

Article

Archaeal Community and Function Disturbed Significantly in Surrounding Soil by Coal Gangue Stockpiling

Bianhua Zhang ^{1,2}, Dongsheng Jin ^{2,3,4,*}, Qiang Zhang ^{2,3,4}, Huijuan Bo ^{2,3,4} and Wei Wang ^{2,3,4}

¹ Department of Geography, Xinzhou Normal University, Xinzhou 034000, China; sxdxzbh@126.com

² Soil Health Laboratory in Shanxi Province, Shanxi Agricultural University, Taiyuan 030031, China; sxsnytf@163.com (Q.Z.); xinzhucej@163.com (H.B.); ww255286@126.com (W.W.)

³ College of Resources and Environment, Shanxi Agricultural University, Taiyuan 030031, China

⁴ Engineering Technology Innovation Center for Ecological Protection and Restoration in the Middle Yellow River, Shanxi Agricultural University, Taiyuan 030031, China

* Correspondence: sxdxjds@126.com

Abstract

Coal gangue (CG) dumped in open-air piles significantly impacts the surrounding soil environment. To investigate the effects of prolonged CG dumping on soil archaeal communities and their ecological functions, we used metagenomic sequencing to analyze soil samples, including control soil area not impacted by CG (CSL), undisturbed control sediment (CST), atmospheric dust fall area (ADF), and leachate flow area (LFA) samples. The results showed that the dominant phylum and genus of archaea were *Thaumarchaeota* (30.53–93.39%) and *Candidatus Nitrosocosmicus* (34.44–69.85%) in the different samples. Significant differences were observed in both α - and β -diversity ($p < 0.05$); archaeal community composition was primarily influenced by total nitrogen (TN), electrical conductivity (EC), Cu, As, and Cd. The contribution rate of As was the largest, about 44.8%. The metabolic functions of archaea were predominantly related to amino acid metabolism, and there were significant variations in carbon and nitrogen metabolic pathways in different areas. The *ppdk* gene showed considerable variation between ADF and CSL, and *Euryarchaeota* was the major contributing phylum to carbon fixation. However, for nitrogen metabolism, the *gltB* gene displayed marked differences, and the phylum of *Thaumarchaeota* was the major contributor. This study provides a theoretical foundation for land management and sustainable utilization in CG dump areas.

Keywords: archaeal communities; archaeal functions; disturbance zones soil; coal gangue dump



Academic Editor: Anna De Marco

Received: 23 August 2025

Revised: 3 October 2025

Accepted: 5 October 2025

Published: 14 October 2025

Citation: Zhang, B.; Jin, D.; Zhang, Q.; Bo, H.; Wang, W. Archaeal Community and Function Disturbed Significantly in Surrounding Soil by Coal Gangue Stockpiling.

Sustainability **2025**, *17*, 9094. <https://doi.org/10.3390/su17209094>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Coal gangue (CG) is a solid waste byproduct of coal mining, rich in mineral components, organic matter, and heavy metals [1–3]. By 2023, China's accumulated stockpile of CG had reached 7 billion metric tons [2,4]. Long-term open-air storage of CG releases coal dust, heavy metals, and other pollutants, which alter the physicochemical properties of surrounding soils, such as pH, organic carbon content, and heavy metal concentrations through atmospheric deposition, rainfall, infiltration, and leaching [5,6]. These changes further impact the structure and function of soil microbial communities in adjacent areas [7].

Numerous studies have investigated the ecological effects of CG storage sites, focusing primarily on heavy metals in nearby soils [8,9], water systems [10], and vegetation [11]. For instance, research reported that the total concentration of 16 priority PAHs (polycyclic

aromatic hydrocarbons) in surface soil exhibited a negative correlation with distance from coal gangue piles [12]. Leachate from CG sites increased water hardness and led to excessive levels of metals such as manganese and sodium [13], causing widespread contamination of surrounding surface water [10]. However, heavy metal pollution declined rapidly during the migration of gangue leachate [14]. Plants near CG sites demonstrate some capacity for heavy metal accumulation and translocation [15]; microbial community structure and function in adjacent soils vary significantly [7]. Though the abundance of archaea is relatively low in soil, they play ecologically significant roles as important functional species, such as methanogenic archaea mediating the methane cycle in soil and river sediments [16] and ammonia-oxidizing archaea in the global nitrogen cycle in extreme environments [17]. However, reports on the structure and function of archaeal communities in the soil surrounding CG storage sites remain limited. Therefore, we hypothesized that archaeal community function in the soil changes significantly across soils disturbed by CG storage, with community composition closely linked to environmental factors such as heavy metal content and nutrients. In this study, we analyzed the structure, functional genes, and metabolic pathways of archaeal communities in the surrounding soil and sediment influenced by CG dumps in Tunlan coal mine in Shanxi province in China, using metagenomic sequencing to provide a theoretical basis for ecological restoration, soil remediation, and riverine environmental management in CG storage.

2. Materials and Methods

2.1. Study Area Overview

Tunlan Mine (112°4' E, 37°54.6' N) is located in Gujiao City in Shanxi Province in China; it experiences a temperate continental climate with approximately 2808 h of annual sunshine, an average annual temperature of 9.5 °C, and an average annual precipitation of 420 mm primarily concentrated between July and September. The region is windy during winter and spring, with prevailing northwesterly winds. The city is rich in high-quality coal resources, with an approved annual production capacity of 16.5 million tons. Its major coal mines include Tunlan, Dongqu, Xiqu, Malan, and Zhenchengdi. Among these mines, Tunlan mine has the highest annual production capacity at approximately 4 million tons. Assuming a 15% gangue discharge rate during raw coal production, the annual output of its CG amounts to about 700,000 tons. It covered an area of approximately 32 hectares. The primary chemical components of its coal gangue include Al_2O_3 , SiO_2 , Fe_2O_3 , and various heavy metal elements such as Zn, As, Hg, Pb, Cr, and so on.

2.2. Sample Collection

Sampling was conducted in the vicinity of the CG dump site in Tunlan coal mine in Gujiao City. Four distinct disturbance zones were established: the control soil area (CSL) which encompassed unaffected surface soil surrounding the CG dump; atmospheric dust fall area (ADF), which was impacted by dry and wet atmospheric deposition; control sediment area (CST), which was an unaffected sediment zone; and leachate flow area (LFA) which was influenced by leachate runoff (Figure 1). Due to soil heterogeneity, each zone was subdivided into five sampling points with three replicate samples collected per point and homogenized into a single composite sample. All soil samples were immediately placed into separate, sterile bags, sealed, and temporarily stored in dry ice containers, then transported to the laboratory for processing within 24 h or preserved at $-80\text{ }^{\circ}\text{C}$ until analysis.

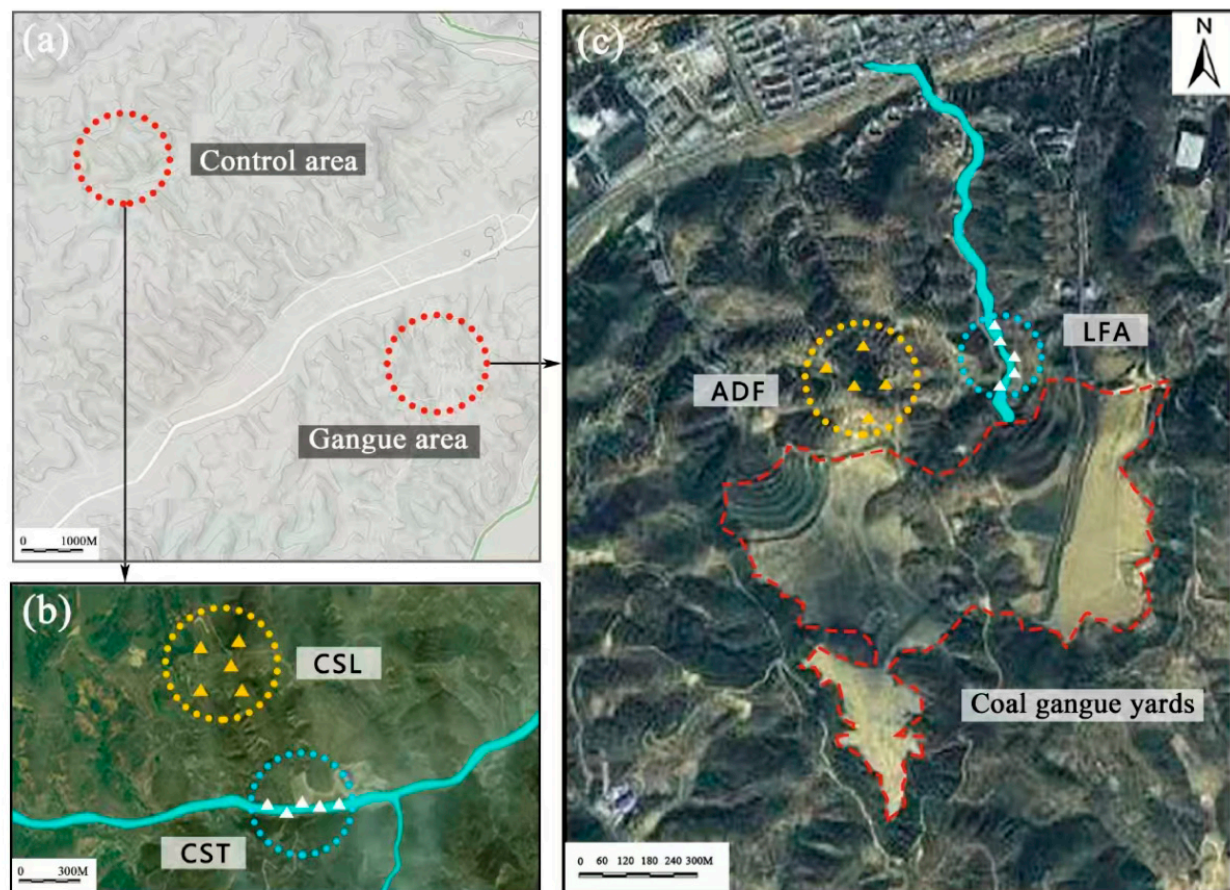


Figure 1. Location and division of the study area (a); enlarged view of the sampling locations of the control areas (b) and the disturbed areas (c). Circles are the sampling areas of each treatment zone, and triangles are the distribution of each sampling. CSL refers to the control soil area which was not impacted by CG, CST refers to the undisturbed control sediment area, ADF refers to the atmospheric dust fall area, LFA refers to the leachate flow area.

2.3. Measurement of Soil Characteristics

Soil pH value was measured by pH meter with a water-to-soil ratio of 2.5:1 [18]. The electrical conductivity (EC) was measured by the method of 5:1 water-to-soil ratio with an FE30 Plus instrument [19]. Soil available potassium (AK) content was measured by $\text{CH}_3\text{COONH}_4$ extraction with a flame photometer [20]. The available phosphorus (AP) was determined by UV spectrophotometer with NaHCO_3 molybdenum extraction [21]. Total potassium (TK) was determined by flame photometer using a NaOH melting method [22]. Soil organic carbon (SOC) was determined by a potassium dichromate volumetric method [23]. Total phosphorus (TP) was measured by NaOH molybdenum blue colorimetry and ultraviolet spectrophotometer [24]. Total nitrogen (TN) was measured using a nitrogen analyzer using the Kjeldahl nitrogen determination method [25]. Soil heavy metal content, including Cr, Cu, Zn, As, Cd, and Pb, was determined using an inductively coupled plasma emission spectrometer [26].

2.4. Metagenomic Sequencing

Metagenome DNA extraction and shotgun sequencing: Total microbial genomic DNA samples were extracted using the MagBeads FastDNA Kit for Soil (116564384) (MP Biomedicals, Santa Ana, CA, USA), following the manufacturer's instructions. The quantity and quality of extracted DNAs were measured using a QubitTM 4 Fluorometer, with WiFi: Q33238 (QubitTM Assay Tubes: Q32856; QubitTM 1X dsDNA HS Assay Kit: Q33231) (Invit-

rogen, Carlsbad, CA, USA) and agarose gel electrophoresis, respectively. The extracted microbial DNA was processed to construct metagenome shotgun sequencing libraries with insert sizes of 400 bp by using an Illumina TruSeq Nano DNA LT Library Preparation Kit. Each library was sequenced by Illumina NovaSeq platform (Illumina, San Diego, CA, USA) with PE150 strategy from Personal Biotechnology Co., Ltd. (Shanghai, China).

Raw sequencing reads were subjected to processing to yield quality-filtered reads suitable for further analysis. This involved adapter removal and trimming of low-quality reads using fastp, with the following parameters “-l 50-g-W 5-3-5-q 20-u 30”. Upon obtaining quality-filtered reads, taxonomic classification of each sample was executed using Kaiju software in greedy-5 mode against an NCBI database.

For assembly, Megahit (v1.1.2) was employed with the parameter settings “-k-list 33,55,77,99,127-min-contig-len 300” for each sample. Reads that remained unaligned to the assembled contigs, as determined by minimap2 with the “-ax sr” option, were pooled and reassembled. Gene prediction was carried out on the generated contigs from each sample using prodigal with the “-p meta” flag. The protein sequences translated from the predicted genes were subsequently pooled and clustered using mmseqs2 in “easylinclust” mode, with the following parameters: “-c 0.9-cov-mode 1-min-seq-id 0.95-cluster-mode 2”. The lowest common ancestor taxonomy of the non-redundant genes was ascertained using mmseqs2 in “taxonomy” mode, with parameters set to “-lca-mode 3-s 2”, by aligning them against an NCBI database. To assess gene abundances, high-quality reads from each sample were mapped onto the contigs using minimap2 with the “-ax sr” option, and read counts were computed using htseq. Abundance values in metagenomes were normalized using CPM (copies per kilobase per million mapped reads). The functional annotation of the non-redundant genes was accomplished through mmseqs2 against KEGG databases. Archaeal community composition was analyzed from the whole metagenome sequencing data using the archaea flags in the Kaiju taxonomic classification tool, which allows for domain-specific assignment of reads.

2.5. Data Processing

Statistical analysis of soil physicochemical properties and heavy metal content across different disturbed areas was performed by one-way ANOVA using SPSS 26 Statistics. And means were declared significant at ($p < 0.05$) using Tukey’s honestly significant different test [27]. Principal coordinate analysis (PCoA) based on the Bray–Curtis dissimilarity was used to assess similarities in archaea community structure and function among samples. Hierarchical clustering used the unweighted pair-group method with arithmetic means (UPGMA) to analyze the clustering effect. Redundancy analysis (RDA) was used to obtain the correlation between archaea community structure and soil environmental factors by Canoco5.0 software [28]. Chao1 (richness), Shannon (diversity), and Pielou_e (evenness) indices were calculated to assess α -diversity. The Kruskal–Wallis test was used to compare phylum and genus taxonomic units and functional genes across sample groups, with $p < 0.05$ indicating statistically significant differences. LEfSe (LDA effect size) analysis identified robust biomarkers between groups using the platform Galaxy. Non-redundant gene sequences were annotated and classified by comparing them against the KEGG database for taxonomic and functional profiling. After the normalization of the obtained data, archaeal function was analyzed.

3. Results

3.1. Soil Physicochemical Properties

The heavy metal contents of soils under different treatments are presented in Figure 2. The stockpiling of coal gangue increased heavy metal concentrations in surrounding soils,

with the following trends: Cr, Cu, and Zn were LFA > ADF > CST > CSL; As was LFA > ADF > CSL > CST; Cd was CSL > CST > ADF > LFA; and Pb was LFA > CST > CSL > ADF. As shown in Figure 3, only the soil in the LFA was slightly acidic, while the soils in the other areas were alkaline. The sediments in the LFA exhibited the highest levels of EC, SOC, and TN, while the soil in the ADF had the highest AP content.

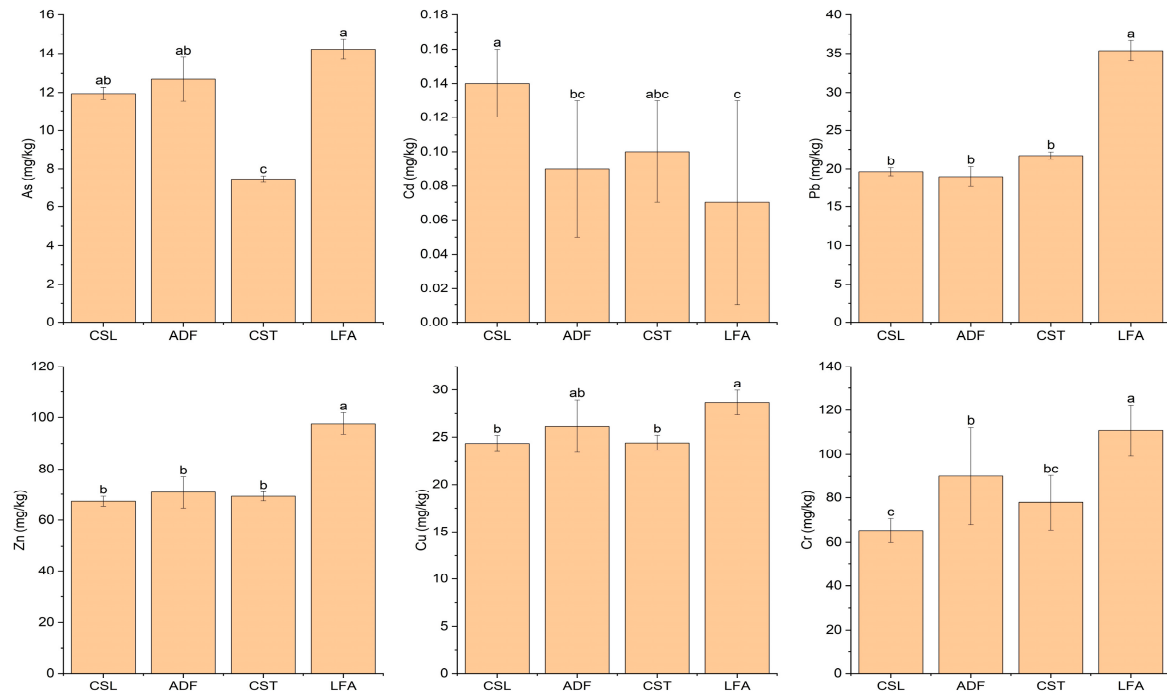


Figure 2. Contents of heavy metals in soil from different disturbed areas near CG dump; error bars represent standard deviation (SD) and different letters indicate significant differences among treatment zones. CSL refers to the control soil area which was not impacted by CG, CST refers to the undisturbed control sediment area, ADF refers to the atmospheric dust fall area, LFA refers to the leachate flow area. This is the same for all figures below.

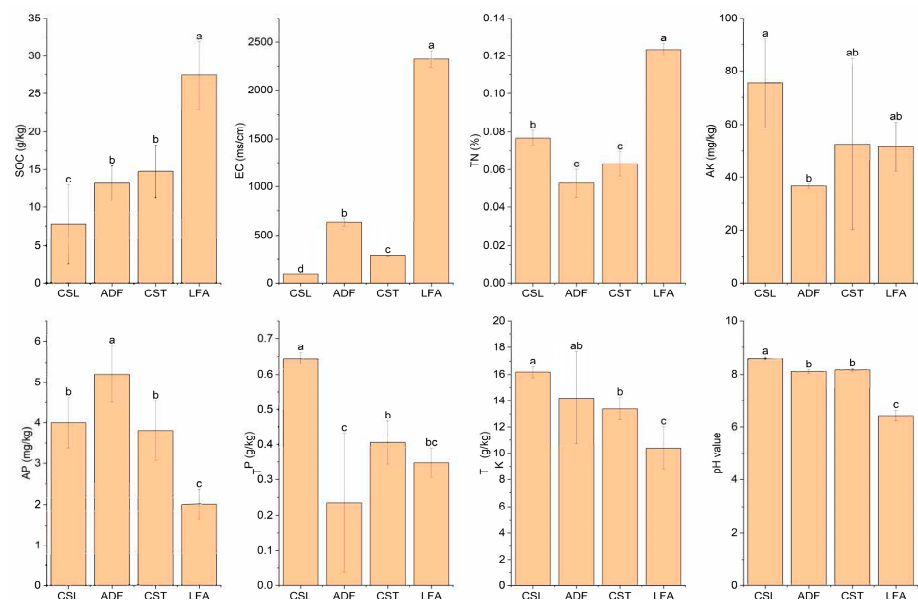


Figure 3. Soil chemical properties in different disturbed areas near CG dump; error bars represent standard deviation (SD) and different letters indicate significant differences among treatments.

3.2. Archaeal Community Diversity in Different Disturbed Areas

The proportion of valid sequences obtained by metagenomic sequencing was 96.85–97.51% under different treatment zones (Table S1). Significant differences ($p < 0.05$) were observed among treatment zones, as shown in Figure 4a–c, with the following hierarchy for all diversity indices: CST > LFA > ADF > CSL. Sediments in the LFA showed substantially altered archaeal communities due to strong coal gangue disturbance. From Figure 4d, the LFA has the highest number of unique species, while the CSL has the lowest, only seven. These results demonstrate that the storage of coal gangue created spatially heterogeneous effects on soil archaeal ecosystems, with the influence of leachate and atmospheric deposition differing significantly.

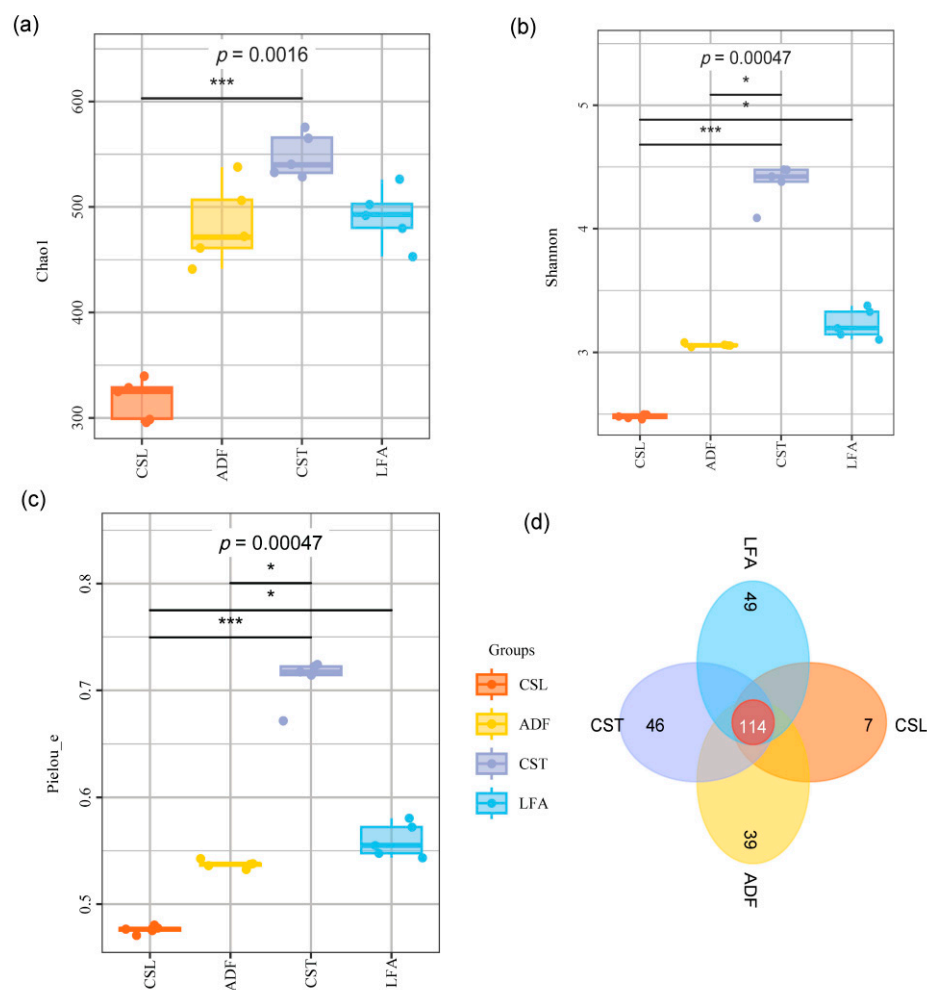


Figure 4. Alpha diversity analysis of archaeal communities in different areas disturbed by CG dump, Chao 1 richness index (a); Shannon diversity index (b); Pielou_e evenness index (c); Venn diagram showing the overlap of the number of archaeal species (d). p -value represents the result of the Kruskal–Wallis test, horizontal line represents the post hoc test between the corresponding two groups; *** represents $p < 0.001$, * represents $p < 0.05$.

The archaeal community of the ADF, CSL, LFA, and CST were clustered into four categories according to similarity, indicating that the disturbance of CG led to differences in archaeal community structures (Figure 5a). From the results of PCoA (Figure 5b,c), we could find that there were significant differences in the β -diversity of archaeal communities between the ADF and CSL, as well as a similar result between the LFA and. This explained why the archaeal community structure in the areas disturbed by the CG dump was significantly changed compared with that in the undisturbed areas.

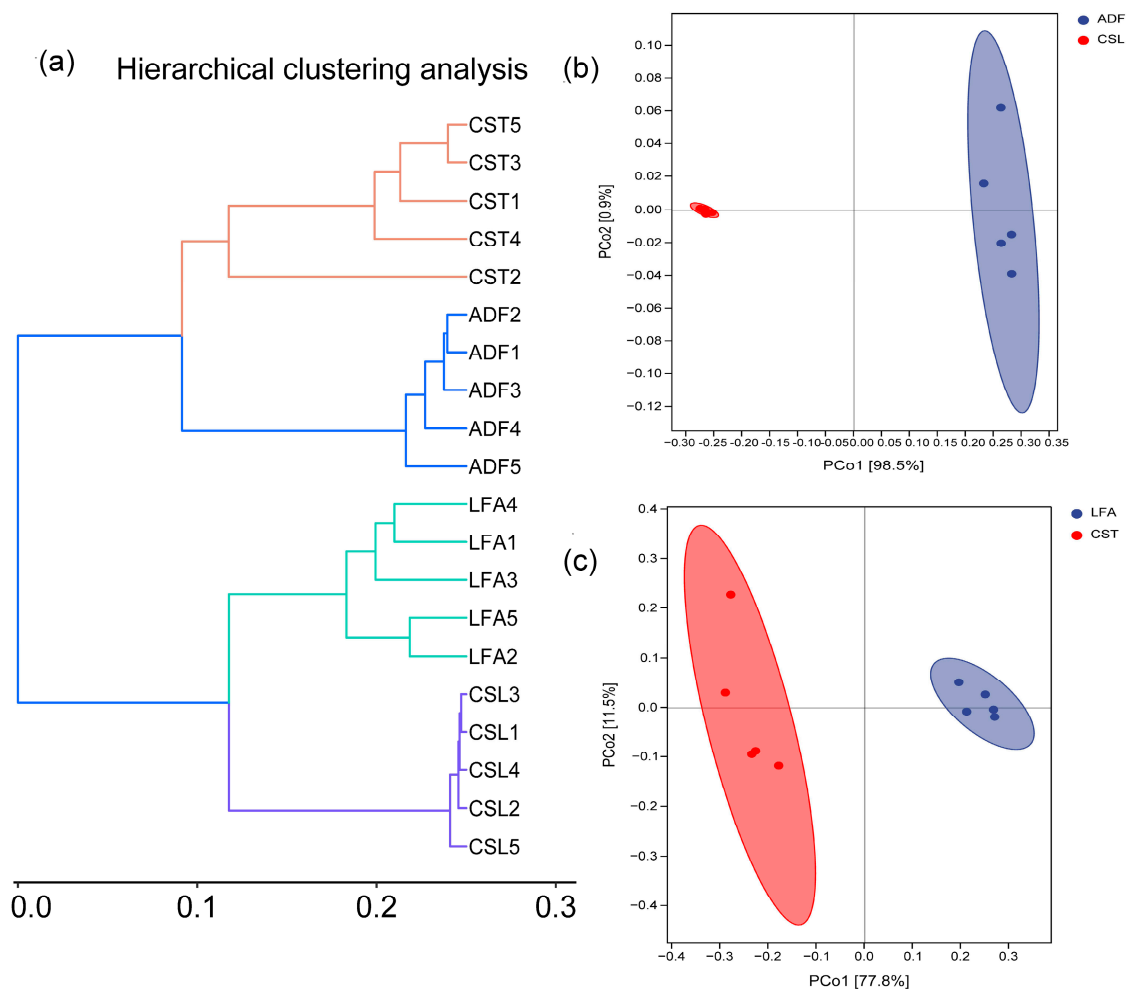


Figure 5. Beta diversity of archaeal communities in different areas disturbed by CG dump, hierarchical clustering tree diagram (a); principal coordinate analysis (PCoA) (b,c); the circle represents a 95% confidence ellipse.

3.3. Archaeal Community Structure in Different Disturbed Areas

Figure 6a revealed that the archaeal phyla with high relative abundance in the different areas were *Thaumarchaeota* (30.53–93.39%), *Euryarchaeota* (3.18–24.83%), *Candidatus Thermoplasmatota* (1.1–6.4%), *Candidatus Woesearchaeota* (0.07–15.29%), and *Candidatus Bathyarchaeota* (0.38–5.07%). *Thaumarchaeota* emerged as the predominant phylum; there were significant differences ($p < 0.05$) between disturbed areas and undisturbed areas (Figure 6c). At the genus level (Figure 6b), archaeal genera with high relative abundance in the CSL and ADF were *Nitrososphaera* (43.99%, 22.48%) and *Candidatus Nitrosocosmicus* (35.55%, 53.28%); in the CST, *Candidatus Nitrosocosmicus* (34.44%) and *Methanothrix* (18.20%); in the LFA, *Candidatus Nitrosocosmicus* (69.85%) and *Nitrososphaera* (11.84%). There were significant differences in the abundance of *Candidatus Nitrosopolaris* and *Methanocella* between disturbed and undisturbed areas ($p < 0.05$); there were larger differences in species between the LFA and CST than between the ADF and CSL. Lefse analysis (Figure 6e) identified 18 archaeal biomarkers with significant differences, 8 in CSL, 7 in ADF, 1 in CST, and 2 in LFA. These findings demonstrated restructuring of archaeal communities in CG-affected soils. From Figure 7, which is about environmental factors and archaeal community composition, we can see the disturbed areas (ADF and LFA) and undisturbed areas are in different quadrants, with the X-axis and Y-axis dimensions explaining 56.22% and 36.79% of the total variance, respectively. Key environmental factors significantly related to archaeal

community structure changes included TN, EC, As, Cu, and Cd, their contributions are, respectively, 7.80%, 37.70%, 44.80%, 5.40%, and 1.10% (Table 1).

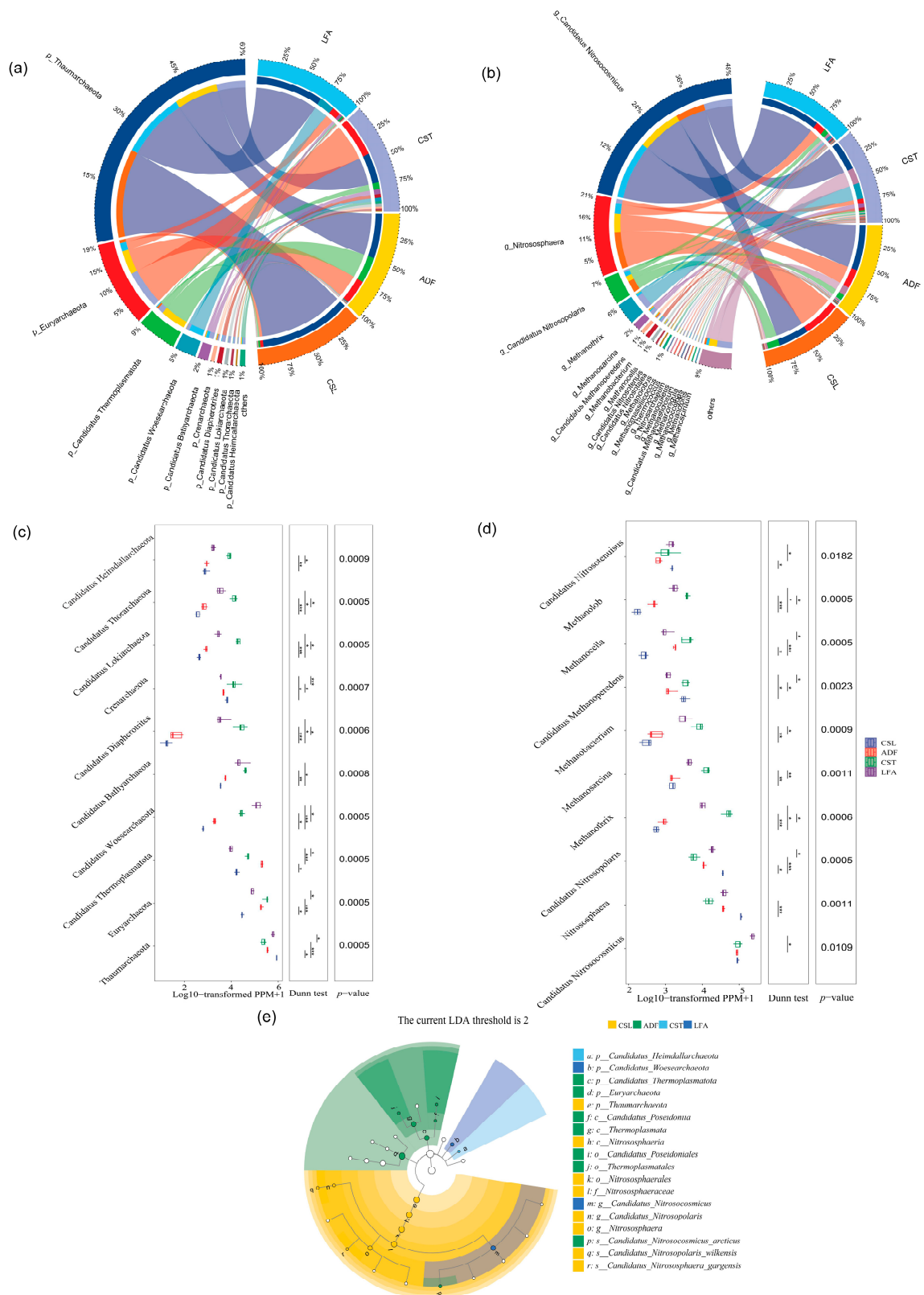


Figure 6. Archaeal community composition in different disturbed areas of CG dump; composition of archaea phyla (a) and genus (b), *** represents $p < 0.001$, ** represents $p < 0.01$, * represents $p < 0.05$; differences in the composition of archaea phyla (c) and genus (d) by Kruskal–Wallis test; Lefse analysis result (e).

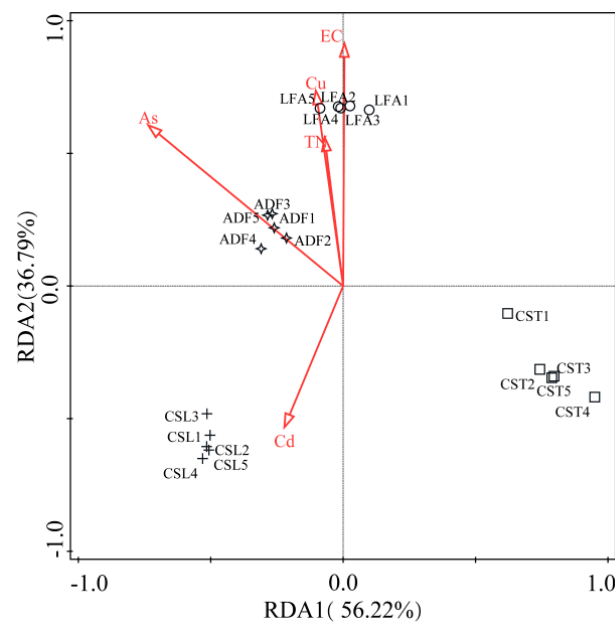


Figure 7. Redundancy analysis (RDA) of soil archaeal community composition and physical and chemical properties in different disturbed areas of CG dump.

Table 1. Results of RDA of environmental factors and archaeal community composition.

Name	Explains %	Contribution %	Pseudo-F	<i>p</i>
As	44	44.80	14.10	0.002
EC	37	37.70	33.20	0.002
TN	7.60	7.80	10.80	0.002
Cu	5.30	5.40	13.40	0.002
Cd	1.10	1.10	3	0.040

3.4. Archaeal Community Function in Different Disturbed Areas

3.4.1. Functional Genetic Composition of Archaeal Communities

Functional annotation using the KEGG database identified six major functional categories across samples. Metabolism (67.87–80.69%) was predominant, followed by genetic information processing (13.99–24.77%), and organismal systems (0.43–2.01%) accounted for the smallest proportion (Figure S1). At the secondary level, amino acid metabolism showed the highest abundance, followed by carbohydrate metabolism and metabolism of cofactors and vitamins among significantly different archaeal functions across disturbance zones ($p < 0.01$) (Figure 8a). The functional categories with a higher abundance of archaea included Ko0970, Ko00290, Ko00250, Ko3010, and Ko3030, which displayed significant variation (Figure 8b). Archaeal communities shared 122 metabolic functions in common across all soil samples. The LFA contained the highest number of unique functions, with 13, followed by the ADF with 7, the CST with 3, while the CSL had none (Figure 8c). LfSe analysis identified differentially abundant functional genes in different areas: 7 in CSL, 2 in ADF, 6 in CST, and 5 in LFA (Figure 8d). These results demonstrate significant alterations in archaeal gene function under coal gangue disturbance.

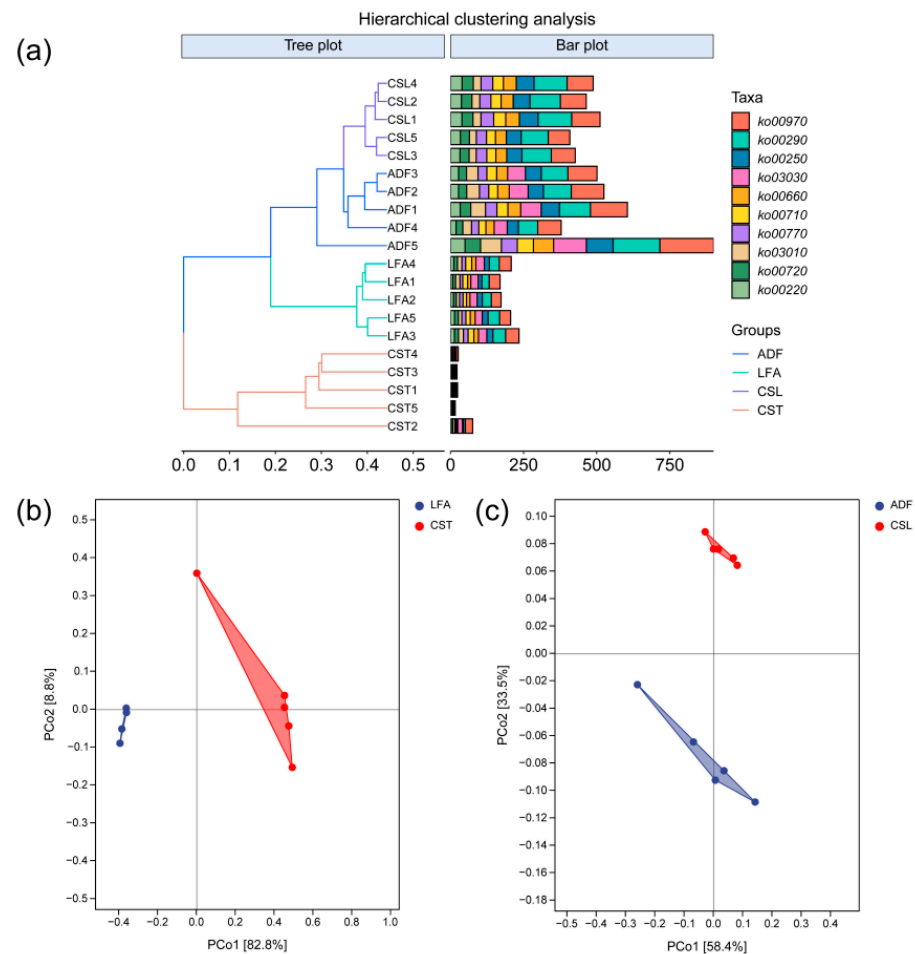


Figure 9. Functional diversity of archaeal communities in different disturbed areas of CG dump; represents result by hierarchical cluster (a); results by PCoA analysis (b,c).

(1) Carbon fixation function

Carbon plays a crucial role in regulating soil nutrient cycling and enhancing microbial activity. Using the KEGG database (map00720), the main microbial carbon fixation pathways include the Calvin cycle, reductive tricarboxylic acid (rTCA) cycle, dicarboxylate/4-hydroxybutyrate cycle, 3-hydroxypropionate/4-hydroxybutyrate cycle, 3-hydroxypropionate cycle, and reductive acetyl-CoA pathway. As can be seen from Figure 10, there were significant differences in carbon fixation genes between soil and sediment. In CSL and ADF soils, the most abundant genes included ACSS, accC, E5.4.99.2A, E2.3.1.9, ACO, E4.2.1.2B, mdh, sdhA, and sucC. Notably, the ppdK gene was significantly less abundant in the ADF compared to CSL. In the CST and LFA, highly expressed genes included MUT, acnB, E4.2.1.2A, pps, por, and fold, with ppc being the key differential gene significantly reduced in the LFA versus CST. The major archaeal genera contributing to carbon fixation (Figure 11) were *g_Unclassified_p_Thaumarchaeota* in the CSL and LFA; the contribution rates were 58.02% and 33.77%, respectively. The major archaeal genera were *g_Unclassified_p_Euryarchaeota* (25%) in the ADF, *g_Unclassified_d_Archaea* (34.92%) in the CST, *g_Candidatus Nitrosocosmicus* (28.81%) in the LFA. The results reflected distinct microbial adaptations to different disturbance conditions.

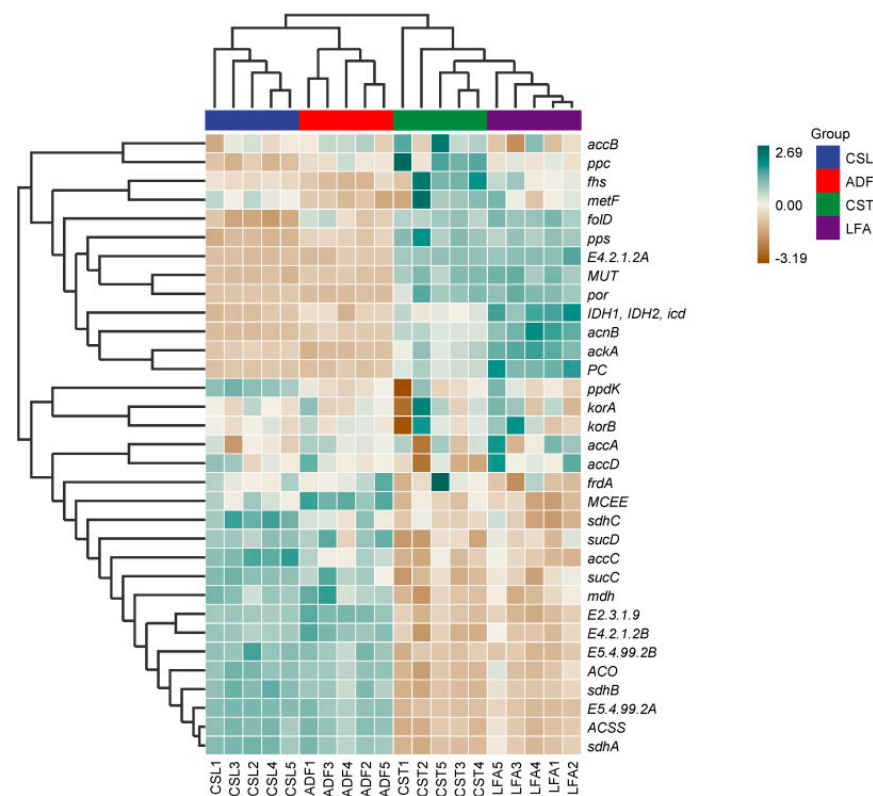


Figure 10. Differences in genes related to carbon fixation function in different disturbed areas of CG dump.

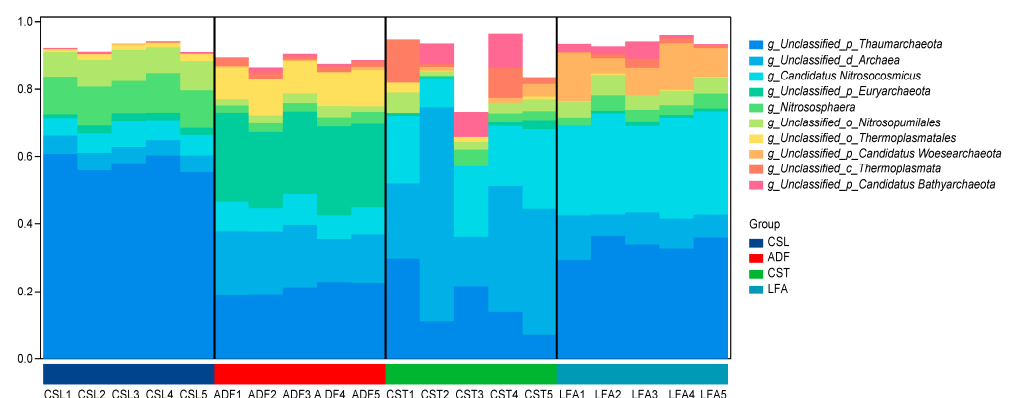


Figure 11. Contribution rate of archaea at the genus level to carbon fixation function in different disturbed areas of CG dump.

(2) Nitrogen Metabolism

Using the KEGG database (map00910), we analyzed archaeal genes involved in nitrogen metabolism and their relative contributions in different areas. From Figure 12, we can see the gene abundance of GLUD1_2, nirA, nirD, glnA, and GLUL in both the CSL and ADF were higher, but cah, gltD, nrfA, nrfH, NRT, nirS, and nrfH were higher in both the CST and LFA. In the ADF, the gltB gene increased while pmoC-amoC decreased significantly compared to in the CSL; in the LFA, arcC increased while cynT gene decreased significantly compared to the CST. The dominant archaeal genus contributing to nitrogen metabolism in the different areas was *Candidatus Nitrosocosmicus* (22.88–49.90%). And the main contributors in the CSL were *g_Unclassified_p_Thaumarchaeota* (30.77%), *g_Unclassified_d_Archaea* (24.60%), and *Candidatus Nitrosocosmicus* (22.88%); in ADF, they were *Candidatus Nitrosocosmicus* (26.24%) and *g_Unclassified_d_Archaea* (22.38%); it was *Candidatus Nitrosocosmicus* (25.85%)

in CST; and they were *Candidatus Nitrosocosmicus* (49.90%) and *g_Unclassified_d_Archaea* (21.19%) in LFA (Figure 13). These results demonstrated distinct archaeal involvement in nitrogen cycling processes across different areas, with *Candidatus Nitrosocosmicus* emerging as a key player, particularly in leachate-affected sediments.

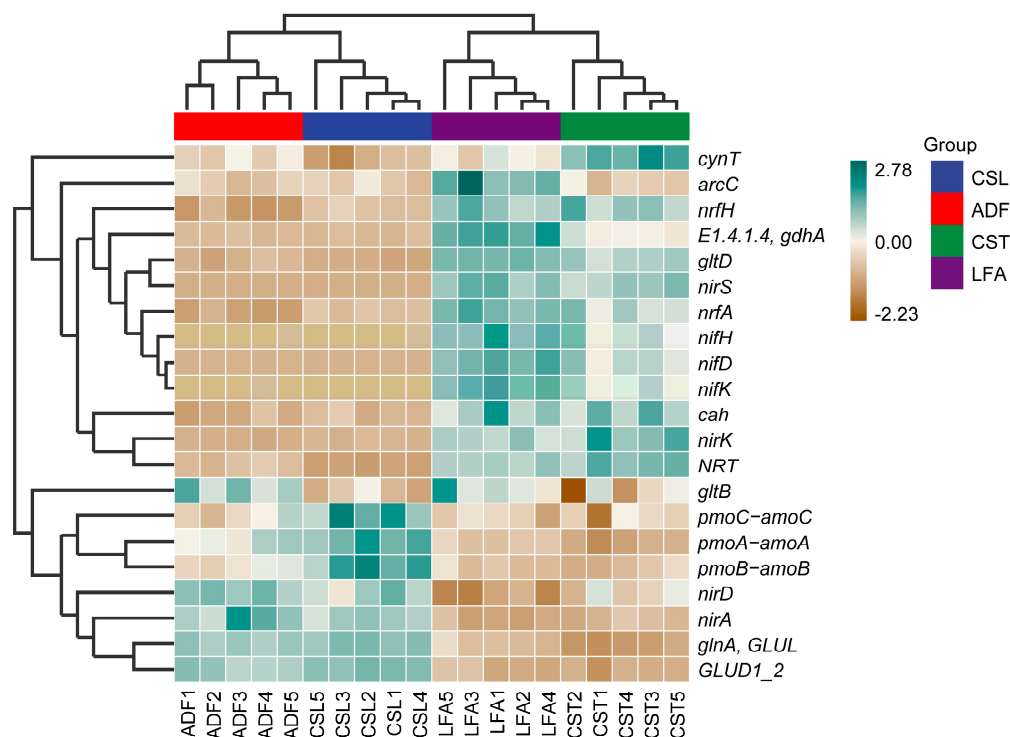


Figure 12. Differences in genes related to nitrogen metabolism function in different disturbed areas of CG dump.

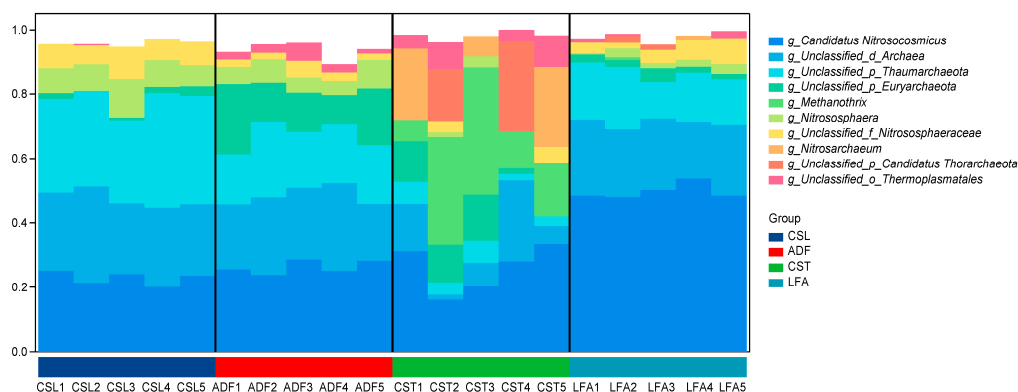


Figure 13. Contribution rate of archaea at the genus level to nitrogen metabolism function in different disturbed areas of CG dump.

4. Discussion

4.1. Variations in Soil Physicochemical Properties in Different Disturbed Areas of CG Dump

This study found that heavy metal concentrations in disturbed zones remained below background values (Table S2), which was consistent with findings by Yin et al. (2024) [29]. The leachate flow area (LFA) exhibited the highest levels of Cr, Cu, Zn, and Pb, with weakly acidic soil conditions ($\text{pH} < 7$). This pattern occurred because heavy metals from CG typically migrated via leaching [30], acidic conditions could promote dissolution of gangue-derived metals (Cr^{3+} , Cu^{2+} , Zn^{2+}) [31], and reduced soil adsorption capacity enhanced metal mobility and accumulation [32]; the result aligned with the report by

Tan et al. (2019) [33]. CG amendments alter soil properties (e.g., pH and EC) through redox processes [34]. Our data showed the LFA had significantly elevated EC, SOC, and TN, while the ADF showed notably increased AP, which demonstrated differential impacts of CG via leachate percolation and atmospheric deposition [31]. However, Yin et al. (2024) reported few differences between gangue-affected and surrounding soils [29], potentially due to geographic location, climatic conditions, and local soil parent material.

4.2. Archaeal Community Composition in Different Disturbed Areas of CG Dump

Common archaeal phyla in soils include *Crenarchaeota*, *Euryarchaeota* [35], and *Thaumarchaeota* [36]. *Euryarchaeota* (16.3–66.7%) dominates archaeal communities in estuarine sediments [37–39], and *Candidatus_Thermoplasmatota* (14.69–84.49%) and *Crenarchaeota* (2.10–23.43%) in soils surrounding CG piles [7]. However, we found *Thaumarchaeota* was the dominant phylum in soil and sediment, in comparison to the CSL, which was possibly due to regional climatic conditions and gangue composition. Archaeal richness, diversity and evenness in the LFA were significantly lower than those in the CST, which was consistent with reported results [7], but the ADF displayed higher diversity indices than the CSL, aligning with the results reported by Xie et al. (2012) [40]. These differences likely resulted from substantial alterations in soil physicochemical properties caused by various disturbance mechanisms such as dust deposition or leaching, which either promote or inhibit specific archaeal colonization during environmental adaptation [41]. For example, EC enhanced archaeal community diversity [42]; copper concentrations ≤ 100 mg/kg benefit bacterial abundance and diversity and concentrations ≥ 500 mg/kg inhibit them [43]; *Nitrososphaera* thrives in high-pH soda saline-alkali soils [38] and drives nitrification in acidic soils [44], potentially through metabolic adaptations that mitigate low-pH toxicity [45]. *Methanosarcina* preferentially colonizes acetate-rich environments [46]. Our study showed that EC, TN, Cu, As, and Cd had a great correlation with archaeal community structure, indicating these soil properties altered competitive relationships within archaeal communities, allowing tolerant species to emerge; alternatively the development of protective resistance mechanisms among some microorganisms in long-term disturbed environments by CG benefited other community members.

4.3. Composition and Diversity of Archaeal Function in Different Disturbed Areas of CG Dump

Archaeal communities occupy distinct ecological niches in nature, with unique function and mechanisms of change [47]. Our study showed that archaeal functions in disturbed soils were dominated by amino acid metabolism, followed by carbohydrate metabolism and cofactor/vitamin metabolism, which contrasts with reported findings where global/overview maps (28.26%) preceded amino acid metabolism [7], likely because of differences in archaeal community composition in the study area. Significant variations in functional gene profiles were observed among different areas ($p < 0.05$), and the LFA contained the highest number of unique functional gene categories, followed by the ADF; LEfSe and PCoA analyses confirmed pronounced differences between the CSL/ADF and CST/LFA pairs, indicating substantial genetic adaptations by archaea to leachate infiltration and atmospheric deposition disturbances [48].

4.4. Carbon Fixation and Nitrogen Metabolic Function of Archaeal Community in Different Disturbed Areas of CG Dump

The *ppdk* gene encodes a crucial enzyme that catalyzes the reaction between phosphoenolpyruvate and CO_2 to form oxaloacetate [49,50]. We found the abundance of the *ppdk* gene in the ADF was lower than in the CSL and the abundance of the *ppc* gene was lower in the LFA than in the CST, suggesting pollutants from the CG dump through atmospheric deposition, leaching, and infiltration processes have altered the archaeal community structures

in soils and sediments [51], consequently impairing the carbon fixation capacity of archaea. Dominant carbon-fixing archaea primarily from *Euryarchaeota*, which participate in anaerobic methane oxidation, organic matter degradation, and hydrocarbon transformation [52], differed substantially; in the ADF, there were predominantly *g_Unclassified_p_Euryarchaeota* but in the CSL there were *g_Unclassified_p_Thaumarchaeota*, which was different from the reported findings of *Thermoplasma* and *Metallosphaera* [7], which can be attributed to specific characteristics of the gangue piles and local environmental conditions.

Soil archaea participate in various nitrogen metabolic pathways including nitrogen fixation [53], nitrification [54], denitrification [55,56], dissimilatory nitrate reduction, assimilatory nitrate reduction, and complete nitrification [7]. *Crenarchaeota* contains numerous hyperthermophiles [57] and has been recognized as a significant contributor to nitrogen cycling [58]. Our study found that the key functional genes involved in nitrogen metabolism were *nirA*, *nirD*, and *amoA* in soils, while they were *nirS* and *NRT* in sediments. Archaeal genera contributing over 20% to nitrogen metabolism primarily included *Candidatus Nitrosocosmicus* from the *Thaumarchaeota* phylum due to their crucial role as ammonia-oxidizing archaea with high diversity and importance in the nitrogen cycle [35,38], enabling adaptation to diverse environmental conditions [52,59]. We observed the *gltB* gene, which is involved in nitrogen assimilation and amino acid metabolism, showed higher abundance in the ADF than CSL; the gene of *cynT*, which assists acidophilic/acid-tolerant archaea in pH adaptation, was more prevalent in the CST than LFA, while the *ArcC* gene, which participates in arginine degradation to release NH_3 and regulate pH, was significantly enriched in the LFA compared to CST. These findings also demonstrated that distinct disturbances led to adaptive changes in archaeal functional gene expression and metabolic strategies, reflecting archaeal community responses to environmental stressors induced by CG through various transport pathways.

5. Conclusions

Our metagenomic analysis identified key archaeal taxa and functional genes that could serve as potential targets for soil remediation. The study revealed significant differences in archaeal diversity across coal gangue-affected areas, with notable variations in archaeal community structures. The dominant archaeal phylum was *Thaumarchaeota*, and the most abundant genus was *Candidatus Nitrosocosmicus*. Archaeal community composition showed strong correlations with TN, EC, Cu, As, and Cd. Key functional differences included the *ppdk* gene, which exhibited significant variation between the ADF and CSL, with *Euryarchaeota* being the major contributing phylum to carbon fixation; the *gltB* gene, which displayed marked differences and was primarily associated with *Thaumarchaeota* and nitrogen metabolism. The long-term storage of CG may exert prolonged and cumulative effects on soil archaeal communities, necessitating further research on its ecological consequences. Future studies should integrate transcriptomic and metabolomic analyses to better elucidate microbial activity and their actual contributions to carbon and nitrogen cycling in these disturbed ecosystems, as well as long-term observations from a seasonal perspective to assess the stability of, and seasonal variations in, archaeal communities.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/su17209094/s1>: Figure S1: Functional composition of archaeal communities in different disturbed areas of CG dump.; Table S1: Clean data and qcSummary in different samples; Table S2: Risk Control Standard for soil contamination (mg/kg); Graphical Abstract S1: Archaeal community and function disturbed significantly in surrounding soil by coal gangue stockpiling.

Author Contributions: Writing—original draft, data curation, methodology, B.Z.; conceptualization, methodology, administration, D.J.; supervision, methodology, Q.Z.; data curation, methodology,

H.B.; methodology, conceptualization, data curation, W.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the fundamental Research Program of Shanxi Province, grant number 202203021221223; Major Project of Science and Technology in Shanxi Province, grant number 202201140401028; Soil health laboratory in Shanxi Province, grant number 2024003; Xinzhou Science and Technology Bureau key research and development plan project, grant number 20240301; and Project of Engineering Technology Innovation Center for Ecological Protection and Restoration in the Middle Yellow River, grant number HHZY-2025-007.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author(s).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADF	atmospheric dust fall area
AK	available potassium
AP	available phosphorus
CSL	control soil area
CST	control sediment area
EC	electrical conductivity
LFA	leachate flow area.
SOC	soil organic carbon
TK	total potassium
TP	total phosphorus
TN	total nitrogen

References

1. Chang, J.; Du, G.J.; Du, J.; Shi, X. Current situation of the comprehensive utilization of coal gangue in China and the related problems and recommendations. *China Environ. Prot. Ind.* **2022**, *2022*, 13–17. [\[CrossRef\]](#)
2. He, Z.; He, S.; Wang, G.; Li, S.; Chen, G.; Gao, M.; Li, H. Research and progress of comprehensive utilization technologies of China's coal gangue in the context of carbon peaking and carbon neutrality. *Conserv. Util. Miner. Resour.* **2024**, *44*, 1–14. [\[CrossRef\]](#)
3. Gong, W.; Hui, H.; Ma, S.; Ji, J.; Jiang, H. Preparation of Coal Gangue-Based Artificial Soil and Investigation of the Mechanism of Aggregate Structure Formation. *Sustainability* **2025**, *17*, 3318. [\[CrossRef\]](#)
4. Shi, L.; Peng, J.; Xu, D.; Tian, J.; Liu, T.; Jiang, B.; Zhang, F. Leaching characteristics and pollution risk assessment of potentially harmful elements from coal gangue exposed to weathering for different periods of time. *Environ. Sci. Pollut. Res.* **2023**, *30*, 63200–63214. [\[CrossRef\]](#)
5. Liu, F.; Song, H.; Wu, H.; Feng, Z.; Chai, C.; Zhang, J.; Zhou, J. Research progress on the utilization of coal gangue for soil remediation and as soil amendment agents. *Conserv. Util. Miner. Resour.* **2023**, *43*, 14–26. [\[CrossRef\]](#)
6. Deng, Y.; Wu, H.; Zhao, T.; Shi, C.; Zhang, Y.; Li, F. Characteristics of atmospheric dustfall fluxes and particle size in an open pit coal mining area and surrounding areas. *Sci. Rep.* **2025**, *15*, 9597. [\[CrossRef\]](#)
7. Chen, C.; Chen, L.; Li, Y.; Tang, X.; Wang, L.; Zhao, R. Structural and functional characteristics of microbial communities of typical coal gangue dumps in Guizhou province. *Microbiol. China* **2023**, *50*, 5300–5319. [\[CrossRef\]](#)
8. Shang, Y.; Sang, N. Pollution characteristics and phytotoxicity of heavy metals in the soil around coal gangue accumulation area. *Environ. Sci.* **2022**, *43*, 3773–3780. (In Chinese) [\[CrossRef\]](#)
9. Wang, Y.; Li, X.; Li, J.; Li, W.; Zhang, M.; Zhao, C.; Zhang, L.; Li, W. Heavy Metal Pollution Characteristics and Risk Evaluation of Soil Around Coal Gangue Stockpile Area. *Bull. Chin. Ceram. Soc.* **2021**, *40*, 3464–3471+3478. [\[CrossRef\]](#)
10. Lu, H.; Min, W.; Jiang, S.; Lu, Y.; Gou, F.; Cai, T. Analysis of Impact of Coal Gangue Leaching on Surrounding Water in Pangzhuang Coal Mine. *Adm. Tech. Environ. Monit.* **2023**, *35*, 64–67. [\[CrossRef\]](#)
11. Li, N.; Geng, X.; Kang, Y. Investigation on crop planting safety in surrounding environment of coal gangue pile in Jingxi coal mine. *Appl. Chem. Ind.* **2021**, *50*, 2909–2912. [\[CrossRef\]](#)
12. Wang, X.; Zhong, N.; Han, X. Impacts of coal gangue stockpiling on polycyclic aromatic hydrocarbons pollution in soil environment. *Acta Sci. Circumstantiae* **2013**, *33*, 3092–3100. [\[CrossRef\]](#)

13. Liu, Z.; Zhang, Y.; Wang, J. Research of influence of leaching water of gangue on environments. *J. Liaoning Technol. Univ.* **2005**, *24*, 280–283. [\[CrossRef\]](#)
14. Zhao, Z.; Sha, H.; Huang, J.; He, X.; Pan, Q.; Liu, X. Study on the characteristics and causes of groundwater pollution in coal gangue dumps. *J. Environ. Eng. Technol.* **2023**, *13*, 1604–1613. [\[CrossRef\]](#)
15. Yan, H.; Song, L.; Jiang, X.; He, C.; Yan, K.; Zhan, F.; Li, B.; Li, Y. Accumulation characteristics and ecological risk assessment of heavy metals in dominant species of different plant communities in coal gangue dump. *Chin. J. Ecol.* **2025**, *44*, 1–11. Available online: <https://www.cje.net.cn/EN/10.13292/j.1000-4890.202506.040> (accessed on 10 June 2025).
16. She, C.; Tong, C. Vertical distribution of methanogen community structures in *Phragmites australis* marsh soil in the Min River estuary. *Acta Ecol. Sin.* **2012**, *32*, 5299–5308. [\[CrossRef\]](#)
17. Zhou, L.; Wang, S.; Zou, Y.; Xia, C.; Zhu, G. Species, abundance and function of ammonia-oxidizing Archaea in Inland Waters across China. *Sci. Rep.* **2015**, *5*, 15969. [\[CrossRef\]](#)
18. Pan, X.; Zhang, S.; Zhong, Q.; Gong, G.; Wang, G.; Guo, X.; Xu, X. Effects of soil chemical properties and fractions of Pb, Cd, and Zn on bacterial and fungal communities. *Sci. Total Environ.* **2020**, *715*, 136904. [\[CrossRef\]](#)
19. Lu, R. *Analytical Methods of Soil Agrochemistry*; Agricultural Science and Technology Press: Beijing, China, 2000.
20. Chen, M.; Zhang, S.; Liu, L.; Wu, L.; Ding, X. Combined organic amendments and mineral fertilizer application increase rice yield by improving soil structure, P availability and root growth in saline-alkaline soil. *Soil Tillage Res.* **2021**, *212*, 105060. [\[CrossRef\]](#)
21. Tang, S.; Nguyen-Sy, T.; Cheng, W.; Sato, C.; Tawaraya, K.; Shiono, H.; Kumagai, K. Long-term application of fused magnesium phosphate and calcium silicate change soil chemical properties, C decomposition and N mineralization in a single rice paddy field of north eastern Japan. *Soil Sci. Plant Nutr.* **2022**, *68*, 149–157. [\[CrossRef\]](#)
22. Wang, X.; Zhang, Z.; Yu, Z.; Shen, G.; Cheng, H.; Tao, S. Composition and diversity of soil microbial communities in the alpine wetland and alpine forest ecosystems on the Tibetan Plateau. *Sci. Total Environ.* **2020**, *747*, 141358. [\[CrossRef\]](#)
23. Wu, X.; Xu, H.; Liu, G.; Ma, X.; Mu, C.; Zhao, L. Bacterial communities in the upper soil layers in the permafrost regions on the Qinghai-Tibetan plateau. *Appl. Soil Ecol.* **2017**, *120*, 81–88. [\[CrossRef\]](#)
24. Yang, T.; Tang, G.; Li, L.; Ma, L.; Zhao, Y.; Guo, Z. Interactions between bacteria and eukaryotic microorganisms and their response to soil properties and heavy metal exchangeability nearby a coal-fired power plant. *Chemosphere* **2022**, *302*, 134829. [\[CrossRef\]](#)
25. Zhao, M.; Cong, J.; Cheng, J.; Qi, Q.; Sheng, Y.; Ning, D.; Lu, H.; Wyckof, K.N.; Deng, Y.; Li, D.; et al. Soil microbial community assembly and interactions are constrained by nitrogen and phosphorus in broadleaf forests of southern China. *Forests* **2020**, *11*, 285. [\[CrossRef\]](#)
26. Liu, B.; Yao, J.; Ma, B.; Chen, Z.; Zhao, C.; Zhu, X.; Li, M.; Cao, Y.; Pang, W.; Li, H.; et al. Microbial community profiles in soils adjacent to mining and smelting areas: Contrasting potentially toxic metals and co-occurrence patterns. *Chemosphere* **2021**, *282*, 130992. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Nyuma, H.T.; Njoroge, R.; Otinga, A.N. Agroforestry adoption and its influence on soil quality under smallholder maize production systems in western Kenya. *PLoS ONE* **2025**, *20*, e0313385. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Guo, Y.; Tian, J.C.; Wang, Z. Composition and functional diversity of soil and water microbial communities in the rice-crab symbiosis system. *PLoS ONE* **2025**, *20*, e0316815. [\[CrossRef\]](#)
29. Yin, M.; Sheng, W.; Zhang, X.; Wu, Y.; Ma, X.; Cui, Z.; Bo, H.; Zheng, G.; Liu, L.; Guo, W. Effects of long-term coal gangue dumping on soil chemical environment and microbial community in an abandoned mine. *Land. Degrad. Dev.* **2024**, *35*, 4923–4934. [\[CrossRef\]](#)
30. Jiang, B.; Zhang, B.; Li, L.; Zhao, Y.; Shi, Y.; Jiang, Q.; Jia, L. Analysis of microbial community structure and diversity in surrounding rock soil of different waste dump sites in Fushun western opencast mine. *Chemosphere* **2021**, *269*, 128777. [\[CrossRef\]](#)
31. Feng, Q.; Liu, G. Harmful microelements contained in gangue from Yanzhou coal mine and their influence to the soil. *China Min. Mag.* **2002**, *11*, 67–69. [\[CrossRef\]](#)
32. Ribeiro, J.; Silva, E.; Li, Z.; Ward, C.; Flores, D. Petrographic, mineralogical and geochemical characterization of the Serrinha coal waste pile (Douro Coalfield, Portugal) and the potential environmental impacts on soil, sediments and surface waters. *Int. J. Coal Geol.* **2010**, *83*, 456–466. [\[CrossRef\]](#)
33. Tan, L.; Feng, S.; Cheng, C.; Guo, P. Study on Migration of Heavy Metals in Coal Gangue of Zhuxianzhuang under Leaching Action. *Guangzhou Chem. Ind.* **2019**, *47*, 131–133. [\[CrossRef\]](#)
34. Li, J.; Wang, J. Comprehensive utilization and environmental risks of coal gangue: A review. *J. Clean. Prod.* **2019**, *239*, 117946. [\[CrossRef\]](#)
35. Zhang, L.; He, J. A novel archaeal phylum: Thaumarchaeot—A review. *Acta Microbiol. Sin.* **2012**, *52*, 411–421. [\[CrossRef\]](#)
36. Linta, R. Metagenome-assembled genomes reveal unique metabolic adaptations of a basal marine Thaumarchaeota lineage. *ISME J.* **2020**, *14*, 2105–2115. [\[CrossRef\]](#)
37. Li, X.F.; Hou, L.J.; Liu, M. Archaeal community structure and diversity in intertidal sediments of the Yangtze River Estuary. *China Environ. Sci.* **2019**, *39*, 1744–1752. [\[CrossRef\]](#)

38. Yu, B.; Yang, S.; Zhu, L.; Gao, F.; Zeng, C.; Liu, S.; Cui, W. Diversity characteristic of archaeal community responding to soil anion in saline-alkali soil. *China Environ. Sci.* **2018**, *38*, 2731–2739. [CrossRef]
39. Niu, Y.; Yuan, R. Effects of water diversion and seasonal change on archaea community in river sediments. *China Environ. Sci.* **2020**, *40*, 1294–1304. [CrossRef]
40. Xie, X.; Fan, F.; Yuan, X.; Zhu, W.; Liu, N.; Liu, J. Impact on microbial diversity of heavy metal pollution in soils near Dexing copper mine tailings. *Microbiol. China* **2012**, *39*, 624–637. [CrossRef]
41. Lin, Y.; Dang, C.; Zhong, S.; Wang, J.; Zheng, T.; Ni, J. Community Characteristics of Dominant Archaea before and after the Danjiangkou Dam. *Acta Sci. Nat. Univ. Peking* **2020**, *56*, 509–517. [CrossRef]
42. Zhang, Y.; Kong, Q.; Guo, D.; Jing, J.; Jiang, A.; Guo, J. The response of soil archaeal community structure to halophyte vegetation succession in the Yellow River Delta. *China Environ. Sci.* **2016**, *36*, 2162–2168. [CrossRef]
43. Li, D.; Chen, J.; Zhang, H.; Li, J. Effects of copper pollution on soil bacterial community structure and heavy-metal resistance genes. *Acta Sci. Circumstantiae* **2021**, *41*, 1082–1090. [CrossRef]
44. Li, W.; Zheng, M.; Wang, C.; Shen, R. Nitrososphaera may be a major driver of nitrification in acidic soils. *Soils* **2021**, *53*, 13–20. [CrossRef]
45. Lehtovirta-Morley, L.; Sayavedra-Soto, L.; Gallois, N.; Schouten, S.; Stein, L.; Prosser, J.; Nicol, G. Identifying potential mechanisms enabling acidophily in the ammonia-oxidizing archaeon “candidatus nitrosotalea devanaterre”. *Appl. Environ. Microbiol.* **2016**, *82*, 2608–2619. [CrossRef]
46. Tourna, M.; Stieglmeier, M.; Spang, A.; Könneke, M.; Schintlmeister, A.; Urich, T.; Engel, M.; Schlöter, M.; Wagner, M.; Richter, A.; et al. *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8420–8425. [CrossRef]
47. Marchal, I. Mining archaea for antibiotics. *Nat. Biotechnol.* **2025**, *43*, 1425. [CrossRef]
48. Chen, L.; Hu, M.; Huang, L.; Hua, Z.; Kuang, J.; Li, S.; Shu, W. Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME J.* **2015**, *9*, 1579–1592. [CrossRef] [PubMed]
49. Wang, Z.; Li, H.; He, Y.; Zeng, H. Advances in plant pyruvate, orthophosphate dikinase. *Plant Physiol. J.* **2012**, *48*, 949–957. [CrossRef]
50. Fang, L.; Ding, Z.; Zhao, M. Characteristics of Drought Tolerance in ppc Overexpressed Rice Seedlings. *Acta Agron. Sin.* **2008**, *34*, 1220–1226. Available online: <https://zwxb.chinacrops.org/CN/Y2008/V34/I07/1220> (accessed on 12 July 2008). [CrossRef]
51. Shi, G.; Peng, J.; Song, Y.; Yuan, L.; Jiang, S. Spatial distribution of methanogenic archaea diversity in the sediments of Hun River. *Acta Sci. Circumstantiae* **2016**, *36*, 1766–1773. [CrossRef]
52. Baker Brett, J.; De Anda, V.; Seitz Kiley, W.; Dombrowski, N.; Santoro Alyson, E.; Lloyd Karen, G. Diversity, ecology and evolution of Archaea. *Nat. Microbiol.* **2020**, *5*, 887–900. [CrossRef]
53. Arnold, W.; Rump, A.; Klipp, W.; Priefer, U.; Pühler, A. Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*. *J. Mol. Biol.* **1988**, *203*, 715–738. [CrossRef]
54. Liu, J.; Wu, W.; Ding, Y.; Shi, D.; Chen, Y. Ammonia-oxidizing archaea and their important roles in nitrogen biogeochemical cycling. *Chin. J. Appl. Ecol.* **2010**, *21*, 2154–2160. [CrossRef]
55. Chen, Q.; Fan, J.; Ming, H.; Su, J.; Wang, Y.; Wang, B. Effects of environmental factors on denitrifying bacteria and functional genes in sediments of Bohai Sea, China. *Mar. Pollut. Bull.* **2020**, *160*, 111621. [CrossRef] [PubMed]
56. Regan, K.; Stempfhuber, B.; Schlöter, M.; Rasche, F.; Prati, D.; Philippot, L.; Boeddinghaus, R.; Kandrlar, E.; Marhan, S. Spatial and temporal dynamics of nitrogen fixing, nitrifying and denitrifying microbes in an unfertilized grassland soil. *Soil Biol. Biochem.* **2017**, *109*, 214–226. [CrossRef]
57. Takai, K.; Olda, H.; Suzuki, Y.; Hirayama, H.; Nakagawa, S.; Nunoura, T.; Inagaki, F.; Nealson, K.; Horikoshi, K. Spatial distribution of marine crenarchaeota group I in the vicinity of deep-sea hydrothermal systems. *Appl. Environ. Microbiol.* **2004**, *70*, 2404–2413. [CrossRef] [PubMed]
58. Yakimov, M.; Cono, V.; Smedile, F.; Deluca, T.; Juárez, S.; Ciordia, S.; Fernández, M.; Albar, J.; Ferrer, M.; Golyshin, P.; et al. Contribution of crenarchaeal autotrophic ammonia oxidizers to the dark primary production in Tyrrhenian deep waters (Central Mediterranean Sea). *ISME J.* **2011**, *5*, 945–961. [CrossRef]
59. Laanbroek, H.; Keuskamp, J.; Lourenço, K.; Costa, O.; Hefting, M. Changing communities of ammonia-oxidizing Betaproteobacteria and Thaumarchaea at discrete vegetation zones along a hill slope grassland gradient. *Appl. Soil Ecol.* **2025**, *214*, 106393. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.