



Soil Microbiome and its Functional Attributes Under the Gradient of Arsenic Contamination in Paddy Soils

Mariya Naseem · Praveen C. Verma · Richa Raghuvanshi · Vivek K. Gaur ·
Manisha Singh · Srishti Seth · Pankaj Kumar Srivastava

Received: 30 April 2024 / Accepted: 1 August 2024 / Published online: 8 August 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract The build-up of arsenic in agricultural soil through contaminated irrigational groundwater is a concern. Metagenomic analysis of such contaminated sites may provide a remarkable opportunity to extensively investigate the responses and adaptation of microbial communities to different levels of arsenic. The study focuses on the comparative analysis of microbial community composition and associated functions in paddy soil samples with a gradient of arsenic contamination (4.88 to 43.67 mg kg⁻¹). Actinobacteria was found to dominate the bacterial phyla in all four samples with abundance ranging from 39.77% to 49.39% followed by Proteobacteria (20.71–38.24%). Whereas the fungal phylum Ascomycota (92.42–95.29%) dominated all the samples studied. In the study, bacteria were found to be abundant in

the mid-level (15.89 and 24.84 mg kg⁻¹) of arsenic, which included genera *Gaiella*, *Nocardioides*, *Solirubrobacter*, *Microvirga*, and *Nitrospira*. Fungi *Beauveria*, *Talaromyces*, *Aspergillus*, *Pyrenophora*, and *Valsa* were higher in relative abundance corresponding to soil arsenic concentration. *Verticillium*, previously reported for Pb, Zn, and Cd removal, was found in the soil sample with the mid-arsenic concentration (15.89 mg kg⁻¹). The relative abundance of the arsenic metabolizing/ transport genes of native microbial communities also varied with the soil arsenic gradient, the genes *arsJ*, *arsM*, *aioR*, *arsH*, and *arsC* being the most affected. The study is the first report of *Gaiella*, *Solirubrobacter*, *Beauveria*, and *Verticillium* present in arsenic-contaminated soil, and further studies are required to explore their potential role in arsenic bioremediation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11270-024-07412-x>.

M. Naseem · R. Raghuvanshi
Department of Botany, Institute of Science, Banaras Hindu University, Varanasi, India

M. Naseem · P. C. Verma · M. Singh · S. Seth ·
P. K. Srivastava (✉)
Environmental Technologies Division, CSIR-National Botanical Research Institute, Lucknow, India
e-mail: drpankajk@gmail.com

V. K. Gaur
School of Energy and Chemical Engineering, UNIST, Ulsan 44919, Republic of Korea

Keywords Arsenic · Metagenomics · Microbes · Diversity · Genes · Agriculture

1 Introduction

Arsenic (As) is a ubiquitous element in the earth's crust. It is a metalloid that exists in different allotropic forms including carbonate, sulfide, and elemental forms (Genchi et al., 2022a) Globally, the As concentration in the soil varies viz., Bangladesh (4 to 137.9 mg kg⁻¹), United States (2.8 to 73 mg kg⁻¹), China (1.9 to 36.0 mg kg⁻¹), Japan

(8.5–10.30 mg kg⁻¹), South Korea, (2–4.6 mg kg⁻¹), India (3.7 to 423 mg kg⁻¹) (Khan et al., 2010; Kumar et al., 2016; Mandal & Suzuki, 2002; Ori et al., 1993; Patel et al., 2005; Srivastava & Sharma, 2013; Zhou et al., 2018).

Contamination and exposure to As have been reported to severely affect the environment and human health. It is categorized as a group I human carcinogen causing skin, lungs, bladder, liver, and kidney cancer (Genchi et al., 2022b). Exposure to As can cause skin lesions, cardiovascular diseases, birth defects, and cognitive impairment (Domingo-Relloso et al., 2022; Hamadani et al., 2011; Monteiro De Oliveira et al., 2021; Rudnai et al., 2014). Millions of people spanning 70 different countries are affected by the use of As-contaminated groundwater (Shrivastava et al., 2015). The As-tainted groundwater also contaminates agricultural soil, leading to toxicity and reduced growth in plants viz., straight-head disease in paddy (Mishra et al., 2021).

Soil-dwelling microorganisms can potentially detoxify As through the process of absorption, precipitation, accumulation, and further chemical transformation through redox and/or methylation (Singh et al., 2015). These microbial As detoxification/resistance mechanisms can be exploited for As bioremediation at the contaminated sites. Various bacterial and fungal species including *Bacillus subtilis*, *Bacillus cereus*, *Acidithiobacillus ferrooxidans*, *Saccharomyces cerevisiae*, and *Aspergillus niger* manifest arsenic biosorption (Chandraprabha & Natarajan, 2011; Giri et al., 2013; MasudHossain & Anantharaman, 2006; Mitra et al., 2017; Zoroufchi Benis et al., 2020). Some microbes such as *Trametes versicolor* are known to hyper-accumulate arsenic within their cells (Adeyemi, 2009). *Bacillus* sp. strain DJ-1 is reported to accumulate up to 9.8 mg of As per g of its dry weight (Joshi et al., 2009). *Westerdykella aurantiaca*, *Neosartorya fischeri*, *Aspergillus* sp., *Rhizopus* sp., and *Humicola* sp. have been reported to methylate As, suggesting their role in arsenic detoxification by formation of organic As⁺⁵ derivatives which are relatively less toxic (Čerňanský et al., 2009; Srivastava et al., 2011; Tripathi et al., 2020). Intracellular methylation of arsenate has been demonstrated in *Trichoderma asperellum*, *Penicillium janthinellum*, and *Fusarium oxysporum* (Su et al., 2012).

Genomic studies have led to the identification of novel microbial species and enzymes involved in the

detoxification of As, which can further be used in As bioremediation. Metagenomics is one such approach that can serve as an effective tool to analyze native microbial communities and to provide comprehensive information on soil community composition, and their potential functional traits (Castro-Severyn et al., 2021). Unlike the culture-dependent approach, metagenomics targets all the microbes and their associated genes hence providing a holistic picture. Xiao et al. (2016) conducted a metagenomic study in five paddy soils contaminated with As (< 16 mg kg⁻¹) and reported the relative abundance of distinct As metabolizing genes in different samples, among which the arsenate reduction genes were found to be dominant. In another study conducted on As-contaminated mine samples, a high abundance of Proteobacteria with Gammaproteobacteria dominating all the samples was found. Bacterial phylum Chlorobi and Bacteroidetes were observed to occur only in the soil samples with high As concentration (~ 821.23 mg kg⁻¹). Among the functional diversity, higher relative abundance and diversity of As-responsive genes including *arsC*, *arrA*, *aioA*, *arsB*, and ACR3 were reported in soil samples with high concentrations of As and Sb (Antimony). The gene *arsC*, responsible for arsenate reduction had the highest mean relative expression, whereas arsenic methylating gene *arsM* showed the lowest expression (Luo et al., 2014).

An understanding of microbial dynamics at As-contaminated sites is needed to explore novel microbes expressing novel genes, enzymes, and processes that may prove helpful in As bioremediation, especially in agricultural ecosystems. The agricultural ecosystem is directly responsible for the health and well-being of the population dependent on it. Therefore, arsenic contamination in these areas may have serious implications. Since microbes inhabiting the rhizosphere of the crops grown in As contaminated regions influence As geochemical cycling, its chemical form, and bioavailability, it becomes essential to study the impact of changing soil arsenic concentration on the native microbial diversity and abundance (Gu et al., 2017). This study thus highlights the role of the As gradient in shaping the microbial community and their functions in the contaminated paddy soil using a comprehensive metagenomic-driven approach. It was hypothesized that the variation of soil arsenic contamination may have a differential impact on the soil microbial communities,

their taxonomic diversity, and their functions. This study divulges microbial community compositions and functions at different levels of As-contaminated soils (DNA level), which provides a comprehensive understanding of the relationship between microbial community characteristics and As contamination in the soil.

2 Experimental Procedures

2.1 Sample Collection, Chemicals, and Reagents

Soil samples were collected from four distant sites namely, Ghorhat (SKNNNGH), Siswania (SKNNNSIS), Sakhi (SKNNNSK), and Majhaua (SKNNNMJ) of an administrative block Nathnagar in the district Sant Kabir Nagar (26.55N – 26.66N, 82.95E – 83.08E) lying in the North Eastern Plains of Uttar Pradesh, India (Fig. S1). The district Sant Kabir Nagar was selected based on our previous soil and groundwater arsenic contamination mapping data. Block Nathnagar was chosen within the district based on availability of a gradient of soil arsenic contamination and groundwater As concentration. The groundwater total As concentration ($\mu\text{g l}^{-1}$) in samples SKNNNGH, SKNNNSIS, SKNNNSK, and SKNNNMJ were 19.93, 86.90, 98.66, and 143.93, respectively. The samples observed with the gradient of arsenic contamination were then further used for metagenomic studies.

Soil samples were collected from the rhizosphere of the paddy fields at the time of harvesting at physiological maturity. The same paddy variety was grown by the farmers at all four sites of sampling. The soil samples were drawn from the topsoil (0–10 cm) of agricultural land using an iron core (60 mm internal diameter). At a site, nine parallel samples from paddy fields located within a 1 ha area were collected and mixed thoroughly to form a site-wise representative sample. The samples were stored in sterile polyethylene bags and aseptically transferred to the laboratory for sieving through a 2 mm sieve. One subsection of the fresh soil sample was stored at $-20\text{ }^{\circ}\text{C}$ for further analysis. The remaining soil was air-dried and used to analyze soil physico-chemical parameters. Based on soil As contents, samples were classified into three ranges of arsenic gradient viz., low

(4.88 mg kg^{-1}), mid (15.89 and 24.84 mg kg^{-1}), and high (43.67 mg kg^{-1}).

The multielement standards were procured from Thermofisher Scientific, Germany. Trace metal grade Nitric acid and Hydrogen Peroxide were used for digesting the samples for ICP-MS analysis. The chemicals used for soil physicochemical and biological analysis were of reagent-grade procured from Himedia Pvt Ltd.

2.2 Soil physico-chemical and biological Analysis

The physicochemical and biological properties of soil samples were analyzed initially to get baseline information on soil characteristics. The soil parameters like bulk density (BD) and soil texture were analyzed using the method given by Jackson (1969). Analysis of Water Holding Capacity (WHC) was performed by Keen's box method. The soil pH and EC were measured using Multi 9630 IDS multimeter (WTW, Germany) in a soil suspension with a 1:5 ratio of soil to ultrapure Milli Q water. Microbial biomass Carbon (MBC) was analyzed by the chloroform-fumigation extraction method and Total Organic Carbon (TOC) by the Walkley and Black method (Vance et al., 1987; Walkley, 1947). The soil available nitrogen and phosphorus were analyzed by the Kjeldahl and Olsen methods, respectively (Bremner, 1960; Olsen, 1954). The available potassium was analyzed using a flame photometer (Systronics 128). The activity of soil enzyme Dehydrogenase was analyzed by spectroscopic quantification of its reaction product Triphenylformazan (TPF), formed by reduction of 2,3,5 Triphenyl tetrazolium chloride (TTC) when used as substrate (Małachowska-Jutysz & Matyja, 2019). The soil Fluorescein diacetate (FDA) hydrolysis activity was analyzed by measuring the concentration of fluorescein ($\mu\text{g g}^{-1}\text{ h}^{-1}$) released upon hydrolysis of FDA (a colorless substrate) by several classes of soil enzymes such as lipases, esterases, and proteases (J. Schn rer & T. Rosswall, 1982).

All soil samples were analyzed in triplicate and results were shown as Mean \pm SE.

2.3 Analysis of total soil As using ICP-MS

The total As, Manganese (Mn), Zinc (Zn), Selenium (Se), and Iron (Fe) content of the soil samples was analyzed using ICP-MS (iCAP TQ Thermo Fisher

Scientific). The dried soil samples were sieved (2 mm) and 500 mg of each sample was digested using 69% nitric acid and 30% hydrogen peroxide in a microwave digester (Mars 6 CEM). The digestion was carried out according to the EPA method 3051A (Element, 2007). The digested samples were diluted and filtered through 0.2 µm syringe filters and then analyzed by ICP-MS. The analysis of As (75) and Se (80) was done in Triple Quadrupole (TQ) mode using oxygen as the reacting gas and analyzed as $75\text{As}.16\text{O}$ and $80\text{Se}.16\text{O}$ respectively. The elements Mn, Zn, and Fe were analyzed in Single Quadrupole-Kinetic Energy Discrimination (SQ-KED) mode using Helium (He) gas in the collision cell. The mass calibration of the instrument was performed using the iCAP Q/Qnova calibration solution (Thermo Scientific). The intermediate performance of the instrument was checked using iCAP TQ Tune Solution BRE0009578 (Thermo Scientific) and BMSCIENTIFIC-442 (Inorganic Ventures) tune solution. The calibration and quality assurance of the samples were ensured by repeated analysis ($n=5$) of CRMs with the recovery of the elements in the range of 98–105%. The detection limit for As was $1 \mu\text{g L}^{-1}$.

2.4 Metagenomic Sequencing

2.4.1 DNA Extraction and Library Construction

Total DNA from the As-contaminated soil samples was isolated using a Nucleospin soil kit (MACH-EREY-NAGEL, Germany) as per the manufacturer's instruction. The purity, concentration of DNA, and library construction were done as reported previously by Kaur et al. (2021). Briefly, the fragmented DNA ~350 bp (100 ng) was used for the library preparation using NEB Next® Ultra™ II DNA Library Prep Kit (New England Biolabs, UK). During this end repair, A-tailing and adapter ligation were done sequentially. Following this, the selected fragments were enriched and purified and the quantitation of the library was done by Qubit DNA.

2.4.2 Sequencing Quality Control and Assembly of Reads

Paired End Illumina libraries were loaded onto the Illumina platform for cluster generation and sequencing. For quality control, low-quality reads (≤ 40 on

the phred scale) and adapter-contaminated reads were removed as a part of pre-processing. The QC-passed samples were assembled using optimized SOAPdenovo protocol or MEGAHIT for Soil Water (Brum et al., 1979; Scher et al., 2013). The assembly was done using different K values to get the largest N50. This was then taken for analysis and the Scaffigs were assembled using Soap 2.21 (Mende et al., 2012; Nielsen et al., 2014).

2.4.3 Taxonomic and Functional Annotation

Taxonomy classification was done using the Kaiju aligner against the latest database of NCBI-(nr+euk), with parameters (SEG low complexity filter: Yes; Run mode=greedy; minimum match length=11; Minimum match score=75; allowed mismatch=5) (Menzel et al., 2016). The output file generated from Kaiju for each respective sample contains the number of assigned reads per taxon. The functional classification was performed through the SEED subsystem database within MG-RAST. The analysis was done with default settings after the normalization of raw counts. Furthermore, the analysis was carried out by referring to different databases viz., KEGG, eggNOG, and CAZy. The relative abundance of the As-responsive genes in different soil samples was obtained by annotating the metagenome assembly (with bit scores ≥ 40) with the AsgeneDB, a manually curated database for As metabolizing genes (<https://github.com/XinweiSong/Asgene>) (Song et al., 2022). The gene prediction was carried out by the MetaGeneMark based on the scaffigs assembled by single and mixed samples. The predicted genes were then pooled together for dereplication to construct a gene catalog (Fu et al., 2012; Li & Godzik, 2006). Based on the Clean Data of each sample from the gene catalog, the abundance of information from the gene catalog for each sample was obtained.

3 Results and Discussion

3.1 Soil Arsenic and Micronutrient Quantification

The total As concentration of the soil samples SKNNNGH was 4.88 mg kg^{-1} and was categorized as low in this study, similarly, samples SKNNNSIS and SKNNNSK where As concentration was 15.89

and 24.84 mg kg⁻¹ were categorized as mid-level and finally soil sample SKNNNMJ with the As concentration 43.67 mg kg⁻¹, was categorized as high. This categorization was employed to explain the findings of this study. Numerous studies have reported the formation of As gradient naturally in contaminated areas, these gradients may form naturally over a long period due to the topography of the area (Liu et al., 2023; Morosini et al., 2023; Valverde et al., 2011; Yang et al., 2024). In most of the cases, the impact of As on the soil microbial communities is studied in isolation, with each study area having its unique edaphic and climatic conditions. Even if these are similar, it is less likely that the vegetation on the sites in question is also the same. Since these factors are important in shaping and influencing the existing microbial communities their influence on the structure and functions of the microbial communities in the As contaminated areas cannot be ruled out. In such conditions, the studies involving As gradient at the contaminated sites are important to comprehend the effect of different As concentrations in shaping the microbial communities. Thus, in our study, we selected the sampling sites based on the similarity in the soil physicochemical and biological properties, the similarity in the climatic conditions, and the vegetation cover. Considering the direct impact of As in agricultural soil on human health through its entry into the food chain, such gradient studies on agricultural soils become of prime importance. One of the studies carried out at a gradient of arsenic at Terrubias mine reported a shift of microbial community towards Firmicutes as the dominant phyla with an increase in soil arsenic concentration (Valverde et al., 2011). Another study carried out at the gradient of As contamination reported the abundance and increase in expression of As resistant genes in paddy fields with increasing As concentration (Zhang et al., 2021).

Apart from As other micronutrients were also analysed as they may also influence microbial diversity. The total Mn concentration in the soil samples SKNNNGH, SKNNNSIS, SKNNNSK, and SKNNNMJ were 33.7, 52.20, 56.17, and 60.51 mg kg⁻¹, Zn concentrations were 12.31, 13.14, 11.21, and 14.90 mg kg⁻¹, Se concentration were 4.31, 5.05, 3.91, and 4.24 mg kg⁻¹, and Fe concentrations were 90.06, 74.81, 89.06, and 60.5 mg kg⁻¹, respectively. The concentrations of these micronutrients in different sampling sites did not show significant

differences and hence the effect on the microbial community structure and function can be attributed to changing arsenic concentrations.

3.2 Soil Physico-Chemical Properties

The soil samples were sandy clay in texture with bulk density in the range of 1.31–1.36 g cc⁻¹ and water holding capacity in the range of 48.9–51.2%. The soil samples were near neutral to slightly alkaline with a pH ranging from 7.41 to 7.61 and electrical conductivity in the range of 225–231 $\mu\text{S cm}^{-1}$ when estimated at 28°C. Soil total organic carbon content in the samples SKNNNGH, SKNNNSIS, SKNNNSK, and SKNNNMJ ranged between 0.23–0.27% with their respective MBC values 527.2, 551.25, 549.45, and 544.1 $\mu\text{g g}^{-1}$, respectively. The soil available N, P, and K ranged between 0.95–1.04%, 72.21–75.16 mg kg⁻¹, and 78.92–82.15 mg kg⁻¹, respectively. The Dehydrogenase activity in all four soil samples SKNNNGH, SKNNNSIS, SKNNNSK, and SKNNNMJ was found to be 4.14, 3.81, 3.94, and 3.87 $\mu\text{g TPF g soil}^{-1} \text{ h}^{-1}$, respectively. Likewise, the soil FDA activity was in the range of 299.8 to 311.8 $\mu\text{g Fluorescein g soil}^{-1} \text{ h}^{-1}$. Soil physicochemical properties of the four samples were not significantly different except for the soil As contents (Table S1). A correlation analysis was performed between total soil arsenic content and other soil physicochemical properties, as well as with soil enzyme activities (data not shown). The total As content of soil samples was positively correlated to the soil pH. This can be attributed to the fact that the increase in pH may decrease the adsorption of As⁺⁵ (Huang et al., 2006). On the contrary, a significant ($p < 0.05$) negative correlation was found between total soil arsenic content and the soil electrical conductivity. The EC of the soil is an indicator of dissolved ions, and with an increase in the ionic strength As absorption may decrease due to an increase in the net negative charges on the plane of sorption which may facilitate leaching of As (Kim et al., 2021; Smith et al., 1999). Similarly, a negative correlation was found between the total soil As and microbial biomass carbon (MBC). It has been earlier shown that chronic As exposure can adversely affect the MBC in contaminated soils (Ghosh et al., 2004). The As concentration was correlated with the changes in soil properties showing some direct or indirect

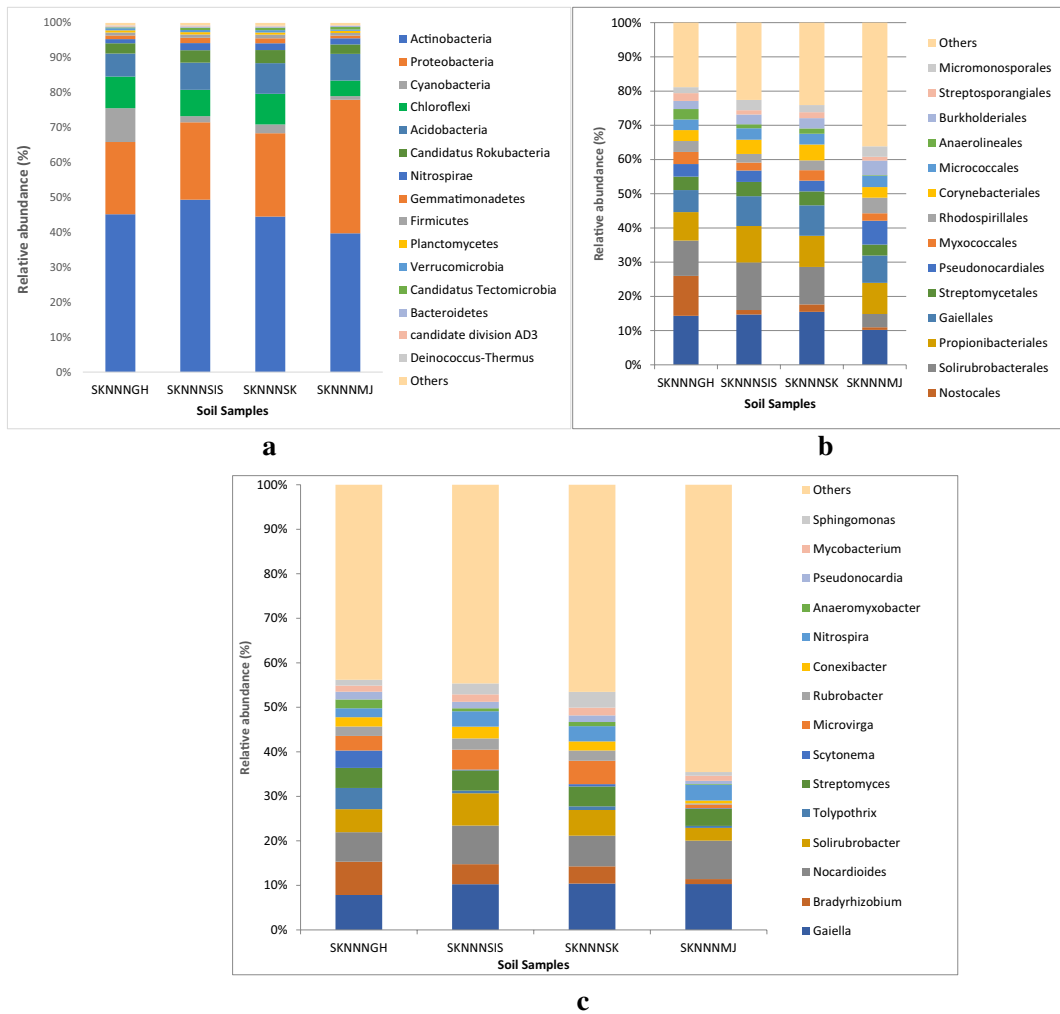


Fig. 1 Bacterial community composition at a gradient of arsenic contamination (a) Relative abundance of bacteria at the phylum level, (b) Relative abundance of bacteria at the order

level, (c) Relative abundance of bacteria at genus level in soil samples contaminated with different concentrations of arsenic

influence on the soil Arsenic concentration and its bioavailability to crops.

3.3 Taxonomic Composition of Bacterial Community

The analysis of microbial community composition in soil samples with different As concentrations was done to unveil the microbes with suitable functional traits to cope with As toxicity, which can be employed as a potential bioremediation agent in contaminated agricultural fields. In the present study, the dominant bacterial phylum observed in all the samples in the

order of their abundance were Actinobacteria, Proteobacteria, Cyanobacteria, Chloroflexi, and Acidobacteria (Fig. 1a). Upon PCoA analysis (data not shown), it was found that the bacterial phylum compositions of two soil samples with the mid-As concentration were more similar to each other than to those with either high or low As concentration. This indicates that As has some role in shaping the bacterial communities in the contaminated soil. Among the bacterial orders, the most abundant ones were, Rhizobiales, Solirubrobacterales, Propionibacteriales, Gaiellales, Streptomycetales, and Pseudonocardiales (Fig. 1b). To get a better picture of the bacterial community composition of

the soil samples contaminated with different levels of As relative abundance at the genus level was studied (Fig. 1c). This may eventually provide an overview of the bacterial genera that can be explored for their As remediating potential. The genera *Bradyrhizobium*, *Tolypothrix*, *Scytonema*, and *Anaeromyxobacter* were found to be sensitive with a decrease in their relative abundance to the increase in soil As concentration. The two genera *Scytonema* and *Anaeromyxobacter* were absent in samples with high As concentrations indicating their sensitivity. Most of the genera such as *Gaiella*, *Nocardioides*, *Solirubrobacter*, *Microvirga*, and *Nitrospira* were observed in the mid-As concentrations. The diversity of bacterial genera reported in the present study for As-contaminated soil samples was similar to those earlier reported from the contaminated soil of lead (Pb) and zinc (Zn) mines. The common genera found were *Solirubrobacter* and *Sphingomonas*, and the similarity was more pronounced at the phylum level (Hemmat-Jou et al., 2018). Cui et al. (2018) conducted a study to evaluate the effect of high concentrations of heavy metals and metalloids on bacterial diversity using 16S rRNA gene sequencing and found *Arthrobacter*, *Nocardioides*, *Aeromicrobium*, *Solirubrobacter*, *Blastococcus*, *Microvirga*, *Gaiella*, and *Candidatus*, tolerant to heavy metals as well as alkalinity stress. This coincided to an extent with the bacterial genera observed in the present study. The presence of bacterial genera *Gaiella* in the soil contaminated with heavy metals has previously been reported (Duan et al., 2021; Hu et al., 2021). *Gaiella* was also found to be positively correlated to Chromium (Cr) concentration in the soil irrigated with the treated wastewater that exhibited Cr bioremediation (Xi et al., 2021). Thus, *Gaiella* might have the mechanism to tolerate high concentrations of heavy metals, but its specific response to As in culture condition and the detoxification mechanism involved therein is yet to be explored. *Nocardioides* showed a similar trend, with its higher relative abundance in soil samples with mid- and high-As concentrations. *Nocardioides* sp. L-37a has been reported to express arsenate reductase (*arsC*) with increased expression by four folds on exposing the strain to As^{+5} (Bagade et al., 2016). The genus *Solirubrobacter* of the phylum Actinobacteria was observed to have a higher relative abundance in soil samples with the mid-As concentration. The genus has previously been reported in the rhizospheric soil of the Cd/Zn hyperaccumulating

Sedum plant and was also found to be positively correlated with the bioaccumulation factor of Cd and Zn in the *Sedum* plant (Wu et al., 2022). It was also reported from the rhizosphere of *Echinocactus platyacanthus* growing in soil with higher zinc concentrations up to 800 mg kg^{-1} and was found to be higher in relative abundance, indicating its tolerance to Zn and probably other heavy metals (Sarria Carabalí et al., 2019). *Microvirga*, an α -proteobacteria was observed to be higher in its relative abundance in the soil samples with the mid-As concentration. One of its strains *Microvirga indica* S-MI1b sp. nov. was known for its As^{+3} oxidizing potential, with an ability to oxidize $15 \text{ mM } As^{+3}$ in 39 h. The oxidation of As^{+3} might be due to the presence of the *aioA* gene and its product arsenite oxidase, which is induced by the build-up of As concentration in soil. Apart from this, the strain was also capable of resisting other heavy metals such as Pb (1000 mg l^{-1}), Hg (10 mg l^{-1}), Sb^{+3} (100 mg l^{-1}), Cd (20 mg l^{-1}), Cr^{+6} (300 mg l^{-1}), Ni (400 mg l^{-1}), and Cu (30 mg l^{-1}) (Tapase & Kodam, 2018). In general, it is seen that various bacterial genera adopt different mechanisms to combat As toxicity and some of these bacteria can be explored and used for bioremediation of As-contaminated agricultural soil. Suitable arsenate-reducing bacteria can also be used along with As hyperaccumulating plants to increase As bioavailability to the As-accumulating plants for successful implementation of phytoremediation in As-contaminated soils.

3.4 Taxonomic Composition of Fungal Community

Fungi can be exploited for As bioremediation as they are metabolically robust, having good surface area and biomass which may aid in efficient bioremediation. Fungal cell surface has several metal-binding functional groups and their hyphae extend deep into the soil, penetrating complex soil structures. They are predominantly present at sites contaminated with heavy metals including arsenic (Singh et al., 2015). The fungal counts observed at the different levels of soil As concentration in contaminated soil samples, their predicted gene statistics, and the number of As gene hits for each sample are summarised in Table S2. In the present study, maximum fungal counts obtained after Kaiju analysis were 532 and 460 in the soil samples SKNNNSIS and SKNNNSK, respectively in the mid-As

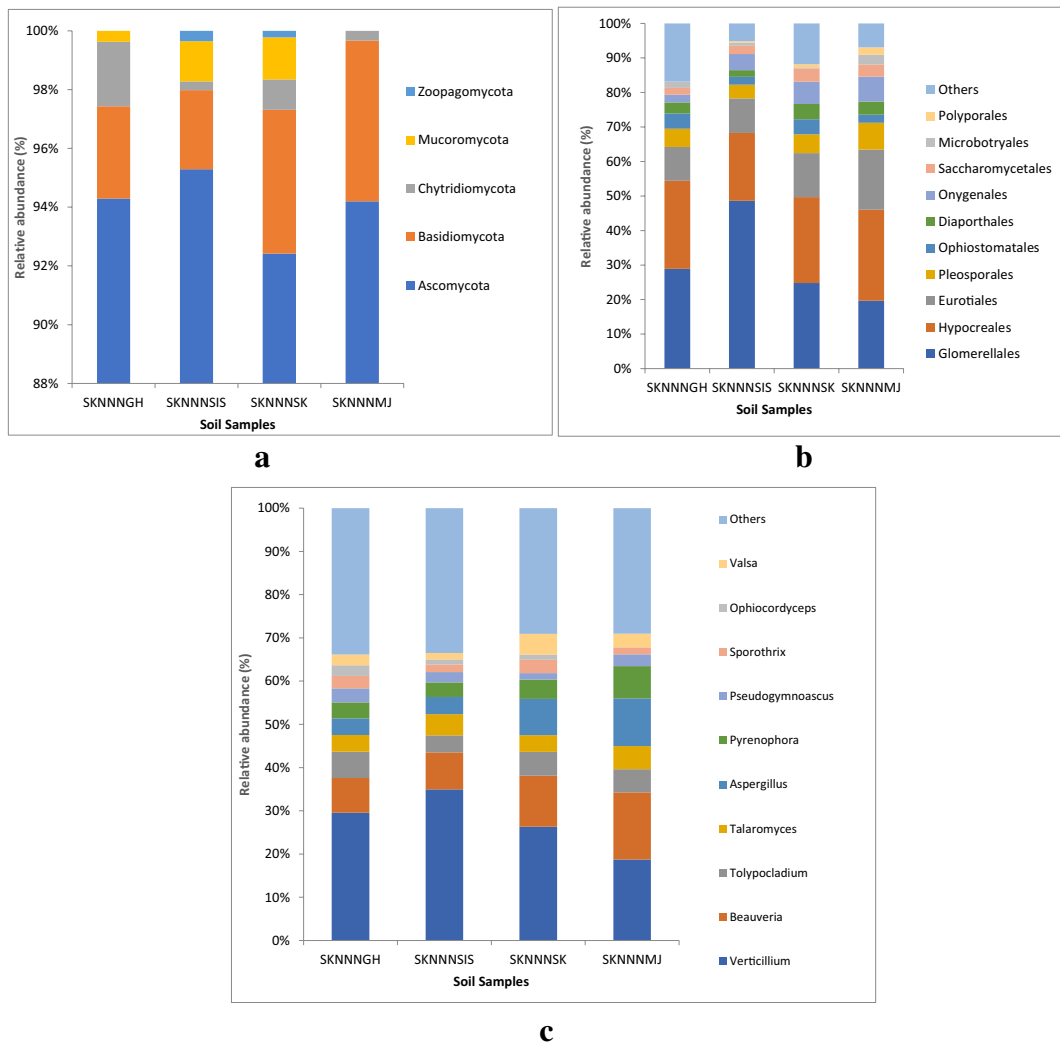


Fig. 2 Fungal community composition at a gradient of arsenic contamination. (a) Relative abundance of fungi at the phylum level, (b) Relative abundance of fungi at the order level (c) Relative abundance of fungi at the genus level in soil samples contaminated with different concentrations of arsenic

concentration. The least number of fungal counts (388) were reported in the sample SKNNNMJ with the high-As concentration, and moderate fungal counts (426) were observed for sample SKNNNGH with the low-As concentration. On the contrary, the maximum number of hits for As metabolizing genes as per the AsgeneDB were observed in the soil sample SKNNNMJ with the highest As concentration. From this, it can be concluded that the high concentration of As in the soil triggers microbial communities to express genes having As remediating potential to survive under the As stress.

Relative abundance of fungi at genus level in soil samples contaminated with different concentrations of arsenic

The phylum Ascomycota and Basidiomycota dominated the fungal community in the studied samples. Basidiomycota displayed a higher relative abundance in samples with higher As concentrations (Fig. 2a). The relative abundance of Ascomycota was found to be on the higher side in the sample with the mid-As concentration (15.89 mg kg^{-1}) as compared to samples having low- and high-As concentrations (Fig. 2a). The dominant orders observed in the fungal communities of As-contaminated soil samples included Glomerellales, Hypocreales, Eurotiales, and Pleosporales (Fig. 2b). In this study, it was observed

that the fungal genera, *Verticillium*, *Beauveria*, *Tolypocladium*, *Talaromyces*, *Aspergillus*, and *Pyrenophora* were found to be the most abundant. The fungal genera *Beauveria*, *Talaromyces*, *Aspergillus*, *Pyrenophora*, and *Valsa* were found to be higher in relative abundance with the As concentration in different soil samples. The genus *Ophiocordyceps* was found to be lower gradually with the build-up of the As concentration, while the genera *Beauveria*, *Talaromyces*, and *Aspergillus* showed higher abundance in a similar trend of increase in soil As concentrations (Fig. 2c). The genus *Beauveria* was reported to produce metallothioneins (metal-binding proteins) under metal (Cu^{+2} and Cd^{+2}) stress (Kameo et al., 2000). In a study, *Beauveria bassiana* was reported for bioaccumulation and removal of heavy metals viz., Ni^{+2} , Cu^{+2} , Zn^{+2} , Cd^{+2} , Cr^{+6} , and Pb^{+2} with a removal percentage of 75, 74.1, 67.8, 63.4, 61.1, and 58.4%, respectively (Gola et al., 2018). Furthermore, *Talaromyces* sp. was earlier reported to be As hyper-tolerant and to biosorb As from aqueous medium (Nam et al., 2019). In a study, *Talaromyces helices* was reported for the removal of Cu, as well as exhibiting co-tolerance to other heavy metals like Co, Cd, and Pb. This type of co-tolerance development and enhanced removal efficiencies are crucial in remediating sites contaminated with multiple heavy metals (Romero et al., 2006). *Aspergillus* has been earlier reported by many researchers as the promising strain for metal(loid) bioremediation. *Aspergillus niger* was studied for As, Cu, and Cr removal efficiency from the waste wood treated with chromated copper arsenate (CCA), used as a wood preservative to increase the durability of the wood. It was reported to remove 97% of arsenic, 49% of copper, and 55% of chromium from the CCA-treated wood chips in 10d (Kartal et al., 2004). In another study, *Aspergillus nidulans* isolated from arsenic-contaminated soil was found resistant to 500 mg kg^{-1} of As and showed 84.3% of arsenic adsorption from the contaminated soil after 11 d of cultivation, indicating its potential role in arsenic remediation of the contaminated soil (Maheswari & Murugesan, 2009). In the present study, it was seen that the relative abundance of *Verticillium* was higher in the soil sample with 15.89 mg kg^{-1} of total As which subsequently lowered as the As concentration in the soil samples built up further. Although the genus has been previously reported for the removal of Pb, Zn, and Cd, the findings here suggest that the

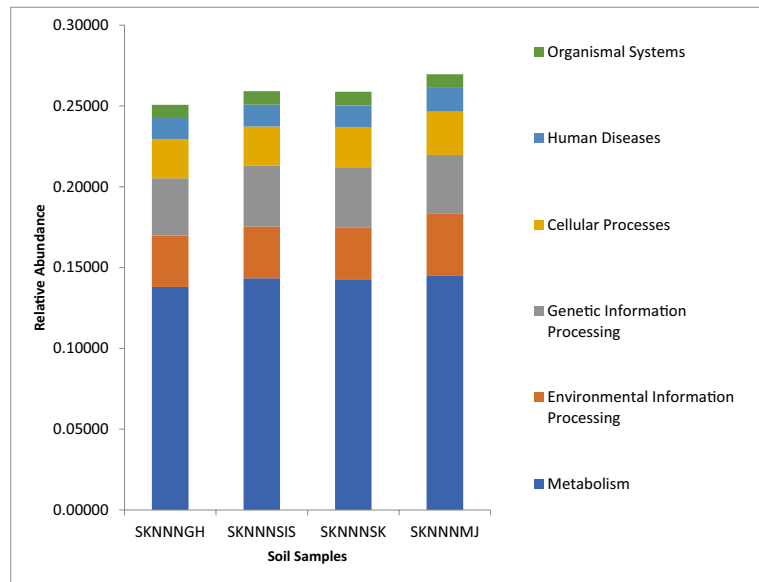
genus is sensitive to higher concentrations of As in soil. The genera *Gaiella*, *Solirubrobacter*, *Beauveria*, and *Verticillium* have been first time reported in As-contaminated soil samples, and understanding of their functions in As-bioremediation is subjected to further experimentation. This study opens up an avenue to use these fungal strains in the field condition for reducing the bioavailability of As to the plants and decreasing its entry into the food chain (Henrique et al., 2019).

3.5 Functional Flux Against Metal Contamination

In the current metagenomic study, the overall community was assessed for its physiology and the collective functions performed by the microbes constituting the community as a function of the As concentration gradient in different agricultural soil samples analyzed. Insights into the functional diversity of a community can be obtained by annotating the metagenomic sequences with the functions (Cantarel et al., 2009; Feng et al., 2015; Li et al., 2014; Qin et al., 2012). According to the functional abundance obtained for each sample, various analyses were performed. On average, the samples SKNNNSIS and SKNNNSK with the mid-level of As concentrations had the maximum number of predicted genes, 1,068,529 and 895,459, respectively. The lowest number of total predicted genes (591,738) was observed in the soil sample with the high-As concentration Table S2.

When considering the database Kyoto Encyclopedia of Genes and Genomes (KEGG), overall the samples displayed the maximum number of genes involved in the metabolism (Fig. 3). This was followed by environmental information processing genes in which the genes for membrane transport were more in number. The genes for cellular processes, genetic information on processing, and organismal systems were reported in all the samples to a relatively lesser extent. Processes like metabolism, environmental information processing, cellular processes, human diseases, and organismal systems were observed to increase in the soil samples with high-As concentration as compared to the samples with low-As concentration. The relative abundance of genes through the KEGG database was analyzed and a graph of the functional abundance for each sample was drawn (Fig. 3).

Fig. 3 Total Functional abundance based on KEGG



3.5.1 Abundance of Arsenic Metabolizing Genes

Microorganisms inhabiting different environments with As contamination are known to possess several As metabolizing genes. The mechanism for As detoxification includes the oxidation of more toxic and more mobile As^{+3} species to less toxic and less mobile As^{+5} species catalyzed by the *aioA* gene. As^{+5} may also be used for energy production by microbes, wherein, As^{+5} reduction is coupled to ATP production catalyzed by respiratory As^{+5} reductase (Arr) (Slyemi & Bonnefoy, 2012). In another detoxification mechanism, As^{+5} is reduced to As^{+3} by a reductase (ArsC) and effluxed out of the cell by ArsB and Acr3 permease (Rosen n.d.). As^{+3} can be methylated by the ArsM, a microbial As(III) S-adenosylmethionine (SAM) methyltransferase, that can convert the inorganic As species to their volatile derivatives or into less toxic As^{+5} organic derivatives. (Verma et al., 2016).

Considering the importance of As metabolizing genes and their role in As detoxification, their diversity in the soil and their relative abundance were studied under the gradient of soil As concentration (Fig. 4). Several genes involved in As transport pathways like *ACR3*, *arsA*, *arsB*, *arsD*, *arsJ*, *arsP*, *GET3*, *glpF*, *pgpA*, *PiT*, *pstA*, *pstB*, *pstC*, and *pstS* were found in all the soil samples contaminated with As (Fig. 4a). The *pstB* gene coding for a high-affinity phosphate uptake system, involved in the uptake

of As^{+5} was the most abundant among the As transport genes. Likewise, several genes responsible for As^{+3} oxidation, *aioA*, *aioB*, *aioR*, *aioS*, *aioX*, *aoxC*, *arsH*, *arxA*, *arxB*, *arxR*, *arxS*, and *arxX* were found to occur in all the soil sample contaminated with As, with *arsH* being the most abundant in all the samples (Fig. 4b). Few genes of the As^{+5} respiratory pathway such as *GstB*, *mgsR*, *arrA*, and *arrB* were found in the soil samples studied. The genes involved in the As^{+5} reduction pathways like *moeA*, *arsC*, and *arsR* were also observed in the As-contaminated samples among which the *arsR* gene was the most abundant (Fig. 4c).

The As metabolizing/ transport genes showing variation across the gradient of As concentration in soil based on the ASgeneDB is given in Fig. 4d. It was found that the As transporter genes *arsJ* and *pgpA* were higher in the soil sample with higher As concentration. The gene *arsJ* was reported to provide As^{+5} resistance in *Pseudomonas aeruginosa* by efflux of As^{+5} as a phosphoglycerate derivative (Chen et al., 2016). Therefore, the higher relative abundance of *arsJ* with the increasing As concentration could detoxify As^{+5} . Similarly, the higher relative abundance of the *pgpA* gene may be due to its role in As detoxification by transporting As^{+3} as a glutathione conjugate ($As(GS)_3$) into the cellular vacuoles (Légaré et al., 2001). The gene *arsM* responsible for As methylation was also found to be higher with the increase in the As concentration along the gradient indicating that it provides As resistance to the

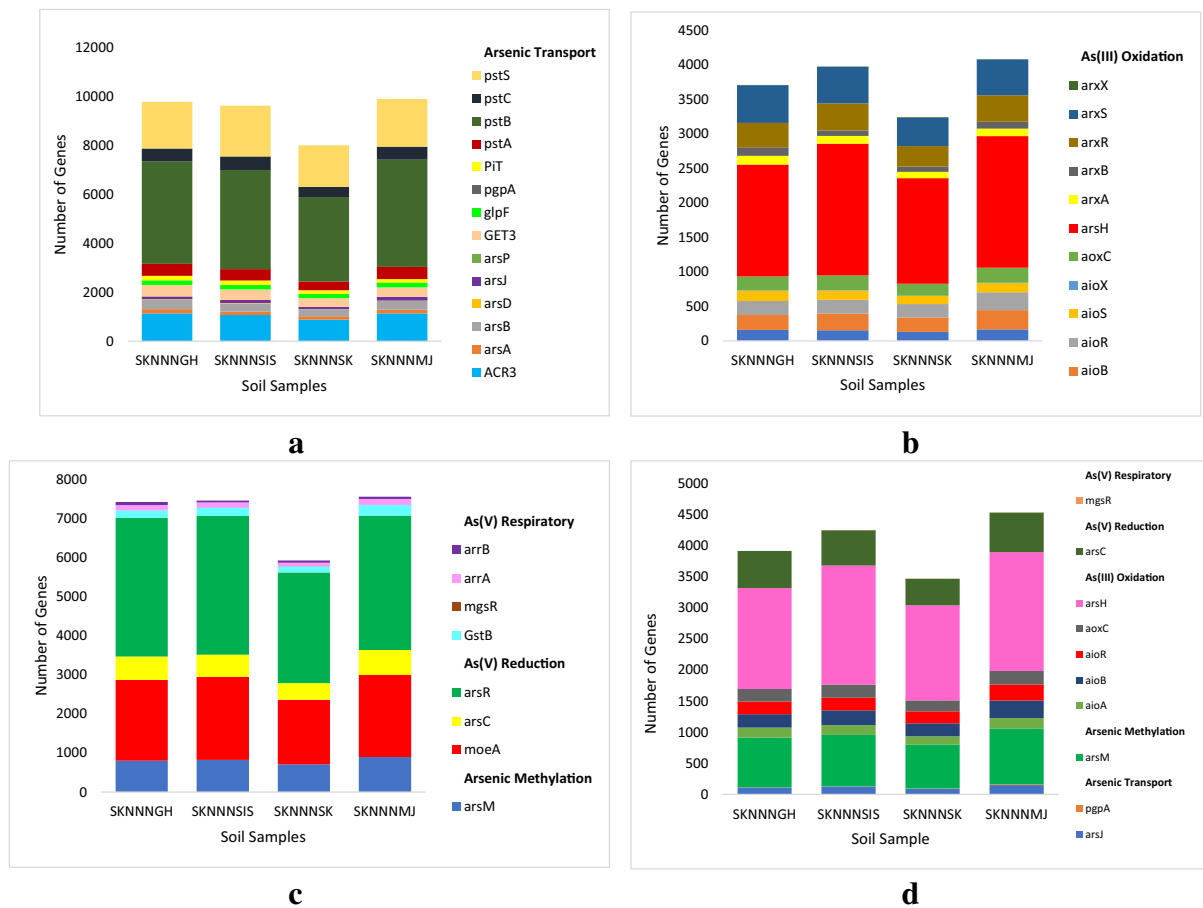


Fig. 4 Number of arsenic metabolizing/transport genes w.r.t. the AsgeneDB in soil samples contaminated with a gradient of As concentration {(a) Arsenic Transport, (b) As (III) Oxida-

tion, (c) As (V) respiratory, As (V) reduction, and As methylation} (d) major genes given in (a), (b), and (c) showed significant variation across the As concentration gradient

microbes by converting inorganic As species into its volatile derivatives (Qin et al., 2006). A few genes of the As^{+3} oxidation pathway such as *aioA*, *aioB*, *aioR*, and *arsH* were also found to be higher in the samples with high As concentration probably to provide As resistance. The *aioA* and *aioB* genes code for the large and small subunits respectively of the Arsenite oxidase catalyzing the oxidation of As^{+3} to relatively less toxic As^{+5} (Li et al., 2021). The product of *aioR* gene is a regulator that interacts with the RNA polymerase to initiate the expression of *aio* genes (Cai et al., 2013). The relative abundance of *arsH* gene, an organoarsenical oxidase, is correlated with the higher abundance of *arsM* gene mainly because it functions to oxidize the toxic As^{+3} organoarsenicals to As^{+5} organoarsenicals and hence may provide resistance to

the microbes harboring these two genes (Chen et al., 2015). The *arsC* gene involved in the As^{+5} reduction pathway is also observed to be higher in relative abundance at high As concentration. This gene is involved in the conversion of As^{+5} taken up by the cell to As^{+3} which is then acted upon by the other As detoxification pathways such as As methylation pathway (Cai et al., 2013). Hence its concomitant higher level with other As detoxification genes is justified. Like *arsC*, *gstB* gene coding for glutathione S-transferase B is also involved in the reduction of As^{+5} to As^{+3} , but unlike *arsC*, the reduction mediated by *gstB* is not dependent on glutaredoxins. The As^{+5} entering the cell is reduced using the reduced form of glutathione as an electron donor (Chrysostomou et al., 2015). *mgsR* is the gene expressed during oxidative

stress conditions which in turn induces the expression of MgsR subregulon consisting of more than 50 genes (Reder et al., 2012). The higher relative abundance of *mgsR* with the build-up in As concentration might be to combat the oxidative stress induced by As.

4 Conclusion

This work presented a comprehensive report on the changes in microbial community dynamics as a function of soil arsenic concentration. We found Actinobacteria and Ascomycota respectively dominating the bacterial and fungal communities across the gradient. We also reported the occurrence of *Gaiella*, *Solirubrobacter*, *Beauveria*, and *Verticillium* in the arsenic-contaminated soil which can be further explored for their potential role in arsenic bioremediation. The study also focused on the changes in the abundance of microbial genes responsible for providing tolerance against As to the resident microbial community exposed to increasing As stress, although further studies at the transcriptional level (RNA) are required to enhance our understanding of the response mechanism of soil microbial community to the As contamination. In this study we found *arsJ*, *arsM*, *aior*, *arsH*, and *arsC* genes to be the ones most influenced by the changing As concentration across the gradient. The study concludes that As contamination of the agricultural soil has a major influence in reducing the taxonomic diversity as well as the composition of the microbial species in the soil microbial community. The same is true for the normal functional pathways operating in the soil, which are altered due to As stress probably to enhance the microbial capability to combat As stress. The microbes that are abundantly found in the As-contaminated soil can further be explored for their As bioremediation potential and their field applicability.

Authors' contributions Mariya Naseem: Conceptualization, Data Curation, Writing—Original Draft, Visualization, Investigation, Validation Vivek Gaur: Data Curation, Writing—Review & Editing Manisha Singh: Data Curation Srishti Seth: Data Curation Richa Raghuvanshi: Supervision, Writing—Review & Editing Praveen C. Verma: Conceptualization, Methodology, Supervision, Writing—Review & Editing Pankaj Kumar Srivastava: Conceptualization, Methodology, Supervision, Writing—Review & Editing

Funding The authors are thankful to the Director, CSIR – NBRI for institutional and financial support to carry out the study (OLP 116). Mariya Naseem is thankful to the Council of Scientific & Industrial Research (CSIR) India for Senior Research Fellowship {31/008 (0391)/2019-EMR-I}

Data Availability Data available on request from the authors.

Declarations

Ethics approval statement Not applicable.

Conflict of interest disclosure The authors have no conflict of interest to declare.

Partial consent statement Not applicable.

Clinical trial registration Not applicable.

Permission to reproduce material from other sources Not applicable.

References

- Adeyemi, A. O. (2009). Bioaccumulation of arsenic by fungi. *American Journal of Environmental Sciences*, 5, 364–370.
- Bagade, A. V., Bachate, S. P., Dholakia, B. B., Giri, A. P., & Kodam, K. M. (2016). Characterization of Roseomonas and Nocardioides spp. for arsenic transformation. *J Hazard Mater*, 318, 742–750. <https://doi.org/10.1016/j.jhazmat.2016.07.062>
- Bremner, J. M. (1960). Determination of nitrogen in soil by the Kjeldahl method. *Journal of Agricultural Science*, 55, 11–33.
- Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doucier, G., Acinas, S. G., Alberti, A., Chaffron, S., Cruaud, C., De Vargas, C., & Gasol, J. M. (1979). Patterns and ecological drivers of ocean viral communities. *Science*, 348(2015), 1261498.
- B.P. Rosen, Biochemistry of arsenic detoxification, n.d.
- Cai, L., Yu, K., Yang, Y., Chen, B. W., Li, X. D., & Zhang, T. (2013). Metagenomic exploration reveals high levels of microbial arsenic metabolism genes in activated sludge and coastal sediments. *Applied Microbiology and Biotechnology*, 97, 9579–9588. <https://doi.org/10.1007/s00253-012-4678-8>
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research*, 37, D233–D238.
- Castro-Severyn, J., Pardo-Esté, C., Mendez, K. N., Fortt, J., Marquez, S., Molina, F., Castro-Nallar, E., Remonsellez, F., & Saavedra, C. P. (2021). Living to the high extreme: Unraveling the composition, structure, and functional insights of bacterial communities thriving in

- the arsenic-rich Salar de Huasco altiplanic ecosystem. *Microbiol Spectr*, 9, e00444–e521.
- Čerňanský, S., Kolenčík, M., Ševc, J., Urík, M., & Hiller, E. (2009). Fungal volatilization of trivalent and pentavalent arsenic under laboratory conditions. *Bioresource Technology*, 100, 1037–1040. <https://doi.org/10.1016/j.biortech.2008.07.030>
- Chandraprabha, M. N., & Natarajan, K. A. (2011). Mechanism of arsenic tolerance and bioremoval of arsenic by *Acidithiobacillus ferrooxidans*. *Journal of Biochemical Technology*, 3, 257–265.
- Chen, J., Bhattacharjee, H., & Rosen, B. P. (2015). ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone. *Molecular Microbiology*, 96, 1042–1052. <https://doi.org/10.1111/mmi.12988>
- Chen, J., Yoshinaga, M., Garbinski, L. D., & Rosen, B. P. (2016). Synergistic interaction of glyceraldehyde-3-phosphate dehydrogenase and ArsJ, a novel organoarsenical efflux permease, confers arsenate resistance. *Molecular Microbiology*, 100, 945–953. <https://doi.org/10.1111/mmi.13371>
- Chrysostomou, C., Quandt, E. M., Marshall, N. M., Stone, E., & Georgiou, G. (2015). An alternate pathway of arsenate resistance in *E. coli* mediated by the glutathione S-transferase GstB. *ACS Chem Biol*, 10, 875–882. <https://doi.org/10.1021/cb500755j>
- Cui, H., Le Liu, L., Dai, J. R., Yu, X. N., Guo, X., Yi, S. J., Zhou, D. Y., Guo, W. H., & Du, N. (2018). Bacterial community shaped by heavy metals and contributing to health risks in cornfields. *Ecotoxicology and Environmental Safety*, 166, 259–269. <https://doi.org/10.1016/j.ecoenv.2018.09.096>
- C. Morosini, E. Terzaghi, G. Raspa, M. Grotti, S. Armiraglio, S. Anelli, A. Di Guardo, Arsenic movement and fractionation in agricultural soils which received wastewater from an adjacent industrial site for 50 years, *Science of the Total Environment* 898 (2023). <https://doi.org/10.1016/j.scitotenv.2023.165422>.
- Domingo-Relloso, A., Makhani, K., Riffo-Campos, A. L., Tellez-Plaza, M., Klein, K. O., Subedi, P., Zhao, J., Moon, K. A., Bozack, K., Haack, K. (2022) Arsenic Exposure, Blood DNA Methylation, and Cardiovascular Disease. *Circ Res*. 10–1161.
- Duan, R., Lin, Y., Zhang, J., Huang, M., Du, Y., Yang, L., Bai, J., Xiang, G., Wang, Z., & Zhang, Y. (2021). Changes in diversity and composition of rhizosphere bacterial community during natural restoration stages in antimony mine. *PeerJ*, 9, 1–22. <https://doi.org/10.7717/peerj.12302>
- D.R. Mende, A.S. Waller, S. Sunagawa, A.I. Järvelin, M.M. Chan, M. Arumugam, J. Raes, P. Bork, Assessment of metagenomic assembly using simulated next generation sequencing data, *PLoS One* 7 (2012). <https://doi.org/10.1371/journal.pone.0031386>.
- Element, C. A. S. (2007). Method 3051A microwave assisted acid digestion of sediments, sludges, soils, and oils, Z. Für. *Analytical Chemistry*, 111, 362–366.
- E. Smith, R. Naidu, A.M. Alston, (1999) Chemistry of Arsenic in Soils: I. Sorption of Arsenate and Arsenite by Four Australian Soils, n.d.
- Feng, Q., Liang, S., Jia, H., Stadlmayr, A., Tang, L., Lan, Z., Zhang, D., Xia, H., Xu, X., & Jie, Z. (2015). Gut microbiome development along the colorectal adenoma–carcinoma sequence. *Nature Communications*, 6, 1–13.
- Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28, 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Genchi, G., Lauria, G., Catalano, A., Carocci, A., & Sinicropi, M. S. (2022b). Arsenic: A Review on a Great Health Issue Worldwide. *Applied Sciences*, 12, 6184. <https://doi.org/10.3390/app12126184>
- G. Genchi, G. Lauria, A. Catalano, A. Carocci, M.S. Sinicropi, Arsenic: A Review on a Great Health Issue Worldwide, *Applied Sciences (Switzerland)* 12 (2022a). <https://doi.org/10.3390/app12126184>.
- Ghosh, A. K., Bhattacharyya, P., & Pal, R. (2004). Effect of arsenic contamination on microbial biomass and its activities in arsenic contaminated soils of Gangetic West Bengal, India. *Environment International*, 30, 491–499. <https://doi.org/10.1016/j.envint.2003.10.002>
- Giri, A. K., Patel, R. K., Mahapatra, S. S., & Mishra, P. C. (2013). Biosorption of arsenic (III) from aqueous solution by living cells of *Bacillus cereus*. *Environmental Science and Pollution Research*, 20, 1281–1291.
- Gola, D., Malik, A., Namburath, M., & Ahammad, S. Z. (2018). Removal of industrial dyes and heavy metals by *Beauveria bassiana*: FTIR, SEM, TEM and AFM investigations with Pb(II). *Environmental Science and Pollution Research*, 25, 20486–20496. <https://doi.org/10.1007/s11356-017-0246-1>
- Hamadani, J. D., Tofail, F., Nermell, B., Gardner, R., Shiraji, S., Bottai, M., Arifeen, S. E., Huda, S. N., & Vahter, M. (2011). Critical windows of exposure for arsenic-associated impairment of cognitive function in pre-school girls and boys: A population-based cohort study. *International Journal of Epidemiology*, 40, 1593–1604.
- Hemmat-Jou, M. H., Safari-Sinegani, A. A., Mirzaie-Asl, A., & Tahmourespour, A. (2018). Analysis of microbial communities in heavy metals-contaminated soils using the metagenomic approach. *Ecotoxicology*, 27, 1281–1291. <https://doi.org/10.1007/s10646-018-1981-x>
- Henrique, L., Guimarães, S., Roberta, F., Tonani, L., Regina, M., Lisboa, J., Augusto, L., Ferreira, F., & Lemos, B. (2019). Arsenic volatilization by *Aspergillus sp.* and *Penicillium sp.* isolated from rice rhizosphere as a promising eco-safe tool for arsenic mitigation. *J Environ Manage*, 237, 170–179. <https://doi.org/10.1016/j.jenvman.2019.02.060>
- Hu, X., Wang, J., Lv, Y., Liu, X., Zhong, J., Cui, X., Zhang, M., Ma, D., Yan, X., & Zhu, X. (2021). Effects of Heavy Metals/Metalloids and Soil Properties on Microbial Communities in Farmland in the Vicinity of a Metals Smelter. *Frontiers in Microbiology*, 12, 1–13. <https://doi.org/10.3389/fmicb.2021.707786>
- H.B. Nielsen, M. Almeida, A.S. Juncker, S. Rasmussen, J. Li, S. Sunagawa, D.R. Plichta, L. Gautier, A.G. Pedersen, E. Le Chatelier, E. Pelletier, I. Bonde, T. Nielsen,

- C. Manichanh, M. Arumugam, J.M. Batto, M.B. Quintanilha Dos Santos, N. Blom, N. Borruel, K.S. Burgdorf, F. Boumezeur, F. Casellas, J. Doré, P. Dworzynski, F. Guarner, T. Hansen, F. Hildebrand, R.S. Kaas, S. Kennedy, K. Kristiansen, J.R. Kultima, P. Leonard, F. Levenez, O. Lund, B. Moumen, D. Le Paslier, N. Pons, O. Pedersen, E. Prifti, J. Qin, J. Raes, S. Sørensen, J. Tap, S. Tims, D.W. Ussery, T. Yamada, P. Renault, T. Sicheritz-Ponten, P. Bork, J. Wang, S. Brunak, S.D. Ehrlich, A. Jamet, A. Mérieux, A. Cultrone, A. Torrejon, B. Quinquis, C. Brechot, C. Delorme, C. M'rini, W.M. de Vos, E. Maguin, E. Varela, E. Guedon, F. Gwen, F. Haimet, F. Artiguenave, G. Vandemeulebrouck, G. Denariáz, G. Khaci, H. Blottière, J. Knol, J. Weissenbach, J.E.T. van Hylckama Vlieg, J. Torben, J. Parkhill, K. Turner, M. van de Guchte, M. Antolin, M. Rescigno, M. Kleerebezem, M. Derrien, N. Galleron, N. Sanchez, N. Grarup, P. Veiga, R. Oozeer, R. Dervyn, S. Layec, T. Bruls, Y. Winogradski, Z. Erwin, Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat Biotechnol* 32 (2014) 822–828. <https://doi.org/10.1038/nbt.2939>.
- Huang, R. Q., Gao, S. F., Wang, W. L., Staunton, S., & Wang, G. (2006). Soil arsenic availability and the transfer of soil arsenic to crops in suburban areas in Fujian Province, southeast China. *Science of the Total Environment*, 368, 531–541. <https://doi.org/10.1016/j.scitotenv.2006.03.013>
- I.H. Nam, K. Murugesan, J. Ryu, J.H. Kim, Arsenic (As) removal using *Talaromyces* sp. KM-31 isolated from As-contaminated mine soil, *Minerals* 9 (2019) 1–11. <https://doi.org/10.3390/min9100568>.
- Joshi, D. N., Flora, S. J. S., & Kalia, K. (2009). *Bacillus* sp. strain DJ-1, potent arsenic hypertolerant bacterium isolated from the industrial effluent of India. *J Hazard Mater*, 166, 1500–1505.
- J. Luo, Y. Bai, J. Liang, J. Qu, Metagenomic approach reveals variation of microbes with arsenic and antimony metabolism genes from highly contaminated soil, *PLoS One* 9 (2014). <https://doi.org/10.1371/journal.pone.0108185>.
- J. Qin, B.P. Rosen, Y. Zhang, G. Wang, S. Franke, C. Rensing, Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase, 2006. www.pnas.org/cgi/doi/10.1073/pnas.0506836103
- Kartal, S. N., Kakitani, T., & Imamura, Y. (2004). Bioremediation of CCA-C treated wood by *Aspergillus niger* fermentation. *Holz Als Roh - und Werkstoff*, 62, 64–68. <https://doi.org/10.1007/s00107-003-0445-y>
- Kaur, I., Gaur, V. K., Regar, R. K., Roy, A., Srivastava, P. K., Gaur, R., Manickam, N., & Barik, S. K. (2021). Plants exert beneficial influence on soil microbiome in a HCH contaminated soil revealing advantage of microbe-assisted plant-based HCH remediation of a dumpsite. *Chemosphere*, 280, 130690. <https://doi.org/10.1016/j.chemosphere.2021.130690>
- Khan, M. A., Stroud, J. L., Zhu, Y.-G., McGrath, S. P., & Zhao, F.-J. (2010). Arsenic bioavailability to rice is elevated in Bangladeshi paddy soils. *Environmental Science and Technology*, 44, 8515–8521.
- Kumar, M., Ramanathan, A. L., Rahman, M. M., & Naidu, R. (2016). Concentrations of inorganic arsenic in groundwater, agricultural soils and subsurface sediments from the middle Gangetic plain of Bihar, India. *Science of the Total Environment*, 573, 1103–1114.
- K. Zoroufchi Benis, A. Motalebi Damuchali, K.N. McPhedran, J. Soltan, Treatment of aqueous arsenic – A review of biosorbent preparation methods, *J Environ Manage* 273 (2020). <https://doi.org/10.1016/j.jenvman.2020.111126>.
- Légaré, D., Richard, D., Mukhopadhyay, R., Stierhof, Y. D., Rosen, B. P., Haimeur, A., Papadopoulou, B., & Ouellette, M. (2001). The Leishmania ATP-binding Cassette Protein PGPA is an Intracellular Metal-Thiol Transporter ATPase. *Journal of Biological Chemistry*, 276, 26301–26307. <https://doi.org/10.1074/jbc.M102351200>
- Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22, 1658–1659.
- Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J. R., Prifti, E., & Nielsen, T. (2014). An integrated catalog of reference genes in the human gut microbiome. *Nature Biotechnology*, 32, 834–841.
- Liu, Y., Dai, W., Yao, D., Wang, N., Liu, M., Wang, L., Tian, W., Yan, P., Huang, Z., & Wang, H. (2023). Arsenic pollution from human activities drives changes in soil microbial community characteristics. *Environmental Microbiology*, 25, 2592–2603.
- Maheswari, S., & Murugesan, A. G. (2009). Remediation of arsenic in soil by *Aspergillus nidulans* isolated from an arsenic-contaminated site. *Environmental Technology*, 30, 921–926. <https://doi.org/10.1080/09593330902971279>
- Małachowska-Jutcz, A., & Matyja, K. (2019). Discussion on methods of soil dehydrogenase determination. *International Journal of Environmental Science and Technology*, 16, 7777–7790.
- Mandal, B. K., & Suzuki, K. T. (2002). Arsenic round the world: A review. *Talanta*, 58, 201–235.
- MasudHossain, S., & Anantharaman, M. (2006). Studies on Bacterial Growth and Arsenic (III) Biosorption Using *Bacillus subtilis*. *Chemical and Biochemical Engineering Quarterly*, 20, 209–216.
- Mishra, R. K., Tiwari, S., Patel, A., & Prasad, S. M. (2021). Arsenic contamination, speciation, toxicity and defense strategies in plants. *Revista Brasileira De Botanica*, 44, 1–10. <https://doi.org/10.1007/s40415-020-00694-5>
- Mitra A., Chatterjee S., Gupta D.K., Potential role of microbes in bioremediation of arsenic, in: *Arsenic Contamination in the Environment: The Issues and Solutions*, Springer International Publishing, 2017: pp. 195–213. https://doi.org/10.1007/978-3-319-54356-7_10.
- Monteiro De Oliveira, E. C., Caixeta, E. S., Santos, V. S. V., & Pereira, B. B. (2021). Arsenic exposure from groundwater: Environmental contamination, human health effects, and sustainable solutions. *Journal of Toxicology and Environmental Health. Part b, Critical Reviews*, 24, 119–135. <https://doi.org/10.1080/10937404.2021.1898504>
- M.L. Jackson, Soil chemical analysis-advanced course., *Soil Chemical Analysis-Advanced Course*. (1969).
- M.S. Kim, S.H. Lee, J.G. Kim, Evaluation of factors affecting arsenic uptake by *Brassica juncea* in alkali soil after biochar application using partial least squares path modeling

- (PLS-PM), *Chemosphere* 275 (2021). <https://doi.org/10.1016/j.chemosphere.2021.130095>.
- M.C. Romero, E. H. Reinoso, M.I. Urrutia, A. Moreno Kiernan, Biosorption of heavy metals by *Talaromyces helicus*: a trained fungus for copper and biphenyl detoxification, *Electronic Journal of Biotechnology* 9 (2006) 0–0. <https://doi.org/10.2225/vol9-issue3-fulltext-11>.
- Ori, L. V., Amacher, M. C., & Sedberry, J. E., Jr. (1993). Survey of the total arsenic content in soils in Louisiana. *Communications in Soil Science and Plant Analysis*, 24, 2321–2332.
- Patel, K. S., Shrivastava, K., Brandt, R., Jakubowski, N., Corns, W., & Hoffmann, P. (2005). Arsenic contamination in water, soil, sediment and rice of central India. *Environmental Geochemistry and Health*, 27, 131–145. <https://doi.org/10.1007/s10653-005-0120-9>
- P. Menzel, K.L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, *Nat Commun* 7 (2016). <https://doi.org/10.1038/ncomms11257>.
- P. Tripathi, P. Khare, D. Barnawal, K. Shanker, P.K. Srivastava, R.D. Tripathi, A. Kalra, Bioremediation of arsenic by soil methylating fungi: Role of *Humicola* sp. strain 2WS1 in amelioration of arsenic phytotoxicity in *Bacopa monnieri* L, *Science of the Total Environment* 716 (2020). <https://doi.org/10.1016/j.scitotenv.2020.136758>.
- Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., & Shen, D. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*, 490, 55–60.
- Reder, A., Pöther, D. C., Gerth, U., & Hecker, M. (2012). The modulator of the general stress response, MgsR, of *Bacillus subtilis* is subject to multiple and complex control mechanisms. *Environmental Microbiology*, 14, 2838–2850. <https://doi.org/10.1111/j.1462-2920.2012.02829.x>
- Rudnai, T., Sándor, J., Kádár, M., Borsányi, M., Béres, J., Métneki, J., Marácz, G., & Rudnai, P. (2014). Arsenic in drinking water and congenital heart anomalies in Hungary. *International Journal of Hygiene and Environmental Health*, 217, 813–818.
- S. Kameo, H. Iwahashi, Y. Kojima, H. Satoh, s Induction of metallothioneins in the heavy metal resistant fungus *Beauveria bassiana* exposed to copper or cadmium, (2000).
- S.R. Olsen, Estimation of available phosphorus in soils by extraction with sodium bicarbonate, US Department of Agriculture, 1954.
- Sarria Carabalí, M. M., García-Oliva, F., Cortés Páez, L. E., & López Lozano, N. E. (2019). The response of candy barrel cactus to zinc contamination is modulated by its rhizospheric microbiota. *Rhizosphere*, 12, 100177. <https://doi.org/10.1016/j.rhisph.2019.100177>
- Scher, J. U., Sczesnak, A., Longman, R. S., Segata, N., Ubeda, C., Bielski, C., Rostron, T., Cerundolo, V., Pamer, E. G., & Abramson, S. B. (2013). Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *eLife*, 2, e01202.
- Shrivastava, A., Ghosh, D., Dash, A., & Bose, S. (2015). Arsenic Contamination in Soil and Sediment in India: Sources, Effects, and Remediation, *Curr Pollut Rep*, 1, 35–46. <https://doi.org/10.1007/s40726-015-0004-2>
- Singh, M., Srivastava, P. K., Verma, P. C., Kharwar, R. N., Singh, N., & Tripathi, R. D. (2015). *Soil Fungi for Mycoremediation of Arsenic Pollution in Agriculture Soils*. <https://doi.org/10.1111/jam.12948>
- Slyemi, D., & Bonnefoy, V. (2012). How prokaryotes deal with arsenic. *Environ Microbiol Rep*, 4, 571–586. <https://doi.org/10.1111/j.1758-2229.2011.00300.x>
- Schn rer, J., & Rosswall, T. (1982). Fluorescein diacetate hydrolysis as a measure of total microbial activity in soil and litter. *Appl Environ Microbiol*, 43, 1256–1261.
- Srivastava, S., & Sharma, Y. K. (2013). Arsenic occurrence and accumulation in soil and water of eastern districts of Uttar Pradesh, India. *Environmental Monitoring and Assessment*, 185, 4995–5002.
- Srivastava, P. K., Vaish, A., Dwivedi, S., Chakrabarty, D., Singh, N., & Tripathi, R. D. (2011). Science of the Total Environment Biological removal of arsenic pollution by soil fungi. *Science of the Total Environment*, the, 409, 2430–2442. <https://doi.org/10.1016/j.scitotenv.2011.03.002>
- Su, S. M., Zeng, X. B., Li, L. F., Duan, R., Bai, L. Y., Li, A. G., Wang, J., & Jiang, S. (2012). Arsenate reduction and methylation in the cells of *Trichoderma asperellum* SM-12F1, *Penicillium janthinellum* SM-12F4, and *Fusarium oxysporum* CZ-8F1 investigated with X-ray absorption near edge structure. *Journal of Hazardous Materials*, 243, 364–367.
- Tapase, S. R., & Kodam, K. M. (2018). Assessment of arsenic oxidation potential of *Microvirga indica* S-MI1b sp. nov. in heavy metal polluted environment. *Chemosphere*, 195, 1–10. <https://doi.org/10.1016/j.chemosphere.2017.12.022>
- Valverde, A., González-Tirante, M., Medina-Sierra, M., Santa-Regina, I., García-Sánchez, A., & Igual, J. M. (2011). Diversity and community structure of culturable arsenic-resistant bacteria across a soil arsenic gradient at an abandoned tungsten-tin mining area. *Chemosphere*, 85, 129–134. <https://doi.org/10.1016/j.chemosphere.2011.06.025>
- Vance, E. D., Brookes, P. C., & Jenkinson, D. S. (1987). An extraction method for measuring soil microbial biomass C. *Soil Biology & Biochemistry*, 19, 703–707.
- Verma, S., Verma, P. K., Meher, A. K., Dwivedi, S., Bansiwala, A. K., Pande, V., Srivastava, P. K., Verma, P. C., Tripathi, R. D., & Chakrabarty, D. (2016). A novel arsenic methyltransferase gene of *Westerdykella aurantiaca* isolated from arsenic contaminated soil: Phylogenetic, physiological, and biochemical studies and its role in arsenic bioremediation. *Metallomics*, 8, 344–353. <https://doi.org/10.1039/c5mt00277j>
- Walkley, A. (1947). A critical examination of a rapid method for determining organic carbon in soils—effect of variations in digestion conditions and of inorganic soil constituents. *Soil Science*, 63, 251–264.
- Xi, B., Yu, H., Li, Y., Dang, Q., Tan, W., Wang, Y., & Cui, D. (2021). Insights into the effects of heavy metal pressure driven by long-term treated wastewater irrigation on bacterial communities and nitrogen-transforming genes along vertical soil profiles. *Journal of Hazardous Materials*, 403, 123853. <https://doi.org/10.1016/j.jhazmat.2020.123853>
- Xiao, K. Q., Li, L. G., Ma, L. P., Zhang, S. Y., Bao, P., Zhang, T., & Zhu, Y. G. (2016). Metagenomic analysis revealed highly diverse microbial arsenic metabolism genes in paddy soils with low-arsenic contents. *Environmental*

- Pollution*, 211, 1–8. <https://doi.org/10.1016/j.envpol.2015.12.023>
- X. Li, X. Liu, N. Cao, S. Fang, C. Yu, Adaptation mechanisms of arsenic metabolism genes and their host microorganisms in soils with different arsenic contamination levels around abandoned gold tailings, *Environmental Pollution* 291 (2021). <https://doi.org/10.1016/j.envpol.2021.117994>.
- X. Song, Y. Li, E. Stirling, K. Zhao, B. Wang, Y. Zhu, Y. Luo, J. Xu, B. Ma, AseGeneDB: a curated orthology arsenic metabolism gene database and computational tool for metagenome annotation, *NAR Genom Bioinform* 4 (2022). <https://doi.org/10.1093/nargab/lqac080>.
- Y. Wu, S.S. Santos, M. Vestergård, A.M. Martín González, L. Ma, Y. Feng, X. Yang, A field study reveals links between hyperaccumulating *Sedum* plants-associated bacterial communities and Cd/Zn uptake and translocation, *Science of the Total Environment* 805 (2022). <https://doi.org/10.1016/j.scitotenv.2021.150400>.
- Y. Gu, J.D. Van Nostrand, L. Wu, Z. He, Y. Qin, F.J. Zhao, J. Zhou, Bacterial community and arsenic functional genes diversity in arsenic contaminated soils from different geographic locations, *PLoS One* 12 (2017). <https://doi.org/10.1371/journal.pone.0176696>.
- Yang, Y., Chen, W., Meng, D., Ma, C., & Li, H. (2024). Investigation of arsenic contamination in soil and plants along the river of Xinzhou abandoned gold mine in Qingyuan. *China Chemosphere*, 359, 142350.
- Zhang, S. Y., Xiao, X., Chen, S. C., Zhu, Y. G., Sun, G. X., & Konstantinidis, K. T. (2021). High Arsenic Levels Increase Activity Rather than Diversity or Abundance of Arsenic Metabolism Genes in Paddy Soils. *Applied and Environment Microbiology*, 87, 1–15. <https://doi.org/10.1128/AEM.01383-21>
- Zhou, Y., Niu, L., Liu, K., Yin, S., & Liu, W. (2018). Arsenic in agricultural soils across China: Distribution pattern, accumulation trend, influencing factors, and risk assessment. *Science of the Total Environment*, 616–617, 156–163. <https://doi.org/10.1016/j.scitotenv.2017.10.232>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.