Insights into the Metabolism and Evolution of the Genus *Acidiphilium*, a Typical Acidophile in Acid Mine Drainage

Liangzhi Li,a,b Zhenghua Liu,a,b Min Zhang,a,b Delong Meng,a,b Xueduan Liu,a,b Pei Wang,c Xiutong Li,c Zhen Jiang,c Shuiping Zhong,c,e Chengying Jiang,c Huaqun Yin,a,b

a School of Minerals Processing and Bioengineering, Central South University, Changsha, China
b Key Laboratory of Biometallurgy of Ministry of Education, Central South University, Changsha, China
c State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China
d College of Zijin Mining, Fuzhou University, Fuzhou, China
e National Key Laboratory of Comprehensive Utilization of Low-Grade Refractory Gold Ores, Shanghang, China

**ABSTRACT** Here, we report three new *Acidiphilium* genomes, reclassified existing *Acidiphilium* species, and performed the first comparative genomic analysis on *Acidiphilium* in an attempt to address the metabolic potential, ecological functions, and evolutionary history of the genus *Acidiphilium*. In the genomes of *Acidiphilium*, we found an abundant repertoire of horizontally transferred genes (HTGs) contributing to environmental adaption and metabolic expansion, including genes conferring photosynthesis (*puf*, *puh*), *CO₂* assimilation (*rbc*), capacity for methane metabolism (*mmo*, *mdh*, *frm*), nitrogen source utilization (*nar*, *cyn*, *hmp*), sulfur compound utilization (*sox*, *psr*, *sqr*), and multiple metal and osmotic stress resistance capacities (*czc*, *cop*, *ect*). Additionally, the predicted donors of horizontal gene transfer were present in a cooccurrence network of *Acidiphilium*. Genome-scale positive selection analysis revealed that 15 genes contained adaptive mutations, most of which were multifunctional and played critical roles in the survival of extreme conditions. We proposed that *Acidiphilium* originated in mild conditions and adapted to extreme environments such as acidic mineral sites after the acquisition of many essential functions.

**IMPORTANCE** Extremophiles, organisms that thrive in extreme environments, are key models for research on biological adaption. They can provide hints for the origin and evolution of life, as well as improve the understanding of biogeochemical cycling of elements. Extremely acidophilic bacteria such as *Acidiphilium* are widespread in acid mine drainage (AMD) systems, but the metabolic potential, ecological functions, and evolutionary history of this genus are still ambiguous. Here, we sequenced the genomes of three new *Acidiphilium* strains and performed comparative genomic analysis on this extremely acidophilic bacterial genus. We found in the genomes of *Acidiphilium* an abundant repertoire of horizontally transferred genes (HTGs) contributing to environmental adaption and metabolic ability expansion, as indicated by phylogenetic reconstruction and gene context comparison. This study has advanced our understanding of microbial evolution and biogeochemical cycling in extreme niches.

**KEYWORDS** acid mine drainage, evolution, horizontal gene transfer, comparative genomics, *Acidiphilium*

P rokaryotes occupy almost all environmental niches and have dominated the majority of Earth’s evolutionary history. Extremophiles that thrive in extreme environments represent a key research field in many disciplines, ranging from the adaption to extreme conditions to the cycling of elements in biogeochemistry. Extremophiles also
have important implications for the research on the origin of life and the search for life on other planets (1, 2). Acid mine drainage (AMD), characterized by extreme acidity and high concentrations of metals and sulfate, represents an extreme ecological condition and a major global challenge (3). The primary microbial taxa in AMD include *Acidiphilium*, *Acidisphaera*, *Acidithiobacillus*, and *Leptospirillum* (4). The biological factors that contribute to the formation of this hyperacidic environment as well as the adaptive mechanisms of the organisms inhabiting it are hot topics in current research (3, 5). The genus *Acidiphilium* belongs to the family *Acetobacteraceae*, class *Rhodospirillales*, and appears frequently in AMD environments (6, 7). Members of this genus are Gram-negative, photosynthetic, aerobic and facultative anaerobic, metal-respiring, acidophilic heterotrophs (8–10). They grow at pH 1.5 to 7.5, are able to utilize a wide range of organic and inorganic substrates, and synthesize poly-β-hydroxybutyrate (PHB) for carbon storage (11–13). *Acidiphilium* can also resist multiple harmful stressors such as toxic metals (e.g., Cd, Ni, Cr) and osmotic pressure (14–16). There has also been increased interest in the application of *Acidiphilium* spp. for microbial fuel cells (MFCs) (17), as well as in metal mobilization from minerals or waste both in pure culture and in coculture (18–20). Nevertheless, little is known about whether there are divergences in the functional potential and niche partitioning among *Acidiphilium*-affiliated species. In addition, the evolutionary history of many notable properties such as carbon assimilation and metal resistance in *Acidiphilium* is still elusive. *Acidiphilium* is one of only four genera in the family *Acetobacteraceae* found in acidic mineral sites, with the other three genera being *Acidicaldus*, *Acidisphaera*, and *Acidocella* (21–24). Evolutionary drivers such as horizontal gene transfer (HGT) and selection pressure might have played their parts in the adaptive evolution of *Acidiphilium* that survives in harsh acidic mineral conditions. However, their relative contributions are still ambiguous. Horizontal gene transfer refers to the acquisition of genetic elements from distant lineages for genetic and phenotypic innovations, a process contributing significantly to evolution within challenging environments and during global geologic and/or climatic events (25, 26). Positive selection, on the other hand, mediating survival fitness by adaptive mutations, has also been an indispensable driving force in microbial evolution, and recent investigations have shifted from testing selection on individual genes to the entire genomes (27–30).

To assess the differences in metabolic capacity and niche adaption potential among *Acidiphilium* species, and to unravel the evolutionary history of many fundamental genetic properties of *Acidiphilium*, we performed whole-genome sequencing of three novel strains of *Acidiphilium* isolated from two different AMD sites. Comparative genome analysis was carried out, focusing on understanding the roles of evolutionary processes in shaping the genomes of *Acidiphilium*. For this purpose, we conducted a detailed comparison of *Acidiphilium* species. We performed ancestral genomic reconstruction, cooccurrence analysis, and extensive phylogenetic analyses and explored the genomic arrangements of pathways of interest. We also focused on discovering genes under positive selection.

**RESULTS**

**Genomic features and reclassification of *Acidiphilium***. Three *Acidiphilium* genomes (AccI, AccII, and ZJSH63) were sequenced, resulting in a complete genome of strain AccI (a single chromosome and seven plasmids) and high-quality drafts of strains AccII and ZJSH63, according to the MISAG standards (31). The characteristics of these genomes and the other publicly available genomes of *Acidiphilium* spp. used in this study are shown in Table 1. The visualization of strain AccI chromosome (applying colors based on clusters of orthologous group [COG] classes) and comparative analysis of our three genomes were performed (see Fig. S1 at https://doi.org/10.6084/m9.figshare.12892016.v1). The genome size of *Acidiphilium* was about 4 Mbp. Although strains AccI and AccII were isolated from the same site, strain AccI shared more gene families with strain ZJSH63 than AccII. Strain AccI contained the most unique gene families among our three strains. COG annotations showed that AccII contained more
**TABLE 1** General features of bacterial genomes used in this study

| Organism and strain | GenBank/IMG-ER accession no. | Level | Contig (bp) | Complete (%) | Size (Mb) | Coding density (%) | GC (%) | Clade assigned | No. of genes | No. of proteins | Source | Geographic location |
|---------------------|-----------------------------|-------|-------------|--------------|-----------|-------------------|--------|----------------|--------------|-----------------|--------|---------------------|
| Acidiphilium sp. strain ZJSH63 | 2828882166 | Draft | 301 124,801 | 95.3 | 4.39 90.6 | 66.5 IV | 4.330 | 4.245 | Acid mine drainage | Fujian, China |
| Acidiphilium sp. strain Accl | 2824045439 | Complete | 4,036,204 7 | 100 | 4.18 90.6 | 66.7 IV | 4.304 | 4.224 | Acid mine drainage | Guangdong, China |
| Acidiphilium sp. strain AccII | 2824049744 | Draft | 716 93,659 | 98.0 | 4.69 89.4 | 65.5 IV | 4.352 | 4.347 | Acid mine drainage | Guangdong, China |
| Acidiphilium angustum ATCC 35903 | GCA_000701585.1/256151102 | Draft | 206 71,968 | 98.5 | 4.07 89.6 | 63.6 II | 3.851 | 3.731 | Acid mine | |
| Acidiphilium cryptum JF-5 | GCA_00016725.1/640477101 | Complete | 1 3,389,227 | 8 | 3.96 90.8 | 67.1 IV | 3.747 | 3.574 | Acid mine | |
| Acidiphilium multivorans AU301 | GCA_000202835.1 | Complete | 1 3,749,411 | 8 | 4.21 89.8 | 67.0 IV | 3.991 | 3.803 | Acid mine water | Iwate, Matsuo, Japan |
| Acidiphilium multivorans ATCC 35905 | GCA_000156265.1/2681812815 | Draft | 78 136,338 | 98.6 | 3.98 89.9 | 63.7 II | 3.752 | 3.692 | Acidic coal mine drainage | Pennsylvania, USA |
| Acidiphilium sp. strain 20-67-58 | GCA_002255515.1 | Draft | 119 80,126 | 98.6 | 3.41 89.4 | 66.6 III | 3.199 | 3.073 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 21-60-14 | GCA_002255745.1 | Draft | 132 63,048 | 98.6 | 3.06 90.8 | 60.2 I | 2.940 | 2.828 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 21-62-4 | GCA_002255456.1 | Draft | 433 1,584 | 16.2 | 0.69 85.0 | 61.7 II | 9.169 | 8.32 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 21-66-27 | GCA_00225545 | Draft | 542 1,752 | 11.5 | 1.02 82.6 | 65.7 III | 1.316 | 1.189 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 21-68-69 | GCA_00225545 | Draft | 884 2,201 | 44.6 | 1.81 82.3 | 67.9 III | 2.291 | 1.999 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 34-60-192 | GCA_00225465 | Draft | 89 65,514 | 83.8 | 3.11 90.6 | 60.1 | 3.029 | 2.695 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 34-64-41 | GCA_00225465 | Draft | 180 38,599 | 93.9 | 3.86 88.6 | 63.7 | 3.623 | 3.413 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 37-60-79 | GCA_002253955.1 | Draft | 86 72,002 | 96.6 | 3.07 90.9 | 60.0 | 2.916 | 2.791 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 37-64-53 | GCA_002253955.1 | Draft | 211 43,062 | 97.3 | 4.03 88.3 | 63.5 | 3.799 | 3.620 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain 37-67-22 | GCA_002253955.1 | Draft | 937 3,121 | 62.2 | 2.48 89.9 | 67.0 | 2.942 | 2.632 | Mine wastewater | Ontario, Canada |
| Acidiphilium sp. strain bin8_M5 | 2734482270 | Draft | 74 76,901 | 97.3 | 3.08 92.7 | 68.8 IV | 3.042 | 2.992 | Acid mine drainage | Guangdong, China |
| Acidiphilium sp. strain OAG727 | GCA_000487155.1 | Draft | 116 9,562 | 82.5 | 1.73 45.4 | — | 1.509 | 1.479 | Gut microbiota | |
| Acidiphilium sp. strain JI21A1 | GCA_000772405.1/2571042905 | Draft | 296 44,500 | 98.6 | 4.18 88.4 | 66.9 IV | 4.059 | 3.719 | Acid mine drainage | Lusatia, Germany, Europe |
| Acidiphilium sp. strain PM | GCA_00019295.2 | Draft | 627 12,446 | 91.2 | 3.93 86.4 | 66.4 IV | 3.908 | 3.859 | Acidic, metal-rich water | Rio Tinto, Spain, Europe |

a—, not available.
unique genes with adaptive functions than ZJSH63 and Accl, especially those related to COG category L (replication, recombination, and repair) and COG N (cell motility). State-of-the-art whole-genome average nucleotide identity (ANI) analysis (32) classified all Acidiphilium genomes into four clades (species) based on an ANI cutoff of 95% (see Fig. S2 at https://doi.org/10.6084/m9.figshare.12892016.v1). We found some disagreement between our ANI results and previous nomenclatures of the Acidiphilium strains (mainly based on 16S rRNA sequences) in GenBank/JGI-IMG databases (33–35). For example, our ANI results showed that strains of Acidiphilium cryptum and Acidiphilium multitorum, as well as Acidiphilium angustum and Acidiphilium rubrum, should be classified as the same species (ANI > 95%). The major problem with previous species classification, based on 16S rRNA gene sequencing, was the low resolution, as shown by the low bootstrap values of the phylogenetic tree constructed with 16S rRNA sequences (see Fig. S3 at https://doi.org/10.6084/m9.figshare.12892016.v1). This shortage might be overcome by ANI analysis (32). Thus, the genomes of Acidiphilium spp. were thereafter referred to as clades I to IV according to their new classification based on ANI (Table 1; see Fig. S2 at https://doi.org/10.6084/m9.figshare.12892016.v1). Strain CAG727 (GCA_000437515.1) was determined not to be a member of Acidiphilium, given that it was phylogenetically distant from other Acidiphilium strains (see Fig. S2 and S3 at https://doi.org/10.6084/m9.figshare.12892016.v1), and was therefore excluded from further analyses. The clustering of Acidiphilium strains based on ANI values was mostly congruent with their geographic locations. For example, strains of clades I, II, and III were isolated from North America, while strains of clade IV were isolated from Europe and East Asia (see Fig. S2 at https://doi.org/10.6084/m9.figshare.12892016.v1). Whole-genome synteny analysis of all available complete sequences of Acidiphilium (strains Accl, JF-5, and AIU301, which belong to the same clade) showed that nine conserved locally colinear blocks (LCBs) were present in these strains, but they differed in their order of arrangement and similarity (see Fig. S4A at https://doi.org/10.6084/m9.figshare.12892016.v1). A similarity-based whole-genome comparison of Acidiphilium spp. with strain Accl as the reference showed that many genomic regions were not common to all isolates, many of which harbored hypothetical proteins and mobile genetic elements (see Fig. S5 at https://doi.org/10.6084/m9.figshare.12892016.v1).

Core genome and pangenome of Acidiphilium. Twelve genomes of the genus Acidiphilium, with estimated completeness over 97.0%, were carefully chosen for further pangenome analysis and genomic content reconstruction. The phylogenetic trees based on the concatenated alignment of 133 core genes inferred with neighbor-joining (NJ) methods were congruent with that based on whole-genome sequences (Fig. 1A; see Fig. S3A at https://doi.org/10.6084/m9.figshare.12892016.v1). The pangenome of the 12 Acidiphilium strains possessed 8,845 gene families, while the core genome possessed 1,422 gene families accounting for only 16.1% of the pangenome (Fig. 1C). Core and pangenome analyses of the 12 Acidiphilium genomes revealed an “open” pangenome fitted into a power-law regression function \( P_c(n) = 3,533.18n^{0.375395} \), while the core genome was fitted into an exponential regression \( F_c(n) = 2,725.11e^{-0.073314n} \) (Fig. 1A). The open pangenome suggested that species have undergone considerable gene exchanging to extend their functional profiles (36). Functional COG annotation revealed that the core genome had a higher proportion of genes classified in COG categories J (translation, ribosomal structure, and biogenesis), C (energy production and conversion), O (posttranslational modification, protein turnover, chaperones), F (nucleotide transport and metabolism), and H (coenzyme transport and metabolism), all associated with basic biological functions. The accessory genome and strain-specific genes were biased toward COG categories G (carbohydrate transport and metabolism), L (replication, recombination, and repair), P (inorganic ion transport and metabolism), and N (cell motility) (Fig. 1D), which were probably related to the adaption of Acidiphilium to oligotrophic, metal-laden, and acidic environments that often cause DNA damage. Gene ontology (GO) enrichment analyses produced similar results (see Table S1 in the supplemental material). Detailed metabolic recon-
(A) The evolutionary timeline of *Acidiphilium* was estimated (left) using RelTime on top of the rooted NJ tree based on the concatenated alignment of 133 core genes. Ancestral genome content reconstruction of *Acidiphilium* was performed with Count software, and the color depth (Continued on next page)
FIG 2  Overview of metabolic potentials in *Acidiphilium* as predicted from genome annotation; core/specific metabolic features are shown using different colors, and pathways containing predicted horizontally transferred genes are marked with black rectangles.

Construction of *Acidiphilium* was also performed; the core/specific metabolic features are shown in Fig. 2 using different colors, and pathways containing predicted horizontally transferred genes are marked with black rectangles.

**MGEs and CRISPR-Cas systems.** Mobile genetic elements (MGEs), such as insertion sequences, transposases, genomic islands (GIs), plasmids, and phages, are known signals of HGT events, and the number of MGEs correlates positively with the frequency of HGT (37). MGEs in the genomes of *Acidiphilium* were identified in this study (Table S2). The average number of transposon sequences per genome was 307, with *A. multivorum* AIU301 harboring the greatest number (841). Members of the ISAi7,
ISGalb1, ISMex27, ISAan1, and ISAcr4 families were most common. The average number of sequences located in GIs per genome was 527, with *Acidiphilium* sp. strain ZJSH63 containing the most (1,055). The average number of prophages and/or prophage remnants per genome was 23, with *Acidiphilium* sp. strain Accl harboring the most (58, with a total size of 69.9 kb). The number of plasmids in *Acidiphilium* could reach eight (*A. cryptum* JF-5 and *A. multivorum* AIU301). The functional gene profiles of plasmids from these three completely sequenced strains were also compared (see Fig. S4B to D at https://doi.org/10.6084/m9.figshare.12892016.v1). Approximately 66 gene families were shared among plasmids from these three strains, and certain degrees of collinearity were observed. COG L (replication, recombination, and repair) functions were enriched in the plasmid genomes. Type I-C/E/V and II-C CRISPR-Cas systems were also found in *Acidiphilium* spp., with *Acidiphilium* sp. strain PM containing the most CRISPR-Cas-related genes or spacers (38). The abundant MGEs present in genomes of *Acidiphilium* indicated that HGT might have contributed significantly to the genomic evolution of *Acidiphilium* species during niche adaption, while the CRISPR-Cas system would also help protect the genomes of *Acidiphilium* by eliminating harmful genomic intrusion events, balancing genomic stability and functional investments (39). A recent study also revealed that spacer sequences of the CRISPR-Cas system could not only specify the targets of Cas nucleases but also facilitate HGT (40).

**Evolutionary analyses of *Acidiphilium***. The evolutionary timeline of *Acidiphilium* was also estimated on the rooted core gene tree (Fig. 1A). Overall, gene families undergoing gain events outnumbered those undergoing loss events by approximately three times (6,319 versus 2,231), and gene families undergoing expansion events outnumbered those undergoing contraction events by approximately 20 times (1,173 versus 55) in the genomes of *Acidiphilium*. Our analyses suggested that there has been an ongoing increase in genomic content throughout the evolutionary history of this genus, from an estimated 2,026 gene families in the common ancestor to over 3,000 gene families. Predicted gain events of over 400 gene families occurred at nodes 3, 4, 6, and 9, accounting for approximately 14% to 25% of gene families at the corresponding nodes. Of all gene families undergoing gain events, about half encoded hypothetical proteins. A considerable proportion of gain events were related to COG category G (carbohydrate transport and metabolism, 6.1%) and COG category E (amino acid transport and metabolism, 5.0%), and a notable proportion of gene families undergoing expansion events were also related to COG categories G (carbohydrate transport and metabolism, 8.0%), E (amino acid transport and metabolism, 7.2%), C (energy production and conversion, 7.2%), K (transcription, 7.0%), and L (replication, recombination, and repair, 7.0%) (Fig. 1B). It seems that the COG categories involved are carbohydrate metabolism and transport as well as amino acid metabolism and transport, which reflect the adaptive strategies of *Acidiphilium*, including the expansion of metabolic abilities to utilize a variety of potential nutrients, while the acquisition of efficient repair mechanisms is in response to damage of biological molecules possibly caused by harsh environments such as AMD sites. This was in line with previous work which showed that larger genomes preferentially accumulated genes associated with metabolism, regulation, and energy conversion (41). We also found that 8.3% of gene families undergoing loss events belonged to COG category J (translation, ribosomal structure, and biogenesis) and that 12.7% of gene families undergoing contraction events belonged to COG category X (mobilome: prophages, transposons) (Fig. 1B), which were probably related to a holistic adjustment toward a more efficient operational and survival mode of these heterotrophs. The most recent common ancestor (MRCA) of *Acidiphilium* spp. was estimated to have emerged around 60.3 million years ago (Mya) (Fig. 1A, left), not long after a recorded strong asteroid impact, which we postulated to be one of the possible courses of significant changes to the Earth’s atmosphere, since it coincided with a decrease in atmospheric O₂ and increase in atmospheric CO₂.

We extracted HGT events predicted with the IMG Annotation Pipeline (Table S3). Results showed that a notable set of genes were identified as being acquired via HGT,
accounting for up to 18.9% of genes among tested Acidiphilium genomes, indicating the chimeric nature of these genomes. Cross-order HGT events from Rhizobiales were the most frequent, accounting for ~32% of total HGT events, followed by cross-class HGT events from Gammaproteobacteria (~11%) and Betaproteobacteria (~7%) and cross-order HGT events from Rhodobacterales (~7%) and Sphingomonadales (~6%). A considerable proportion (~4%) of genes were derived via cross-class HGT from the typical AMD autotroph Acidithiobacillia (Fig. 3). The above-mentioned HGT donors were almost all consistently present together with Acidiphilium in the cooccurrence network based on 16S rRNA gene amplicon data sets generated from AMD samples (Fig. 4). In the cooccurrence network, Acidiphilium accounted for 0.2% of the nodes, while the HGT donors occupied 0.3 to 22.4% of the nodes. Furthermore, most of these HGT donors were present in the first-neighbor network of Acidiphilium. Function annotations of putative horizontally transferred genes (HTGs) based on COG classes showed that these
Cooccurrence network based on correlation analysis of 16S rRNA amplicon sequencing data sets of AMD samples (n = 205). Each node denotes a microbial OTU at a 97% cutoff. The first neighbors of Acidiphilium nodes (highlighted by a red rectangle) were selected using the tool “first neighbors of selected nodes” in Cytoscape.
genes were biased toward COG categories E (amino acid transport and metabolism), K (transcription), C (energy production and conversion), and G (carbohydrate transport and metabolism), which are associated with metabolic and energy production processes. COG categories L (replication, recombination, and repair) and V (defense mechanisms), all associated with defense and repair mechanism, COG category X (mobilome: prophages, transposons), involved in the mobilization of genome fragments, and COG category M (cell wall/membrane/envelope biogenesis), related to enhanced cellular barriers against external disturbance. COG category N (cell motility) was found to account for up to 37% of Rhodobacterales-derived HGT (Fig. 3), which might facilitate Acidiphilium to swim away from harmful environments and/or toward nutrients. We further performed detailed analyses of gene synteny and phylogeny for examination and quantification of the cumulative impact of HGT on Acidiphilium evolution.

Environmental stress adaption. Acidiphilium was predicted to originate at a time when the O$_2$ concentration present in the atmosphere reached its peak (22.7%) (Fig. 1A, left). As Acidiphilium evolved, the atmospheric O$_2$ concentration decreased to about 20.6% in a consistent manner. In contrast, solar luminosity gradually increased to current levels (100 L), and the atmospheric CO$_2$ concentrations first increased by approximately 0.1%, followed by a decrease of about 0.05% (Fig. 1A, left). The bd-type oxidase encoded by cydAB for oxygen-reducing energy production was present in all clades of Acidiphilium but was not uniform with gene synteny and discrepancy of phylogeny compared with the species tree (see Fig. S3, S6, and S7 at https://doi.org/10.6084/m9.figshare.12892016.v1). This suggested that independent HGT events contributed to the acquisitions of cydAB after the speciation of clades I and II but before the divergence of clades III and IV. The acquisitions of cydAB genes were probably in adaption to decreasing atmospheric O$_2$ concentrations, considering the high affinity of bd-type oxidase even at low O$_2$ concentrations (42). Additionally, cytochrome bo$_3$-type ubiquinol oxidase (CyoABCD), which facilitates growth at low pH and low O$_2$ concentrations (43), was found in Acidiphilium clades I and IV, which were probably acquired from species sharing a habitat with Acidiphilium, such as Acidithiobacillus, Acidihalobacter, and Acidiferro bacter species (see Fig. S8 at https://doi.org/10.6084/m9.figshare.12892016.v1). The nuo gene cluster in Acidiphilium that encodes NADH:ubiquinone oxidoreductase (complex I), and functions preferentially under aerobic conditions (44–46), was also HGT derived (see Fig. S9 at https://doi.org/10.6084/m9.figshare.12892016.v1). These acquisitions might explain the facultative anaerobic ability of Acidiphilium. The HGT-derived Calvin cycle-related gene cluster prk-rbcLS-cbbX was present in clades II and IV of Acidiphilium (see Fig. S10 to S14 at https://doi.org/10.6084/m9.figshare.12892016.v1), which might help Acidiphilium overcome oligotrophic AMD conditions through CO$_2$ assimilation. Carbon monoxide dehydrogenase HTGs (CoxLMS) were detected in clades III and IV (see Fig. S15 and S16 at https://doi.org/10.6084/m9.figshare.12892016.v1). This suggested that Acidiphilium might utilize CO that was present in mine areas (47) as an energy supplement and source of CO$_2$, since the atmospheric CO$_2$ concentration dropped to a lower level upon the diversification of Acidiphilium (Fig. 1A, left). Photosystem II-type photosynthetic reaction centers (Puf-BALMC and PuhA) were found in all clades of Acidiphilium, probably transferred to Acidiphilium after the speciation of clade III (see Fig. S17 to S19 at https://doi.org/10.6084/m9.figshare.12892016.v1), coinciding with the acquisition of gene clusters prk-rbcLS-cbbX and coxLMS. Ni/Fe hydrogenase HTGs were detected only in clade IV (see Fig. S20 at https://doi.org/10.6084/m9.figshare.12892016.v1). AMD sites that Acidiphilium inhabits are hyperosmotic and rich in various metals due to the corrosion of minerals by sulfuric acid and chemosynthetic microbes (48). We discovered a set of HGT-derived heavy metal resistance genes as well as osmotic pressure resistance genes in the genomes of Acidiphilium (see Fig. S21 to S46 at https://doi.org/10.6084/m9.figshare.12892016.v1). For example, apcA (49), arsH (50), which encodes an NADPH-dependent metal-reducing cytochrome/protein, merA, encoding mercuric reductase,
**Metabolism and Evolution of the Genus *Acidiphilium***

*chrA*, encoding a chromate transporter, and *arsRCB*, which confers arsenic resistance, were present in all clades of *Acidiphilium*. It was notable that two of our strains, AccI and AccII, harbored more copies of ferric ion-reducing *apcA* genes (49) than strain ZJSH63, probably due to the higher ferric ion concentration in the sites that AccI and AccII inhabited. HGT-derived toxic divalent cation (e.g., cadmium, zinc, cobalt, and copper) resistance genes *czcABC*, *czcD*, *copA*, and *mco* and regulatory genes *czrRS* were also present in all clades of *Acidiphilium*, while HGT-derived Cu homeostasis genes *copCD* (51) were detected only in some strains of clade IV. Chelation of metals with polyphosphate (polyP) is also an effective metal resistance mechanism of acidophiles (52). HGT-derived alkaline phosphatases (Alp), which release phosphate groups from various compounds, and phosphate import systems encoded by *pstSABC* and *pitA* were present in all clades of *Acidiphilium*. Compatible solute uptake and biosynthesis as well as potassium uptake are known as common strategies to counteract osmotic stress (53, 54). The HGT-derived biosynthetic pathway of the compatible solute hydroxyectoine conferred by the gene cluster *ectRABCD-ask* (55) was detected in the genomes of clades III and IV. HGT-derived *kdpABCDE*, which confer resistance to osmotic stress by uptake of K⁺, were present in all clades of *Acidiphilium*, likely acquired via individual HGT. HGT-derived gene cluster *proXYV-beta-soxBDAG*, involved in the uptake and biosynthesis of glycine betaine, is present in all clades of *Acidiphilium* except for clade I. This gene cluster was likely gained before diversification of clades III and IV but after speciation of clade II. A standalone HTG, *betC*, also involved in glycine betaine synthesis, was present only in clades III and IV. Motility and chemotaxis conferred by flagellar and sensory proteins might help microbes swim away from harmful environments and toward favorable chemicals or other nutrients. We found operons involved in flagellar biosynthesis such as *flg*, *flh*, and *fli* located in identified genomic islands (GIs) (Table S2C), and the HGT-derived chemotaxis operon *cheD/Y/B/R/W/X/Y* was also found in all clades of *Acidiphilium*.

**Metabolic potential expansion through HGT.** Many parts of the sulfur, nitrogen, and carbon metabolic pathways in *Acidiphilium* were also acquired through HGT. The Sox multienzyme complex (encoded by *soxCDYZZAB* and the regulatory *soxH* and *soxR* genes), which oxidizes thiosulfate to sulfate (56), was present in all clades of *Acidiphilium* and was likely derived via independent HGT events in clade I, clade II, and the MRCA of clades III and IV (see Fig. S47 and S48 at https://doi.org/10.6084/m9.figshare.12892016.v1). HTG *sqr*, encoding sulfide:quinone oxidoreductase, was found in clades III and IV (see Fig. S49 at https://doi.org/10.6084/m9.figshare.12892016.v1). HGT-derived polysulfide reductase (PsrABC) was found in clade IV, while thiosulfate sulfurtransferase (*Tst*) was detected only in clades II and IV (see Fig. S50 to S52 at https://doi.org/10.6084/m9.figshare.12892016.v1). Homologues of *Sdo1* in *Acidithiobacillus caldus* MTH-04 (A5904_0421), a new sulfur dioxygenase associated with tetrathionate oxidation (57), were detected in all clades of *Acidiphilium* except clade II, likely acquired after the diversification of clades III and IV (see Fig. S53 and S54 at https://doi.org/10.6084/m9.figshare.12892016.v1). Thiosulfate dehydrogenase/tetrathionate reductase (encoded by *tsdA*), which mediates the flow of electrons into respiratory or photosynthetic electron chains (58), was also detected in *Acidiphilium* outside clade I (see Fig. S55 at https://doi.org/10.6084/m9.figshare.12892016.v1), which was likely acquired after the emergence of clade III (see Fig. S56 at https://doi.org/10.6084/m9.figshare.12892016.v1). Reversible assimilatory sulfate reduction (conferred by *cysND* and *cysJHI*), found in all clades of *Acidiphilium*, was inferred to be acquired before diversification of clades III and IV but after the speciation of clade II (see Fig. S57 to S59 at https://doi.org/10.6084/m9.figshare.12892016.v1). Nitrogen metabolism enzymes for nitrate/nitrite transporter (*NarK*), dissimilatory nitrate reductase (*NarGHJI*), nitric oxide dioxygenases (*Hmp*), cyanate lyase (*CynS*), hydroxylamine reductase (*Hcp*), assimilatory nitrate reductase (*NasA*), and nitrite reductase (*NirBD*) were all derived via HGT. However, they were only sparsely present in *Acidiphilium* (mostly clade IV) (see Fig. S60 to S64 at https://doi.org/10.6084/m9.figshare.12892016.v1).
Abundant HTGs involved in carbon metabolism were also detected, including those coding for a complete methane catabolic pathway and many key genes involved in hydrocarbon utilization. An HGT-derived methane catabolism pathway (conferring by genes *mmoXCYB*, *fdhA*, *mdh*, *gfa*, *frmA*, and *frmB*) was found in almost all clades of *Acidiphilium* (see Fig. S65 to S70 at https://doi.org/10.6084/m9.figshare.12892016.v1). Soluble methane monooxygenase (sMMO) converts methane to methanol, which could be further converted to formaldehyde by methanol dehydrogenase (*Mdh*). Formaldehyde is eventually converted to CO2 via enzymes encoded by *gfa*, *frmA*, and *frmB* (59, 60). Multiple genes related to glycolysis and gluconeogenesis (*fba*, *fbp*, GNL), the pentose phosphate pathway (*tkt*, *tal*, *pgd*, *rpe*, G6PD, *prsA*, and *xfp*), the Entner-Doudoroff pathway (*gdh*, *ddgk*, and *pglI*, methylglyoxal metabolism (*megR*), tricarboxylic acid cycle or glyoxylate bypass (*fumC*, *mdh*, *mdh*, *mls*, *icl*), acetogenesis (*spxB*, *pta*, PDHA1, *actP*), and the previously mentioned *prk-rbcLS-cbbX* involved in the Calvin cycle were proposed to have been acquired via HGT (see Fig. S71 to S92 at https://doi.org/10.6084/m9.figshare.12892016.v1). In addition, the *rha* operon, involved in metabolism of L-rhamnose, was present in the genomes of clades II and IV, likely acquired by HGT (see Fig. S93 at https://doi.org/10.6084/m9.figshare.12892016.v1). For sugar alcohol metabolism, genes involved in the metabolism of erythritol and inositol were found only in clades II and IV. The HGT-derived cluster *iolBDGC-iolKEF-iolGEG2G2-iolBCG3*, conferring inositol catabolic ability, was detected in clade IV, while a similar but differently arranged gene cluster, *iolDEG1G2-inoEKF-iolBCG2*, was detected in clade II (see Fig. S94 at https://doi.org/10.6084/m9.figshare.12892016.v1). In addition, the *ery* operon, involved in erythritol utilization, was presented in clades II and IV of *Acidiphilium*, likely acquired via cross-order HGT (see Fig. S95 at https://doi.org/10.6084/m9.figshare.12892016.v1).

Multiple genes involved in the biodegradation and metabolism of aromatic compounds showed patchy distribution throughout the genus *Acidiphilium*. Many were gained via HGT (see Fig. S96 to S103 at https://doi.org/10.6084/m9.figshare.12892016.v1), including genes *fadA*, *fadB*, *fadD* (n-phenylalkanoic acid degradation), gene cluster *catFJ* (chloroaromatic degradation), gene cluster *benABCD-catBCA* (benzoate degradation), and genes *pcaGH*, *pcaB*, *pcaC*, *pcaD*, *pcaT*, *pcaQ*, and *pcaU* (beta-ketoisocaprate pathway), together with *pobA* and *pcaK* (hydroxybenzoate degradation), *iqaAB* (n-heterocyclic aromatic compound degradation), and *fahA* and *faaH* (styrene degradation).

**Positive selection analyses.** Genome-scale positive selection analyses were exhaustively performed on all 21 genomes of *Acidiphilium* sp. Accl was used as the anchor genome. Results showed that 15 genes were identified as being under positive selection, including Ap_37 (exoribonuclease), Ap_279 (acyl coenzyme A [acyl-CoA] dehydrogenase), Ap_346 (H+/Cl− exchange transporter), Ap_374 (membrane component of nitrite reductase), Ap_696 (hypothetical protein), Ap_835 (N-acetyltransferase), Ap_1029 (hypothetical protein), Ap_1038 (Holliday junction resolvase), Ap_1473 (3-hydroxy acid dehydrogenase), Ap_1766 (hypothetical protein), Ap_1855 (asparaginase synthetase), Ap_2034 (transposase), Ap_2355 (pyridoxine/pyridoxamine 5′-phosphate oxidase), Ap_2619 (cysteine synthase), and Ap_2636 (thioredoxin) (Table S4).

**DISCUSSION**

In this study, we present a detailed analysis of the metabolic capabilities and evolutionary history of the genus *Acidiphilium*. Abundant HGT events were found to contribute substantially to the genomic contents of *Acidiphilium*, providing this genus with unprecedented elasticity to counteract harsh conditions such as those found in AMD. HGT may also have had a great impact on the diversity of *Acidiphilium* gene repertoires. Genome size dynamics (“Why are some genomes really big and others quite small?”) and the occurrence of horizontal gene transfer (“Why does lateral transfer occur in so many species and how?”) were listed as two world-class scientific questions that awaited answers by the editorial of the journal *Science* entitled “So much more to
know” (61), and we believe that the present study might provide some clues for these questions. Our results showed that the genome size of Acidiphilium was relatively large (~4 Mbp), with overwhelming gene family gains predicted across its evolution. This is in sharp comparison with its endosymbiotic Acetobacteraceae relative (genome size of ~2 Mbp) that underwent genome reduction (62, 63). Microbes tend to evolve relatively large genomes with higher nutrient uptake and metabolic potential as a means to compensate for fluctuating and inhospitable environments (64–66). This theory could be applied to Acidiphilium spp. that inhabit hyperacidic, metal-laden, nutrient-depleted AMD environments, which are quite different from the stable environments (with plentiful easy-to-metabolize resources) in which their endosymbiotic Acetobacteraceae relatives dwell (67). The gene repertoire of microbes might evolve rapidly, with HGT being a major source of gene acquisitions in microbial genomes, and a number of genomic analyses have shown that microorganisms adopted new a “lifestyle” via HGT in the colonization of new niches (68). In addition, HGT occurs mainly in the form of horizontal operon transfer (HOT) (69), since many functional modules require a contiguous gene cluster, as exemplified by the acquisition of the photosynthetic operon in Rhodobacteraceae (70). Consistent with this, our results showed that HGT of functional genes (cluster) might have conferred to Acidiphilium better environmental adaptations as well as the expansion of a wide range of metabolic abilities. For example, the coincided acquisitions of the photosystem II-type photosynthetic reaction center (RC), Calvin cycle enzymes, and carbon monoxide dehydrogenases might have conferred adaptive benefits to Acidiphilium by taking advantage of the high CO levels (~50 ppm) in mining areas as an energy source (47, 71) and the increasing solar luminosity for enhanced CO₂ assimilation and/or energy production (72). The methane (CH₄), metal sulfides, hydrogen (H₂), and hydrogen sulfide (H₂S) that are present in mining areas (47, 73) are also potential energy sources for Acidiphilium. Hydrogen (H₂) might be formed in AMD areas through the acid dissolution of metals and minerals, and Ni/Fe hydrogenses might exploit this as an electron donor to support chemolithotrophic growth (74). Correspondingly, Acidiphilium acquired a nearly complete repertoire of methane and sulfur metabolic genes, as well as genes encoding Ni/Fe hydrogenase. Evidence of expression of the above-mentioned pathways has been shown in a Rhodospirillales relative (75). This evidence together with previously observed related phenotypes of Acidiphilium (see the introduction) suggests that these HGT-derived genes may also be functional in Acidiphilium. Numerous HGT-derived resistance genes for AMD adaption were present in Acidiphilium, similar to other AMD inhabitants (76, 77), reflecting diverse strategies of Acidiphilium to avert the deleterious effects of toxic metals and osmotic pressure in AMD environments. The acquired metabolic capacity of organic compounds and hydrocarbon in Acidiphilium suggested a mutualistic interaction of autotrophic acidophiles in AMD. For instance, metabolotrophic organic byproducts excreted by autotrophs (e.g., Acidithiobacillus) might be utilized by chemoorganotrophs (e.g., Acidiphilium), a process that in turn might stimulate the metabolic processes of the autotrophs (78–80). AMD microbial communities tend to form biofilms on mineral substrates for better metabolic cooperation and enhanced resistance against harsh environments (80–82). The microbes in biofilms are usually active, and the high community density, with increased proximity of microbes encapsulated in biofilm, might create more numerous opportunities for the efficient occurrence of HGT. In addition, MGES, such as plasmids, might contribute to the development, stabilization, and expansion of biofilm (83–85). Consistent with this, our results showed that the predicted donors of HGT were also present in the cooccurrence network of Acidiphilium. Cooccurrence networks in microbial communities may provide hints for ecological interactions between species, on which HGT might have an influence (86, 87). Positive selection was also found to be an important driving force for adaptive evolution of Acidiphilium. Genes can be changed by positive selection for fixation of beneficial variants in a population/species over time if they increase survival fitness, which might help fine-tune gene expression in adaption to changing environmental conditions (27–30). Those genes under positive selection in Acidiphilium were prone to play key
multifunctional roles, of which even small adaptive changes in their coding sequences might influence multiple pathways, bringing considerable benefits for survival of microbes during evolution in response to changing global conditions and shifting of niches. For example, the positively selected gene (PSG) cysK (Ap_2619, encoding cysteine synthase) might perform functions related to sulfide utilization, tellurite resistance, and growth inhibition (88–90); the PSG pdxH (Ap_2355), encoding pyridoxine/pyridoxamine 5′-phosphate oxidase, might act as a potent quencher of reactive oxygen intermediates and as an essential cofactor in amino acid metabolism (91, 92); and, finally, the PSG trxA (Ap_2636) encodes thioredoxin, a small redox protein that may play important roles in electron transfer, transcriptional regulation, immune response, and oxidative stress defense (93–96). However, further experiments are required to confirm their actual functions in Acidiphilium. The genus Acidiphilium is one of only four genera in the family Acidobacteriaceae reported to colonize metal-rich AMD sites thus far, with the other genera of Acetobacteraceae found primarily in more moderate environments such as vinegar production environments and breweries (38, 97, 98). We suggest that the ancestor of Acidiphilium may have originated in mild or moderate conditions but then adapted to extreme environments, such as AMD niches, with the help of HGT and probable positive selection on the genes, similar to what has been found in acidophilic archaeal lineages such as Thermoplasmales and Sulfolobales, which seemed to have evolved independently from moderately acidophilic ancestors (99). It is foreseeable that as more Acidiphilium strains are isolated and sequenced, the panorama of Acidiphilium evolution will gradually unfold before us.

Conclusions. Extremophiles that thrive in extremely acidic environments are research model organisms for microbial adaption and evolution. In this study, we provided evidence that Acidiphilium is characterized by a complex lifestyle granted by HGT. By way of gene acquisitions, Acidiphilium has greatly expanded its genetic diversity, resulting in functional divergence. Acidiphilium has acquired multiple abilities via HGT, such as photosynthesis, CO2 assimilation, metal resistance, and organic compound metabolism, which would facilitate beneficial interactions with cohabitant autotrophs. In addition, the predicted donors of HGT were present in the cooccurrence network of Acidiphilium. Positive selection on new mutations was also an important driving force in the evolution of Acidiphilium. We further proposed that microorganisms originating under mild conditions can adapt to extreme environments such as AMD sites after the acquisition of multiple adaptive functions. Taken together, this study has shed light on the ecological roles and evolutionary scenario of Acidiphilium and is a good example of research on the adaption and evolution of extremophiles.

MATERIALS AND METHODS

DNA extraction, genome sequencing, and assembly. Strains Accl and Accll were isolated from an acid mine drainage (AMD) water sample obtained in the Mangzi mining area (formed by oxidizing dissolution of pyrite, characterized by a high ferric ion concentration), Yunnan Province, China (long 103.5, lat 23.3, altitude 1,847 m). Strain ZJSH63 was isolated from an AMD water sample obtained in the heap leaching area for copper ore in the Zjinshan Gold and Copper Mine, Fujian Province (lat 25.2, long 116.4, altitude 282.6 m), China. Genomic DNA of strains Accl, Accl, and ZJSH63 were extracted using the Qiaqen genomic DNA extraction kit (Qiagen, Hilden, Germany) in accordance with the manufacturer’s instructions. After the DNA sample passed quality testing, the large fragments were subjected to agarose recovery using a BluePippin automatic nucleic acid recovery instrument (SAGE Science). The DNA was damaged and repaired; after purification, the DNA fragments were end repaired and linked with adenine.

After a purification and ligation reaction, Qubit was used to accurately quantify the constructed DNA library by following official protocol (http://support/software-downloads/). The DNA library was subjected to the PacBio Sequel platform for sequencing at Guangdong Magigene Biotechnology Co. Ltd. (Guangzhou, China). After sequencing, SMRT Link v5.1.0 (https://www.pacb.com/) was utilized for correction and assembly.

ANI and whole-genome alignments. JSpecies v1.2.1 was used to calculate average nucleotide identity (ANI) based on the BLASTN algorithm with default parameters (100). BLASTN-based whole-genome comparison of Acidiphilium strains (completeness > 97%) was performed and represented with BRIG-0.95 (101). We utilized Circos (102) for construction and visualization of the multiple genome alignments of strains with completely sequenced genomes, including Accl, JF-5, and AIU301.

Pangenome analyses and gene family evolution analyses of Acidiphilium. A summary of features for the Acidiphilium genomes involved in this study are listed in Table 1. BUSCO (103) was used to estimate the completeness of each genome against its bacterial core gene set. Gene family clustering of
12 Acidiphilium genomes (completeness > 97%) together with UniProt search, GO Slim annotation, and GO enrichment analyses (default cutoff P value, 0.05) was performed via OrthoVenn2 (104) with default parameters. The BPGA pipeline (105) was used to perform model extrapolations of the Acidiphilium pangenome/core genome by applying default parameters. We applied COUNT (106) under the Wagner parsimony algorithm for ancestor genome size estimation and for detecting the gain, loss, expansion, and contraction events of gene families with the penalty ratio set to 1.

Phylogenic analyses and divergence time estimation. Phylogenetic trees based on 133 concatenated core genes and 16S rRNA gene sequences of Acidiphilium were constructed with the neighbor-joining (NJ) method using MEGA-X (107) with 1,000 bootstrap replicates. A chronogram for Acidiphilium species with branch lengths reflecting divergence times was inferred on the core gene tree of Acidiphilium using the RelTime method (108) implemented in MEGA-X with the JTT matrix-based model as described previously (109). The TimeTree reference data (110) that integrated data of asteroid impacts (Earth Impact Database, http://www.impact-structures.com/database-of-earth-impact-structures/), solar luminosity (111), and fluctuations in atmospheric O\textsubscript{2} (112) and CO\textsubscript{2} (113–116) amount were displayed synchronously with divergence times in the form of time panels. Phylogenetic trees based on protein sequences of functional genes were constructed using PhyML (117) with the maximum likelihood (ML) method and 1,000 bootstrap replicates, followed by visualization with iTOL (118). Sequences were aligned with Muscle (119) and trimmed with Gblocks (120) before tree construction.

Genome annotation and horizontally transferred gene prediction. We applied RAST (121), KEGG (122), and COG (123) databases (BLASTP cutoff, E value < 10^{-5}) for genome annotation. We also extracted information of putative horizontally transferred genes from IMG Annotation results. Genome neighbor (context) visualizations were conducted with the EFI-GNT tool (124). Identification of putative horizontally transferred genes (HTGs) in the genomes of Acidiphilium was performed via the Integrated Microbial Genomes (IMG) system (125), which defined genes as being putative lateral transfers by the following principle: genes that have their best BLAST hits (best bit scores) or >90% of the best hits outside the taxonomic lineage of the genome (i.e., to genomes from another phylum, class, etc.) but with lower-scoring hits or no hits within the lineage.

Prediction of mobile genetic elements. We applied the ISfinder (126) to predict and classify insertion sequences (IS) and transposases within Acidiphilium genomes with BLASTP (cutoff E value, 1e^{-5}). IslandViewer 4 (127) was used to detect putative genomic islands (GIs) distributed within Acidiphilium genomes. PHASTER (128) was applied to detect prophage and prophage remnant sequences within Acidiphilium genomes. We also applied CRISPICasFinder (129) for detection of CRISPRs and Cas within Acidiphilium genomes.

Construction of cooccurrence network. To identify the associations between Acidiphilium and other microbes in AMD environments, 16S rRNA amplicon sequencing data sets of AMD samples (n = 205) were collected from the Sequence Read Archive (SRA) database (see Table S5 in the supplemental material). The QIIME (130) pipeline was applied to analyze these data sets. Sequences were clustered into operational taxonomic units (OTUs) at the 97% similarity level with the “closed reference OTU picking” strategy against the QIIME formatted Greengenes v.13.8 reference database (http://greengenes.lbl.gov). Rare OTUs, with fewer than five occurrences, were removed before network construction. The cooccurrence network was constructed using CoNet (131), which was implemented in Cytoscape v.3.6.1 based on the OTU occurrence frequency. Pairwise scores between OTUs were calculated using Spearman rank correlations applying a threshold rho of >0.6 and a P value of <0.01. The cooccurrence network was visualized with Organic layout in Cytoscape v.3.6.1 (132).

Genome-wide detection of positively selected genes. We used the PosiGene pipeline (133) for genome-wide detection of positively selected genes in the above-mentioned strains of Acidiphilium spp., in which Acidiphilium sp. AccI was used as the anchor species, reference, and target species. Genes were considered under positive selection if the branch-wide test resulted in false discovery rates (FDR) of <0.05 and adjusted P values of <0.05.

Data availability. The genome sequences of Acidiphilium strains AccI, AccII, and ZJSH63 have been deposited in the JGI IMG-ER database under ER Genome IDs 2824045439, 2824049744, and 2828862166, respectively.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TEXT S1, DOCX file, 0.01 MB.

TABLE S1, XLSX file, 0.02 MB.

TABLE S2, XLSX file, 0.5 MB.

TABLE S3, XLSX file, 0.04 MB.

TABLE S4, DOCX file, 0.01 MB.

TABLE S5, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

We thank Han Zhou for assistance in data processing and figure creation and the NCBI database and IMG-ER database for providing the genome sequences of Acidiphilium spp.

This work was funded by the National Natural Science Foundation of China (grant...
REFERENCES

1. Merino N, Aronson HS, Bojanova DP, Feyhl-Buska J, Wong ML, Zhang S, Giovannelli D. 2019. Living at the extremes: extremophiles and the limits of life in a planetary context. Front Microbiol 10:780. https://doi.org/10.3389/fmicb.2019.00780.

2. Rothschuld LJ, Mancinelli RL. 2001. Life in extreme environments. Nature 409:1092-1101. https://doi.org/10.1038/35059215.

3. Hua ZS, Chen LX, Liu J, Hu M, Li SJ, Kuang JL, Chain PS, Huang LN, Shu WS. 2015. Ecological roles of dominant and rare prokaryotes in acid mine drainage revealed by metagenomics and metatranscriptomics. ISME J 9:1280–1294. https://doi.org/10.1038/ismej.2014.212.

4. Kadnikov KV, Ivensko DA, Beletsky AV, Mardanov AV, Danilova EV. 2013. Iron-oxidizing bacterium capable of coupling the reduction of Fe(III) to the oxidation of glucose. Appl Environ Microbiol 69:6134–6144. https://doi.org/10.1128/AEM.01065-13.

5. Kuang JL, Huang L, Chen L, Hu M, Li J, Shu W. 2013. Contemporary environmental variation determines microbial diversity patterns in acid mine drainage. ISME J 7:1038–1050. https://doi.org/10.1038/ismej.2012.139.

6. Sievers M, Ludwig W, Teuber M. 1994. Phylogenetic positioning of Acetobacter, Gloconobacter, Rhodopila and Acidiphilium species as a branch of acidophilic bacteria in the a-subclass of Proteobacteria based on 16s ribosomal DNA sequences. Syst Appl Microbiol 17:189–196. https://doi.org/10.1007/s00792-008-0214-2.

7. Cummings DE, Fendorf S, Singh N, Sani RK, Peyton BM, Magnuson TS. 2010. Limiting factors for microbial community development in acid mine drainage. Appl Environ Microbiol 65:3633–3640. https://doi.org/10.1127/aem.65.3633-3640.1999.

8. Tomi T, Shibata Y, Ikeda Y, Taniguchi S, Haik C, Mataga N, Shimada K, Haik C, Mataga N, Shimada K, et al. 2007. Reduction of Cr(VI) under acidic conditions by the facultative Fe(III)-reducing bacterium Acidiphilium cryptum JF-5. Extremophiles 11:299–308. https://doi.org/10.1007/s00792-007-0566-0.

9. Küsel K, Dorsch T, Acker G, Stackebrandt E. 1999. Microbial reduction of inorganic sulfur compound oxidation of heterotrophic Acidiphilium rubrum containing Zn-oligotrophic conditions. mSphere 4:e00011-19.

10. Mahapatra NR, Banerjee PC. 1996. Extreme tolerance to cadmium and high resistance to copper, nickel and zinc in different Acidiphilium strains. Lett Appl Microbiol 23:393–397. https://doi.org/10.1111/j.1742-765X.1996.tb03444.x.

11. San Martin-Uriz P, Mirote S, Alcolea PJ, Gomez MJ, Amils R, Gonzalez-Pastor JE. 2014. Nickelfastenkeim in Acidiphilium sp. PM identfied by genome-wide functional screening. PLoS One 9:e95041. https://doi.org/10.1371/journal.pone.0095041.

12. Borole AP, O’Neill H, Tsorius C, Cesar S. 2008. A microbial fuel cell operating at low pH using the acidophile Acidiphilium cryptum. Bio- technol Lett 30:1367–1372. https://doi.org/10.1007/s10529-008-9700-y.

13. Xu A-L, Xia J-I, Zhang S, Yang Y, Nie Z-Y, Qiu G-Z. 2010. Bioleaching of chalcopyrite by UV-induced mutagenized Acidiphilium cryptum and Acidithiobacillus ferrooxidans. Trans Nonferrous Metals Soc China 20:315–321. https://doi.org/10.1016/j.tnf.2009.10.040.

14. González E, Rodríguez JM, Muñoz JA, Blázquez ML, Ballester A, González F. 2018. The contribution of Acidiphilium cryptum to the dissolution of low-grade manganese ores. Hydrometallurgy 175:312–318. https://doi.org/10.1016/j.hydromet.2017.12.008.

15. Priya A, Hait S. 2017. Feasibility of bioleaching of selected metals from electronic waste by Acidiphilium cryptum. Waste Biomass Valoriz. 9:871–877.

16. Almeida WI, Vieira RP, Cardoso AM, Silveira CB, Costa RG, Gonzalez AM, Paranhos R, Medeiros JA, Freitas FA, Albano RM, Martins OB. 2009. Archaeal and bacterial communities of high sulphur metal contaminated acidic waters from zinc mine residues in Sepetiba Bay. Extremophiles 13:263–271. https://doi.org/10.1007/s00792-009-0214-2.

17. Borole AP, O’Neill H, Tsorius C, Cesar S. 2008. A microbial fuel cell operating at low pH using the acidophile Acidiphilium cryptum. Bio- technol Lett 30:1367–1372. https://doi.org/10.1007/s10529-008-9700-y.

18. Li et al. 2020. Gene expansion and positive selection as bacterial adaptations to oligotrophic conditions. mSphere 4:e00011-19. https://doi.org/10.1128/mSphereDirect.00011-19.
Metabolism and Evolution of the Genus Acidiphilium

30. Zhang Y, Jalan N, Zhou X, Gees J, Jones JB, Setubal JC, Deng X, Wang N. 2015. Positive selection is the main driving force for evolution of citrus canker-causing Xanthomonas. ISME J 9:2138–2138. https://doi.org/10.1038/ismej.2015.15.

31. Bowers RM, Kypridines NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jaret J, Rivers AR, Eloe-Fadroz EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Sakurai Y, Shiota H. 1994. Acidiphilium multivorum sp. nov., an acidophilic heterotrophic bacteria isolated from acidic coal mine drainage. J Gen Appl Microbiol 40:143–159. https://doi.org/10.10373/jgam.40.143.

32. Vicilacz PL, Unz RF, Langworthy TA. 1986. Acidiphilium angustum sp. nov., a new acidophilic bacterium from the acid mine drainage. J Gen Appl Microbiol 36:197–201. https://doi.org/10.1046/j.1365-2958.2001.02100.x.

33. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jaret J, Rivers AR, Eloe-Fadroz EA, Tringe SG, Ivanova NN, Copeland A, Sakurai Y, Shiota H. 1994. Acidiphilium multivorum sp. nov., an acidophilic heterotrophic bacteria isolated from acidic coal mine drainage. J Gen Appl Microbiol 40:143–159. https://doi.org/10.10373/jgam.40.143.

34. Govantes F, Albrecht JA, Gunsalus RP. 2000. Oxygen regulation of the Escherichia coli cytochrome d oxidase (cydAB) operon: roles of multiple promoters and the Fnr-1 and Fnr-2 binding sites. Mol Microbiol 37: 1456–1469. https://doi.org/10.1046/j.1365-2958.2000.02100.x.

35. Varble A, Meaden S, Barrangou R, Westra ER, Marraffini LA. 2009. Recombination between phages and CRISPR-cas loci facilitates horizontal gene transfer in staphylococci. Nat Microbiol 12:53–58. https://doi.org/10.1036/j.tim.2003.12.010.

36. Wei W, Pang X, Lin J, Liu X, Wang R, Lin J, Chen L. 2017. Discovery of a new subgroup of sulfur dioxygenases and characterization of sulfur dioxygenases in the sulfur metabolic network of Acidithiobacillus caldus. PLoS One 12:e0183668. https://doi.org/10.1371/journal.pone.0183668.

37. Bentkowski P, Van Oosterhout C, Mock T. 2015. A model of genome size evolution for prokaryotes in stable and fluctuating environments. Geomicrobiol J 6:912–920. https://doi.org/10.1093/gbe/evu062.

38. El I, Ileghems K, Van Kerrebroeck S, Borremans W, Cleenwerck I, De Vuyst L, Vandamme P. 2016. Whole-genome sequencing of two phylogenetically distinct bacteria and a new protein involved in thiosulfate oxidation. FEBS Lett 583:1281–1286. https://doi.org/10.1016/j.febslet.2009.03.020.

39. Xie J, He Z, Liu X, Liu X, Van Nostrand JD, Deng Y, Wu L, Zhou J, Qiu G. 2011. GeoChip-based analysis of the functional gene diversity and metabolic potential of microbial communities in acid mine drainage. Environ Microbiol 13:421–456. https://doi.org/10.1111/j.1462-2695.2012.08068.x.

40. Ozmen I, Aksoy E. 2015. Respiratory emergencies and management of mining accidents. Turk Thorac J 16(Suppl 1):S18–S20.

41. Xue J, He Z, Liu X, Liu X, Van Nostrand JD, Deng Y, Wu L, Zhou J, Qiu G. 2011. GeoChip-based analysis of the functional gene diversity and metabolic potential of microbial communities in acid mine drainage. Appl Environ Microbiol 77:991–999. https://doi.org/10.1128/AEM.01798-10.

42. Magnuson TS, Swenson MW, Paszczynski AJ, Deobald LA, Kerk D, Cunnings DE. 2010. Proteogenomic and functional analysis of chorate reduction in Acidiphilium cryophilum JF-5, an Fe(III)-respiring acidiphile. Biometals 23:1129–1138. https://doi.org/10.1007/s10534-010-9360-y.

43. Sharma P, Mattos MJTD, Hellingwerf KJ, Bekker M. 2012. On the function of the various quinone species in Escherichia coli. FEBS J 279: 3364–3373. https://doi.org/10.1111/j.1742-4658.2012.08608.x.

44. Oszmian P, van Oosterhout C, Mock T. 2015. A model of genome size evolution for prokaryotes in stable and fluctuating environments. Genome Biol Evol 7:2344–2351. https://doi.org/10.1093/gbe/evv148.

45. Guiseppi E, Wuerzt S. 2012. Metabolically versatile large-genome prokaryotes. Curr Opin Biotechnol 23:467–473. https://doi.org/10.1016/j.copbio.2011.12.022.

46. Bentkowski P, van Oosterhout C, Ashby B, Mock T. 2017. The effect of November/December 2020 Volume 5 Issue 6 e00867-20 msystems.asm.org 17
extrinsic mortality on genome size evolution in prokaryotes. ISME J 11:1011–1018. https://doi.org/10.1038/ismej.2016.156.

67. Brown BP, Wernigreen JJ. 2019. Genomic erosion and extensive horizontal gene transfer in gut-associated Acetobacteraceae. BMC Genomics 20:472. https://doi.org/10.1186/s12864-019-5844-5.

68. Vos M, Hesselman MC, Te Beek TA, van Passel MJW, Eyre-Walker A. 2015. Rates of lateral gene transfer in prokaryotes: high but why? Trends Microbiol 23:598–605. https://doi.org/10.1016/j.tim.2015.07.002.

69. Lindsey AR, Newton ILG. 2019. Some like it HOT: horizontal operon transfer. Cell 176:1243–1245. https://doi.org/10.1016/j.cell.2019.02.007.

70. Brinkmann H, Göker M, Klobilzek M, Wagner-Döbler I, Petersen J. 2018. Horizontal operon transfer, plasmids, and the evolution of photosynthesis in Rhodobacteraceae. ISME J 12:1994–2010. https://doi.org/10.1038/s41396-018-0150-9.

71. Yelton AP, Comolli LR, Justice NB, Castelle C, Denef VJ, Thomas BC, Banfield JF. 2013. Comparative genomics in acid mine drainage biofilm communities reveals metabolic and structural differentiation of co-occurring archaea. BMC Genomics 14:485. https://doi.org/10.1186/1471-2164-14-445.

72. Kishimoto N, Fukaya F, Inagaki K, Sugio T, Tanaka H, Tano T. 1995. Distribution of bacteriochlorophyll a among aerobic and acidophilic bacteria and iron-enhanced CO2 incorporation in Acidithiobacillus rufus. FEMS Microbiol Ecol 16:291–296. https://doi.org/10.1111/j.1574-6941.1995.tb00293.x.

73. Ly T, Wright JR, Weit N, McLimans CJ, Ulrich N, Tokarev V, Valkanas MM, Trun N, Rummel S, Grant CJ, Lamendella R. 2019. Microbial communities associated with passive acidic abandoned coal mine remediation. Front Microbiol 10:1955. https://doi.org/10.3389/fmicb.2019.01955.

74. Hedrich S, Johansson D. 2013. Aerobic and anaerobic oxidation of hydrogen by acidophilic bacteria. FEMS Microbiol Lett 349:40–45. https://doi.org/10.1111/1574-6966.12290.

75. Orlova MV, Tarlachkov SV, Dubinina GA, Belousova GA, Tutukina MN, Hedrich S, Johnson DB. 2013. Aerobic and anaerobic oxidation of hydrogen by acidophilic bacteria. FEMS Microbiol Lett 349:40–45. https://doi.org/10.1111/1574-6966.12290.

76. Li L, Liu Z, Meng D, Liu X, Li X, Zhang M, Tao J, Gu Y, Zhong S, Yin H. 2019. Comparative genomic analysis reveals distribution, organization, and evolution of metal resistance genes in the genus Acidithiobacillus. AppI Environ Microbiol 85:e02153-18. https://doi.org/10.1128/AEM .02153-18.

77. Navarro CA, Von Bernath D, Jerez CA. 2013. Heavy metal resistance strategies of acidophilic bacteria and their acquisition: importance for the evolution of metal resistance genes in the genus Acidithiobacillus. AppI Environ Microbiol 85:e02153-18. https://doi.org/10.1128/AEM .02153-18.

78. De Roos J, De Vuyst L. 2018. Acetic acid bacteria in fermented foods and beverages. Curr Opin Biotechnol 49:115–119. https://doi.org/10.1016/j.copbio.2017.03.076.

79. Möller MC, Hederstedt L. 2008. Extracytoplasmic processes impaired by the thioredoxin antioxidant system. Free Radic Res 42:472. https://doi.org/10.1080/10715769.2008.1093312.

80. Yi W, Coleman-Derr D, Chen G, Yong QG. 2015. OrthoVenn: a web server for generating Venn diagrams and performing set comparisons of viral, bacterial, and eukaryotic genomes. Nucleic Acids Res 43:7568–7573. https://doi.org/10.1093/nar/gku1130.

81. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2011. BUSCO: assessing genome assembly and annotation completeness of prokaryotic genomes. Genome Res 21:3219–3228. https://doi.org/10.1101/gr.092759.109.

82. Colman DR, Poudel S, Hamilton TL, Havig JR, Selensky MJ, Shock EL, Alikhan NF, Petty NK, Zakour NLB, Beatson SAA. 2011. BLAST Ring Alignment Tool (BRAT): a web server for visualizing BLAST alignments. Nucleic Acids Res 39:e53. https://doi.org/10.1093/nar/gkr057.

83. Alikhan NF, Petty NK, Zakour NLB, Beatson SAA. 2011. BLAST Ring Alignment Tool (BRAT): a web server for visualizing BLAST alignments. Nucleic Acids Res 39:e53. https://doi.org/10.1093/nar/gkr057.

84. Li et al. 2014. Circos: an information aesthetic for visualizing complex data structures of genomic variation and function. Genome Res 24:1531–1540. https://doi.org/10.1101/gr.153138.114.

85. Matxain JM, Padro D, Ristilä M, Strid A, Eriksson LA. 2009. Evidence of extracytoplasmic processes impaired by the thioredoxin antioxidant system. Free Radic Biol Med 46:565–571. https://doi.org/10.1016/j.freeradbiomed.2008.11.003.

86. Fang P, Peng F, Gao X, Xiao P, Yang J. 2019. Decoupling the effects of pyridoxamine 5'-phosphate oxidase family. FEBS Lett 590:453–460. https://doi.org/10.1002/1873-3468.12480.

87. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, T, Hirakawa H, Kuhara S, Matsushita K, Fujita N, Shirai M. 2009. Identification and cloning of a high-affinity iron uptake system in Acidithiobacillus. Biochemistry 48:11243–11251. https://doi.org/10.1021/bi900611w.

88. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, T, Hirakawa H, Kuhara S, Matsushita K, Fujita N, Shirai M. 2009. Identification and cloning of a high-affinity iron uptake system in Acidithiobacillus. Biochemistry 48:11243–11251. https://doi.org/10.1021/bi900611w.

89. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, T, Hirakawa H, Kuhara S, Matsushita K, Fujita N, Shirai M. 2009. Identification and cloning of a high-affinity iron uptake system in Acidithiobacillus. Biochemistry 48:11243–11251. https://doi.org/10.1021/bi900611w.

90. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, T, Hirakawa H, Kuhara S, Matsushita K, Fujita N, Shirai M. 2009. Identification and cloning of a high-affinity iron uptake system in Acidithiobacillus. Biochemistry 48:11243–11251. https://doi.org/10.1021/bi900611w.

91. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, T, Hirakawa H, Kuhara S, Matsushita K, Fujita N, Shirai M. 2009. Identification and cloning of a high-affinity iron uptake system in Acidithiobacillus. Biochemistry 48:11243–11251. https://doi.org/10.1021/bi900611w.

92. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, T, Hirakawa H, Kuhara S, Matsushita K, Fujita N, Shirai M. 2009. Identification and cloning of a high-affinity iron uptake system in Acidithiobacillus. Biochemistry 48:11243–11251. https://doi.org/10.1021/bi900611w.
