



Article

Utilizing a Metagenome Assembled Genome Approach Revealed Further Insights into Microbially Mediated Heavy-Metal Resistance in Soils from a Former Nuclear Materials Production Facility

Navya Kommu ¹, Paul Stothard ², Christian Chukwujindu ³, Ashish Pathak ¹ and Ashvini Chauhan ^{1,*}

¹ School of the Environment, Florida A&M University, 1515 S. Martin Luther King Blvd., Tallahassee, FL 32307, USA; navyakommu311@gmail.com (N.K.); ashish1.pathak@famu.edu (A.P.)

² Department of Agricultural, Food and Nutritional Science, University of Alberta, 2-31 General Services Bldg, Edmonton, AB T6G 2H1, Canada; stothard@ualberta.ca

³ Material & Energy Technology Department, Projects Development Institute, Emene Industrial Layout, Enugu 400104, Nigeria; engrchrischukwujindu@gmail.com

* Correspondence: ashvini.chauhan@famu.edu; Tel.: +850-412-5119; Fax: +850-561-2248

Abstract: Soils and sediments from the Savannah River Site (SRS), located in the USA are known to have a long history of co-contamination with radionuclides (mainly uranium) and heavy metals. To better understand the bacterial taxonomic and genomic characteristic of the SRS soil habitat, shotgun metagenomes were obtained from three different levels of contaminated soil—high, medium, and low. Sequences were then assembled and annotated to generate metagenome-assembled genomes (MAGs) using toolkits within the nf-core/mag. The initial analysis resulted in a total of 254 MAGs. After bin refinement and de-replication, 55 MAGs which met the quality standard with a completeness > 75% and contamination < 25%, accounting for 21.67% of all the MAGs, were reconstructed. Further refinement with completeness > 90% and contamination < 10% yielded 24 MAGs (18 from the winter season and 6 from the summer season) spanning 6 bacterial phyla, predominantly Actinomycetota, Proteobacteriota, Bacteroidota, and Cyanobacteria. Overall, the *Arthrobacter* MAG was found to be robust for further analysis, with over 1749 genes putatively involved in the crucial metabolism of elements viz. nitrogen, phosphorous, and sulfur, and 598 genes encoding enzymes for the resistance of metals including cadmium, zinc, chromium, arsenic, and copper. In summary, this project enhances our understanding of genes conferring resistance to heavy metals in uranium-contaminated soils.

Keywords: metagenome assembled genomes (MAGs); uranium; savannah river site; metagenomics



Citation: Kommu, N.; Stothard, P.; Chukwujindu, C.; Pathak, A.; Chauhan, A. Utilizing a Metagenome Assembled Genome Approach Revealed Further Insights into Microbially Mediated Heavy-Metal Resistance in Soils from a Former Nuclear Materials Production Facility. *Appl. Microbiol.* **2024**, *4*, 376–389. <https://doi.org/10.3390/applmicrobiol4010026>

Academic Editor: Ian Connerton

Received: 28 December 2023

Revised: 6 February 2024

Accepted: 8 February 2024

Published: 13 February 2024



Copyright: © 2024 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

Uranium is a naturally occurring radionuclide associated with volcanic rocks, black shales, or granite at a concentration between 2 and 4 mg/kg of soil, and it can be rendered mobile upon exposure to oxidative groundwater [1]. The US Environmental Protection Agency sets the maximum allowable uranium contamination level in groundwater to not exceed 30 µg/L; however, anthropogenic activities, including mining, nuclear fuel production, and testing, have historically led to the release of uranium into groundwater and soil, thus escalating environmental health concerns. Uranium exposure causes pathological alterations to the kidney (nephrotoxic) in both humans and animals [2], and it accumulates in bone, leading to a widespread osteotoxicity in exposed cohorts, among other serious health issues [3].

Over 50 uranium companies were involved in uranium processing and refining in the US during the cold war; however, plans for the treatment of the toxic byproducts were minimal, leading to the seepage of radioactive materials into nearby communities, streams, and aquifers [4]. One such former nuclear production facility is the Savana River Site

(SRS), which covers about 310 square miles and was built during the 1950s in the upper Atlantic coastal plain of South Carolina. Particularly, Steed Pond, an abandoned farm pond at the Northwest of the SRS, served as a settling basin for contaminants originating from the target production activities and received an estimated 44,000 kg of depleted uranium, leading to high uranium tailing diffusing into the environment, including the surface and groundwater; Punshon et al. [5] and Sowder et al. [6] reported that uranium contamination from the Steed Pond sediments could be as high as 2500 mg/kg of soil, about 4 orders of magnitude above the background concentration of 4 mg/kg from the uncontaminated areas. Therefore, investigating approaches to remediate uranium contamination from the SRS soil habitat has been a large focus of our project.

In this context, microbially mediated remediation strategies are cost-effective, eco-friendly, and efficient compared with other remediation strategies [7]. Previous studies have observed relatively high bacterial diversity in samples with different levels of uranium contamination, which suggests that bacterial communities can acclimatize and colonize uranium-contaminated soils and may also exert remediation influences, mainly by immobilizing uranium through adsorption, precipitation, intermolecular interactions, and/or redox transformations [8,9]. As part of our broader-scale project, we have evaluated bacterial and fungal communities within the SRS contaminated soils using culture-based and culture-independent techniques, which has revealed that certain phyla are predominant in the soil, and they possess genes encoding enzymes for the resistance of heavy metals.

Our ongoing culture-dependent and culture-independent work on SRS metalliferous soils has demonstrated several different bacterial groups, predominantly from the following genera: *Arthrobacter* spp., *Burkholderia* spp., *Bacillus* spp., *Bradyrhizobium* spp., *Pseudomonas* spp., *Lysinibacillus* spp., *Paenibacillus* spp., *Stenotrophomonas* spp., and *Serratia* spp. [10–14]. It is noteworthy that researchers from another DOE site in Washington State (the U-contaminated Hanford site), which has a similar uranium contamination history to the SRS site, reported *Arthrobacter* spp. to be predominant [15]. *Arthrobacter* spp. have also been isolated from ecosystems under extreme environmental stresses, such as a nuclear waste plume and U mined site, as well as heavy-metal-contaminated habitats. It is very likely that *Arthrobacter* spp. have acquired robust genomic traits to facilitate their survival under heavy-metal stress. In this regard, Suzuki and Banfield [16] showed that *Arthrobacter* spp. accumulate uranium intracellularly in the form of precipitates closely associated with polyphosphate granules, thus performing natural attenuation on sites laden with uranium contamination. Similarly, Bader et al. [17] and Banala et al. [7] showed that *Actinomycetes*, such as *Arthrobacter* spp., have the highest uranium sorption capacity of up to 971 mg U/g among several tested strains. *Actinomycetes* are gram-positive, ubiquitous in soil ecosystems where they perform critical ecological roles, including the decomposition of organic materials, such as lignin and chitin [18]. A better understanding of the basis of metal-microbial interactions could be instrumental for the successful restoration of nuclear-legacy contaminated environments.

The difficulty of culturing *Actinomycetota* and other microbial species that play significant roles in uranium remediation underscores the need to apply metagenomics assembled genome (MAG) techniques to obtain a comprehensive understanding of microorganisms and their environmentally relevant functions represented in shotgun genome sequence libraries, such as the one available through our project from samples collected in summer and winter, from low, medium, and high levels of uranium contamination. Furthermore, reference genomes for many un-culturable microbes are essential for a detailed functional and taxonomic characterization of species in a microbial population of a given niche; however, the vast majority of microbes cannot be isolated for individual sequencing; thus, the MAG technique is a viable alternative [19]. Therefore, the main objective of this study was to reconstruct microbial genomes from shotgun metagenome sequence data we have obtained through our ongoing project focused on obtaining a better understanding of heavy-metal-resistant genes in historically contaminated soils [12,13,20].

2. Materials and Methods

2.1. Sample Collection

Samples were collected from the SRS Steeds Pond-Tims Branch riparian stream system, as shown in our previous report [12]. Surface soil samples were collected in duplicate from three Steed Pond locations during the winter (January 2016) and summer (July 2017) seasons. Samples were stored in sterile containers and transported to FAMU for microbial analysis.

2.2. Shotgun Metagenomics and Bioinformatics Processing

DNA was extracted from the duplicate samples representing high, medium, and low uranium contamination using a DNeasy Powersoil Kit (QIAGEN Inc., Germantown, MD, USA) per the manufacturer's instructions. A NanoDrop 1000 (NanoDrop Technologies, Wilmington, DE, USA) was used to quantify the total DNA of the extracted samples by measuring the concentrations (ng/μL) using the absorbance ratios A260/280 and A260/230. An aliquot of high-quality genomic DNA was processed for the shotgun sequencing in accordance with protocols described by Illumina Inc. (San Diego, CA, USA). The equimolar libraries were pooled and sequenced on an Illumina HiSeq platform using 2 × 150 bp chemistry.

The processing of the obtained raw reads involved mapping to the NCBI non-redundant (NR) protein database using DIAMOND and default parameters [21] followed by taxonomic assignment using MEGAN's least common ancestor algorithm [22]. The raw read counts were transformed to obtain the relative abundance, which was then subjected to filtration based on the taxonomic rank of kingdom, e.g., bacteria, archaea, viruses, fungi, and non-fungal eukaryotes, which yielded the respective abundance tables. The functional profiling of these datasets was carried out using SUPER-FOCU to obtain the estimated abundance of genes and subsystem pathways, which were used to represent the persistence of different pathways and their specificity.

2.3. Bioinformatics Analysis to Reconstruct Metagenome Assembled Genomes (MAGs)

The raw shotgun sequence reads were downloaded from the NCBI genome repository (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA436168>, accessed on 20 July 2023), as stated below:

- (1) Soil from summer with high U contamination: SRR6788480, SRR6788481, and SRR6788482;
- (2) Soil from summer with medium U contamination: SRR6788478, SRR6788479, and SRR6788487;
- (3) Soil from summer with low U contamination: SRR6788483, SRR6788484, and SRR6788485;
- (4) Soil from winter with high U contamination: SRR6788486;
- (5) Soil from winter with medium U contamination: SRR6788488;
- (6) Soil from winter with low U contamination: SRR6788489.

The MAG pipeline nf-core (<https://nf-co.re/mag>, accessed on 20 July 2023) was run on the soil data, and the genome assembly was then performed using MEGAHIT. Due to the size, the data had to be processed in groups based on the sample type. The assembled sequences generated by MEGAHIT were evaluated using the computational tool QUAST. The genome binning was then performed using MetaBAT2, and the quality of the bins, including the completeness and the contamination levels, was evaluated using CheckM. The CheckM analysis was also used to perform the taxonomic classification of the bins.

Once the sequence binning of the MAGs was complete, we performed an additional process using the program dRep, which compared the obtained MAGs, enabling us to choose the best representative for each genome. dRep was also utilized to observe which MAGs are closely related and to further analyze genomes of interest. Prokka and Prodigal were used to predict the protein coding genes of the relevant MAG. The bins were assigned a taxonomic classification using the GTDB-Tk and CAT toolkits. Other annotation workflows such as rapid annotation using subsystems technology (RAST) [23], the Pathosystems

Resource Integration Center (PATRIC) [24], and IslandViewer [25] were utilized for further genomic analysis on MAGs of specific interest.

3. Results and Discussion

3.1. Characteristics of the MAGs Recovered from Shotgun Metagenomes Using the CheckM Pipeline

Genes assembled using MAGs have lower quality compared to the isolated genomes because MAGs are a mix of genes from different microbial populations; however, this technique has contributed to the discovery of a vast array of novel organisms and functions from diverse samples, including guts, rumens, sediments, and soil, with a three-step basic conventional workflow: preprocessing of sequence reads, construction of MAGs, and quantification and annotation of MAGs [19,26]. Using a MAG approach, a total of 55,681 Mbytes of sequences were obtained from 12 soil samples, with a variable level of uranium contamination. The binning of the metagenomics assembled sequences was carried out with the MetaBAT pipeline and the quality assessment of the obtained bins was performed by CheckM, all within the nf-core domain. CheckM does a phylogenetic placement of the bin genomes into its separate species tree, which allows the computation of the universal and additional single copy marker genes specific for a particular lineage, giving rise to the completeness of the bins [27]. About 254 individual bins were recovered from the SRS soil samples; however, due to the sheer volume of data retrieved, the data table was limited to MAGs produced with >75% completeness and <25% contamination (Figure S1) that possessed unique marker lineages representative of the range of genomes present in the samples. Genome binning was then performed using MetaBAT2, and the quality of the bins, including the completeness and the contamination levels, was evaluated using CheckM, thus providing an estimate of the completeness and redundancy of the processed genomic data based on universal single-copy orthologs (Table S1).

Furthermore, as shown in Table 1, the CheckM analysis represented important features of metagenome assembled genomes (MAGs) reconstructed from the SRS soil samples with >90% completeness, and <10% contamination. Specifically, the quality assessment with CheckM reduced the number of bins to 24, distributed at 1, 5, 9, 5, and 4 for the summer high, summer low, winter high, winter low, and winter medium uranium-contaminated soil samples, respectively. Moreover, out of the 24 bins that passed the completeness threshold, five were classified to the phylum level, four to the class level, and the rest were at the kingdom level, suggesting that the MAG technique for the assessment of the SRS soil samples likely captured un-culturable bacteria not represented in the database. In addition, two of the five phyla belonged to *Actinobacteria*, which was identified in our previously published findings on the uranium-contaminated SRS ecosystem [13,14]; thus, the MAG further highlights the potential role of this phyla in heavy-metal resistance processes.

On a broader scale, phylogenomic analysis based on sets of single-copy marker genes universal to the bacterial or archaeal domains showed that the 254 MAGs reconstructed from SRS soils regardless of the season or contamination level, were distributed in the following taxonomic groups: *Alphaproteobacteria*, *Gammaproteobacteria*, *Deltaproteobacteria*, *Betaproteobacteria*, *Euryarchaeota*, *Archaea*, *Acidobacteria*, *Actinobacteria*, and *Rhizobiales*. These findings are in line with our previous reports on the taxonomic assessment of the SRS soils [12,13,20].

Table 1. Metagenome-assembled genome (MAG)-based characteristics from the shotgun metagenomes obtained from soils from the Savannah River Site (SRS) with variable levels of uranium contamination.

Genome	Marker Lineage	Completeness	Contamination	Strain_Heterogeneity	Length	N50
soil-summer-high-uranium-results-MEGAHIT-group-0.13.fa	p__Actinobacteria (UID1454)	95.73	1.42	33.33	3,083,233	58,255
soil-summer-low-uranium-results-MEGAHIT-group-0.12.fa	k__Bacteria (UID1452)	92.75	2.42	50	3,127,523	20,566
soil-summer-low-uranium-results-MEGAHIT-group-0.21.fa	p__Actinobacteria (UID1454)	94.44	3.16	14.29	2,333,606	75,210
soil-summer-low-uranium-results-MEGAHIT-group-0.24.fa	p__Euryarchaeota (UID3)	99.2	1.6	0	2,661,188	139,475
soil-summer-low-uranium-results-MEGAHIT-group-0.3.fa	p__Euryarchaeota (UID49)	98.53	2.61	0	2,895,156	27,231
soil-summer-low-uranium-results-MEGAHIT-group-0.6.fa	k__Archaea (UID2)	91.75	2.91	0	2,256,961	33,599
soil-winter-high-uranium-results-MEGAHIT-group-0.17.fa	c__Betaproteobacteria (UID3971)	95.79	3.48	36.36	4,120,679	30,345
soil-winter-high-uranium-results-MEGAHIT-group-0.23.fa	k__Bacteria (UID3187)	91.88	5.27	12.5	8,053,684	54,917
soil-winter-high-uranium-results-MEGAHIT-group-0.33.fa	p__Bacteroidetes (UID2591)	94.88	2.64	37.5	3,768,066	14,449
soil-winter-high-uranium-results-MEGAHIT-group-0.35.fa	k__Bacteria (UID1452)	91.36	4.63	16.67	2,805,076	12,991
soil-winter-high-uranium-results-MEGAHIT-group-0.40.fa	c__Gammaproteobacteria (UID4202)	92.77	3.83	23.81	3,661,217	32,786
soil-winter-high-uranium-results-MEGAHIT-group-0.5.fa	c__Alphaproteobacteria (UID3305)	96.3	3.07	30	4,186,502	26,896
soil-winter-high-uranium-results-MEGAHIT-group-0.51.fa	k__Bacteria (UID1452)	91.67	1.57	0	4,442,876	15,948
soil-winter-high-uranium-results-MEGAHIT-group-0.52.fa	k__Bacteria (UID3187)	90.17	3.47	16.67	4,854,587	19,436
soil-winter-high-uranium-results-MEGAHIT-group-0.61.fa	k__Bacteria (UID3187)	92.07	9.96	31.82	5,211,487	10,544
soil-winter-low-uranium-results-MEGAHIT-group-0.12.fa	k__Bacteria (UID3187)	97.15	2.56	0	6,711,185	35,231
soil-winter-low-uranium-results-MEGAHIT-group-0.38.fa	k__Bacteria (UID203)	91.23	5.5	75	5,025,612	9893
soil-winter-low-uranium-results-MEGAHIT-group-0.47.fa	c__Gammaproteobacteria (UID4202)	94.57	8.5	20	3,319,549	31,643
soil-winter-low-uranium-results-MEGAHIT-group-0.56.fa	k__Bacteria (UID1452)	97.73	9.05	7.69	4,126,511	52,840
soil-winter-low-uranium-results-MEGAHIT-group-0.9.fa	k__Bacteria (UID3187)	94.02	6.41	0	8,036,188	19,973
soil-winter-medium-uranium-results-MEGAHIT-group-0.2.fa	k__Bacteria (UID3187)	95.22	6.09	0	6,095,384	47,944
soil-winter-medium-uranium-results-MEGAHIT-group-0.3.fa	k__Bacteria (UID1452)	91.09	3.47	75	2,893,218	16,581
soil-winter-medium-uranium-results-MEGAHIT-group-0.37.fa	k__Bacteria (UID3187)	94.87	4.7	16.67	6,098,353	118,469
soil-winter-medium-uranium-results-MEGAHIT-group-0.39.fa	k__Bacteria (UID1452)	91	1.39	0	2,674,289	7171

3.2. Statistical Analysis of the MAGs from SRS Uranium-Contaminated Soils

The recovered MAGs were analyzed statistically using MASH to evaluate which MAGs are closely related and were plotted using a dendrogram (Figure 1). The MASH tool extends the MinHash dimensionality-reduction technique to include a pairwise mutation and p-value significance test for the efficient clustering of sequence collections [28]. However, the dendrogram analysis revealed that the MAGs did not cluster based on the uranium concentration or the sampling season. This dendrogram was constructed after the program dRep was used to identify sets of genomes that are similar and choose the best representative genome from each set (referred to as the “winning genome”). dRep

is a metagenomics tool that identifies groups of essentially identical genomes from independently assembled genomes of metagenomic samples and selects the best genome from each replicate set [29]. The dRep algorithm combines different metagenomics tools, such as CheckM for completeness-based genome filtering, Mash for the fast grouping of similar genomes, gANI for accurate genome comparison, and Scipy for hierarchical clustering [29]. The 254 MAGS from SRS shotgun metagenomes were compared and processed to choose the winning genomes. The analysis of the winning genomes revealed the predominance of *Arthrobacter oryzae*, thus pointing to the relevance of this soil-borne bacterial genera in the SRS historically contaminated soils.

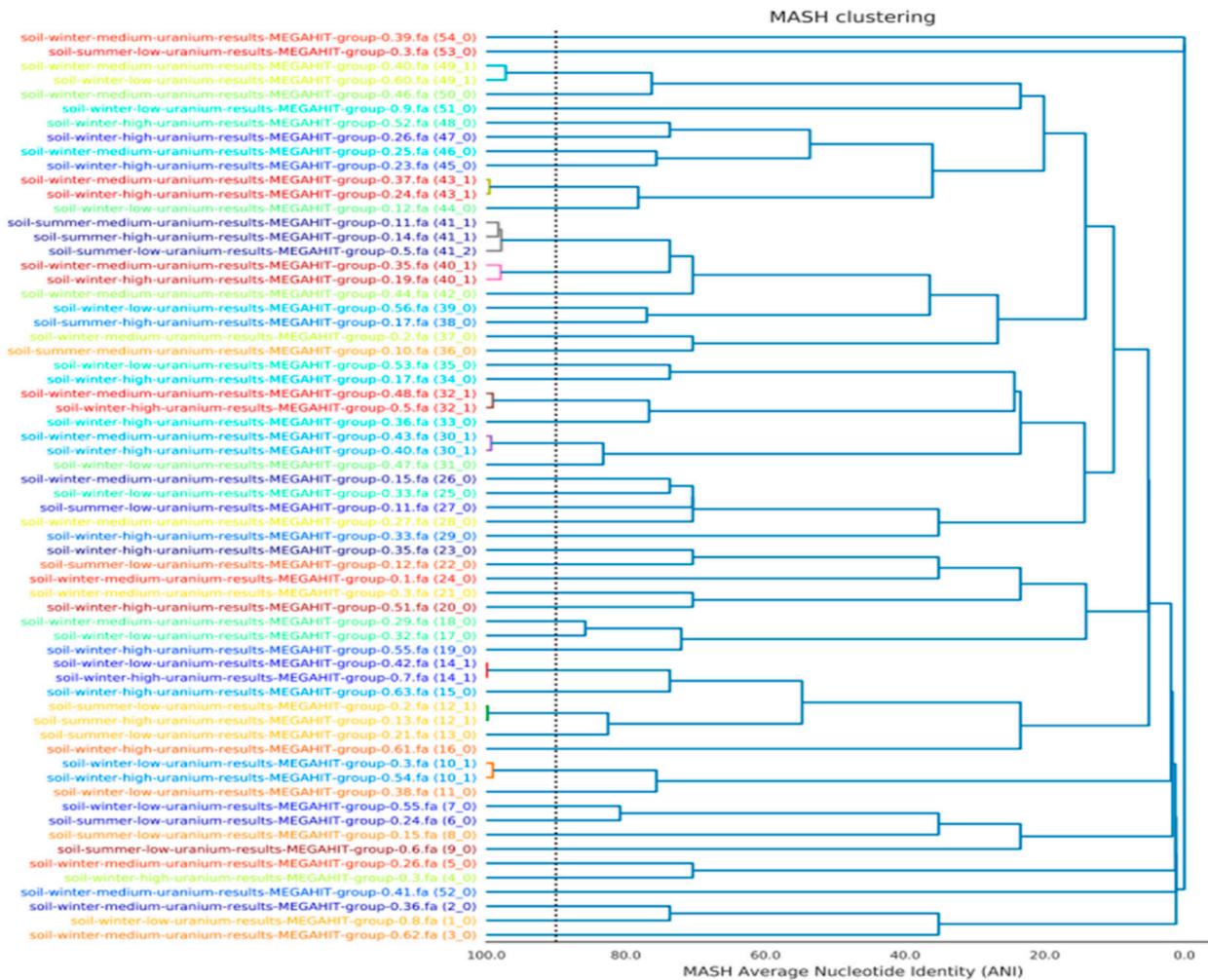


Figure 1. Dendrogram of the best representative MAGs from SRS uranium-contaminated soil samples. The dendrogram was constructed using MASH based on the average nucleotide identity (ANI) to illustrate which MAGs are closely related. The dendrogram revealed that the MAGs did not cluster based on uranium concentration or the sampling season.

3.3. Assessment of Metabolic Traits of the Winning MAG—*Arthrobacter oryzae*

RAST is a fully automated service for annotating complete or nearly complete bacterial or archaeal genomes by identifying protein-encoding and rRNA and tRNA genes, assigning functions to the genes, and predicting which subsystems are represented in the genomes [23]. Figure 2 shows a metabolic model constructed from the SRS MAG belonging to *Arthrobacter oryzae*. This MAG was specifically chosen for further genomic analysis as *Arthrobacter* is one bacterial genus previously identified to perform critical role(s) in heavy-metal resistance within the SRS soil habitat [14]. This model obtained using the RAST in silico workflow [30] revealed that only 26% of the genes could be assigned to

known functions and 74% were found to not have any affiliation with biologically known functions. This indicates that contaminated soil environments could potentially be driving the evolution of microbes with genomic traits that were hitherto unrepresented in the annotation databases.

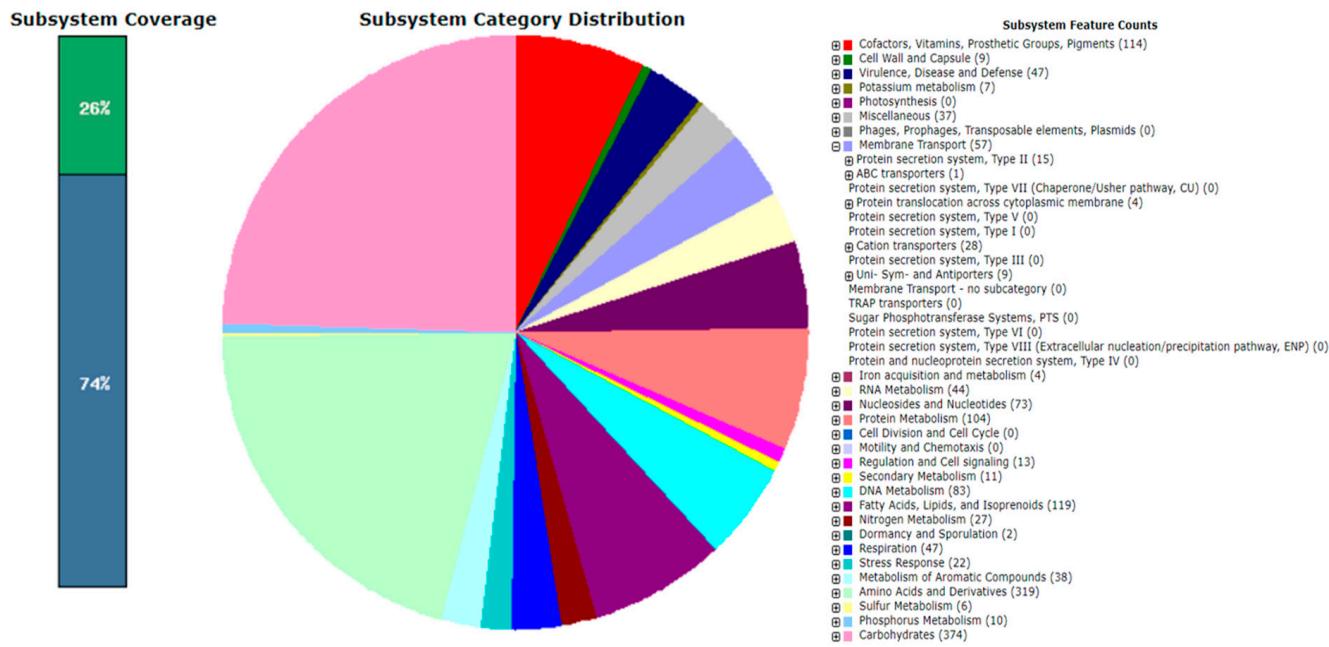


Figure 2. Metabolic traits of the MAG (*Arthrobacter oryzae*) from SRS uranium-contaminated soil. The Metabolic trait of *Arthrobacter oryzae* was computed using Rapid Annotations using Subsystems Technology (RAST). The model chart indicates that 26% of the genes could be assigned to known functions, with 74% with no affiliation to known functions. Carbohydrate metabolism, amino acid and derivatives, protein metabolism, cofactor and vitamin metabolism, and membrane transport are among the major gene classes identified in this species.

Most notably, carbohydrate metabolism, protein metabolism, cofactor and vitamin metabolism, and membrane transport were among the major gene classes identified in the MAG (Figure 2). These features suggest the tendency of *Arthrobacter oryzae* to participate in high metabolic and catabolic activity in its contaminated habitat, which likely plays a key role in its survival and ability to degrade harmful pollutants. Membrane transport is especially important for bioremediation because it allows microorganisms to regulate their internal environment by extruding toxic chemicals, such as uranium, thereby maintaining the cytoplasmic content and homeostatic balance necessary for growth [31].

3.4. Gene Functional Analysis of the *Arthrobacter oryzae* MAG

Gene functional analysis was performed on the *Arthrobacter oryzae* MAG using the One Codex bioinformatics pipeline, and the results are shown in Figure 3. It is clear from this analysis that the ATP-binding-cassette family (ABC), as well as the major facilitator superfamily (MFS) are among the two most overrepresented genes in the SRS *Arthrobacter* MAG. ABC transporter proteins and the MFS are two of the five major families of efflux transporters (ref), and are thus likely important for rendering heavy-metal resistance to the native SRS soil microbiota. According to Akhtar and Turner [32], the ABC transporters superfamily are transporters found in all domains of life, translocating various molecules, such as toxins and xenobiotics to the outside of the cell's cytoplasm, or they can even import various nutrients for cellular survival. In the same way, the major facilitator superfamily (MFS) is another membrane transporter that plays a key role in the maintenance of cellular elements. Pao et al. [33] reported the occurrence of several distinct families within the MFS, each of which generally transports a single class of organic and inorganic anions and

cations. Note that our previous analysis of heavy-metal-contaminated soils also showed the ABC and MFS to be predominant in historically contaminated soils. Moreover, the analysis performed on the SRS *Arthrobacter* MAG using the KBase bioinformatics pipeline yielded an abundance of genes notable to DNA recombination and repair, as well as defense and virulence responses, as shown in Table S2. A plethora of defense and repair mechanisms is essential to counter the damage this organism may face from constant exposure to a heavy-metal-contaminated environment. Furthermore, this species was shown to contain several metal-transporting ATPases and metal-efflux proteins, that are important adaptations for survival in a heavy-metal-contaminated environment and enhance the heavy-metal-resistance systems this organism harbors [34]. If the metabolic efficacy of the above stated gene classes is enhanced in contaminated environments, perhaps the bioremediation of heavy metals can be rapidly accomplished for the better management of contaminated habitats.

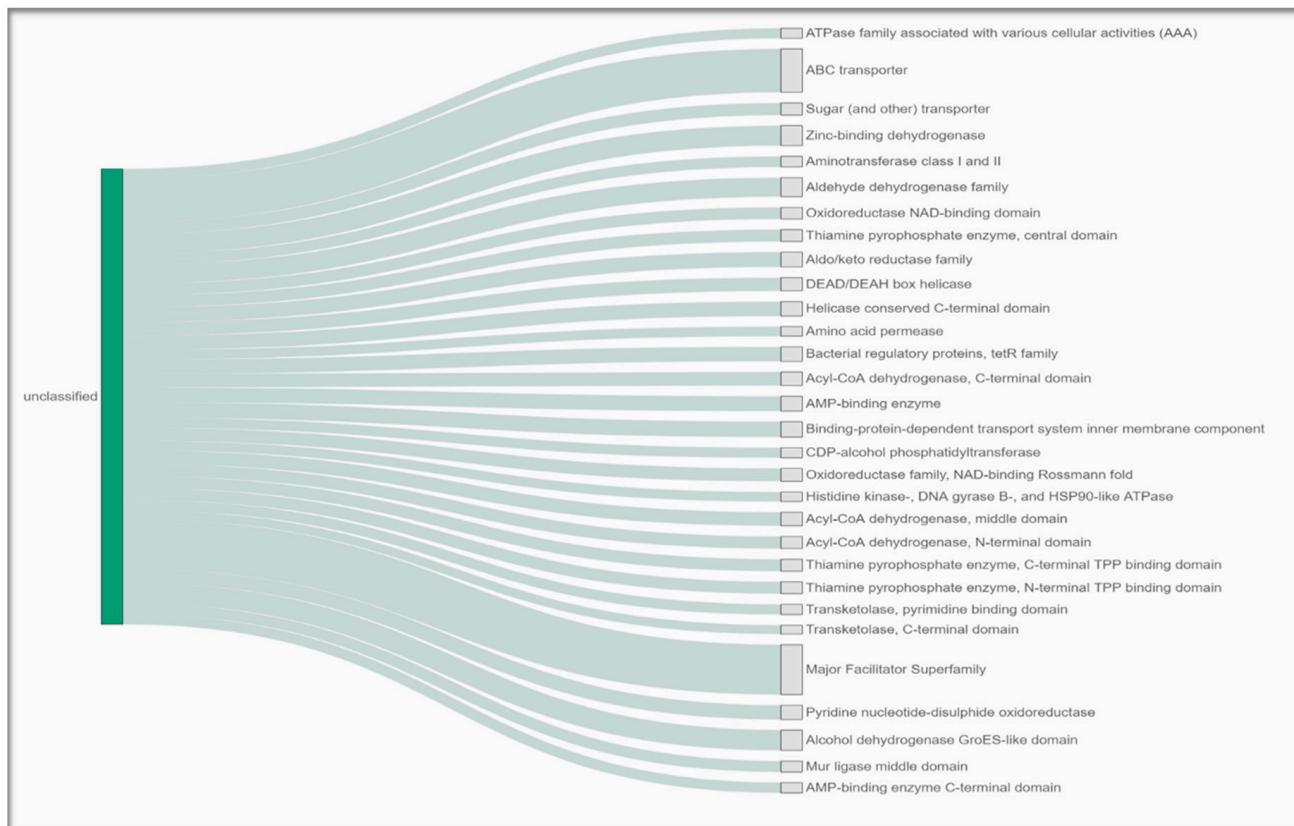


Figure 3. Functional analysis of the *Arthrobacter oryzae* genome to examine specific genes that aid in uranium bioremediation. The uranium bioremediation functional analysis was carried out using One Codex bioinformatics pipeline. The ATP-binding-cassette family (ABC), and the major facilitator superfamily are among the most overrepresented genes in this species. ABC transporter proteins and the major facilitator superfamily are two of the five major families of efflux transporters. Efflux transporters are important for bioremediation because they not only prevent contaminants from entering a cell, but also convert harmful compounds within a cell to a more hydrophilic form.

3.5. Genome Map of *Arthrobacter oryzae* MAG Reconstructed from the SRS Uranium-Contaminated Soil

Figure 4 shows the circular genome map of *Arthrobacter oryzae* MAG depicting DNA contigs, AMR Genes, VF Genes, drug targets, GC content, GC skew, protein families, efflux pumps, permease proteins, and transporters, obtained using the Pathosystems Resource Integration Center (PATRIC) pipeline. PATRIC is a bacterial bioinformatics pipeline that can be used to run a variety of genomic analyses including a genomic assembly using different strategies, genomic annotation using a RAST tool kit, and the provision of several annotated genomes for comparative studies [35]. The Circular Viewer tool in PATRIC provides the circular genomic map using all the contigs from an assembly via the Circos technology (www.patricbrc.org). The genomic map shown in Figure 4 consists of 13 circles from the outside in; the first shows the size in Mbp, followed by DNA contigs, CDS forward and reverse, AMR Genes, VF Genes, drug targets, GC contents, GC skew, protein families, efflux pumps, permease proteins, and transporters.

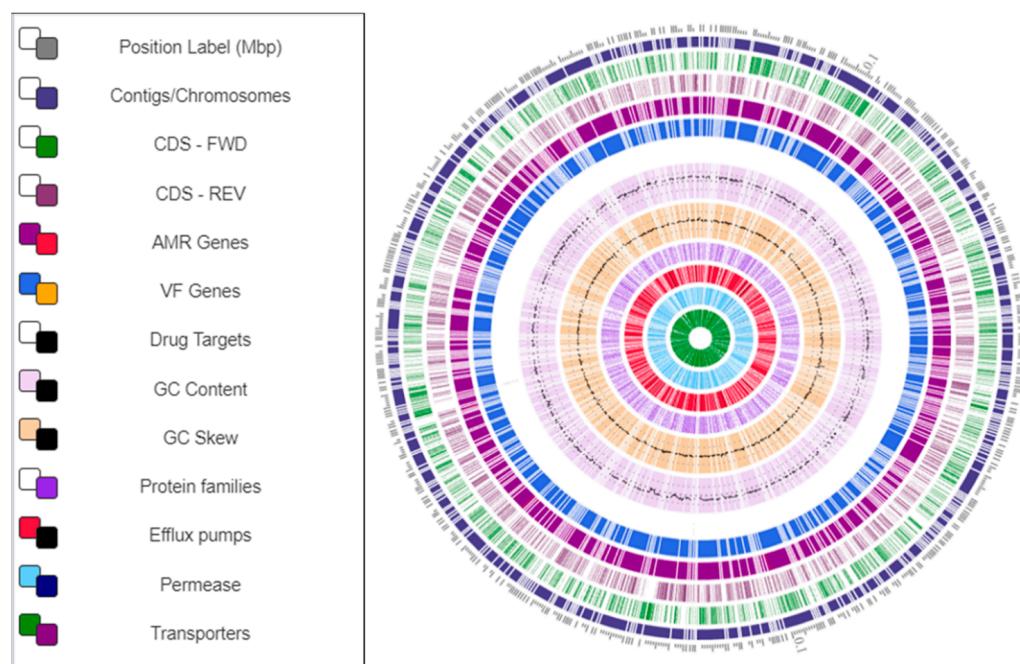


Figure 4. Circular genome map using PATRIC default settings to visualize the genomic features of *Arthrobacter oryzae* species from SRS uranium-contaminated soil. The map consists of 13 circles from outside in; the first shows size in Mbp, followed by DNA contigs, CDS forward and reverse, AMR Genes, VF Genes, drug targets, GC contents, GC skew, protein families, efflux pumps, permease proteins, and transporters.

In the genomic map of the MAG, of particular interest was the identification of several protein families, efflux pumps, and permease proteins, as shown by the white/purple, red/black, and light/dark blue respectively. As stated above, efflux pumps are transport proteins involved in the extrusion of hazardous substrates from within the cells into the external environment, and they are linked to bioremediation mechanisms [36]. In addition, permease proteins are essential membrane-transport mechanisms that allow the introduction of stronger proteins into the bacterial cell through genetic engineering to enhance bioremediation traits [37]. Interestingly, Cho et al. [38] drafted the whole genome of an *Arthrobacter oryzae* strain from a site not affected by anthropogenic activities and reported that the genome contains many heavy-metal-resistant genes, so bioremediation may be a central genome-enabled trait of this genera.

Moreover, this study also identified several antimicrobial resistance (AMR) genes, depicted by the purple/red lines on the genomic map. Specifically, the number and types of specialized genes found in the MAG of *Arthrobacter oryzae* are shown in Supplementary Table S3, thus further supporting our observation that the SRS contaminated site is likely “priming” the evolution of antibiotic resistant microbes; this contributes to the public health threat that these contaminated sites pose.

3.6. Genomic Islands (GEIs) in the *Arthrobacter oryzae* MAG

Genomic Islands (GEIs) are clusters of genes in a microbial genome acquired by horizontal gene transfer, carrying genes that are important for genome evolution and adaptation to niches, such as genes involved in heavy-metal resistance [39]. Figure 5 shows the identification of several genomic islands (GEIs) in the MAG of *Arthrobacter oryzae*. The outer green circle represents the scale line in Mbps and GEIs obtained from each of the following methods are shown in color: SIGI-HMM (orange), IslandPath-DIMOB (blue), and integrated detection (red). GEI-prediction methods usually exploit sequence composition and sporadic phylogenetic distribution as the main features of GEI in a genome; based on these features, the GEI methods can be either composition based or comparative-genomic based [40]. The SIGI-HMM method uses the Hidden Markov Models to predict the GEIs [41], while the IslandPath-DIMOB and integrated detection methods are based on the GEI structure of the genome [39]. The visual in Figure 5 shows several potential genomic islands within the *Arthrobacter* MAG and further analysis of the GEI regions using BLAST showed gene homologs with functions of metal resistance and the biodegradation of contaminants. Molecular transport mechanisms, such as ABC transporters and MFS genes, as mentioned earlier, can facilitate metabolic processes that transform toxic metal compounds into less detrimental forms, as well as the development of stronger cell walls and membranes used for protection against the toxic effects of heavy metals. These traits can be acquired via a horizontal gene transfer processes [42]. Moreover, it has been shown that heavy-metal-resistance mechanisms can differ greatly between different taxa, even between those within the same community [43]. Thus, it is possible that this *Arthrobacter oryzae* species acquired taxa-specific metal-resistance traits to allow it to survive in its stressful environment through horizontal-gene-transfer mechanisms. This suggests that acquiring GEIs likely facilitates the survival of microorganisms in a heavy metal and aromatic-compound-contaminated environment, such as the Savannah River Site (SRS); these findings mirror our previous genome-enabled studies on isolated organisms, such as *Arthrobacter*, *Serratia*, *Stenotrophomonas*, and *Burkholderia* from the SRS site [10–14,44]. In summation, this MAG-based study lends support to our previous studies conducted on the SRS heavy-metal-resistant microbiota and further enhances our understanding of genes conferring resistance to heavy metals in uranium-contaminated soils. These findings can be applied to develop innovative methods to restore heavy-metal-contaminated sites such as the SRS.

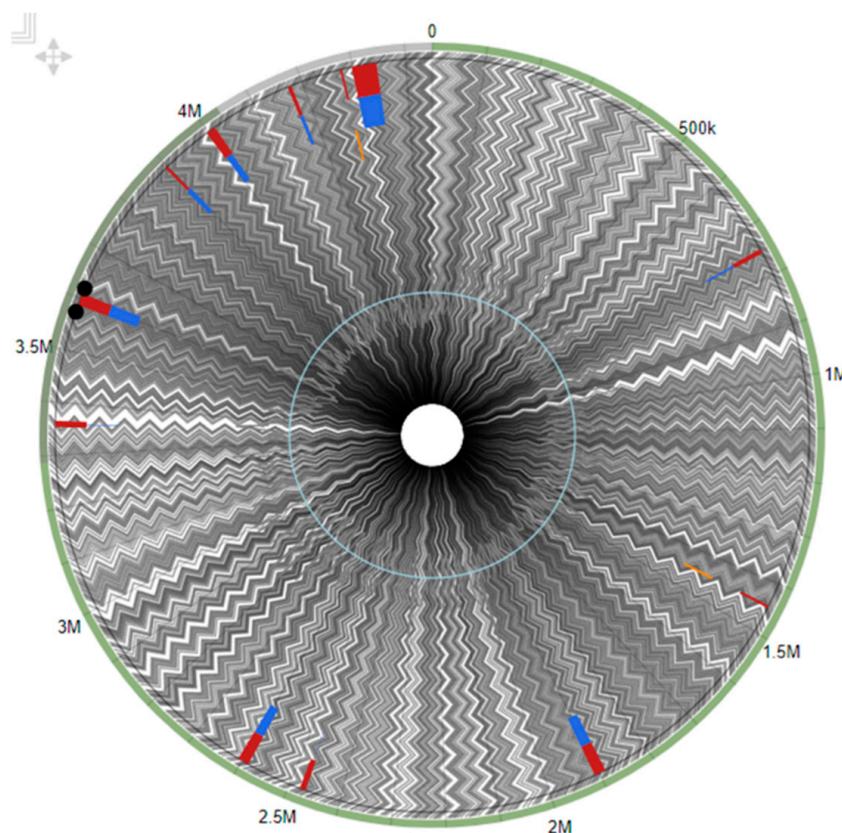


Figure 5. The Genomic Island (GEI) view of *Arthrobacter oryzae* MAG obtained with Islandviewer 4. GEIs obtained from each of the following methods are shown in color: SIGI-HMM (orange), IslandPath-DIMOB (blue), and integrated detection (red).

4. Conclusions

The Metagenomics assembled genome (MAG) technique was applied to enhance the understanding of bacterial taxonomic characteristics and genes conferring resistance to heavy metals at the uranium-contaminated Savana River Site. The MAG techniques yielded 254 genomes from 12 samples collected from the metalliferous SRS, from which we selected 24 MAGs based on the 90% threshold completeness and less than 10% contamination level and derePLICATION and refinement. Analysis of the winning MAGs revealed the predominance of *Arthrobacter oryzae*, a species previously known to provide heavy-metal-resistance roles in the SRS soil habitat. Focusing on the *A. oryzae* MAG, the gene functional analysis laid bare the ATP-binding-cassette family (ABC) and the major facilitator superfamily (MFS) as the two most overrepresented genes, which are, obviously, among the efflux transporters, necessary for rendering heavy-metal resistance to the SRS soil microbiome. Moreover, this study uncovered several potential genomic islands within the *Arthrobacter* MAG showing gene homologs with functions of metal resistance and the biodegradation of contaminants. This revelation has provided more insight into the microbiome community at the SRS site and can contribute to the ongoing efforts toward the SRS restoration.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/applmicrobiol4010026/s1>, Figure S1: Overall winner MAGs produced from the SRS soils; Table S1: Winning Genome information with Marker lineage reconstructed after dRep analysis on the MAG generated from SRS samples with varying uranium concentrations collected at different seasons; Table S2: Analysis performed on the MAGs of *Arthrobacter oryzae* using the KBase bioinformatics pipeline; Table S3: Number and types of specialized genes found in the MAG of *Arthrobacter oryzae*.

Author Contributions: Conceptualization A.C. and P.S.; methodology, P.S.; software, P.S.; validation, A.C., P.S., N.K. and C.C.; formal analysis, A.C., N.K., A.P. and C.C.; data curation, A.C. and A.P.; writing—original draft preparation, A.C., N.K., A.P., C.C. and P.S.; writing—review and editing, A.C., N.K., C.C. and P.S.; visualization, P.S.; supervision, A.C.; project administration, A.C.; funding acquisition, A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was jointly supported by the following grants: National Science Foundation (awards 1901371 and 2200615); the Department of Energy (DOE) Minority Serving Institution Partnership Program (MSIPP) under task order agreement 0000602538; Department of Energy's University Training & Research Program University Coal Research (UCR) and Historically Black Colleges and Universities and Other Minority Institutions (HBCU-OMI) award #DE-FE0032198; Department of Defense contract #W911NF2210145.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw shotgun sequence reads were obtained in a related study.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

- Nolan, J.; Weber, K.A. Natural Uranium Contamination in Major U.S. Aquifers Linked to Nitrate. *Environ. Sci. Technol. Lett.* **2015**, *2*, 215–220. [[CrossRef](#)]
- Vicente-Vicente, L.; Quiros, Y.; Pérez-Barriocanal, F.; López-Novoa, J.M.; López-Hernández, F.J.; Morales, A.I. Nephrotoxicity of uranium: Pathophysiological, diagnostic and therapeutic perspectives. *Toxicol. Sci.* **2010**, *118*, 324–347. [[CrossRef](#)] [[PubMed](#)]
- Kurttio, P.; Komulainen, H.; Leino, A.; Salonen, L.; Auvinen, A.; Saha, H. Bone as a Possible Target of Chemical Toxicity of Natural Uranium in Drinking Water. *Environ. Health Perspect.* **2005**, *113*, 68–72. [[CrossRef](#)]
- Olalde, M.; Simon, M.; Mierjeski, A. The Cold War Legacy Lurking in U.S. Groundwater. PROPUBLICA. 2022. Available online: <https://www.propublica.org/article/uranium-mills-pollution-cleanup-us> (accessed on 11 February 2024).
- Punshon, T.; Gaines, K.F.; Jenkins, R.K. Bioavailability and Trophic Transfer of Sediment-Bound Ni and U in a Southeastern Wetland System. *Arch. Environ. Contam. Toxicol.* **2003**, *44*, 30–35. [[CrossRef](#)] [[PubMed](#)]
- Sowder, A.G.; Bertsch, P.M.; Morris, P.J. Partitioning and availability of uranium and nickel in contaminated riparian sediments. *J. Environ. Qual.* **2003**, *32*, 885–898. [[CrossRef](#)]
- Banala, U.K.; Indradyumna Das, N.P.; Toleti, S.R. Microbial interactions with uranium: Towards an effective bioremediation approach. *Environ. Technol. Innov.* **2021**, *21*, 101254. [[CrossRef](#)]
- Yan, X.; Luo, X.; Zhao, M. Metagenomic analysis of microbial community in uranium-contaminated soil. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 299–310. [[CrossRef](#)] [[PubMed](#)]
- Merroun, M.L.; Selenska-Pobell, S. Bacterial interactions with uranium: An environmental perspective. *J. Contam. Hydrol.* **2008**, *102*, 285–295. [[CrossRef](#)] [[PubMed](#)]
- Agarwal, M.; Pathak, A.; Rathore, R.S.; Prakash, O.; Singh, R.; Jaswal, R.; Seaman, J.; Chauhan, A. Proteogenomic Analysis of *Burkholderia* Species Strains 25 and 46 Isolated from Uraniferous Soils Reveals Multiple Mechanisms to Cope with Uranium Stress. *Cells* **2018**, *7*, 269. [[CrossRef](#)]
- Pathak, A.; Chauhan, A.; Stothard, P.; Green, S.; Maienschein-Cline, M.; Jaswal, R.; Seaman, J. Genome-centric evaluation of *Burkholderia* sp. strain SRS-W-2-2016 resistant to high concentrations of uranium and nickel isolated from the Savannah River Site (SRS), USA. *Genom. Data* **2017**, *12*, 62–68. [[CrossRef](#)]
- Pathak, A.; Jaswal, R.; Xu, X.; White, J.R.; Edwards, B.; Hunt, J.; Brooks, S.; Rathore, R.S.; Agarwal, M.; Chauhan, A. Characterization of Bacterial and Fungal Assemblages from Historically Contaminated Metalliferous Soils Using Metagenomics Coupled with Diffusion Chambers and Microbial Traps. *Front. Microbiol.* **2020**, *11*, 1024. [[CrossRef](#)]
- Jaswal, R.; Pathak, A.; Edwards, B., III; Lewis, R., III; Seaman, J.C.; Stothard, P.; Krivushin, K.; Blom, J.; Rupp, O.; Chauhan, A. Metagenomics-Guided Survey, Isolation, and Characterization of Uranium Resistant Microbiota from the Savannah River Site, USA. *Genes* **2019**, *10*, 325. [[CrossRef](#)]
- Chauhan, A.; Pathak, A.; Jaswal, R.; Edwards, B., III; Chappell, D.; Ball, C.; Garcia-Sillas, R.; Stothard, P.; Seaman, J. Physiological and Comparative Genomic Analysis of *Arthrobacter* sp. SRS-W-1-2016 Provides Insights on Niche Adaptation for Survival in Uraniferous Soils. *Genes* **2018**, *9*, 31. [[CrossRef](#)] [[PubMed](#)]
- Katsenovich, Y.; Carvajal, D.; Guduru, R.; Lagos, L.; Li, C.-Z. Assessment of the resistance to uranium (vi) exposure by *Arthrobacter* sp. Isolated from hanford site soil. *Geomicrobiol. J.* **2013**, *30*, 120–130. [[CrossRef](#)]

16. Suzuki, Y.; Banfield, J.F. Resistance to, and Accumulation of, Uranium by Bacteria from a Uranium-Contaminated Site. *Geomicrobiol. J.* **2004**, *21*, 113–121. [[CrossRef](#)]
17. Bader, M.; Müller, K.; Foerstendorf, H.; Schmidt, M.; Simmons, K.; Swanson, J.S.; Reed, D.T.; Stumpf, T.; Cherkouk, A. Comparative analysis of uranium bioassociation with halophilic bacteria and archaea. *PLoS ONE* **2018**, *13*, e0190953. [[CrossRef](#)] [[PubMed](#)]
18. Sullivan, D.C.; Chapman, S.W. Bacteria that masquerade as fungi: Actinomycosis/nocardia. *Proc. Am. Thorac. Soc.* **2010**, *7*, 216–221. [[CrossRef](#)]
19. Zhou, Y.; Liu, M.; Yang, J. Recovering metagenome-assembled genomes from shotgun metagenomic sequencing data: Methods, applications, challenges, and opportunities. *Microbiol. Res.* **2022**, *260*, 127023. [[CrossRef](#)] [[PubMed](#)]
20. Jaswal, R.; Pathak, A.; Chauhan, A. Metagenomic Evaluation of Bacterial and Fungal Assemblages Enriched within Diffusion Chambers and Microbial Traps Containing Uraniferous Soils. *Microorganisms* **2019**, *6*, 324. [[CrossRef](#)]
21. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60. [[CrossRef](#)]
22. Huson, D.H.; Auch, A.F.; Qi, J.; Schuster, S.C. MEGAN analysis of metagenomic data. *Genome Res.* **2007**, *17*, 377–386. [[CrossRef](#)] [[PubMed](#)]
23. Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genom.* **2008**, *9*, 75. [[CrossRef](#)] [[PubMed](#)]
24. Wattam, A.R.; Davis, J.J.; Assaf, R.; Boisvert, S.; Brettin, T.; Bun, C.; Conrad, N.; Dietrich, E.M.; Disz, T.; Gabbard, J.L.; et al. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res.* **2017**, *45*, D535–D542. [[CrossRef](#)]
25. Bertelli, C.; Laird, M.R.; Williams, K.P.; Lau, B.Y.; Hoad, G.; Winsor, G.L.; Brinkman, F.S.L. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* **2017**, *45*, W30–W35. [[CrossRef](#)]
26. Saheb Kashaf, S.; Almeida, A.; Segre, J.A.; Finn, R.D. Recovering prokaryotic genomes from host-associated, short-read shotgun metagenomic sequencing data. *Nat. Protoc.* **2021**, *16*, 2520–2541. [[CrossRef](#)]
27. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [[CrossRef](#)] [[PubMed](#)]
28. Ondov, B.D.; Treangen, T.J.; Melsted, P.; Mallonee, A.B.; Bergman, N.H.; Koren, S.; Phillippy, A.M. Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **2016**, *17*, 132. [[CrossRef](#)]
29. Olm, M.; Brown, C.; Brooks, B.; Banfield, J.F. dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **2017**, *11*, 2864–2868. [[CrossRef](#)]
30. Surachat, K.; Kantachote, D.; Deachamag, P.; Wonglapsuwan, M. In silico genomic analysis of Rhodopseudomonas palustris strains revealed potential biocontrol agents and crop yield enhancers. *Biol. Control* **2022**, *176*, 105085. [[CrossRef](#)]
31. Guan, L. Structure and mechanism of membrane transporters. *Sci. Rep.* **2022**, *12*, 13248. [[CrossRef](#)]
32. Akhtar, A.A.; Turner, D.P.J. The role of bacterial ATP-binding cassette (ABC) transporters in pathogenesis and virulence: Therapeutic and vaccine potential. *Microb. Pathog.* **2022**, *171*, 105734. [[CrossRef](#)] [[PubMed](#)]
33. Pao, S.S.; Paulsen, I.T.; Saier, M.H., Jr. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 1–34. [[CrossRef](#)] [[PubMed](#)]
34. Hemme, C.; Deng, Y.; Gentry, T.; Fields, M.W.; Wu, L.; Barua, S.; Barry, K.; Tringe, S.G.; Watson, D.B.; He, Z.; et al. Metagenomic insights into evolution of a heavy metal-contaminated groundwater microbial community. *ISME J.* **2010**, *4*, 660–672. [[CrossRef](#)] [[PubMed](#)]
35. Wattam, A.R.; Brettin, T.; Davis, J.J.; Gerdes, S.; Kenyon, R.; Machi, D.; Mao, C.; Olson, R.; Overbeek, R.; Pusch, G.D.; et al. Assembly, Annotation, and Comparative Genomics in PATRIC, the All-Bacterial Bioinformatics Resource Center. *Methods Mol. Biol.* **2018**, *1704*, 79–101. [[CrossRef](#)] [[PubMed](#)]
36. Sharma, S.; Kaushik, V.; Kulshrestha, M.; Tiwari, V. Different Efflux Pump Systems in Acinetobacter baumannii and Their Role in Multidrug Resistance. *Adv. Exp. Med. Biol.* **2023**, *1370*, 155–168. [[CrossRef](#)] [[PubMed](#)]
37. Rafeeq, H.; Afsheen, N.; Rafique, S.; Arshad, A.; Intsar, M.; Hussain, A.; Bilal, M.; Iqbal, H.M. Genetically engineered microorganisms for environmental remediation. *Chemosphere* **2023**, *310*, 136751. [[CrossRef](#)] [[PubMed](#)]
38. Cho, Y.J.; Cho, A.; Hong, S.G.; Choi, H.G.; Kim, O.S. Draft Genome Sequence of Arthrobacter oryzae TNBS02, a Bacterium Containing Heavy Metal Resistance Genes, Isolated from Soil of Antarctica. *Microbiol. Resour. Announc.* **2019**, *8*, e01501-18. [[CrossRef](#)] [[PubMed](#)]
39. Lu, B.; Leong, H.W. Computational methods for predicting genomic islands in microbial genomes. *Comput. Struct. Biotechnol. J.* **2016**, *14*, 200–206. [[CrossRef](#)]
40. Langille, M.; Hsiao, W.; Brinkman, F. Detecting genomic islands using bioinformatics approaches. *Nat. Rev. Microbiol.* **2010**, *8*, 373–382. [[CrossRef](#)]
41. Waack, S.; Keller, O.; Asper, R.; Brodag, T.; Damm, C.; Fricke, W.F.; Surovcik, K.; Meinicke, P.; Merkl, R. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. *BMC Bioinform.* **2006**, *7*, 142. [[CrossRef](#)]
42. Li, L.; Meng, D.; Yin, H.; Zhang, T.; Liu, Y. Genome-resolved metagenomics provides insights into the ecological roles of the keystone taxa in heavy-metal-contaminated soils. *Front. Microbiol.* **2023**, *14*, 1203164. [[CrossRef](#)] [[PubMed](#)]

43. Lin, Y.; Wang, L.; Xu, K.; Li, K.; Ren, H. Revealing taxon-specific heavy metal-resistance mechanisms in denitrifying phosphorus removal sludge using genome-centric etaproteomics. *Microbiome* **2021**, *9*, 67. [[CrossRef](#)] [[PubMed](#)]
44. Pathak, A.; Jaswal, R.; Xu, X.; Chauhan, A. Draft Genome Sequence of *Serratia* sp. Strain S1B, Isolated from Mercury-Contaminated Soil. *Genome Announc.* **2018**, *6*, e00534-18. [[CrossRef](#)] [[PubMed](#)]

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.