Hazen Terry Clyde (Orcid ID: 0000-0002-2536-9993)
Stahl David (Orcid ID: 0000-0003-3051-841X)
Adams Michael (Orcid ID: 0000-0003-0881-8870)

Ecophysiological and genomic analyses of a representative isolate of highly abundant *Bacillus cereus* strains in contaminated subsurface sediments

Jennifer L. Goff[1], Elizabeth G. Szink[1], Michael P. Thorgersen[1], Andrew D. Putt[3], Yupeng Fan[4], Lauren M. Lui[2], Torben N. Nielsen[2], Kristopher A. Hunt[5], Jonathan P. Michael[4], Yajiao Wang[4], Daliang Ning[4], Ying Fu[4], Joy D. Van Nostrand[4], Farris L. Poole II[1], John-Marc Chandonia[2], Terry C. Hazen[3,6,7], David A. Stahl[5], Jizhong Zhou[4,8,9,10], Adam P. Arkin[2,11], and Michael W.W. Adams[1,†]

[1]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, USA
[2]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[3]Earth and Planetary Sciences, University of Tennessee, Knoxville, TN, USA
[4]Institute for Environmental Genomics, University of Oklahoma, Norman, OK, USA
[5]Civil and Environmental Engineering, University of Washington, Seattle, WA, USA
[6]Genome Sciences Division, Oak Ridge National Lab, Oak Ridge, TN, USA
[7]Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, USA
[8]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA
[9]School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK, USA
[10]Earth and Environmental Sciences, Lawrence Berkley National Laboratory, Berkeley, CA, USA
[11]Department of Bioengineering, University of California at Berkeley, Berkeley, CA, USA
[†]Address correspondence to adamsm@uga.edu

Corresponding Author:
Michael W. W. Adams
(ORCID ID: 0000-0002-9796-5014)
Department of Biochemistry & Molecular Biology
University of Georgia
Athens, GA 30602, USA
Email: adamsm@uga.edu
Phone: 706 542 2060

Running title: Mixed waste tolerance of novel *B. cereus* isolate

Key words: *Bacillus*, nitrate, metals, genomics, contamination, pangenome, adaptation

Originality-Significance Statement: We isolated a *Bacillus cereus* strain (CPTF) that represents a dominant community member of a mixed waste contaminated subsurface site as determined by 16S rRNA gene (V4 region) amplicon sequence variant profiling. Pangenomic analysis of the *B. cereus* group revealed significant genomic expansion in strain CPTF that was largely attributable to plasmid acquisition and transposable element mobilization. We further

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1462-2920.16173

provide evidence for its adaptation to the unique geochemistry of this contaminated site through integrated pangenomic and physiological analyses.

SUMMARY

Bacillus cereus strain CPT56D-587-MTF (CPTF) was isolated from the highly contaminated Oak Ridge Reservation (ORR) subsurface. This site is contaminated with high levels of nitric acid and multiple heavy metals. Amplicon sequencing of the 16S rRNA genes (V4 region) in sediment from this area revealed an amplicon sequence variant (ASV) with 100% identity to the CPTF 16S rRNA sequence. Notably, this CPTF-matching ASV had the highest relative abundance in this community survey, with a median relative abundance of 3.77% and comrprised 20-40% of reads in some samples. Pangenomic analysis revealed that strain CPTF has expanded genomic content compared to other *B. cereus* species—largely due to plasmid acquisition and expansion of transposable elements. This suggests that these features are important for rapid adaptation to native environmental stressors. We connected genotype to phenotype in the context of the unique geochemistry of the site. These analyses revealed that certain genes (e.g. nitrate reductase, heavy metal efflux pumps) that allow this strain to successfully occupy the geochemically heterogenous microniches of its native site are characteristic of the *B. cereus* species while others such as acid tolerance are mobile genetic element associated and are generally unique to strain CPTF.

INTRODUCTION

Industrial and agricultural activities have created many sites globally that are contaminated by multiple waste types including heavy metals, radionuclides, nitrogen oxides, and volatile organic compounds (Riley and Zachara, 1992, Baeseman *et al.*, 2006, Keskin, 2010, Abdelouas *et al.*, 1998, Boguslavsky *et al.*, 2020). Co-contaminants have the potential to influence the biogeochemical processes that are critical for transformation and removal of other components of the contamination, making the management of these sites a challenge (Gumaelius *et al.*, 1996, Ge *et al.*, 2019, Thorgersen *et al.*, 2015, Thorgersen *et al.*, 2017a, Spain and Krumholz, 2011, Arjoon *et al.*, 2013, Duxbury, 1985). Frequently, heavy metals and nitrate co-exist at sites of industrial waste discharge (Riley and Zachara, 1992, Baeseman *et al.*, 2006, Moon *et al.*, 2020). While nitrate can be removed naturally by resident denitrifying (Foglar *et al.*, 2005) and ammonifying (Burgin and Hamilton, 2007) microbial populations, metal co-contaminants can potentially inhibit nitrate removal either through direct inhibition of enzymes in the pathway (Sato *et al.*, 1998, Thorgersen *et al.*, 2017a) or through general inhibition of microbial growth

(Thorgersen *et al.*, 2019, Lu *et al.*, 2019, Bollag and Barabasz, 1979). Thus, determining how heavy metals impact the physiology of field-relevant, nitrate-respiring isolates is essential to modeling the impact of co-contaminants on nitrogen cycling at field sites (Lui *et al.*, 2021).

The US Department of Energy (DOE) Oak Ridge Reservation (ORR) in Oak Ridge, Tennessee contains a subsurface site contaminated by nitric acid and a mixture of heavy metals (Spain and Krumholz, 2011). This contamination is the result of discharge of liquid wastes from uranium processing operations at the on-site Y-12 National Security Complex into waste ponds (referred to as the S-3 ponds) between 1951 and 1983 (Moon *et al.*, 2020). Waste disposal at this site ceased nearly four decades ago, and the S-3 ponds have subsequently been drained, capped, and covered by a parking lot. However, there are still five large contamination plumes of groundwater extending from the site of the former S-3 ponds (Revil *et al.*, 2013). As a result, the near-source groundwater is highly acidic and contains high levels of uranium (U) and nitrate, as well as other metal contaminants, such as nickel (Ni), cadmium (Cd), copper (Cu), aluminum (Al), manganese (Mn), and iron (Fe) (Smith *et al.*, 2015, Moon *et al.*, 2020). Prior studies of the ORR groundwater have found that this mixture of heavy metals and nitric acid wastes are major drivers of community assemblage and genetic potential, suggesting that these co-contaminants are significant controllers of ecological processes occurring at the site (Hemme *et al.*, 2010, Carlson *et al.*, 2019, Moreels *et al.*, 2008). Metagenome sequencing of the highly contaminated ORR groundwater revealed an enrichment of genes involved in nitrogen metabolism and heavy metal efflux pumps (Hemme *et al.*, 2010). These stressors have also promoted rapid adaptive evolution of microorganisms living at the site via horizontal gene transfer and gene duplication events (Casacuberta and González, 2013, Martinez *et al.*, 2006, Liang *et al.*, 2011, Coombs and Barkay, 2004).

Here, we present the isolation and characterization of a nitrate-respiring *Bacillus cereus* strain that represents a dominant species within the ORR subsurface sediments as determined by 16S rRNA gene (V4 region) community analysis. This contrasts with many prior studies that have focused primarily on isolates enriched from groundwater pumped from wells at the site rather than those isolated directly from the subsurface sediments (Fields *et al.*, 2005, Thorgersen *et al.*, 2019, Prakash *et al.*, 2021). The objective of this study was to examine the genomic features of this sediment-associated strain that facilitate its persistence at this mixed-waste subsurface site. We performed a pangenomic analysis to explore the genome features predicted to mediate strain resilience at the site within the broader context of the *B. cereus* species. To connect genome observations to relevant resistance phenotypes, we used site-

informed geochemical parameters to conduct laboratory growth experiments to explore the viability of this isolate during exposure to low pH, elevated concentrations of nitrogen oxides, and mixed metals. This allowed us to constrain the environmental conditions that may be permissive to its activity in the ORR subsurface.

RESULTS AND DISCUSSION

Phylogenetic and geospatial characterization of *B. cereus* strain CPTF.

Strain CPTF was isolated from a sediment sample collected from the saturated zone of the southeastern region of ORR Area 3 (long. -84.27335º, lat. 35.977268º, depth 535.94 cm) immediately adjacent to the contamination source, the former S-3 ponds. Whole-genome comparison using the Genome Taxonomy Database Toolkit (Chaumeil *et al.*, 2020) identified CPTF as a *Bacillus cereus* species with an average nucleotide identity (ANI) to the reference strain ATCC 14579 of 98.58%. Intriguingly, the strain CPTF 16S rRNA gene V4 region has 100% sequence identity to the most abundant amplicon sequence variant (ASV) observed in our 16S rRNA gene (V4 region) survey of sediment samples collected from the Area 3 subsurface **(Fig. S1, Fig. S2, Fig. 1).** The relative abundance of this CPTF-matching ASV in the samples range from 0.02 – 40.52% with a median relative abundance of 3.77% (**Fig. 1).** Importantly, the relative abundance of this ASV exceeds 5% of the total reads in 12/32 samples. Thus, strain CPTF is a member of a dominant species at the site (Callahan *et al.*, 2017). Further discussion of the distribution of strain CPTF can be found in the **Supporting Information.**

Pangenome analysis reveals significant acquisition and expansion of mobile genetic elements in strain CPTF.

Pangenomic structure. We used the complete genome sequence of strain CPTF (Goff *et al.*, 2022b) to provide insight into the adaptations contributing to the resilience of this strain at the site. For comparison, we assembled a reference dataset of 41 finished *B. cereus* genomes in addition to the strain CPTF genome (42 genomes in total) **(Fig. 2A, Table S1)**. It is important to note that the taxonomic classification of *B. cereus* species and its close relatives (e.g. *B. anthracis* and *B. thuringiensis*) is controversial (Fayad *et al.*, 2019). It is challenging to delineate these closely related species—although several recent studies have successfully done so by incorporating mobile genetic elements in their analyses (Liu *et al.*, 2015, Felten *et al.*, 2017). For this study, we chose to analyze completed genomes annotated as "*B. cereus*" in the Integrated Microbial Genomes & Microbiomes (IMG/M) database. Our reference dataset includes strains

isolated from both environmental and clinical sources. At 6.5 Mbp in length, the strain CPTF genome is substantially larger than the average *B. cereus* genome in our reference dataset (5.6 Mbp) **(Fig. 2B)**. The second largest genome in our dataset was 6.1 Mbp in length. Genomic analyses suggested that this size difference is largely attributable to the increased mobile genetic element (MGE) content of strain CPTF: specifically plasmids and transposable elements including insertion sequence (IS) elements and Tn3 family transposons. In our *B. cereus* reference dataset, the average genome contains two plasmids (range: 0-7) **(Fig. 2B, Table S1)**. By comparison, the CPTF genome contains eight plasmids (designated pCPTF 1-8) of varying size and transferability **(Table S2)**.The plasmids themselves comprise 0.879 Mbp, or ~13% of the genome length. This increased plasmid content is notable as horizontal gene transfer can mediate the acquisition of metal resistance determinants at contaminated sites allowing for rapid adaptation to fluctuating environmental stressors (Hemme *et al.*, 2016, Coombs and Barkay, 2004, Martinez *et al.*, 2006).

The strain CPTF genome also contains 294 complete and 180 partial IS elements, respectively **(Table S3)**. In contrast, the average genome in the *B. cereus* reference dataset contains 30 (range: 10-63) and 13 (range: 1-47), respectively **(Fig. 2C, Table S1)**. The IS elements are distributed across the CPTF chromosome and six of the eight plasmids (pCPTF 1-6) (**Fig. S3).** Approximately, 4.8% of the CPTF chromosomal genetic material consists of IS elements, both complete and partial. Typically, the density of IS elements in bacterial chromosomes is less than 3% except in extreme instances of IS expansion (Siguier *et al.*, 2006). Indeed, the average chromosomal density of IS elements in the *B. cereus* reference dataset was only 0.63% (range: 0.29 – 1.14%) **(Fig. 2C, Table S4)**. The high IS element content of the strain CPTF genome is suggestive of IS expansion events. IS expansion is commonly observed in bacteria that have recently adopted host-restricted lifestyles as transposition events cause gene disruptions and other genomic rearrangements which lead to gene inactivation, gene elimination, and ultimately genomic streamlining (Siguier *et al.*, 2014). As such, the occurrence of IS expansion within a free-living environmental isolate is puzzling. However, the hypermutability of a genome undergoing active transposition may provide a selective advantage under certain conditions. IS elements can modulate genomic content and gene expression in several ways including gene disruption, insertion of upstream promoter sequences, mobilization of neighboring genes in the form of composite transposons, and promotion of homologous recombination (Vandecraen *et al.*, 2017). Active transposition has been observed under various

stress conditions such as oxidative stress (Drevinek *et al.*, 2010) heavy metals, (Mijnendonckx *et al.*, 2011), and radiation (Pasternak *et al.*, 2010).

Pangenome gene content. To contextualize the gene content of strain CPTF, we computed the pangenome of the *B. cereus* reference dataset plus strain CPTF (42 genomes in total). The pangenome contained 16,254 unique orthologous gene clusters **(Table S5)**. These gene clusters were further categorized as core (present in 100% of genomes), soft-core (present in 95-99% of genomes), shell (present in 16-94% of genomes), or cloud (present in 1-15% of genomes). The *B. cereus* pangenome contains 3,577 core gene clusters (22% of the pangenome), 275 soft-core gene clusters (2%), 2,400 shell gene clusters (15%), and 10,003 cloud gene clusters (62%). Generally, the core genome includes all genes necessary for the basic biological functioning of a species as well as other genes that encode functions characteristic of the species. By contrast, the cloud genome includes genes typically associated with mobile genetic elements that are non-essential for growth but that confer selective advantages under different conditions, allowing for adaptation to new or fluctuating environmental niches (Medini *et al.*, 2005). The ORR Area 3 sediments and associated porewater are unique in that multiple physiological stressors exist simultaneously: namely, high concentrations of nitrate, a mixture of heavy metals, and low pH values (Smith *et al.*, 2015, Moon *et al.*, 2020). We sought to connect phenotypic features of strain CPTF that are crucial for survival at this contaminated site to either the core *B. cereus* genome or the mobile genetic element-associated cloud genome.

Strain CPTF respires nitrate and has high nitrogen oxide resistance.

The sediment samples collected for the community analysis described above were also meausured for pH and associated porewater nitrate concentrations. Within the porewaters of the Area 3 sediments, nitrate concentrations range from 0.007 to 609 mM with a median of 2.1 mM. Notably, the CPTF-matching ASV identified in the sediment portion of these samples is observed across the full range of associated porewater nitrate concentrations **(Fig. 3A)**. We monitored the growth of strain CPTF with a range of nitrate concentrations under anoxic conditions to determine its level of nitrate resistance. We found that strain CPTF has a high level of nitrate resistance with an IC_{50} value of 579 mM **(Fig. 3A)**. Strain CPTF also has high tolerance to nitrite with an IC_{50} of 139 mM **(Fig. 3A).** Generally, nitrite is highly toxic to many environmental microorganisms even at relatively low concentrations (i.e. < 5 mM) (Spain and Krumholz, 2011). While high concentrations of nitrite are not routinely detected at the ORR,

nitrite can transiently spike to levels >100 mM during massive microbial blooms, such as those that occurred during prior biostimulation efforts (Spain and Krumholz, 2011). Nitrite levels could also increase through a decoupling of nitrite production and reduction on a more local scale (Philips *et al.*, 2002).

The genome of strain CPTF encodes two enzyme complexes involved in nitrogen oxide reduction, including one copy of the *nar* operon encoding the NarGHI nitrate reductase **(Fig. 3B)**. Across the *B. cereus* pangenome, *narGHI* were near-ubiquitous: present in 41/42, 40/42, and 40/42 analyzed genomes, respectively, suggesting that nitrate respiration is a fundamental phenotypic feature of the species Interestingly, prior metagenomic studies at the ORR Area 3 site have found that genes involved in respiring nitrogen oxides, including respiratory nitrate reductases, are enriched in the native microbial community relative to all sequenced microbial genomes (Hemme *et al.*, 2010). This suggests that the nitrate contamination at the site has imposed a strong selective pressure on the microbial community, favoring the proliferation of nitrate-respiring organisms like strain CPTF. In line with our genomic observations, we found that strain CPTF grows by nitrate respiration at a concentration (10 mM nitrate) that is close to the median (~2 mM nitrate) at the site **(Fig. 3C)**. With glucose (20 mM) as the carbon source, nitrite and ammonia are the major products of nitrate respiration with a small fraction of nitrous oxide (0.26 ± 0.02 mM) detected at an end-point measurement, representing less than 5% of the total inorganic nitrogen. The end-point nitrate, ammonium, nitrite, and nitrous oxide can fully account for the total inorganic N added to the system (N balanced to 107.2 ± 1.5%) **(Fig. 3C)**.

Immediately adjacent to the *nar* operon are the genes encoding cytoplasmic NADH-dependent siroheme nitrite reductase NasDE (also known as NirBD in other organisms) (Sun *et al.*, 2016) **(Fig. 3B)**. Similar to the nitrate reductase genes, *nasDE* are found in the soft core of the *B. cereus* pangenome with both present in 40/42 genomes. In other *Bacillus* species, the NasDE nitrite reductase is involved in both assimilatory and fermentative dissimilatory nitrite reduction to ammonium (Nakano *et al.*, 1998, Sun *et al.*, 2016). Consistent with these genomic observations, we found that CPTF consumes nitrite during anoxic fermentative growth with glucose as a carbon source **(Fig. S4A,B)**. Addition of nitrite to cultures enhanced growth on glucose compared to cultures grown with glucose alone **(Fig. S4B)**. In further agreement, CPTF was unable to grow under anoxic conditions with nitrite using non-fermentable carbon sources **(Fig. S4B)**.

Environmental factors control pH tolerance of strain CPTF.

The CPTF-matching ASV is observed in the contaminated ORR subsurface environment across a broad range of sediment pH values (3.1 to 7.1) **(Fig. 4A)**. However, when grown in batch culture with glucose (20 mM) under anoxic nitrate-respiring (10 mM nitrate) conditions, we only observed robust growth at pH values of 6.0 and 7.0 with slightly higher growth observed at pH 7.0 **(Fig. 4B)**. Substantially lower growth was observed at pH 5.0. We predicted that the accumulation of nitrite during nitrate respiration in our batch cultures increased the sensitivity of the strain to lower pH due to the increased proportion of the protonophore nitrous acid (Vadivelu *et al.*, 2006). When CPTF was grown under oxygenated conditions without added nitrate, robust growth was observed down to pH 5.0 and the pH optimum shifted to 6.0 **(Fig. 4B).** Addition of nitrate or nitrite (7 mM, the typical amount accumulated during nitrate respiration) to the cultures under these same oxygenated conditions increased the sensitivity of the cells to the lower pH and growth was significantly diminished at pH 5.0. Thus, local accumulation of nitrite in the environment may increase the sensitivity of the strain to more acidic growth conditions. This could occur, for example, during massive blooms of nitrate-respiring organisms at the site (Spain and Krumholz, 2011).

Even under conditions where nitrate/nitrite were not added (i.e. under aerobic growth conditions) no growth was observed by 36 hours at either pH 3.0 or 4.0. However, we were able to recover viable cells up to 120 hours by transferring from these low pH cultures to LB broth at various time points post-inoculation. Further recovery was not monitored post-120 hours. These results suggest that, in the field, the cells may persist in a slow-growing state at these lower recorded pH values. Alternatively, microheterogeneity of the subsurface sediment may result in pockets of higher pHs values that are not reflected by bulk pH measurements (Sumner, 1994). The strain CPTF genome encodes several proteins that may facilitate surface attachment and stasis within a sediment microniche conducive towards its growth including flagellar proteins and surface adhesins. Nonetheless, there is also suggestion of genetically-encoded mechanisms of acid tolerance in the CPTF genome. For example, the pCPTF1 megaplasmid carries a copy of the *ureDABCEFG* urease operon which may play a role in pH tolerance. Interestingly, the complete *ureDABCEFG* operon was only observed in 2 out of the 42 analyzed *B. cereus* genomes including strain CPTF. Both gram-negative and gram-positive human GI tract commensal microorganisms rely on urease activity to buffer against acidity in their local environments (Merrell and Camilli, 2002, Huang *et al.*, 2014). As urea is ubiquitous in natural environments, urease may be an adaptive feature for pH tolerance in subsurface sediment environments (Wang *et al.*, 2008). Other possible features of the acid-tolerance response

observed in the strain CPTF genome include an arginine decarboxylase (a core genome feature) and an arginine deiminase (observed in 29/42 genomes) (Carlin *et al.*, 2010).

Strain CPTF is adapted to growth with mixed-metals contaminants.

We hypothesized that strain CPTF would have extensive metal homeostasis systems to protect against the multiple metals stress of its native environment (Hemme *et al.*, 2010). Comparison of the CPTF genome against the BacMet antibacterial, biocide, and metal resistance genes database (Pal *et al.*, 2014) revealed 138 putative metal homeostasis determinants, including those annotated for Al (1 genetic determinant), Mn (8), Fe (40), Cd (8), Cu (18), Co (14), and Ni (2) homeostasis (**Table S6)**. These seven metals are all major components of the ORR contaminant plume (Thorgersen *et al.*, 2015, Smith *et al.*, 2015) **(Table S7)**. The remaining determinants are annotated as being involved in homeostasis of other metals/metalloids that are not significant components of the ORR contamination plume such as As, Sb, Se, Te, and Hg.

Several of these genetic determinants were ubiquitous among *B. cereus* species and clustered into three regions of the CPTF chromosome. The first region contained three copies of the gene encoding the tellurite resistance determinant TerD (all three copies are present in 42/42 *B. cereus* genomes). TerD has been suggested to have functions relating to general stress responses and tellurite resistance is unlikely to be the main function of these proteins (Daigle *et al.*, 2015). The *terD* genes clustered with a cation-transporting ATPase (42/42 genomes) and a general stress response protein (42/42 genomes). The second region included the *feuABC* genes which are involved in the uptake of the siderophore bacillibactin (42/42 genomes).These genes were found immediately downstream from a gene encoding a ferrichrome transport protein (42/42 genomes) and a trilactone hydrolase (42/42 genomes) which releases iron from bacillibactin in the cytosol. Maintenance of iron homeostasis appears to be important in the cellular response to divalent heavy metal cation stress, e.g. Co^{2+},Cu^{2+},Cd^{2+}, Ni^{2+} (Chandrangsu *et al.*, 2017). Included in this region was a gene encoding a metal efflux P-type ATPase (42/42 genomes). P-type ATPases are involved in the membrane transport of metal cations commonly found in the ORR subsurface (e.g. Cd^{2+}, Cu^{+}) (Nies, 2003). Also present was a CsoR repressor (42/42 genomes) which regulates activity of CopZA, a Cu^{2+}-transport system. The third region includes a gene encoding a second metal efflux P-type ATPase (42/42 genomes) and a gene encoding a Cd^{2+} efflux accessory protein (42/42 genomes). Thus, *B. cereus* species appear to contain a core set of heavy metal resistance genes that may allow for adaptation to fluctuating soil environments.

By contrast, the eight plasmids of strain CPTF contained a number of metal resistance determinants that are largely unique to strain CPTF. The pCPTF1 megaplasmid (4.5 Mbp) contains two metal-responsive ArsR family transcriptional regulators (both present in 1/42 genomes; unique to strain CPTF) and one metal-responsive MerR family transcriptional regulator (4/42 genomes). Characterized members of the ArsR and MerR families regulate cellular responses to various metal cations including those found in the ORR subsurface such as Ni^{2+}, Co^{2+}, and Cd^{2+} (Busenlehner *et al.*, 2003, Brown *et al.*, 2003). Also present on pCPTF1 is an RND family cation/drug efflux pump (1/42 genomes; unique to strain CPTF) (Du *et al.*, 2018); a TerC-like integral membrane protein that has been shown to mediate Mn resistance in *B. subtilis* (Paruthiyil *et al.*, 2020) (unique to strain CPTF genome); and a NixA Ni/Co transporter (3/42 genomes) (Dosanjh and Michel, 2006). Among the remaining seven smaller plasmids, heavy metal resistance-related genes included ArsR family transcriptional regulators (3/7 plasmids) and MerR family transcriptional regulators (2/7 plasmids), all unique to the strain CPTF genome. Gene regulation evolves rapidly and new transcription factors will often arise via gene duplication events that can result in novel regulatory interactions (Price *et al.*, 2007, Teichmann and Babu, 2004). A larger repertoire of transcription factors may be an adaptive feature that allows microorganisms living in high-stress and fluctuating environments to integrate a larger number of environmental signals (Cases *et al.*, 2003). An interesting example of a suspected gene duplication event in the strain CPTF genome is the five plasmid-borne ArsR homologs.These plasmid ArsR regulators form a phylogenetically distinct cluster from their eleven chromosomal ArsR homologs **(Fig. S5A)**. We observed cryptic IS element scars on one or both flanking regions of these plasmid-localized sequences in addition to preserved neighboring gene synteny. **(Fig S5B)**. We propose that this metal-responsive regulatory element has undergone multiple IS element-mediated duplication events across the plasmids.

We next examined if this heavy metal resistance gene content was reflected in a broad heavy metal resistance phenotype for strain CPTF. We used previously-published measurements of metal concentrations in the nearby groundwater from Area 3 wells for determination of the metal resistance of strain CPTF (Thorgersen *et al.*, 2015, Thorgersen *et al.*, 2019, Smith *et al.*, 2015) as the sediments of the Area 3 subsurface are frequently saturated by this groundwater (Moon *et al.*, 2020). Consistent with our genomic observations, strain CPTF was resistant to site-relevant concentrations of Cd (up to 10 μM), Cu (up to 20 μM), Al (up to 2000 μM), Mn (up to 200 μM), and Fe (up to 20 μM) **(Table S7, Fig. S6)**. While Ni and Co inhibited the growth of strain CPTF at higher site-relevant concentrations of 200 and 60 μM,

respectively, these cultures still produced measurable growth. Additionally, while no U resistance determinants were identified in the strain CPTF genome due to their absence in the BacMet database, we observed resistance to site relevant concentrations of U (up to 200 μM). This is likely because a major mechanism of U resistance in bacteria is adsorption to the cell surface (Pinel-Cabello *et al.*, 2021, Fowle *et al.*, 2000, Thorgersen *et al.*, 2017b). Finally, we quantified strain CPTF resistance to a site-relevant mixture of metals. Previously, we developed a contaminated ORR environmental metal mixture (COMM) designed to reflect the average concentrations of the major metal contaminants of the Area 3 groundwater (**Table S7**) (Thorgersen *et al.*, 2019). We monitored the growth of strain CPTF with glucose (20 mM) as a carbon source under nitrate-respiring conditions with dilutions of the COMM **(Fig. 5A,B)**. We calculated an IC_{50} for the mixture of 0.84X **(**see **Fig. 5B** for individual metal concentrations**)**. Notably, growth was still observed even at the 1X concentration. Interestingly, we also observed that the COMM inhibited nitrite reduction in the cultures **(Fig. S7)**.This is in line with prior studies that show the divalent metal cations Cd^{2+}, Cu^{2+}, Ni^{2+}, and Zn^{2+} cause an accumulation of nitrite in cultures of denitrifying microorganisms (Gui *et al.*, 2017, Ramírez *et al.*, 2018).

CONCLUDING REMARKS

Our results suggest that core *B. cereus* genes and mobilomic flexibility have allowed for strain CPTF to successfully occupy the heterogenous, contaminated microniches of the ORR Area 3 subsurface sediments. There are two major caveats to the conclusions presented here. First, it will always be challenging to extrapolate laboratory observations of metal toxicity to the field (Duxbury, 1985). Numerous factors that control metal speciation and solubility—including pH, inorganic ions, and carbon compounds—are difficult to match between field and laboratory conditions (Duxbury, 1985, Gadd, 2010). This challenge is further amplified within a solid matrix, as particle surface sites will adsorb metals (Bradl, 2004). However, we partially remedy this issue here by assessing metal toxicity in the presence of two of the major categories of co-contaminants: (1) nitrate and (2) other metals. Second, the 16S rRNA gene V4 region does not allow for strain-level resolution (Edgar, 2018). Thus, the CPTF-matching ASV may represent a collection of closely related *B. cereus* strains. We suggest, though, that these strains will share the same core set genes identified in this study as well as many of the same cloud genes due to their close physical proximity.

While it is apparent that mobile genetic elements have contributed to the resilience of strain CPTF at this site, an intriguing question that lingers is whether the environmental

stressors promoted horizontal gene transfer and IS expansion in the strain CPTF genome? Or, did the waste contaminants at ORR select for a native organism whose mobilome had already undergone extensive expansion prompted by an unrelated cause? How genome evolution, particularly the mobility of certain genetic elements within microbial populations, facilitates resilience in the face of environmental stressors is a critical topic of research when studying microbial processes at contaminated sites.

MATERIALS AND METHODS

Field sampling and geochemical measurements. Sediment samples were collected from Area 3 of the Oak Ridge Reservation (ORR) Oak Ridge, TN which lies adjacent (southwest) to the former S-3 waste-disposal pond (Moon *et al.*, 2020). Sediment samples were collected from the target depth using a 1.5" diameter sediment sampler. Once at the surface, the sample casing was removed and samples were aseptically collected from the inner sections of the core using a sterile plastic scraper and stored in sterile 50 mL conical tubes at 4˚C. Samples for 16S rRNA gene (V4 region) sequencing were stored at -80°C before use. Extractable porewater was separated from 10 g of sediment in an Amicon Ultra 100 K MWCO (Merck Millipore) 50mL filtering device centrifuged at 3,000 x g at 4°C for 1 hour. Porewater nitrate was determined from a working range of 0.1-200 mg/L using an electrical conductivity detector on a Dionex™ ICS-5000+ series system (ThermoFisher Scientific, Waltham, MA) with an IonPac AS11HC analytical and AS11 guard column at 35˚C with a KOH effluent gradient of 0 – 60 mM at 1.3 ml/min (Moon *et al.*, 2020). Soil pH was determined using a 1:1 soil to deionized water saturation method (EPA, 2004) and measured with stirring on an Oakton® pHTestr® 50S double junction pH electrode.

16S rRNA gene (V4 region) amplicon sequencing. Microbial genomic DNA was extracted by a freeze-grinding method as described previously (Zhou *et al.*, 1996) and purified by using DNeasy PowerSoil DNA isolation Kit (QIAGEN; bead tubes were not used). DNA quality was checked by NanoDrop, and double-strand DNA concentration was measured by Quant-iT™ PicoGreen™ dsDNA Assay Kit. A two-step PCR was used for library preparation as described previously (Wu *et al.*, 2015). In the first step, the standard primers were used to amplify the V4 region (253 bp in length) of prokaryotic 16S rRNA genes (515F [5'-GTGCCAGCMGCCGCGGTAA-3'] and 806R [5'-GGACTACHVGGGTWTCTAAT-3']). In the

second step, phasing primers were designed and used to increase the base diversity in sequences of sample libraries. The PCR was performed for 11 cycles in the first and 22 cycles in the second step. Sample libraries were then sequenced by MiSeq platform (Illumina, San Diego, CA, USA) as described previously (Caporaso *et al.*, 2012). QIIME2 (v2021.2) (Bolyen *et al.*, 2019) was used to process the sequencing data. After barcode and primer sequences were trimmed with zero error, exact amplicon sequence variants (ASV) were identified by DADA2 (Callahan *et al.*, 2016). The method 'consensus' was used to remove chimeras; samples are denoised independently; forward and reverse truncate positions are 163 and 130, respectively, after quality trim with a quality score 2; maximal expected error score is 2.0 when combining forward and reverse sequences. Taxonomy was assigned to ASVs using QIIME2 pre-fitted sklearn-based taxonomy classifier (Bokulich *et al.*, 2018, Pedregosa *et al.*, 2011) against the Silva 138 99% OTUs reference sequences (Quast *et al.*, 2013), at the default confidence level (0.7). After classification, ASVs identified as Chloroplast, Mitochondria, Eukaryota, or unclassified kingdom were removed.

The full-length 16S rRNA gene sequence was retrieved from the strain CPTF genome (Goff *et al.*, 2022b)**.** The strain CPTF 16S rRNA gene V4 region sequence was compared to the ORR sediment 16S rRNA gene V4 region amplicon sequence variants (ASV) using a BLASTn search with default parameters. The relative abundance of the CPTF-matching ASV (ID: bdf8a26094624622d68509a87fa75ba7) across the different samples was calculated using the number of ASV reads per sample divided by the total number of reads per sample. To rank the overall relative abundance of the strain CPTF-matching ASV, the total reads across all samples were determined for each ASV in the dataset. The strain CPTF-matching ASV was found to have the greatest number of total reads across the entire subsurface sediment survey. The complete ASV dataset with associated IDs is available via a published KBase narrative (Goff *et al.*, 2022a)

Media and culture conditions. *Bacillus cereus* str. CPT56D-587-MTF (referred to as strain CPTF) was isolated as described previously in Goff *et al.* (2022b)*.* CPTF was routinely streaked out on LB plates and grown overnight at 30 ºC. A single isolated colony was inoculated into LB broth and grown overnight shaking at 200 rpm at 30 ºC. The overnight culture was diluted 100-fold into a defined medium for further experimentation. Unless otherwise specified, the growth medium contained, per liter, 2.5 g $NaHCO_3$, 0.6 g NaH_2PO_4, 0.5 g casamino acids, and 40 mL of a 25X stock solution of salts composed of 5 g/L NH_4Cl, 2.5 g/L KCl, 12.32 g/L $MgSO_4 \cdot 7H_2O$, 0.367 g/L $CaCl_2 \cdot 2\ H_2O$, and 0.25 g/L NaCl. Additionally, the medium contained, per liter, 1 mL of

1000X vitamin mix and 10 mL of a 100X DL minerals stock solution, prepared as described previously by Widdel and Bak (1992). The medium was routinely supplemented with 20 mM glucose, and 10 mM $NaNO_3$ as a carbon source and terminal electron acceptor, respectively. The pH was adjusted to 7.2 using either HCl or NaOH. For anaerobic cultivation, the medium was purged of oxygen and a bottle headspace of 80%/20% N_2/CO_2 was used. The cultures were incubated at 30ºC. For aerobic cultivation, cultures were shaken at 30ºC at 200 rpm.

To determine the pH range permissive to strain CPTF growth, a modified medium from the one described above was used to allow for buffering from pH 3 to 8 using a 50 mM citrate phosphate buffer. This modified medium contained, per liter, 11.5 g citric acid monohydrate, 7.1 g Na_2HPO_4, 0.5 g casamino acids, 40 mL of the 25X stock solution described above, 1 mL of the1000X vitamin mix described above, and 10 mL of the 100X DL minerals stock solution described above. For anaerobic cultivation, the cultures were amended with 20 mM glucose and 10 mM nitrate or 7 mM nitrite. The concentration of nitrite was selected based on the amount that typically accumulates in nitrate-respiring cultures. For aerobic cultivation, 20 mM glucose was added with additional amendments of 10 mM nitrate or 7 mM nitrite. We confirmed that this medium produced near-identical CPTF growth at pH 7.2 as the standard medium described previously. Likewise, in the absence of glucose, only minimal growth was observed. To adjust the pH of the medium, concentrated HCl and NaOH pellets were used.

To determine the growth of strain CPTF in the presence of site-relevant concentrations of metals, a contaminated ORR environmental metal mix (COMM) was used (Thorgersen *et al.*, 2019). The 1X COMM contains 1000 µM $AlK(SO_4)_2 \cdot 12H_2O$, 100 µM $UO_2(CH_3COO)_2$, 100 µM $MnCl_2 \cdot 4H_2O$, 100 µM $NiSO_4 \cdot 6H_2O$, 30 µM $CoCl_2 \cdot 6\ H_2O$, 10 µM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, 10 µM $CuCl_2 \cdot 2H_2O$, and 5 µM $Cd(CH_3COO)_2 \cdot 2H_2O$. The Al and U were added separately to the medium as filter-sterilized stock solutions. The remaining metals were added using a 100X filter-sterilized stock. The same metal salts at various concentrations were used for experiments with individual metals additions.

Growth curves. To monitor utilization and production of nitrogen oxides, cultures were grown in Balch tubes with 10 mL of the standard medium or modified medium for pH experiments with various amendments. Growth was monitored spectrophotometrically at 600 nm (OD_{600}). Samples were collected regularly for nitrate and nitrite determination as described below. The headspace was sampled directly for nitrous oxide measurements as detailed below.

High-throughput growth curves were conducted using a Bioscreen C automated microbiology growth curve analysis system (Growth Curves Ab Ltd.,Turku, Finland) contained in a hard glove box with a 95%/5% Ar/H_2 atmosphere (Plas Labs, Lansing, MI). Cultures (400 µL) were grown in 100-well honeycomb plates using the standard defined growth medium defined above with various amendments. The growth curves were conducted at 30ºC with continuous agitation to avoid settling at the bottom of the plate. OD_{600} was measured every 30 minutes. Cultures growing under these standard growth conditions reached early stationary phase after approximately 10 hours, the time point was selected for comparison of the extent of growth between cultures.

Nitrogen compound measurements during cell growth. For the N mass balance shown in Figure 2B, nitrate and nitrite were quantified using a Dionex 5000+ (Thermo Scientific, Carlsbad, CA) equipped with a single pump, a AS11-HC column, ADRS 600 4 mm suppressor, EGC 500 KOH cartridge, and a gradient method as described previously (Moon *et al.*, 2020). Ammonia was determined using the Ammonia Dichloroisocyanurate (DIC) assay in a Gallery (Thermo Fisher Scientific Oy, Vantaa, Finland). Nitrous oxide was quantified by direct headspace injection into a gas chromatograph (SRI, Torrance, CA) equipped with an electron capture detector and a 6 in Haysep D column with N_2 as the carrier gas operating isothermally at 70°C. For all other nitrate and nitrite measurements, supernatants were collected for colorimetric analysis using the Griess assay (Miranda *et al.*, 2001). Nitrite was measured directly at 540 nm. Nitrate was first reduced to nitrite using 8 g/L VCl_3. Nitrate concentrations were calculated from the difference between [nitrate + nitrite] and [nitrite].

Genomic analysis.

Genome sequencing and phylogeny. The strain CPTF genome was previously sequenced using a combination of long-read nanopore technology and Illumina short reads to obtain a high quality completed genome (Goff *et al.*, 2022b). Phylogeny was determined by average nucleotide identity (ANI) using the “Classify Microbes with GTDB-Tk”. (v1.7.0) application (Chaumeil *et al.*, 2020, Arkin *et al.*, 2018) implemented in the Department of Energy (DOE) KnowledgeBase (KBase).

Pangenomic analyses. To create a *B. cereus* pangenome. The Joint Genome Institution (JGI) Integrated Microbial Genomes & Microbiomes (IMG/M) database was searched for high-quality, finished *B. cereus* genomes. A total of 41 genomes were identified that met these criteria.

These genomes were selected for further analyses. Plasmid counts, assembly lengths, and isolation sources were extracted from the metadata downloaded from IMG/M. Genomes were reannotated in KBase using the "Annotate Multiple Microbial Assemblies with RASTtk" application (v1.073) using the default methods (Brettin *et al.*, 2015, Aziz *et al.*, 2008, Overbeek *et al.*, 2014). The pangenome was computed in KBase using the "Compute Pangenome" (v0.0.7) application with default parameters. Insertion sequence (IS) elements were identified in the genome set using ISSaga with default parameters (Varani *et al.*, 2011). The annotation tables for the main chromosome for all of the genomes were used to calculate the chromosomal density of IS elements. Chromosomal density was calculated by summing the length (in bp) of all chromosomal IS elements (both partial and complete) and dividing by the total length (in bp) of the chromosome.

Phylogenetic tree construction. A multi-locus species tree was constructed using the strain CPTF genome plus the 41 *B. cereus* genomes used for the pangenome analysis. Automated Multi-Locus Species Tree (autoMLST) was used for the concatenated alignment and tree construction (Alanjary *et al.*, 2019). Alignment and tree construction was performed in denovo mode with default parameters except for the addition of an IQ-Tree Ultrafast Bootstap analysis (1000 replicates). The resulting concatenated alignment used 80 conserved genes given in **Table S8.** The final tree was visualized using Interactive Tree of Life v6 (iTOL) (Letunic and Bork, 2021).

Strain CPTF genome analysis. The ACLAME database v0.4 (Leplae *et al.*, 2010), and oriTfinder v1.1 (Li *et al.*, 2018) in comibation with manual review of annotations were used to identify plasmid features involved in mobilization and conjugal transfer in the strain CPTF genome. Plasmids were categorized as "conjugative", "mobilizable", and "non-mobilizable" based on a classification system described by Douarre et. al. (Douarre *et al.*, 2020). The CPTF genome was analyzed for heavy metal resistance genes using Geneious Prime (v2021.2.2) to perform a BLASTp search against the BacMet Antibacterial Biocide & Metal Resistance Genes Predicted Database (v2) (Pal *et al.*, 2014). Curation of the results followed: (1) results with an e-value > 1E-10 were discarded; (2) results with a query coverage < 70% were discarded; (3) results with a pairwise identity < 25% were discarded; and (4) results with 25% < pairwise identity < 80% were manually assessed for likelihood of a positive hit based on sequence length, analysis of surrounding genomic region, and similarity of the protein annotation (Cai *et al.*, 2019, Arango-Argoty *et al.*, 2018).

DATA AVAILABILITY

Processed 16S rRNA amplicon sequencing reads are available in the Department of Energy (DOE) KnowledgeBase (KBase) as a static narrative: https://doi.org/10.25982/112150.61/1863566 (Goff *et al.*, 2022a). Geochemical data are also available in this narrative. Additionally, raw 16S amplicon sequencing reads are available in the Sequence Read Archive (SRA) under the BioProject number PRJNA827189.

ACKNOWLEDGEMENTS

We thank three anonymous reviewers for their detailed and constructive commentary which have significantly enhanced this work. We kindly thank Dominique C. Joyner, Erin Kelly, Kenneth Lowe and Miguel Rodriguez Jr. for assistance with field sampling and shipping. This work was made possible by the support of our Radiological Control Technicians Mike Cooke, Joel Miller, and Alex Patton as well as our CPT rig operators Bobby Brandt and Tyrone Sanders of Fugro. We would like to thank Astrid Terry at Lawrence Berkeley National Laboratory; Ryan Post and Firas Mishu at M&W Drilling; and Dennis Stauffer at Fugro for their support and effort in securing contracts to complete this work. This material by ENIGMA (Ecosystems and Networks Integrated with Genes and Molecular Assemblies) (http://enigma.lbl.gov), a Science Focus Area Program at Lawrence Berkeley National Laboratory, is based on work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract DE-AC02-05CH11231.

REFERENCES

Abdelouas, A., Lutze, W.,and Nuttall, E. (1998) Chemical reactions of uranium in ground water at a mill tailings site. *J Contam Hydrol* **34:** 343-361.

Alanjary, M., Steinke, K.,and Ziemert, N. (2019) AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res* **47:** W276-W282.

Arango-Argoty, G., Garner, E., Pruden, A., Heath, L.S., Vikesland, P.,and Zhang, L. (2018) DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome* **6:** 1-15.

Arjoon, A., Olaniran, A.O.,and Pillay, B. (2013) Co-contamination of water with chlorinated hydrocarbons and heavy metals: challenges and current bioremediation strategies. *Int J Environ Sci Technol (Tehran)* **10:** 395-412.

Arkin, A.P., Cottingham, R.W., Henry, C.S., Harris, N.L., Stevens, R.L., Maslov, S., Dehal, P., Ware, D., Perez, F., Canon, S., Sneddon, M.W., Henderson, M.L., Riehl, W.J., Murphy-Olson, D., Chan, S.Y., Kamimura, R.T., Kumari, S., Drake, M.M., Brettin, T.S., Glass, E.M., Chivian, D., Gunter, D., Weston, D.J., Allen, B.H., Baumohl, J., Best, A.A., Bowen, B., Brenner, S.E., Bun, C.C., Chandonia, J.-M., Chia, J.-M., Colasanti, R., Conrad, N., Davis, J.J., Davison, B.H., Dejongh, M., Devoid, S., Dietrich, E., Dubchak, I., Edirisinghe, J.N., Fang, G., Faria, J.P., Frybarger, P.M., Gerlach, W., Gerstein, M., Greiner, A., Gurtowski, J., Haun, H.L., He, F., Jain, R., Joachimiak, M.P., Keegan, K.P., Kondo, S., Kumar, V., Land, M.L., Meyer, F., Mills, M., Novichkov, P.S., Oh, T., Olsen, G.J., Olson, R., Parrello, B., Pasternak, S., Pearson, E., Poon, S.S., Price, G.A., Ramakrishnan, S., Ranjan, P., Ronald, P.C., Schatz, M.C., Seaver, S.M.D., Shukla, M., Sutormin, R.A., Syed, M.H., Thomason, J., Tintle, N.L., Wang, D., Xia, F., Yoo, H., Yoo, S.,and Yu, D. (2018) KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* **36:** 566-569.

Aziz, R.K., Bartels, D., Best, A.A., Dejongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., Mcneil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A.,and Zagnitko, O. (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9:** 75.

Baeseman, J.L., Smith, R.L.,and Silverstein, J. (2006) Denitrification potential in stream sediments impacted by acid mine drainage: effects of pH, various electron donors, and iron. *Microb Ecol* **51:** 232-241.

Boguslavsky, A.E., Gaskova, O.L., Naymushina, O.S., Popova, N.M.,and Safonov, A.V. (2020) Environmental monitoring of low-level radioactive waste disposal in electrochemical plant facilities in Zelenogorsk, Russia. *Appl Geochem* **119:** 104598-104609.

Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., Huttley, G.A.,and Gregory Caporaso, J. (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6:** 90.

Bollag, J.M.,and Barabasz, W. (1979) Effect of heavy metals on the denitrification process in soil. *J Environ Qual* **8:** 196-201.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.-X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., Mcdonald, D., Mciver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., Van Der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., Von Hippel, M., Walters, W., et al. (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37:** 852-857.

Bradl, H.B. (2004) Adsorption of heavy metal ions on soils and soils constituents. *J Colloid Interface Sci* **277:** 1-18.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A., Stevens, R., Vonstein, V., Wattam, A.R.,and Xia, F. (2015) RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* **5:** 8365.

Brown, N.L., Stoyanov, J.V., Kidd, S.P.,and Hobman, J.L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27:** 145-163.

Burgin, A.J.,and Hamilton, S.K. (2007) Have we overemphasized the role of denitrification in aquatic ecosystems? A review of nitrate removal pathways. *Front Ecol Environ* **5:** 89-96.

Busenlehner, L.S., Pennella, M.A.,and Giedroc, D.P. (2003) The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. *FEMS Microbiol Rev* **27:** 131-143.

Cai, X., Zheng, X., Zhang, D., Iqbal, W., Liu, C., Yang, B., Zhao, X., Lu, X.,and Mao, Y. (2019) Microbial characterization of heavy metal resistant bacterial strains isolated from an electroplating wastewater treatment plant. *Ecotoxicol Environ Saf* **181:** 472-480.

Callahan, B.J., Mcmurdie, P.J.,and Holmes, S.P. (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal* **11:** 2639-2643.

Callahan, B.J., Mcmurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A.,and Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13:** 581-583.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G.,and Knight, R. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6:** 1621-1624.

Carlin, F., Brillard, J., Broussolle, V., Clavel, T., Duport, C., Jobin, M., Guinebretière, M.-H., Auger, S., Sorokine, A.,and Nguyen-Thé, C. (2010) Adaptation of *Bacillus cereus*, an ubiquitous worldwide-distributed foodborne pathogen, to a changing environment. *Food Research International* **43:** 1885-1894.

Carlson, H.K., Price, M.N., Callaghan, M., Aaring, A., Chakraborty, R., Liu, H., Kuehl, J.V., Arkin, A.P.,and Deutschbauer, A.M. (2019) The selective pressures on the microbial community in a metal-contaminated aquifer. *ISME J* **13:** 937-949.

Casacuberta, E.,and González, J. (2013) The impact of transposable elements in environmental adaptation. *Mol Ecol* **22:** 1503-17.

Cases, I., De Lorenzo, V.,and Ouzounis, C.A. (2003) Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol* **11:** 248-253.

Chandrangsu, P., Rensing, C.,and Helmann, J.D. (2017) Metal homeostasis and resistance in bacteria. *Nat Rev Microbiol* **15:** 338-350.

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P.,and Parks, D.H. (2020) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36:** 1925-1927.

Coombs, J.M.,and Barkay, T. (2004) Molecular evidence for the evolution of metal homeostasis genes by lateral gene transfer in bacteria from the deep terrestrial subsurface. *Appl Environ Microbiol* **70:** 1698-1707.

Daigle, F., Lerat, S., Bucca, G., Sanssouci, É., Smith, C.P., Malouin, F., Beaulieu, C.,and Schneewind, O. (2015) A terD domain-encoding gene (SCO2368) is involved in calcium homeostasis and participates in calcium regulation of a DosR-like regulon in *Streptomyces coelicolor*. *J Bacteriol* **197:** 913-923.

Dosanjh, N.S.,and Michel, S.L.J. (2006) Microbial nickel metalloregulation: NikRs for nickel ions. *Curr Opin Chem Biol* **10:** 123-130.

Douarre, P.-E., Mallet, L., Radomski, N., Felten, A.,and Mistou, M.-Y. (2020) Analysis of COMPASS, a new comprehensive plasmid database revealed prevalence of multireplicon and extensive diversity of IncF plasmids. *Front Microbiol* **11:** 483.

Drevinek, P., Baldwin, A., Lindenburg, L., Joshi, L.T., Marchbank, A., Vosahlikova, S., Dowson, C.G.,and Mahenthiralingam, E. (2010) Oxidative stress of *Burkholderia cenocepacia* induces insertion sequence-mediated genomic rearrangements that interfere with macrorestriction-based genotyping. *J Clin Microbiol* **48:** 34-40.

Du, D., Wang-Kan, X., Neuberger, A., Van Veen, H.W., Pos, K.M., Piddock, L.J.,and Luisi, B.F. (2018) Multidrug efflux pumps: structure, function and regulation. *Nat Rev Microbiol* **16:** 523-539.

Duxbury, T. 1985. Ecological aspects of heavy metal responses in microorganisms. In Marshall, K. C. (ed.) *Adv Microb Ecol.*

Edgar, R.C. (2018) Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* **34:** 2371-2375.

Epa, U. 2004. SW-846 test method 9045D: soil and waste pH. *US EPA, Washington DC.*

Fayad, N., Kallassy Awad, M.,and Mahillon, J. (2019) Diversity of Bacillus cereus sensu lato mobilome. *BMC Genomics* **20:** 436.

Felten, A., Guillier, L., Radomski, N., Mistou, M.-Y., Lailler, R.,and Cadel-Six, S. (2017) Genome Target Evaluator (GTEvaluator): A workflow exploiting genome dataset to measure the sensitivity and specificity of genetic markers. *PLoS One* **12:** e0182082.

Fields, M.W., Yan, T., Rhee, S.-K., Carroll, S.L., Jardine, P.M., Watson, D.B., Criddle, C.S.,and Zhou, J. (2005) Impacts on microbial communities and cultivable isolates from groundwater contaminated with high levels of nitric acid–uranium waste. *FEMS Microbiol Ecol* **53:** 417-428.

Foglar, L., Briški, F., Sipos, L.,and Vuković, M. (2005) High nitrate removal from synthetic wastewater with the mixed bacterial culture. *Bioresour Technol* **96:** 879-888.

Fowle, D.A., Fein, J.B.,and Martin, A.M. (2000) Experimental study of uranyl adsorption onto Bacillus subtilis. *Environ Sci Technol* **34:** 3737-3741.

Gadd, G.M. (2010) Metals, minerals and microbes: geomicrobiology and bioremediation. *Microbiology* **156:** 609-643.

Ge, X., Vaccaro, B.J., Thorgersen, M.P., Poole, F.L., 2nd, Majumder, E.L., Zane, G.M., De León, K.B., Lancaster, W.A., Moon, J.W., Paradis, C.J., Von Netzer, F., Stahl, D.A., Adams, P.D., Arkin, A.P., Wall, J.D., Hazen, T.C.,and Adams, M.W.W. (2019) Iron- and aluminium-induced depletion of molybdenum in acidic environments impedes the nitrogen cycle. *Environ Microbiol* **21:** 152-163.

Goff, J., Putt, A., Fan, Y., Michael, J., Wang, Y., Ning, D., Fu, Y., Van Nostrand, J., Chandonia, J.-M., Hazen, T., Zhou, J., Arkin, A.,and Adams, M. (2022a). Geochemical and sequencing data from Goff et. al. 2022. https://doi.org/10.25982/112150.61/1863566

Goff, J.L., Lui, L.M., Nielsen, T.N., Thorgersen, M.P., Szink, E.G., Chandonia, J.-M., Poole, F.L., Zhou, J., Hazen, T.C., Arkin, A.P.,and Adams., M.W.W. (2022b) Complete Genome Sequence of *Bacillus cereus* Strain CPT56D587-MTF, Isolated from a Nitrate- and Metal-Contaminated Subsurface Environment. *Microbiology Resource Announcements* **5:** e00145-22.

Gui, M., Chen, Q., Ma, T., Zheng, M.,and Ni, J. (2017) Effects of heavy metals on aerobic denitrification by strain *Pseudomonas stutzeri* PCN-1. *Appl Microbiol Biotechnol* **101:** 1717-1727.

Gumaelius, L., Smith, E.H.,and Dalhammar, G. (1996) Potential biomarker for denitrification of wastewaters: effects of process variables and cadmium toxicity. *Water Res* **30:** 3025-3031.

Hemme, C.L., Deng, Y., Gentry, T.J., Fields, M.W., Wu, L., Barua, S., Barry, K., Tringe, S.G., Watson, D.B., He, Z., Hazen, T.C., Tiedje, J.M., Rubin, E.M.,and Zhou, J. (2010) Metagenomic insights into evolution of a heavy metal-contaminated groundwater microbial community. *ISME J* **4:** 660-672.

Hemme, C.L., Green, S.J., Rishishwar, L., Prakash, O., Pettenato, A., Chakraborty, R., Deutschbauer, A.M., Nostrand, J.D.V., Wu, L., He, Z., Jordan, I.K., Hazen, T.C., Arkin, A.P., Kostka, J.E., Zhou, J.,and Huffnagle, G.B. (2016) Lateral gene transfer in a heavy metal-contaminated-groundwater microbial community. *mBio* **7:** e02234-15.

Huang, S.-C., Burne, R.A.,and Chen, Y.-Y.M. (2014) The pH-dependent expression of the urease operon in *Streptococcus salivarius* is mediated by CodY. *Appl Environ Microbiol* **80:** 5386-5393.

Keskin, T.E. (2010) Nitrate and heavy metal pollution resulting from agricultural activity: a case study from Eskipazar (Karabuk, Turkey). *Environ Earth Sci* **61:** 703.

Leplae, R., Lima-Mendez, G.,and Toussaint, A. (2010) ACLAME: a CLAssification of Mobile genetic Elements, update 2010. *Nucleic Acids Res* **38:** D57-61.

Letunic, I.,and Bork, P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49:** W293-W296.

Li, X., Xie, Y., Liu, M., Tai, C., Sun, J., Deng, Z.,and Ou, H.-Y. (2018) oriTfinder: a web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. *Nucleic Acids Res* **46:** W229-W234.

Liang, B., Wang, G., Zhao, Y., Chen, K., Li, S.,and Jiang, J. (2011) Facilitation of bacterial adaptation to chlorothalonil-contaminated sites by horizontal transfer of the chlorothalonil hydrolytic dehalogenase gene. *Appl Environ Microbiol* **77:** 4268-4272.

Liu, Y., Lai, Q., Göker, M., Meier-Kolthoff, J.P., Wang, M., Sun, Y., Wang, L.,and Shao, Z. (2015) Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci Rep* **5:** 14082.

Lu, Z., Gan, L., Lin, J.,and Chen, Z. (2019) Aerobic denitrification by *Paracoccus* sp. YF1 in the presence of Cu(II). *Sci Total Environ* **658:** 80-86.

Lui, L.M., Majumder, E.L.W., Smith, H.J., Carlson, H.K., Von Netzer, F., Fields, M.W., Stahl, D.A., Zhou, J., Hazen, T.C., Baliga, N.S., Adams, P.D.,and Arkin, A.P. (2021) Mechanism across scales: a holistic modeling framework integrating laboratory and field studies for microbial ecology. *Front Microbiol* **12:** 642422.

Martinez, R.J., Wang, Y., Raimondo, M.A., Coombs, J.M., Barkay, T.,and Sobecky, P.A. (2006) Horizontal gene transfer of PIB-type ATPases among bacteria isolated from radionuclide- and metal-contaminated subsurface soils. *Appl Environ Microbiol* **72:** 3111-8.

Medini, D., Donati, C., Tettelin, H., Masignani, V.,and Rappuoli, R. (2005) The microbial pan-genome. *Curr Opin Genet Dev* **15:** 589-94.

Merrell, D.S.,and Camilli, A. (2002) Acid tolerance of gastrointestinal pathogens. *Curr Opin Microbiol* **5:** 51-55.

Mijnendonckx, K., Provoost, A., Monsieurs, P., Leys, N., Mergeay, M., Mahillon, J.,and Van Houdt, R. (2011) Insertion sequence elements in *Cupriavidus metallidurans* CH34: Distribution and role in adaptation. *Plasmid* **65:** 193-203.

Miranda, K.M., Espey, M.G.,and Wink, D.A. (2001) A rapid, simple spectrophotometric method for simultaneous detection of nitrate and nitrite. *Nitric Oxide* **5:** 62-71.

Moon, J.-W., Paradis, C.J., Joyner, D.C., Von Netzer, F., Majumder, E.L., Dixon, E.R., Podar, M., Ge, X., Walian, P.J., Smith, H.J., Wu, X., Zane, G.M., Walker, K.F., Thorgersen, M.P., Poole Ii, F.L., Lui, L.M., Adams, B.G., De León, K.B., Brewer, S.S., Williams, D.E., Lowe, K.A., Rodriguez, M., Mehlhorn, T.L., Pfiffner, S.M., Chakraborty, R., Arkin, A.P., Wall, J.D., Fields, M.W., Adams, M.W.W., Stahl, D.A., Elias, D.A.,and Hazen, T.C. (2020) Characterization of subsurface media from locations up- and down-gradient of a uranium-contaminated aquifer. *Chemosphere* **255:** 126951.

Moreels, D., Crosson, G., Garafola, C., Monteleone, D., Taghavi, S., Fitts, J.P.,and Van Der Lelie, D. (2008) Microbial community dynamics in uranium contaminated subsurface sediments under biostimulated conditions with high nitrate and nickel pressure. *Environ Sci Pollut Res* **15:** 481-491.

Nakano, M.M., Hoffmann, T., Zhu, Y.,and Jahn, D. (1998) Nitrogen and oxygen regulation of *Bacillus subtilis* nasDEF encoding NADH-dependent nitrite reductase by TnrA and ResDE. *J Bacteriol* **180:** 5344-50.

Nies, D.H. (2003) Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol Rev* **27:** 313-339.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F.,and Stevens, R. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* **42:** D206-D214.

Pal, C., Bengtsson-Palme, J., Rensing, C., Kristiansson, E.,and Larsson, D.G. (2014) BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res* **42:** D737-43.

Paruthiyil, S., Pinochet-Barros, A., Huang, X., Helmann, J.D.,and Henkin, T.M. (2020) *Bacillus subtilis* TerC family proteins help prevent manganese intoxication. *J Bacteriol* **202:** e00624-19.

Pasternak, C., Ton-Hoang, B., Coste, G., Bailone, A., Chandler, M.,and Sommer, S. (2010) Irradiation-induced *Deinococcus radiodurans* genome fragmentation triggers transposition of a single resident insertion sequence. *PLoS Genet* **6:** e1000799.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M.,and Duchesnay, É. (2011) Scikit-learn: machine learning in python. *J Mach Learn Res* **12:** 2825–2830.

Philips, S., Laanbroek, H.J.,and Verstraete, W. (2002) Origin, causes and effects of increased nitrite concentrations in aquatic environments. *Rev Env Sci Biotechnol* **1:** 115-141.

Pinel-Cabello, M., Jroundi, F., López-Fernández, M., Geffers, R., Jarek, M., Jauregui, R., Link, A., Vílchez-Vargas, R.,and Merroun, M.L. (2021) Multisystem combined uranium resistance mechanisms and bioremediation potential of *Stenotrophomonas bentonitica* BII-R7: Transcriptomics and microscopic study. *J Hazard Mater* **403:** 123858.

Prakash, O., Green, S.J., Singh, P., Jasrotia, P.,and Kostka, J.E. (2021) Stress-related ecophysiology of members of the genus *Rhodanobacter* isolated from a mixed waste contaminated subsurface. *Front Environ Sci Eng* **15:** 1-9.

Price, M.N., Dehal, P.S.,and Arkin, A.P. (2007) Orthologous transcription factors in bacteria have different functions and regulate different genes. *PLoS Comput Biol* **3:** e175.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J.,and Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41:** D590-D596.

Ramírez, J.E., Rangel-Mendez, J.R., Limberger Lopes, C., Gomes, S.D., Buitrón, G.,and Cervantes, F.J. (2018) Denitrification of metallurgic wastewater: mechanisms of inhibition by Fe, Cr and Ni. *J Chem Technol Biotechnol* **93:** 440-449.

Revil, A., Skold, M., Karaoulis, M., Schmutz, M., Hubbard, S.S., Mehlhorn, T.L.,and Watson, D.B. (2013) Hydrogeophysical investigations of the former S-3 ponds contaminant plumes, Oak Ridge Integrated Field Research Challenge site, Tennessee. *Geophysics The Journal of Biochemistry* **78:** EN29-EN41.

Riley, R.G.,and Zachara, J.M. 1992. Chemical contaminants on DOE lands and selection of contaminant mixtures for subsurface science research. US Department of Energy, Office of Energy Research, Subsurface Science Program.

Sato, K., Okubo, A.,and Yamazaki, S. (1998) Characterization of a multi-copper enzyme, nitrous oxide reductase, from *Rhodobacter sphaeroides* f. sp. *denitrificans*. *J Biochem* **124:** 51-54.

Siguier, P., Filée, J.,and Chandler, M. (2006) Insertion sequences in prokaryotic genomes. *Curr Opin Microbiol* **9:** 526-531.

Siguier, P., Gourbeyre, E.,and Chandler, M. (2014) Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev* **38:** 865-891.

Smith, M.B., Rocha, A.M., Smillie, C.S., Olesen, S.W., Paradis, C., Wu, L., Campbell, J.H., Fortney, J.L., Mehlhorn, T.L., Lowe, K.A., Earles, J.E., Phillips, J., Techtmann, S.M., Joyner, D.C., Elias, D.A., Bailey, K.L., Hurt, R.A., Preheim, S.P., Sanders, M.C., Yang, J., Mueller, M.A., Brooks, S., Watson, D.B., Zhang, P., He, Z., Dubinsky, E.A., Adams, P.D., Arkin, A.P., Fields, M.W., Zhou, J., Alm, E.J., Hazen, T.C.,and Lindow, S.E. (2015) Natural Bacterial Communities Serve as Quantitative Geochemical Biosensors. *mBio* **6:** e00326-15.

Spain, A.M.,and Krumholz, L.R. (2011) Nitrate-reducing bacteria at the nitrate and radionuclide contaminated Oak Ridge Integrated Field Research Challenge site: a review. *Geomicrobiol J* **28:** 418-429.

Sumner, M. (1994) Measurement of soil pH: problems and solutions. *Commun Soil Sci Plant Anal* **25:** 859-879.

Sun, Y., De Vos, P.,and Heylen, K. (2016) Nitrous oxide emission by the non-denitrifying, nitrate ammonifier *Bacillus licheniformis*. *BMC Genomics* **17:** 68.

Teichmann, S.A.,and Babu, M.M. (2004) Gene regulatory network growth by duplication. *Nat Genet* **36:** 492-496.

Thorgersen, M.P., Ge, X., Poole, F.L., 2nd, Price, M.N., Arkin, A.P.,and Adams, M.W.W. (2019) Nitrate-utilizing microorganisms resistant to multiple metals from the heavily contaminated Oak Ridge Reservation. *Appl Environ Microbiol* **85:** e00896-19.

Thorgersen, M.P., Lancaster, W.A., Ge, X., Zane, G.M., Wetmore, K.M., Vaccaro, B.J., Poole, F.L., Younkin, A.D., Deutschbauer, A.M., Arkin, A.P., Wall, J.D.,and Adams, M.W.W. (2017a) Mechanisms of chromium and uranium toxicity in *Pseudomonas stutzeri* RCH2 grown under anaerobic nitrate-reducing conditions. *Front Microbiol* **8**.

Thorgersen, M.P., Lancaster, W.A., Rajeev, L., Ge, X., Vaccaro, B.J., Poole, F.L., Arkin, A.P., Mukhopadhyay, A.,and Adams, M.W. (2017b) A highly expressed high-molecular-weight S-layer complex of *Pelosinus* sp. strain UFO1 binds uranium. *Appl Environ Microbiol* **83:** e03044-16.

Thorgersen, M.P., Lancaster, W.A., Vaccaro, B.J., Poole, F.L., Rocha, A.M., Mehlhorn, T., Pettenato, A., Ray, J., Waters, R.J., Melnyk, R.A., Chakraborty, R., Hazen, T.C.,

Deutschbauer, A.M., Arkin, A.P.,and Adams, M.W.W. (2015) Molybdenum availability Is key to nitrate removal in contaminated groundwater environments. *Appl Environ Microbiol* **81:** 4976-4983.

Vadivelu, V.M., Yuan, Z., Fux, C.,and Keller, J. (2006) The inhibitory effects of free nitrous acid on the energy generation and growth processes of an enriched nitrobacter culture. *Environ Sci Technol* **40:** 4442-4448.

Vandecraen, J., Chandler, M., Aertsen, A.,and Van Houdt, R. (2017) The impact of insertion sequences on bacterial genome plasticity and adaptability. *Crit Rev Microbiol* **43:** 709-730.

Varani, A.M., Siguier, P., Gourbeyre, E., Charneau, V.,and Chandler, M. (2011) ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol* **12:** R30.

Wang, W.-H., Köhler, B., Cao, F.-Q.,and Liu, L.-H. (2008) Molecular and physiological aspects of urea transport in higher plants. *Plant Sci* **175:** 467-477.

Widdel, F.,and Bak, F. 1992. Gram-negative mesophilic sulfate-reducing bacteria. In Balows, A., Trüper, H. G., Dworkin, M., Harder, W.,and Schleifer, K.-H. (eds.) *The Prokaryotes: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications.* New York, NY: Springer New York.

Wu, L., Wen, C., Qin, Y., Yin, H., Tu, Q., Van Nostrand, J.D., Yuan, T., Yuan, M., Deng, Y.,and Zhou, J. (2015) Phasing amplicon sequencing on Illumina Miseq for robust environmental microbial community analysis. *BMC Microbiol* **15:** 125.

Zhou, J., Bruns, M.A.,and Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62:** 316-22.

FIGURES

Figure 1. Geospatial distribution of the CPTF-matching ASV across the ORR Area 3 subsurface sediments. Upper: The relative percent abundance of the CPTF-matching ASV is shown across all subsurface sediment samples. Dots on the scatter plot represent the spatial location of sampling locations while the dot color is representative of the relative abundance of the ASV within that sample (scale is shown in the legend). Lower: A map of the sampling site is also shown with corresponding regions to the scatter plot marked out with the boxes with the grey dashed lines. Sampling locations are marked by pins. Note that while only latitude-longitude coordinates are reflected on this map, multiple depths were sampled at each location. The sampling location where strain CPTF was isolated from (yellow pin) is marked on the map and on the scatter plot.

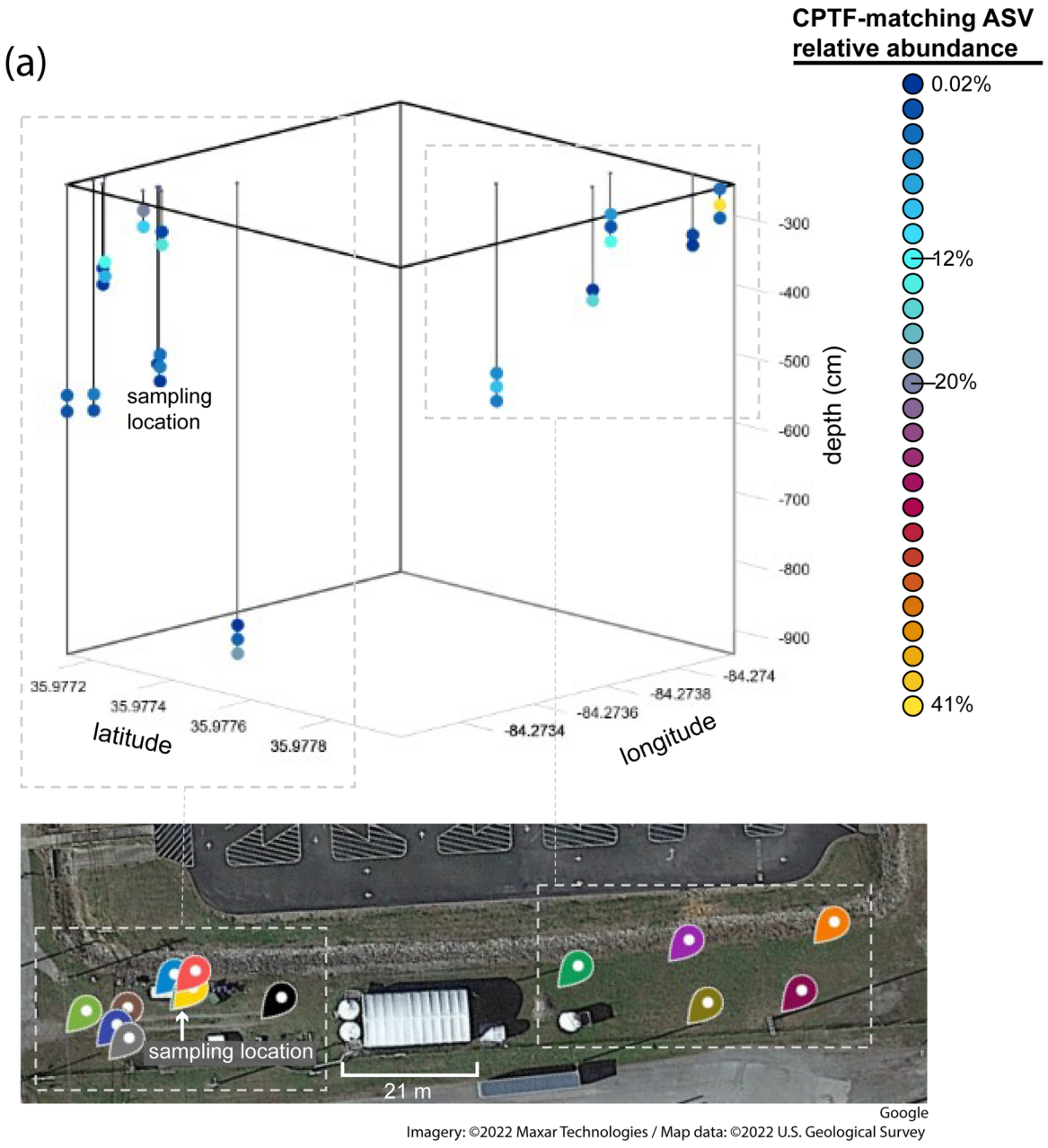
Figure 2. Phylogeny and structural features of 42 completed *B. cereus* genomes including strain CPTF. (a) Multi-locus (n=80) sequence alignment of 42 *B. cereus* genomes used for the pangenome analysis (marked with circles) plus additional reference genomes (no circles). Sequences used for the alignment are listed in **Table S8.** Circle colors indicate the source of the strain: red (clinical), blue (environmental), black (unable to be determined from metadata or literature search). Colors of branches indicate boostrap support (1000 bootstraps). Boostrap values (%) are given in the legend. **(b)** Genome plasmid content (y-axis) is plotted against completed genome assembly length (Mbp) (x-axis). Circles represent individual

genomes. Strain CPTF is labeled for clarity. Circle colors indicate the source of the strain: red (clinical), blue (environmental), black (unable to be determined from metadata or literature search). **(c)** The total number of partial IS elements (y-axis) is plotted against the total number of complete IS elements (x-axis) predicted in each genome. The bubble size is scaled to the chromosomal density (given as a percentage of total chromosomal bp) of the chromosomally-located IS elements. Bubble colors indicate the source of the strain: red (clinical), blue (environmental), black (unable to be determined from metadata or literature search).

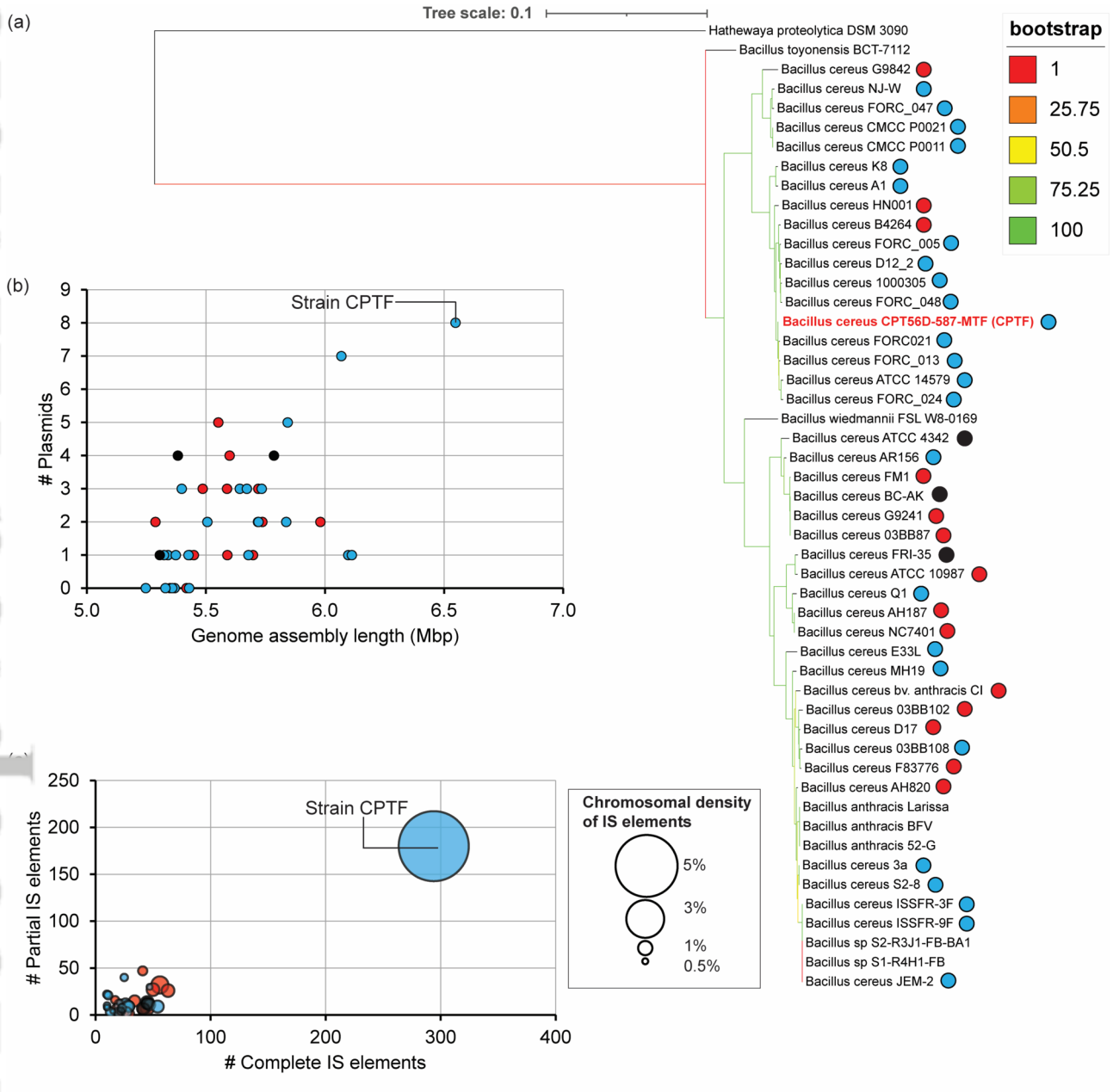
Figure 3. Nitrogen oxide resistance and metabolism of strain CPTF. (a) Relative abundance of the CPTF-matching ASV as a function of porewater nitrate concentrations (open circles). The relative ASV abundance is shown on the primary vertical axis. To examine the limits of survival of strain CPTF with high nitrate/nitrite concentrations, nitrate- (red circles) and nitrite- (green circles) dependent growth are plotted along the secondary vertical axis. These nitrate/nitrite stress experiments were conducted by growing strain CPTF at a range of environmentally relevant nitrate concentrations as well as an identical range of nitrite concentrations with 20 mM glucose as the carbon source. Growth (OD_{600}) was monitored and the extent of growth at 10 h (early stationary phase) was normalized against control cultures grown with no added nitrate/nitrite under anoxic conditions with 20 mM glucose. Individual circles represent the average of three biological replicates and error bars represent ±SD. **(b)** Genomic region of the NarGHI nitrate reductase and the NasDE nitrite reductase. *fnr* encodes an anaerobic metabolism regulator. *narT* encodes a nitrate/nitrite transporter. *ric* encodes an iron-sulfur cluster repair protein. Genes shaded in black have annotated functions unrelated to nitrogen metabolism. **(c)** Assessment of nitrate utilization during the growth of strain CPTF with 10 mM nitrate and 20 mM glucose under anoxic conditions. Nitrate (red), nitrite (green), ammonium (yellow), and total inorganic nitrogen (grey) concentrations were measured. Growth was determined by monitoring OD_{600} (black circles). Individual circles represent the average of four biological replicates and error bars represent ±SD.

Figure 4. Correlation of CPTF-matching ASV range with pH-dependent growth of strain CPTF. (a) The relative abundance of the CPTF-matching ASV is plotted on the vertical axis against the pH of the associated sediment sample. **(b)** Bars represent the maximum OD_{600} of strain CPTF at varying pH values under different growth conditions in batch culture. The figure legend indicates the pH of the culture: pH 3 (green), 4 (orange), 5 (grey), 6 (yellow), 7 (blue). Glucose (20 mM) was added as the carbon source for all experiments. pH 3 data was only collected for the anoxic nitrate-respiring cultures. Graphed data are the average of three biological replicates and error bars represent ±SD.

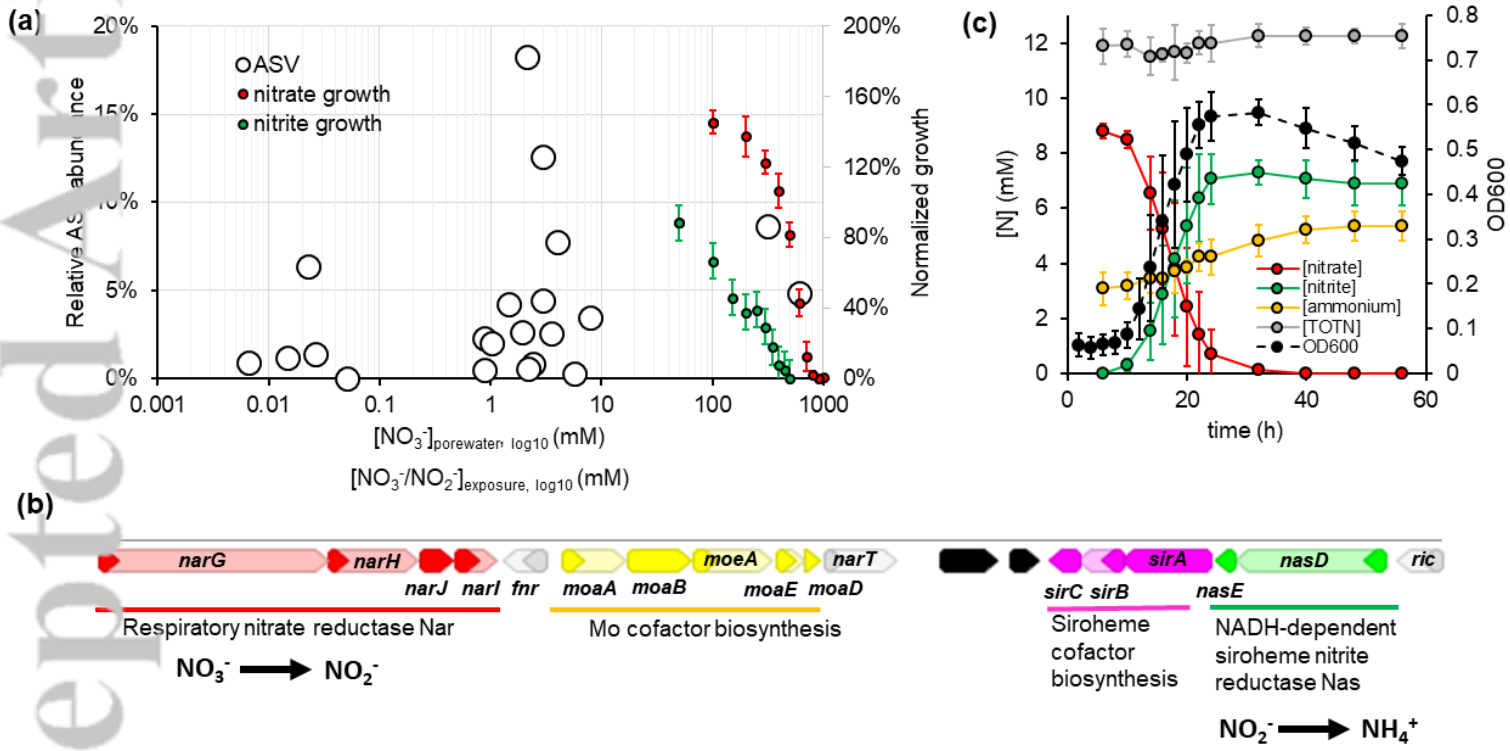
Figure 5. Impact of mixed metals on growth and nitrate respiration of strain CPTF. (a) Relative growth at 10 h (early stationary phase) of strain CPTF under anoxic nitrate-respiring conditions with different dilutions (1X to 0X) of the contaminated ORR metal mix (COMM). Growth is normalized against growth of the control (no COMM addition, (0X)). Cultures contain 10 mM nitrate and 20 mM glucose. Individual points represent the average of three biological replicates and error bars represent ±SD. **(b)** The concentrations of the individual components of the 1X COMM mixture are displayed next to the median ORR Area 3 groundwater (GW) concentrations. The groundwater range is reported in parentheses.The concentrations of the individual components of the 0.84X mix (the IC_{50} value) are displayed in the final column.



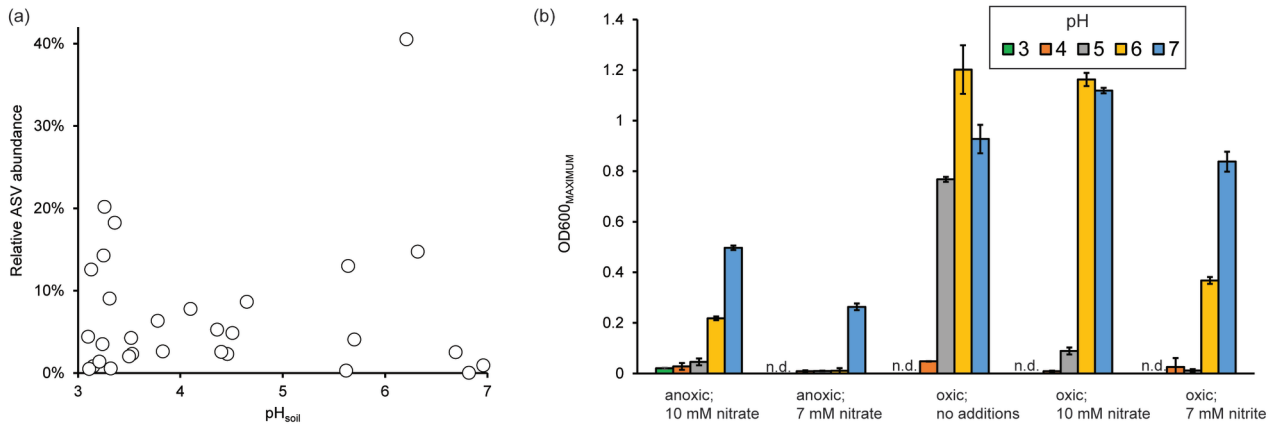
EMI_16173_Figure_1-revised.tif



EMI_16173_Figure 2.tif

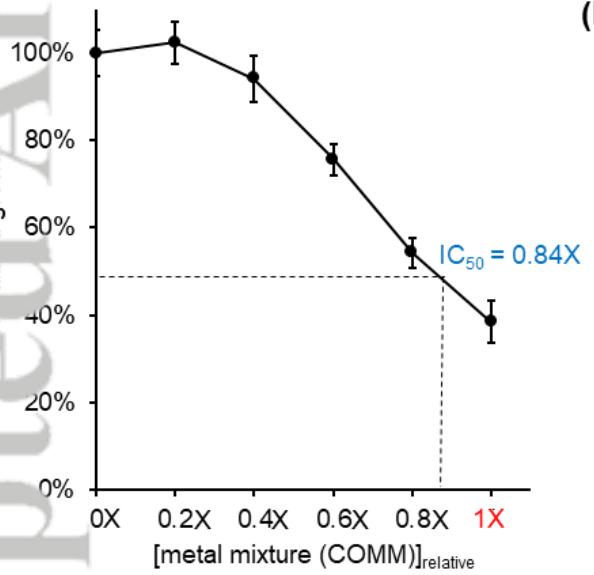


EMI_16173_Figure 3_revised.tif



EMI_16173_Figure_4_revised.tif

(a)



(b)

Metal	[ORR GW Median] (µM)	[1X COMM] (µM)	[0.84X COMM] (µM)
Al	3385 (215-20700)	1000	840
U	110 (40-576)	100	84
Mn	1202 (446-3150)	100	84
Ni	59 (19-157)	100	84
Co	8 (2-30)	30	25.2
Cu	1 (0.2-15)	10	8.4
Fe	1 (0.6-9.9)	10	8.4
Cd	1 (0.6-10)	5	8.4

EMI_16173_Figure 5_revised.tif