



Unexpected Dominance of Elusive Acidobacteria in Early Industrial Soft Coal Slags

Carl-Eric Wegner^{1,2} and Werner Liesack^{1*}

¹ Department of Biogeochemistry, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany, ² Aquatic Geomicrobiology, Institute of Ecology, Friedrich Schiller University Jena, Jena, Germany

OPEN ACCESS

Edited by:

Etienne Yergeau,
Institut National de la Recherche
Scientifique, Canada

Reviewed by:

Julien Tremblay,
National Research Council Canada
(NRC-CNRC), Canada
Nils-Kaare Birkeland,
University of Bergen, Norway

*Correspondence:

Werner Liesack
liesack@mpi-marburg.mpg.de

Specialty section:

This article was submitted to
Terrestrial Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 15 February 2017

Accepted: 22 May 2017

Published: 08 June 2017

Citation:

Wegner C-E and Liesack W (2017)
Unexpected Dominance of Elusive
Acidobacteria in Early Industrial Soft
Coal Slags. *Front. Microbiol.* 8:1023.
doi: 10.3389/fmicb.2017.01023

Acid mine drainage (AMD) and mine tailing environments are well-characterized ecosystems known to be dominated by organisms involved in iron- and sulfur-cycling. Here we examined the microbiology of industrial soft coal slags that originate from alum leaching, an ecosystem distantly related to AMD environments. Our study involved geochemical analyses, bacterial community profiling, and shotgun metagenomics. The slags still contained high amounts of alum constituents (aluminum, sulfur), which mediated direct and indirect effects on bacterial community structure. Bacterial groups typically found in AMD systems and mine tailings were not present. Instead, the soft coal slags were dominated by uncharacterized groups of Acidobacteria (DA052 [subdivision 2], KF-JG30-18 [subdivision 13]), Actinobacteria (TM214), Alphaproteobacteria (DA111), and Chloroflexi (JG37-AG-4), which have previously been detected primarily in peatlands and uranium waste piles. Shotgun metagenomics allowed us to reconstruct 13 high-quality Acidobacteria draft genomes, of which two genomes could be directly linked to dominating groups (DA052, KF-JG30-18) by recovered 16S rRNA gene sequences. Comparative genomics revealed broad carbon utilization capabilities for these two groups of elusive Acidobacteria, including polysaccharide breakdown (cellulose, xylan) and the competence to metabolize C1 compounds (ribulose monophosphate pathway) and lignin derivatives (dye-decolorizing peroxidases). Equipped with a broad range of efflux systems for metal cations and xenobiotics, DA052 and KF-JG30-18 may have a competitive advantage over other bacterial groups in this unique habitat.

Keywords: soft coal, slags, mineral leaching, acid mine drainage, microbiome, acidobacteria, metagenomics, microbial dark matter

1. INTRODUCTION

Mining frequently produces enormous amounts of tailings, which often pose the danger of generating acid mine drainage (AMD) (Rohwerder et al., 2003). AMD is a consequence of sulfur- and iron-rich minerals, primarily pyrite (FeS₂), being exposed to oxygen and humidity (Druschel et al., 1999). Pyrite oxidation leads to the formation of sulfuric acid and ferrous iron. Sulfuric acid lowers the pH and ferrous iron is further oxidized yielding ferric iron. The gradually decreasing pH increases the solubility of ferric iron, which is a stronger oxidant of pyrite than oxygen (Druschel et al., 1999). Microbial activity accelerates AMD generation tremendously. Microbial iron and sulfur oxidation are orders of magnitude faster than the abiotic oxidation of these two elements. The microbial activity easily leads to a release of acidic and metal-contaminated

discharges. Early studies targeting AMD microbial communities revealed the prevalence of a few taxonomic groups, including *Leptospirillum* and *Acidithiobacillus* among the bacteria and *Ferroplasma* and other *Thermoplasma*-related groups (A-, E-, and G-Plasma) within the archaea (Baker and Banfield, 2003). Metagenomic (Tyson et al., 2004; Dick et al., 2009; Yelton et al., 2011, 2013), metatranscriptomic (Lehembre et al., 2013; Hua et al., 2015; Chen et al., 2015) and metaproteomic (Denef et al., 2010; Mueller et al., 2011) studies showed that iron- and sulfur-cycling are the dominating microbial activities in AMD environments. Most AMD studies focused either on tailings as source material (Radeva and Selenska-Pobell, 2004; Senko et al., 2008; Urbanová et al., 2011; Korehi et al., 2014) or on AMD discharges as the end product (Baker and Banfield, 2003; Tyson et al., 2004; Dick et al., 2009; Xie et al., 2011; Kuang et al., 2013). The most common sources of AMD are mine tailings, and iron and coal deposits.

Though coal is nowadays almost exclusively mined for its use as fossil fuel, mineral-rich coal is mined to win scarce and precious minerals such as germanium (Arroyo and Fernández-Pereira, 2008). In the nineteenth century, sulfur mineral-rich soft coal was mined in Western Germany, in the surroundings of Bonn, to extract these minerals by a smelting-like process (Supplementary Figure 1). Among these minerals, especially alum ($KAl(SO_4)_2 \times 12H_2O$) was commonly used as mordant and precipitant in early textile and paper industry. Alum was extracted from soft coal in a four-step process. Soft coal was mined, sheared and subsequently smoldered to enrich present alum. The resulting ash (one third of the original coal, alum content 15%) was leached and the leachate boiled to yield crystalline alum. Overall, alum leaching was a low-yield process. One ton of coal had to be processed to win 50 kg of crystalline alum, thereby leading to 300–400 kg of leached slag as by-product. At peak year level, 500,000 tons of coal were mined, producing 1,250 tons of alum and 150,000–200,000 tons of slag. In the middle of the nineteenth century, the alum extraction plants and soft coal mines in the surroundings of Bonn were the biggest of their kind in former Prussia. Given a mining period of roughly 70 years (1805–1875), tremendous amounts of leached slag were produced and dumped into the environment without precautions and remediation.

Here we investigated the microbiology of soft coal slag deposits. Given their industrial history and leached nature, we hypothesized that the slags represent an endpoint AMD environment that is characterized by a unique bacterial community composition. In particular, we were interested in exploring to which extent the remaining alum and other metals affect the genome coding potential of the slag-deposit-inhabiting microbial communities. An additional challenge faced by these communities is the recalcitrant nature of their carbon sources. Our research involved a thorough geochemical analysis, bacterial community profiling, and shotgun metagenomics. In addition to samples from three slag deposit sites, sediment from a drainage pond and nearby undisturbed forest soil were examined for comparison.

2. MATERIALS AND METHODS

2.1. Sampling

Detailed information about the sampling site, sampling and alum leaching from soft coal can be found in the Supplementary Methods and Supplementary Figure 1. Three individual sites were sampled at the slag deposit “Red Hill” (abbreviated: RH1 [50° 44′21.45″N, 7° 10′27.87″E], RH2 [50° 44′21.85″N, 7° 10′27.78″E], RH3 [50° 44′22.23″N, 7° 10′27.46″E]) in September 2014. After removing a cover layer (approx. 5 cm) of slowly humifying biomass, the underlying slag was collected in 1 L plastic containers. Two additional sites were sampled for reference and comparison: Sediment of a nearby pond, which collects the drainage water of slag deposits upon rain events (abbreviated: RHP [50° 44′22.32″N, 7° 10′28.32″E]) and nearby undisturbed forest soil (abbreviated: Ref [50° 43′49.63″N, 7° 10′30.44″E]). Three replicate samples were collected within a 4 m² area surrounding each individual sampling site. Replicates were pooled to obtain composite samples that were aliquoted and subjected to individual analyses. Geochemical analyses were carried out as described in the Supplementary Methods.

2.2. DNA Extraction

DNA was extracted using the FastDNA[®] SPIN kit for soil (MP Biomedicals, Eschwege, Germany). Quality and integrity of the DNA was assessed by agarose gel electrophoresis. DNA concentrations were determined fluorometrically. Extracts were aliquoted and stored at –20° C for downstream analyses, including qPCR, amplicon sequencing and shotgun metagenomics.

2.3. 16S rRNA Gene-Targeted qPCR

Bacterial 16S rRNA gene copies were determined by SybrGreen-based quantitative PCR (Stubner, 2004). Standard curves were generated using genomic DNA from *E. coli* (calibration range: 10 – 10⁶ copies). Quantitative PCR was carried out using a CFX Connect Real-Time PCR detection system (Bio-Rad, Munich, Germany). The PCR efficiency was at least 85% ($R^2 \geq 0.99$). Melt curve analyses were carried out to check for unspecific products.

2.4. Amplicon Library Preparation and Sequencing

PCR amplicons for subsequent amplicon library preparation were generated using a bacteria-specific (341f [S-D-Bact-0341-b-S-17, 5′-CCTACGGGNGGCWGCAG-3′]/805r [S-D-Bact-0785-a-A-21, 5′-GACTACHVGGGTATCTAATCC-3′]) primer pair (Klindworth et al., 2013). The PCR assay is further described in the Supplementary Methods. PCR products were checked by agarose gel electrophoresis, gel-excised and purified using the Wizard[®] SV Gel and PCR clean-up system (Promega, Mannheim, Germany). Library preparation was done using the NEBNext[®] Ultra DNA Library Prep Kit for Illumina[®] (New England BioLabs, Frankfurt/Main, Germany), with a final library size of 550 bp. The size distribution was checked by running DNA 12K chips on a Bio-Rad Experion[®] (Bio-Rad, Munich, Germany). Sequencing was carried out at the Max Planck

Genome Centre Cologne using an Illumina[®] MiSeq platform in paired-end mode (2 × 300 bp).

2.5. Shotgun Metagenome Library Preparation and Sequencing

Extracts of total DNA were sheared to a fragment size of approximately 200–250 bp using a Bioruptor[®] Plus (Diagenode, Denville, New Jersey, USA). Shearing involved 15 cycles of 30 s/30 s (on/off) at high intensity. Library preparation was done as described above. Size selection was modified to yield a final library size of approximately 250–300 bp. Libraries were sequenced using an Illumina[®] HiSeq 2500 device in paired-end mode (2 × 100 bp) at the Max Planck Genome Centre Cologne.

2.6. Amplicon Data Processing

Amplicon data were quality-controlled and pre-processed as described in the Supplementary Methods. OTUs (operational taxonomic units) were clustered using USEARCH (v. 7.0) (Edgar, 2010) and a sequence identity threshold of 97%. A compatible OTU table was assembled to facilitate amplicon data analysis using QIIME (v. 1.8) (Caporaso et al., 2010) and PHYLOSEQ (v. 1.10) (McMurdie and Holmes, 2013).

2.7. Metagenome Data Processing and Computational Analysis

Quality control of metagenome reads was done as described for amplicon datasets (Supplementary Methods), except that initial paired-end assembly was omitted, and reads were not filtered based on expected errors but truncated from the first position showing a quality score below Q20. The taxonomic affiliation of quality-controlled reads prior to metagenome assembly was determined using KRAKEN (v. 0.10.5) (Wood and Salzberg, 2014) and UBLAST (an optimized BLAST (Altschul et al., 1997) algorithm implemented in USEARCH) searches against NCBI NR (e -value $1e^{-3}$) (Pruitt et al., 2012). Full-length 16S rRNA gene sequences were reconstructed using EMIRGE (Miller et al., 2011) and matched against the SILVA database (v. 123) (Quast et al., 2013) using ARB (v. 6.0.2) (Ludwig et al., 2004). The nearly complete 16S rRNA gene sequences allowed us to validate sufficient sequencing depth by detecting abundant groups previously identified by community profiling (Supplementary Figures 2, 3). Metagenome assembly was done with MEGAHIT (v. 1.0.1) (Li et al., 2014). After examining a variety of different k-mer lengths, a minimum k-mer length of 33 and a maximum k-mer length of 73 were applied with 10-mer iterations. Contigs greater than 1000 bp were taxonomically binned using PHYLOPYTHIA S+ (Gregor et al., 2016). This involved the analysis of marker genes and k-mer profiles. The coverage of individual contigs was calculated using BBMAP (v. 35.02) (<http://sourceforge.net/projects/bbmap/>). Coding genes were predicted using PRODIGAL (v. 2.6.2) (Hyatt et al., 2012).

2.8. Recovery of Metagenome-Assembled Genomes

A two-step binning strategy was applied for the recovery of metagenome-assembled genomes (MAGs) (Supplementary

Figure 4). The procedure is described in detail in the Supplementary Methods. MAGs were examined for completeness and possible sequence contamination by CHECKM (v. 1.0.3) (Parks et al., 2015). MAGs more complete than 85% were considered high-quality draft genomes, while those more complete than 70% were considered good-quality draft genomes. The categorization was based on the presence of lineage-specific single-copy marker genes and a sequence contamination of less than 10%. High- and good-quality draft genomes were subjected to genome annotation. MAGs of lesser quality were processed to determine their taxonomic affiliation. Genome annotations were done using PROKKA (v. 1.11) (Seemann, 2014) and RAST (v. 2.0) (Aziz et al., 2008; Overbeek et al., 2014) and validated against each other. Annotated genomes were analyzed using ARTEMIS (v. 16.0) (Carver et al., 2012). Gene predictions were examined by parsing the output from UBLAST searches in MEGAN (v. 5) (Huson et al., 2011) and by generating KEGG (Kyoto Encyclopedia of Genomes and Genes) annotations by subjecting coding genes to GHOSTKOALA (v. 2.0) (Kanehisa et al., 2015). The phylogenetic affiliation of recovered MAGs was determined by identified 16S rRNA genes and by concatenated alignments of 31 single-copy marker genes using AMPHORA (v. 2.0) (Wu and Scott, 2012). A maximum-likelihood tree was calculated using FASTTREE (v. 2.1.3) (Price et al., 2010) and the WAG model (Whelan and Goldman, 2001). Nearest neighbor interchange was applied for optimizing tree topology and the CAT approximation was used to account for evolutionary rate heterogeneity (Stamatakis, 2014). Genes of interest (e.g., genes encoding Czc efflux pumps) were extracted from the genomes (nucleotide sequences and deduced amino acid sequences) and queried against Interpro (v. 62.0), Pfam (v. 30.0) and Tigrfam (v. 15.0) for more detailed analyses (Haft et al., 2003; Mitchell et al., 2015; Finn et al., 2016).

2.9. Statistical Analyses

Statistical analyses were carried out using the R software framework (v. 3.2.2) (R Development Core Team, 2008) and the packages APE (v. 3.3) (Paradis et al., 2004), METAGENOMESEQ (v. 1.13.02) (Paulson et al., 2016), PHYLOSEQ (v. 1.10) (McMurdie and Holmes, 2013), and VEGAN (v. 2.3.1) (Oksanen et al., 2016), including their respective dependencies. Individual analyses are described in more detail in the Supplementary Methods.

2.10. Figure Generation

Figures were generated using the R packages GGPlot2 (<http://ggplot2.org>) (Wickham, 2009), as well as the python module MATPLOTLIB (<http://matplotlib.org/>) (Hunter, 2007).

2.11. Sequence Data Deposition

Amplicon sequencing and shotgun metagenomics data, and genome bins of interest were deposited at NCBI SRA and NCBI Genomes and are accessible under the following BioProject/Genome accessions: PRJNA384238 (amplicon sequencing data), PRJNA384362 (shotgun metagenomics data), NFUN00000000 (RH2 MAG17b), NFUO00000000 (RH1 MAG20).

3. RESULTS

3.1. Bacterial Community Profiling

Both 16S rRNA gene-targeted quantitative PCR and DNA extraction efficiency showed that the recovery of microbial biomass was 10-fold lower for slag samples RH1-3 ($2.20 - 2.89 \times 10^8$ copies/g dry weight [DW] soil; $4.45-8.24 \mu\text{g DNA/g DW}$) than for the undisturbed forest soil (Ref: 2.21×10^9 copies/g DW, $45.18 \mu\text{g DNA/g DW}$). The corresponding values for RHP were as follows: 1.51×10^8 copies/g DW, $22.53 \mu\text{g DNA/g DW}$ (Supplementary Figure 5). Chao1 and Shannon indices revealed that alpha diversity and evenness of the bacterial communities in slag deposits and pond sediment were lower than in undisturbed forest soil (Supplementary Figure 6). Lowest alpha diversity and evenness values were calculated for RH2 and RH3. Beta diversity analysis showed that the bacterial communities in samples RH1-3, RHP, and Ref differed in structural composition (Supplementary Figure 6). In particular, the communities in slag deposits (RH1-3) clustered clearly distinct from those in pond sediment (RHP) and forest soil (Ref), as revealed by Jensen-Shannon divergences.

Basic characteristics of amplicon datasets are given in Supplementary Table 1. Bacterial communities were dominated by four phyla (range of averaged relative abundances in sample-specific amplicon sequencing datasets are given in parenthesis): Acidobacteria (6.8–26.8%), Actinobacteria (14.6–33.0%), Chloroflexi (5.3–26.8%) and Alphaproteobacteria (18.3–42.0%) (Figure 1A). Acidobacteria and Actinobacteria showed a divergent pattern, with Acidobacteria being enriched in site RH1 (ANOVA, $p \leq 0.05$), while Actinobacteria were most prevalent in sites RH2 and RH3. Chloroflexi showed increased abundances in slag deposits (RH1-3, $p \leq 0.05$), while Alphaproteobacteria were strongly enriched in pond sediment (RHP). Less-abundant phyla included Planctomycetes (1.8–7.6%), Cyanobacteria (0.1–1.6%), and Bacteroidetes (0.1–6.0%). The latter groups, however, were detected only in sites RHP and Ref. A more detailed taxonomic analysis showed that all four dominant phyla were represented mainly by uncharacterized order- and family-level groups (Figure 1B). These groups included: DA052, KF-JG30-18 (Acidobacteria [subdivision 2, subdivision 13]), TM214 (Actinobacteria [Acidimicrobiales]), DA111 (Alphaproteobacteria [Rhodospirillales]), and JG37-AG-4 (Chloroflexi). Both DA052 and KF-JG30-18 were significantly enriched in all three RH sites ($p \leq 0.05$), with DA052 being most abundant in site RH1. TM214 was detected at all sites, but most pronounced at sites RH2, RH3, and RHP ($p \leq 0.05$). Like DA052 and KF-JG30-18, DA111 and JG37-AG-4 were significantly enriched and showed high abundances in all slag samples.

In order to gain a more resolved phylogenetic analysis, we made the attempt to reconstruct full-length 16S rRNA gene sequences from our metagenomic datasets. We recovered nearly complete (≥ 1200 bp) 16S rRNA gene sequences for all the most abundant groups: DA052 (9 sequences), KF-JG30-18 (2), DA111 (10), JG37-AG-4 (13), and TM214 (22). Phylogenetic analysis showed that all five groups—DA052, KF-JG30-18, DA111, JG37-AG-4, and TM214—were related to 16S rRNA gene sequences

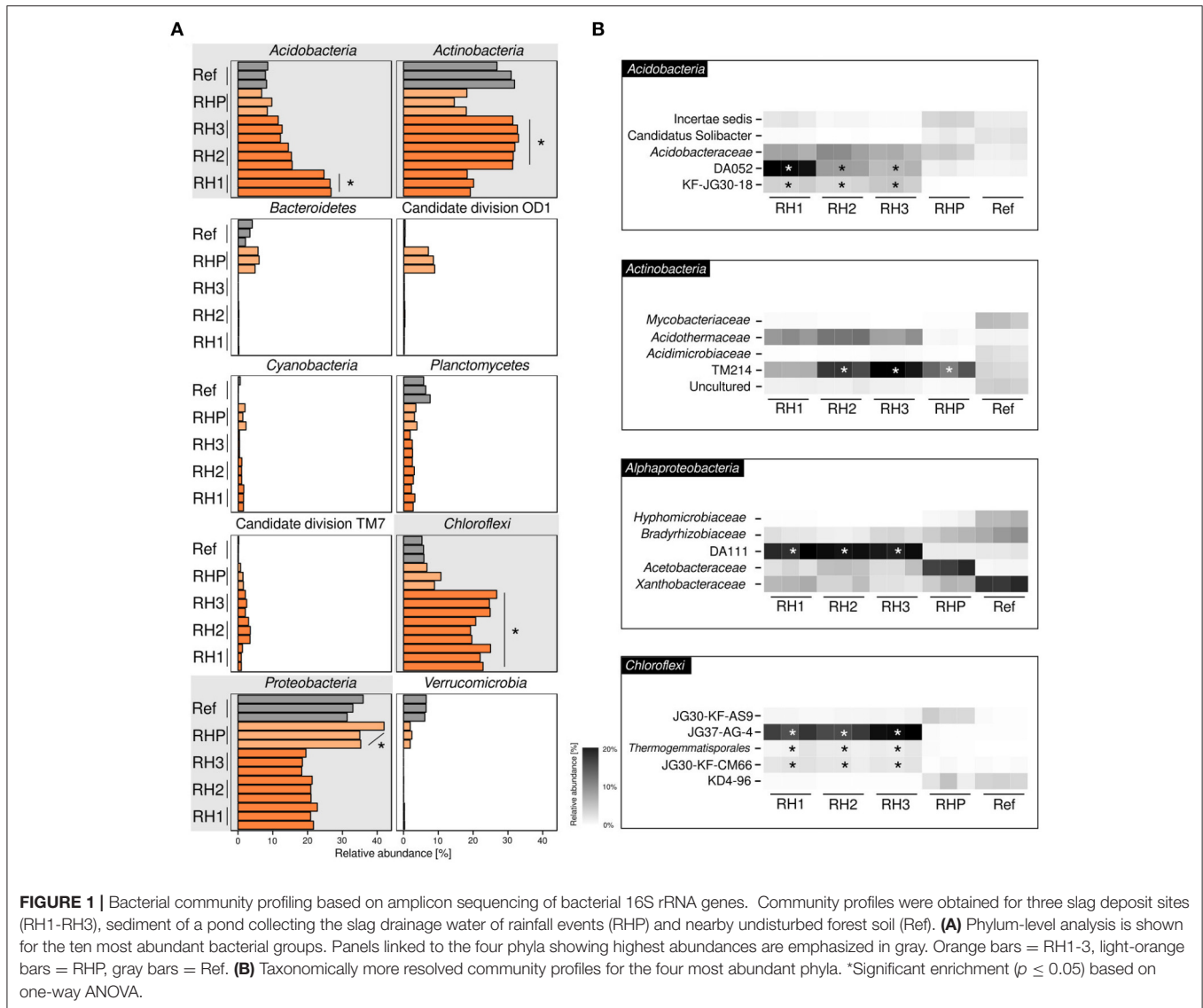
previously recovered from acidic forest soil. DA052, KF-JG30-18 and TM214 also showed a close affiliation to uncultured bacteria detected in volcanically impacted soil systems and, in addition, TM214 to populations in iron mineral-rich environments. Uncultured relatives of DA052, DA111, and JG37-AG-4 were also detected in metal-contaminated sites (Supplementary Table 2).

3.2. Geochemical Characteristics

The differences in bacterial community structure between RH1-3 and RHP/Ref prompted us to collect various metadata to place the community profiles in a broader context. Given the unusual industrial history of the slag deposits, we were especially interested in the metal content of the slag and the possible presence of xenobiotics. Elemental analysis revealed that the content of sulfur (37.33–46.31 mg/g DW soil), aluminum (21.57–27.23 mg/g), and calcium (24.12–33.26 mg/g) was significantly (ANOVA, $p \leq 0.05$) increased in the slag deposits (RH1-3), relative to the undisturbed forest soil (Ref) (sulfur: 0.77 mg/g, aluminum: 9.85 mg/g, calcium: 3.32 mg/g) (Figure 2A). The pond site showed a strongly elevated level of iron (48.84 mg/g) ($p \leq 0.05$) (RH1-3: 13.88–17.52 mg/g, Ref: 3.36 mg/g). BTEX compounds (benzene, toluene, ethylbenzene and xylene), PAHs (polycyclic aromatic hydrocarbons) and heavy metals were not detected. Samples from the slag deposits (RH1-3) had a low pH of 3.4–3.6. The pond sediment (RHP) was less acidic (Figure 2B).

The attempt to relate the observed alpha diversity at the individual sampling sites to particular environmental parameters (aluminum, calcium, iron, manganese, sulfur) showed no or poor correlations, with the exception of aluminum (correlation = 0.44) (Figure 2B). We, however, detected significant correlations between the abundance of particular bacterial phyla and environmental parameters, such as sulfur, iron and aluminum ($p \leq 0.05$, $r^2 \geq 0.99$), based on multiple regression analyses (Figures 2C,D). The pH was not an independent parameter, but significantly linked to the available sulfur and aluminum ($p \leq 0.05$, $r^2 \geq 0.99$). As revealed by Jensen-Shannon divergences, the correlations between the observed differences in bacterial community structure and environmental parameters were in the range of 0.3–0.48 (Figure 2C). Combining the different environmental parameters into a best-fit model, however, yielded a strong correlation of 0.96.

Given the dominance of yet-uncharacterized groups (Figure 1B) and greatest differences in both bacterial community profiles and environmental parameters between slag deposits (RH1-3) and non-slag samples (RHP + Ref), we determined Kendall's tau coefficients to identify links between particular taxonomic groups and single environmental parameters (Figure 2D). Among the Acidobacteria, the higher relative abundance of most subphylum groups, including DA052 and KF-JG30-18, was found to be linked to aluminum, iron and manganese ($p \leq 0.05$ [corrected according to Benjamini and Hochberg, 1995]). DA111 (Alphaproteobacteria), but also JG30-KF-CM66 and KD4-96 (Chloroflexi), showed a strong correlation with aluminum and iron ($p \leq 0.05$). The Actinobacteria showed high correlations with all recorded environmental parameters. The significance of these correlations, however, has to be



considered low, given that members of this phylum showed no distinct distribution pattern between the different sampling sites (Figures 1A,B). The high aluminum content at the three slag deposition sites (RH1-3) (Figure 2A) and the relatively low pH (Figure 2B) may point to a high bioavailability of toxic aluminum species (Supplementary Figures 7, 8).

3.3. Reconstruction of Metagenome-Assembled Genomes

An overview of processed metagenome datasets is given in Supplementary Table 3. Due to the co-dominance of yet-uncharacterized Acidobacteria groups on subphylum level (Figure 1), we were interested in elucidating both their metabolic potential and their specific adaptations favoring their persistence in slag deposits. Application of a two-step binning strategy (Supplementary Figure 4) allowed us to reconstruct 13 Acidobacteria MAGs (RH1: 6, RH2: 5, RHP: 2) (Supplementary Table 4). A major proportion of MAGs was of high quality

in terms of completeness (8 out of 13, >80% complete) and contamination (9 out of 13, <10% contaminated) (Figure 3). Based on treeing analysis of a concatenated alignment of 31 single-copy marker genes, a group of 5 MAGs (RH2 MAG 17b, RH2 MAG 17c, RH1 MAG 19a, RH1 MAG 19b, and RH2 MAG21) had an uncertain affiliation with no close relationship to any described genus. 16S rRNA gene sequences could be identified for 6 MAGs. Based on identified 16S rRNA gene sequences it was possible to link two genomes to uncharacterized taxonomic groups (DA052 and KF-JG30-18) that were found to be dominant in the bacterial community profiles (Figure 1B). These two genomes were analyzed in greater detail (RH1 MAG 20 [completeness: 94%]: DA052 [Acidobacteria, Subdivision 2]; RH2 MAG 17b [completeness: 89%]: KF-JG30-18 [Acidobacteria, Subdivision 13]). The focus of our genome analysis was on carbon utilization capabilities and adaptations favoring colonization of Red Hill slag deposits.

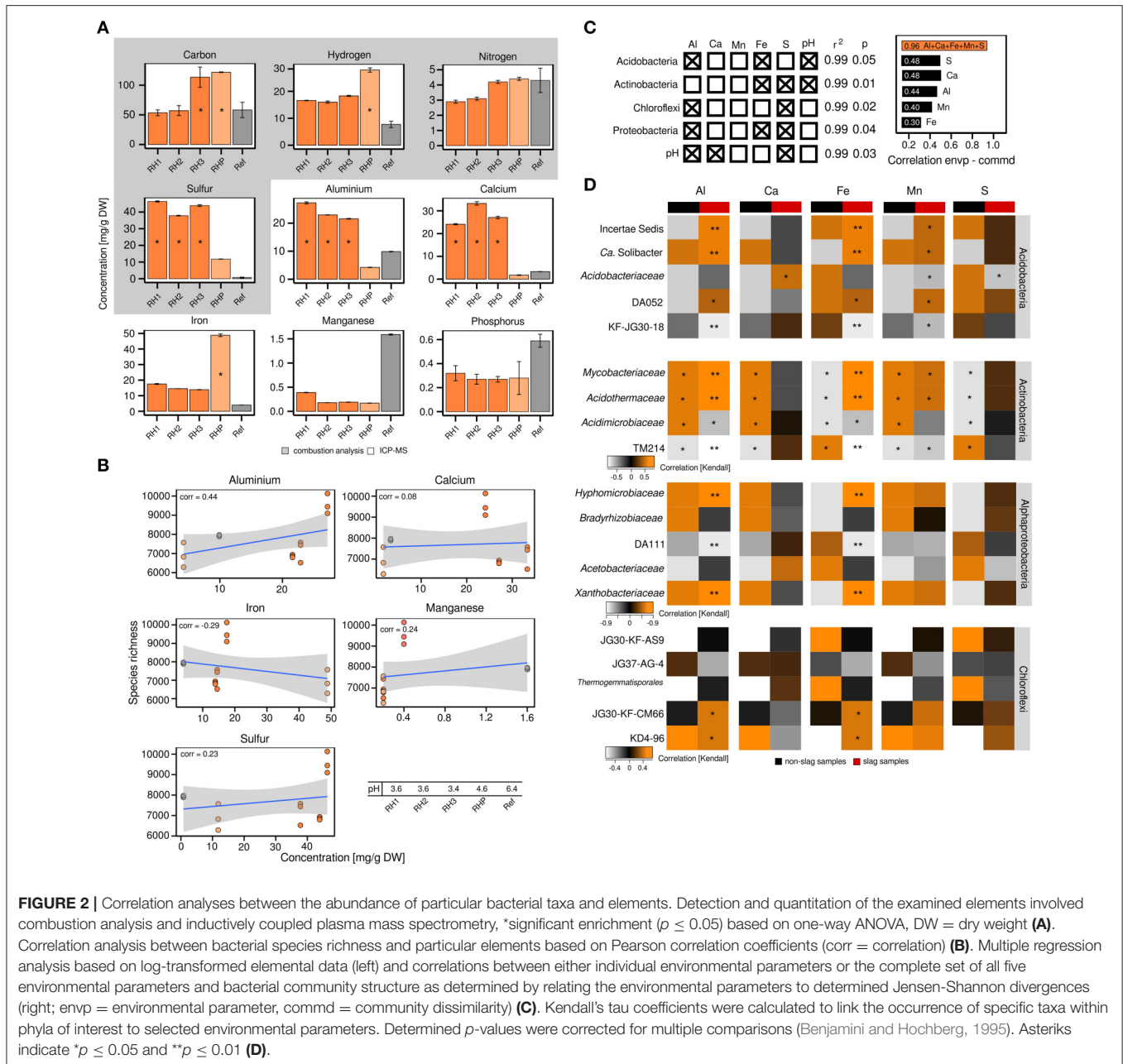


FIGURE 2 | Correlation analyses between the abundance of particular bacterial taxa and elements. Detection and quantitation of the examined elements involved combustion analysis and inductively coupled plasma mass spectrometry, *significant enrichment ($p \leq 0.05$) based on one-way ANOVA, DW = dry weight (A). Correlation analysis between bacterial species richness and particular elements based on Pearson correlation coefficients (corr = correlation) (B). Multiple regression analysis based on log-transformed elemental data (left) and correlations between either individual environmental parameters (corr = correlation) or the complete set of all five environmental parameters and bacterial community structure as determined by relating the environmental parameters to determined Jensen-Shannon divergences (right; envp = environmental parameter, commd = community dissimilarity) (C). Kendall's tau coefficients were calculated to link the occurrence of specific taxa within phyla of interest to selected environmental parameters. Determined p -values were corrected for multiple comparisons (Benjamini and Hochberg, 1995). Asterisks indicate * $p \leq 0.05$ and ** $p \leq 0.01$ (D).

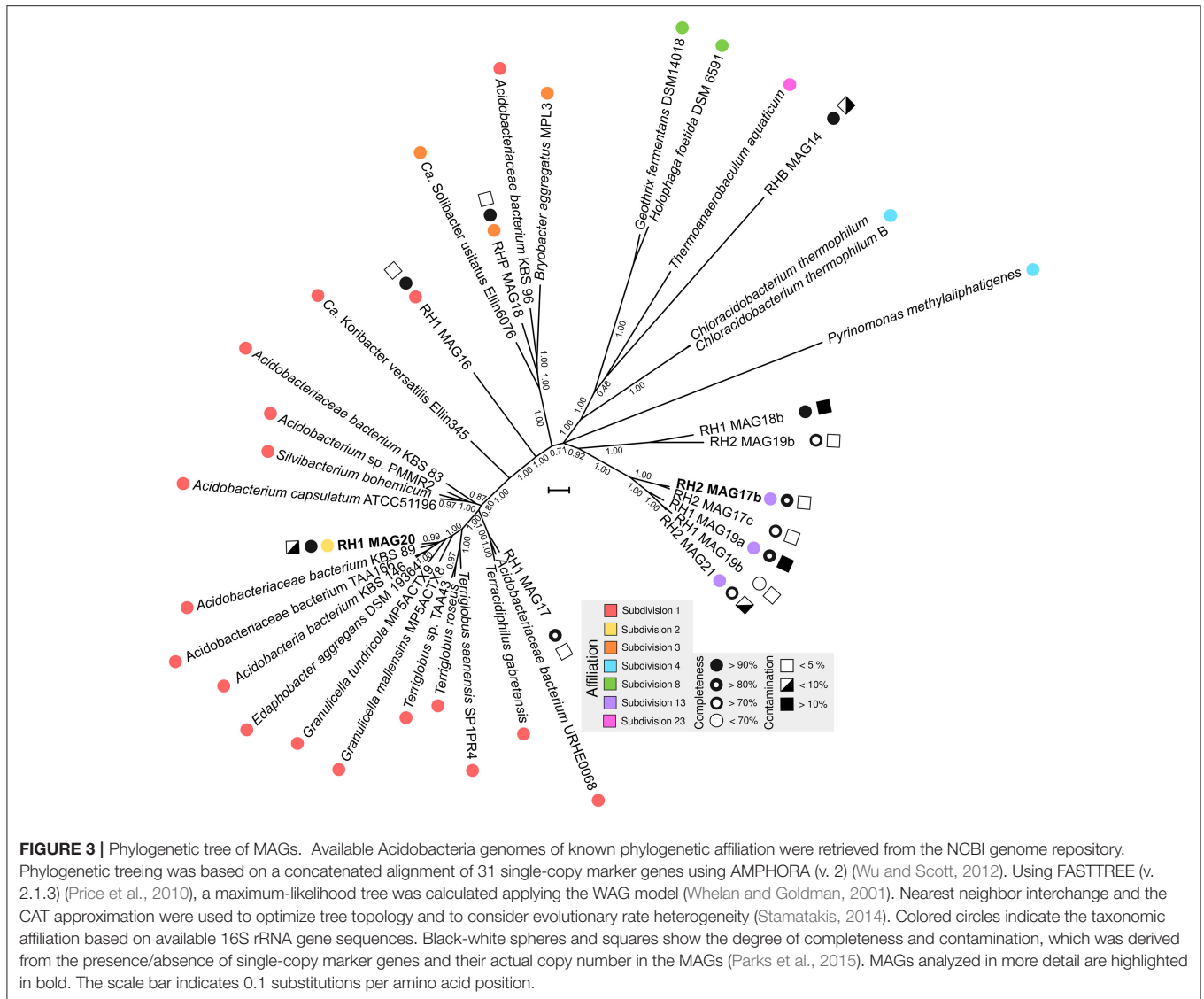
3.4. MAG Analysis of KF-JG30-18 and DA052

Both Acidobacteria MAGs encode key pathways for central carbohydrate metabolism, most pathways for the synthesis of proteinogenic amino acids and complete respiratory chains, thereby suggesting a heterotrophic lifestyle (Figure 4). In case of KF-JG30-18, the failure to detect the complete gene sets encoding glycolysis and TCA cycle can be attributed to the incompleteness of its MAG (Supplementary Table 4).

3.4.1. KF-JG30-18

RH2 MAG 17b showed a total length of 2.8 Mb with a GC content of 58% (Table 1). The KF-JG30-18 representative has the genetic potential to utilize the polysaccharides cellulose

and starch as carbon source and nitrite as additional nitrogen source (Figure 4A). While no sulfate adenylylkinase (EC 2.7.1.25) gene is present, KF-JG30-18 carries all the other genes for assimilatory sulfate reduction. The presence of genes encoding CusS/CusR and HydH/HydG two-component systems enables sensing of extracellular metal cations. BaeS/BaeR and LiaS/LiaR two-component systems allow detection of cell envelope stress. KF-JG30-18 possesses 11 genes linked to CzcCBA efflux systems, which belong to the multifunctional family of type VI secretion systems. It also contains various transporter genes including those for molybdate, cobalt/nickel, branched-chain amino acid uptake systems, and heme export systems. In addition, genes encoding high-affinity phosphate uptake systems and genes encoding enzymes facilitating polyphosphate formation were



detected for KF-JG30-18. Finally, two CRISPR arrays were identified (Table 1).

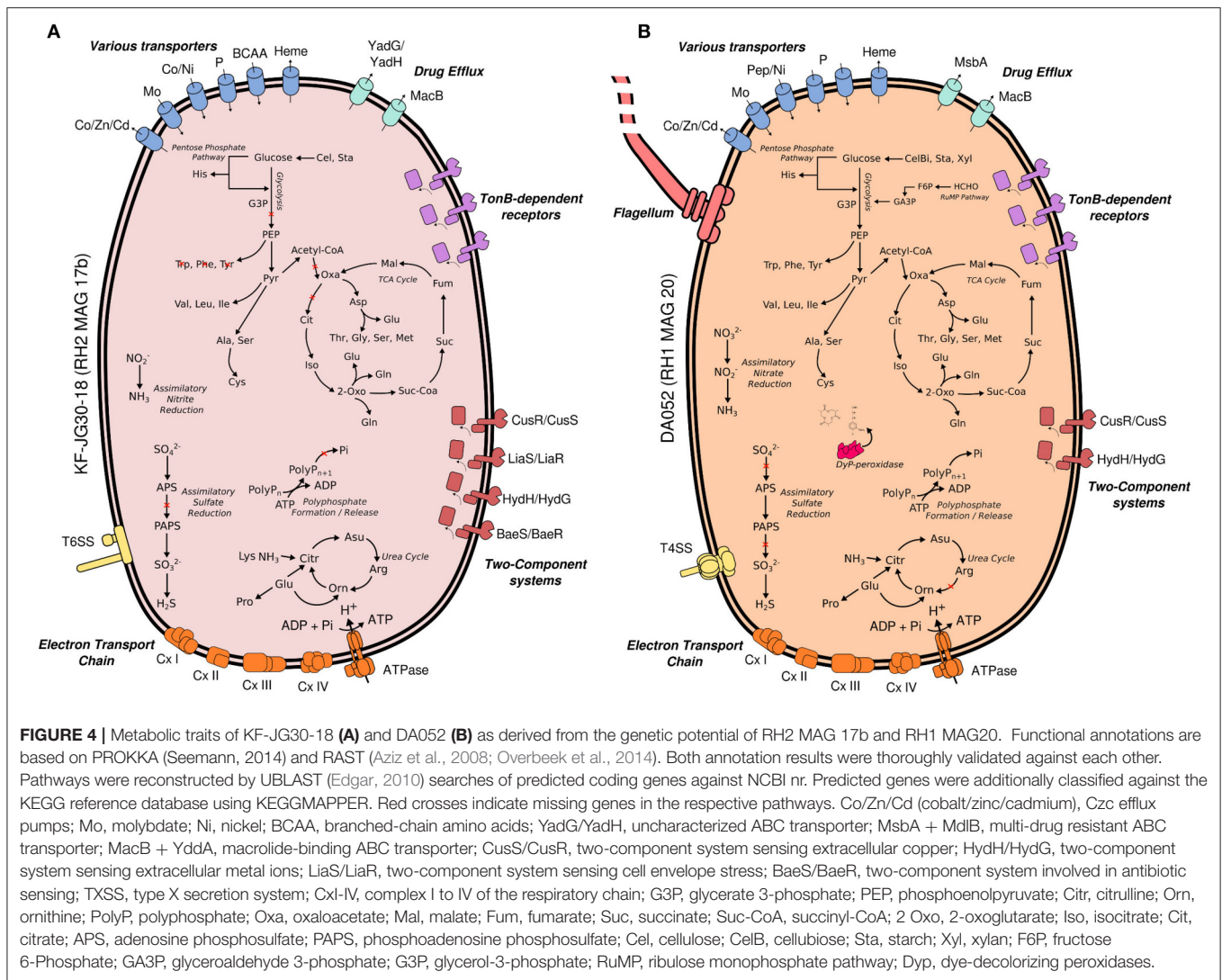
3.4.2. DA052

The RH1 MAG 20 genome was substantially bigger in comparison to KF-JG30-18 and showed a total length of 4.9 Mb featuring a GC content of 59% (Table 1). In contrast to KF-JG30-18, the DA052 representative (RH1 MAG 20) has the capacity to express the ribulose monophosphate pathway, suggesting the potential to metabolize C1 compounds by formaldehyde fixation. The RH1 MAG 20 also possesses genes involved in the breakdown of cellobiose, starch and xylan. In addition, genes that encode dye-decolorizing peroxidases were detected. Like KF-JG30-18, DA052 has the potential to utilize oxidized nitrogen species such as nitrite and nitrate as additional nitrogen sources (Figure 4B). Albeit no adenylyltransferase (EC 2.7.7.4) gene could be identified, DA052 is likely able to carry out assimilatory sulfate reduction. The complete set of genes required

for flagellar motility and type IV secretion was identified in RH1 MAG 20. It also encodes high-affinity phosphate uptake systems, is equipped with the potential to form and release polyphosphate, and contains multiple genes that code for phosphate-binding subunits. Genes involved in sensing extracellular metal cations were detected, such as those encoding CusS/CusR and HydH/HydG two-component systems. DA052 contains even more genes associated with CzcCBA efflux systems (29 genes) than KF-JG30-18 (11). Worth mentioning is also the high number of transposase genes and genes encoding mobile element proteins (29) (Table 1).

4. DISCUSSION

The slag samples contained high amounts of iron and sulfur (Figure 2A), presumably in mineral form. This was unexpected given the age (150 years) and leached nature of the slags.



Assuming a continuous exposure to oxygen and humidity, an oxidation of present sulfide-mineral components was anticipated. Sulfide-rich mine tailings show a zonation over time: a sulfide-depleted zone; an active, transitory oxidation zone; and a primary zone representing the original conditions (Diaby et al., 2007; Tan et al., 2008; Huang et al., 2011). In our study, the slags originated from soft coal with a high content of pyrite. Organic carbon in soft coal and, as a consequence, in the studied slags primarily consists of humic and fulvic acids (Peuravuori et al., 2006). Humic acids were repeatedly reported to be efficient in pyrite passivation (Lalvani et al., 1996; Belzile et al., 1997; Chen et al., 1999), either by chelating released ferric iron with their negatively charged functional (carboxylic, phenolic) groups or by direct pyrite adsorption. Both processes may explain why the oxidation of sulfide minerals in the slag deposits was mitigated. This low oxidation level clearly differentiates the Red Hill slag deposits from mine tailings and other AMD environments.

Aluminum was found to be a major driver of bacterial community structure in the leached slags (Figures 2B–D). The

high aluminum content highlights the inefficiency of the early-industrial leaching process. Given the low pH (3.4–3.6) of the slags, aluminum was likely present as cations (Supplementary Figure 7). Cationic aluminum is highly toxic because of potential interferences with ATP metabolism, lipid metabolism and membrane transport processes (Supplementary Figure 8). In soils, aluminum toxicity can be mitigated by clay minerals and organic acids, which act as a buffer system due to their cation exchange capacity (Turpault et al., 1996).

The chelating capabilities of humic and fulvic acids may lead to the immobilization of aluminum cations as described above for iron-sulfur minerals. Mohan and Chander (2006) showed that brown coal is an efficient adsorbent for metal cations in AMD discharges. Given a passivation of iron-sulfur minerals and aluminum, their role as drivers for bacterial community structure can be considered indirect, albeit toxic effects of aluminum due to minor releases may occur. Previous research had revealed that the retention of metal cations can alter the accessibility and utilizability of humic and fulvic acids as potential carbon sources

TABLE 1 | Characteristics of RH1 MAG 20 (DA052) and RH2 MAG 17b (KF-JG30-18).

	RH1 MAG20 (DA052)	RH2 MAG 17b (KF-JG30-18)
Total length [bp]	4,885,294	2,798,462
No. of recruited reads	404,720 406,649	254,272 254,736
Coverage	18	18
No. contigs	760	169
N50 [bp]	58,735	25,655
GC content [%]	59	58
Completeness [%]	93.97	89.39
Contamination [%]	5.15	2.14
Classification*	Nearly complete genome with medium contamination	Substantial genome with low contamination
No. of coding genes	5,311	2,822
No. of rRNA operons	1	1
No. of tRNAs	44	35
Mobile elements**	29	3
CRISPR arrays	–	2
No. of encoded Czc efflux pumps	29	11
No. of encoded Acriflavine resistance proteins	5	6

Recruited reads refer to the number of reads incorporated during the assembly of respective contigs. Numbers represent forward and reverse reads. Coverage was determined by determining the average coverage of genome bin contigs. Genome completeness and contamination have been assessed using CHECKM (Parks et al., 2015). The number of coding genes, rRNA operons, tRNAs, mobile elements, CRISPR/CAS repeats and repeat regions is given based on genome annotations done with PROKKA (v. 1.11) (Seemann, 2014) and RAST (v. 2.0) (Aziz et al., 2008; Overbeek et al., 2014) followed by manual refining. *According to Parks et al. (2015); **transposons and insertion sequences.

for microbes (Jones and Edwards, 1998; Jones et al., 2003). Considering that humic and fulvic acids are naturally rather inert and recalcitrant carbon sources (Kirk and Farrell, 1987), metal complexation would increase the supposed oligotrophic nature of the studied slag deposits.

In the leached slags, both microbial biomass and bacterial cell numbers were significantly lower than in the undisturbed forest soil (Supplementary Figure 5), but in the range previously determined for sulfidic mine tailings (Korehi et al., 2014). Five yet-uncultured subphylum groups were identified to be dominant in the slag deposits. These belonged to Acidobacteria, Actinobacteria, Alphaproteobacteria, and Chloroflexi (Figures 1A,B). None of the five groups have previously been reported to be highly abundant in AMD systems. Their first detection was in peatlands (DA052, DA111, TM214) (Rheims et al., 1996; Felske et al., 1998) and uranium waste piles (KF-JG30-18, JG37-AG-4, DA111) (Selenska-Pobell et al., 2001; Selenska-Pobell, 2002). Phylogenetic analysis of nearly complete 16S rRNA gene sequences also revealed relationships to populations present in acidic, metal-rich, and contaminated soil environments (Supplementary Table 2). In accordance with the geochemical differences discussed above, non-occurrence of bacteria known to be involved in iron- and sulfur-cycling again

clearly differentiates the slag deposits from AMD systems and mine tailings.

Given that cultured representatives are not yet available for the dominant groups identified in the Red Hill slag deposits, their metabolic capabilities remained elusive. Therefore, high abundances of Acidobacteria subdivisions 2 (DA052) and 13 (KF-JG30-18) prompted us to conduct a metagenomic approach in order to get first insights into their genetic potential. Acidobacteria are known to be abundant members of terrestrial microbiomes (Janssen et al., 2002; Sait et al., 2006) and to be phylogenetically highly diverse (Barns et al., 2007). A total of 22 genera have been described, which cover only 8 of 26 subdivisions proposed until today. Genome sequences are available for members of 5 subdivisions (1, 3, 4, 8, and 23; Kielak et al., 2016). Acidobacteria are believed to have an oligotrophic (K-strategist) lifestyle. This perception, however, may be biased due to the low number of acidobacterial isolates of which most belong to subdivision 1. Members of subdivisions 2 and 13 have repeatedly been detected in higher abundances in U-contaminated environments (Selenska-Pobell et al., 2001; Barns et al., 2007). The occurrence of subdivision 2 populations in Amazonian forest soils was correlated to the microbial availability of CO₂, Fe, and Al³⁺. Thus, it was hypothesized that members of subdivision 2 are tolerant to aluminum (Navarrete et al., 2013; Catão et al., 2014). Adaptation and tolerance to Al³⁺ could also be one major reason for the high relative abundance of DA052 in our sampling sites, especially in RH1. Amazonian forest soils are commonly rich in biochar whose chemical characteristics are remotely similar to those of studied soft coal slags. However, a recent study of biochar microbiomes and metagenomes revealed a rather low abundance of Acidobacteria (Noyce et al., 2016). Given the limited knowledge of yet uncharacterized Acidobacteria groups, the MAGs obtained for subdivisions 2 (DA052) and 13 (KF-JG30-18) allowed us to infer valuable information on the physiology and metabolic potential of these two Acidobacteria subdivisions.

As of November 2016, 27 genome sequences were available for Acidobacteria, with most of them being affiliated with subdivision 1 Acidobacteria. This number has recently more than doubled by deep metagenomic sequencing of an aquifer system (Anantharaman et al., 2016). However, to the best of our knowledge, no genomes of Acidobacteria subdivisions 2 and 13 have yet been reported, highlighting the novelty of the draft genomes reconstructed for DA052 (RH1 MAG20) and KF-JG30-18 (RH2 MAG 17b). The genome size of DA052 is comparable to those of subdivision 1 Acidobacteria such as *Terriglobus roseus* (4.9 Mb, 60% GC) (Rawat et al., 2014). The KF-JG30-18 genome is comparably small, being in the range of the only subdivision 23 genome (*Thermoanaerobaculum aquaticum* 2.7 Mb) (Stamps et al., 2014). Apparently, members of DA052 and KF-JG30-18 are able to utilize polysaccharides such as cellulose and xylan (Figure 4). Comparative genomics suggests that the ability of polysaccharide breakdown is widely distributed among the Acidobacteria (Kielak et al., 2016), but to date experimental evidence for polysaccharide utilization was obtained only for members of subdivisions 1, 3, and 4 (Dedysh et al., 2012; Pascual et al., 2015; Garcia-Fraile et al., 2016; Jiang et al., 2016; Lladó et al., 2016). Xylan degradation appears to

be a rather common metabolic trait of the Acidobacteria, while cellulose breakdown has been shown only for *Telmatobacter bradus* (Pankratov et al., 2008). Plant cell wall constituents such as cellulose or xylan are commonly identified in soft coal in varying quantities dependent on the degree of coalification (del Río et al., 1994; Rumpel et al., 1998). The DA052 draft genome RH1 MAG20 contains all the genes required for formaldehyde fixation via the cyclic ribulose monophosphate pathway (Figure 4). Thus, members of DA052 may be able to utilize C1 compounds as carbon source. These are present in coal as methyl- and methoxyl-groups or as formate (Stafford, 1988; Stout et al., 1988; Hatcher and Clifford, 1997). A survey of the available Acidobacteria genomes revealed key genes of the ribulose monophosphate pathway (3-hexulose-6-phosphate synthase and 6-phospho-3-hexuloisomerase) in *Edaphobacter aggregans* and *Terriglobus roseus*, but both organisms have not yet been shown to utilize C1 compounds (Eichorst et al., 2007; Koch et al., 2008). In the DA052 representative, we also identified the genetic potential to produce lignin-degrading dye-decolorizing peroxidases. This family of enzymes is of increasing interest as it represents the bacterial counterpart of lignin-cleaving peroxidases and laccases known from fungi (de Gonzalo et al., 2016). The ability to utilize rather inert, lignin-derived carbon sources would give DA052 a selective advantage as lignin-derived carbon is assumed to be abundantly present in the studied slags.

Although most aluminum in the Red Hill slag deposits is assumed to be immobilized by humic and fulvic acids, some bioavailability cannot be excluded. Both draft genomes, DA052 and KF-JG30-18, contain genes encoding high-affinity phosphate uptake systems, while the potential to form and release polyphosphate was detected specifically for DA052. The intracellular cleavage of polyphosphate and release of phosphate moieties was described for bacteria and archaea (Remonsellez et al., 2006; Rao et al., 2009; Navarro et al., 2013) as a measure to inactivate toxic metal cations by precipitation. Since inorganic phosphate has manifold roles in cellular metabolism, the sole presence of these genes cannot be exclusively attributed to metal resistance. The draft genomes of both DA052 and KF-JG30-18, however, harbor an unexpectedly high number of genes encoding Czc efflux pumps (Figure 4). These mediate resistance to cobalt, zinc, and cadmium (Nies, 1995; Chen et al., 2015). Using InterPro and PFAM, a closer inspection of the encoded proteins revealed that 6 and 5 sequences for, respectively, KF-JG30-18 and DA052 are closely related to acriflavine resistance proteins, which, like Czc efflux pumps, belong to the resistance/nodulation/division superfamily of solute transporters (Table 1). Acriflavine resistance proteins are multi-drug efflux systems with a tremendously broad substrate range, including different classes of antibiotics, detergents and small organic molecules (Blair and Piddock, 2009; Pos, 2009). Acriflavine was first extracted from smoldered coal tar by precipitation with diluted sulfuric acid, a procedure that is reminiscent of the early industrial mineral leaching from soft coal (Supplementary Figure 1). Paul Ehrlich found acriflavine to be bacteriocidal (Wainwright, 2001; Kumar et al., 2012) and intercalation into DNA was identified as its mode of action (Lerman, 1961). It has been shown for iron-oxidizing bacteria

that the number of acriflavine resistance genes can have a strong effect on their fitness when being exposed to metals, metalloids, or hazardous organic compounds (Emerson et al., 2013). It is thus tempting to hypothesize that the presence of such substances in the leached slags selects for acriflavine-tolerant bacteria. In consequence, the presence of multiple genes encoding acriflavine resistance proteins could thus represent a competitive advantage, favoring the colonization of studied slags by DA052 and KF-JG30-18.

In summary, our analyses revealed that the microbiology of the studied slags greatly differs from commonly studied AMD and mine tailing environments. The unusual geochemistry, including the proposed mitigation of mineral oxidation and subsequent metal release, increases the oligotrophic nature of the Red Hill slag deposits. Members of the elusive Acidobacteria subdivisions 2 (DA052) and 13 (KF-JG30-18) showed a dominant occurrence not observed in any other previously studied environment. Metagenome-assembled genomes allowed us to identify broad carbon utilization capabilities and, more intriguingly, pronounced metal and xenobiotic detoxification mechanisms. These may explain the dominant occurrence of DA052 and KF-JG30-18. Our findings call for further research into the microbiology of the Red Hill slag deposits, in particular for activity-centered approaches such as metatranscriptomics and -proteomics that allow to link both genome coding potential and actual metabolic activity of identified groups in this unique, early-industrial, man-made habitat.

AUTHOR CONTRIBUTIONS

CW and WL designed research. CW carried out research, analyzed data, wrote code. CW and WL wrote the manuscript together.

FUNDING

Funding was provided by the German Research Foundation (DFG SFB 987, Microbial Diversity in Environmental Signal Response, Project A05 to WL).

ACKNOWLEDGMENTS

The authors wish to express their sincere gratitude to Walter Wilhelm Jungmann (Department of Geology) for experimental advice, and to Christoph Weihrauch (Department of Geology), Daniel Schindler (SYNMIKRO) and Heike Mallinger (Department of Chemistry) for excellent technical assistance. All of them are members of the Philipps-Universität Marburg (Germany). In addition, the authors would like to thank Horst Wolfgarten for in-depth information about soft coal mining in the surroundings of Bonn.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01023/full#supplementary-material>

REFERENCES

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., et al. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* 7, 1–11. doi: 10.1038/ncomms13219
- Arroyo, F., and Fernández-Pereira, C. (2008). Hydrometallurgical recovery of germanium from coal gasification fly ash. Solvent extraction method. *Ind. Eng. Chem. Res.* 47, 3186–3191. doi: 10.1021/ie7016948
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75
- Baker, B. J., and Banfield, J. F. (2003). Microbial communities in acid mine drainage. *FEMS Microbiol. Ecol.* 44, 139–152. doi: 10.1016/S0168-6496(03)00028-X
- Barns, S. M., Cain, E. C., Somerville, L., and Kuske, C. R. (2007). Acidobacteria phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum. *Appl. Environ. Microbiol.* 73, 3113–3116. doi: 10.1128/AEM.02012-06
- Belzile, N., Maki, S., Chen, Y. W., and Goldsack, D. (1997). Inhibition of pyrite oxidation by surface treatment. *Sci. Total Environ.* 196, 177–186. doi: 10.1016/S0048-9697(96)05410-1
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* 57, 289–300.
- Blair, J. M., and Piddock, L. J. (2009). Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. *Curr. Opin. Microbiol.* 12, 512–519. doi: 10.1016/j.mib.2009.07.003
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303
- Carver, T., Harris, S. R., Berriman, M., Parkhill, J., and McQuillan, J. A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28, 464–469. doi: 10.1093/bioinformatics/btr703
- Catão, E. C. P., Lopes, F. A. C., Araújo, J. F., De Castro, A. P., Barreto, C. C., Bustamante, M. M. C., et al. (2014). Soil acidobacterial 16S rRNA gene sequences reveal subgroup level differences between savanna-like cerrado and atlantic forest brazilian biomes. *Int. J. Microbiol.* 2014:156341. doi: 10.1155/2014/156341
- Chen, L.-X., Hu, M., Huang, L.-N., Hua, Z.-S., Kuang, J.-L., Li, S.-J., et al. (2015). Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME J.* 9, 1579–1592. doi: 10.1038/ismej.2014.245
- Chen, Y.-W., Belzile, N., and Goldsack, D. E. (1999). Passivation of pyrite oxidation by organic compounds. *Sudbury* 99, 1063–1071.
- de Gonzalo, G., Colpa, D. I., Habib, M. H. M., and Fraaije, M. W. (2016). Bacterial enzymes involved in lignin degradation. *J. Biotechnol.* 236, 110–119. doi: 10.1016/j.jbiotec.2016.08.011
- Dedysh, S. N., Kulichevskaya, I. S., Serkebaeva, Y. M., Mityaeva, M. A., Sorokin, V. V., Suzina, N. E., et al. (2012). *Bryocella elongata* gen. nov., sp. nov., a member of subdivision I of the Acidobacteria isolated from a methanotrophic enrichment culture, and emended description of *Edaphobacter aggregans* Koch et al. 2008. *Int. J. Syst. Evol. Microbiol.* 62(Pt 3), 654–664. doi: 10.1099/ijs.0.031898-0
- del Río, J. C., González-Vila, F. J., Martín, F., and Verdejo, T. (1994). Characterization of humic acids from low-rank coals by ¹³C-NMR and pyrolysis-methylation. Formation of benzenecarboxylic acid moieties during the coalification process. *Org. Geochem.* 22, 885–891. doi: 10.1016/0146-6380(94)90026-4
- Denef, V. J., Kalnejais, L. H., Mueller, R. S., Wilmes, P., Baker, B. J., Thomas, B. C., et al. (2010). Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2383–2390. doi: 10.1073/pnas.0907041107
- Diaby, N., Dold, B., Pfeifer, H. R., Holliger, C., Johnson, D. B., and Hallberg, K. B. (2007). Microbial communities in a porphyry copper tailings impoundment and their impact on the geochemical dynamics of the mine waste. *Environ. Microbiol.* 9, 298–307. doi: 10.1111/j.1462-2920.2006.01138.x
- Dick, G. J., Andersson, A. F., Baker, B. J., Simmons, S. L., Thomas, B. C., Yelton, A. P., et al. (2009). Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* 10:R85. doi: 10.1186/gb-2009-10-8-r85
- Druschel, G. K., Baker, B. J., Gihring, T. M., and Banfield, J. F. (1999). Acid mine drainage biogeochemistry at Iron Mountain, California. *Geochem. Trans.* 5:13. doi: 10.1063/1.1769131
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461
- Eichorst, S. A., Breznak, J. A., and Schmidt, T. M. (2007). Isolation and characterization of soil bacteria that define *Terriglobus* gen. nov., in the phylum Acidobacteria. *Appl. Environ. Microbiol.* 73, 2708–2717. doi: 10.1128/AEM.02140-06
- Emerson, D., Field, E. K., Chertkov, O., Davenport, K. W., Goodwin, L., Munk, C., et al. (2013). Comparative genomics of freshwater Fe-oxidizing bacteria: implications for physiology, ecology, and systematics. *Front. Microbiol.* 4:254. doi: 10.3389/fmicb.2013.00254
- Felske, A., Wolterink, A., Van Lis, R., and Akkermans, A. D. L. (1998). Phylogeny of the main bacterial 16S rRNA sequences in Drentse A grassland soils (The Netherlands). *Appl. Environ. Microbiol.* 64, 871–879.
- Finn, R. D., Cogill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344
- García-Fraile, P., Benada, O., Cajthaml, T., Baldrian, P., and Lladó, S. (2016). *Terracidiphilus gabretensis* gen. nov., sp. nov., an abundant and active forest soil Acidobacterium important in organic matter transformation. *Appl. Environ. Microbiol.* 82, 560–569. doi: 10.1128/AEM.03353-15
- Gregor, I., Dröge, J., Schirmer, M., Quince, C., and McHardy, A. C. (2016). *PhyloPythiaS+*: a self-training method for the rapid reconstruction of low-ranking taxonomic bins from metagenomes. *PeerJ* 4:e1603. doi: 10.7717/peerj.1603
- Haft, D. H., Selengut, J. D., and White, O. (2003). The TIGRFAMs database of protein families. *Nucleic Acids Res.* 31, 371–373. doi: 10.1093/nar/gkg128
- Hatcher, P. G., and Clifford, D. J. (1997). The organic geochemistry of coal: from plant materials to coal. *Org. Geochem.* 27, 251–274. doi: 10.1016/S0146-6380(97)00051-X
- Hua, Z.-S., Han, Y.-J., Chen, L.-X., Liu, J., Hu, M., Li, S.-J., et al. (2015). Ecological roles of dominant and rare prokaryotes in acid mine drainage revealed by metagenomics and metatranscriptomics. *ISME J.* 9, 1280–1294. doi: 10.1038/ismej.2014.212
- Huang, L. N., Zhou, W. H., Hallberg, K. B., Wan, C. Y., Li, J., and Shu, W. S. (2011). Spatial and temporal analysis of the microbial community in the tailings of a Pb-Zn mine generating acidic drainage. *Appl. Environ. Microbiol.* 77, 5540–5544. doi: 10.1128/AEM.02458-10
- Hunter, J. D. (2007). Matplotlib: a 2d graphics environment. *Comput. Sci. Eng.* 9, 90–95. doi: 10.1109/MCSE.2007.55
- Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N., and Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* 21, 1552–1560. doi: 10.1101/gr.120618.111
- Hyatt, D., LoCascio, P. F., Hauser, L. J., and Uberbacher, E. C. (2012). Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* 28, 2223–2230. doi: 10.1093/bioinformatics/bts429
- Janssen, P. H., Yates, P. S., Grinton, B. E., Taylor, P. M., and Sait, M. (2002). Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia*. *Appl. Environ. Microbiol.* 68, 2391–2396. doi: 10.1128/AEM.68.5.2391-2396.2002
- Jiang, Y. W., Wang, J., Chen, M. H., Lv, Y. Y., and Qiu, L. H. (2016). *Acidipila dinghuensis* sp. nov., an acidobacterium isolated from forest soil. *Int. J. Syst. Evol. Microbiol.* 66, 76–83. doi: 10.1099/ijssem.0.000676
- Jones, D. L., Dennis, P. G., Owen, A. G., and Van Hees, P. A. W. (2003). Organic acid behavior in soils - misconceptions and knowledge gaps. *Plant Soil* 248, 31–41. doi: 10.1023/A:1022304332313

- Jones, D. L., and Edwards, A. C. (1998). Influence of sorption on the biological utilization of two simple carbon substrates. *Soil Biol. Biochem.* 30, 1895–1902. doi: 10.1016/S0038-0717(98)00060-1
- Kanehisa, M., Sato, Y., and Morishima, K. (2015). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006
- Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A., and Kuramae, E. E. (2016). The ecology of Acidobacteria: moving beyond genes and genomes. *Front. Microbiol.* 7:744. doi: 10.3389/fmicb.2016.00744
- Kirk, T. K., and Farrell, R. L. (1987). Enzymatic “combustion”: the microbial degradation of lignin. *Annu. Rev. Microbiol.* 41, 465–501. doi: 10.1146/annurev.mi.41.100187.002341
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41:e1. doi: 10.1093/nar/gks808
- Koch, I. H., Gich, F., Dunfield, P. F., and Overmann, J. (2008). *Edaphobacter modestus* gen. nov., sp. nov., and *Edaphobacter aggregans* sp. nov., acidobacteria isolated from alpine and forest soils. *Int. J. Syst. Evol. Microbiol.* 58, 1114–1122. doi: 10.1099/ijs.0.65303-0
- Korehi, H., Blöthe, M., and Schippers, A. (2014). Microbial diversity at the moderate acidic stage in three different sulfidic mine tailings dumps generating acid mine drainage. *Res. Microbiol.* 165, 713–718. doi: 10.1016/j.resmic.2014.08.007
- Kuang, J.-L., Huang, L.-N., Chen, L.-X., Hua, Z.-S., Li, S.-J., Hu, M., et al. (2013). Contemporary environmental variation determines microbial diversity patterns in acid mine drainage. *ISME J.* 7, 1038–1050. doi: 10.1038/ismej.2012.139
- Kumar, R., Kaur, M., and Kumari, M. (2012). Acridine: a versatile heterocyclic nucleus. *Acta Pol. Pharm. Drug Res.* 69, 3–9.
- Lalvani, S., Zhang, G., and Lalvani, L. (1996). Coal pyrite passivation due to humic acids and lignin treatment. *Fuel Sci. Technol. Int.* 14, 1291–1313. doi: 10.1080/08843759608947640
- Lehembre, F., Doillon, D., David, E., Perrotto, S., Baude, J., Foulon, J., et al. (2013). Soil metatranscriptomics for mining eukaryotic heavy metal resistance genes. *Environ. Microbiol.* 15, 2829–2840. doi: 10.1111/1462-2920.12143
- Lerman, L. S. (1961). Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.* 3, 18–30. doi: 10.1016/S0022-2836(61)80004-1
- Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2014). MEGAHT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033
- Lladó, S., Benada, O., Cajthaml, T., Baldrian, P., and García-Fraile, P. (2016). *Silvibacterium bohemicum* gen. nov. sp. nov., an acidobacterium isolated from coniferous soil in the Bohemian Forest National Park. *Syst. Appl. Microbiol.* 39, 14–19. doi: 10.1016/j.syapm.2015.12.005
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371. doi: 10.1093/nar/gkh293
- McMurdie, P. J., and Holmes, S. (2013). PhyloSeq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8:e61217. doi: 10.1371/journal.pone.0061217
- Miller, C. S., Baker, B. J., Thomas, B. C., Singer, S. W., and Banfield, J. F. (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol.* 12:R44. doi: 10.1186/gb-2011-12-5-r44
- Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., et al. (2015). The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* 43, D213–D221. doi: 10.1093/nar/gku1243
- Mohan, D., and Chander S. (2006). Single, binary, and multicomponent sorption of iron and manganese on lignite. *J. Colloid Interface Sci.* 299, 76–87. doi: 10.1016/j.jcis.2006.02.010
- Mueller, R. S., Dill, B. D., Pan, C., Belnap, C. P., Thomas, B. C., VerBerkmoes, N. C., et al. (2011). Proteome changes in the initial bacterial colonist during ecological succession in an acid mine drainage biofilm community. *Environ. Microbiol.* 13, 2279–2292. doi: 10.1111/j.1462-2920.2011.02486.x
- Navarrete, A. A., Kuramae, E. E., de Hollander, M., Pijl, A. S., van Veen, J. A., and Tsai, S. M. (2013). Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. *FEMS Microbiol. Ecol.* 83, 607–621. doi: 10.1111/1574-6941.12018
- Navarro, C. A., von Bernath, D., and Jerez, C. A. (2013). Heavy metal resistance strategies of acidophilic bacteria and their acquisition: importance for biomining and bioremediation. *Biol. Res.* 46, 363–371. doi: 10.4067/S0716-97602013000400008
- Nies, D. H. (1995). The cobalt, zinc, and cadmium efflux system CzcABC from *Alcaligenes eutrophus* functions as a cation-proton antiporter in *Escherichia coli*. *J. Bacteriol.* 177, 2707–2712. doi: 10.1128/jb.177.10.2707-2712.1995
- Noyce, G. L., Winsborough, C., Fulthorpe, R., and Basiliko, N. (2016). The microbiomes and metagenomes of forest biochars. *Sci. Rep.* 6:26425. doi: 10.1038/srep26425
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., et al. (2016). *vegan: Community Ecology Package*. Available online at: <https://CRAN.R-project.org/package=vegan>
- Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42, 206–214. doi: 10.1093/nar/gkt1226
- Pankratov, T. A., Serkebaeva, Y. M., Kulichevskaya, I. S., Liesack, W., and Dedysh, S. N. (2008). Substrate-induced growth and isolation of Acidobacteria from acidic *Sphagnum* peat. *ISME J.* 2, 551–560. doi: 10.1038/ismej.2008.7
- Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Pascual, J., Wüst, P. K., Geppert, A., Foessel, B. U., Huber, K. J., and Overmann, J. (2015). Novel isolates double the number of chemotrophic species and allow the first description of higher taxa in Acidobacteria subdivision 4. *Syst. Appl. Microbiol.* 38, 534–544. doi: 10.1016/j.syapm.2015.08.001
- Paulson, J. N., Talukder, H., Pop, M., and Bravo, H. C. (2016). *MetagenomeSeq: Statistical Analysis for Sparse High-throughput Sequencing*. *Bioconductor Package: 1.12.1*. Available online at: <http://cpcb.umd.edu/software/metagenomeSeq>
- Peuravuori, J., Žbáňková, P., and Pihlaja, K. (2006). Aspects of structural features in lignite and lignite humic acids. *Fuel Process. Technol.* 87, 829–839. doi: 10.1016/j.fuproc.2006.05.003
- Pos, K. M. (2009). Drug transport mechanism of the AcrB efflux pump. *Biochim. Biophys. Acta* 1794, 782–793. doi: 10.1016/j.bbapap.2008.12.015
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. doi: 10.1371/journal.pone.0009490
- Pruitt, K. D., Tatusova, T., Brown, G. R., and Maglott, D. R. (2012). NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* 40, D130–D135. doi: 10.1093/nar/gkr1079
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219
- Radeva, G., and Selenska-Pobell, S. (2004). Bacterial diversity in water samples from two uranium mill tailings as revealed by 16S rDNA retrieval. *Comptes Rendus l'Acad. Bulg. Sci.* 57, 85–90.
- Rao, N. N., Gómez-García, M. R., and Kornberg, A. (2009). Inorganic polyphosphate: essential for growth and survival. *Annu. Rev. Biochem.* 78, 605–647. doi: 10.1146/annurev.biochem.77.083007.093039
- Rawat, S. R., Männistö, M. K., Starovoytov, V., Goodwin, L., Nolan, M., Hauser, L., et al. (2014). Complete genome sequence of *Granulicella tundricola* type strain MP5ACTX9(T), an Acidobacteria from tundra soil. *Stand. Genomic. Sci.* 9, 449–461. doi: 10.4056/sigs.4648353
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org>
- Remonsellez, F., Orell, A., and Jerez, C. A. (2006). Copper tolerance of the thermoacidophilic archaeon *Sulfolobus metallicus*: possible role of polyphosphate metabolism. *Microbiology* 152, 59–66. doi: 10.1099/mic.0.28241-0

- Rheims, H., Spröer, C., Rainey, F. A., and Stackebrandt, E. (1996). Molecular biological evidence for the occurrence of uncultured members of the actinomycete line of descent in different environments and geographical locations. *Microbiology* 142, 2863–2870. doi: 10.1099/13500872-142-10-2863
- Rohwerder, T., Gehrke, T., Kinzler, K., and Sand, W. (2003). Bioleaching review part A: progress in bioleaching: fundamentals and mechanisms of bacterial metal sulfide oxidation. *Appl. Microbiol. Biotechnol.* 63, 239–248. doi: 10.1007/s00253-003-1448-7
- Rumpel, C., Knicker, H., Kögel-Knabner, I., Skjemstad, J. O., and Hüttel, R. F. (1998). Types and chemical composition of organic matter in reforested lignite-rich mine soils. *Geoderma* 86, 123–142. doi: 10.1016/S0016-7061(98)00036-6
- Sait, M., Davis, K. E. R., and Janssen, P. H. (2006). Effect of pH on isolation and distribution of members of subdivision 1 of the phylum Acidobacteria occurring in soil. *Appl. Environ. Microbiol.* 72, 1852–1857. doi: 10.1128/AEM.72.3.1852-1857.2006
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Selenska-Pobell, S. (2002). “Diversity and activity of bacteria in uranium waste piles, chapter 8” in *Interactions of Microorganisms with Radionuclides*, eds M. J. Keith-Roach and F. R. Livens (Munich: Elsevier B.V.), 225–254.
- Selenska-Pobell, S., Kampf, G., Flemming, K., Radeva, G., and Satchanska, G. (2001). Bacterial diversity in soil samples from two uranium waste piles as determined by rep-APD, RISA and 16S rDNA retrieval. *Antonie van Leeuwenhoek* 79, 149–161. doi: 10.1023/A:1010237711077
- Senko, J. M., Wanjugi, P., Lucas, M., Bruns, M. A., and Burgos, W. D. (2008). Characterization of Fe(II) oxidizing bacterial activities and communities at two acidic Appalachian coalmine drainage-impacted sites. *ISME J.* 2, 1134–1145. doi: 10.1038/ismej.2008.60
- Stafford, H. A. (1988). Proanthocyanidins and the lignin connection. *Phytochemistry* 27, 1–6. doi: 10.1016/0031-9422(88)80583-1
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stamps, B. W., Losey, N. A., Lawson, P. A., and Stevenson, B. S. (2014). Genome sequence of *Thermoanaerobaculum aquaticum* MP-01T, the first cultivated member of Acidobacteria subdivision 23, isolated from a hot spring. *Genome Announc.* 2:e00570-14. doi: 10.1128/genomeA.00570-14
- Stout, S. A., Boon, J. J., and Spackman, W. (1988). Molecular aspects of the peatification and early coalification of angiosperm and gymnosperm woods. *Geochim. Cosmochim. Acta* 52, 405–414. doi: 10.1016/0016-7037(88)90096-8
- Stubner, S. (2004). Quantification of Gram-negative sulphate-reducing bacteria in rice field soil by 16S rRNA gene-targeted real-time PCR. *J. Microbiol. Methods* 57, 219–230. doi: 10.1016/j.mimet.2004.01.008
- Tan, G. L., Shu, W. S., Hallberg, K. B., Li, F., Lan, C. Y., Zhou, W. H., et al. (2008). Culturable and molecular phylogenetic diversity of microorganisms in an open-dumped, extremely acidic Pb/Zn mine tailings. *Extremophiles* 12, 657–664. doi: 10.1007/s00792-008-0171-9
- Turpault, M. P., Bonnaud, P., Fichter, J., Ranger, J., and Dambrine, É. (1996). Distribution of cation exchange capacity between organic matter and mineral fractions in acid forest soils (Vosges mountains, France). *Eur. J. Soil Sci.* 47, 545–556. doi: 10.1111/j.1365-2389.1996.tb01854.x
- Tyson, G., Chapman, J., and Hugenholtz, P. (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428, 37–43. doi: 10.1038/nature02340
- Urbanová, M., Kopecký, J., Valášková, V., Ságová-Marečková, M., Elhottová, D., Kyselková, M., et al. (2011). Development of bacterial community during spontaneous succession on spoil heaps after brown coal mining. *FEMS Microbiol. Ecol.* 78, 59–69. doi: 10.1111/j.1574-6941.2011.01164.x
- Wainwright, M. (2001). Acridine—a neglected antibacterial chromophore. *J. Antimicrob. Chemother.* 47, 1–13. doi: 10.1093/jac/47.1.1
- Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18, 691–699. doi: 10.1093/oxfordjournals.molbev.a003851
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.
- Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46
- Wu, M., and Scott, A. J. (2012). Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* 28, 1033–1034. doi: 10.1093/bioinformatics/bts079
- Xie, J., He, Z., Liu, X., Liu, X., Van Nostrand, J. D., Deng, Y., et al. (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. doi: 10.1128/AEM.01798-10
- Yelton, A. P., Comolli, L. R., Justice, N. B., Castelle, C., Deneff, V. J., Thomas, B. C., et al. (2013). Comparative genomics in acid mine drainage biofilm communities reveals metabolic and structural differentiation of co-occurring archaea. *BMC Genomics* 14:485. doi: 10.1186/1471-2164-14-485
- Yelton, A. P., Thomas, B. C., Simmons, S. L., Wilmes, P., Zemla, A., Thelen, M. P., et al. (2011). A semi-quantitative, synteny-based method to improve functional predictions for hypothetical and poorly annotated bacterial and archaeal genes. *PLoS Comput. Biol.* 7:e1002230. doi: 10.1371/journal.pcbi.1002230

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

Copyright © 2017 Wegner and Liesack. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.