# BIOINFORMATICS ANALYSIS OF MICROBIAL ABUNDANCE AND DIVERSITY IN ACID MINE DRAINAGE FROM THE SOLOMON MINE NEAR CREEDE, COLORADO

RYAN K. MILLER[1], ROBERT M. KIRKHAM[2], AND ADAM J. KLEIN-SCHMIT[1]*

[1]DEPARTMENT OF BIOLOGY AND EARTH SCIENCES, ADAMS STATE UNIVERSITY, ALAMOSA, COLORADO
[2]GEOLOGICAL SOLUTIONS, ALAMOSA, COLORADO

## CORRESPONDING AUTHOR

*Adam J. Kleinschmit , Adams State University, Alamosa, Co Usa
akleinschmit@adams.edu

## KEYWORDS

- Acid Mine Drainage
- Microbial Ecology
- 16S rRNA Marker Gene Survey
- High Throughput Sequencing
- Environmental Microbiota

# ABSTRACT

This study focused on characterizing the phylotypic composition of acid mine drainage (AMD) communities associated with the Solomon Mine near Creede, Colorado, and its relative diversity compared to microbial communities found in the East Willow Creek (EWC) watershed. AMD from the Solomon Mine adit flows into an existing passive bioremediation wetland system located next to the Solomon Mine adit that currently is ineffective and is under consideration for renovation. We are interested in gaining an understanding of the baseline microbial communities present in AMD/EWC and to monitor them during future wetland renovation. Prokaryotic community profiling was approached using SSU 16S rRNA marker gene amplification coupled with next generation sequencing. Bioinformatics analysis included raw read preprocessing, data visualization, and statistical testing using a combination of USEARCH and QIIME-based scripts. A pH and conductivity gradient were observed for water moving through the currently inefficient wetland system at the Solomon Mine. The EWC microbiomes had statistically significant higher alpha diversity compared to the AMD microbiomes. Beta diversity analysis parsed the sample locations into statistically significant groups including core AMD microbiomes, the wetland Cell 3 microbiome, and EWC microbiomes using multidimensional scaling. Taxa driving beta diversity included the phylum Proteobacteria for the core AMD microbiomes, the phyla Firmicutes and Chloroflexi for the constructed wetland Cell 3, and the phyla Bacteroidetes and Verrucomicrobia for EWC. Our data suggests that the microbial community in constructed wetland Cell 3 is likely where limited sulfate reduction activity is operating at low capacity, which will be further investigated via shotgun metagenomic analysis.

# INTRODUCTION

Microbes populate every known niche on earth including physiologically challenging environments for biological organisms, such as deep–sea hydrothermal vents and subglacial Antarctic lakes, where extremophiles thrive (38, 41). Environmentally, microbes are essential in the ecological biogeochemical cycling of elements (e.g., carbon, nitrogen, sulfur, phosphorus) to allow other biological organisms access to tangible forms of these elements, thus making life sustainable (18). The shear abundance of microbes makes them a major force that keeps elements circulating between biotic and abiotic entities in an ecosystem. For example, chemolithotrophs can "eat rocks" to obtain energy and electrons, while utilizing carbon dioxide as a carbon source. Thus, chemolithotrophs do not need light or organic matter to proliferate and can live in extreme habitats such as deep–sea hydrothermal vents and underground caves (33).

Oxidation of inorganic matter by chemolithotrophs associated with volcanic caves or even exposed rock that is highly mineralized with metal sulfides can be the limiting factor for the generation of acid rock drainage (ARD) (39). ARD is a result of an abiotic reaction typically involving iron pyrite that can be accelerated by microbes and produces sulfuric acid as byproduct, thus dramatically lowering the pH within a watershed downstream of the ARD source. Under these low pH conditions, other heavy metals are easily solubilized from metallic minerals and can lead to environmental degradation within the watershed (13).

ARD generation can be expedited and/or be directly generated as a result of anthropogenic activity such as civil engineering projects and mining. When ARD is associated with mining activity it is more commonly referred to as acid mine drainage (AMD). Abandoned mine adits and associated tailing piles can bring sulfide–rich rock into contact with abiotic factors (e.g., oxygen and moisture) along with biotic factors (e.g., chemolithotrophs), which can in turn produce AMD for thousands of years. This undesired impact from mining can lead to both economic and conservation–based environmental issues for biological organisms living downstream of an acidified heavy–metal leaching water source.

Studying microbial ecology associated with AMD sources has historically been fruitful in gaining insight into microbial interactions, as the harsh conditions yield an ecosystem with low complexity (3, 22). Knowledge gained from understanding microbial interactions in AMD has the potential to be applied to promising fields associated with the mining industry, including biomining and bioremediation of AMD (6). Biomining has the potential to extract trace metals via bioleaching in an efficient and environmentally conscious way and perhaps even remine minerals that historically were not at a sufficient concentration to warrant previous efforts. Bioremediation, on the other hand, can be harnessed as a desired passive approach to

naturally address AMD through microbial reduction of metal oxides, thus increasing pH and precipitating heavy metals. This passive approach is generally more economically sustainable compared to intensive active approaches (e.g., lime treatment, filtering) (24).

For this project, we are specifically interested in characterizing the microbial ecology of microbes associated with AMD produced from the abandoned Solomon Mine along East Willow Creek (EWC) near the town of Creede, Colorado. The Solomon Mine was established in 1890 along EWC in the South San Juan Mountains (Figure 1). The mine has been worked sporadically with the primary minerals mined at the site including galena (lead sulfide), sphalerite (zinc sulfide), pyrite (iron sulfide), and minor chalcopyrite (copper and iron sulfide) up until its formal decommission in 1971 (15, 23). The mine also produced minor gold and even less silver. In 1991, a passive mine drainage treatment system was constructed by the Colorado Mined Land Reclamation Division (now called Colorado Division of Reclamation, Mining, and Safety) to reduce the load of heavy metals moving into EWC. The system consisted of three constructed wetland cells. AMD from the adit flowed by gravity through the cells. As originally constructed, Cell 1 was filled with alfalfa hay, Cell 2 contained wood chips, and a mixture of spent mushroom compost and wood chips were placed in Cell 3.

The system worked well for a few years, but a lack of funds for upkeep and maintenance dramatically reduced the efficiency of the constructed wetland system and only very

minor levels of dissolved metals currently are being removed from AMD that passes through the system (47).

Willow Creek Reclamation Committee (WCRC) is contemplating the renovation or rebuilding of the constructed wetland system. Local geochemistry greatly influences AMD microbial ecosystems; thus, it is important to characterize this site to understand the dynamics within the Solomon Mine AMD microbiome for future intervention efforts (5, 27). We undertook an initial survey of the pre-intervention microbial phylotype composition and diversity, with the aim of identifying focused investigatory questions and monitoring future microbial phylotype evolution longitudinally after the installation of a functioning passive bioremediation pond system.

Experimentally, we approached the characterization of the AMD phylotype composition and diversity using a culture-independent method to gain an inclusive profile of the present microbial communities (16). Preliminary community profiling was performed by targeting the SSU 16S rRNA marker gene with next generation sequencing of the resulting amplicons (14) coupled with bioinformatic analysis using USEARCH (12) and Quantitative Insights into Microbial Ecology (QIIME)–based scripts on the resulting sequence reads (7, 32). Here we describe the baseline microbial community of Solomon Mine AMD and EWC drainage by using the SSU 16S rRNA marker gene to taxonomically profile associated sediment samples.
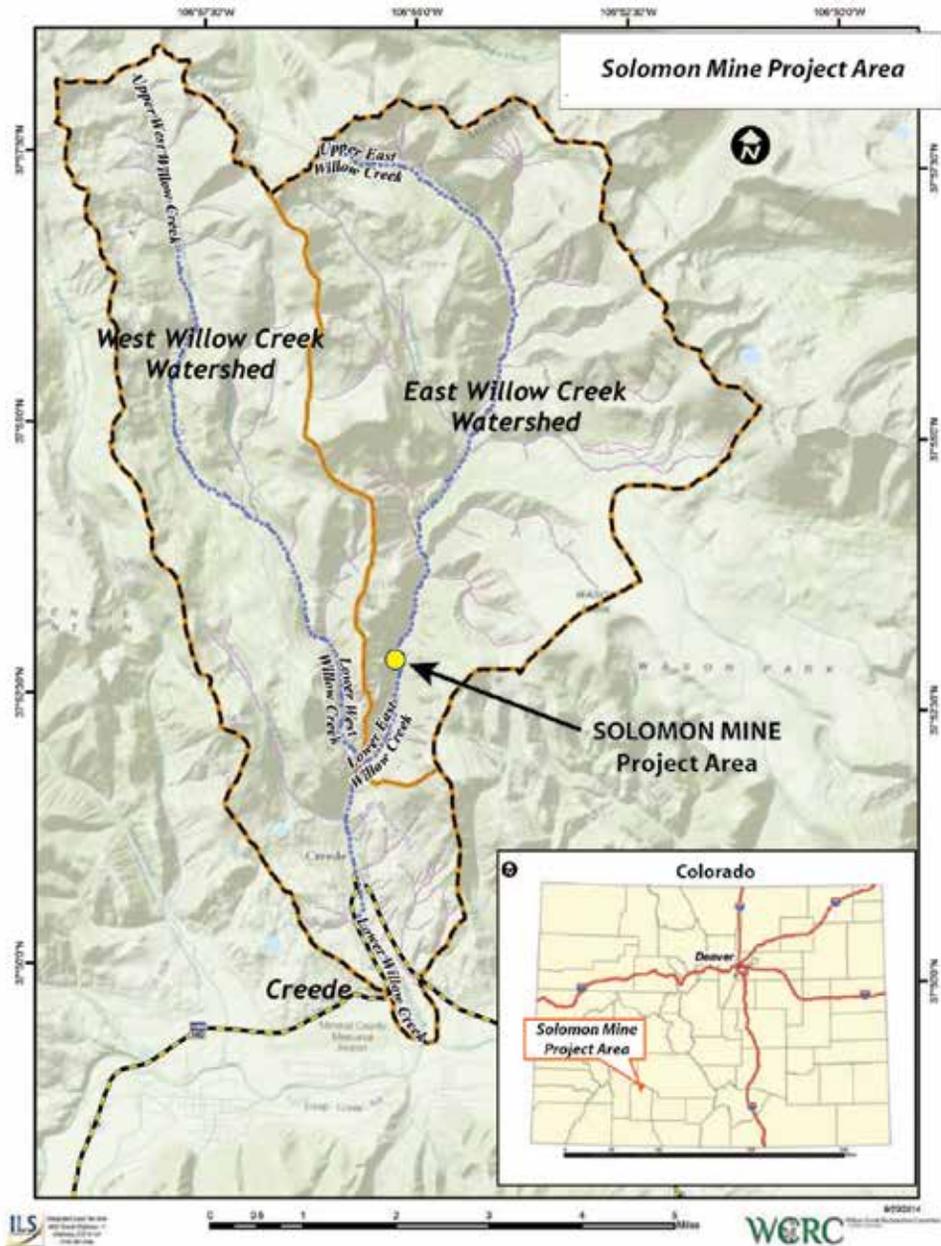
Figure 1. Hillshade map depicting the East Willow Creek Watershed in southwest Colorado and the Solomon Mine Project Area, which includes both the Solomon Mine adit and associated constructed wetlands along East Willow Creek.

# MATERIALS AND METHODS

## SAMPLE COLLECTION

Water and Sediment samples were collected on July 9, 2015 in the East Willow Creek drainage basin. The following sample locations mirrored locations used by Willow Creek Reclamation Committee for water-quality monitoring (marked and labeled by wood stakes) along East Willow Creek at Site A (EW-A), Site G (EW-G), from AMD at the Solomon Mine adit portal just inside and adjacent to the bulkhead (SMA), and the Solomon Mine Wetland Discharge (SWD). Additional samples were obtained within the Solomon Mine Adit ~20 meters from the bulkhead (SMT) and from a central position within each of the currently defunct constructed wetland system cells (Cell 1, Cell 2, Cell 3), and in East Willow Creek immediately upstream of the constructed wetlands (EW-USM). The samples associated with the Solomon Mine were obtained with permission from Willow Creek Reclamation Committee.  For the collection of water samples, sterile Nalgene polypropylene bottles were rinsed three times with sample water prior to sample collection in undisturbed water. Sediment samples were collected aseptically from the upper oxic zone at three points at 1 m distances triangulated and pooled as a single sample in sterile polypropylene conical tubes. Water and sediment samples were immediately placed on ice in a cooler for transport to the laboratory. Sediment sample aliquots were stored at −80°C in the laboratory until processed for DNA extraction.

## SAMPLE METADATA COLLECTION

Air and water temperature, pH, and electrical conductivity were measured in the field using a handheld combination multimeter. A handheld GPS unit determined sample locations in NAD83. Laboratory pH and oxidation/reduction potential were measured ~3 hours after sample collection using a VWR Symphony multimeter device. Historical East Willow Creek water-quality monitoring data was obtained from Willow Creek Reclamation Committee (47).

## 16S rRNA GENE AMPLICON SEQUENCING

Total genomic DNA was extracted from 0.25 g of sediment per sample using a MoBio PowerSoil Kit following the manufacturer's instructions (Qiagen, Carlsbad, CA, USA). Illumina 515F and 806R iTag Primers were designed based on Caporaso et al. (8) and used to amplify the V4 region of the 16S rRNA gene. PCR reactions were composed of a final concentration of 5-10 ng of template DNA, 0.8 mM dNTP's, 0.2 µM forward and reverse primers, 1x PCR buffer, and 1x Promega GoTaq with use of a Bio-Rad MyCycler. Duplicate PCR reactions were pooled to address PCR drift (42), were gel-purified using a Qiagen kit (Qiagen, Frederick, MD,

USA), and quantified using a high sensitivity dsDNA kit with a Qubit 3.0 fluorometer (Life Technologies, Carlsbad, CA, USA). Processed sediment samples were sent overnight on dry ice for sequencing preparation and service by Wright Laboratories (affiliated with Juniata College, Huntingdon, PA). Wright Laboratories quality checked the amplicon libraries using a 2100 Bioanalyzer DNA 1000 chip and performed 150 bp paired-end sequencing reads using the Illumina MiSeq platform.

## BIOINFORMATICS ANALYSIS AND DATA VISUALIZATION

The sequencing service carried out by Wright Laboratories provided the authors with the raw sequence reads *via* a set of demultiplexed FASTQ files for the forward and reverse reads. The authors remotely utilized Juniata College's HHMI Computing Cluster Server to run a custom Unix-based USEARCH/Quantitative Insights into Microbial Ecology (QIIME) pipeline to process the raw reads. Briefly, paired end reads were merged using a minimal overlap of 46 bp and a minimum quality score of 36. Quality filtering of the paired reads was performed by trimming reads to 252 bp and a maximum average expected error rate of 1% using USEARCH v7 (12). USEARCH v7 was also used for sequence dereplication, de novo OTU picking at 97% identity, and removal of chimeras. A combination of QIIME 1.9.0 and USEARCH v7 scripts were used to generate an OTU table, assign taxonomy using the Greengenes Reference Database (13_8 release) clustered at 97% identity (11), and to generate

a phylogenetic tree (7). The pipeline output files were further analyzed using QIIME 1.9.0 via Oracle VM Virtualbox software on a PC for downstream OTU table filtering, diversity analysis, and multivariate statistical tests. Circos plots were generated using a CSS normalized filtered OTU table with relative abundance values transformed into integers by multiplying values by 10,000 and rounding up to the nearest whole value. Relative abundance values for sample location groups were averaged and formatted for Circos Online (26). Alpha Diversity box plots were generated using BoxPlotR, a web-based tool running R packages on a shiny server (40).

## STATISTICAL TESTS

All statistical tests were performed in QIIME. Significance of alpha diversity between AMD and EWC samples was obtained *via* a non-parametric t-test using Monte Carlo permutation to calculate the p-value. A beta diversity non-parametric multivariate statistical test was calculated with Adonis (modified PERMANOVA). Beta diversity distance plot p-values were generated using a paired Bonferroni-corrected t-test. Resulting p-values were adjusted *via* false discovery rate (FDR) or Bonferroni method for multiple comparisons as indicated in the results. Differential abundance analysis was performed using the metagenomeSeq and DESeq2 package *via* QIIME (34).

# RESULTS

## ENVIRONMENTAL PARAMETERS OF SOLOMON MINE AMD AND EAST WILLOW CREEK

Water-quality measurements and sediment samples were taken from the Solomon Mine site and neighboring EWC at selected WCRC water-quality sampling sites, and from the mine adit and constructed wetland cells (Figure 2). AMD water and sediment samples collected from the mine adit, wetland Cell 1, and wetland Cell 2 were stained with iron oxides with no obvious biological organisms present, while wetland Cell 3 exhibited notable algal and aquatic plant growth (Figure 3). Generally, an increasing pH gradient was detected from the AMD source point, through the constructed wetland pond cells, and into EWC (Table 1). Additionally, an increasing gradient of conductivity and oxidation/reduction potential correlated with the pH gradient as AMD moved *via* gravity flow through the constructed wetlands and into EWC (Table 1). EWC samples were within about a half pH unit of 7 and were 2-3 orders of magnitude higher than the AMD measurements. Additionally, conductivity and oxidation/reduction potential measurements in the EWC samples were markedly lower than the AMD samples (Table 1) in the EWC samples were markedly lower than the AMD samples (Table 1).

Table 1. Solomon Mine AMD/East Willow Creek sampling site water-quality and location data.

| Sample ID* | Category | GPS Coordinates (Degrees – NAD83) | Elevation (m) | Water Depth (cm) | Air Temp (°C) | Water Temp (°C) | Field pH | Lab pH (@ 25°C) | Field Conductivity (µS/m) | Lab ORP (RmV @ 25°C) |
|---|---|---|---|---|---|---|---|---|---|---|
| EW-A | Willow Creek | 37.86444 -106.92524 | 2,736 | 11 | 14.6 | 7.8 | 5.9 | 7.24 | 618 | -29.1 |
| EW-G | Willow Creek | 37.87832 -106.91769 | 2,832 | 8 | 19.6 | 8.4 | 6.6 | 7.52 | 57.7 | -31.9 |
| EW-USM | Willow Creek | 37.88198 -106.91772 | 2,844 | 10 | 14.8 | 8.3 | 6.4 | 6.89 | 52.7 | -56.6 |
| SWD | AMD | 37.88162 -106.91777 | 2,838 | 6 | 16.5 | 13.9 | 4.7 | 4.96 | 691 | 111.6 |
| CELL 3 | Bioremediation | 37.88154 -106.91778 | 2,838 | 72 | 17.0 | 15.3 | 4.4 | 4.84 | 694 | 121.7 |
| CELL 2 | AMD | 37.88123 -106.91773 | 2,839 | 86 | 17.4 | 15.5 | 4.3 | 4.18 | 694 | 161.5 |
| CELL 1 | AMD | 37.88092 -106.91764 | 2,840 | 80 | 13.7 | 10.2 | 4.2 | 4.18 | 719 | 159.5 |
| SMA | AMD | 37.88093 -106.91782 | 2,884 | 7 | 8.6 | 8.5 | 4.2 | 4.06 | 722 | 166.2 |
| SMT | AMD | Underground | 2,884 | 52 | 8.8 | 8.5 | 4.1 | 4.09 | 742 | 163.9 |

* Sample ID acronyms key: East Willow Creek Site A (EW-A) or Site G (EW-G) = WCRC sample site A, G, or immediately

* Sample ID acronyms key: East Willow Creek Site A (EW-A) or Site G (EW-G) = WCRC sample site A, G, or immediately upstream of Solomon Mine (USM); SWD = Solomon Mine Constructed Wetland Discharge; SMA = Solomon Mine Adit just inside the bulkhead; SMT = inside Solomon Mine Adit, ~20 meters from bulkhead.
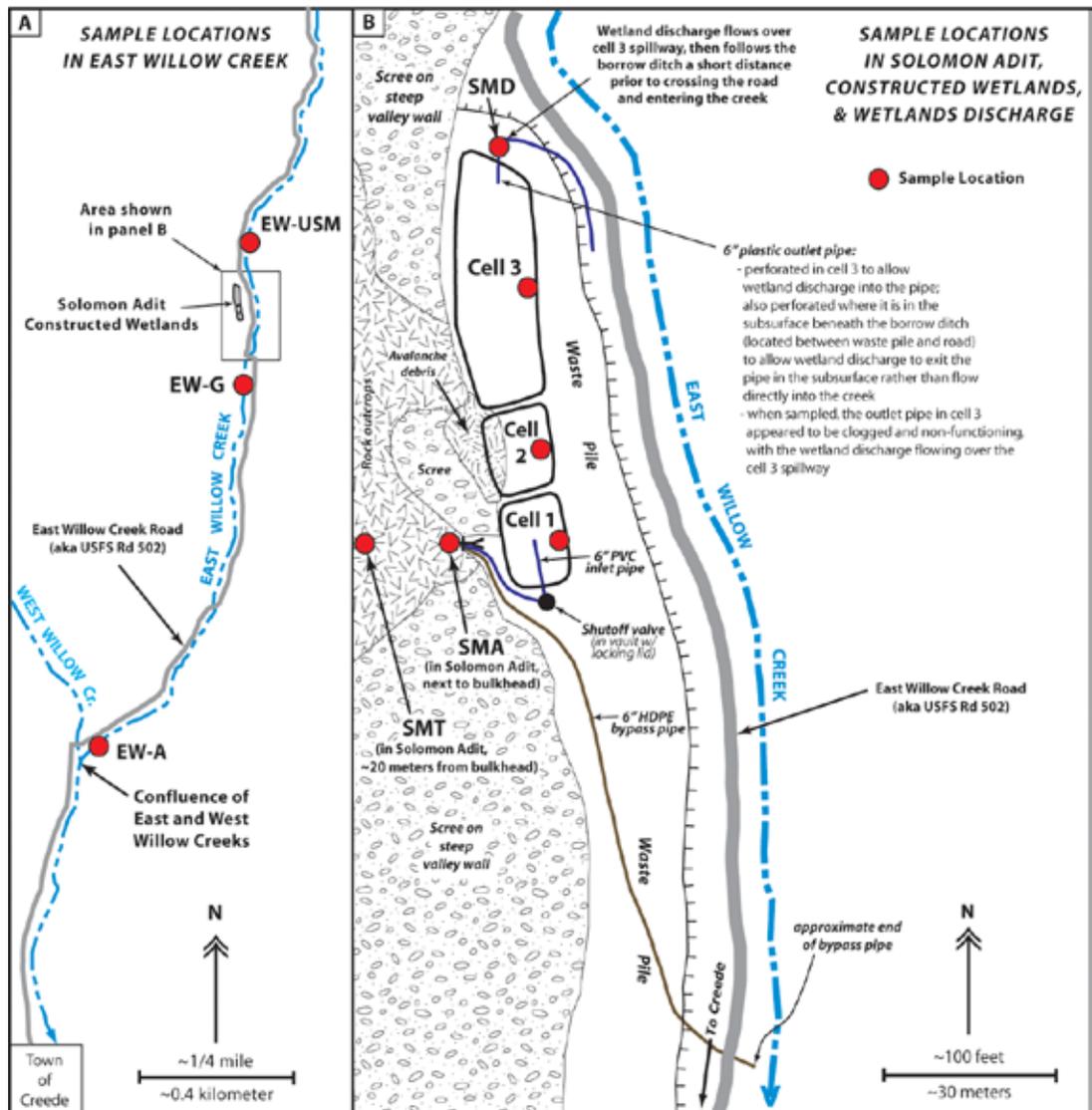
Figure 2. East Willow Creek and Solomon Mine acid mine drainage (AMD) sample locations (red circles). (A) East Willow Creek sample locations immediately upstream (EW–USM) and downstream (EW–G) of the Solomon Mine and immediately upstream of the confluence of East and West Willow Creek (EW–A). (B) Solomon Mine and constructed wetlands (outlined region in panel "A"). SMT (Solomon Mine Adit, ~20 meters from the bulkhead); SMA (Solomon Mine Adit, just inside and adjacent to the bulkhead); SMD (Solomon Mine Wetland Discharge).
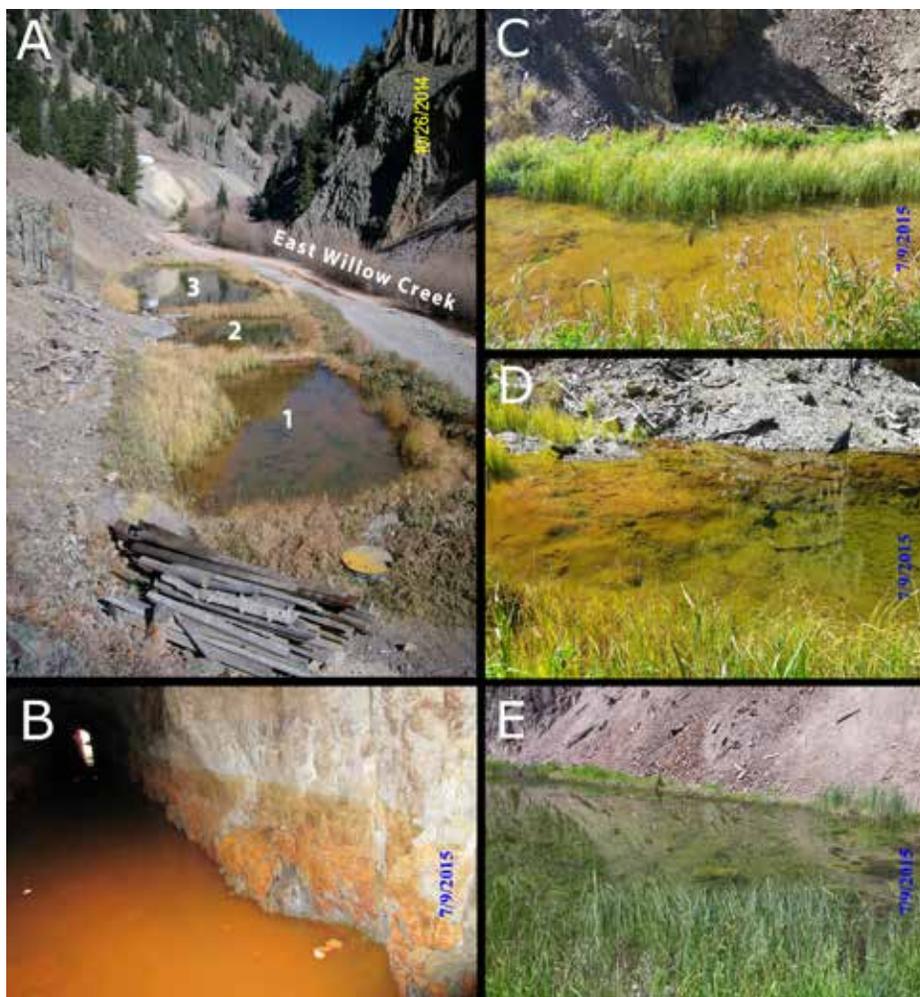
Figure 3. Acid Mine Drainage (AMD) sample location images. (A) Overview photograph of the constructed wetland system looking north. Portal of the Solomon Adit is in the hillside immediately beyond the lower left side of the photograph. Water draining from the Solomon Adit flows into the cells via a buried pipe that extends from the portal to the shutoff valve located below ground in the vertical culvert with metal lid (seen in the photograph above the right end of the pile of lumber and timbers). The buried pipe then runs from the shutoff valve into the bottom of Cell 1, where the mine drainage enters the wetland system. (B) Inside the Solomon Mine adit, ~20 meters from the bulkhead. (C–E) Images of the constructed wetland Cell 1 (C), Cell 2 (D), and Cell 3 (E) prior to sampling. Note the iron–oxide "yellow boy" in Cell 1 and –2 (C and D respectively), while Cell 3 (E) exhibited noticeable aquatic plant and algal growth.

Table 2. Sequence read counts associated with each processing step with the Juniata College USEARCH/QIIME Pipeline.

| Data Processing Step | # of sequences for analysis |
| --- | --- |
| Total Sequences Produced in MiSeq Run | 9,679,736 |
| Post Sequence Demultiplexing (# of project specific sequences) [a] | 302,235 |
| Converted Quality Filtered Paired End Reads | 291,330 |
| Post Quality Filtering (252 bp length & Average Error 1%) | 282,183 |
| Post Sequence Dereplication | 18,025 |
| Post Chimera Removal | 17,261 |
| Post Clustering Sequences into OTUs (97%) & Singleton Removal [b] | 3,832 |

[a] Demultiplexed sequences are split across nine sample locations

[b] Rarefaction for the Chao1 alpha diversity metric utilized a OTU table without singletons removed

## TAXONOMIC COMPOSITION

Raw Illumina MiSeq reads were preprocessed to merge paired-end reads, quality filter to trim or remove low quality reads, add QIIME sample identifiers, dereplicate, cluster into operational taxonomic units (OTUs), remove chimeras and singletons, and assign OTU taxonomy

(Table 2). Across the nine samples, 3,832 OTUs were observed in the final OTU table. Depending on the specific downstream analysis, the table was either rarefied or underwent CSS normalization to account for the varying read count depth across samples, which ranged from 21,994–37,949 reads/ sample location.

To obtain a visual representation of the relative abundance differences between the nine samples for the 3,832 OTUs, a log-transformed heatmap was generated with the OTUs sorted by phylogeny along the y-axis (Figure 4). The visual OTU relative abundance for the AMD sample subgroup exhibited a relative low abundance of OTUs indicated by color intensity across the set of 3,832 OTUs compared to the EWC sample subset. This result is indicative of low AMD alpha diversity and more generally of a relatively different phylotype composition between the two environments.

At the phylum level, AMD and EWC subgroups exhibited differential patterns of relative abundance of major phyla (phyla with >2% relative abundance after CSS normalization) (Figure 5).

AMD samples generally exhibited a higher relative abundance of Proteobacteria,

Actinobacteria, Chloroflexi, and Firmicutes phyla. The EWC subgroup generally exhibited a higher relative abundance of Bacteroidetes and Verrucomicrobia phyla. These differences in the relative abundance of phyla were also observed to drive differences in beta diversity between AMD and EWC sample location groups when visualized via 3-D PCoA biplots (data not shown).

Differential abundance analysis was performed on the AMD and EWC subgroups to identify over/under-represented taxa between the sample groups. A single statistically significant OTU (Bonferroni corrected; p=0.034), Acidobacteria-6 iii1-15, which was identified down to the order level, was absent in AMD samples, but present at low levels in the EWC samples. Although not significant, the Bdellovibrio and Gallionella genera were notably over-represented in AMD samples with Gallionella more abundant in the Solomon Mine adit sample locations than within the other AMD sampling locations.
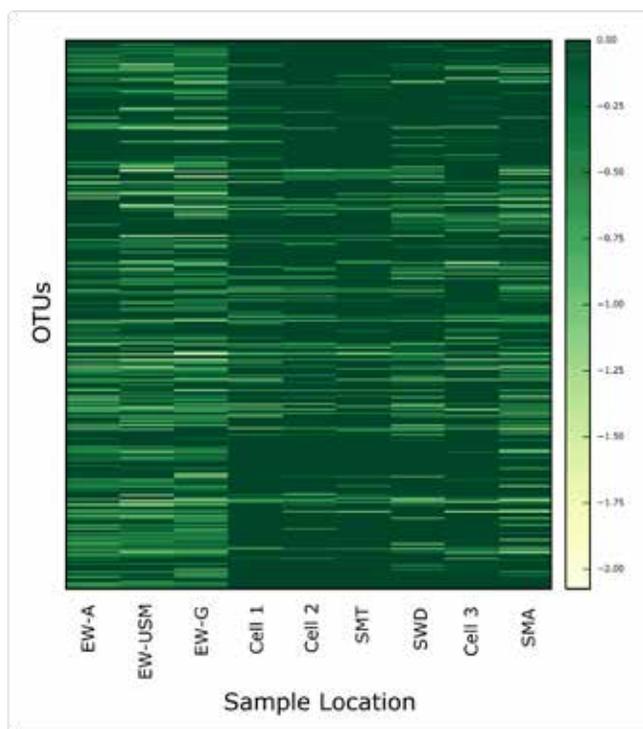


Figure 4. Relative abundance heat map of the 3,832 USEARCH clustered taxa. A CSS normalized relative abundance OTU table was used to plot log-transformed relative abundance intensity for each of the 3,832 clustered OTUs for each sample location. Note the relative abundance of OTUs in the EWC sample location microbiomes (EW-A, EW-USM, & EW-G) compared to the AMD sample location microbiomes (Cells 1 to 3, SMT, SWD, & SMA).
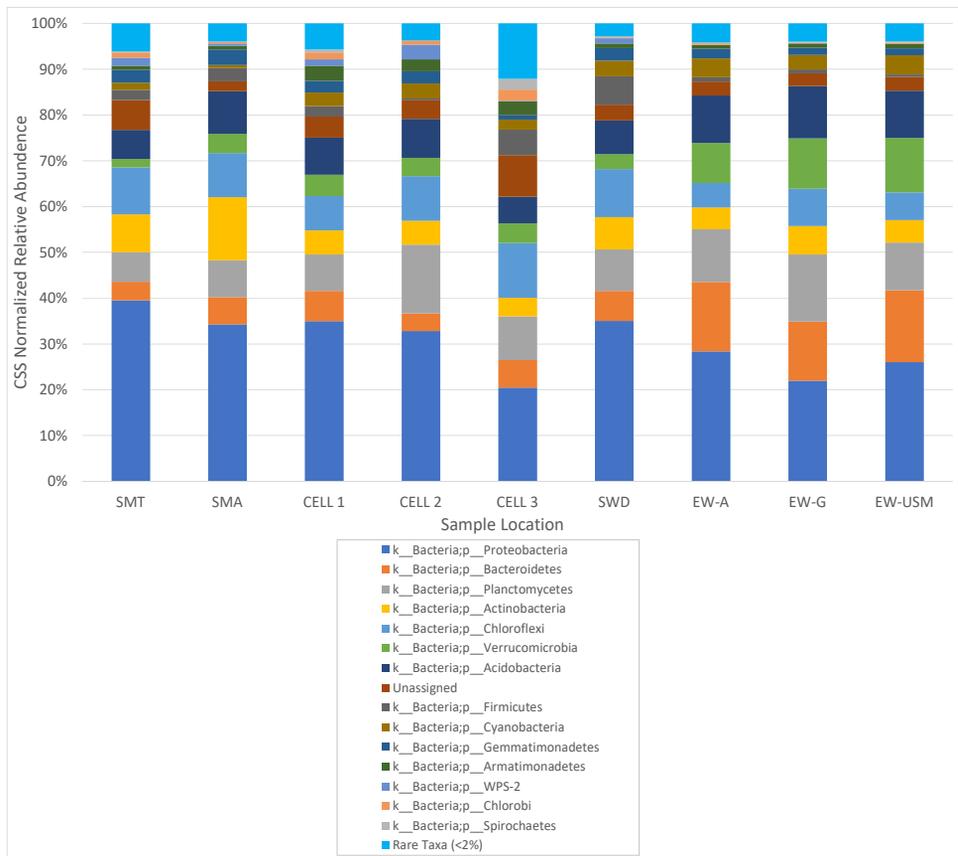
Figure 5. Relative abundance of bacterial phyla by sample location. A CSS normalized and singleton filtered OTU table was utilized to show phyla that make up >2% relative abundance in at least one of the sample locations. AMD sample group: SMT, SMA, Cell 1, Cell 2, Cell 3, SWD. EWC sample group: EW–A, EW–G, EW–USM.

## DIVERSITY ANALYSIS

The amount of diversity within each sample site was determined using a collection of alpha diversity metrics and indices. Multiple rarefactions were performed to normalize the data set for alpha diversity analysis and to validate that sequencing was done at an adequate depth for fair representation of the microbial community. The raw read depth ranged from 21,994–37,949 reads/sample, thus the OTU table was rarefied without replacement to a depth of 21,600. The resulting rarefactions were collated and a rarefaction plot was generated for the sample locations. The resulting rarefaction curves were indicative of quality representative samples as the curves were close to reaching the asymptote (Figure 6). Observed species (p=0.013), phylogenetic diversity index (p=0.021), Chao1 richness estimator (p=0.006), and the Shannon index (p=0.039) exhibited

statistically significant differences in alpha diversity between AMD and EWC sample location microbiomes (Figure 7). Heips evenness index ($p$=0.123) did not indicate a statistically significant difference between AMD and EWC suggesting that although there is a marked difference in species richness between AMD and EWC sample locations, that a portion of those additional taxa in EWC may not be as relatively abundant. The input OTU table used for Chao1 alpha diversity analysis was not filtered for singletons due to the Chao1 metric requirement of needing to take rare taxa into
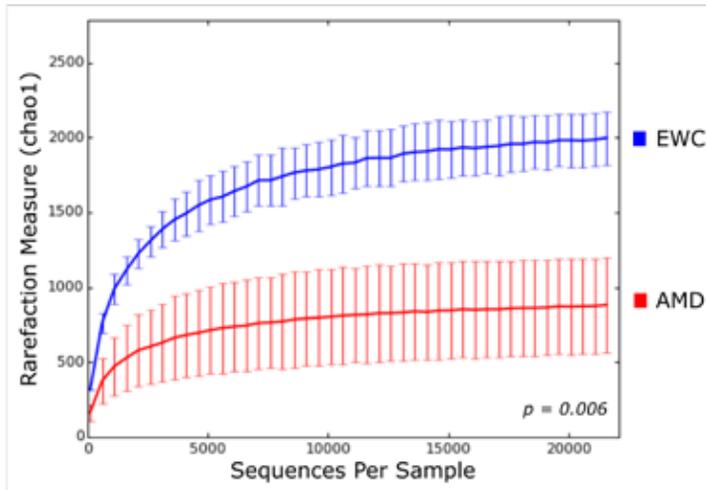


Figure 6. Chao1 rarefaction curve for AMD and EWC samples. Solomon Mine Acid Mine Drainage (AMD –red) and East Willow Creek (EWC – blue). An unfiltered OTU table was subsampled without replacement for 20 iterations using 500 sequence steps with a maximum of 21,600 sequences. Chao1 alpha diversity was calculated on the directory of rarified tables, and the multiple rarefactions were collated to produce the plot. AMD sample group: SMT, SMA, Cell 1, Cell 2, Cell 3, SWD. EWC sample group: EW–A, EW–G, EW–USM.
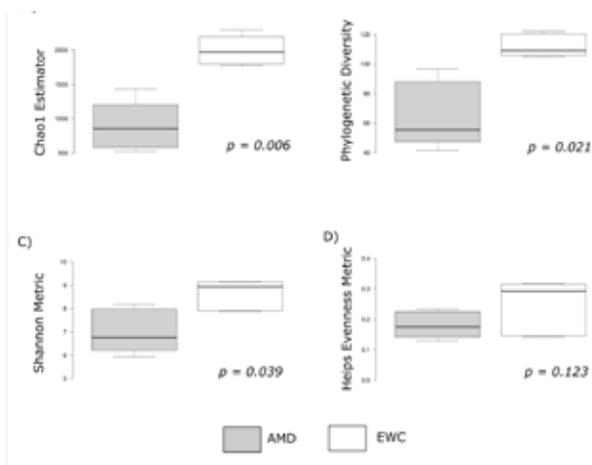


Figure 7. Alpha diversity metric box plots. Error bars represent +/– 1.58*IQR/sqrt(n). An unfiltered OTU table was subsampled without replacement for 20 iterations using 500 sequence steps with a maximum of 21,600 sequences. (A) Chao1 estimator using an OTU table with the presence of singletons. (B) Phylogenetic Diversity, (C) Shannon Index, (D) Heips Evenness Index.

account for modeling.

The amount of diversity between sample sites was determined through beta diversity analysis. The original OTU table was filtered to remove singletons and rare taxa (<0.005% relative abundance) that might inflate the data set (4). To account for sequencing depth, the filtered OTU table was normalized via cumulative sum scaling (34) and used as an input for the calculation of a weighted Unifrac distance metric (29). This data reduction step was then visualized via a principle coordinate analysis (PCoA) plot along with the principle coordinates summarized via a Scree plot. The Scree plot indicated that the samples generated 100% of the variance within 7 dimensions, with >86% of the variance observed in the first 3 dimensions (Supplemental Figure S1). The AMD and the EWC sample microbiomes clustered tightly in the first two dimensions of the PCoA plot, with the exception being the constructed wetland Cell 3 which diverged from the other AMD samples along the second principal coordinate (Figure 8). Beta diversity between sample location microbiomes exhibited statistical significance via the Adonis permutational multivariate analysis of variance statistical test using the three clustered sample group microbiomes as an input (p<0.001). Furthermore, distance comparison boxplots representing all the dimensions of variance from the distance matrix indicated a statistically significant difference when comparing diversity within and between the three categories (p<0.001) (data not shown).
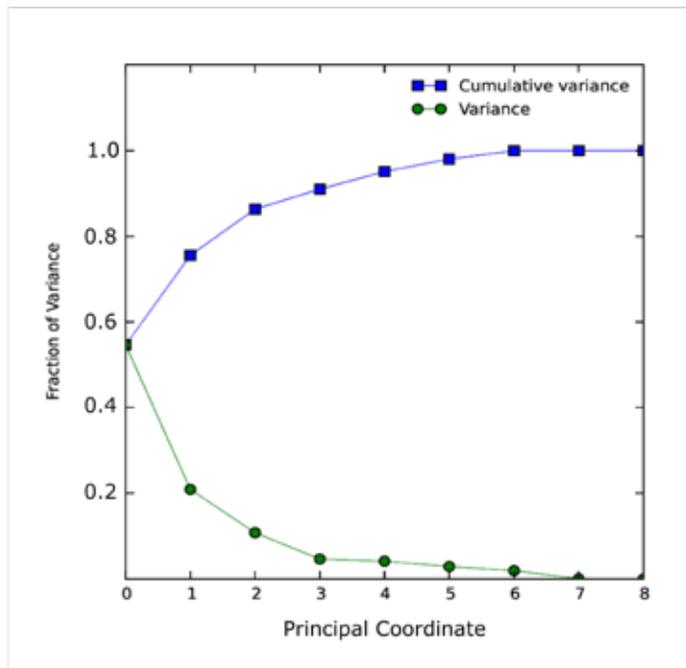


Supplemental Figure S1. Scree Plot of Principal Coordinate Analysis. 100% of the variance can be explained within seven coordinates.
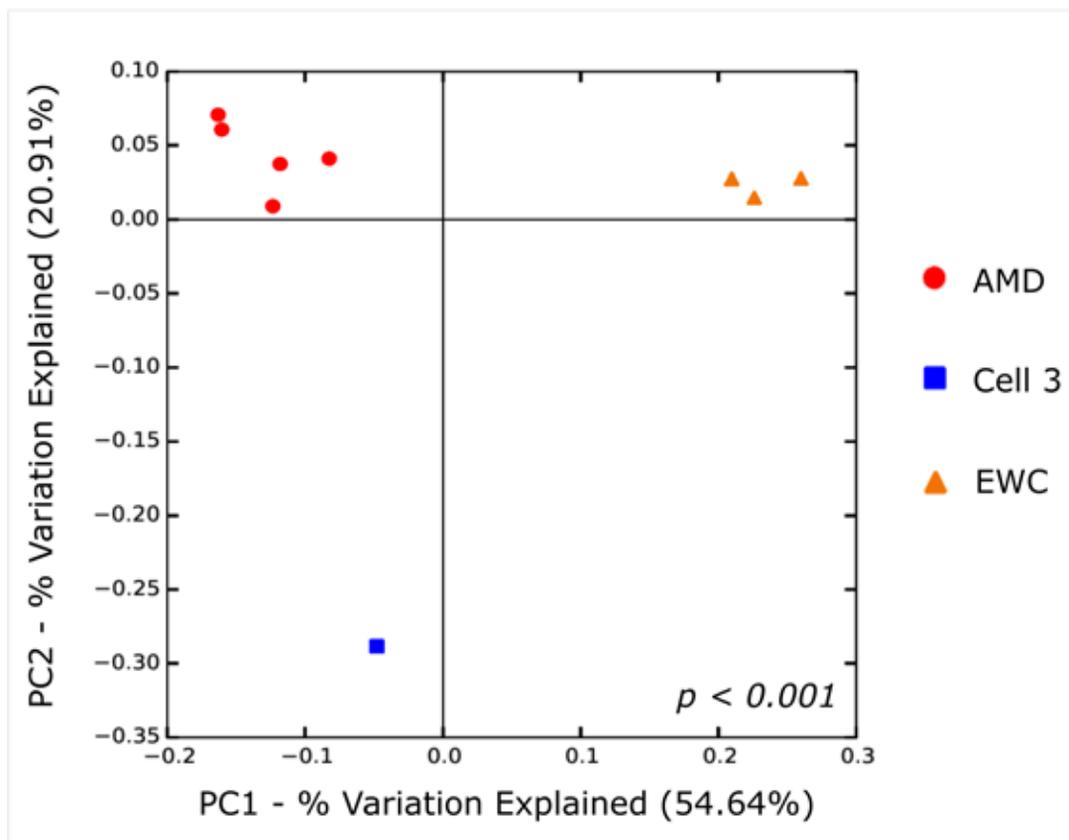
Figure 8. First two Principal Coordinates of a Principle Coordinate Analysis (PCoA) Plot. Solomon Mine Acid Mine Drainage (AMD – red circles), Constructed Wetland Cell 3 Acid Mine Drainage (Cell 3 – blue box), East Willow Creek (EWC – orange triangles).

Analysis of the major taxa and their relative abundance driving the beta diversity clustering of the samples was performed via 3–D PCoA biplots. This data supports the general observed differences in relative abundance at the phylum level. Further exploration of PCoA biplots down to the family level, reveals major taxa in terms of abundance that drive differences in microbial community composition between AMD, EWC, and constructed wetland Cell 3. AMD sample locations were associated with the relative abundance of the family–level taxa *Acetobacteraceae* and *Rhodospirillaceae*, which are both hierarchically placed within the Proteobacteria phylum. Constructed wetland Cell 3 beta diversity was driven by the phylum Chloroflexi and the family Isophaeraceae and a group of unassigned taxa. Lastly, the Planctomycete phylum family member Germataceae and Chitinophagaceae family member of the Bacteroidetes phylum are particularly notable in driving beta diversity in EWC samples (Figure 9).
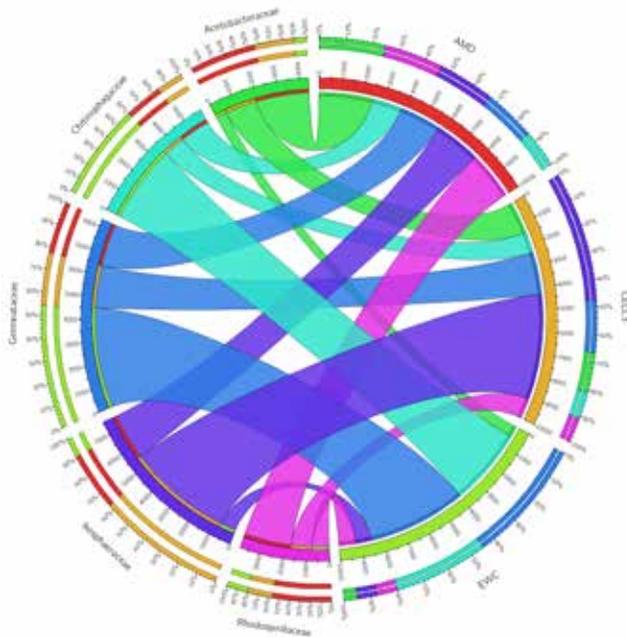
Figure 9. Circos Plot of family-level taxa driving beta diversity between AMD, Cell 3, and EWC sampling locations. Each family-level taxon indicated on the left side of the plot is represented with a colored ribbon. The width of each thread of the split ribbon directly correlates with the number of normalized average OTU reads associated with the respective sample location that the ribbon bridges to on the right side of the figure. The outer ring segments are color coded along a percentage scale of total reads for the family member that associate with each sample location. CSS Normalized OTU table was positive filtered for the respective family-level taxa. Read counts were averaged for each grouped microbiome.

Interestingly, a group of unassigned OTUs was prominent across all the sample locations after taxonomy was assigned using the Greengenes reference database. In the 3-D beta diversity biplots, the unassigned taxonomy group appears to be driving beta diversity in the AMD samples along principal coordinate 1 as well as forcing the wetland Cell 3 sample location to diverge along principal coordinate 2. To further investigate this group, we tabulated the total number of normalized reads across samples and sorted the unassigned OTUs by number of reads, followed by sorting the top unassigned OTUs with maximum relative distribution across all samples. Next, to gain insight into the unassigned identities, we used the reference sequence for each respective unassigned OTU as an input for the Ribosome Database Project (RDP) Classifier algorithm (44). The majority of the top Greengenes unclassified OTUs had a higher relative abundance within the AMD sample location subgroup. Furthermore, the predominant phylum predicted by RDB classifier for the majority of these samples was Chloroflexi, which already was shown to be driving variation along the second principal coordinate. Thus, in conclusion, the unassigned taxa group is composed to a large degree with Chloroflexi taxa, which dominate AMD sample locations, especially wetland Cell 3.

# DISCUSSION

Alpha diversity analysis indicates that the diversity within samples is relatively higher for EWC compared to AMD sample locations, as is expected considering the more challenging physiological conditions associated with low pH and higher soluble concentrations of heavy metals. The lower alpha diversity in AMD samples is significantly supported beyond the number of OTUs by species estimators and indices that took species richness, evenness, and/or evolutionary distance into account. For example, the Chao1 species richness estimator attempts to predict species saturation at the asymptote of the species accumulation curve by considering the number of rare taxa (9). This estimator is thought to provide a more realistic picture than observed species alone. The Shannon Index takes into account both taxa richness and evenness (differential abundance), as a community may have a large number of taxa but be dominated by a few select taxa (37). The Heips evenness alpha diversity index does not yield statistical significance, suggesting that a proportion of additional taxa in EWC may be rare taxa. Recently there has been renewed interest in rare taxa with the evidence suggesting that rare taxa can dictate and influence the dynamics of microbial communities (2). Thus, evenness may be a good descriptive index but not a cut and dry indicator of alpha diversity significance.

Beta diversity analysis indicates that the diversity between sample locations was substantially different between EWC and AMD subgroups along the first principal coordinate (54.64% of explained variance). The second principal coordinate (20.91% of explained variance) further separated the Solomon Mine constructed wetland Cell 3 from the other sample locations, suggesting a large change in community composition compared to the core AMD (SMT, SMA, Cell 1, Cell 2, and SMD) and EWC groups. PCoA biplots indicate that Firmicutes play a role in driving this positioning of the constructed wetland Cell 3 community within the second principal coordinate. Cell 3 also physically differed from the other constructed wetland cells in that it had noticeable algal and aquatic plant growth. Constructed wetland Cells 1 and 2 were heavily stained with iron oxides and lacked noticeable biological activity to the naked eye (Figure 3). The distance matrix comparison, which took all seven collective principal coordinates into account, exhibited a statistically significant difference between AMD and EWC sample location groups, suggesting that the variance between samples were statistically closer to each other than between sample locations.

Historically, the constructed wetland system has had only a minor effect on reducing metal loads and increasing pH after the first three years of service (47). Although our budget did not allow for heavy metal analysis of the water samples, our water-quality data collected during field work (pH and conductivity), along with recent

Table 3. Comparison of our field parameters for Solomon Mine AMD and East Willow Creek water-quality data compared to historical data [a].

| Sample | pH | | Conductivity (µS/cm) | |
|---|---|---|---|---|
| | 1999 [b] | 2015 | 1999 [b] | 2015 |
| SMA | 4.3 | 4.2 | – | 722 |
| SWD | 5.1 | 4.7 | 660 | 691 |
| EW-USM [c] | 6.3 | 6.4 | 50 | 52.7 |
| EW-G | 7.3 | 6.6 | 52 | 57.7 |
| EW-A | 7.2 | 6.3 | 53 | 61.8 |

[a] 1999 estimated flow rate from Solomon Mine adit bulkhead flume was 0.035 CFS [b] and 0.040 CFS in 2015

[b] Data obtained from WCRC, 2004

[c] 1999 data sample site was located ~90 meters upstream of EW-USM (labeled as EW-H by WCRC). A closer EWC sampling site, relative to the Solomon Mine, was chosen due to the close proximity of tailings piles from a separate upstream mining site close to EW-H

water-quality data by the WCRC suggest that heavy metals continue to leach from minerals exposed in the mine workings and enter EWC (Table 3).

This study's budgetary restrictions and experimental design resulted in low statistical power, but interesting changes in relative abundance data of microbes suggests future investigation is warranted. For example, the parasitic bacterium Bdellovibrio was observed to be higher in abundance in AMD samples compared to EWC (FDR adjusted p=0.685). This bacterium is interesting from the standpoint that it has the potential to be used as a means for biological control of AMD production by reducing populations of Gram-negative chemolithotrophs through predation (20). Gallionella is another interesting genus associated with AMD as Gallionella species are iron-oxidizing chemolithotrophs (FDR adjusted, p=0.615) (45). This microbe was observed at higher relative abundance levels within the mine adit, where chemolithotrophs can thrive (FDR adjusted, p=0.311).

More generally, the phylum Proteobacteria was enriched within the AMD sample locations, which makes sense in that they are commonly found in acidic ecosystems (31). The phylum Proteobacteria includes the *Acetobacteraceae* and *Rhodospirillaceae* family members (21). These Alphaproteobacteria and Betaproteobacteria family taxa, respectively, drove beta diversity separation between AMD and EWC sample locations. *Acetobacteraceae* are acid-tolerant aerobes that typically ferment

organic material on the surface of water and are likely to thrive in AMD outside of the mine adit, while *Rhodospirillaceae* family members are purple non–sulfur bacteria that can utilize anoxygenic photosynthesis using hydrogen or low concentrations of sulfide while living in anaerobic aquatic environments, such as the AMD sediment. Constructed wetland Cell 3 contained an abundance of bacteria within the phylum Chloroflexi, which are generally composed of purple and green bacteria that may use sulfide as an electron donor for anoxygenic photosynthesis and have previously been associated with AMD (48). Additionally, the phylum Firmicutes strongly drives the microbial community in Cell 3 along the second principal coordinate axis. There are many sulfate–reducing bacteria that belong to the Firmicutes, such as the Alicyclobacillaceae family found in Cell 3, which likely are reducing sulfate and thus reducing pH and helping precipitate heavy metals. The aquatic bacterial family Isosphaeraceae, which includes a number of characterized moderately acidophilic genera, was also associated with wetland Cell 3 (25, 28). Acidophilic *Isophaeraceae* members are chemohetrorophs that rely on carbohydrates yielded from plant and microbial sources, thus they may be important in the degradation of the abundant organic material observed within Cell 3 (10).

Limitations of this preliminary analysis include low statistical power, which increases the risk of type II error, and thus our ability to identify major taxa that are over/ under–represented in the sample groups and are likely ecosystem biomarkers. Due to budgetary limitations, biological replicates from three triangulated sediment samples for each location were pooled prior to PCR. This technique has been demonstrated to often miss locally rare prokaryotes, but yielded spatially ubiquitous phylotypes at higher taxonomic classifications (30), which was our aim. Future examination of the microbial populations at this project site should address this issue through use of true biological and technical replicates. A series of replicates combined with deeper sequencing will shed light on the role of rare microbial community members on dynamic interactions as well as provide us with a better idea of spatial niches (2, 35). Additionally, we are interested in how seasonal changes may affect the taxonomic profile of the sampling sites. We hypothesize that microbial community composition will be correlated with geochemical variables including pH and metal concentrations as flow rates drop during the late summer/early fall as has been observed in other AMD studies (1). Likewise, temperature fluctuation and changes in dissolved oxygen levels during the winter in this high–elevation, alpine environment will likely drive changes in the taxonomic profile due to reduced photosynthetic activity and diffusion from the atmosphere resulting from ice and snow that cover the constructed wetland cells.

Taxonomic resolution is also a limiting factor in our study. We used the V4 hypervariability region of the SSU 16S rRNA gene (8) with our MiSeq run with paired–end (PE) 150 bp reads that provided 48 bp overlap. Although sufficient for pairing reads to spanning the 252 bp V4

hypervariability region, we would expect to increase taxonomic resolution by sequencing with longer PE reads (250+ bp), which would retain a greater percentage of reads by boosting quality through full overlap of the V4 hypervariability region. Alternatively, longer reads could be used with a new primer set constructed to span both the V3 and V4 hypervariability regions for longer PE reads to increase the data set size and resolution (17). Either approach (increased quality or amount of data), obtained with MiSeq PE 250+ bp reads would likely provide a desired data set through more accurate clustering, assignment of taxonomy, and generation of a more accurate phylogenetic tree (46).

Biases that should be considered when interpreting our data include limitations of using a single marker gene for inferring phylogeny, in addition to biases introduced during PCR. The SSU 16S rRNA gene is a well-documented marker gene and has extensive database records to associate with the data, but is likely not always the best measure of phylogeny between all organisms. PCR primer binding fidelity and the stochastic processes that occur early during amplification can introduce biases in the amplicons produced to be sequenced (36). PCR enzyme choice also has been demonstrated to effect accuracy of microbial community profiling (19). Future studies could address these limitations by taking a true shotgun metagenomic approach to community DNA analysis. Not only would this approach provide taxonomic information through the combination of multiple marker genes, but it would also provide us with

information on the functional potential of the microbes within the ecosystem (14).

For future study, we are interested in investigating the taxonomic composition of biofilms found within the Solomon Mine adit in addition to looking at differences between AMD microbial communities within the adit (no light or source of significant organic matter) to those directly towards the end of the first constructed wetland cell (exposed to light with organic matter). We hypothesize that phototrophs may play a larger role in generating organic matter, while chemolithotrophs are likely to have a reduced role in the bioremediation cells considering the more abundant environmental oxygen and nitrogen sources from organic debris present in those cells. To gain insight into the ecological role of these microbes, we plan to correlate microbe presence with dissolved oxygen levels and sources of carbon and nitrogen. Further shotgun metagenomic analysis could also reveal the relative abundance of pathways for carbon, nitrogen fixation, and energy generation expected for this extreme environment. Association of these geochemical parameters with differentially abundant microbes will allow us to understand symbiotic relationships between taxa in such an ecosystem (43).

Secondly, we are interested in determining the metagenome or metatranscriptome of the microbes found within constructed wetland Cell 3 to better understand the metabolic potential or functionality of the microbial community. Visual inspection alone differentiates the cell from the other AMD sampling location sites considering

the extensive algal growth (Figure 3) and the water-quality measures of an increase in pH and reduction in conductivity suggesting sulfate-reducing bacteria may be at work. The beta diversity analysis also indicates that microbial community composition is significantly different from that of the other AMD sample locations and in EWC. In conclusion, further investigation is necessary to determine if current microbial communities will likely be sufficient to successfully seed a functional renovation of the constructed wetland system.

# ACKNOWLEDGEMENTS

# REFERENCES

1. Auld RR, Mykytczuk NCS, Leduc LG, Merritt TJS. 2016. Seasonal variation in an acid mine drainage microbial community. *Can. J. Microbiol.* 63:137–52.

2. Bachy C, Worden AZ. 2014. Microbial ecology: finding structure in the rare biosphere. *Curr. Biol.* 24:R315–R317.

3. Baker BJ, Banfield JF. 2003. Microbial communities in acid mine drainage. FEMS *Microbiol. Ecol.* 44:139–152.

4.Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA, Caporaso JG. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods.* 10:57–59.

5. Bond PL, Druschel GK, Banfield JF. 2000. Comparison of acid mine drainage microbial communities in physically and geochemically distinct ecosystems. *Appl. Environ. Microbiol.* 66:4962–4971.

6. Brune KD, Bayer TS. 2012. Engineering microbial consortia to enhance biomining and bioremediation. *Front. Microbiol.* 3:203.

7. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI. 2010. QIIME allows analysis of high-throughput

community sequencing data. *Nat. Methods.* 7:335–336.

8. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* 108:4516–4522.

9. Chao A. 1984. Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.*, pp. 265–270

10. Dedysh SN, Kulichevskaya IS. 2013. Acidophilic Planctomycetes: expanding the horizons of new planctomycete diversity. In Planctomycetes: Cell Structure, Origins and Biology, pp. 125–139. Springer

11. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72:5069–5072.

12. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods.* 10:996–998.

13. Egiebor NO, Oni B. 2007. Acid rock drainage formation and treatment: a review. *Asia-Pac. J. Chem. Eng.* 2:47–62.

14. Eisen JA. 2007. Environmental shotgun sequencing: its potential and challenges for studying the hidden world of microbes. *PLoS Biol.* 5:e82.

15. Emmons WH, Larsen ES. 1923. Geology and ore deposits of the Creede district, Colorado. US Government Printing Office

16. Epstein SS. 2013. The phenomenon of microbial uncultivability. *Curr. Opin. Microbiol.* 16:636–642.

17. Fadrosh DW, Ma B, Gajer P, Sengamalay N, Ott S, Brotman RM, Ravel J. 2014. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome.* 2:6.

18. Fenchel T, Blackburn TH, King GM. 2012. Bacterial biogeochemistry: the ecophysiology of mineral cycling. Academic Press

19. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, Gould TJ, Clayton JB, Johnson TJ, Hunter R, others. 2016. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nature.* 201:6.

20. Harini K, Ajila V, Hegde S. 2013. Bdellovibrio bacteriovorus: A future antimicrobial agent? *J. Indian Soc. Periodontol.* 17:823.

21. Hell R, Dahl C, Knaff DB, Leustek T, eds. 2008. Sulfur Metabolism in Phototrophic Organisms. Dordrecht: Springer

22. Huang LN, Kuang JL, Shu WS. 2016. Microbial ecology and evolution in the acid mine drainage model system. *Trends Microbiol.* 24:581–593.

23. Huston RC. 2005. A silver camp called Creede: A century of mining. Western Reflections Publishing Company

24. Johnson DB, Hallberg KB. 2005. Acid

mine drainage remediation options: a review. *Sci. Total Environ.* 338:3–14.

25. Joseph SJ, Hugenholtz P, Sangwan P, Osborne CA, Janssen PH. 2003. Laboratory cultivation of widespread and previously uncultured soil bacteria. *Appl. Environ. Microbiol.* 69:7210–7215.

26. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res.* 19:1639–1645.

27. Kuang JL, Huang LN, Chen LX, Hua ZS, Li S-J, Hu M, Li JT, Shu WS. 2013. Contemporary environmental variation determines microbial diversity patterns in acid mine drainage. *ISME J.* 7:1038.

28. Kulichevskaya IS, Ivanova AA, Suzina NE, Rijpstra WIC, Damsté JSS, Dedysh SN. 2016. Paludisphaera borealis gen. nov., sp. nov., a hydrolytic planctomycete from northern wetlands, and proposal of Isosphaeraceae fam. nov. Int. *J. Syst. Evol. Microbiol.* 66:837–844.

29. Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 71:8228–8235.

30. Manter DK, Weir TL, Vivanco JM. 2010. Negative effects of sample pooling on PCR-based estimates of soil microbial richness and community structure. *Appl. Environ. Microbiol.* 76:2086–2090.

31. Méndez-García C, Peláez AI, Mesa V, Sánchez J, Golyshina OV, Ferrer M. 2015. Microbial diversity and metabolic networks in acid mine drainage habitats. *Front. Microbiol.* 6:

32. Navas-Molina JA, Peralta-Sánchez JM, González A, McMurdie PJ, Vázquez-Baeza Y, Xu Z, Ursell LK, Lauber C, Zhou H, Song SJ, others. 2013. Advancing our understanding of the human microbiome using QIIME. *Methods Enzymol.* 531:371.

33. Oren A. 2009. Chemolithotrophy. eLS

34. Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods.* 10:1200–1202.

35. Pedrós-Alió C. 2012. The rare bacterial biosphere. *Annu. Rev. Mar. Sci.* 4:449–466.

36. Polz MF, Cavanaugh CM. 1998. Bias in template-to-product ratios in multitemplate PCR. *Appl. Environ. Microbiol.* 64:3724–3730.

37. Shannon CE, Weaver W. 1998. The mathematical theory of communication. University of Illinois press

38. Shtarkman YM, Koçer ZA, Edgar R, Veerapaneni RS, D'Elia T, Morris PF, Rogers SO. 2013. Subglacial Lake Vostok (Antarctica) accretion ice contains a diverse set of sequences from aquatic, marine and sediment-inhabiting bacteria and eukarya. *PLoS One.* 8:e67221.

39. Singer PC, Stumm W. 1970. Acidic mine drainage: the rate-determining step. Science. 167:1121–1123.

40. Spitzer M, Wildenhain J, Rappsilber J, Tyers M. 2014. BoxPlotR: a web tool for generation of box plots. *Nat. Methods.*

11:121–122.

41. Takai K, Komatsu T, Inagaki F, Horikoshi K. 2001. Distribution of archaea in a black smoker chimney structure. *Appl. Environ. Microbiol.* 67:3618–3629.

42. Thomas MC, Thomas DK, Kalmokoff ML, Brooks SP, Selinger LB, others. 2012. Molecular methods to measure intestinal bacteria: a review. *J. AOAC Int.* 95:5–23.

43. Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram, Rachna J., Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature.* 428:37.

44. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73:5261–5267.

45. Weber KA, Achenbach LA, Coates JD. 2006. Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction. *Nat. Rev. Microbiol.* 4:752–764.

46. Werner JJ, Zhou D, Caporaso JG, Knight R, Angenent LT. 2012. Comparison of Illumina paired–end and single–direction sequencing for microbial 16S rRNA gene amplicon surveys. *ISME J.* 6:1273.

47. Willow Creek Reclamation Committee. 2004. Report of surface and mine water sampling and monitoring in Willow Creek watershed, Mineral County, CO (1999-2002)

48. Yamada T, Sekiguchi Y. 2009. Cultivation of uncultured chloroflexi subphyla: significance and ecophysiology of formerly uncultured chloroflexi'subphylum i'with natural and biotechnological relevance. *Microbes Environ.* 24:205–216.