Metals Alter Membership but Not Diversity of a Headwater Stream Microbiome

Brian A. Wolff,a,b William H. Clements,a,b Ed K. Halla,c,d

aGraduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado, USA
bDepartment of Fish, Wildlife and Conservation Biology, Colorado State University, Fort Collins, Colorado, USA
cDepartment of Ecosystem Science and Sustainability, Colorado State University, Fort Collins, Colorado, USA
dNatural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado, USA

ABSTRACT  Metal contamination from mining or natural weathering is a common feature of surface waters in the American west. Advances in microbial analyses have created the potential for routine sampling of aquatic microbiomes as a tool to assess the quality of stream habitat. We sought to determine if microbiome diversity and membership were affected by metal contamination and identify candidate microbial taxa to be used to indicate metal stress in stream ecosystems. We evaluated microbiome membership from sediments at multiple sites within the principal drainage of an EPA superfund site near the headwaters of the Upper Arkansas River, Leadville, CO. From each sample, we extracted DNA and sequenced the 16S rRNA gene amplicon on the Illumina MiSeq platform. We used the remaining sediments to simultaneously evaluate environmental metal concentrations. We also conducted an artificial stream mesocosm experiment using sediments collected from two of the observational study sites. The mesocosm experiment had a two-by-two factorial design: (i) location (upstream or downstream of contaminating tributary), and (ii) treatment (metal exposure or control). We found no difference in diversity between upstream and downstream sites in the field. Similarly, diversity changed very little following experimental metal exposure. However, microbiome membership differed between upstream and downstream locations and experimental metal exposure changed microbiome membership in a manner that depended on origin of the sediments used in each mesocosm.

IMPORTANCE Our results suggest that microbiomes can be reliable indicators of ecosystem metal stress even when surface water chemistry and other metrics used to assess ecosystem health do not indicate ecosystem stress. Results presented in this study, in combination with previously published work on this same ecosystem, are consistent with the idea that a microbial response to metals at the base of the food web may be affecting primary consumers. If effects of metals are mediated through shifts in the microbiome, then microbial metrics, as presented here, may aid in the assessment of stream ecosystem health, which currently does not include assessments of the microbiome.

KEYWORDS 16S, bioassessment, mesocosms, metals, stream

Streams in the western United States are frequently impaired from elevated metal concentrations due to a combination of historical mining activities and, to a lesser extent, natural weathering processes. In Colorado, there are approximately 23,000 abandoned mines (1) resulting in approximately 23% of Colorado streams qualifying as impaired (2). Currently, various protocols use stream macroinvertebrate community composition to evaluate stream water quality and ecosystem health (3–5). Macroinvertebrates have a series of characteristics that have proven useful for stream
bioassessments, including ubiquity, high diversity, restricted ranges, short generation times, small size, and their role as important food sources for both aquatic and terrestrial consumers (6). We hypothesized that microbiomes may potentially be better indicators of water quality than macroinvertebrates because they are even more ubiquitous and dynamic and thus may report even subtler differences in water quality. For instance, typical bacterial generation times (i.e., doubling time) occur over hours or days (7) compared to weeks or months for macroinvertebrates (8). The spatial scale at which microbiomes operate is also much smaller than for stream macroinvertebrates, creating the potential to identify small pockets of contamination in heterogeneous stream ecosystems. We now know that microbial biofilms are formed by complex, non-random assemblages of algae, bacteria, and fungi (9) and that these diverse microbiomes can be shaped by physical properties like stream velocity (10) and substrate type (11), as well as chemical properties such as pH (12, 13). It is also clear that metals affect the function of microbiomes, including evidence for metals decreasing stream nitrification (14) and reducing rates of microbial respiration (15). Metals also affect microbiome membership, including evidence where specific subphyla increased (gammaproteobacteria) or decreased (betaproteobacteria) with metal exposure (16).

To test the potential for microbiomes to act as indicators of metal contamination, we evaluated the stream microbiome of the Upper Arkansas River near Leadville, CO, USA (Fig. 1). The Upper Arkansas River has been impaired by metal pollution due to historical mining since the mid-1800s (17). By the late 1990s, implementation of water treatment facilities and removal of floodplain mine tailings resulted in significant improvements in water quality, including decreased dissolved metals—principally cadmium (Cd), copper (Cu), and zinc (Zn)—downstream from where California Gulch enters the main stem of the Upper Arkansas River (17). After significant improvements in water quality, macroinvertebrate species richness increased at the impaired downstream site to similar levels observed at the upstream reaches (17). However, despite improved richness, community membership has remained different between upstream reference sites and sites downstream of California Gulch (18).

To assess how microbiomes were affected by metal exposure in the Upper Arkansas River, we chose to focus on the bacterial component of the stream microbiome because (i) sediment biofilms are primarily composed of bacterial biomass (from 90 to 99%) (19, 20); (ii) many stream macroinvertebrates spend a significant portion of their life cycle in aquatic environments; and (iii) bacteria are known to play a key role in nutrient cycling and energy transfer in freshwater ecosystems. In this study, we explored the hypothesis that differences in bacterial community composition could be used to track metal pollution in the Upper Arkansas River, and we compared the microbiomes of sites upstream and downstream of California Gulch to determine if metal exposure affected bacterial diversity and community structure.
life cycle grazing on biofilm in sediments (21) such that changes in microbiome may have effects on higher trophic levels; and (iii) 16S rRNA gene sequences have better developed sequence libraries compared to analogous phylogenetic markers for other groups, such as the 18S rRNA gene for eukaryotic microbes (22). We collected samples at locations upstream and downstream of California Gulch during both spring and fall seasons. From each sample, we used 16S rRNA gene amplicon sequencing of the sediment biofilm to determine how metals influenced microbiome diversity and membership in the Upper Arkansas River. We complemented our field observations with experiments that exposed microbial communities sampled from both upstream (reference) and downstream (metal impaired) of California Gulch to elevated metal concentrations. The purpose of this study was to examine (i) how microbiome diversity and membership differed in an ecosystem that has elevated metal levels, (ii) if these differences could be attributed to exposure to metals, and (iii) if certain members of the microbiome were consistently enriched or depleted in response to metal exposure and therefore may be candidates for indicators of stream water quality. The last goal is an important step toward identifying mechanistic responses of individual bacteria, aiding in the development of using certain groups as sensitive “indicators” of metal stress. We chose to focus on differences between the upstream and downstream components of the main stem of the Upper Arkansas, where the metal contamination is subtle, rather than on the microbiome of the tributary called California Gulch, where it is clear that metals still significantly impact the ecosystem and the river has habitat features (e.g., narrower channel, less riparian vegetation, smaller substrates, and warmer temperatures) that are not comparable to the upstream or downstream locations of the upper Arkansas River. Additionally, the downstream portion of the Upper Arkansas River is thought to have been restored and we were interested in assessing this from the perspective of the microbiome; thus we focused on differences upstream and downstream of where the California Gulch intersects the Arkansas River.

RESULTS

Observational study. To assess the effect of metals on the Upper Arkansas River microbiome during the spring and fall of 2017, we analyzed surface water parameters (e.g., alkalinity, hardness, pH, specific conductance, and temperature), sediment metal concentrations, and sediment 16S rRNA gene amplicons from California Gulch and sites upstream and downstream of California Gulch in the main stem of the Upper Arkansas River. Of all sampling locations during both seasons, California Gulch had the highest sediment metal concentrations for all three metals, followed by downstream sites, with upstream sites consistently having the lowest sediment metal concentrations (Table 1). At downstream sites there was a statistically significant increase of Cu and Zn in the spring compared to fall, but not for Cd (Table 1). However, we did not observe statistically different metal concentrations between fall and spring for any of the metals at the upstream sites (consistently low) or in California Gulch (consistently high, Table 1).

<table>
<thead>
<tr>
<th>Location</th>
<th>Season</th>
<th>pH</th>
<th>Sediment Cd</th>
<th>Sediment Cu</th>
<th>Sediment Zn</th>
<th>Observed OTUs</th>
<th>Shannon index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>StdDev</td>
<td>Mean</td>
<td>StdDev</td>
<td>Mean</td>
<td>StdDev</td>
</tr>
<tr>
<td>Upstream</td>
<td>Spring</td>
<td>8</td>
<td>7.63 a</td>
<td>0.21</td>
<td>1.59 a</td>
<td>0.41</td>
<td>4.29 a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CalGulch</td>
<td>Spring</td>
<td>4</td>
<td>7.12 b</td>
<td>-</td>
<td>23.66 b</td>
<td>10.15</td>
<td>331.30 b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Downstream</td>
<td>Spring</td>
<td>13</td>
<td>7.41 c</td>
<td>0.05</td>
<td>6.29 c</td>
<td>4.09</td>
<td>53.56 c</td>
</tr>
<tr>
<td>Upstream</td>
<td>Fall</td>
<td>8</td>
<td>7.75 A</td>
<td>0.53</td>
<td>1.74 A</td>
<td>0.58</td>
<td>4.19 A</td>
</tr>
<tr>
<td>CalGulch</td>
<td>Fall</td>
<td>4</td>
<td>7.50 B</td>
<td>-</td>
<td>33.45 B</td>
<td>16.58</td>
<td>382.49 B</td>
</tr>
<tr>
<td>Downstream</td>
<td>Fall</td>
<td>14</td>
<td>7.49 B</td>
<td>0.07</td>
<td>5.14 C</td>
<td>2.81</td>
<td>18.13 C</td>
</tr>
</tbody>
</table>

*Different letters refer to a statistically significant difference (α = 0.05) among locations (upstream, California Gulch [CalGulch], and downstream samples) in the spring (lowercase) and in the fall (uppercase). Asterisks refer to a statistically significant difference (α = 0.05) between spring and fall for each location. OTU, operational taxonomic unit; -, not applicable (sample size of 1).
We also examined surface water pH at the time of sediment sample collection, since the addition of metals can lower pH of the receiving waters and, more importantly, lower pH can make metals more bioavailable to aquatic organisms. While differences in pH among locations were minimal (they ranged between 7.12 to 7.75) in general, samples with higher metal concentrations had statistically significantly lower pH (Table 1). There was also a statistically significant season by location interaction for pH ($P = 0.002$). Spring pH values were different among all locations, with upstream sites having the highest pH values followed by downstream sites, and then California Gulch (Table 1). In the fall, pH at upstream sites was significantly higher than California Gulch and all downstream sites, but there was no statistical difference between California Gulch and downstream locations (Table 1). For all locations, pH was lower in the spring (average range 7.12 to 7.63) than in the fall (average range 7.49 to 7.75), and although differences were statistically significant, the differences were relatively small and water was generally neutral during both seasons.

At each site we assessed the 16S rRNA gene amplicons from all sediment samples. Richness (i.e., number of observed operational taxonomic units [OTUs]) was significantly different among locations and between seasons (Table 1). California Gulch had the lowest richness among all sites in both spring and fall; however, upstream and downstream locations had similar richness in both seasons (Table 1). At downstream locations, richness was significantly lower ($P = 0.0365$) in the spring than the fall, but we observed no significant difference in microbiome richness between seasons at the upstream sites or in California Gulch (Table 1). The Shannon diversity index values were lowest at California Gulch in both spring and fall (Table 1), and there was no statistical difference in diversity between upstream and downstream locations (Table 1). However, unlike richness results, there was no difference in Shannon index values between seasons, or any significant season by location interactions (Table 1).

We also evaluated microbiome membership to determine if composition of the microbiome (i.e., membership) could be explained by metal concentrations, even though indices of alpha diversity may not have differed. We found that microbiome membership was significantly different among locations ($P = 0.001$) and between seasons ($P = 0.015$). The membership of the California Gulch microbiome was different than the membership of the microbiome at upstream and downstream locations for all sampling dates (Fig. 2). Because the difference in membership between California Gulch and either location in the main stem of the Arkansas River was so pronounced, we also performed a PERMANOVA that excluded California Gulch from the data set to focus on differences between upstream and downstream sites. We found a significant difference in membership ($P = 0.003$) between locations and a significant within-site difference between seasons ($P = 0.016$). However, the season by location (i.e., upstream versus downstream) effect was not significant ($P = 0.670$), suggesting that microbiome membership differed between seasons but that changes occurred at both locations. The canonical analysis of principal components (CAP) analysis showed that differences in microbiome membership between locations were statistically correlated, with higher sediment Cu concentrations ($P = 0.034$) in downstream sites, whereas sediment Zn and Cd differences were not significant (Fig. 3). However, other surface water covariates such as pH ($P = 0.011$), specific conductance ($P = 0.035$), and hardness ($P = 0.028$) were also statistically correlated with separation between upstream and downstream microbiomes (Fig. 3). Many of these surface water parameters are correlated with elevated metal concentrations in the upper Arkansas River. Although reductions in pH can increase metal solubility in freshwaters, this is observed in more acidic conditions and downstream pH was neutral during all sampling periods. Therefore, we do not think that pH had an important effect on metal solubility in the downstream locations during the course of this study.

Mesocosm study. We did not find significant effects of location or metals, or an interaction between location and metals, on richness in any of the mesocosms after the 10-day exposure period (Table 2). However, Shannon index values were
consistently higher for the microbiomes sourced from upstream sites compared to the microbiomes sourced from downstream sites for both the control and metal-enriched treatments. When we assessed the effect of the metal treatment on diversity within a location, we found that Shannon index values were significantly lower in the metal-treated samples for the downstream location, but we did not see a similar change in diversity in response to metals for the upstream location (Table 2).

Interestingly, in the metal-treated samples, more Cu and Zn were retained in the downstream sediments compared to the upstream sediments over the course of the experiment (Table 2). Although downstream sediments likely began with greater metal concentrations (inferred from the observational study), we did not see a similar difference in concentrations in the control samples between upstream and downstream sediments. This suggested that the metal-treated sediments retained or accumulated metals at a higher rate than upstream sediments during the course of the experiment. This also mirrored the response of the stream microbiome, where the effect of metals depended on the location from which sediments were sourced. Over the course of the incubation, microbiomes from the downstream site showed a more pronounced change in membership to metal exposure than the upstream microbiome (Fig. 4). This result was supported by a PERMANOVA that identified significant differences in membership between location ($P=0.001$), treatment ($P=0.001$), and the location by treatment interaction term ($P=0.007$).

**Family-level responses between observational and mesocosm studies.** In order to assess if microbiome membership was altered by the presence of metals similarly in our observational and experimental studies, we used log$_2$-fold plots to evaluate changes in bacteria to the phylogenetic level of family (all OTUs identified to a common family were aggregated) between the two components of the study. For the observational study we focused on the upstream (AR1) and downstream (AR5) sites that were used to seed the mesocosm experiments. The upstream site was significantly ($P \leq 0.01$) enriched...
in the following families: Blastocatellaceae; Pyrinomonadaceae; a family from Subgroup 9 (Acidobacteria); a family from 1–20 and an unclassified family from Chloroflexi; an unclassified family in Omnitrophiceota; bacteriap25 and Methyomonaceae (Proteobacteria); Methylomirabiaceae and a family from Rokubacteriales (Rokubacteria); and a family from Latesciabacteria (Fig. 5). The downstream site was significantly (P ≤ 0.01) enriched in the following families: Oligoflexaceae and C2U (Proteobacteria); Leptolyngbyaceae and Oxyphotobacteria (Cyanobacteria); an unclassified family in Verrucomicrobiales (Verrucomicrobia); and Amoebophilaceae (Bacteroidetes) (Fig. 5).

Both downstream and upstream microbiomes were enriched in the family env.OPS 17 (Bacteroidetes) in the metal-treated mesocosms (Fig. 6). Conversely, unclassified families from Flavobacteriales (Bacteroidetes) and Proteobacteria were significantly enriched in control mesocosms (Fig. 6). None of the significantly enriched families in the control samples matched families from the upstream samples when comparing the log2-fold plots. However, families from Verrucomicrobia and Oxyphotobacteria were significantly enriched in metal-treated samples and also in the downstream-sourced samples.

**DISCUSSION**

In the observational study of the Upper Arkansas River, measures of microbiome alpha diversity (richness and Shannon index values) were not different between sites upstream and downstream of California Gulch despite consistently higher sediment metal concentrations of Cd, Cu, and Zn at downstream sites. However, the California Gulch microbiome did have significantly lower richness and Shannon index values than either the upstream or downstream locations. These results suggest that metal concentrations may need to exceed a higher concentration before having significant effects on microbiome diversity, or that there are other environmental factors in
California Gulch (e.g., physiochemical and/or habitat availability) that limit microbiome diversity. The reported effects of metal contamination on diversity have been mixed, where some studies have reported decreased richness or diversity in the presence of heavy metals (23–25) while others have reported increases in richness or diversity (26–28). Some previous research has also shown that mining activities or metal-impaired streams have very little impact on microbiome diversity, particularly when the metal contamination does not have a pronounced impact on pH (29, 30), as was the case in this study. Similar effects of metal exposure on macroinvertebrate diversity have been reported from these same locations in the Upper Arkansas River. Specifically, overall species richness for macroinvertebrates upstream and downstream of California Gulch was found to be similar, both within and between seasons (17). However, membership for the macroinvertebrate community was different between upstream and downstream locations (18), with the metals affecting some taxa more than others.

Because metal delivery to the stream is most pronounced during spring runoff, we hypothesized that differences in microbiome membership between locations would be greatest in the springtime, and downstream membership would change the most between seasons. However, seasonal variation in microbiome membership occurred at both upstream and downstream locations. When we examined microbiome membership among sites, the majority (39.6%) of the variance in membership explained by the first two principal coordinates was driven by the differences between the California Gulch microbiome and the two main stem locations (Fig. 3). Although subtle, membership of the downstream microbiome was more similar to the California Gulch microbiome and this similarity was more pronounced in spring than fall, consistent with the idea of metals exerting an influence on the membership of downstream microbiomes. However, California Gulch metal concentrations remain above current EPA metal standards (31), and the goal of this study was to see if the microbiome would report on environmental impact even when other criteria were met. For these reasons we focused our analyses on differences between the upstream and downstream sections of the Upper Arkansas River because the downstream reach passes all current criteria for water quality. We found significant differences in membership between upstream and downstream locations (Fig. 3). The differences in microbiome membership between upstream and downstream sites are notable because concentration of metals in the surface water downstream of California Gulch are typically below EPA chronic aquatic life criteria (31).

In the mesocosm experiment, we expected to see greater changes in microbiome membership of upstream microbiomes in response to the metal treatments, since this site has historically had lower metal exposure and we anticipated the microbes would be more sensitive to metal stress. However, we observed a greater change in downstream microbiome membership in response to metal exposure, in contrast to our expectations. One potential explanation for the differences in treatment effect between locations is that the samples from the downstream location had metal-resistant bacteria present within the microbiome, whereas the upstream samples did not or had fewer. Thus, the downstream microbiome had a more rapid response to metal exposure.
exposure than the upstream community after the 10-day treatment. This mechanism is supported by the lower evenness in samples sourced from the downstream location following metal exposure compared to upstream microbiomes (Table 2), suggesting an enrichment of metal-tolerant organisms altered the rank abundance of the downstream microbiome. A recent study examining the effects of the antimicrobial drug ciprofloxacin also reported more pronounced differences in microbiome membership from experimental exposure to ciprofloxacin along a gradient of urbanized streams in New York (32). The greatest differences in microbiome membership were observed in stream reaches with the highest ambient concentrations of ciprofloxacin.

Previous investigations at these same sites in the Upper Arkansas River have shown that experimental metal exposure resulted in much greater effect on composition from macroinvertebrate communities sourced from sites upstream of California Gulch (18). We posit that the discrepancy between community responses in macroinvertebrates versus microbiomes in response to metal exposure was due to the relative timescale of our study. For instance, over a 10-day period of metal exposure, observed differences in macroinvertebrate membership are likely driven by mortality within the original community, whereas microbiomes may experience multiple generations during that same time period. Thus, microbiome membership was likely not only altered by mortality but also by enrichment of metal-tolerant taxa, which was more pronounced at the downstream compared to the upstream sites.

It is also possible that the physiological mechanisms may be an important factor for macroinvertebrate community and microbiome response to metal exposure. Macroinvertebrate responses include detoxification (e.g., metallothioneine and glutathione proteins), sequestration (e.g., extracellular capsules), and elimination (33). It has been demonstrated that certain caddisfly families found in the Arkansas River (e.g., Hydropsychidae and Rhyacophilidae) are much more efficient at eliminating metals than members of the mayfly families (e.g., Ephemereillidae and Heptageniidae).
Resistance responses of bacteria to metal exposure include active transport (efflux pumps), sequestration (intracellular and extracellular), detoxification, reduction of metal ions, reduced sensitivity of cellular targets, and production of permeable extracellular barriers, (e.g., extracellular polymeric substance [EPS]) (35). Many of the bacterial responses are generically shared by macroinvertebrates, but important differences exist. For instance, metallothionine production by macroinvertebrates is a relatively common response (36). However, relatively few bacteria appear to possess genes required for metallothionine production (e.g., bmtA and mymT families) and efflux pumps are a more common response mechanism (37). Production of EPS (35) and horizontal gene transfer (38) facilitate rapid responses to metals that are exhibited solely by microbiomes. Additionally, differences within bacteria (e.g., autotrophs and heterotrophs) can determine cellular metal requirements. For example, metal demands of cyanobacteria are typically higher than other bacterial groups to facilitate nitrogen fixation (e.g., Fe, Mo, and V) and photosynthesis (e.g., Cu, Mg, and Mn) (39). These physiological differences could explain much of the different macroinvertebrate and microbial community responses to metal exposure. The timescale for these responses also takes much longer to take effect. For instance, the time for macroinvertebrate tissues to reach steady state following exposure can take 50 to 450 days (40), whereas for microbes, equilibrium can be reached in as little as 1 h (41).

We also observed differences in the amount of metals retained in sediments sourced from different locations. Although statistically different, metals in control samples at the end of the experiment were only 1.8-fold (Cu) and 1.2-fold (Zn) higher in downstream samples, whereas metals in treatment samples at the end of the experiment were 4.0-fold (Cu) and 3.0-fold (Zn) higher in downstream samples (Fig. 7). Of the bacterial resistance mechanisms, an increase in metal retention in sediments through greater production of extracellular polymeric substance (EPS) from downstream-sourced samples is best supported by the data presented here. Stream biofilm EPS can retain metals in proteins, polysaccharides, and aromatic components (42), and bacterial EPS production can be enhanced in the presence of...
metals (43). Therefore, microbiome metal tolerance may further exacerbate metal exposure to higher trophic levels by retaining metals in their biomass, including the EPS produced by stream biofilms. Recent studies have shown that macroinvertebrates derive much of their metals from diet and not just from aqueous exposure (44–46). Interestingly, in a recent experimental study, Zn concentrations in periphyton samples at similar levels as the downstream metal-treated biofilms caused dramatic reduction (>75%) in mayfly abundance (47). Additionally, metal-resistant populations of oligochaetes in Foundry Cove, New York, increased metal exposure to higher trophic levels by production of metal-binding proteins in their tissues (48). If the downstream microbiome did produce more EPS in response to metal

---

**FIG 6** Log$_2$-fold change plot of phylogenetic families between metal-treated versus control microbiomes (sediment + periphyton floc) from upstream (upper panel) and downstream (lower panel) communities. Families with a positive value were more enriched in metal-treated (MET) samples, while those with negative values were more enriched in control (CON) samples. The color of each dot is the phylogenetic order.
exposure (thus retaining more metals), this would suggest a mechanism for how the microbial response to environmental stress (increased EPS production) altered the diets of the next trophic level (i.e., stream macroinvertebrates). Dietary exposure to metals, or decrease in resource quality, provides a mechanism that could explain the differences in macroinvertebrate membership between upstream and downstream locations previously reported for this same location (18). We did not directly measure EPS in sediment samples in field observation samples or from the mesocosm experiment, so this explanation, although plausible, remains speculative and merits caution as a causal mechanism.

**Similarity in response between the observational and mesocosm studies.** When we compared the membership of microbiomes in the field to those incubated in the experimental streams for ~10 days we did not observe strong association between samples from observational and mesocosm studies at the family level. In fact, aggregated OTUs from only two groups (*Oxyphotobacteria* and *Verrucomicrobia*) were found in both metal-treated mesocosm and downstream observational samples. However, mesocosm results were helpful in identifying how certain groups responded to metal treatment in a controlled setting. Aggregated OTUs from unclassified families in *Flavobacteriales* and *Proteobacteria* were found to be enriched in control mesocosms from both upstream and downstream locations. Similarly, aggregated OTUs from the family within *env.OPS 17* were found in metal-treated microbiomes from both locations. Because of their dynamic response to metal enrichment in the mesocosm experiments, the above-mentioned taxa represent the best candidates from our study to assess the impact of metals on stream ecosystems.

Whereas the mesocosm experiment was designed to isolate the effect of metals on the stream microbiome, our experimental design may have also introduced other confounding factors. For example, comparison between the field observations and the mesocosm experiments were complicated by differences in source water chemistry. Mesocosms received natural water inoculum from the hypolimnion of a large reservoir (Horsetooth Reservoir) and not Upper Arkansas River water. Therefore, it is possible that differences in membership between the mesocosm and the field samples may be due in part to differences in source water and the microorganisms that were associated with the water from each ecosystem. In addition, the sites in the Upper Arkansas River are open canopy and the downstream site is located downstream from the town of Leadville, CO (pop. ~3,000), whereas the reservoir water is comparatively lower in nutrients and sourced from the aphotic hypolimnion. Thus, differences in light environment and water chemistry between the field and the laboratory may also contribute to enrichment of certain families due to other factors that were not influenced in the same way in the mesocosm studies. Additionally, California Gulch is likely enriched in ammonia and other nutrients from the wastewater treatment process that may be responsible for some observed differences.

**Microbially mediated metal effects on stream ecosystems.** The results we present here suggest that the microbiomes in the Upper Arkansas River downstream of

![FIG 7 Zinc concentrations of sediment + periphyton floc after the 10-day exposure period. All metal-treated samples (Metal) were dosed at a target concentration of 650 μg/liter Zn.](image-url)
California Gulch are still responding to metal stress and may be influencing the previously observed (18) differences in composition of higher trophic levels in the same location. Even though diversity metrics suggest similarities between upstream and downstream microbiomes, differences in membership indicate that metals are impacting the Upper Arkansas River even after extensive restoration efforts have lowered surface water metals below current U.S. EPA criteria levels (31). However, the patterns we observed between upstream and downstream microbiomes (i.e., comparable alpha diversity but distinct membership) were similar to patterns found for the macroinvertebrate communities at the same site (18). In the mesocosm study, the sediments sourced from the downstream site accumulated more metals during the course of the mesocosm experiments than those sourced from the upstream (reference) site. We did not measure the metal content of the sediment before the start of the experiments, so we do not know the metal concentrations at the beginning of the mesocosm experiment. However, in the control mesocosms, metals from downstream and upstream sediments were not significantly different at the end of the experiment. The increase in metal content in the sediments that have previously been exposed to metals in the metal treatments suggests a possibility that previously exposed microbiomes respond to metal stress by altering biomass in a way that increases metal content. This shift in the composition of the microbiome may alter dietary exposure to metals or decrease nutritive quality (e.g., generation of excess EPS to metal exposure [49–51], or both), and may cause shifts in the macroinvertebrate community composition that was observed in the Upper Arkansas River (18). This mechanism is further supported by the much lower abundance of functional feeding groups in the downstream communities that would be indicative of a dietary shift from biofilm to water column filterers (18). A previous study (18) found that upstream macroinvertebrates were enriched in mayflies and other “scrapers” (scraper is the common name given to insects of the functional feeding type that “scrape” biofilms from rocks as a food source), while downstream communities were enriched in caddisflies and other water column filterer taxa. It is becoming increasingly evident that dietary exposure is as important as direct exposure to aquatic life (52–54) and should be considered when assessing the impact of metals on stream ecosystems. If the dietary exposure is mediated through shifts in the microbiome, then microbial metrics, as presented here, may provide a better alternative to assess the impact of metals on stream ecosystems.

In conclusion, we show that even after extensive restoration efforts, low levels of metal exposure may alter membership of stream microbiomes. We also show that microbiomes previously exposed to metals have a differential response to metals than those that have not been previously exposed. We cannot conclusively link the response of the microbiome to metals to changes in diet quality of their primary consumers, stream macroinvertebrates. However, several results presented in this study in concert with results of previous studies (17, 18, 55) on this same ecosystem are consistent with the idea that a microbial response to metals at the base of the food web may be affecting consumers one trophic level above. If this is indeed the case, then it suggests that the current aquatic life criteria that uses chronic exposure of aquatic macroinvertebrates to assess stream health (a threshold below which is thought to be protective of ~95% of the aquatic community) is insufficient to assess the impact of metals on stream ecosystems. Our research suggests current best practice guidelines of stream water quality (e.g., EPA aquatic life criteria) may miss important impacts of metal contamination on stream ecosystems and additional assessments (e.g., dietary exposure, microbial metrics) should be included in future assessment of stream health as these standards are improved.

MATERIALS AND METHODS

Study site. We conducted our observational study on the Upper Arkansas River, located near the town of Leadville, approximately 100 km west of Denver, Colorado. This area of the Upper Arkansas has been monitored since 1989 and the site conditions are well characterized in previous studies (17). Briefly, this area is approximately 2,820 m above sea level and typically receives ~30 cm of precipitation.
annually. The Arkansas River has a snowmelt-driven hydrograph, with peak discharges in May or June, normally reaching baseflow by July or August. Variable runoff alters streamflow and contributions of solutes (including metals) from the watershed, resulting in higher metal concentrations recorded during spring (i.e., during snowmelt runoff) compared to the fall (i.e., at baseflow) (17). In the study reach, the stream substrate was primarily composed of medium to large cobbles in a matrix of gravel and sand. Most riparian vegetation was composed of sagebrush (Artemisia spp.), grasses, and willows (Salix spp.) trees.

**Observational study.** We sampled sediment bacteria communities in the main stem of the Arkansas River at three locations: two sites upstream (AR1 and AR2), four sites downstream (AR3, AR4, AR4G, and AR5), and one site within the principal metal-contributing tributary, California Gulch (Fig. 1). At each site we collected samples to be analyzed for metal concentration and 16S rRNA gene amplicon sequencing in spring (first week of May) and fall (first week of October) of 2017. For each sampling event, we collected 3 to 4 sediment samples in riffle habitat (with a water column depth of ~0.25 m) at each of the seven sites. For taking sediment samples, we removed a large cobble (~0.3 m diameter) and scooped underlying sediments into separate clean 50-ml centrifuge tubes. Sediment samples were treated as independent samples and not homogenized for a single site.

**Experimental mesocosm.** To more explicitly test the effects of metals on stream microbiomes, we designed an artificial mesocosm experiment using samples from an upstream and downstream location. Specifically, we tested if experimentally manipulated metal concentrations would result in similar effects on the microbiome as seen from the metal gradient in the Upper Arkansas River. The observational metal gradient study was conducted in spring and fall; however, the mesocosm experiments were conducted only in the fall because we were primarily interested in the differences in communities under stable conditions (e.g., base flow) and less by short-term seasonal effects from spring snowmelt. The design and parameters of the mesocosm experiments have been described elsewhere (56). Briefly, biofilms for the experiments were collected by placing plastic trays containing clean (scrubbed and air-dried) cobbles in the river for 31 days (5 September 2017 to 6 October 2017) to allow for sediment deposition and colonization of the cobble within each tray by macroinvertebrates and microbial biofilms. All cobbles used in the trays were from a common source and therefore not different between locations. Trays were deployed at one reference site upstream of California Gulch (AR1; called “upstream”) and one site downstream of California Gulch (AR5, called “downstream”). The sediments were derived from within the stream during colonization. Periphyton, sediments, and bacteria (and aquatic insects) were all colonizing during duration of stream deployment. Upon retrieval, four colonized trays were collected from each site and placed into individual coolers filled with ambient stream water and then immediately transported to the Colorado State University (CSU) Stream Research Laboratory (~3 h from the sampling site). The four trays from each cooler were then placed into an individual experimental “racetrack” stream and, after an equilibration period (~24 h), were randomly assigned to a control (no metals added) or treatment (metal exposure).

Each artificial stream received source water from the hypolimnion of a mesotrophic reservoir (Horsetooth Reservoir) that was delivered at a rate of 1.0 liter min⁻¹, resulting in a residence time of approximately 20 min for each mesocosm. Characteristics of the source water (e.g., pH, conductivity, temperature, dissolved oxygen) were typical of nonpolluted mountain streams in Colorado (57). We implemented a two-by-two (location-by-treatment) factorial experimental design: (i) control-upstream, (ii) control-downstream, (iii) metals-upstream, and (iv) metals-downstream. Each control and treatment were replicated four times for a total of 16 experimental streams. We started metal additions after a 24-h acclimation period. Peristaltic pumps delivered stock solutions of a metal mixture from a 20 liter concentration carboy at a rate of 10 ml min⁻¹ to obtain a targeted concentration of 25 μg liter⁻¹ Cu and 650 μg liter⁻¹ Zn for each treatment. Paddlewheels provided a constant flow of 0.35 m s⁻¹ to each mesocosm. Metals concentrated in the carboys were refreshed daily during the 10-day experiment. We checked water and peristaltic pump flows twice daily to ensure consistent delivery of metal solutions among treatments. We measured ambient metal concentrations from each mesocosm by filtering (0.45 μm) 15-ml water samples on day 2, day 4, and day 10 of the experiment. On day 10, all trays from each stream (16 streams; 4 trays combined per stream) were sieved (350 μm) to remove macroinvertebrates. All material (e.g., periphyton and sediments) that passed through the sieve was poured into a clean, 2-gal plastic bucket. Buckets were then decanted and the sediments and periphyton floc was transferred into 50-ml centrifuge tubes and frozen at ~80°C until DNA extraction and metal analysis.

**DNA preparation and 16S rRNA gene amplicon sequencing.** We extracted DNA from each sediment sample using a MoBio PowerSoil DNA isolation kit (Qiagen, Hilden, Germany) using standard protocols. Sediment grain sizes were not normalized for the observational field samples. Our rationale for such an approach was to integrate natural sediment communities and thus did not differentiate based on any particular sediment size. However, in order to remove macroinvertebrate samples from mesocosm samples, we had to use a 350-μm sieve. We removed sediment from vials in the exact manner for all samples, but we acknowledge that some samples had variable size composition. The 16S rRNA gene (V4 region) was amplified using 515F and 806R universal primers with the forward primer barcoded following the Earth Microbiome Project protocols (58). The forward primer 515F included the unique sample barcode following Parada et al. (59), and both primers included degeneracies as described in Parada et al. (59) and Apprill et al. (60). For each sample, we ran a 50-μl PCR using a Platinum Hot Start PCR Master Mix (Invitrogen, Carlsbad, CA, USA) with 10 μl of DNA. The PCR product was quantified and then pooled into a single pool in equimolar concentrations and cleaned using a MinElute PCR purification kit (Qiagen, Hilden, Germany). Cleaned, pooled DNA was sequenced with a MiSeq reagent v2 500 cycle kit.
on the MiSeq (Illumina, San Diego, CA, USA) platform at the CSU Next Generation Sequencing Core facility.

Sequence reads were analyzed using mothur (61) standard operating procedure. Operational taxonomic units (OTUs) were constructed at a 97% similarity of the sequence using the OptiClust algorithm. Generated OTUs were then aligned to a SILVA reference file (62). After sequences were processed through the mothur pipeline, we imported the data into R studio (63) for statistical analyses and visualization utilizing the package Phyloseq. Sequences were preprocessed to remove chloroplasts, mitochondria, and OTUs that were not counted at least 3 times in 20% of the samples. Following the preprocessing steps, raw OTU counts were transformed to relative abundances. We did not rarefy sample OTU counts since that approach potentially excludes important data and does not appropriately address issues of uneven sample library sizes (64). However, we did examine Shannon index values and OTU counts after rarefying to ensure there was no bias from uneven reads among samples and found there was no discernible difference between rarefied and nonrarefied data.

For DESeq2 analyses, we did not normalize count data to relative abundances because DESeq2 algorithms require raw sequence count data inputs. We also aggregated all OTUs that shared the same family before performing DESeq2 analysis. We visualized DESeq2 results with log2-fold change plots analyses.

**Metal preparation.** We extracted metals from the remaining sediment samples that were subsampled for DNA preparation using a modification of EPA Method 3050B (65). We dried sediments in a drying oven at 60°C for at least 24 h with periodic weighing of each sample until no more mass was lost and the sample remained at a constant weight. A small amount of sediment (0.14 to 0.25 g) was then weighed and transferred into 15-ml sterile centrifuge tubes. Next, 1 ml of trace-element grade nitric acid and the sample remained at a constant weight. A small amount of sediment (0.14 to 0.25 g) was then drying oven at 60°C for at least 24 h with periodic weighing of each sample until no more mass was lost.

To determine if clusters from each location/season were statistically different from each other, we used a PERMANOVA model. For the in situ Arkansas River bacterial communities, we tested the effect of location (upstream versus downstream), season (spring versus fall), and their interaction. PERMANOVA tests were also used for pairwise comparisons (e.g., upstream-spring versus downstream-spring; upstream-fall versus downstream-fall, etc.). We performed a canonical analysis of principal components (CAP) analysis (67) on downstream and upstream microbiomes to test if microbiome group separation could be explained by sediment metals or by other surface water measurements (e.g., pH, hardness, alkalinity, and specific conductance).

After evaluating whole community membership differences, we used log2-fold plots to visualize what, if any, microbiome families were significantly enriched or depleted in the upstream or downstream locations. Note, we had performed the same analysis at a lower taxonomic resolution, but most bacteria were unclassified at the genus level. We used a low alpha value (α ≤ 0.01) for Benjamini-Hochberg post hoc adjusted P values to protect against type-I error when identifying families that were differentially enriched between upstream and downstream locations.

If certain families were more enriched at downstream locations (positive log2-fold value), we considered them as potential indicators of metal-tolerant bacteria. Similarly, families more enriched at upstream sites (negative log2-fold value) we considered to be indicators of metal-sensitive bacteria. To aid in comparison with mesocosm results, we only included samples sourced directly from AR1 (upstream) and AR5 (downstream) locations for log2-fold plots.

**Mesocosm statistics.** For the mesocosm experiment samples, we used a two-by-two factorial design that tested the effect of metal treatment (control versus treatment) and location (upstream versus downstream). Each treatment (control and metal treated) had four downstream (ARS) and four upstream (AR1) independent replicate streams. We used many of the same statistical and visual analyses as described above for the observational study. One notable difference is that for mesocosm log2-fold plots, we compared metal-treated groups (positive log2-fold values) to control groups (negative log2-fold values) separately by each location.
ACKNOWLEDGMENTS

Portions of this work were supported by NSF IOS award DEB number 1456959 awarded to E.K.H., the Warner College of Natural Resources, the Colorado Water Center, and the Colorado Mountain Club, Colorado Water Institute, and SETAC Student Training Exchange Opportunity awards to B.A.W.

We are grateful to Mike Wilkins and three anonymous reviewers who provided valuable feedback that greatly improved the manuscript. We thank Colorado State University Natural Resource Ecology Laboratory and Aquatic Ecotoxicology Laboratory students and staff for laboratory assistance.

REFERENCES

20. Reis MP, Barbosa FA, Chartone-Souza E, Nascimento AM. 2013. The prokaryotic community of a historically mining-impacted tropical stream


Poteat MD, Buchwalter DB. 2013. Four reasons why traditional metal toxicity testing with aquatic insects is irrelevant, p 887–888. In Environmental science and technology. ACS Publications, Washington, DC.


