrogen, and phosphorus (P) cycling, and seven soil enzyme activities. A study-defined exploratory direct coupling index (DCI) provided a complementary within-sample summary of matching between predefined functional-gene markers and corresponding enzyme activities.

Results Grain yield was highest at 300 kg N ha⁻¹ (14,519.5 kg ha⁻¹; 52.6% above the control), with the highest agronomic nitrogen-use efficiency (ANUE; 16.68 kg grain kg⁻¹ N). At 375 kg N ha⁻¹, yield and ANUE declined by 9.5% and 42.1%, while NO₃⁻-N and electrical conductivity increased. At 300 kg N ha⁻¹, most measured enzymes showed higher activities, selected nutrient-cycling genes were comparatively abundant, and Integrated DCI was highest; it declined at 375 kg N ha⁻¹. Bacterial community composition also differed among N treatments.

Conclusion Among the tested rates, 300 kg N ha⁻¹ produced the highest yield and was associated with favorable rhizosphere nutrient conditions, higher enzyme

activities, selected gene responses, and stronger gene-enzyme matching rather than greater mineral-N accumulation alone.

Keywords: Maize, Nitrogen fertilization, Functional genes, Soil enzymes, Gene-enzyme coupling, Nitrogen use efficiency

Abbreviations

AK, available potassium; ANUE, agronomic nitrogen-use efficiency; AP, available phosphorus; DCI, direct coupling index; DOC, dissolved organic carbon; EC, electrical conductivity; FDR, false discovery rate; NMDS, non-metric multidimensional scaling; OOB, out-of-bag; PERMANOVA, permutational multivariate analysis of variance; PLS-SEM, partial least squares structural equation modeling; RMSE, root mean squared error; SOM, soil organic matter; TPM, transcripts per million.

1. Introduction

Maize is one of the major cereal crops in China and plays an important role in national grain production and livestock feed supply. Its high productivity is strongly dependent on nitrogen (N) inputs, particularly under intensive irrigated production systems. However, crop productivity and nitrogen-use efficiency (NUE) commonly show diminishing returns once N inputs exceed crop demand (Zhang et al. 2022). When N supply exceeds crop requirements, nitrate accumulation and soil electrical conductivity may increase without a proportional yield benefit (Mikha et al. 2019),

with potential consequences for long-term soil functioning. Clarifying the rhizosphere processes underlying the transformation and utilization of applied N is therefore important for understanding maize responses to N supply and improving fertilization strategies.

The rhizosphere harbors diverse microbial communities that contribute substantially to nutrient transformation and plant–soil interactions (Ling et al. 2022). Nitrogen input can alter maize root exudation and rhizosphere bacterial composition, thereby favoring different bacterial groups across N gradients (Zhu et al. 2016). In the present study, genus-level community analyses were restricted to bacteria, whereas fungal community responses were not evaluated separately; this taxonomic scope should therefore be considered when interpreting the microbial results. Long-term fertilization can shape the functional recruitment of maize rhizobacteria, with recruited taxa depending on soil context (Zhang et al. 2024). However, shifts in bacterial community composition alone do not demonstrate changes in realized soil function (Chen et al. 2022). Their ecological significance therefore needs to be interpreted together with microbial functional-gene potential and measured soil enzyme activity.

Functional genes provide a complementary representation of microbial genetic potential for nutrient transformation. To characterize major C-, N-, and P-cycling processes, the target genes were selected a priori based on previous rhizosphere and nutrient-cycling studies and their ecological or biochemical relevance to nutrient-cycling processes and, where applicable, to the measured enzyme activities,

rather than on their abundance or statistical significance in the present dataset. The C-cycling markers *bglB*, *sacA*, and *bglX* were used as indicators of carbohydrate transformation (Enebe and Babalola 2021a); the N-cycling markers *narG*, *napA*, and *ureC* were selected as indicators of nitrate reduction and urea hydrolysis (Enebe and Babalola 2021a; Zhao et al. 2022); and the P-cycling markers *phoD*, *phoA*, and *appA* represented alkaline-phosphatase- or phytase-mediated organic P mineralization (Dai et al. 2020; Enebe and Babalola 2021b). The measured soil enzyme activities considered in relation to these functional groups included sucrase (S-SC), urease (S-UE), nitrate reductase (S-NR), acid phosphatase (S-ACP), and alkaline phosphatase (S-AKP/ALP). Importantly, functional-gene abundance represents genetic potential rather than a direct measurement of process rates, and the strength of this relationship varies among systems (Rocca et al. 2015). Measured enzyme activity also depends on enzyme production and persistence, substrate availability, and soil environmental conditions (Burns et al. 2013; Schimel et al. 2017). Consequently, functional-gene abundance and corresponding enzyme activity need not respond synchronously to changes in N supply.

Empirical studies have evaluated associations between functional-gene abundance and corresponding enzyme activities across environmental gradients (Trivedi et al. 2016; Wang et al. 2023). Other work has compared predicted microbial functional capacity with measured biochemical activity to evaluate correspondence between functional potential and realized function (Breitkreuz et al. 2021). These approaches are valuable for identifying covariation and broader associations across

87 samples or environmental gradients, but few are designed to provide a sample-level
88 summary of the matching between a predefined functional-gene marker and its
89 corresponding measured enzyme activity. We therefore used a study-defined direct
90 coupling index (DCI) as a complementary exploratory metric that integrates the joint
91 magnitude of the paired measurements while reducing the index value when the two
92 components diverge.

93 To address this knowledge gap, we conducted a field experiment with five N
94 application rates (0, 150, 225, 300, and 375 kg N ha⁻¹) in a drip-irrigated maize system.
95 We characterized rhizosphere physicochemical properties, bacterial community
96 composition, C-, N-, and P-cycling functional-gene abundance, corresponding soil
97 enzyme activities, and study-defined gene-enzyme coupling indices. The DCI was
98 formulated as a complementary exploratory metric that assigns higher values when
99 normalized functional-gene abundance and the corresponding enzyme activity are
100 both high and closely matched. We hypothesized that: (1) an intermediate N rate
101 would be associated with greater coordination among rhizosphere nutrient status,
102 bacterial community composition, functional-gene potential, and enzyme activity than
103 low or high N inputs; (2) the highest N input would be associated with weaker
104 matching between functional-gene abundance and measured enzyme activity,
105 reflected in lower coupling indices; and (3) because N was the experimentally
106 manipulated nutrient, the N-cycling functional module would show more pronounced
107 associations within the soil–microbe–enzyme framework than the C- or P-cycling
108 modules.

2. Methods

2.1 Experimental site

The field experiment was conducted in 2025 in Ganggou Village, Yining City, Xinjiang, China (44.00 °N, 81.22 °E), and maize was monitored throughout the growing season. The experimental site has a temperate continental semi-arid climate, with an annual precipitation of 365.5 mm, mean annual evaporation of 1613.6 mm, mean annual sunshine duration of 2776.3 h, and a frost-free period of approximately 170 days. The soil is classified as gray desert soil, characterized by a sandy texture and low organic matter content. The experimental field was relatively uniform in topography, irrigation conditions, and prior management history, and no discrete spatial blocking factor was established before treatment allocation. Spring maize (*Zea mays* L.) was sown in mid-May and harvested in early September of the same year.

Before fertilizer application, initial topsoil samples were collected from the 0–20 cm soil layer. The initial soil contained 9.95 g kg⁻¹ soil organic matter (SOM), 37.8 mg kg⁻¹ hydrolyzable N, 3.01 mg kg⁻¹ available P (AP), and 93 mg kg⁻¹ available K (AK). Initial soil pH was 8.00, electrical conductivity (EC) was 15.4 mS m⁻¹, dissolved organic carbon (DOC) was 0.055 g kg⁻¹, NH₄⁺-N was 3.33 mg kg⁻¹, and NO₃⁻-N was 2.36 mg kg⁻¹. Detailed initial soil physicochemical properties are provided in Supplementary Table S14.

2.2 Experimental design

The field experiment was arranged in a completely randomized design with five nitrogen (N) fertilizer treatments and three replicate plots per treatment, yielding 15 independent experimental plots. The experimental field was relatively uniform in

topography, irrigation conditions, and prior management history, and no predefined spatial blocking factor was incorporated into the original field layout. Treatments were therefore assigned randomly across the field according to a completely randomized design. The subsequent statistical analysis followed this original randomization structure, with N treatment treated as the fixed factor and no block term. Each plot measured 100 m² (20 m × 5 m). Maize was cultivated using a mulched drip-irrigation system with a wide–narrow row configuration. Wide and narrow row spacings were 0.7 and 0.4 m, respectively, with an intra-row plant spacing of 0.17 m, corresponding to a planting density of 106,500 plants ha⁻¹. A single drip-irrigation tape was positioned in the center of each narrow row, 0.2 m from the adjacent maize rows, to supply water and nutrients.

Five total N application rates were established: 0 (CK), 150 (N22), 225 (N28), 300 (N34), and 375 (N40) kg N ha⁻¹. The N gradient was selected to span N omission, progressively increasing N supply, and a high-N treatment intended to represent potential over-fertilization under the local field conditions. The 375 kg N ha⁻¹ treatment was included as an intentionally high-N boundary treatment to test whether further N input would produce diminishing crop responses and altered rhizosphere conditions; it was not assumed a priori to represent a yield-reduction threshold. To isolate N supply as the primary experimental factor, phosphorus and potassium were applied uniformly to all plots at rates equivalent to 85.1 kg P ha⁻¹ and 137.0 kg K ha⁻¹, respectively. Irrigation and fertilizer applications were divided among maize growth stages to correspond with seasonal crop nutrient demand. Total irrigation was 5550 m³ ha⁻¹. The detailed water and fertilizer application schedule is provided in Table 1.

Table 1. Maize water and fertilizer management scheme

Growth stages	Irrigation quota (m ³ ha ⁻¹)	N rate (kg ha ⁻¹)					P rate (kg ha ⁻¹)	K rate (kg ha ⁻¹)
	I	N22	N28	N34	N40	CK		
Emergence	385	25	37.5	50	62.5	0	32.7	24.9
8-9 Leaf Stage	540	18.7	28.1	37.5	46.9	0	9.8	18.7
11-12 leaf stage	617	18.7	28.1	37.5	46.9	0	13.1	18.7
17-18 leaf stage	617	18.7	28.1	37.5	46.9	0	9.8	18.7
Silking	617	15.6	23.4	31.25	39.1	0	6.5	18.7
7 DAS	617	15.6	23.4	31.25	39.1	0	6.5	18.7
15 DAS	463	12.5	18.8	25	31.3	0	6.5	18.7
25 DAS	463	12.5	18.8	25	31.3	0	0	0
35 DAS	463	6.3	9.4	12.5	15.6	0	0	0
45 DAS	463	6.3	9.4	12.5	15.6	0	0	0
55 DAS	308	0.0	0.0	0.0	0.0	0	0	0
Total	5550	150.0	225.0	300.0	375.0	0	85.1	137.0

Note: N, nitrogen; P, phosphorus; K, potassium. All nutrient application rates are expressed on an elemental basis (kg ha⁻¹), DAS, days after silking.

2.3 Sampling and Measurement

2.3.1 Soil sampling and physicochemical analysis

Soil and plant samples were collected at the maize physiological maturity stage.

To obtain rhizosphere soil, five representative plants from each plot were carefully excavated along an “S”-shaped transect. The roots were not washed before rhizosphere soil collection. Roots were gently shaken by hand to remove loosely attached bulk soil, and soil tightly adhering to the root surfaces within approximately 4 mm of the roots was then collected using a sterile brush. Rhizosphere soil from the five plants in each plot was thoroughly mixed to form one composite sample per plot, resulting in three biological replicate samples for each N treatment. A separate quantitative repeatability test was not performed; however, the same rhizosphere separation and composite-sample preparation procedures were applied consistently

across all plots to reduce sampling bias. Composite samples were immediately passed through a 2-mm sieve to remove fine roots and debris and divided into three aliquots: one was stored at -80 °C for subsequent microbial DNA extraction, one was kept at 4 °C for fresh-soil analyses of available nutrients and enzyme activities, and the remaining soil was air-dried for basic physicochemical analyses.

Basic soil physicochemical properties were measured according to standard soil agrochemical procedures (Bao 2000). Soil pH and electrical conductivity (EC) were determined in soil–water suspensions, soil organic matter (SOM) by dichromate oxidation, available phosphorus (AP) colorimetrically, and hydrolyzable N by the alkali-hydrolysis diffusion method. Dissolved organic carbon (DOC) was measured using a TOC analyzer (TOC-L CPH), available potassium (AK) using inductively coupled plasma optical emission spectrometry (ICP-OES), and NH_4^+ -N and NO_3^- -N using UV–Vis spectrophotometry.

2.3.2 Soil enzyme activity assays

Activities of soil sucrase (S-SC), cellulase (S-CL), urease (S-UE), acid phosphatase (S-ACP), nitrate reductase (S-NR), dehydrogenase (S-DHA), and alkaline phosphatase (S-AKP/ALP) were determined in rhizosphere soil using commercial enzyme activity assay kits supplied by Suzhou Mengxi Biomedical Technology Co., Ltd. (Suzhou, China; www.michybio.com), according to the manufacturer's instructions and standard colorimetric soil enzyme assay principles (Tabatabai 1994). The catalogue numbers were M1408A for S-SC, M1405A for S-CL, M1401A for S-UE, M1409A for S-ACP, M1407A for S-NR, M1412A for S-DHA, and M1411A for S-AKP/ALP. The assays were based on reducing-sugar production for S-SC and S-CL, NH_4^+ -N release for S-UE, phenol release for S-ACP and S-AKP/ALP, nitrite formation for S-NR, and reduction of

2,3,5-triphenyltetrazolium chloride to triphenyl formazan for S-DHA. Detailed kit information and assay principles are provided in Supplementary Table S13.

2.3.3 Plant sampling and agronomic traits

Concurrently, five representative maize plants were selected from each plot for agronomic trait measurements. Plant height was measured from the soil surface to the tassel tip, and stem diameter was measured at the third internode. At physiological maturity, ear length, ear diameter, barren-tip length, kernel number per row, and row number per ear were determined using five representative ears from each plot, using a ruler or digital caliper, as appropriate. The sampled plants were then separated into stems, leaves, and ears. Plant tissues were heated at 105 °C for 30 min and subsequently dried at 80 °C to constant mass. After drying, stems, leaves, and ears were weighed separately and then finely ground. Total N concentration in each plant component was determined separately using the Kjeldahl digestion–distillation method. Shoot N concentration was calculated as the dry-mass-weighted mean of the N concentrations in stems, leaves, and ears and was expressed on a dry-mass basis (g kg^{-1}). Grain yield was determined by manually harvesting ears from the central rows of each plot, threshing the kernels, and expressing grain mass on an area basis (kg ha^{-1}) after adjustment to a standard grain moisture content of 14%.

2.3.4 Nitrogen use efficiency

The agronomic nitrogen use efficiency (ANUE) index was calculated as follows:
Agronomic Nitrogen Use Efficiency (ANUE, $\text{kg grain kg}^{-1} \text{ N}$) = (Yield in N-fertilized plot – Yield in control plot) / N application rate.

2.4 DNA Extraction and Metagenomic Sequencing

Total microbial genomic DNA samples were extracted using the OMEGA Mag-Bind Soil DNA Kit (M5635-02) (Omega Bio-Tek, Norcross, GA, USA),

following the manufacturer's instructions, and stored at -20 °C prior to further assessment. The quantity and quality of extracted DNA were assessed using a Qubit™ 4 Fluorometer, with WiFi: Q33238 (Qubit™ Assay Tubes: Q32856; Qubit™ 1X dsDNA HS Assay Kit: Q33231) (Invitrogen, USA) and agarose gel electrophoresis, respectively. The extracted microbial DNA was processed to construct metagenomic shotgun sequencing libraries with an insert size of 400 bp using the Illumina TruSeq Nano DNA LT Library Preparation Kit. Each library was sequenced by Illumina NovaSeq platform (Illumina, USA) using a PE150 strategy at Personal Biotechnology Co., Ltd. (Shanghai, China).

2.5 Metagenomic Analysis

Raw sequencing reads were quality-filtered using fastp (v0.23.2; parameters: -l 50 -g -W 5 -5 -q 20 -u 30). Taxonomic annotation of high-quality reads was performed using Kraken2 v2.0.8-beta with paired-end mode and a confidence threshold of 0.5 against a locally constructed Kraken2 database based on the NCBI nucleotide (nt) database downloaded in June 2024. According to the sequencing-provider response, no host-removal step was applied. High-quality reads were error-corrected using BBCMS from BBTools v38.98 (tossjunk=t, ecc=f, mincount=2, hcf=0.4, bits=4) and assembled with MEGAHIT v1.1.2 using --k-list 33,55,77,99,127 and --min-contig-len 300. Contigs shorter than 300 bp were excluded before downstream analysis, consistent with a previously reported metagenomic workflow (Hu et al. 2024). Protein-coding genes were predicted from retained contigs using Prodigal v2.6.3 with the -p meta option. Predicted protein sequences were

functionally annotated using MMseqs v13.45111 (parameters: -s 5.7 --max-seq-len 100000 --max-accept 1 --max-seqs 100) against KEGG (20200920), eggNOG (v5.0.0), GO (20221015), CAZy (07292021), NCBI-nr (2022-11-12), and Swiss-Prot (20221015). KOBAS/KOBAS-i was not used in the present analysis; KEGG annotation was performed solely using MMseqs2 against the specified KEGG database release. Gene abundance was estimated by aligning effective reads to the non-redundant gene catalogue using Strobealign v0.12.0, sorting BAM files with samtools v1.8, and summarizing gene-level read counts with featureCounts v2.0.3. Gene abundance was normalized as transcripts per million (TPM), which accounts for both sequencing depth and gene length. Detailed bioinformatic procedures, quality-control summaries, database information, and provider-confirmed version details are provided in Supplementary Method S1. Sequencing quality metrics and clean-data summaries are reported in Supplementary Tables S1-S3.

Functional genes used for the gene indices and DCI calculations were selected before statistical analysis based on published evidence for their roles in C, N, and P cycling and on their biochemical correspondence with the measured enzyme assays. This selection was not based on treatment significance, Random Forest ranking, or relative abundance in the present dataset. For direct pairwise coupling, we retained only gene-enzyme pairs with an interpretable process-level correspondence: *sacA*-S-SC, *ureC*-S-UE, *narG*-S-NR, *napA*-S-NR, *appA*-S-ACP, *phoA*-S-AKP/ALP, and *phoD*-S-AKP/ALP. The C-cycling markers *bglB* and *bglX* were retained in the broader C functional gene index but were not used as direct cellulase-coupling pairs

because their correspondence with the S-CL assay was less direct at the level of gene annotation and assay specificity.

2.6 Analytical Methods

All statistical analyses were conducted in R version 4.3.2. Data are presented as the mean $\pm$ standard deviation of three biological replicates. Treatment effects on maize agronomic traits, grain yield, soil physicochemical properties, enzyme activities, functional gene indices, and gene-enzyme coupling indices were evaluated using one-way analysis of variance (ANOVA), with nitrogen treatment as the main factor, because the field trial followed a completely randomized design without a block term. When significant treatment effects were detected, mean comparisons were performed using Tukey's HSD test at $P < 0.05$. Functional gene and enzyme indices were calculated from sample-level standardized values.

For the direct coupling analysis, each functional gene and its corresponding enzyme activity were normalized to the 0-1 range across all 15 samples. The direct coupling index (DCI) for each gene-enzyme pair was calculated as $DCI = \sqrt{(G_{norm} \times E_{norm}) \times (1 - |G_{norm} - E_{norm}|)}$, where G_{norm} and E_{norm} represent min-max-normalized gene abundance and enzyme activity, respectively. The geometric-mean term assigns higher values to pairs in which both components are high, whereas the mismatch term decreases the index when gene potential and enzyme activity diverge. This construction differs from a simple gene-enzyme correlation because it is calculated for each sample and gives low values when only one component is high. Module-level C DCI, N DCI, P DCI, and Integrated DCI were

then calculated from the predefined gene-enzyme pairs. Because DCI is a study-defined exploratory index, it was interpreted together with the original gene abundance, enzyme activity, correlation analysis, and robustness checks. These checks compared DCI with alternative matching metrics, including the geometric mean only, product, minimum value, harmonic mean, and agreement-only index; evaluated treatment-rank consistency under alternative normalization approaches; repeated the calculation using leave-one-out sample omission; and tested sensitivity to small perturbations of G_{norm} and E_{norm} . The functional-gene groupings and predefined gene-enzyme pairs are listed in Supplementary Tables S4 and S5, and the robustness checks are summarized in Supplementary Table S6.

Bacterial alpha and beta diversity analyses, Spearman correlation analysis, exploratory Random Forest regression, and exploratory PLS-SEM were used to examine associations among nitrogen input, rhizosphere properties, bacterial and functional attributes, gene-enzyme coupling, and maize productivity. For the cross-module Spearman correlation analysis, P values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure. Random Forest analysis was used only as an exploratory relative-importance analysis because of the limited sample size: 15 samples for grain yield and 12 fertilized samples for ANUE. PLS-SEM was interpreted as an exploratory, hypothesis-guided path model rather than a confirmatory causal model. Diagnostic outputs, including Random Forest stability summaries, R^2 values, VIF values, and bootstrap confidence intervals for indirect and total effects, are reported in the supplementary materials where applicable.

The detailed index calculations, multivariate parameters, Random Forest settings, and PLS-SEM diagnostics are described in Supplementary Method S2.

3. Results

3.1 Effects of Nitrogen Fertilizer Gradients on Maize Yield and Related Traits

Nitrogen fertilizer application significantly affected maize grain yield, shoot N concentration, and agronomic nitrogen use efficiency (ANUE) (Table 2). Grain yield increased progressively from CK to N34 and then declined at the highest N rate. Among the five tested treatments, N34 produced the highest grain yield (14,519.50 kg ha⁻¹), representing a 52.6% increase relative to CK. Shoot N concentration increased with N application up to N34, reaching 53.83 g kg⁻¹, and remained statistically comparable under N40. N34 also showed the highest ANUE (16.68 kg grain kg⁻¹ N). Increasing N from 300 to 375 kg N ha⁻¹ did not further increase yield; instead, grain yield and ANUE declined by 9.5% and 42.1%, respectively, under N40 compared with N34.

Several ear-related traits differed among N treatments (Table 2). Relative to CK, N34 increased ear length and ear diameter by 30.5% and 24.1%, respectively, and increased kernel number per row by 53.9%. N34 also had the numerically highest number of rows per ear, although ear row number did not differ significantly among treatments, whereas barren-tip length was 56.6% lower than that under CK. These trait differences accompanied the highest grain yield observed under N34.

Table 2. Effects of Nitrogen Fertilizer Gradients on Maize Yield and Related Traits

Treat ment	Ear length (cm)	Ear diamete r (mm)	Barren tip length (mm)	Grains per row	Ear rows	Grain yield (kg ha ⁻¹)	Shoot N (g kg ⁻¹)	ANUE (kg grain kg ⁻¹ N)
CK	17.00±1. 05b	39.88± 1.46c	38.14± 5.14a	30.37± 2.72d	14.03±0 .72a	9517.00±86. 24d	25.75±0 .57d	--
N22	18.33±0. 93b	43.91± 0.78b	24.47± 1.76b	35.53± 5.21c	14.30±0 .30a	10969.50±2 13.00c	39.62±0 .82c	9.68±1. 42c
N28	21.31±0. 84a	47.93± 1.72a	18.03± 3.33b	42.77± 1.25b	15.30±1 .54a	12438.83±3 57.90b	49.77±1 .38b	12.99± 1.59b
N34	22.18±0. 95a	49.49± 1.05a	16.57± 2.04c	46.73±1. 85a	15.73±0. 46a	14519.50±2 21.70a	53.83± 0.75a	16.68± 0.74a
N40	21.92±1. 01a	48.91± 1.47a	17.36± 1.38c	45.60±2. 45a	14.87±0. 31a	13135.00±3 74.02b	53.04± 0.74a	9.65±1. 00c

Note: Values are expressed as means ± SD (n = 3). Within the same column, means followed by different lowercase letters are significantly different according to Tukey's HSD test at $P < 0.05$. Shoot N, dry-mass-weighted shoot N concentration; ANUE, agronomic nitrogen use efficiency.

3.2 Effects of Nitrogen Fertilizer Gradients on Soil Physicochemical Properties

Significant treatment differences were observed in several rhizosphere soil physicochemical properties (Table 3). Soil pH varied within a relatively narrow range (8.50–8.72), with N22 showing a significantly higher value than the other treatments. AK was highest under N34 (247.67 mg kg⁻¹), representing a 68.8% increase relative to CK. Among the measured N properties, NH₄⁺-N reached its highest concentration under N34 (4.75 mg kg⁻¹), 114.9% higher than CK, whereas NO₃⁻-N was highest under N40 (78.37 mg kg⁻¹), corresponding to a 165.1% increase relative to CK. Hydrolyzable N also reached its highest numerical value under N34 (121.67 mg kg⁻¹; 33.6% above CK), although this increase was not significant relative to CK. DOC and SOM were highest under N34, increasing by 33.3% and 26.7%, respectively, relative to CK. AP showed relatively high values under N22, N34, and N40, with the maximum observed under N40 (20.70 mg kg⁻¹; 45.0% above CK), although these treatments did not differ significantly from CK. EC was significantly higher under N34 and N40 than under CK and reached its highest value under N40 (27.37 mS m⁻¹; 29.3% above CK).

Table 3. Effects of Nitrogen Fertilizer Gradients on Soil Physicochemical Properties

Treatm ent	pH	AK (mg kg ⁻¹)	DOC (g kg ⁻¹)	NO ₃ ⁻ - N (mg kg ⁻¹)	Hydrol yzable N (mg kg ⁻¹)	AP (mg kg ⁻¹)	NH ₄ ⁺ - N (mg kg ⁻¹)	SOM (g kg ⁻¹)	EC (mS m ⁻¹)
CK	8.59 ± 0.06b	146.67 ± 1.20c	0.021 ±0.001 b	29.57 ± 5.06b	91.10 ± 12.04a b	14.28 ± 5.16ab	2.21 ± 0.13b	13.87 ± 0.56b	21.17 ± 0.73b
N22	8.72 ± 0.08a	234.33 ± 59.22a bc	0.027 ± 0.005a b	33.53 ± 2.56b	104.07 ± 6.51a	20.33 ± 4.21a	2.73 ± 0.19b	13.77 ± 0.38b	21.90 ± 0.31b
N28	8.56 ± 0.08b	206.67 ± 37.25b c	0.022 ± 0.002b	28.20 ± 5.35b	78.00 ± 6.27b	8.62 ± 1.46b	3.34 ± 0.12b	13.83 ± 0.29b	20.47 ± 0.75b
N34	8.57 ± 0.04b	247.67 ± 19.32a b	0.028 ± 0.003a b	38.80 ± 0.38b	121.67 ± 1.33a	20.20 ± 0.15a	4.75 ± 0.58a	17.57 ± 0.15a	24.93 ± 0.22a
N40	8.50 ± 0.03b	230.67 ± 58.70a bc	0.024 ± 0.002a b	78.37 ± 0.65a	95.30 ± 0.98ab	20.70 ± 0.12a	2.93 ± 0.76b	14.63 ± 0.20b	27.37 ± 0.35a

Note: Values are expressed as mean ± SD (n = 3). Different lowercase letters within the same column indicate significant differences among treatments at $P < 0.05$ according to Tukey's HSD test. AK, available potassium; DOC, dissolved organic carbon; NO₃⁻-N, nitrate nitrogen; AP, available phosphorus; NH₄⁺-N, ammonium nitrogen; SOM, soil organic matter; EC, electrical conductivity.

3.3 Effects of Nitrogen Gradients on Rhizosphere Bacterial Community

Composition

3.3.1 Effects of Nitrogen Fertilizer Gradients on Rhizosphere Bacterial Community Alpha and Beta Diversity

Rhizosphere bacterial alpha-diversity indices differed in their responses among N treatments (Fig. 1). Chao1 and ACE showed no significant differences among treatments ($P > 0.05$), whereas Simpson and Shannon differed significantly among

360 treatments. CK showed the highest values of both diversity indices. Among the
361 fertilized treatments, N34 showed comparatively high Simpson and Shannon values
362 and did not differ significantly from CK, whereas N22 and N28 showed lower values;
363 N40 showed intermediate values that were significantly lower than CK but did not
364 differ significantly from N34. Thus, bacterial richness indices remained relatively
365 stable across treatments, whereas Shannon and Simpson diversity showed significant
366 treatment-related variation.

367 Non-metric multidimensional scaling (NMDS) based on Bray – Curtis
368 dissimilarity revealed distinct treatment-related differences in bacterial community
369 composition (Fig. 1e; stress = 0.027). Permutational multivariate analysis of variance
370 (PERMANOVA) indicated significant differences in bacterial community
371 composition among N treatments ($R^2 = 0.686$, $P = 0.001$). CK and N22 were clearly
372 separated from the higher-N treatments, whereas N28, N34, and N40 were positioned
373 more closely in ordination space.

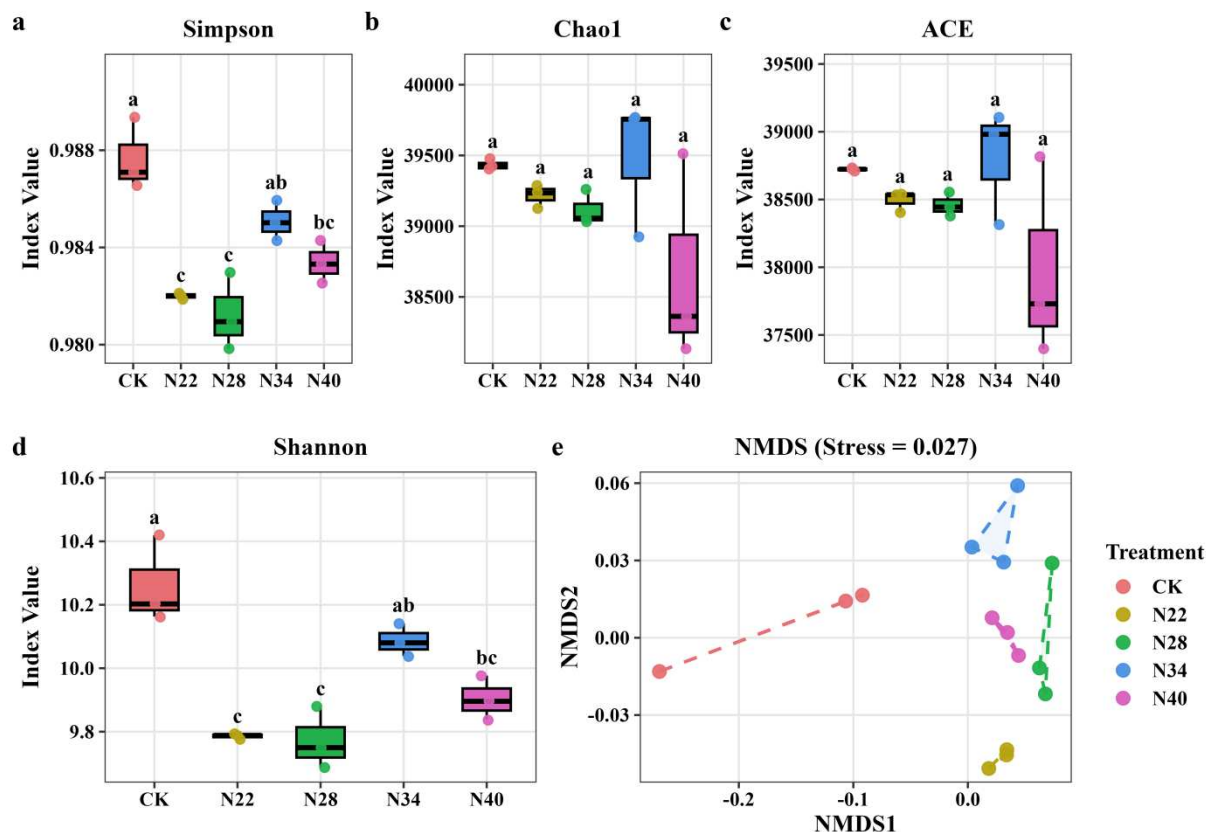


Fig. 1 Rhizosphere bacterial alpha and beta diversity under nitrogen fertilizer gradients

Note: Boxplots show alpha diversity indices among CK, N22, N28, N34, and N40, including (a) Simpson, (b) Chao1, (c) abundance-based coverage estimator (ACE), and (d) Shannon indices. The central line indicates the median, box limits indicate the interquartile range (IQR), whiskers extend to $1.5 \times \text{IQR}$, and dots represent biological replicates. Different lowercase letters indicate significant differences among treatments based on one-way ANOVA followed by Tukey's HSD test ($P < 0.05$). (e) Non-metric multidimensional scaling (NMDS) based on Bray-Curtis dissimilarity shows beta diversity of bacterial community structure. Dashed convex hulls enclose samples from the same treatment. PERMANOVA indicated significant differences among N treatments ($R^2 = 0.686$, $P = 0.001$). NMDS axes and alpha diversity indices are dimensionless; the stress value is shown in the panel title.

3.3.2 Effects of Nitrogen Fertilizer Gradients on Rhizosphere Bacterial Community Composition

Descriptive analysis of the top 20 bacterial genera showed treatment-related variation in rhizosphere bacterial community composition across the N gradient (Fig.

391 2). CK was characterized by comparatively higher relative abundances of
392 *Nocardioides*, *Streptomyces*, *Pseudonocardia*, *Kribbella*, *Lentzea*, and *Devosia* than
393 the N-fertilized treatments. N22 showed relatively higher abundances of
394 *Bradyrhizobium* and *Steroidobacter*, whereas N28 was characterized by higher
395 abundances of *Blastococcus*, *Luteitalea*, and *Mesorhizobium* relative to CK. N34
396 showed comparatively high abundances of *Arthrobacter*, *Agromyces*, *Ramlibacter*,
397 *Microvirga*, *Mesorhizobium*, and *Solirubrobacter*. Under N40, the relative abundances
398 of several genera that were prominent under N34, particularly *Ramlibacter* and
399 *Microvirga*, were lower than those under N34. Overall, the dominant genus-level
400 composition showed descriptive variation among N treatments, with several bacterial
401 genera reaching comparatively high relative abundances under N34.

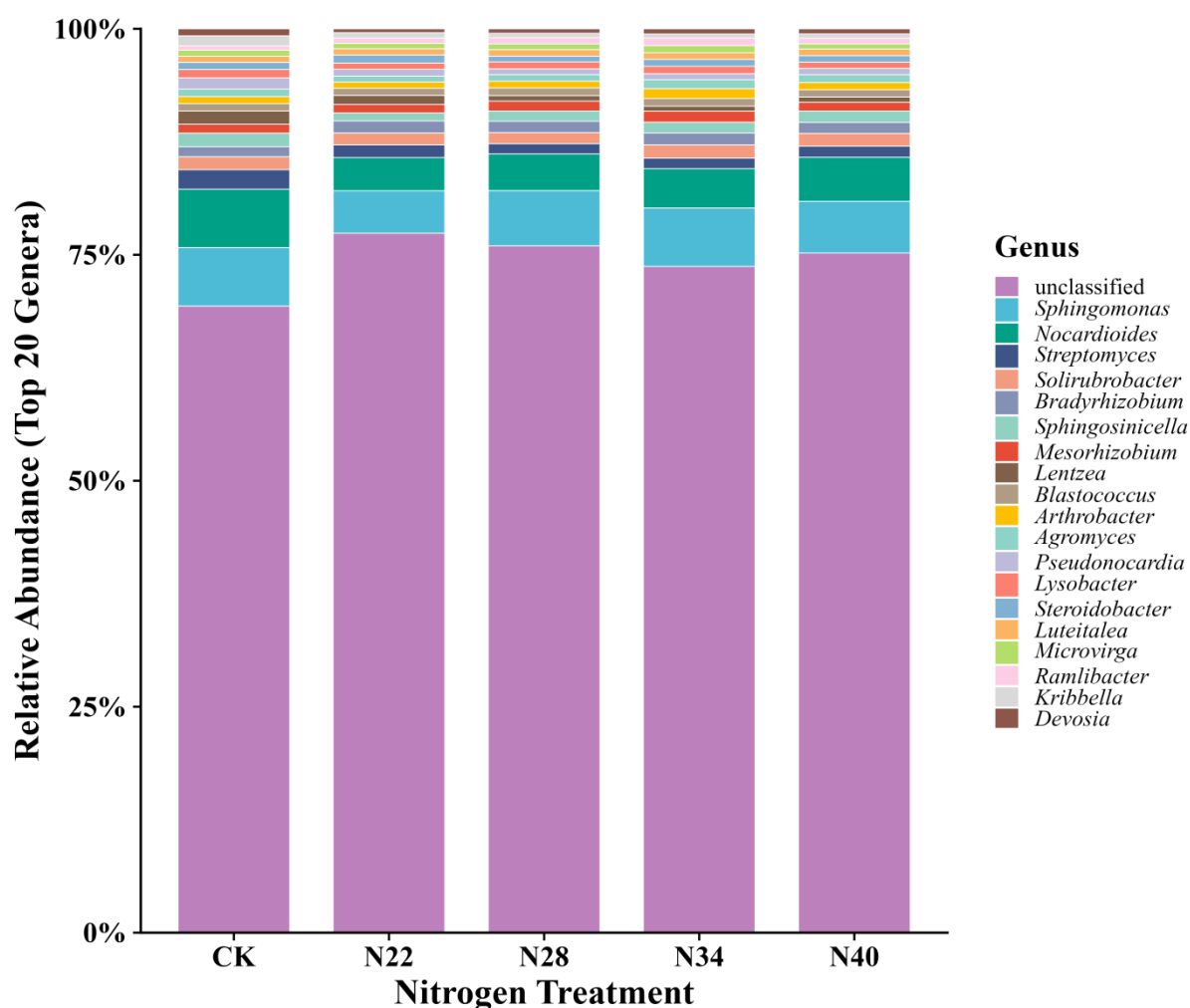


Fig. 2 Relative abundance of dominant bacterial genera under nitrogen fertilizer gradients

Note: Bars show the relative abundance (%) of the top 20 dominant bacterial taxonomic categories at the genus level. The displayed proportions were scaled to 100% within these 20 categories to emphasize compositional shifts among nitrogen treatments. Genus names are shown in italics.

3.4 Effects of Nitrogen Fertilizer Gradients on Rhizosphere Enzyme Activities

Significant treatment differences were observed for six of the seven measured rhizosphere enzyme activities, whereas S-CL did not differ significantly among treatments (Fig. 3). S-SC increased from CK and N22 to N28, reached its highest activity under N34, and declined under N40. S-UE, S-ACP, S-DHA, and S-AKP/ALP showed similar responses, with significantly higher activities under N34 than under

the other treatments. S-NR also reached its highest numerical value under N34; however, N34 did not differ significantly from N22 or N28, whereas both CK and N40 showed lower values. Overall, N34 showed the highest numerical activities for six of the seven measured enzymes, whereas these activities were lower under N40.

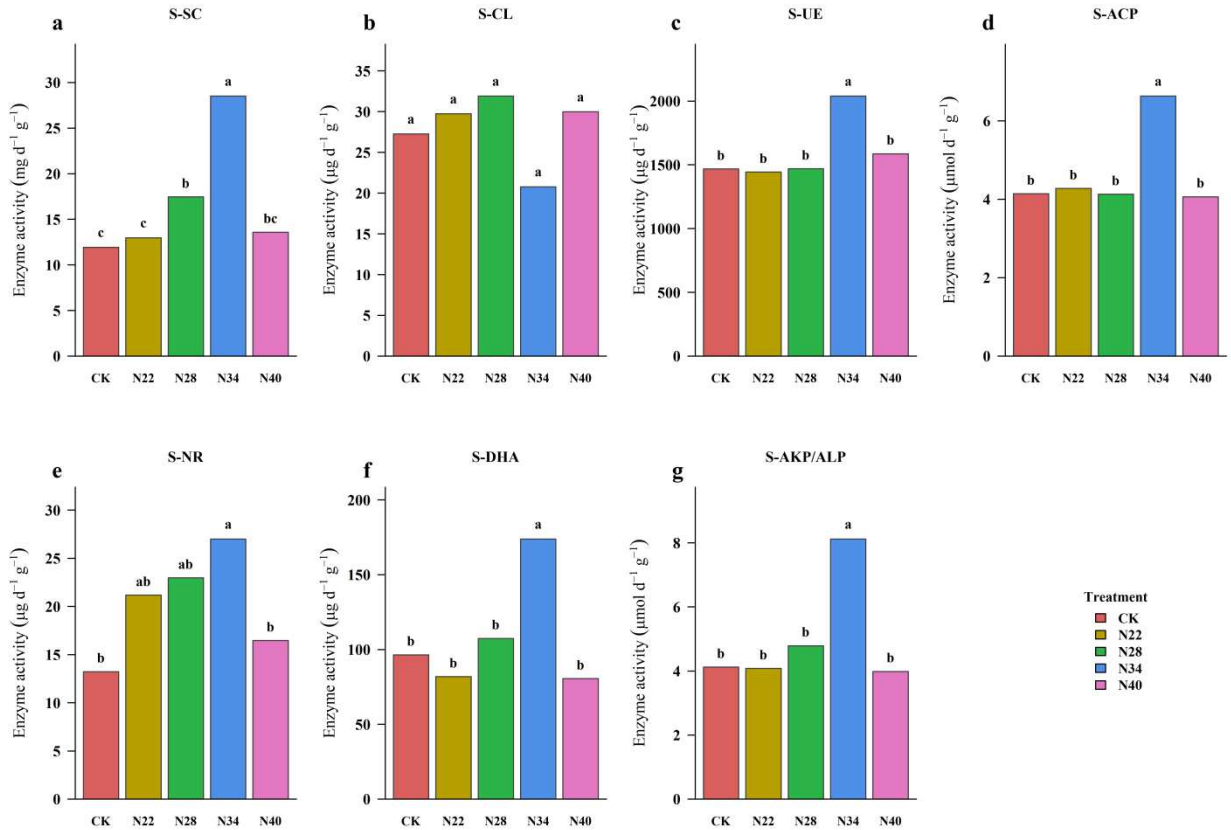


Fig. 3 Rhizosphere enzyme activities under nitrogen fertilizer gradients
Note: Bars show treatment means (n = 3) for maize rhizosphere soil enzyme activities. Different lowercase letters above bars indicate significant differences among treatments at $P < 0.05$ based on one-way ANOVA followed by Tukey's HSD test. S-SC, sucrase (mg d⁻¹ g⁻¹); S-CL, cellulase (μg d⁻¹ g⁻¹); S-UE, urease (μg d⁻¹ g⁻¹); S-ACP, acid phosphatase (μmol d⁻¹ g⁻¹); S-NR, nitrate reductase (μg d⁻¹ g⁻¹); S-DHA, dehydrogenase (μg d⁻¹ g⁻¹); S-AKP/ALP, alkaline phosphatase (μmol d⁻¹ g⁻¹).

3.5 Responses of C-, N-, and P-cycling functional genes to nitrogen fertilizer gradients

Standardized profiles of rhizosphere functional genes involved in C, N, and P

cycling varied among N treatments (Fig. 4a). N34 showed comparatively high standardized values for several genes, including *sacA* and *bglX* in the C-cycling group, *napA* in the N-cycling group, and *phoD* and *appA* in the P-cycling group. In contrast, N28 showed relatively low standardized values for several of these genes. N40 retained relatively high values for some N-cycling genes, particularly *narG* and *napA*, but showed less consistent responses among C- and P-cycling genes.

Functional-gene indices further summarized these treatment responses (Fig. 4b). The C-cycling and N-cycling indices showed no significant differences among treatments, although the N-cycling index was numerically higher under N34 and N40. The P-cycling and integrated indices were highest under N34 and lowest under N28, with N34 significantly higher than N28; CK, N22, and N40 showed intermediate values. These results indicate that the strongest treatment-level differences in the composite functional-gene indices occurred primarily in the P-cycling and integrated responses.

CV-based variability analysis further showed that individual functional genes differed in their responses to the N gradient (Fig. 4c). *appA* showed the highest between-treatment variability, followed by *sacA*, *napA*, and *narG*, whereas *bglX*, *phoD*, *phoA*, and *ureC* showed comparatively lower variability. Thus, the magnitude of response to N fertilization differed among individual functional genes and among the C-, N-, and P-cycling groups. The functional-gene groupings used to calculate the C-, N-, P-, and integrated indices are provided in Supplementary Table S4.

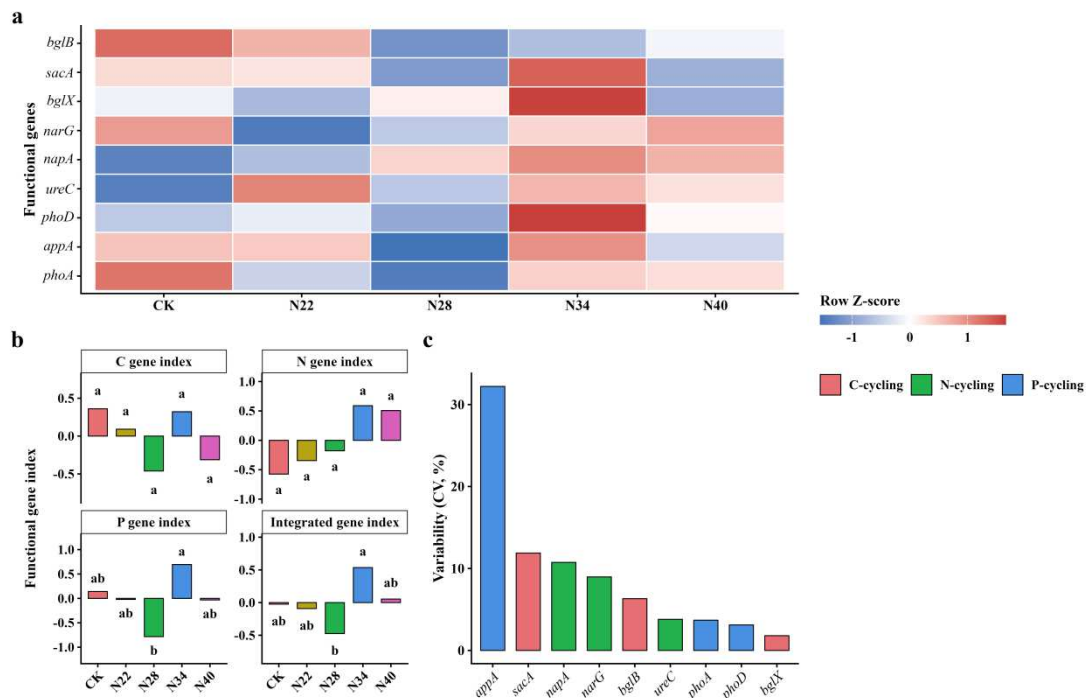


Fig. 4 Responses of C-, N-, and P-cycling functional genes to nitrogen fertilizer gradients

Note: (a) Heatmap showing row-standardized Z-scores of treatment-mean functional gene abundance grouped by C-, N-, and P-cycling processes. Functional gene abundance was normalized using TPM before downstream index calculations. Red and blue colors indicate higher and lower relative abundance, respectively; the color scale is unitless. (b) C-, N-, P-, and integrated functional gene indices calculated from Z-score-standardized gene abundance. Different lowercase letters indicate significant differences among treatments within the same index at $P < 0.05$ based on one-way ANOVA followed by Tukey's HSD test. (c) Variability of individual functional genes across nitrogen fertilizer treatments, expressed as the coefficient of variation (CV, %) of treatment-mean gene abundance.

3.6 Associations of Functional Genes with Soil Properties, Enzyme Activities, and Bacterial Genera

Spearman correlation analysis based on sample-level data ($n = 15$) showed varying associations between functional genes and rhizosphere soil properties, enzyme activities, and bacterial genera (Fig. 5). Across the complete set of 357 cross-module tests, nine associations remained significant after Benjamini–Hochberg FDR correction (FDR-adjusted $P < 0.05$; Supplementary Table S16). Among the gene-centered associations displayed in Fig. 5, two remained significant after FDR

470 correction. *Agromyces* was positively correlated with *narG* ($\rho = 0.782$; FDR-adjusted
471 $P = 0.0320$), whereas S-CL was negatively correlated with *phoD* ($\rho = -0.779$;
472 FDR-adjusted $P = 0.0320$).

473 Several additional gene-centered relationships showed relatively high absolute
474 correlation coefficients, including those of *Ramlibacter* with *napA*, S-SC with *napA*,
475 *Microvirga* with *sacA*, SOM with *bglX*, and NO_3^- -N with *napA*. However, none of
476 these relationships remained significant after FDR correction. Across the complete set
477 of cross-module tests, additional FDR-significant correlations were observed between
478 S-SC and *Ramlibacter*, S-DHA and *Mesorhizobium*, S-DHA and *Microvirga*, S-SC
479 and *Mesorhizobium*, NH_4^+ -N and S-SC, hydrolyzable N and S-ACP, and S-NR and
480 *Ramlibacter* (Supplementary Table S16).

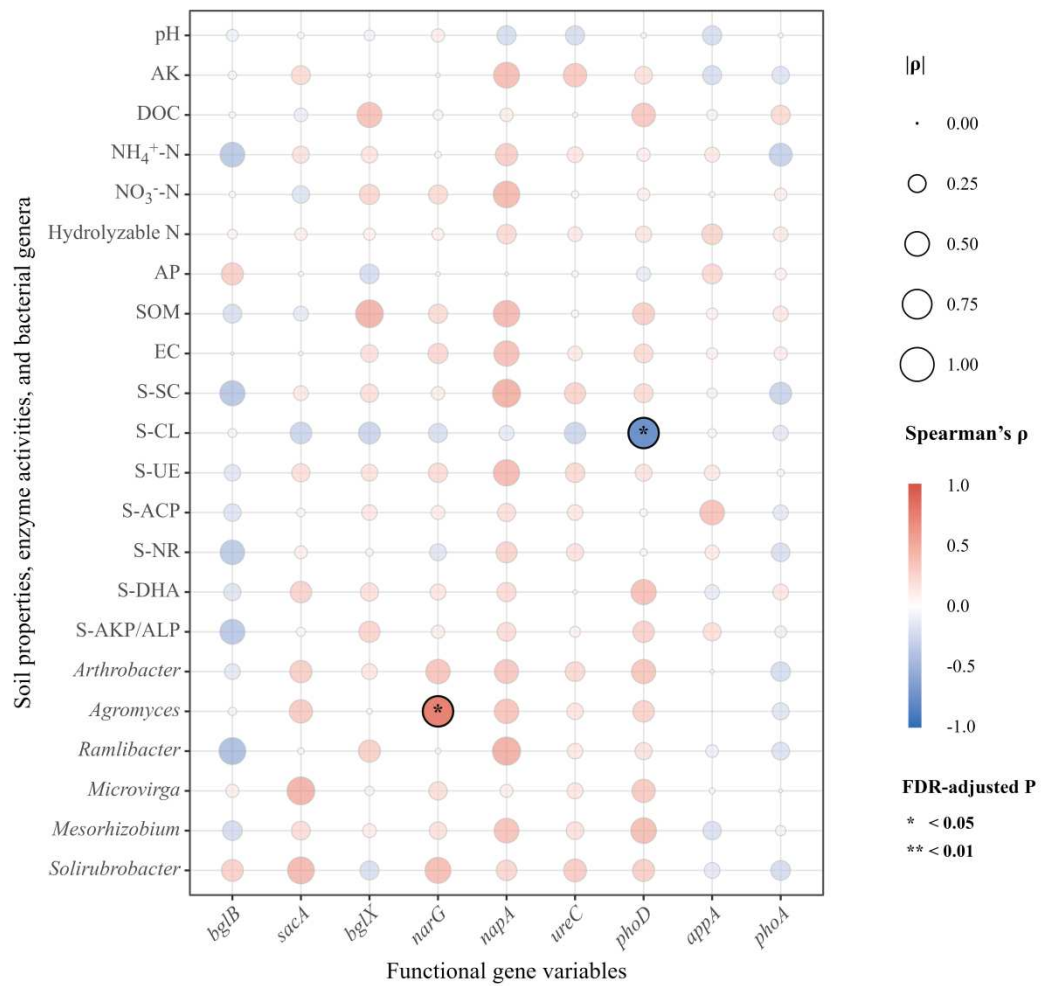


Fig. 5 Associations of functional genes with rhizosphere soil properties, enzyme activities, and bacterial genera

Note: Spearman correlations were calculated using sample-level data ($n = 15$). Bubble color indicates Spearman's ρ , with red and blue representing positive and negative associations, respectively, and bubble size represents $|\rho|$. Functional genes are shown on the x-axis, and soil properties, enzyme activities, and bacterial genera are shown on the y-axis. P values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure across the complete set of 357 cross-module tests. Asterisks indicate associations remaining significant after FDR correction (*FDR-adjusted $P < 0.05$, **FDR-adjusted $P < 0.01$). Unmarked bubbles are retained to show the direction and magnitude of associations but are not interpreted as robust pairwise evidence. Functional-gene and bacterial-genus names are shown in italics.

3.7 Direct Coupling Between Functional Genes and Corresponding Enzyme Activities

DCI values varied across the N treatments at both the pairwise and module levels

(Fig. 6). At the pairwise level, N34 showed the highest numerical DCI for all seven predefined gene-enzyme pairs. Significant treatment differences were observed for *sacA*-S-SC, *ureC*-S-UE, *napA*-S-NR, *appA*-S-ACP, and *phoD*-S-AKP/ALP, with N34 generally showing higher values than CK and N22. In contrast, *narG*-S-NR and *phoA*-S-AKP/ALP did not differ significantly among treatments despite their numerically higher values under N34. For most pairs, DCI declined from N34 to N40.

At the module level, C DCI, N DCI, P DCI, and Integrated DCI differed significantly among treatments, with the highest values consistently observed under N34 (Fig. 6b). Integrated DCI increased numerically from CK and N22 to N28, reached its maximum under N34, and declined under N40. Thus, the highest-yielding treatment was associated with the strongest gene-enzyme matching among the tested N rates, whereas this pattern was not maintained at the highest N input.

Internal robustness analyses indicated that the treatment-level pattern of Integrated DCI was stable across alternative analytical specifications. Integrated DCI showed high agreement with alternative matching metrics (median sample-level Spearman's $|\rho| = 0.96$; treatment-mean rank $|\rho| = 1.00$), and treatment rankings were unchanged under alternative normalization ($|\rho| = 1.00$) and leave-one-out recalculation (minimum $|\rho| = 1.00$). Across 1000 small-perturbation iterations of G_{norm} and E_{norm} , N34 remained the treatment with the highest Integrated DCI in all iterations. The predefined gene-enzyme pairs and complete robustness results are provided in Supplementary Tables S5 and S6.

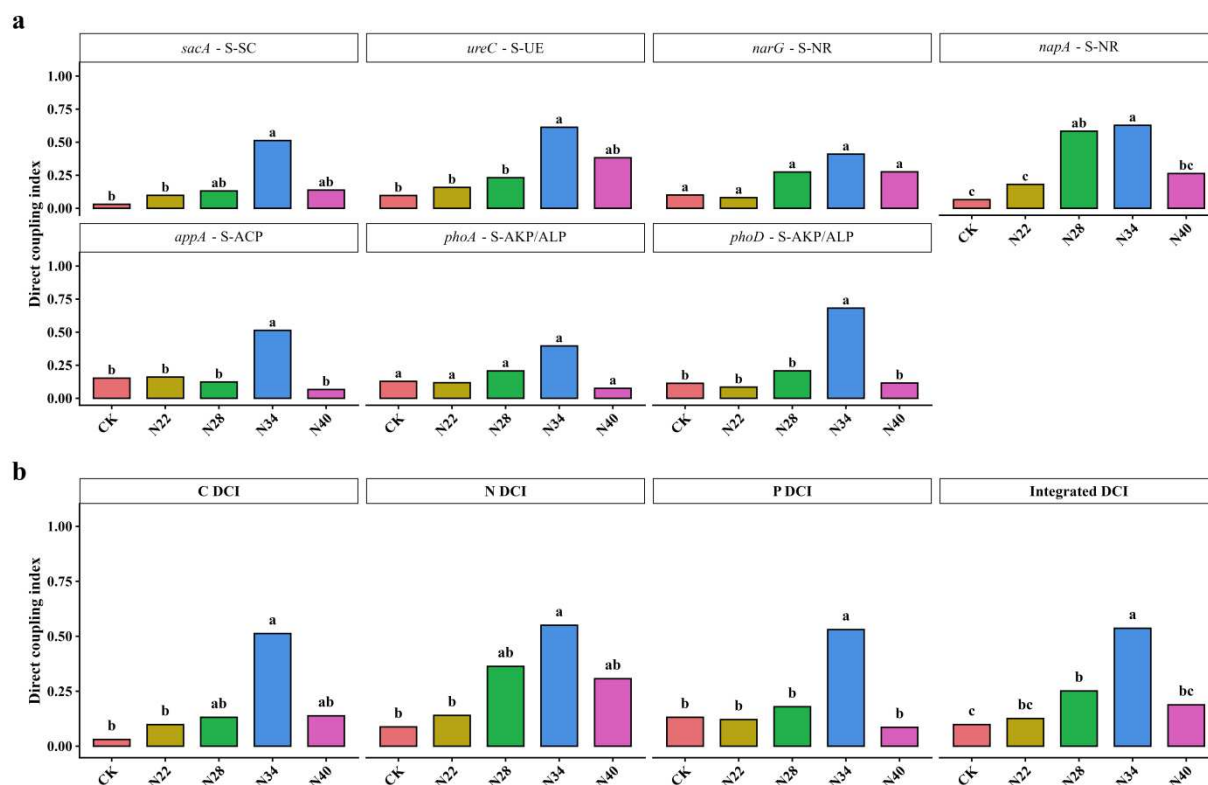


Fig. 6 Gene-enzyme coupling indices under N fertilizer gradients.

Note: (a) Pairwise direct coupling index (DCI; unitless) values for seven predefined gene-enzyme pairs involved in C, N, and P cycling: *sacA*-S-SC, *ureC*-S-UE, *narG*-S-NR, *napA*-S-NR, *appA*-S-ACP, *phoA*-S-AKP/ALP, and *phoD*-S-AKP/ALP. (b) C DCI, N DCI, P DCI, and Integrated DCI. Different lowercase letters indicate significant differences among treatments based on one-way ANOVA followed by Tukey's HSD test ($P < 0.05$). Higher DCI values indicate stronger within-sample matching between normalized functional-gene abundance and the corresponding measured enzyme activity.

3.8 Rhizosphere Factors Associated with Grain Yield and Agronomic Nitrogen Use Efficiency

Spearman correlation analysis after Benjamini-Hochberg FDR correction showed that grain yield was positively associated with NH_4^+ -N, SOM, EC, S-SC, S-UE, C DCI, N DCI, and Integrated DCI (Fig. 7a). For ANUE, CK was excluded because ANUE is undefined under zero N input; among the fertilized treatments, Spearman correlation analysis showed that ANUE was positively associated with NH_4^+ -N, S-SC, S-DHA, S-AKP/ALP, C DCI, N DCI, and Integrated DCI. Exploratory Random Forest analysis

yielded OOB R^2 values of 0.51 for grain yield and 0.44 for ANUE (Fig. 7b). Across 1000 repeated runs, *napA*, S-SC, *Ramlibacter*, EC, and $\text{NH}_4^+\text{-N}$ showed the highest median relative importance for grain yield, followed by S-UE, S-ACP, *Arthrobacter*, SOM, and *Agromyces*. For ANUE, S-SC, *Mesorhizobium*, and $\text{NH}_4^+\text{-N}$ ranked highest, followed by S-AKP/ALP, S-DHA, AP, SOM, *bglX*, hydrolyzable N, and Microvirga. The percentile intervals indicated appreciable variability in the importance of several predictors across repeated runs. Detailed Random Forest settings, model performance, and stability results are provided in Supplementary Method S2 and Tables S7-S8.

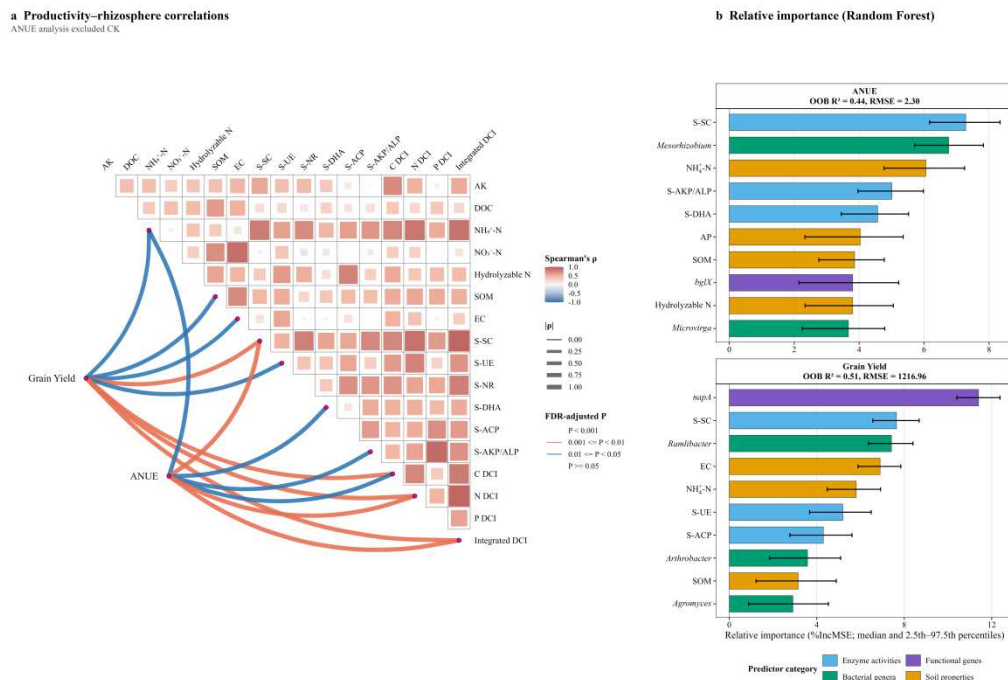


Fig. 7 Associations and relative importance of rhizosphere attributes in relation to grain yield and agronomic nitrogen use efficiency under nitrogen fertilizer gradients

Note: (a) Pairwise associations among selected soil properties, enzyme activities, and direct gene-enzyme coupling indices were evaluated using Spearman correlation. Red and blue matrix colors indicate positive and negative Spearman's ρ values, respectively. Edges link grain yield (kg ha^{-1}) and agronomic nitrogen use efficiency (ANUE; $\text{kg grain kg}^{-1} \text{N}$) with rhizosphere variables; edge width represents $|\rho|$ and edge color denotes the Benjamini–Hochberg false discovery rate (FDR)-adjusted significance level. Only associations with FDR-adjusted $P < 0.05$ are shown. Grain-yield correlations used all 15 samples, whereas ANUE-related analyses used the 12 N-fertilized samples because ANUE is undefined for CK. (b) Relative importance of predictors for

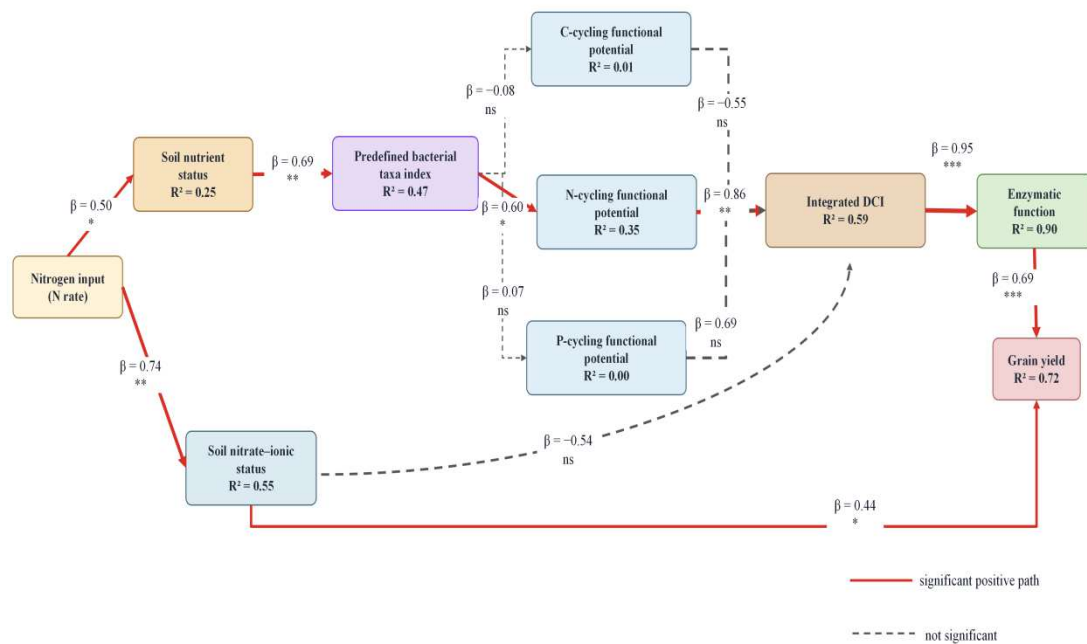
grain yield and ANUE based on exploratory Random Forest regression. Bars show the median percentage increase in mean squared error (%IncMSE) across 1,000 repeated runs, and horizontal intervals show the 2.5th-97.5th percentile range. Out-of-bag (OOB) R^2 and root mean squared error (RMSE) are shown for each model. Colors indicate predictor categories shown in the panels: soil properties, enzyme activities, functional genes, and bacterial genera.

3.9 Exploratory PLS-SEM Linking Nitrogen Input, Rhizosphere Functional Processes, and Grain Yield

The exploratory PLS-SEM showed that N input was positively associated with soil nutrient status ($\beta = 0.50$, $P = 0.015$; $R^2 = 0.25$) and soil nitrate-ionic status ($\beta = 0.74$, $P = 0.003$; $R^2 = 0.55$) (Fig. 8). Soil nutrient status was positively associated with the predefined bacterial taxa index ($\beta = 0.69$, $P = 0.002$; $R^2 = 0.47$). The bacterial taxa index was positively associated with N-cycling functional potential ($\beta = 0.60$, $P = 0.011$; $R^2 = 0.35$), whereas its associations with C- and P-cycling functional potential were not significant.

Among the three functional modules, N-cycling functional potential was positively associated with Integrated DCI ($\beta = 0.86$, $P = 0.008$), whereas the corresponding paths from the C- and P-cycling modules were not significant ($\beta = -0.55$, $P = 0.236$ and $\beta = 0.69$, $P = 0.095$, respectively). Soil nitrate-ionic status also showed a negative but non-significant association with Integrated DCI ($\beta = -0.54$, $P > 0.05$). Integrated DCI was positively associated with enzymatic function ($\beta = 0.95$, $P < 0.001$; $R^2 = 0.90$), which was positively associated with grain yield ($\beta = 0.69$, $P < 0.001$). Soil nitrate-ionic status was also positively associated with grain yield ($\beta = 0.44$, $P = 0.048$). The model accounted for 72.2% of the observed variation in grain yield.

Among the three parallel functional branches, only the N-cycling-specific indirect association from N input to grain yield was significant (indirect effect = 0.115, 95% CI = 0.004–0.301, $P = 0.034$), whereas the corresponding C- and P-cycling indirect pathways were not significant. Given the small sample size ($n = 15$), these relationships were interpreted as exploratory associations within the prespecified conceptual model. All structural VIF values were below 3 (range: 1.000–1.977), and the total indirect association of N input with grain yield was not significant (effect = 0.196, 95% CI = -0.183 – 0.707 , $P = 0.263$); complete bootstrap estimates for the indirect and total effects are provided in Supplementary Tables S15 and S17. Detailed PLS-SEM settings, bootstrap confidence intervals, R^2 values, VIF diagnostics, leave-one-out stability, branch-specific indirect effects, and complete total indirect and total effects are provided in Supplementary Method S2 and Tables S9–S12, S15, and S17.



Exploratory PLS-SEM; coupling represented by Integrated DCI; bootstrap = 5,000; n = 15

Fig. 8 Exploratory PLS-SEM linking nitrogen input, rhizosphere functional processes, and grain yield

Note: The predefined bacterial taxa index comprised *Agromyces*, *Ramlibacter*, and *Mesorhizobium* and should not be interpreted as representing the full bacterial community. C-, N-, and P-cycling functional potentials were modeled as three parallel modules and linked to Integrated DCI. Red arrows indicate significant positive paths, whereas grey dashed arrows indicate non-significant paths. Arrow width is proportional to the absolute standardized path coefficient. Values on arrows are standardized β coefficients, and R^2 values are shown inside endogenous variables. Significance levels are indicated as * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$ based on 5,000 bootstrap resamples; ns indicates not significant. The direct path from nitrogen input to grain yield was not included. Bootstrap = 5,000; n = 15; exploratory PLS-SEM.

4. Discussion

4.1 The highest-yielding N rate was associated with favorable rhizosphere conditions and nitrogen-use efficiency

Grain yield increased up to N34 and declined under N40, showing that the additional N input from 300 to 375 kg N ha⁻¹ did not result in a proportional increase in maize productivity. N34 achieved the highest grain yield and ANUE, whereas N40

maintained a shoot N concentration comparable to N34 but showed lower yield and markedly reduced ANUE. The comparable shoot N concentrations under N34 and N40, together with the decline in yield and ANUE under N40, indicate that greater N input did not translate into a proportional improvement in crop N utilization or grain production. Similar nonlinear yield and N-use efficiency responses to increasing N supply have been reported in maize (Zhang et al. 2022; Zou et al. 2024).

The higher productivity under N34 coincided with relatively high NH_4^+ -N, hydrolyzable N, AK, SOM, and DOC, whereas N40 was characterized by the highest NO_3^- -N and EC. Thus, the additional accumulation of NO_3^- -N and the higher ionic status under N40 were not accompanied by further gains in yield or ANUE. Rather than indicating a simple benefit of greater mineral-N accumulation, the contrast between N34 and N40 suggests that the highest-yielding treatment was associated with a different balance of nutrient availability and rhizosphere ionic conditions. Despite small treatment-related variation, pH remained within a narrow alkaline range and was therefore considered a relatively stable background condition in this experiment.

Bacterial richness, represented by Chao1 and ACE, did not differ significantly among treatments, whereas Shannon and Simpson diversity varied across the N gradient. N34 maintained comparatively high diversity among the fertilized treatments and coincided with relatively high DOC and nutrient availability, whereas N40 combined the highest NO_3^- -N and EC with lower Shannon and Simpson diversity than CK. However, because CK itself showed the highest diversity values, the

observed diversity pattern cannot be attributed to nutrient availability or ionic conditions alone. These relationships should therefore be interpreted as treatment-level associations rather than evidence for a single controlling mechanism.

The exploratory PLS-SEM further distinguished soil nutrient status from nitrate-ionic status (Fig. 8). N input was positively associated with both modules, whereas only soil nutrient status was positively associated with the predefined bacterial taxa index comprising *Agromyces*, *Ramlibacter*, and *Mesorhizobium*. Although nitrate-ionic status showed a positive direct association with grain yield within the exploratory PLS-SEM, this across-sample relationship should not be interpreted as evidence that continued NO_3^- -N and EC accumulation was beneficial, because yield declined from N34 to N40. Accordingly, the PLS-SEM result is best interpreted as a gradient-level association that complements, rather than overrides, the treatment-specific contrast between N34 and N40.

4.2 Asynchronous responses of C-, N-, and P-cycling genes and enzymes were linked to variation in functional coupling

The contrast between N34 and N40 was also evident at the microbial functional level, where functional-gene potential and measured enzyme activity did not respond synchronously to increasing N supply. N34 combined relatively high abundances of several selected functional genes with the highest activities of most measured enzymes and the highest DCI values, whereas N40 retained relatively high abundances of several N-cycling genes but showed lower activities of most measured

enzymes and lower gene-enzyme matching (Fig. 3, 4, and 6). This divergence indicates that greater functional-gene abundance does not necessarily correspond to an equivalent increase in measured enzyme activity. Functional-gene abundance represents genetic potential, whereas measured enzyme activity also depends on enzyme production and persistence, substrate availability, and soil environmental conditions (Burns et al. 2013; Schimel et al. 2017).

Among the three nutrient-cycling branches, the N-related response showed the most coherent pattern across functional levels. Although the overall N gene index did not differ significantly among treatments, individual N-cycling genes varied across the N gradient, and N34 was accompanied by high activities of N-related enzymes and the highest N DCI. At the FDR-corrected level, *narG* was positively associated with *Agromyces* (Fig. 5). The exploratory PLS-SEM was consistent with this pattern, showing positive associations from the predefined bacterial taxa index to N-cycling functional potential and from N-cycling potential to Integrated DCI (Fig. 8). Together, these results suggest that the N-cycling branch showed the most coherent cross-level association under the imposed N gradient. Fertilization-related variation in maize rhizosphere N-cycling genes has also been reported by Enebe and Babalola (2021a). Studies across other soil environmental gradients have examined relationships between N-cycling genes and N transformation processes (Petersen et al. 2012; Wang et al. 2022).

The P-cycling branch showed a less consistent cross-level response. The P gene index, *appA* abundance, and phosphatase activities were highest or comparatively

high under N34, whereas the P-specific indirect association was not significant in the exploratory PLS-SEM (Supplementary Table S15). Thus, P-related variables responded to N fertilization, but these responses did not form a statistically supported indirect branch to grain yield within the present model. This inconsistency may partly reflect the dependence of P-cycling processes on P availability and organic-P substrates (Dai et al. 2020; Enebe and Babalola 2021b; Liu et al. 2024). The predefined bacterial taxa index may also not capture all bacterial groups associated with P cycling, which is a limitation of the present model.

The C-cycling branch showed the greatest asynchrony. The C gene index did not differ significantly among treatments, whereas S-SC was highest under N34 and S-CL showed no significant treatment response. C DCI should also be interpreted cautiously because it was represented only by the predefined *sacA*–S-SC pair, whereas the broader C functional gene index included additional C-cycling genes (Supplementary Tables S4 and S5). Therefore, the non-monotonic C DCI pattern may reflect both biological asynchrony among individual C-related components and the limited representation of C cycling by a single predefined gene-enzyme pair. The weaker cross-level coherence of the C branch may also reflect the fact that N, rather than C, was the experimentally manipulated resource. Although N34 coincided with relatively high DOC and SOM, the C gene index remained statistically unchanged and S-CL showed no significant treatment response, suggesting that variation in these measured C-related pools did not translate into a coordinated response across the measured C functions. The three C markers, *bglB*, *sacA*, and *bglX*, represent distinct

carbohydrate-transforming functions, and their responses may therefore depend on substrate identity and on the balance between C supply and microbial N and P demand rather than on N input alone. Because microbial biomass C:N:P stoichiometry, root-exudate composition, and substrate-specific C pools were not measured, this substrate-stoichiometric explanation remains a hypothesis rather than a demonstrated mechanism.

4.3 Gene-enzyme coupling provided an exploratory link between rhizosphere functional processes and maize yield

The asynchronous responses of functional genes and enzymes described above provide the rationale for evaluating their joint matching rather than interpreting either component alone. Previous studies have examined gene-enzyme relationships using correlations between functional-gene abundance and enzyme activities (Trivedi et al. 2016; Wang et al. 2023) and have discussed comparisons between microbial functional potential and measured extracellular enzyme activity (Breitkreuz et al. 2021; Chen and Sinsabaugh 2021). These approaches quantify covariation or prespecified relationships across variables, whereas the study-defined DCI provides a relative sample-level summary that incorporates both the magnitude of paired functional-gene abundance and enzyme activity and their agreement. DCI should therefore be regarded as complementary to, rather than a replacement for, conventional gene-, enzyme-, correlation-, or path-based analyses. In the present experiment, C DCI, N DCI, P DCI, and Integrated DCI were highest under N34 and declined under N40 (Fig. 6), indicating that the highest-yielding treatment was also

associated with stronger within-sample gene-enzyme matching. This treatment-level pattern remained stable across alternative matching metrics, normalization procedures, leave-one-out recalculation, and perturbation tests, supporting the within-study robustness of the DCI response (Supplementary Table S6).

The broader associations between rhizosphere functional attributes and crop performance were further examined using correlation, Random Forest, and exploratory PLS-SEM analyses. After FDR correction, grain yield remained positively associated with N DCI and Integrated DCI, together with several soil nutrient and enzyme variables, whereas ANUE showed similar associations with N-related soil, enzyme, and coupling metrics (Fig. 7a). Random Forest analysis likewise identified predictors from multiple functional levels, including soil properties, enzyme activities, functional genes, and bacterial genera (Fig. 7b). The exploratory PLS-SEM further integrated these relationships by linking soil nutrient status with the predefined bacterial taxa index, N-cycling functional potential, Integrated DCI, enzymatic function, and grain yield (Fig. 8). Among the three predefined functional branches, only the N-cycling-specific indirect association with grain yield was significant (indirect effect = 0.115, 95% CI = 0.004–0.301, $P = 0.034$; Supplementary Table S15). Together, these analyses provide convergent exploratory evidence for a comparatively coherent N-related functional response, rather than independent confirmation of a causal mechanism.

The decline in DCI under N40 provides an important contrast with N34. Although N40 accumulated the highest NO_3^- -N and EC and retained relatively high

abundances of several functional genes, it showed lower activities of several enzymes, lower DCI values, and reduced yield and ANUE relative to N34. Thus, greater residual N and ionic accumulation at the highest N input were not accompanied by a corresponding increase in gene-enzyme matching or crop performance. One possible explanation is that increasing N supply altered the balance among microbial functional potential, substrate availability, physiological demand, and extracellular-enzyme production, thereby increasing the mismatch between gene abundance and measured enzyme activity; however, this interpretation remains a hypothesis because the present measurements cannot identify the underlying mechanism (Burns et al. 2013). Because the metagenomic measurements quantify gene abundance rather than transcriptional activity, the present study cannot directly determine whether high N decoupled gene transcription from extracellular-enzyme production. Extracellular enzyme activity can persist after the cessation of new enzyme production (Schimel et al. 2017). Such persistence could contribute to a temporal mismatch with contemporaneous genetic measurements, although this possibility was not tested here. However, microbial biomass stoichiometry, transcriptional activity, and direct indicators of ionic stress were not measured. The present data therefore cannot identify whether the lower coupling under N40 was primarily associated with nutrient imbalance, ionic conditions, altered microbial physiology, temporal gene-enzyme mismatch, or a combination of these processes.

Several limitations define how DCI and the multivariate analyses should be interpreted. DCI depends on min–max normalization within the analysed dataset, is

sensitive to the predefined gene-enzyme pairs, and does not account for temporal lags among gene abundance, gene expression, and extracellular-enzyme persistence. Its module representation was also unequal: C DCI was based on one predefined pair, whereas N and P DCI each incorporated three pairs. Moreover, Integrated DCI was calculated partly from enzyme measurements, so its strong association with the enzymatic-function construct in the PLS-SEM is not statistically independent and should be interpreted as functional correspondence rather than an independent causal effect. DCI has not yet been externally validated across independent soil datasets; therefore, the present robustness analyses demonstrate within-study stability rather than general validity. Similarly, the small sample size requires both the Random Forest and PLS-SEM results to be interpreted as exploratory rather than confirmatory. Accordingly, the exploratory PLS-SEM should be interpreted as a set of conditional associations within the prespecified model, rather than as evidence of causal pathways or mediation. With only 15 samples, both analyses remain susceptible to overfitting and unstable model estimates; the repeated Random Forest runs, bootstrap confidence intervals, structural VIF diagnostics, and leave-one-out analyses evaluate internal stability but do not eliminate this limitation or substitute for validation in an independent dataset. Accordingly, DCI is best viewed as a complementary summary of paired gene-enzyme matching that should be interpreted together with its component gene and enzyme measurements. Overall, the highest-yielding N rate among the tested treatments was associated with favorable rhizosphere nutrient conditions, comparatively coherent N-related functional potential, higher measured

enzyme activity, and stronger gene-enzyme matching, whereas the highest N input did not maintain the same degree of rhizosphere functional coordination.

5. Conclusions

This study showed that maize yield responded nonlinearly to nitrogen input, with 300 kg N ha⁻¹ representing the highest-yielding rate among the five tested treatments under the present drip-irrigated field conditions. This yield advantage was associated with favorable rhizosphere nutrient conditions, comparatively high relative abundances of several rhizosphere-associated bacterial genera, relatively strong responses of selected functional genes, and higher activities of most measured soil enzymes. Direct coupling analysis further showed that C DCI, N DCI, P DCI, and Integrated DCI all peaked at this N rate and declined at the highest tested N rate, indicating that the highest-yielding treatment was associated with stronger within-sample gene-enzyme matching. Exploratory PLS-SEM further indicated associations among soil nutrient status, the predefined bacterial taxa index, N-cycling functional potential, Integrated DCI, enzymatic function, and grain yield. Overall, these findings indicate that the highest-yielding N rate among the tested treatments was associated not simply with greater mineral-N accumulation, but with more coordinated changes in rhizosphere nutrient conditions, microbial functional potential, measured enzyme activity, and gene-enzyme matching.

References

- Bao SD (2000) Soil and agricultural chemistry analysis, 3rd edn. China Agriculture Press, Beijing
- Breitkreuz C, Heintz-Buschart A, Buscot F, Wahdan SFM, Tarkka M, Reitz T (2021) Can we estimate

functionality of soil microbial communities from structure-derived predictions? A reality test in agricultural soils. *Microbiol Spectr* 9(1):e00278-21. <https://doi.org/10.1128/spectrum.00278-21>

Burns RG, DeForest JL, Marxsen J et al (2013) Soil enzymes in a changing environment: current knowledge and future directions. *Soil Biol Biochem* 58:216-234. <https://doi.org/10.1016/j.soilbio.2012.11.009>

Chen H, Ma K, Lu C et al (2022) Functional redundancy in soil microbial community based on metagenomics across the globe. *Front Microbiol* 13:878978. <https://doi.org/10.3389/fmicb.2022.878978>

Chen J, Sinsabaugh RL (2021) Linking microbial functional gene abundance and soil extracellular enzyme activity: implications for soil carbon dynamics. *Glob Chang Biol* 27(7):1322-1325. <https://doi.org/10.1111/gcb.15506>

Dai Z, Liu G, Chen H et al (2020) Long-term nutrient inputs shift soil microbial functional profiles of phosphorus cycling in diverse agroecosystems. *ISME J* 14(3):757-770. <https://doi.org/10.1038/s41396-019-0567-9>

Enebe MC, Babalola OO (2021a) Soil fertilization affects the abundance and distribution of carbon and nitrogen cycling genes in the maize rhizosphere. *AMB Express* 11(1):24. <https://doi.org/10.1186/s13568-021-01182-z>

Enebe MC, Babalola OO (2021b) The influence of soil fertilization on the distribution and diversity of phosphorus cycling genes and microbes community of maize rhizosphere using shotgun metagenomics. *Genes (Basel)* 12(7):1022. <https://doi.org/10.3390/genes12071022>

Hu M, Zhou S, Xiong X, Wang X, Sun Y, Meng Z, Hui D, Li J, Zhang D, Deng Q (2024) Dynamics of soil microbial communities involved in carbon cycling along three successional forests in southern China. *Front Microbiol* 14:1326057. <https://doi.org/10.3389/fmicb.2023.1326057>

Ling N, Wang T, Kuzyakov Y (2022) Rhizosphere bacteriome structure and functions. *Nat Commun* 13(1):836. <https://doi.org/10.1038/s41467-022-28448-9>

Liu S, Li H, Xie X, Chen Y, Lang M, Chen X (2024) Long-term moderate fertilization increases the complexity of soil microbial community and promotes regulation of phosphorus cycling genes to improve the availability of phosphorus in acid soil. *Appl Soil Ecol* 194:105178. <https://doi.org/10.1016/j.apsoil.2023.105178>

Mikha MM, Obour AK, Kumar V, Stahlman PW (2019) Soil physicochemical properties influenced by nitrogen sources and rates in the Central Great Plains. *J Soil Water Conserv* 74(6):584-593. <https://doi.org/10.2489/jswc.74.6.584>

Petersen DG, Blazewicz SJ, Firestone M, Herman DJ, Turetsky M, Waldrop M (2012) Abundance of microbial genes associated with nitrogen cycling as indices of biogeochemical process rates across a vegetation gradient in Alaska. *Environ Microbiol* 14(4):993-1008. <https://doi.org/10.1111/j.1462-2920.2011.02679.x>

Rocca JD, Hall EK, Lennon JT et al (2015) Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *ISME J* 9(8):1693-1699. <https://doi.org/10.1038/ismej.2014.252>

Schimel J, Becerra CA, Blankinship J (2017) Estimating decay dynamics for enzyme activities in soils from different ecosystems. *Soil Biol Biochem* 114:5-11. <https://doi.org/10.1016/j.soilbio.2017.06.023>

Tabatabai MA (1994) Soil enzymes. In: Weaver RW, Angle S, Bottomley P, Bezdicsek D, Smith S, Tabatabai A, Wollum A (eds) *Methods of Soil Analysis, Part 2: Microbiological and Biochemical*

- Properties. Soil Science Society of America, Madison, pp 775-833.
<https://doi.org/10.2136/sssabookser5.2.c37>
- Trivedi P, Delgado-Baquerizo M, Trivedi C et al (2016) Microbial regulation of the soil carbon cycle: evidence from gene–enzyme relationships. *ISME J* 10(11):2593-2604.
<https://doi.org/10.1038/ismej.2016.65>
- Wang J, He L, Xu X et al (2022) Linkage between microbial functional genes and net N mineralisation in forest soils along an elevational gradient. *Eur J Soil Sci* 73(4):e13276.
<https://doi.org/10.1111/ejss.13276>
- Wang J, Chen G, Ji S et al (2023) Close relationship between the gene abundance and activity of soil extracellular enzyme: evidence from a vegetation restoration chronosequence. *Soil Biol Biochem* 177:108929. <https://doi.org/10.1016/j.soilbio.2022.108929>
- Zhang H, Zhang C, Sun P, Jiang X, Xu G, Yang J (2022) Optimizing planting density and nitrogen application to enhance profit and nitrogen use of summer maize in Huanghuaihai region of China. *Sci Rep* 12:2704. <https://doi.org/10.1038/s41598-022-06059-0>
- Zhang L, Yuan L, Wen Y et al (2024) Maize functional requirements drive the selection of rhizobacteria under long-term fertilization practices. *New Phytol* 242:1275-1288.
<https://doi.org/10.1111/nph.19653>
- Zhao R, Liu J, Xu N, He T, Meng J, Liu Z (2022) Urea hydrolysis in different farmland soils as affected by long-term biochar application. *Front Environ Sci* 10:950482.
<https://doi.org/10.3389/fenvs.2022.950482>
- Zhu S, Vivanco JM, Manter DK (2016) Nitrogen fertilizer rate affects root exudation, the rhizosphere microbiome and nitrogen-use-efficiency of maize. *Appl Soil Ecol* 107:324-333.
<https://doi.org/10.1016/j.apsoil.2016.07.009>
- Zou H, Li D, Ren K et al (2024) Response of maize yield and nitrogen recovery efficiency to nitrogen fertilizer application in field with various soil fertility. *Front Plant Sci* 15:1349180.
<https://doi.org/10.3389/fpls.2024.1349180>

Statements and Declarations

Funding: This research was funded by the “Tianchi Talents” Young Doctors Recruitment Program (2024TCYCQNBS03) and the National Natural Science Foundation of China (32360444).

Competing interests: The authors declare that they have no competing interests.

Author contributions: The study was conceived and designed by Baoguo Yang and Kanghui Liu. The writing of the manuscript and data analysis were performed by Baoguo Yang and Kanghui Liu. Experimental procedures and data collection were carried out by Kanghui Liu, Ning Li, Fei Liang, Qinzheng Xu, Wenxin Ma, Baoqiang Li, Dongmei Zhao, Tongxu Liu, Waseem Hassan, Muqarrab Ali, and Shengguang Zhang. The acquisition of funding was secured by Baoguo Yang.

885 **Data availability:** The processed agronomic, soil physicochemical, enzyme activity, functional
886 gene abundance, microbial abundance, and model-output datasets are not publicly available
887 because they are part of ongoing related studies, but are available from the corresponding author
888 on reasonable request.
889 **Declarations:** The authors declare that they have no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial2.docx](#)