2016

Lateral Gene Transfer in a Heavy Metal-Contaminated-Groundwater Microbial Community

Christopher L. Hemme
University of Rhode Island, hemmecl@uri.edu

Stefan J. Green

See next page for additional authors

Follow this and additional works at: https://digitalcommons.uri.edu/bps_facpubs

Authors

Christopher L. Hemme, Stefan J. Green, Lavanya Rishishwar, Om Prakash, Angelica Pettenato, Romy Chakraborty, Adam M. Deutschbauer, Joy D. Van Nostrand, Liyou Wu, Zhili He, I. King Jordan, Terry C. Hazen, Adam P. Arkin, Joel E. Kostka, and Jizhong Zhou

Lateral Gene Transfer in a Heavy Metal-Contaminated-Groundwater Microbial Community

Christopher L. Hemme,[a,b]* Stefan J. Green,[c,d] Lavanya Rishishwar,[e] Om Prakash,[f] Angelica Pettenato,[g] Romy Chakraborty,[g] Adam M. Deutschbauer,[h] Joy D. Van Nostrand,[a] Liyou Wu,[a] Zhili He,[a] I. King Jordan,[e,i] Terry C. Hazen,[j,k,l,m] Adam P. Arkin,[n] Joel E. Kostka,[e] Jizhong Zhou[a,g,o]

Institute for Environmental Genomics, Department of Microbiology and Plant Sciences, University of Oklahoma, Norman, Oklahoma, USA[a]; Ostrow School of Dentistry, University of Southern California, Los Angeles, California, USA[b]; Department of Biological Sciences[c] and DNA Services Facility,[d] University of Illinois at Chicago, Chicago, Illinois, USA; School of Biology and Earth & Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA[e]; National Centre for Cell Science, Ganeshkhind, Pune, Maharashtra, India[f]; Earth Sciences Division[g] and Physical Biosciences Division,[h] Lawrence Berkeley National Laboratory, Berkeley, California, USA; PanAmerican Bioinformatics Institute, Santa Marta, Magdalena, Colombia[i]; Department of Civil & Environmental Engineering,[j] Department of Earth and Planetary Sciences,[k] and Department of Microbiology,[l] University of Tennessee, Knoxville, Tennessee, USA; Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA[m]; Department of Bioengineering, Lawrence Berkeley National Laboratory, Berkeley, California, USA[n]; State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China[o]

* Present address: Christopher L. Hemme, Department of Biomedical and Pharmaceutical Sciences, University of Rhode Island, Kingston, Rhode Island, USA.

ABSTRACT Unraveling the drivers controlling the response and adaptation of biological communities to environmental change, especially anthropogenic activities, is a central but poorly understood issue in ecology and evolution. Comparative genomics studies suggest that lateral gene transfer (LGT) is a major force driving microbial genome evolution, but its role in the evolution of microbial communities remains elusive. To delineate the importance of LGT in mediating the response of a groundwater microbial community to heavy metal contamination, representative *Rhodanobacter* reference genomes were sequenced and compared to shotgun metagenome sequences. 16S rRNA gene-based amplicon sequence analysis indicated that *Rhodanobacter* populations were highly abundant in contaminated wells with low pHs and high levels of nitrate and heavy metals but remained rare in the uncontaminated wells. Sequence comparisons revealed that multiple geochemically important genes, including genes encoding Fe^{2+}/Pb^{2+} permeases, most denitrification enzymes, and cytochrome c_{553}, were native to *Rhodanobacter* and not subjected to LGT. In contrast, the *Rhodanobacter* pangenome contained a recombinational hot spot in which numerous metal resistance genes were subjected to LGT and/or duplication. In particular, $Co^{2+}/Zn^{2+}/Cd^{2+}$ efflux and mercuric resistance operon genes appeared to be highly mobile within *Rhodanobacter* populations. Evidence of multiple duplications of a mercuric resistance operon common to most *Rhodanobacter* strains was also observed. Collectively, our analyses indicated the importance of LGT during the evolution of groundwater microbial communities in response to heavy metal contamination, and a conceptual model was developed to display such adaptive evolutionary processes for explaining the extreme dominance of *Rhodanobacter* populations in the contaminated groundwater microbiome.

IMPORTANCE Lateral gene transfer (LGT), along with positive selection and gene duplication, are the three main mechanisms that drive adaptive evolution of microbial genomes and communities, but their relative importance is unclear. Some recent studies suggested that LGT is a major adaptive mechanism for microbial populations in response to changing environments, and hence, it could also be critical in shaping microbial community structure. However, direct evidence of LGT and its rates in extant natural microbial communities in response to changing environments is still lacking. Our results presented in this study provide explicit evidence that LGT played a crucial role in driving the evolution of a groundwater microbial community in response to extreme heavy metal contamination. It appears that acquisition of genes critical for survival, growth, and reproduction via LGT is the most rapid and effective way to enable microorganisms and associated microbial communities to quickly adapt to abrupt harsh environmental stresses.

Address correspondence to Jizhong Zhou, jzhou@ou.edu.

This article is a direct contribution from a Fellow of the American Academy of Microbiology.

Abrupt, intense environmental stress can overwhelm the ability of microbial communities to adapt by classical selection/ drift/mutation mechanisms and force communities to rapidly adapt to such a stress by large-scale genomic rearrangements, including lateral gene transfer (LGT) and gene duplication (GD) (1–12). These genomic rearrangement events frequently underlie

the natural attenuation of environmental pollutants, leading to important ecological and economic consequences (13). However, it remains challenging to identify and quantify such events from metagenome data alone (1, 6, 12). Most metagenomes are too diverse and complex to allow for extensive assembly, making it difficult to distinguish between native and laterally transferred genes. To ameliorate this problem, sequencing of the genomes of isolated microbes from the environments in question can serve as references for comparison (14–17). This strategy has been extensively used in the Human Microbiome Project to identify pathways of gene sharing between community members and hosts (10, 14, 17–23). However, for many types of environmental samples, such as soil and subsurface groundwater, few reference genomes are available, and thus, a large proportion of such communities are unrepresented in genomic databases (1).

The Oak Ridge Integrated Field Research Challenge (OR-IFRC) site at the Y-12 Federal Security Complex in Oak Ridge, TN, is a well-characterized experimental field site for studying the environmental impacts of legacy waste on soil and groundwater systems (1, 24–27). The groundwater contamination plume at this site is derived from spent uranium and nitric acid waste which was stored in unlined ponds that were capped in 1983. Near-source-zone groundwater contains chronically high concentrations of radionuclides (e.g., uranium), nitric acid, organics, salts, mercury, and other heavy metals, resulting in an extremely low diversity of subsurface microbial communities (1, 27, 28). Repeated cultivation-independent analyses of community genomic DNA and RNA from groundwater recovered from well FW106 within the near-source zone revealed a stressed microbial community of low diversity, dominated by populations of metal-resistant, denitrifying *Rhodanobacter* (>80% by analysis of rRNA gene abundance) (1, 24, 29–33). The FW106 metagenome showed high relative abundances of genes encoding geochemical resistance functions required for microbial survival in the presence of known environmental contaminants at the site (1), including those associated with denitrification, heavy metal resistance, and organic compound degradation. These genes were more abundant in the FW106 metagenome than in a control groundwater community from well FW301 at the OR-IFRC background site (1, 34). Experimental and metagenomic analyses have shown that LGT of heavy metal resistance genes may partially account for their high abundance in the FW106 metagenome (1). To complement metagenomic analyses, multiple laboratories are involved in isolating and sequencing of reference genomes from these environments, including a number of strains of dominant *Rhodanobacter* populations (24, 32, 35–39). These reference genomes are valuable for delineation of the importance of LGT in driving the adaptation of the groundwater microbial communities in response to extremely heavy environmental contamination.

Overabundance of resistance systems, as revealed by several previous studies, can result from selection favoring naturally resistant populations (30, 40–42), via LGT, and/or via GD (1, 6, 7, 10, 11, 43, 44). Three hypotheses can provide explanations for the overabundance of geochemical resistance genes in the contaminant-stressed community: (i) they provide no selective benefit in the contaminated environment and they are overabundant simply because the host organisms are abundant, (ii) they provide selective benefit to the hosts under contaminated conditions but are not subjected to LGT or GD, and (iii) they provide selective benefit to the host under contaminated conditions and

are subjected to LGT and/or GD. To test these hypotheses, multiple *Rhodanobacter* species were isolated from OR-IFRC sites, and their genomes were sequenced and compared to those of other *Rhodanobacter* reference strains and to OR-IFRC site groundwater metagenomes. The metagenomes and several of the *Rhodanobacter* strains were sequenced as part of previous studies, and seven additional *Rhodanobacter* strains were sequenced as part of this study to complement this data set. Comparative analysis showed that many geochemically important genes were highly abundant in the uranium-contaminated groundwater community due to the dominance of *Rhodanobacter* in the environment. A recombinational hot spot was identified in the *Rhodanobacter* pangenome where multiple metal resistance genes were acquired via LGT and/or GD. These results indicated that LGT could play a critical role in driving the evolution of a groundwater microbial community in response to extreme heavy metal contamination.

RESULTS

Overabundance of *Rhodanobacter* species in contaminated wells. A 16S rRNA gene survey of groundwater microbial communities from 100 wells located within the OR-IFRC site was conducted in the years 2012 to 2013 to identify the microbial community composition across the OR-IFRC research area (45). To identify both the locations and the relative abundances of *Rhodanobacter* strains in these communities, 16S rRNA gene sequences from *Rhodanobacter* isolates were obtained from the Ribosomal Database (46) and compared to the 16S rRNA gene sequences from the OR-IFRC site 100-well survey by BLASTn (47). A total of 370,765 OR-IFRC sequences showed hits to *Rhodanobacter* sequences with >97% sequence identity across the aligned regions (see Table S1 in the supplemental material). Of these sequences, 79.8% were mapped to the three wells located in contaminated areas 1 and 3, with 33.8% to FW106 (area 3), 27.4% to FW109 (area 3), and 18.6% to FW021 (area 1) (see Table S1). These wells all lie near the base of contaminated groundwater plumes emanating from the former S-3 waste disposal ponds and are heavily contaminated with uranium, nitric acid, and heavy metals. In contrast to the contaminated sites, only a small number of hits were seen for background (i.e., uncontaminated) wells (e.g., FW300, FW301, and FW303), with <0.1% for each (see Table S1). Thus, while *Rhodanobacter* species are broadly distributed across the OR-IFRC research area, they are most dominant in low-pH groundwater located near the source of the contaminant plumes.

Abundance of geochemically important genes in the FW106 metagenome compared to *Rhodanobacter* species. Previous comparisons of OR-IFRC contaminated-groundwater metagenomes indicated that numerous geochemically important genes were abundant in the contaminated-groundwater metagenome and that some of the abundant genes could have been acquired by LGT (1, 34). However, LGT events are difficult to distinguish with metagenomic sequences alone. To determine whether these overabundant genes were linked to *Rhodanobacter* populations in the contaminated groundwater community and whether LGT contributed to the abundance of these genes, metagenomes from *Rhodanobacter*-dominated uranium-contaminated groundwater communities (OR-IFRC well FW106) and background groundwater communities (OR-IFRC well FW301) were compared directly to the *Rhodanobacter* reference genomes at the species and genus levels (Table 1 and Fig. 1; see also Table S2 in the supplemental material). Previous comparative analysis of the two met-

TABLE 1 Genomes and metagenomes used in this study

Source of genome or metagenome	DSM no.	OR-IRFC strain	Denitrification from nitrate[a]	Reference or source
Rhodanobacter isolate				
R. denitrificans 2APBS1[T]	DSM23569	Y	Y	31
R. denitrificans 116-2	DSM24678	Y	Y	31
Sp. 115		Y	N	31
Sp. OR87		Y	?	24
Sp. OR92		Y	?	24
Sp. OR444		Y	?	
Sp. FW104-R3		Y	Y	This study
Sp. FW104-R5		Y	?	This study
Sp. FW104-R8		Y	Y	This study
Sp. FW104-T7		Y	Y	This study
Sp. FW510-T8		Y	Y	This study
Sp. FW510-R10		Y	Y	This study
Sp. FW510-R12		Y	Y	This study
R. thiooxydans	DSM18863	N	Y	31
R. fulvis	DSM18449	N	N	31
R. spathiphylli	DSM17631	N	N	31
OR-IFRC metagenomes				
FW106		Y	Y	1
FW301		Y	Y	34

[a] Indicates presence or absence of denitrification from nitrate based on experimental and/or genomic evidence. Denitrification from nitrite may still be possible in strains incapable of nitrate reduction.

agenomes showed that the abundance of *Rhodanobacter* populations was <1% in the background groundwater metagenome compared to >85% in the contaminated groundwater metagenome (30). The abundance of each cluster of orthologous groups (COG) of interest was calculated for each sample and for all sequenced bacterial genomes available as of September 2014, and this information was used to calculate ln odds ratios for each COG (Fig. 1). COGs of interest were divided into four categories: nitrogen metabolism and transport, metal transport and metabolism, hydrogenase, and cytochrome.

All *Rhodanobacter* strains studied showed high abundances of genes related to denitrification, divalent metal ion efflux (*czc*), Fe/Pb permease, and cytochrome c_{553}, similar to the contaminated groundwater metagenome (Fig. 1). While the abundance profile of an isolate is unlikely to exactly match that of even a simple metagenome, the abundance profiles of geochemically important genes (including genes encoding denitrification enzymes, hydrogenases, and some cytochromes) in several *Rhodanobacter* strains resembled those of the contaminated groundwater metagenome (Fig. 1). The abundance profiles of the contaminated-groundwater metagenome most closely mirrored those of OR-IFRC *Rhodanobacter* isolates *R. denitrificans* strain 2APBS1, *Rhodanobacter* species strains OR87, OR92, OR444, FW510-R12, FW104-T7, and FW104-R3 and less closely resembled OR-IFRC isolates *Rhodanobacter* sp. strain 115 and *R. denitrificans* strain 116-2 and the three non-OR-IFRC isolates (Fig. 1). The hydrogenase gene profiles showed that the contaminated-groundwater metagenome was most similar to the genomes of *R. denitrificans* strains 2APBS1 and 116-2, *Rhodanobacter* species strains OR87, OR444, FW104-R8, FW510-R10, FW510-R12, and FW510-T8, and *Rhodanobacter thiooxydans* (Fig. 1). These results show that while the contaminated-groundwater metagenome does resemble the genomes of the dominant *Rhodanobacter* populations, the *Rhodanobacter* populations cannot fully account for all of the metabolic potential in the contaminated-groundwater metagenome.

Identification of LGT events in *Rhodanobacter* genomes. Some of the overabundance of genes in the contaminated-

groundwater metagenome could be attributed to genes naturally present in the dominant *Rhodanobacter* populations. However, resistance genes could also be acquired from other species through LGT. To identify likely causes of resistance gene overabundance in the contaminated-groundwater metagenome, the *Rhodanobacter* genomes were searched for target genes overabundant in the contaminated-groundwater metagenome. Putative LGT events in the genomes were then identified and correlated to the genes of interest. Multiple methods to identify aberrant sequences deviating from the genomic background were used to identify LGT events in the isolate genomes, including phylogenetic analysis, DarkHorse (48), Alien_Hunter (49), and GOHTAM (50, 51) (Fig. 2; see also Fig. S1 and S2 and Table S2 in the supplemental material). Analysis of the complete genome of *R. denitrificans* 2APBS1 revealed several regions containing putative LGT events, many of which matched up to annotated transposase and integrase genes (Fig. 2). Fifty-one genes related to heavy metal resistance were also analyzed (see Table S2). Of these genes, seven were identified by three LGT detection programs as putative LGT events, 13 were detected by at least two programs, 19 were detected by a single program, and 12 were not identified as LGT events by any of the programs used (see Table S2). In particular, most of the genes in adjacent arsenate resistance (*ars*) and copper resistance (*cop*) operons were identified as LGT events by all programs used (see Table S2). Many of these genes were located within a large genomic region near the 1.7-to-2.3-Mb mark noted for a large number of metal resistance genes, including operons encoding mercuric, arsenate, and divalent cation resistance (e.g., *czc*) (Fig. 2). DarkHorse analysis indicated the sources of these genes to be multiple gamma- and betaproteobacteria known to be present at OR-IFRC sites, including *Burkholderia*, *Herbaspirillum*, *Ralstonia*, *Cupriavidus*, *Pseudomonas*, and *Stenotrophomonas* (see Table S2). In the following sections, detailed descriptions of gene abundance and possible LGT events for several geochemically important genes are provided.

Nitrate/nitrite antiporter gene abundance. While most denitrification genes were present in high abundance in the FW106
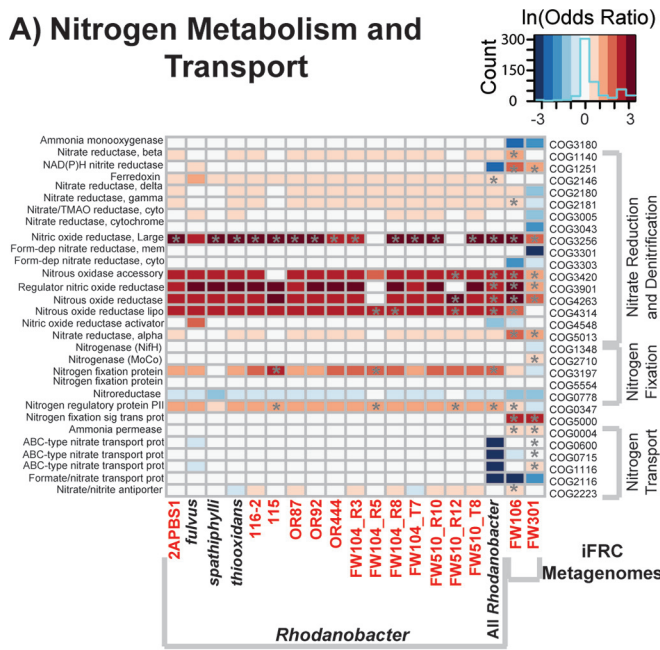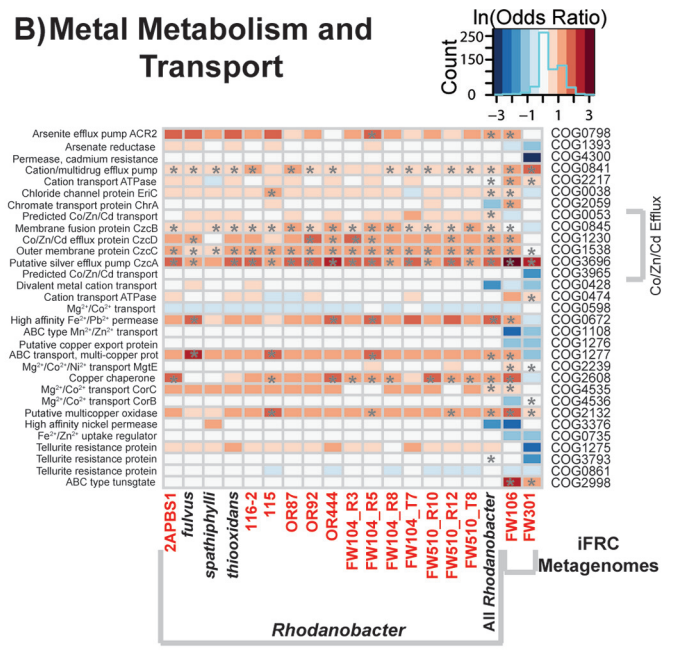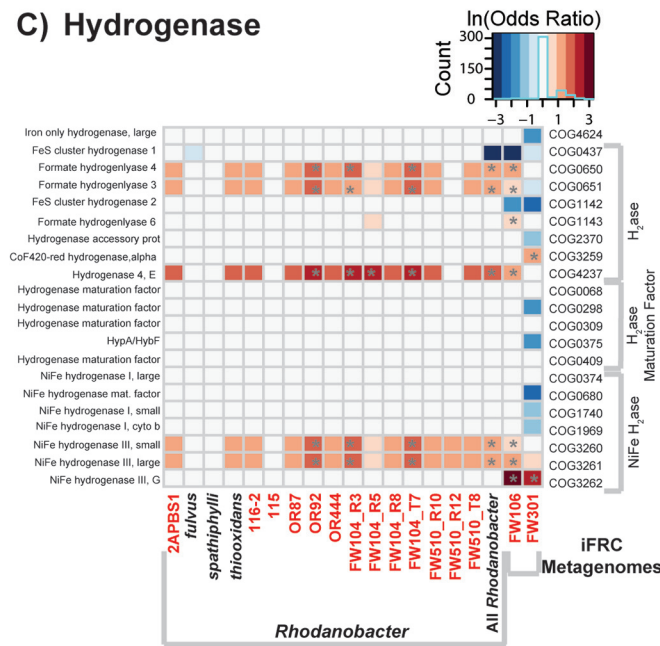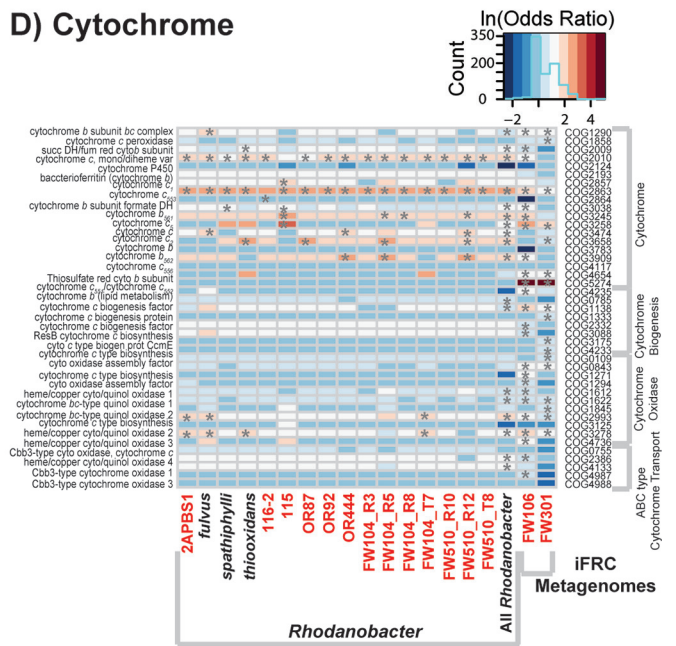
FIG 1 Heat maps of gene abundances of key functional groups in *Rhodanobacter* genomes and OR-IFRC metagenomes. Abundance is measured as relative abundance against all sequenced bacterial genomes and as the odds ratio of each COG number against all COG categories for the selected organism/community. Asterisks indicate significantly overabundant COGs (chi-square test and 1-tailed Fisher's exact test). The *x* axis shows the biological samples, with "All *Rhodanobacter*" representing all *Rhodanobacter* genomes combined and red font indicating OR-IFRC isolates or metagenomes.

metagenome, the putative nitrate/nitrite antiporter (COG2223) genes showed particularly high abundance (Table 1 and Fig. 1A) (1). The numbers of putative nitrate/nitrite antiporter genes varied in the *Rhodanobacter* strains, with 10 strains (*R. denitrificans* 2APBS1, *R. denitrificans* 116-2, *Rhodanobacter* species OR87, OR92, OR444, FW104-R3, FW104-R5, FW104-T7, and FW510-R12, and *R. thiooxydans*) containing genes encoding a pair of nitrate/nitrite transporters associated with denitrification pathways (represented by homologs R2APBS1_2892 and R2APBS1_2893 from *R. denitrificans* 2APB1) (Fig. 3).

Phylogenetic analysis of the nitrate/nitrite transporter (COG2223) genes provided new insights into the population structure of the uranium-contaminated community. Homologs of *R. denitrificans* 2APBS1_2892 were separated into two clades (Fig. 3; see also Fig. S3 in the supplemental material). Ten homologs were found in clade 1a and represented the vertically inherited ancestral *Rhodanobacter* genes (Fig. 3). Three strains (*Rhodanobacter* species strains FW104-R8, FW510-T8, and FW510-R10) lacked this particular suite of genes but had a similar group of genes in a different orientation, appearing to be derived

FIG 2 Genome map of *R. denitrificans* 2APBS1 showing putative LGT events. Regions of anomalous nucleotide composition were computed by using Alien_Hunter on the genome of *Rhodanobacter denitrificans* 2APBS1. The denitrification genes were not found to overlap the putative laterally transferred regions. The HMR gene cluster near the 2-Mb mark was found to overlap the putative laterally transferred regions and was confirmed independently, using DarkHorse, to be putatively laterally transferred (see Fig. S2 in the supplemental material).

from species other than *Rhodanobacter* (represented by clade 1b in Fig. 3). Genes in this cluster were more similar to genes from *Thermomonas*, *Stenotrophomonas*, and *Pseudomonas* species than from *Rhodanobacter* and were associated with transposable elements. Clade 2a (represented by R2APBS1_2893) contains 14 FW106 homologs (Fig. 3). *Rhodanobacter* sp. 115 did not have a clade 2a homolog, and *Rhodanobacter fulvus* contained a distinct homolog representing its own clade (clade 2b) (Fig. 3). A third putative nitrate/nitrite transporter gene (clade 4a) was not associated with denitrification genes and was found in all 14 *Rhodano-*

bacter strains. The region surrounding this gene appeared to be native to *Rhodanobacter*, but the gene itself was most similar to *Burkholderia* homologs. Clade 5 (represented by R2APBS1_2043 from *R. denitrificans* 2APB1) was found in 11 OR-IFRC strains and was absent in the three non-OR-IFRC reference genomes (Fig. 3). This homolog was not associated with denitrification genes but with a cluster of $Cd^{2+}/Zn^{2+}/Co^{2+}$ (*czc*) efflux genes located between a pair of transposons. While the genealogy of the region as a whole was unclear, several of the genes were most similar to betaproteobacterial homologs.
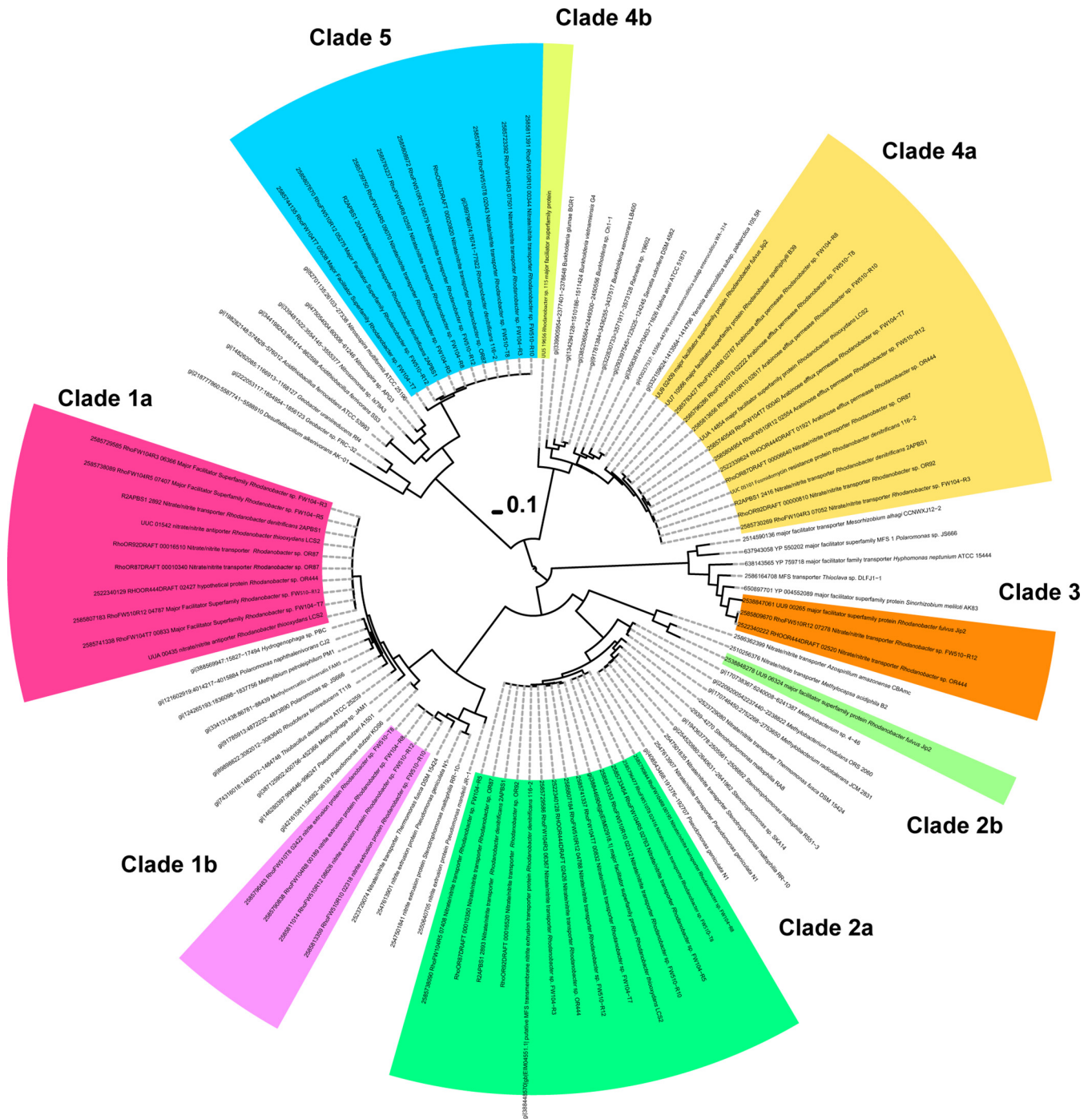
FIG 3 Nitrate/nitrite transporter genes of *Rhodanobacter* species. Sequences corresponding to COG2223 (nitrate/nitrite transporter) were extracted from the *Rhodanobacter* strains, and homologous reference sequences were identified by BLASTp analysis and extracted from GenBank. The sequences separated into 4 clades and 8 subclades, and sequences from *Rhodanobacter* strains are labeled in color. The tree was inferred in MEGA 5.2 using the maximum-likelihood method based on the JTT matrix-based model. The tree with the highest log likelihood ($-13,400.3677$) is shown. A discrete gamma distribution was used to model the evolutionary rate differences among sites (5 categories [$+G$ parameter $= 2.7633$]). The rate variation model allowed for some sites to be evolutionarily invariable ([$+I$], 0.1999% sites). The analysis involved 55 amino acids with a total of 354 positions in the final data set.

Previous analyses of the contaminated-groundwater metagenome suggested the existence of multiple *Rhodanobacter* populations (1). The *Rhodanobacter* COG2223 amino acid sequences were compared to homologs from the contaminated- and background groundwater metagenomes to identify possible popula-

tion structures in these metagenomes (see Fig. S4A in the supplemental material). A partial clade 1a nitrate/nitrite transporter gene sequence from contaminated groundwater was partially reconstructed from four contiguous gene fragments (FW106_100021117–FW106_100021119) on scaffold FW106_c1000211

(19.8 kb), which was most similar to the *R. thiooxydans* homolog (see Fig. S3A). The R2APBS1_2982 homologs in the contaminated-groundwater metagenome homologs formed a distinct subclade within the larger *Rhodanobacter* clade. A similar pattern was observed for the contiguous R2APBS1_2893 homolog, which matched two contaminated-groundwater and one background groundwater metagenome fragment that showed similarity to the genome of *R. thiooxydans* (see Fig. S3B). This evidence supports the hypothesis that multiple *Rhodanobacter* populations exist in the contaminated-groundwater community.

Metal transport and metabolism gene abundance. A variety of metal resistance genes were identified in high abundance in the contaminated-groundwater metagenome, including *mer* operon genes, high-affinity Fe^{2+}/Pb^{2+} permease genes, and $Cd^{2+}/Zn^{2+}/Co^{2+}$ (*czc*) efflux genes, for example (Fig. 1B) (1). Most of these genes were also present in *Rhodanobacter* genomes at variable abundances (Fig. 1B). *czcA* was the most abundant gene, especially in the contaminated-groundwater metagenome and the genomes of *R. denitrificans* 2APBS1, *R. denitrificans* 116-2, *R. thiooxydans*, and *Rhodanobacter* species OR87, OR92, OR444, FW104_R3, FW104_R5, FW104_T7, and FW510_R12. *Rhodanobacter* sp. 115 and *R. spathyphyllii* had lower abundances of *czcD* than other isolates, and FW106 showed a lower abundance of *czcB* than other isolates (Fig. 1B).

Two copies of high-affinity Fe^{2+}/Pb^{2+} permease genes were identified from *Rhodanobacter* genomes (Fig. 1B). The first copy, represented in *R. denitrificans* 2APBS1 by coding sequence (CDS) R2APBS1_2568, was present in all strains except *Rhodanobacter* sp. 115, where this region appeared to have been excised. This region was highly conserved in several genomes (*Rhodanobacter* sp. 115 being the primary exception), and no transposable elements were identified. The second copy, represented in *R. denitrificans* 2APBS1 by CDS R2APBS1_1742, lay in the high-LGT region described above that was identified by DarkHorse analysis. All *Rhodanobacter* species contained a copy of this gene except for *R. spathiphylli* and *Rhodanobacter* sp. 115, and the latter may have been rendered a pseudogene due to a frameshift mutation. As previously mentioned, this region was associated with other metal resistance genes, including Cu^{2+} efflux, Czc efflux, arsenate resistance, and iron efflux genes. *Rhodanobacter* sp. 115 contains an Fe/Pb permease gene that is distinct from other *Rhodanobacter* homologs (Fig. 1B).

Mercury is a major contaminant at the OR-IFRC, and existing experimental evidence has suggested that mercury resistance genes were subjected to LGT in this environment (8). Mercuric resistance genes were present in high abundance in both the contaminated- and background groundwater metagenomes across a wide phylogenetic breadth, and multiple mercuric resistance operons were identified in *Rhodanobacter* genomes (see Fig. S4 to S6 in the supplemental material). One mercuric resistance operon (represented by *merA* gene R2APBS1_1852 of *R. denitrificans* 2APBS1) appeared to be common to *R. denitrificans* 2APBS1, *Rhodanobacter* sp. OR87, *Rhodanobacter* sp. OR92 (all OR-IFRC strains), *R. thiooxydans*, and possibly, OR-IFRC strain *Rhodanobacter* sp. 115. This operon appeared to have been duplicated twice in *R. denitrificans* 2APBS1 (*merA* genes R2APBS1_2075 and R2APBS1_2097) (see Fig. S5 and S6).

Hydrogenase and cytochrome gene abundance. The abundances of hydrogenase systems varied widely between *Rhodanobacter* strains (Fig. 1C). Based on the abundance profiles of hydro-

genase genes, the *Rhodanobacter* strains could be separated into three distinct groups: group I (*R. denitrificans* 2APBS1, *R. denitrificans* 116-2, *R. thiooxydans*, and *Rhodanobacter* species OR87, OR444, FW104_R8, FW510_R10, and FW510_T8), group II (*Rhodanobacter* species OR92, FW104_R3, and FW104_T7), and group III (*Rhodanobacter* sp. 115, *R. fulvus*, and *R. spathiphyli*) (Fig. 1C). Members of groups I and II each contained an operon encoding a formate hydrogen lyase-NiFe hydrogenase III cluster that is absent in cluster III strains. Each of the group I and II strains (except for *R. thiooxydans*) was isolated from OR-IFRC sites, and thus, this system may allow the removal of protons resulting from the metabolism of organic acids present in low-pH environments. The FW106 metagenome showed a hydrogenase gene profile similar to those of the group I and group II *Rhodanobacter* strains, while the FW301 metagenome showed a higher abundance of NiFe hydrogenase I genes (Fig. 1C).

Cytochrome genes were highly abundant in both the FW106 and the FW301 metagenome, with the contaminated-groundwater metagenome showing higher abundances of cytochrome c_{553} (COG2863) genes and cytochrome *c* mono- and diheme variant (COG2010) genes than the background groundwater metagenome (Fig. 1D). The abundance of the cytochrome c_{553} gene in *Rhodanobacter* strains ranged from 6 to 9 copies, and this gene was significantly more abundant in *Rhodanobacter* genomes than in sequenced bacteria overall (Fig. 1D). Several of these homologs (represented by *R. denitrificans* 2APBS1 genes R2APBS1_0257, R2APBS1_1008, R2APBS1_1835, R2APBS1_1836, R2APBS1_3438, and R2APBS1_3645) were present in nearly all of the strains. The R2APBS1_3645 ortholog was missing in strain *Rhodanobacter* sp. 115. All six primary ortholog groups appeared to be native to the genomes and did not appear to be subject to LGT. One copy unique to OR-IFRC strain *Rhodanobacter* sp. 115 (UU5_02112) was associated with copper resistance and *czc* operon genes. Several of the cytochrome c_{553} genes were found in tandem with cytochrome *c* mono-/diheme variant (COG2010) genes that were also overabundant in the FW106 metagenome.

Identification of a recombinational hot spot in *Rhodanobacter* genomes related to heavy metal resistance. DarkHorse analysis and comparison of the *Rhodanobacter* genomes identified a region containing a large number and variety of metal resistance genes linking to multiple putative LGT events (Fig. 2 and 4; see also Fig. S7 in the supplemental material). This region was conserved in every *Rhodanobacter* strain except *R. spathiphyli*, in which the entire region was absent (*R. spathiphyli* is thereby ignored for the remainder of this analysis). Gene order in this region was largely conserved except in *R. thiooxydans*, in which extensive shuffling of gene order appeared to have occurred, and in most strains, the entire region is flanked by transposable elements (Fig. 4; see also Fig. S7). Most genomes (except that of *Rhodanobacter* sp. 115) showed high degrees of assembly in this region, allowing comparisons of both gene content and synteny. The common elements of this region included an orthologous group of Fe^{2+}/Pb^{2+} permease genes (reference CDS R2APBS1_1742), with the *Rhodanobacter* sp. 115 copy annotated as a frameshift and all strains containing an adjacent *czc* operon (including the genes encoding Czc subunits CzcA, -B, -C, and -D and *glnB*) (Fig. 4; see also Fig. S7). Beyond this conserved cluster, the complement of heavy metal resistance genes varied considerably. *R. denitrificans* 2APBS1, *R. denitrificans* 116-2, *Rhodanobacter* sp. OR87, and *R. thiooxydans* all contained a conserved *mer* operon (Fig. 4; see
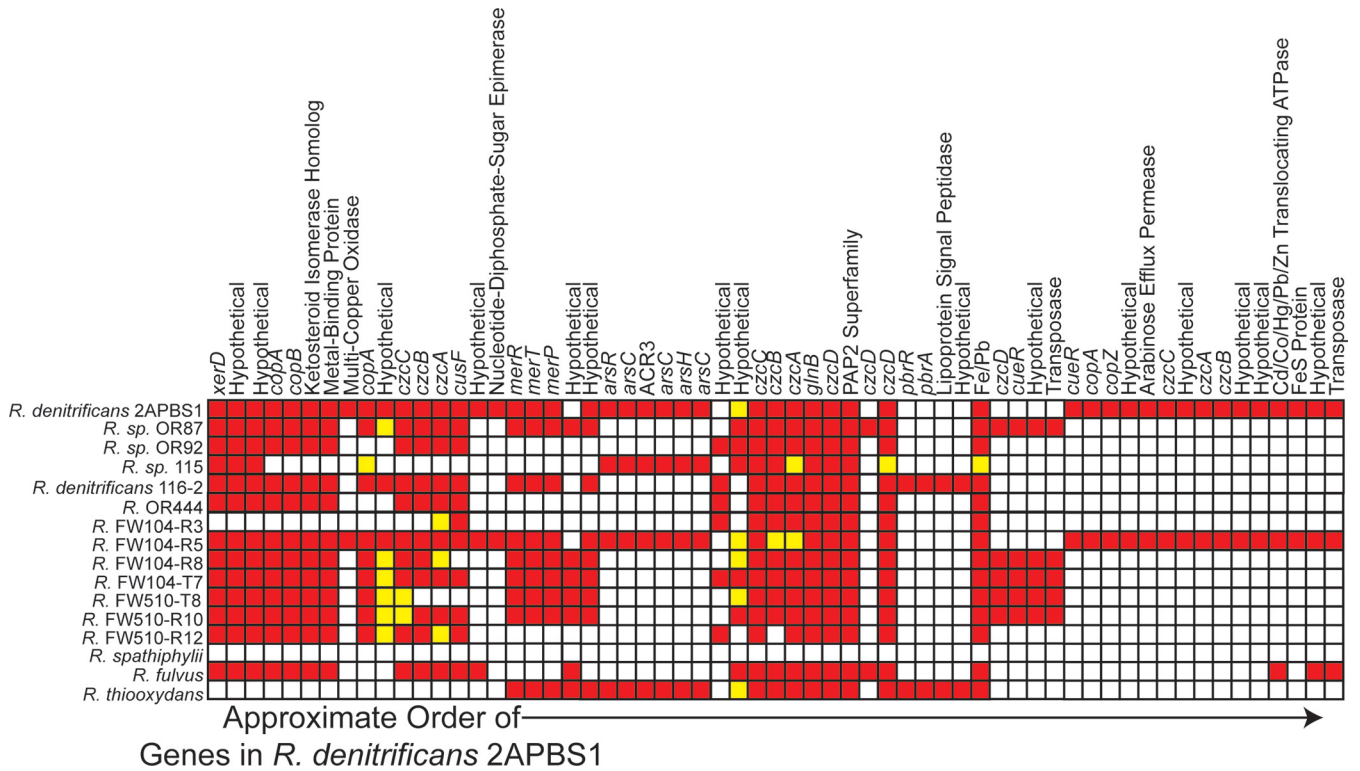
FIG 4 Recombinational hot spot in *Rhodanobacter* genomes that confers heavy metal resistance. The *y* axis shows the organism analyzed, and the *x* axis shows the gene analyzed, with genes organized roughly in the direction of the leading strand of *R. denitrificans*. Cells colored red indicate genes that are present based on annotation, yellow indicates truncated genes or unannotated conserved regions, and white indicates the absence of the gene.

also Fig. S7). *R. denitrificans*, *Rhodanobacter* sp. 115, and *R. thiooxydans* also contained an *ars* arsenic resistance operon (Fig. 5; see also Fig. S7). Multiple instances of *czc* operon genes, copper resistance genes (*copAB* and *cueR*), lead resistance genes (*pbrA*), and uncharacterized heavy metal resistance genes were found throughout the genomes, often in lineage-specific arrangements (Fig. 4; see also Fig. S7). Of the strains isolated from wells FW104 and FW510, FW104_R5 closely resembled *R. denitrificans* 2APBS1 and most of the other strains resembled *Rhodanobacter* sp. OR87 (Fig. 4). Only *R. denitrificans* 2APBS1, *R. thiooxydans*, and *Rhodanobacter* species 115 and FW104_R5 contained arsenate reductase genes while containing the conserved mercuric reductase operon (Fig. 4).

DISCUSSION

The mechanism by which specific bacterial populations adapt to intensive stress resulting from multiple environmental contaminants and extreme conditions remains an open question. Several experimental analyses at the OF-IFRC sites suggest that nitrate concentration, pHs, and heavy metal concentrations are the key drivers of community structure and functions in contaminated sites (30, 41), and many strains have been isolated from these sites that show resistance to these particular environmental stresses (28, 32, 37, 39, 52, 53). Comparative genomics analyses of the *Rhodanobacter* genomes and the OR-IFRC groundwater metagenomes show that many of the genes identified as overabundant in the contaminated-groundwater metagenome are present in most of the *Rhodanobacter* strains. Some of these genes, such as the core denitrification pathway genes, are vertically inherited, with no evidence of lateral gene transfer. Many metal resistance genes, including *czc* and mercuric resistance genes, show evidence of extensive lateral gene transfer and/or gene duplication in *Rhodanobacter* strains. An area of the *Rhodanobacter* pangenome appears to be a recombinational hot spot that allows rapid acquisition and accumulation of these genes in *Rhodanobacter* populations. Thus, the overabundance of putative resistance genes in the contaminated-groundwater population can be attributed to a combination of vertical inheritance, LGT, and GD in the dominant *Rhodanobacter* population (Table 2).

However, it should be noted that the methods used for identifying putative LGT events all use different strategies to detect regions of the genome that deviate from the genomic background. Of the 51 putative laterally transferred genes listed in Table S2 in the supplemental material, 76.5% were identified as positive LGT events by at least one of the programs used (Alien_Hunter, Dark-Horse, and GOHTAM) and 39.2% were identified as positive LGT events by at least two methods. Many of these genes are located in operons, and thus, it can be assumed that if some of the genes in the cluster were identified as laterally transferred, then all genes in the cluster were likely laterally transferred as well. Additional evidence supporting lateral transfer of these genes is their phylogenetic similarity to orthologs from organisms other than *Rhodanobacter* and the association of many of these genes in the *Rhodanobacter* genomes with mobile elements.

Based on the results of this work, a conceptual model was constructed to predict how *Rhodanobacter* populations adapted to contamination and came to dominance in many of the OR-IFRC contaminated-groundwater communities. Comparative genom-
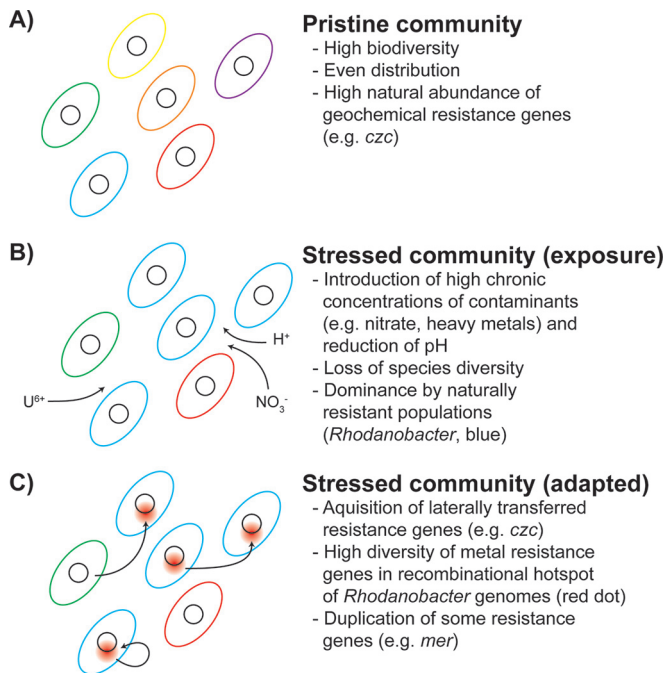
FIG 5 A conceptual model of *Rhodanobacter* population dynamics in contaminated-groundwater systems. Populations are distinguished by different colors, with *Rhodanobacter* shown in blue. Fuzzy red dots indicate the proposed recombinational hot spot in the *Rhodanobacter* pangenome. (A) Background state with high biodiversity and low abundance of *Rhodanobacter*. (B) Contaminated state with chronic exposure to uranium, nitrate, and low pH in which the environment selects for *Rhodanobacter*. (C) Lateral transfer and duplication of geochemical resistance genes (indicated by arrows) within and between *Rhodanobacter* populations, which initiates subsequent rounds of selection.

ics analyses of *Rhodanobacter* isolates suggest a selection-transfer-selection dynamic by which populations of this particular genus survive and thrive in these environments (Fig. 5 and Table 2). The original community prior to contamination is highly diverse, consisting of many different species with different abundances (represented by different colors in Fig. 5A). Following the introduction of contaminants, the community goes through an initial round of selection in which only those species naturally resistant to the contaminants survive (indicated by the increase in the number of *Rhodanobacter* populations, represented by blue in Fig. 5B). In the case of *Rhodanobacter*, vertically inherited genes, such as cytochrome c_{553} genes, denitrification genes, and native metal resistance genes like *czc*, may impart a selective benefit on these populations (Fig. 5B). Simultaneous with or immediately following contaminant introduction, the surviving populations take up exogenous DNA from the environment that may allow rapid adaptation to specific contaminant profiles (Fig. 5C). In the case of *Rhodanobacter* populations, metal resistance genes, such as *czc* and mercuric resistance operons, appear to be the primary targets of lateral transfer and acquisition. Much of this activity appears to occur in the identified recombinational hot spot of the *Rhodanobacter* pangenome. The process of lateral gene transfer is likely ongoing, but once established by subsequent rounds of selection, new populations of *Rhodanobacter* may emerge adapted to highly specific niches within the contaminated-groundwater environment. This is consistent with some recent suggestions that ecologically determined gene transfer networks may exist in microbial communities, allowing rapid sharing of niche-adaptive genes in the flexible (or accessory) genome, and that these networks are influenced by recombination rates within populations (12). The exact nature of the vector for LGT is unknown, but circumstantial evidence from the comparative genomics analyses and other work

TABLE 2 Lateral gene transfer and gene duplication events in *Rhodanobacter* species relevant to geochemical resistance in the FW106 contaminated-groundwater metagenome

Hypothesis	Event[a]	Stressor	Gene or product	Representative CDS(s)[b]
I. No selection in response to presence of stressors, no LGT/GD	VI	None	Housekeeping genes	
IIa. Selection in response to presence of stressors, no LGT/GD	VI	Heavy metals	High-affinity Fe^{2+}/Pb^{2+} permease	R2APBS1_2568
	VI	Heavy metals	Cytochrome c_{553}	R2APBS1_0257
IIb. Selection in response to lack of stressors, no LGT/GD	GL	Nitrate/nitrite	Nitrate/nitrite transporter	R2APBS1_2892
IIIa. Selection in response to presence of stressors, LGT/GD predating contamination	LGT	Heavy metals	High-affinity Fe^{2+}/Pb^{2+} permease	R2APBS1_1742
	LGT	Heavy metals (mercury)	Mercuric resistance operon genes (original)	R2APBS1_1850–1856,[c] R2APBS1_2076–2071,[c] R2APBS1_2095–2101c
IIIb. Selection in response to presence of stressors, LGT/GD following contamination	GD	Heavy metals (mercury)	Mercuric resistance operon genes (duplicates)	R2APBS1_1850–1856,[c] R2APBS1_2076–2071,[c] R2APBS1_2095–2101c
	LGT	Heavy metals (divalent cations)	$Co^{2+}/Zn^{2+}/Cd^{2+}$ efflux genes	R2APBS1_1735–1751
	LGT	Heavy metals (arsenic)	Arsenate reductase operon genes	R2APBS1_1729–1733
	LGT	Nitrate/nitrite	Clade 2b nitrate/nitrite transporters	RhoFW510R12_08626
IV. Abundant genes that cannot currently be characterized		Unknown	NiFe hydrogenase III (large and small subunits)	R2APBS1_2173–2174

[a] VI, vertical inheritance; LGT, lateral gene transfer; GD, gene duplication; GL, gene loss.

[b] The locus tags of representatives of the gene families in the reference genome *R. denitrificans* 2APBS1 are listed.

[c] It is unclear which is original and which are duplicates.

suggests that viral and/or plasmid vectors are the primary means of LGT. It remains to be seen if plasmids specific to *Rhodanobacter* exist or if *Rhodanobacter* species are capable of conjugation. It appears that *Rhodanobacter* may be a key node in such a network in contaminated-groundwater environments.

A second interesting question regards whether the observed LGT/GD events preceded contamination of the site or whether these are recent changes specifically in response to contamination. Some metal resistance genes, for instance, are conserved in most of the strains, suggesting they were acquired prior to contamination. On the other hand, the OR-IFRC strains often show genomic properties distinct from those of the reference strains isolated outside the OR-IFRC sites, particularly in terms of their complement of metal resistance genes. While population structures are evident in the OR-IFRC strains, it is still difficult to determine whether these represent legacy (fixed in the population prior to contamination) or opportunistic (fixed in the population following contamination) events. While definitive evidence of lateral gene transfer is difficult to identify by using sequence data alone, the biogeography and population structures of *Rhodanobacter* within and outside the OR-IFRC show compelling evidence of contamination-driven lateral gene transfer and gene duplication. In particular, the apparent duplication of mercuric resistance operons in *R. denitrificans* species provides the best evidence to date of a genome-scale evolutionary event specifically in response to contamination.

Another interesting question is whether the proposed model of *Rhodanobacter* may be extended to other bacterial populations at the OR-IFRC sites. For instance, do other naturally resistant OR-IFRC populations, such as *Pseudomonas* and *Stenotrophomonas*, exhibit the same selection-LGT-selection dynamics as *Rhodanobacter*? If so, do these populations employ the same mechanisms for accumulating these genes? Analysis of the OR-IFRC background groundwater metagenome from the OR-IFRC background site suggests a reservoir of mobile geochemical resistance genes in groundwater communities (34). If multiple populations in the contaminated-groundwater source zone are drawing from the same pool of mobile genes, different OR-IFRC populations might show similar LGT patterns. Alternatively, individual population genome dynamics may affect the rate of gene transfer and fixation in specific populations, which in turn might affect their relative fitness in these environments. As more OR-IFRC genomes become available, these hypotheses can be tested directly.

The results of this analysis help explain why *Rhodanobacter* species are dominant in many OR-IFRC contaminated environments, but some questions remain. It is known, for instance, that *Rhodanobacter* populations exist in uncontaminated background communities, albeit in extremely low abundances. It is unknown whether the introduction of contaminants allows opportunistic *Rhodanobacter* populations to grow to dominance or whether these environments are colonized by exogenous *Rhodanobacter* populations. It is also not clear why *Rhodanobacter* populations dominate certain contaminated environments and not others. That is, do specific environments select specific populations, or are the dominance patterns observed at the OR-IFRC the result of stochastic effects (54–57)? Further study of this phenomenon may help explain why other, seemingly equally adapted populations are not able to outcompete *Rhodanobacter* at specific sites and vice versa.

In conclusion, comparative genome analysis of *Rhodanobacter*

isolates and metagenomes from contaminated-groundwater sites clarifies the role and extent of LGT and GD in the accumulation of geochemical resistance genes in the population. Metal resistance genes in particular are highly susceptible to lateral transfer and have accumulated in multiple *Rhodanobacter* species due to a putative recombinational hot spot in the *Rhodanobacter* pangenome. This ability to rapidly acquire metal resistance genes via LGT likely explains in part the ability of *Rhodanobacter* to dominate uranium-contaminated sites.

MATERIALS AND METHODS

Comparison of *Rhodanobacter* 16S rRNA gene sequences to the OR-IFRC 100-well-survey 16S rRNA gene data. One hundred wells from various areas selected in the OR-IFRC research area were recently sampled for 16S rRNA gene sequence diversity (45). OR-IFRC sequences were obtained using MiSeq as previously described (58). Aligned *Rhodanobacter* 16S sequences were obtained from the Ribosomal Database (46). *Rhodanobacter* and OF-IFRC samples were compared using BLASTn (47). OR-IFRC sequences resulting in BLAST hits with >97% sequence identity were extracted and counted using *ad hoc* Perl scripts.

Abundance profiling of reference genomes and metagenomes. Abundance profiles for the two metagenomes, available *Rhodanobacter* genomes, and all available bacterial reference genomes based on COG categories were downloaded from the IMG website (59–62). For other *Rhodanobacter* genomes, the COG abundance information was extracted directly from the associated GenBank files using *ad hoc* Perl scripts. Odds ratios (OR) for each COG were calculated using the equation OR = (A/B)/(C/D), based on the following contingency tables: A is the number of genes in the genome/metagenome of interest assigned to COG category i, B is the number of genes in the genome/metagenome assigned to all COG categories, C is the number of genes from all sequenced bacterial genomes assigned to COG category i, and D is the number of genes from all sequenced bacterial genomes assigned to COG categories. Odds ratios and significance tests were calculated using *ad hoc* Perl scripts. If C or D equaled 0, the value was changed to 1 to allow for calculation. Each contingency table was tested for significance using the χ^2 test (df = 1; $P \leq$ 0.05), and the right-tailed Fisher exact test ($P \leq 0.05$) was used to identify COG categories overabundant in the selected genome/metagenome based on all sequenced bacterial genomes. The ln odds ratio values were plotted as heat maps using R version 2.15.0 (63).

***Rhodanobacter* strain isolation, sequencing, assembly, and annotation.** Groundwater samples from wells FW104 and FW510 were collected on 14 November 2012 and 26 November 2012, respectively. The samples were shipped on ice overnight. Thirty-microliter aliquots of the groundwater samples were streaked on Reasoner's 2A (R2A) agar and tryptic soy agar (TSA) plates. Samples from FW510 were streaked on R2A and TSA plates amended with 125 μM Mn. Colonies developed within a week of incubation at room temperature. Individual colonies were picked, restreaked for purity tests, and regrown in the corresponding liquid medium. Overnight liquid cultures were used to extract DNA for 16S-based identification.

In addition to the nine publicly available *Rhodanobacter* genomes (including six OR-IFRC isolates) (24, 36), seven additional OR-IFRC strains were isolated from wells FW104 (four from area 3) and FW510 (three from Y12 West), which were sequenced using Illumina MiSeq at the Institute for Environmental Genomics, University of Oklahoma. In total, 16 *Rhodanobacter* genomes (13 OR-IFRC strains and 3 reference strains) were available for comparative analysis. DNA libraries were prepared using the Nextera DNA library preparation kit (Illumina, Inc., San Diego, CA) according to the manufacturer's protocol (October 2011 version). Briefly, 50 ng DNA (20 μl at 2.5 ng/μl) was fragmented using 5 μl of Tagment DNA enzyme with 25 μl of Tagment DNA buffer. Tagmentation reactions were performed by incubation at 55°C for 5 min, followed by purification of the tagmented DNA using the Zymo Clean & Concentrator-5 kit (Zymo Research, Orange, CA). Purified DNA was

eluted from the column with 25 μl of resuspension buffer. Illumina adapter and index sequences were added to the purified tagmented DNA (20 μl) by limited-cycle PCR (5 cycles) amplification with index 1 and 2 primers in a 50-μl reaction mixture volume according to the Nextera protocol. Amplified DNA was purified using 30 μl of AMPure XP beads (Beckman Coulter, Inc., Australia). The fragment size distribution of the tagmented DNA was analyzed utilizing a 2100 Bioanalyzer with a high-sensitivity DNA assay kit (Agilent Technologies, Santa Clara, CA). DNA libraries, pooled in equal volumes, were normalized to 2 nM. Sample libraries for sequencing were prepared according to the MiSeq reagent kit preparation guide (Illumina, San Diego, CA). Briefly, the pooled sample library (2 nM) was denatured by mixing 10 μl of the library and 10 μl of 0.2 N fresh NaOH and incubating for 5 min at room temperature. Nine hundred eighty microliters of chilled Illumina HT1 buffer was added to the denatured DNA and mixed to make a 20 pM library. The 20 pM library was further adjusted to reach the desired concentration for sequencing: for example, 625 μl of the 20 pM library was mixed with 375 μl of chilled Illumina HT1 buffer to make a 12.5 pM library. The library for sequencing was mixed with about 10% of the Phix library at the same concentration. A 500-cycle version 2 MiSeq reagent cartridge (Illumina) was thawed for 1 h in a water bath, inverted 10 times to mix the thawed reagents, and stored at 4°C for a short time until use. Sequencing was performed for 251, 12, and 251 cycles for forward, index, and reverse reads, respectively, on the MiSeq instrument.

Read libraries were quality trimmed using Btrim (Btrim64 -S -q -t <file> -o <output> -a 20 -w 4 -l 80) (64), and the resulting libraries were assembled using SOAPdenovo (65) using a range of kmer sizes (37, 55, 67, and 75 bp) (SOAPdenovo-127mer all -S <config> -K <kmer> -V -o <prefix> -R). For each genome, the kmer size that minimized the number of contigs and scaffolds and maximized the N_{50} lengths and the lengths of the contigs and scaffolds was chosen for annotation (see Table S3A in the supplemental material). Assembled sequences were uploaded to the IMG expert review pipeline for annotation (see Table S3B) (66). Four of the strains (FW104-R5, FW104-T7, FW510-T8, and FW510-R10) were assembled to draft quality, while the remaining three (FW104-R3, FW510-R10, and FW510-R12) assembled poorly.

Phylogenetic tree construction. For most of the phylogenetic trees, protein alignments were constructed in MEGA 5.2 (67) using Muscle (68). The phylogenetic trees were constructed using the maximum-likelihood method based on the JTT matrix-based model (69) using a discrete gamma distribution with invariable sites allowed to model rate differences among sites. To analyze the distribution of *merA* genes in the OR-IFRC metagenomes, reference MerA protein sequences were obtained for sequenced reference genomes (including all available strains of *Rhodanobacter*) from GenBank. Only those sequences most closely corresponding to true MerA enzymes based on previous studies were maintained (70). A reference tree was created in MEGA 5.1 (67). The evolutionary history was inferred by using the maximum-likelihood method based on the JTT matrix-based model (69). The tree with the highest log likelihood ($-12,814.2808$) is shown. The initial tree(s) for the heuristic search were obtained automatically by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology with superior log likelihood value. A discrete gamma distribution was used to model evolutionary rate differences among sites (5 categories [$+G$, parameter $= 1.1818$]). The rate variation model allowed for some sites to be evolutionarily invariable ([$+I$], 5.0071% sites). The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 199 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, less than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 144 positions in the final data set. The bootstrap consensus tree was derived from 500 replicates (71). Branches corresponding to partitions that were reproduced in less than 50% of bootstrap replicates were collapsed. Metagenome sequences corresponding to putative MerA proteins were iden-

tified by BLASTp analysis (47). Because many of the metagenome fragments did not overlap, a traditional phylogenetic tree could not be constructed. Therefore, metagenome fragments were aligned to the reference alignment using the MAFFT multisequence alignment tool (72) using the following command: mafft --retree 0 --treeout --reorder --localpair --weighti 0 --averagelinkage input. The final tree was formatted for publication using iTOL version 2 (73).

Detection of LGT events. Whole-genome alignments were prepared using Mauve (74) to provide a reference visualization of LGT events. Four complementary methods were used to identify putative LGT events: phylogenetic comparison (identifying deviations from the 16S tree), Alien_Hunter (interpolating variable motifs based on compositional bias), DarkHorse (lineage probability index), and GOHTAM (local variations of genome signatures). Putative LGT events were detected using the GOHTAM software (50) with the sliding windows plus codon usage option and default options. Since the draft genomes consisted of multiple contigs, all contigs were concatenated into a single sequence with appropriate adjustment of feature coordinates prior to GOHTAM analysis. The complete (closed) *R. denitrificans* 2APBS1 genome was screened for the presence of any candidate laterally transferred genes via Alien_Hunter (49) and DarkHorse (48). The threshold score for the program Alien_Hunter was set as 20 instead of the default to minimize the number of false positives after observing the distribution of the scores. The resulting predictions were plotted, along with the genes involved in the denitrification pathway, heavy metal resistance, integrases, and transposases, using the program CGview (Fig. 1) (75). The program DarkHorse was built on the GenBank nr database downloaded on 30 July 2012 and was run with the default filtering threshold of 0.1 on two different "self" sequence definitions, (i) Genus: *Rhodanobacter*, and (ii) Species: *R. sp. 2APBS1*. Since the *Rhodanobacter* GenBank nr data are composed of sequences from the isolates from the same site that were found to display high similarity levels, the genus *Rhodanobacter* was also selected as "self" to avoid false negatives. A comparative distribution obtained upon using the different self terms is shown in Fig. S1 in the supplemental material.

Nucleotide sequence accession numbers. The *Rhodanobacter* species sequenced as part of this study were deposited in NCBI GenBank under BioSample accession numbers SAMN02929088, SAMN02929136, SAMN02929138, SAMN02929139, SAMN02929141, SAMN02929142, and SAMN02929143. The Whole Genome Shotgun projects have been deposited at DDBJ/ENA/GenBank under the accession numbers LVJP00000000-LVJV00000000. The versions described in this paper are versions LVJP01000000-LVJV01000000. All other sequences have been reported previously (1, 24, 34, 36). All sequences are also available through the JGI IMG database (http://img.jgi.doe.gov/).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.02234-15/-/DCSupplemental.

Figure S1, TIF file, 1 MB.
Figure S2, TIF file, 0.9 MB.
Figure S3, TIF file, 1.1 MB.
Figure S4, TIF file, 1.5 MB.
Figure S5, TIF file, 2.7 MB.
Figure S6, TIF file, 1.9 MB.
Figure S7, TIF file, 1.1 MB.
Table S1, DOCX file, 0.02 MB.
Table S2, DOCX file, 0.01 MB.
Table S3, DOCX file, 0.01 MB.

REFERENCES

1. **Hemme CL, Deng Y, Gentry TJ, Fields MW, Wu L, Barua S, Barry K, Tringe SG, Watson DB, He Z, Hazen TC, Tiedje JM, Rubin EM, Zhou J.** 2010. Metagenomic insights into evolution of a heavy metal-contaminated groundwater microbial community. ISME J **4:**660–672. http://dx.doi.org/10.1038/ismej.2009.154.

2. **Aminov RI.** 2011. Horizontal gene exchange in environmental microbiota. Front Microbiol **2:**158. http://dx.doi.org/10.3389/fmicb.2011.00158.

3. **Dougherty K, Smith BA, Moore AF, Maitland S, Fanger C, Murillo R, Baltrus DA.** 2014. Multiple phenotypic changes associated with large-scale horizontal gene transfer. PLoS One **9:**e102170. http://dx.doi.org/10.1371/journal.pone.0102170.

4. **Casacuberta E, González J.** 2013. The impact of transposable elements in environmental adaptation. Mol Ecol **22:**1503–1517. http://dx.doi.org/10.1111/mec.12170.

5. **Baltrus DA.** 2013. Exploring the costs of horizontal gene transfer. Trends Ecol Evol **28:**489–495. http://dx.doi.org/10.1016/j.tree.2013.04.002.

6. **Aminov RI.** 2012. The extent and regulation of lateral gene transfer in natural microbial ecosystems, p 93–130. *In* Francino MP (ed), Horizontal gene transfer in microorganisms. Caister Academic press, Norfolk, United Kingdom.

7. **Liang B, Wang G, Zhao Y, Chen K, Li S, Jiang J.** 2011. Horizontal transfer of the chlorothalonil hydrolytic dehalogenase gene facilitates bacterial adaptation to chlorothalonil-contaminated sites. Appl Environ Microbiol **77:**4268–4272. http://dx.doi.org/10.1128/AEM.02457-10.

8. **Martinez RJ, Wang Y, Raimondo MA, Coombs JM, Barkay T, Sobecky PA.** 2006. Horizontal gene transfer of PIB-type ATPases among bacteria isolated from radionuclide- and metal-contaminated subsurface soils. Appl Environ Microbiol **72:**3111–3118. http://dx.doi.org/10.1128/AEM.72.5.3111-3118.2006.

9. **Nongkhlaw M, Kumar R, Acharya C, Joshi SR.** 2012. Occurrence of horizontal gene transfer of PIB-type ATPase genes among bacteria isolated from an uranium rich deposit of Domiasiat in North East India. PLoS One **7:**e48199. http://dx.doi.org/10.1371/journal.pone.0048199.

10. **Roberts AP, Kreth J.** 2014. The impact of horizontal gene transfer on the adaptive ability of the human oral microbiome. Front Cell Infect Microbiol **4:**124. http://dx.doi.org/10.3389/fcimb.2014.00124.

11. **Sobecky PA, Coombs JM.** 2009. Horizontal gene transfer in metal and radionuclide contaminated soils. Methods Mol Biol **532:**455–472. http://dx.doi.org/10.1007/978-1-60327-853-9_26.

12. **Polz MF, Alm EJ, Hanage WP.** 2013. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. Trends Genet **29:**170–175. http://dx.doi.org/10.1016/j.tig.2012.12.006.

13. **Committee on Metagenomics: Challenges and Functional Applications, National Research Council.** 2007. The new science of metagenomics: revealing the secrets of our microbial planet. National Academies Press, Washington, DC.

14. **Human Microbiome Jumpstart Reference Strains Consortium, Nelson KE, Weinstock GM, Highlander SK, Worley KC, Creasy HH, Wortman JR, Rusch DB, Mitreva M, Sodergren E, Chinwalla AT, Feldgarden M, Gevers D, Haas BJ, Madupu R, Ward DV, Birren BW, Gibbs RA, Methe B, Petrosino JF.** 2010. A catalog of reference genomes from the human microbiome. Science **328:**994–999. http://dx.doi.org/10.1126/science.1183605.

15. **Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, Hooper SD, Pati A, Lykidis A, Spring S, Anderson IJ, D'haeseleer P, Zemla A, Singer M, Lapidus A, Nolan M.** 2009. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature **462:**1056–1060. http://dx.doi.org/10.1038/nature08656.

16. **Kyrpides NC, Woyke T, Eisen JA, Garrity G, Lilburn TG, Beck BJ, Whitman WB, Hugenholtz P, Klenk H-P.** 2014. Genomic encyclopedia of type strains, phase I: the one thousand microbial genomes (KMG-I) project. Stand Genomic Sci **9:**1278–1284. http://dx.doi.org/10.4056/sigs.5068949.

17. **Chen T, Yu W-H, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE.** 2010. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. Database (Oxford) 2010:baq013. http://dx.doi.org/10.1093/database/baq013.

18. **Friedman J, Alm EJ.** 2012. Inferring correlation networks from genomic survey data. PLOS Comput Biol **8:**e1002687. http://dx.doi.org/10.1371/journal.pcbi.1002687.

19. **Liu L, Chen X, Skogerbø G, Zhang P, Chen R, He S, Huang D-W.** 2012. The human microbiome: a hot spot of microbial horizontal gene transfer. Genomics **100:**265–270. http://dx.doi.org/10.1016/j.ygeno.2012.07.012.

20. **Meehan CJ, Beiko RG.** 2012. Lateral gene transfer of an ABC transporter complex between major constituents of the human gut microbiome. BMC Microbiol **12:**248. http://dx.doi.org/10.1186/1471-2180-12-248.

21. **Smillie CS, Smith MB, Friedman J, Cordero OX, David LA, Alm EJ.** 2011. Ecology drives a global network of gene exchange connecting the human microbiome. Nature **480:**241–244. http://dx.doi.org/10.1038/nature10571.

22. **Sommer MO, Church GM, Dantas G.** 2010. The human microbiome harbors a diverse reservoir of antibiotic resistance genes. Virulence **1:**299–303. http://dx.doi.org/10.4161/viru.1.4.12010.

23. **Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI.** 2007. The Human Microbiome Project. Nature **449:**804–810. http://dx.doi.org/10.1038/nature06244.

24. **Bollmann A, Palumbo AV, Lewis K, Epstein SS.** 2010. Isolation and physiology of bacteria from contaminated subsurface sediments. Appl Environ Microbiol **76:**7413–7419. http://dx.doi.org/10.1128/AEM.00376-10.

25. **Van Nostrand JD, Wu L, Wu W-M, Huang Z, Gentry TJ, Deng Y, Carley J, Carroll S, He Z, Gu B, Luo J, Criddle CS, Watson DB, Jardine PM, Marsh TL, Tiedje JM, Hazen TC, Zhou J.** 2011. Dynamics of microbial community composition and function during *in situ* bioremediation of a uranium-contaminated aquifer. Appl Environ Microbiol **77:** 3860–3869. http://dx.doi.org/10.1128/AEM.01981-10.

26. **Van Nostrand JD, Wu WM, Wu L, Deng Y, Carley J, Carroll S, He Z, Gu B, Luo J, Criddle CS, Watson DB, Jardine PM, Marsh TL, Tiedje JM, Hazen TC, Zhou J.** 2009. GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer. Environ Microbiol **11:**2611–2626. http://dx.doi.org/10.1111/j.1462-2920.2009.01986.x.

27. **Xu M, Wu W-M, Wu L, He Z, Van Nostrand JD, Deng Y, Luo J, Carley J, Ginder-Vogel M, Gentry TJ, Gu B, Watson D, Jardine PM, Marsh TL, Tiedje JM, Hazen T, Criddle CS, Zhou J.** 2010. Responses of microbial community functional structures to pilot-scale uranium in situ bioremediation. ISME J **4:**1060–1070. http://dx.doi.org/10.1038/ismej.2010.31.

28. **Fields MW, Yan T, Rhee SK, Carroll SL, Jardine PM, Watson DB, Criddle CS, 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. http://dx.doi.org/10.1016/j.femsec.2005.01.010.

29. **Green SJ, Prakash O, Gihring TM, Akob DM, Jasrotia P, Jardine PM, Watson DB, Brown SD, Palumbo AV, Kostka JE.** 2010. Denitrifying bacteria isolated from terrestrial subsurface sediments exposed to mixed-waste contamination. Appl Environ Microbiol **76:**3244–3254. http://dx.doi.org/10.1128/AEM.03069-09.

30. **Green SJ, Prakash O, Jasrotia P, Overholt WA, Cardenas E, Hubbard D, Tiedje JM, Watson DB, Schadt CW, Brooks SC, Kostka JE.** 2012. Denitrifying bacteria from the genus *Rhodanobacter* dominate bacterial communities in the highly contaminated subsurface of a nuclear legacy waste site. Appl Environ Microbiol **78:**1039–1047. http://dx.doi.org/10.1128/AEM.06435-11.

31. **Kostka JE, Green SJ, Rishishwar L, Prakash O, Katz LS, Mariño-Ramírez L, Jordan IK, Munk C, Ivanova N, Mikhailova N, Watson DB, Brown SD, Palumbo AV, Brooks SC.** 2012. Genome sequences for six Rhodanobacter strains, isolated from soils and the terrestrial subsurface, with variable denitrification capabilities. J Bacteriol **194:**4461–4462. http://dx.doi.org/10.1128/JB.00871-12.

32. **Prakash O, Green SJ, Jasrotia P, Overholt WA, Canion A, Watson DB, Brooks SC, Kostka JE.** 2012. Description of *Rhodanobacter denitrificans* sp. nov., isolated from nitrate-rich zones of a contaminated aquifer. Int J Syst Evol Microbiol **62:**2457–2462 http://dx.doi.org/10.1099/ijs.0.035840-0.

33. **Van Den Heuvel RN, Van Der Biezen E, Jetten MS, Hefting MM, Kartal B.** 2010. Denitrification at pH 4 by a soil-derived Rhodanobacter-

dominated community. Environ Microbiol **12:**3264–3271. http://dx.doi.org/10.1111/j.1462-2920.2010.02301.x.

34. **Hemme CL, Tu Q, Shi Z, Qin Y, Gao W, Deng Y, Van Nostrand JD, Wu L, He Z, Chain PS, Tringe SG, Fields MW, Rubin EM, Tiedje JM, Hazen TC, Arkin AP, Zhou J.** 2015. Comparative metagenomics reveals impact of contaminants on groundwater microbiomes. Front Microbiol **6:**1205. http://dx.doi.org/10.3389/fmicb.2015.01205.

35. **Green SJ, Prakash O, Jasrotia P, Overholt WA, Cardenas E, Hubbard D, Tiedje JM, Watson DB, Schadt CW, Brooks SC, Kostka JE.** 2012. Denitrifying bacteria from the genus Rhodanobacter dominate bacterial communities in the highly contaminated subsurface of a nuclear legacy waste site. Appl Environ Microbiol **78:**1039–1047. http://dx.doi.org/10.1128/AEM.06435-11.

36. **Kostka JE, Green SJ, Rishishwar L, Prakash O, Katz LS, Mariño-Ramírez L, Jordan IK, Munk C, Ivanova N, Mikhailova N, Watson DB, Brown SD, Palumbo AV, Brooks SC.** 2012. Genome sequences for six Rhodanobacter strains, isolated from soils and the terrestrial subsurface, with variable denitrification capabilities. J Bacteriol **194:**4461–4462. http://dx.doi.org/10.1128/JB.00871-12.

37. **Prakash O, Green SJ, Jasrotia P, Overholt WA, Canion A, Watson DB, Brooks SC, Kostka JE.** 2012. Rhodanobacter denitrificans sp. nov., isolated from nitrate-rich zones of a contaminated aquifer. Int J Syst Evol Microbiol **62:**2457–2462. http://dx.doi.org/10.1099/ijs.0.035840-0.

38. **Brown SD, Palumbo AV, Panikov N, Ariyawansa T, Klingeman DM, Johnson CM, Land ML, Utturkar SM, Epstein SS.** 2012. Draft genome sequence for Microbacterium laevaniformans strain OR221, a bacterium tolerant to metals, nitrate, and low pH. J Bacteriol **194:**3279–3280. http://dx.doi.org/10.1128/JB.00474-12.

39. **Spain AM, Peacock AD, Istok JD, Elshahed MS, Najar FZ, Roe BA, White DC, Krumholz LR.** 2007. Identification and isolation of a Castellaniella species important during biostimulation of an acidic nitrate- and uranium-contaminated aquifer. Appl Environ Microbiol **73:**4892–4904. http://dx.doi.org/10.1128/AEM.00331-07.

40. **Cardenas E, Wu W-M, Leigh MB, Carley J, Carroll S, Gentry T, Luo J, Watson D, Gu B, Ginder-Vogel M, Kitanidis PK, Jardine PM, Zhou J, Criddle CS, Marsh TL, Tiedje JM.** 2010. Significant association between sulfate-reducing bacteria and uranium-reducing microbial communities as revealed by a combined massively parallel sequencing-indicator species approach. Appl Environ Microbiol **76:**6778–6786. http://dx.doi.org/10.1128/AEM.01097-10.

41. **Moreels D, Crosson G, Garafola C, Monteleone D, Taghavi S, Fitts JP, 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. http://dx.doi.org/10.1007/s11356-008-0034-z.

42. **Spain AM, Krumholz LR.** 2011. Nitrate-reducing bacteria at the nitrate and radionuclide contaminated Oak Ridge Integrated Field Research Challenge site: a review. Geomicrobiol J **28:**418–429. http://dx.doi.org/10.1080/01490451.2010.507642.

43. **Bouzat JL, Hoostal MJ.** 2013. Evolutionary analysis and lateral gene transfer of two-component regulatory systems associated with heavy-metal tolerance in bacteria. J Mol Evol **76:**267–279. http://dx.doi.org/10.1007/s00239-013-9558-z.

44. **Boyd ES, Barkay T.** 2012. The mercury resistance operon: from an origin in a geothermal environment to an efficient detoxification machine. Front Microbiol **3:**349. http://dx.doi.org/10.3389/fmicb.2012.00349.

45. **Smith MB, Rocha AM, Smillie CS, Olesen SW, Paradis C, Wu L, Campbell JH, Fortney JL, Mehlhorn TL, Lowe KA, Earles JE, Phillips J, Techtmann SM, Joyner DC, Elias DA, Bailey KL, Hurt RA, Preheim SP, Sanders MC, Yang J.** 2015. Natural bacterial communities serve as quantitative geochemical biosensors. mBio **6:**e00326-15. http://dx.doi.org/10.1128/mBio.00326-15.

46. **Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM.** 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res **37:**D141–D145. http://dx.doi.org/10.1093/nar/gkn879.

47. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389–3402. http://dx.doi.org/10.1093/nar/25.17.3389.

48. **Podell S, Gaasterland T.** 2007. DarkHorse: a method for genome-wide

49. prediction of horizontal gene transfer. Genome Biol **8:**R16. http://dx.doi.org/10.1186/gb-2007-8-2-r16.

49. **Vernikos GS, Parkhill J.** 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the salmonella pathogenicity islands. Bioinformatics **22:**2196–2203. http://dx.doi.org/10.1093/bioinformatics/btl369.

50. **Ménigaud S, Mallet L, Picord G, Churlaud C, Borrel A, Deschavanne P.** 2012. GOHTAM: a website for "genomic origin of horizontal transfers, alignment and metagenomics." Bioinformatics **28:**1270–1271. http://dx.doi.org/10.1093/bioinformatics/bts118.

51. **Dufraigne C, Fertil B, Lespinats S, Giron A, Deschavanne P.** 2005. Detection and characterization of horizontal transfers in prokaryotes using genomic signature. Nucleic Acids Res **33:**e6. http://dx.doi.org/10.1093/nar/gni004.

52. **Caldwell BA.** 2005. Enzyme activities as a component of soil biodiversity: a review. Pedobiologia **49:**637–644. http://dx.doi.org/10.1016/j.pedobi.2005.06.003.

53. **Shelobolina ES, Sullivan SA, O'Neill KR, Nevin KP, Lovley DR.** 2004. Isolation, characterization, and U(VI)-reducing potential of a facultatively anaerobic, acid-resistant bacterium from low-pH, nitrate- and U(VI)-contaminated subsurface sediment and description of Salmonella subterranea sp. nov. Appl Environ Microbiol **70:**2959–2965. http://dx.doi.org/10.1128/AEM.70.5.2959-2965.2004.

54. **Zhou J, Deng Y, Zhang P, Xue K, Liang Y, Van Nostrand JD, Yang Y, He Z, Wu L, Stahl DA, Hazen TC, Tiedje JM, Arkin AP.** 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proc Natl Acad Sci U S A **111:**E836–E845. http://dx.doi.org/10.1073/pnas.1324044111.

55. **Zhou J, Liu W, Deng Y, Jiang Y-H, Xue K, He Z, Van Nostrand JD, Wu L, Yang Y, Wang A.** 2013. Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. mBio **4:**e00584-12. http://dx.doi.org/10.1128/mBio.00584-12.

56. **Stegen JC, Lin X, Konopka AE, Fredrickson JK.** 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. ISME J **6:**1653–1664. http://dx.doi.org/10.1038/ismej.2012.22.

57. **Wang J, Shen J, Wu Y, Tu C, Soininen J, Stegen JC, He J, Liu X, Zhang L, Zhang E.** 2013. Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. ISME J **7:**1310–1321. http://dx.doi.org/10.1038/ismej.2013.30.

58. **Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R.** 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J **6:**1621–1624. http://dx.doi.org/10.1038/ismej.2012.8.

59. **Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC.** 2012. IMG: the Integrated Microbial Genomes database and comparative analysis system. Nucleic Acids Res **40:**D115–D122. http://dx.doi.org/10.1093/nar/gkr1044.

60. **Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Anderson I, Lykidis A, Mavromatis K, Ivanova NN, Kyrpides NC.** 2010. The Integrated Microbial Genomes system: an expanding comparative analysis resource. Nucleic Acids Res **38:**D382–D390. http://dx.doi.org/10.1093/nar/gkp887.

61. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278 http://dx.doi.org/10.1093/bioinformatics/btp393:btp393.

62. **Markowitz VM, Ivanova NN, Szeto E, Palaniappan K, Chu K, Dalevi D, Chen I-MA, Grechkin Y, Dubchak I, Anderson I, Lykidis A, Mavromatis K, Hugenholtz P, Kyrpides NC.** 2008. IMG/M: a data management and analysis system for metagenomes. Nucleic Acids Res **36:**D534–D538. http://dx.doi.org/10.1093/nar/gkm869.

63. **Ihaka R, Gentleman R.** 1996. R: a language for data analysis and graphics. J Comput Graph Statist **5:**299–314. http://dx.doi.org/10.1080/10618600.1996.10474713.

64. **Kong Y.** 2011. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. Genomics **98:**152–153. http://dx.doi.org/10.1016/j.ygeno.2011.05.009.

65. **Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW.** 2012. SOAPdenovo2: an empirically improved memory-

efficient short-read de novo assembler. Gigascience **1:**18. http://dx.doi.org/10.1186/2047-217X-1-18.

66. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC**. 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278. http://dx.doi.org/10.1093/bioinformatics/btp393.

67. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S**. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:**2731–2739. http://dx.doi.org/10.1093/molbev/msr121.

68. **Edgar RC**. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5:**113. http://dx.doi.org/10.1186/1471-2105-5-113.

69. **Jones DT, Taylor WR, Thornton JM**. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci **8:**275–282. http://dx.doi.org/10.1093/bioinformatics/8.3.275.

70. **Barkay T, Kritee K, Boyd E, Geesey G**. 2010. A thermophilic bacterial origin and subsequent constraints by redox, light and salinity on the evolution of the microbial mercuric reductase. Environ Microbiol **12:**2904–2917. http://dx.doi.org/10.1111/j.1462-2920.2010.02260.x.

71. **Felsenstein J**. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791. http://dx.doi.org/10.2307/2408678.

72. **Katoh K, Standley DM**. 2013. MAFFT multiple sequence alignment software version 7. Improvements in performance and usability. Mol Biol Evol **30:**772–780.

73. **Letunic I, Bork P**. 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res **39:** W475–W478 http://dx.doi.org/10.1093/nar/gkr201.

74. **Darling AE, Mau B, Perna NT**. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.

75. **Stothard P, Wishart DS**. 2005. Circular genome visualization and exploration using CGView. Bioinformatics **21:**537–539. http://dx.doi.org/10.1093/bioinformatics/bti054.