Keystone Bacterial Taxa Drive Denitrification and N_2O Emission via Adaptive Genomic and Metabolic Strategies in Contrasting Agricultural Soils

Siyu Yu[a, 1], Qiaoyu Wu[a, b, 1], Yimin Ma[a], Saira Bano[a], Xiaojun Zhang[a] *

[a] *State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, and School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, China*

[b] *State Key Laboratory of Pollution Control and Resources Reuse, Shanghai Institute of Pollution Control and Ecological Security, College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China*

[1] These authors contributed equally to this work.

*Corresponding author: Xiaojun Zhang,

Email: xjzhang68@sjtu.edu.cn.

Supplementary Information

Supplementary Materials and Methods

Supplementary Tables

Supplementary Figures

Supplementary Materials and Methods

High-throughput identification of bacterial isolates

To rapidly identify 719 bacterial isolates from black soil (BS) and fluvo-aquic soil (FS), we employed a high-throughput barcoding approach previously described by Zhang et al. [1]. Briefly, DNA extracts from all isolates were systematically arranged into 96-well plates, with each well containing DNA from a single bacterial isolate. The two-sided barcode PCR system was used to amplify the V5–V7 regions of the bacterial 16S rRNA gene with primers 799F and 1193R, followed by Illumina sequencing. The amplicon data were analyzed using the 'Culturome' pipeline (https://github.com/YongxinLiu/Culturome) to achieve precise taxonomic classification of the bacterial isolates.

Phylogenetic analysis

For the phylogenetic reconstruction of bacterial isolates, the representative sequences for each ASVs identified from bacterial isolates were aligned and trimmed using MAFFT v7.505 [2] and trimal v1.4 [3]. IQ-TREE [4] was employed to determine the optimal model for constructing a maximum-likelihood tree, supported by 1000 bootstrap replicates. The phylogenetic tree was visualized and annotated using the Interactive Tree of Life (iTOL) online tool (https://itol.embl.de/).

For the phylogenetic reconstruction of denitrification genes within the *Ensifer* ASV205 isolates, the amino acid sequences of the *napA, nirK, norB* and *nosZ* genes were collected

from all *Ensifer* ASV205 isolates obtained in this study and reference genomes from NCBI GenBank. Phylogenetic trees for each gene were constructed following the same procedure as described above.

Supplementary Tables

Table S1 The information of primers targeting the denitrification genes

Target genes	Primers	Sequence (5' to 3')	Amplicon size (bp)	Reference
narG	narG-F	TCGCCSATYCCGGCSATGTC	173	[5]
	narG-R	GAGTTGTACCAGTCRGCSGAYTCSG		
nirS	cd3aF	AACGYSAAGGARACSGG	413	[6]
	R3cd	GASTTCGGRTGSGTCTTSAYGAA		
nirK	F1aCu	ATCATGGTSCTGCCGCG	473	[6]
	R3Cu	GCCTCGATCAGRTTGTGGTT		
*nosZ*I	nosZ1mod-F	WCSYTSTTCMTSGAYAGCCAG	248-251	[7]
	nosZ1mod-R	ATRTCGATSARCTGVKCRTTYTC		
*nosZ*II	NosZ-II-F	CTIGGICCIYTKCAYAC	689-716	[8]
	NosZ-II-R	GCIGARCARAAITCBGTRC		

Table S2 Genomic features of *Ensifer* ASV205 strains

Strains	Seq ID	Seq Length (bp)	G+C(%)	tRNA genes	rRNA genes	ORF number	ORF/Genome (%)
ABK11	chr	4,108,085	62.36	56	9	4032	85.72%
	plasmid1	2,490,287	61.84	4	3	2311	87.93%
	plasmid2	724,468	59.32	1	0	840	80.60%
	plasmid3	402,162	59.06	0	0	410	86.80%
	plasmid4	139,379	58.81	0	0	158	79.90%
	plasmid5	95,712	58.92	0	0	104	86.81%
ABS24	chr	4,030,277	62.44	58	9	3921	85.68%
	plasmid1	2,429,957	61.73	4	3	2276	87.84%
	plasmid2	749,407	59.47	1	0	843	80.78%
ABS37	chr	4,052,777	62.42	57	9	3934	85.53%
	plasmid1	2,409,568	62.03	3	3	2240	88.06%
	plasmid2	886,307	59.85	1	0	949	83.39%
ABS97	chr	4,015,948	62.5	56	9	3893	85.77%
	plasmid1	2,472,957	61.85	4	3	2292	88.03%
	plasmid2	774,208	59.32	1	0	873	80.56%
	plasmid3	38,174	59.02	0	0	48	85.80%
AFS42	chr	4,131,388	62.29	56	9	4002	85.68%
	plasmid1	2,640,497	61.4	5	3	2460	87.63%
	plasmid2	740,497	59.5	1	0	870	80.90%
	plasmid3	452,795	59.16	1	0	433	85.52%
OBK33	chr	4,049,704	62.42	56	9	3930	85.76%
	plasmid1	2,505,184	61.8	4	3	2320	88.05%
	plasmid2	667,930	59.28	1	0	775	79.75%
	plasmid3	163,983	59.44	0	0	169	86.84%
OBS65	chr	4,046,222	62.38	55	9	3954	85.76%
	plasmid1	2,395,377	61.91	3	3	2229	87.96%
	plasmid2	913,751	59.85	2	0	998	83.52%
OFS103	chr	4,120,005	62.42	57	9	4059	85.71%
	plasmid1	2,399,310	61.8	4	3	2239	87.82%
	plasmid2	665,409	59.28	1	0	773	79.69%
OFS104	chr	4,120,007	62.42	57	9	4061	85.73%
	plasmid1	2,397,751	61.79	4	3	2228	87.81%
	plasmid2	665,410	59.28	1	0	776	79.84%
OFS199	chr	4,046,782	62.08	52	9	3934	85.65%
	plasmid1	1,916,357	61.03	6	6	1823	86.44%
	plasmid2	1,665,499	60.78	2	0	1652	86.33%

Table S3 Basic characteristics of the node ASV205 in the BS and FS networks

Network_property	ASV_205 from BS	ASV_205 from FS
Node degree	49	28
Node betweenness	869.04	1309.01
Node stress	28076.00	31939.00
Node eigenvector centrality scores	0.54	0.16
Node closeness	0.00	0.00
Clustering coefficient	0.50	0.36
Zi	0.42	1.09
Pi	0.63	0.63
Network roles	Connector hubs	Connector hubs

Supplementary Figures

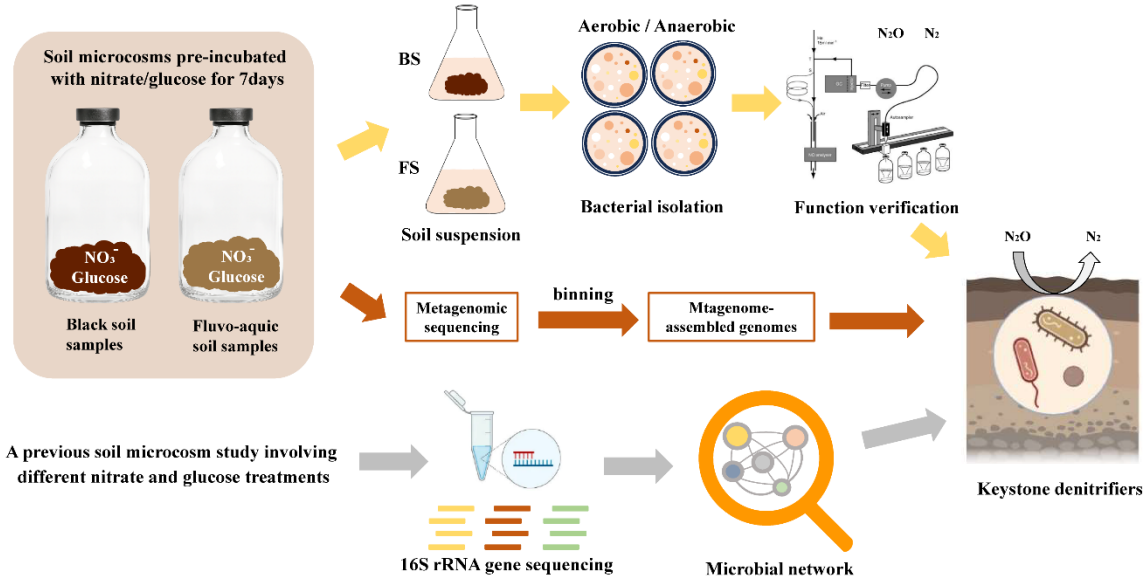

Fig. S1 Integrated workflow for identifying keystone denitrifiers in BS and FS. Soil samples from BS and FS were subjected to short-term anaerobic enrichment incubation with nitrate and glucose to stimulate denitrifying bacterial activity. The enriched soils were then used for bacterial isolation under aerobic and anaerobic conditions, followed by screening for denitrification-related genes. Based on metagenomic sequencing, the community structures of black soil and fluvo-aquic soil were characterized, and high-quality metagenome-assembled genomes (MAGs) were recovered from the enriched soils. Soil samples used for 16S rRNA gene sequencing originated from a previous soil microcosm study involving different nitrate and glucose treatments [9]. Keystone strains were selected and further verified for N_2O/N_2 production.

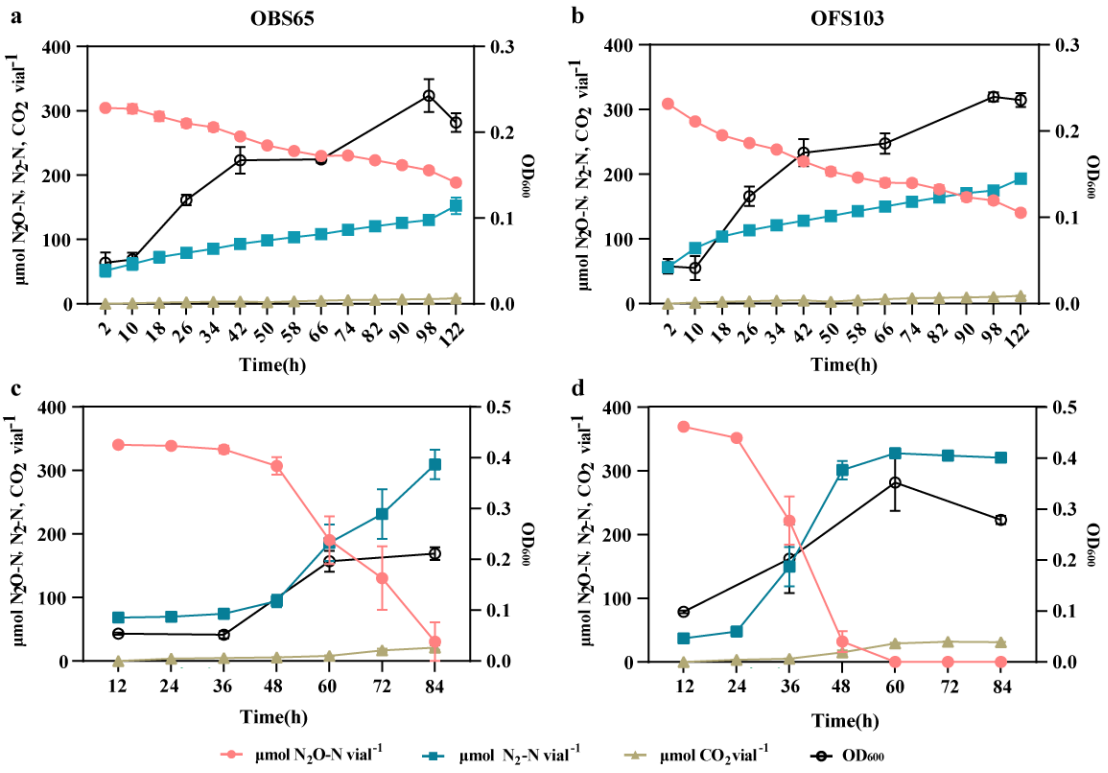

Fig. S2 Growth dynamics and N_2O-reducing performance of *Ensifer* ASV205 strains OBS65 (Group I) and OFS103 (Group III) under different culture conditions. Kinetics of N_2O, N_2, CO_2 and OD_{600} of strains cultured in nitrogen-free DM medium (a-b). Kinetics of N_2O, N_2, CO_2 and OD_{600} of strains cultured in TSB medium (c-d).

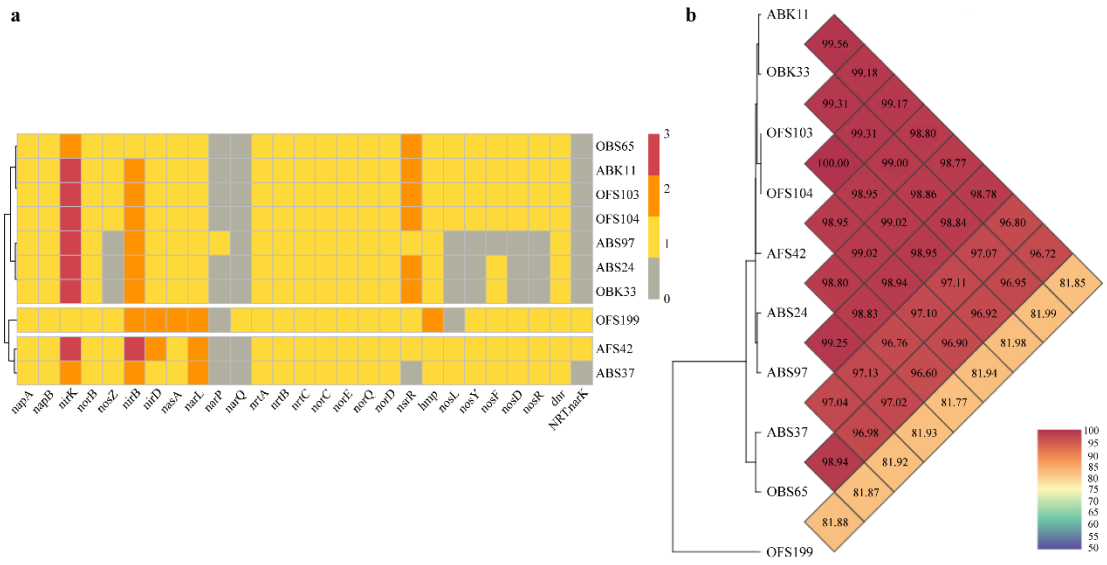

Fig. S3 Nitrogen-cycling genes and ANI value among *Ensifer* ASV205 strains. (a) Heatmap showing the copy number of nitrogen-cycling genes identified in the genomes of 10 *Ensifer* ASV205 strains. (b) ANI matrix among *Ensifer* ASV205 strains, with hierarchical clustering based on genome similarity. Values in diamonds indicate pairwise ANI percentages.

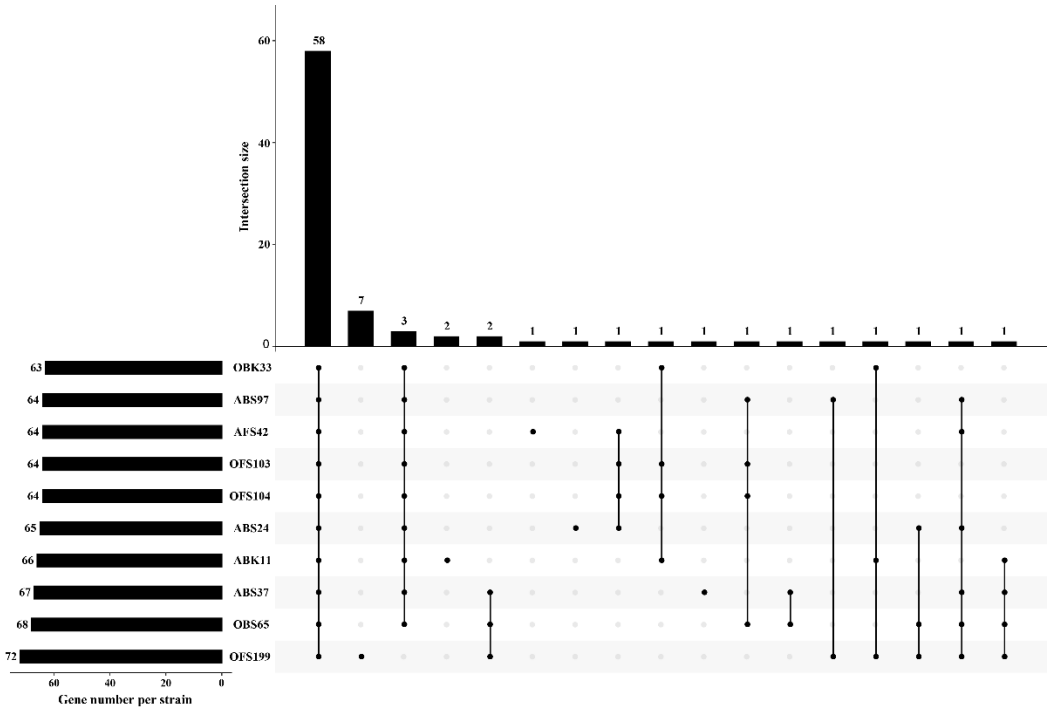

Fig. S4 Comparative analysis of carbohydrate-active enzyme (CAZyme) gene profiles in *Ensifer* ASV205 strains. An UpSet plot showing the distribution of CAZyme-encoding genes among 10 *Ensifer* ASV205 genomes. The horizontal bars on the left indicate the total number of CAZyme genes identified in each strain. The vertical bars above represent the number of shared or unique gene sets corresponding to the intersection matrix below.

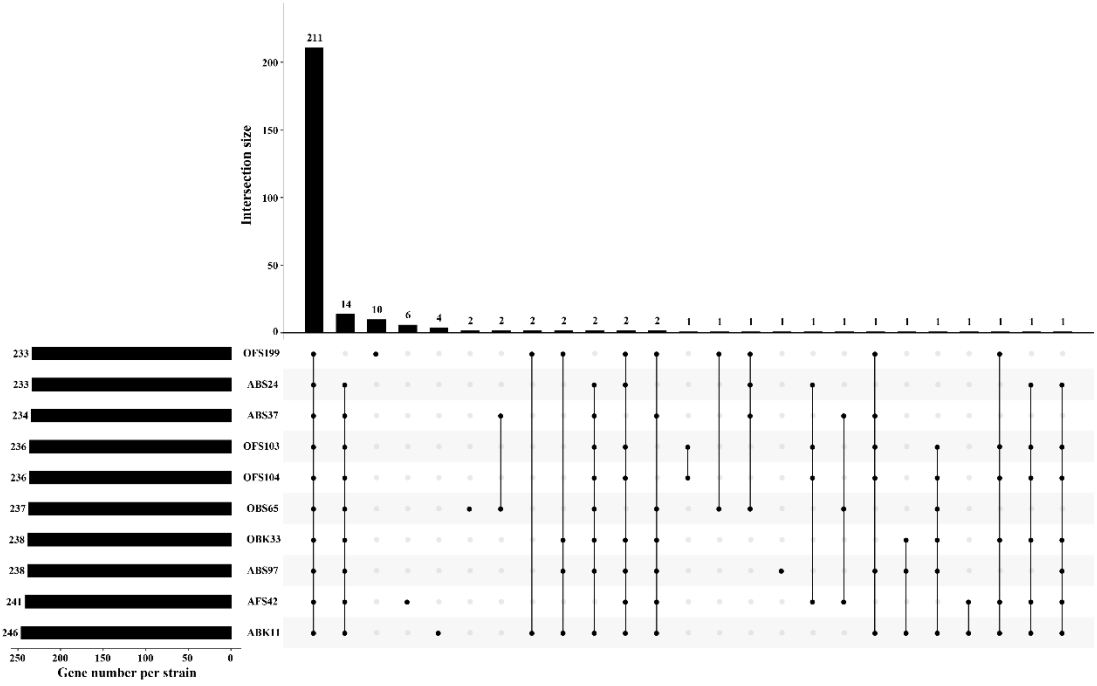

Fig. S5 Comparative analysis of amino acid metabolism gene profiles in *Ensifer* ASV205 strains based on KEGG annotation. An UpSet plot showing the distribution of amino acid metabolism-related genes among 10 *Ensifer* ASV205 genomes. The horizontal bars on the left represent the total number of such genes identified in each strain. The vertical bars above represent the number of shared or unique gene sets corresponding to the intersection matrix below.

Fig. S6 Phylogenetic trees of the denitrification genes *napA* and *nirK* from *Ensifer* strains using the Maximum-likelihood method. Phylogenetic trees were constructed based on the amino acid sequences of the *napA* (top) and *nirK* (bottom) genes. Red labels indicate sequences from *Ensifer* ASV205 strains isolated in this study.

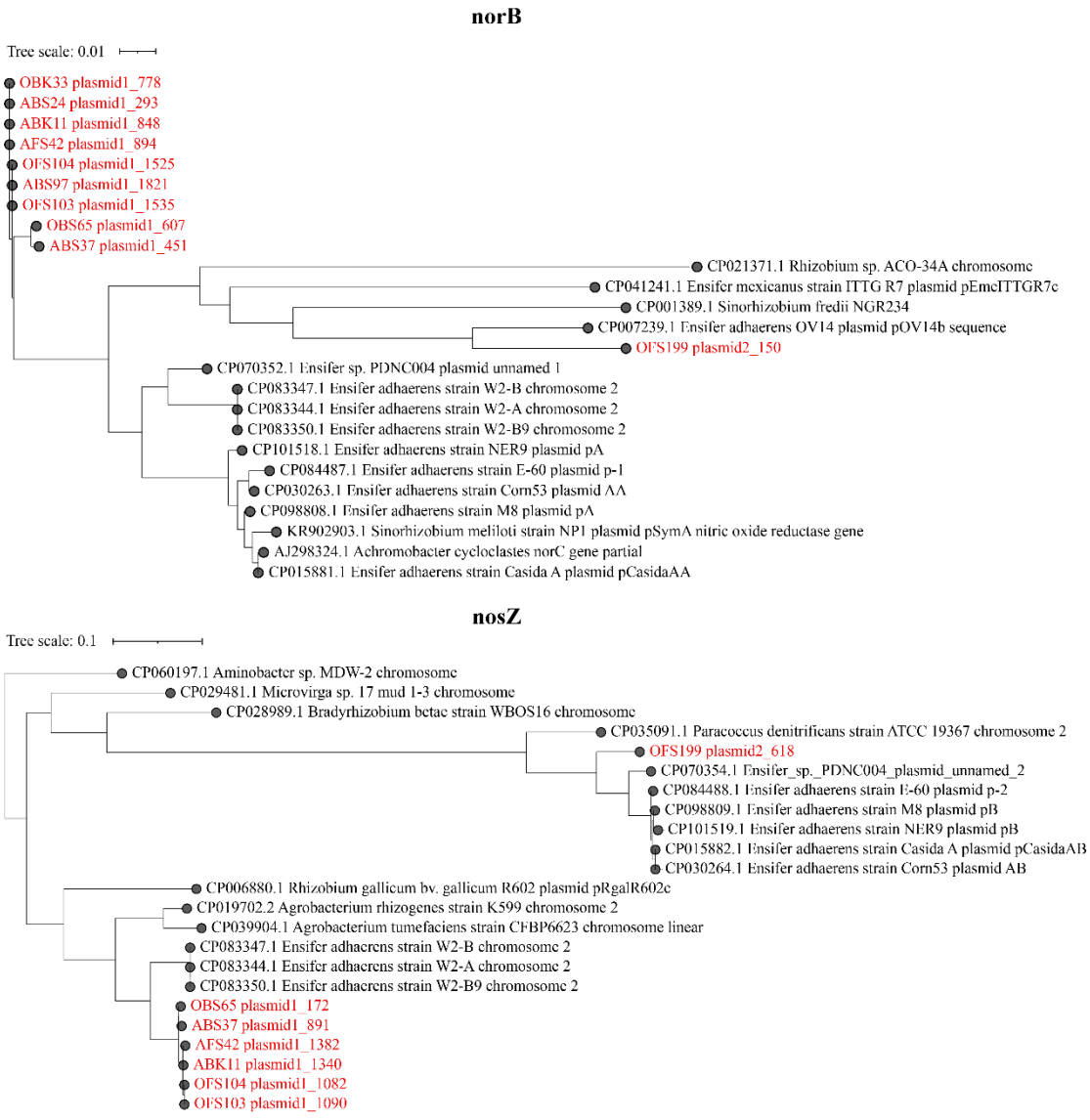

Fig. S7 Phylogenetic trees of the denitrification genes *norB* and *nosZ* from *Ensifer* strains using the Maximum-likelihood method. Phylogenetic trees were constructed based on the amino acid sequences of the *norB* (top) and *nosZ* (bottom) genes. Red labels indicate sequences from *Ensifer* ASV205 strains isolated in this study.

Supplementary Reference

1. Zhang J, Liu Y-X, Guo X, Qin Y, Garrido-Oter R, Schulze-Lefert P, et al. High-throughput cultivation and identification of bacteria from the plant root microbiota. Nat Protoc. 2021;16(2):988–1012; https://doi.org/10.1038/s41596-020-00444-7.

2. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 2013;30(4):772–80; https://doi.org/10.1093/molbev/mst010.

3. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972–3; https://doi.org/10.1093/bioinformatics/btp348.

4. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2019;37:1530 – 4; https://doi.org/10.1093/molbev/msaa015.

5. Bru D, Sarr A, Philippot L. Relative Abundances of Proteobacterial Membrane-Bound and Periplasmic Nitrate Reductases in Selected Environments. Appl Environ Microbiol. 2007;73(18):5971–4; https://doi.org/10.1128/AEM.00643-07.

6. Wei W, Isobe K, Nishizawa T, Zhu L, Shiratori Y, Ohte N, et al. Higher diversity and abundance of denitrifying microorganisms in environments than considered previously. ISME J. 2015;9(9):1954–65; https://doi.org/10.1038/ismej.2015.9.

7. Chee-Sanford JC, Connor L, Krichels A, Yang WH, Sanford RA. Hierarchical

detection of diverse Clade II (atypical) *nosZ* genes using new primer sets for classical- and multiplex PCR array applications. J Microbiol Methods. 2020;172:105908; https://doi.org/10.1016/j.mimet.2020.105908.

8. Jones CM, Graf DRH, Bru D, Philippot L, Hallin S. The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. ISME J. 2013;7(2):417–26; https://doi.org/10.1038/ismej.2012.125.

9. Wu Q, Ji M, Yu S, Li J, Wu X, Ju X, et al. Distinct Denitrifying Phenotypes of Predominant Bacteria Modulate Nitrous Oxide Metabolism in Two Typical Cropland Soils. Microb Ecol. 2023;86(1):509–20; https://doi.org/10.1007/s00248-022-02085-7.