Supplementary Materials

Simultaneous removal of humic acid and nitrate by combining micro-electrolysis and autotrophic denitrification

Guangxin Zhou [1], Wei Xing (✉)[1, 2, 3], Kexin Tian [1], Longsheng Li [1], Hong Yao [1, 2]

1 School of Environment, Beijing Jiaotong University, Beijing 100044, China

2 Tangshan Research Institute of Beijing Jiaotong University, Tangshan 063000, China

3 Beijing Key Laboratory of Novel Materials Genetic Engineering and Application for Rail Transit, Beijing Jiaotong University, Beijing 100044, China

1. 1 DNA extraction and PCR amplification

Total microbial genomic DNA was extracted using the E.Z.N.A.® soil DNA Kit (Omega Bio-tek, Norcross, GA, USA) according to manufacturer's instructions. The quality and concentration of DNA were determined by 1.0% agarose gel electrophoresis and a NanoDrop2000 spectrophotometer (Thermo Scientific, USA) and kept at −80℃ prior to further use. The hypervariable region V4 of the bacterial 16S rRNA gene were amplified with primer pairs ArBa515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′) synthesized by Sangon Biotech Co., Ltd. (>90% purity, Shanghai, China), by T100 Thermal Cycler PCR thermocycler (BIO-RAD, USA). The PCR reaction mixture including 4 μL 5 × Fast Pfu buffer, 2 μL 0.0025 mol/L dNTPs, 0.8 μL each primer (5×10^{-6} mol/L, >90% purity, Sangon Biotech, China), 0.4 μL Fast Pfu polymerase (high-fidelity, TransGen, China), 10 ng of template DNA, and ddH_2O to a final volume of 20 μL. PCR amplification cycling conditions were as follows: initial denaturation at 95℃ for 3 min, followed by 27 cycles of denaturing at 95℃ for 30 s, annealing at 55℃ for 30 s and extension at 72℃ for 45 s, and single extension at 72℃ for 10 min, and end at 4℃. The PCR product was

✉ Corresponding author
E-mail: weix@bjtu.edu.cn

extracted from 2% agarose gel and purified using the PCR Clean-Up Kit (YuHua, Shanghai, China) according to manufacturer's instructions and quantified using Qubit 4.0 (Thermo Fisher Scientific, USA) .

1. 2 Data processing

Raw FASTQ files were de-multiplexed using an in-house perl script, and then quality-filtered by fastp version 0.19.6 and merged by FLASH version 1.2.7 with the following criteria:

(i) the reads were truncated at any site receiving an average quality score of < 20 over a 50 bp sliding window, and the truncated reads shorter than 50 bp were discarded, reads containing ambiguous characters were also discarded; (ii) only overlapping sequences longer than 10 bp were assembled according to their overlapped sequence. The maximum mismatch ratio of overlap region is 0.2. Reads that could not be assembled were discarded; (iii) Samples were distinguished according to the barcode and primers, and the sequence direction was adjusted, exact barcode matching, 2 nucleotide mismatch in primer matching. Then the optimized sequences were clustered into operational taxonomic units (OTUs) using UPARSE 7.0.1090 with 97% sequence similarity level. The most abundant sequence for each OTU was selected as a representative sequence. To minimize the effects of sequencing depth on alpha and beta diversity measure, the number of 16S rRNA gene sequences from each sample were rarefied to 20000, which still yielded an average Good's coverage of 99.09%, respectively.

The taxonomy of each OTU representative sequence was analyzed by RDP Classifier version 2.2 against the 16S rRNA gene database (e.g. Silva v138) using confidence threshold of 0.7.

1. 3 Statistical Analysis

Bioinformatic analysis of the soil microbiota was carried out using the Majorbio Cloud platform. Based on the OTUs information, rarefaction curves and alpha diversity indices including observed OTUs, Chao1 richness, Shannon index and Good's coverage were calculated with Mothur v1.30.1.The similarity among the microbial communities in different samples was determined by Principal Component Analysis (PCA) based on Bray-curtis dissimilarity using Vegan v2.5-3 package. The PERMANOVA test was used to assess the percentage of variation explained by the treatment along with its statistical significance using Vegan v2.5-3 package.

Table S1 Primer sequences for each target gene used in this study.

Target genes	Primer	Sequences* (5'-3')
napA	napA1-F	5'-TCTGGACCATGGGCTTCAACCA-3'
	NapA2-R	5'-ACGACGACCGGCCAGCGCAG-3'
narG	1960m2-F	5'-TAYGTSGGGCAGGARAAACTG-3'
	2050m2-R	5'-CGTAGAAGAAGCTGGTGCTGTT-3'
nirS	cd3a-F	5'-GTSAACGTSAAGGARACSGG-3'
	R3cd-R	5'-GASTTCGGRTGSGTCTTGA-3'
nirK	nirK876-F	5'-ATYGGCGGVCAYGGCGA-3'
	nirK1040-R	5'-GCCTCGATCAGRTTRTGGTT-3'
Geobacter	564-F	5'-AAGCGTTGTTCGGAWTTAT-3'
	840-R	5'-GGCACTGCAGGGGTCAATA-3'
16sRNA	515-F	5'-GTGCCAGCMGCCGCGGTAA-3'
	806-R	5'-GGACTACHVGGGTWTCTAT-3'

*S Base: Represents G or C; R Base: Represents A or G; Y Base: Represents C or T; V Base: Represents G, A, or C; M Base: Represents A or C; W Base: Represents A or T; H Base: Represents A, T, or C.

Table S2 Diversity indices based on total genomic DNA.

Samples	Sobs	Ace	Chao	Shannon	Coverage
IS	1611	1910.5	1884.9	5.6	0.9924
R1	899	1119.4	1102.1	4.4	0.9951
R2	1365	1688.2	1661.0	4.5	0.9925
R3	1263	1577.1	1584.2	4.4	0.9929

Fig. S1 Photograph of the actual experimental reactors used in this study.

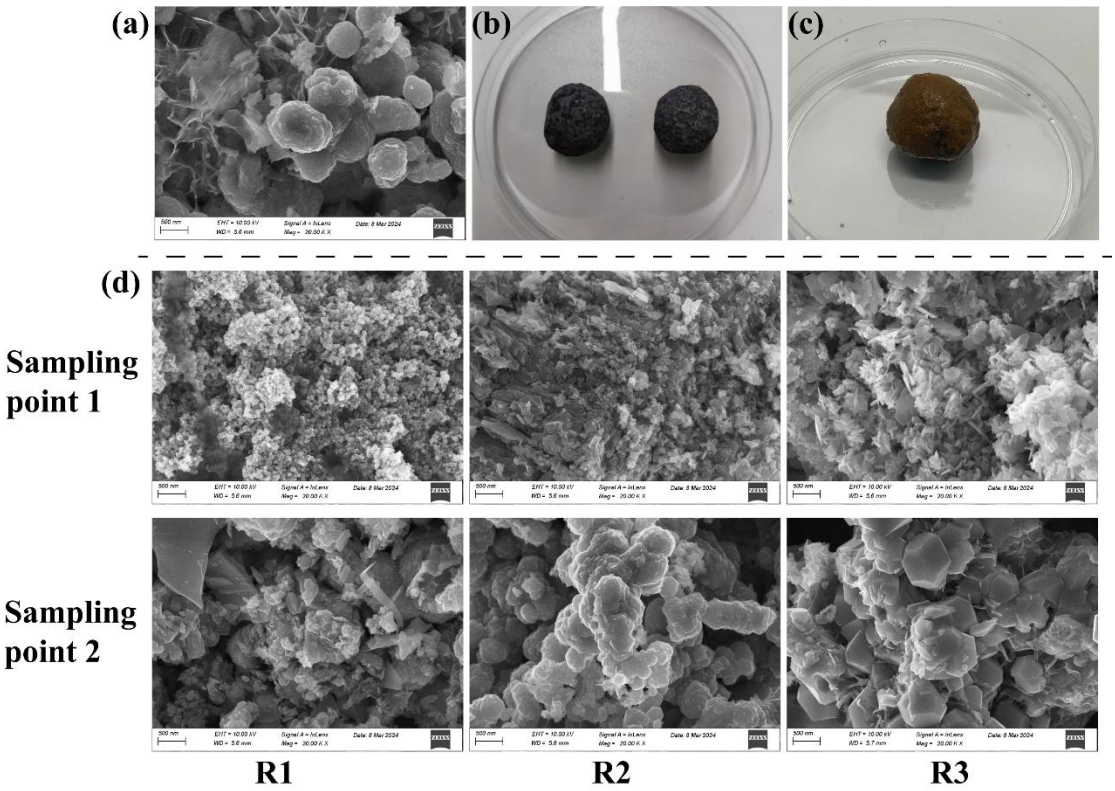

Fig. S2 SEM images and photographs of the iron–carbon carriers before and after long-term operation; (a) SEM image of the fresh iron – carbon carrier before operation; (b) photograph of the fresh carrier; (c) photograph of the carrier after long-term operation; (d) SEM images of the carriers at Sampling Point 1 and Sampling Point 2 after long-term operation in reactors R1, R2, and R3.

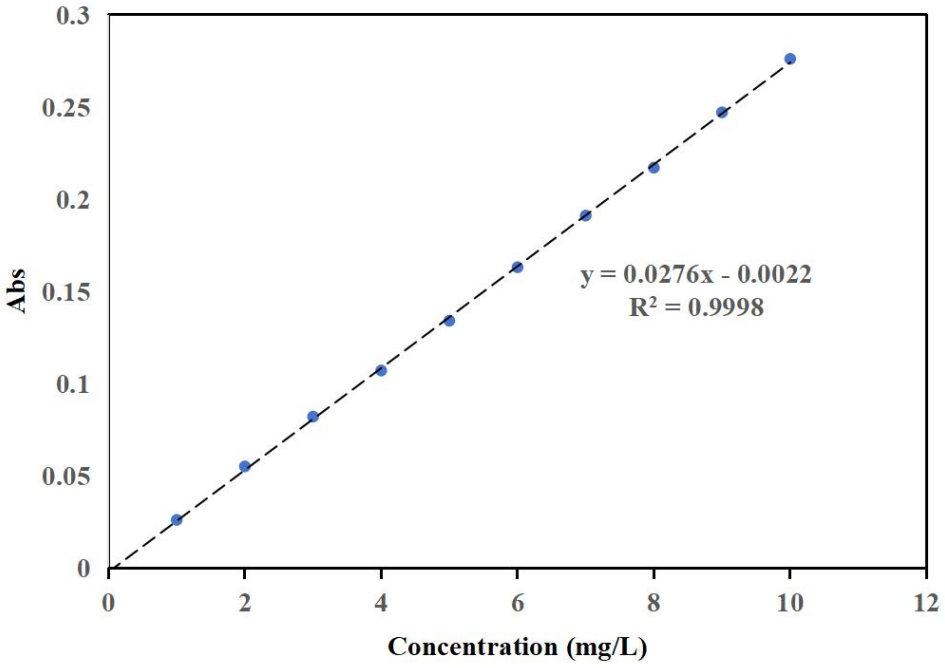

Fig. S3 Standard curve of humic acid concentration quantified by UV254 absorbance using sodium humate (0 - 10 mg/L; R^2 = 0.9998).

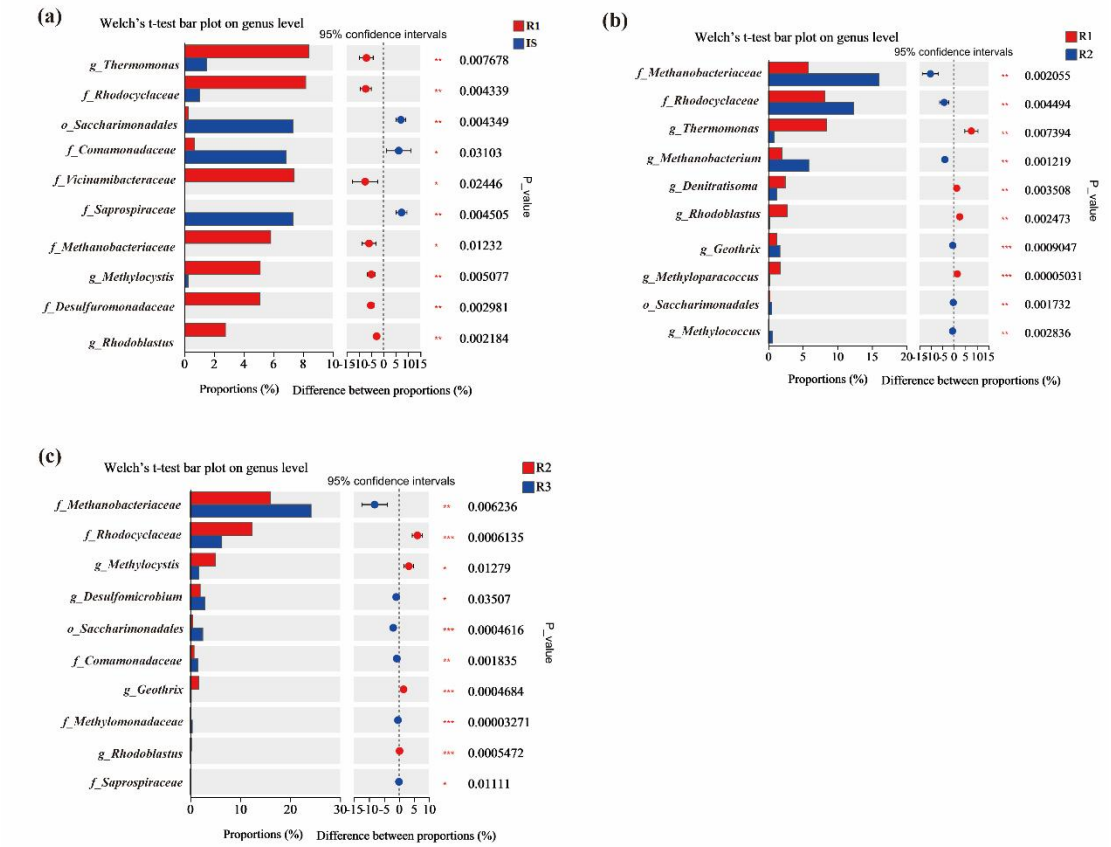

Fig. S4 Welch's *t*-test bar plots on the Genus level, **P* < 0.05; ***P* < 0.01 ;****P* < 0.001. (R1, without HA or glucose; R2, with 15 mg/L HA and without glucose; R3, with 15 mg/L HA and 2 mg/L glucose).