Environmental Toxicology and Chemistry—Volume 37, Number 10—pp. 2575–2592, 2018
Received: 6 December 2017 | Revised: 28 December 2017 | Accepted: 22 June 2018

Environmental Toxicology

Long-Term Effects and Recovery of Streams from Acid Mine Drainage and Evaluation of Toxic Metal Threshold Ranges for Macroinvertebrate Community Reassembly

David B. Herbst,[a,]* R. Bruce Medhurst,[a] and Ned J.P. Black[b]

[a]Sierra Nevada Aquatic Research Laboratory, University of California, Mammoth Lakes, California, USA
[b]Region 9, US Environmental Protection Agency, San Francisco, California, USA

Abstract: Monitoring of benthic invertebrates in streams receiving acidic metal-contaminated water over an 18-yr period revealed both degraded conditions and recovery along a network of downstream locations. Compared with reference streams, and over the course of clean-up remediation efforts below an abandoned open-pit sulfur mine in the central Sierra Nevada of California, improving water quality was accompanied by recovery of benthic communities at some sites. Years of high flow resulted in degraded biological status when acid mine drainage capture was incomplete and metal loading had increased with runoff. Seasonal patterns of recovery evident in the fall after the summer treatment season reverted in the next spring after overwinter periods when sources were not captured. As the metal load has been reduced, phased recovery of community structure, function, and similarity progressed toward that of reference assemblage taxonomic composition. From impacted communities dominated by relatively tolerant midges, reassembly involved an increase in density, return of long-lived taxa, an increased ratio of sensitive-to-tolerant forms, then overall diversity and community composition, and eventually large predators and grazers reappearing along with mayfly, stonefly, and caddisfly richness. Threshold effect levels defined using several analysis methods showed that the response range of biological indicators corresponds to US Environmental Protection Agency guidelines of predicted effects utilizing cumulative criterion units (CCUs) of metal toxicity (i.e., CCU $\sim$ 1). All sites have shown improved function with increased density of some or all trophic groups over time. Although recovery is progressing, year-around treatment may be necessary to fully restore biological integrity in streams nearest the mine. *Environ Toxicol Chem* 2018;37:2575–2592. © 2018 SETAC

Keywords: Acid mine drainage; Benthic macroinvertebrates; Biomonitoring; Cumulative criterion units; Dose–response modeling; Ecotoxicology; Metal toxicity; Mine effluents; Superfund site; Leviathan mine; Sierra Nevada; Headwater streams

INTRODUCTION

Open-pit and strip mining have been used throughout the world to extract mineral ore and coal deposits from under the ground. In the presence of water and oxygen, exposed soils and excavated overburden can undergo chemical and biological oxidation to produce sulfuric acid from pyrite minerals, coal seams, and sulfur deposits (Johnson 2003). With leaching by overland runoff from precipitation and stream flow, substantial amounts of heavy metals, metal oxides, sediments, and acidified waters may enter streams of such mining areas. Acid mine drainage often presents a severe and chronic impact to the biological integrity of stream ecosystems receiving inflow from surface and underground mines (Hogsden and Harding 2012a).

* Address correspondence to david.herbst@lifesci.ucsb.edu
Published online 25 June 2018 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.4217

Evaluation of remedial actions to recover stream ecological integrity in mining areas ideally combines information relating changing metals and water chemistry to ecological indicators as treatment activities are undertaken to reduce acid mine drainage impacts.

The effects of acid mine drainage on water quality and the biological welfare of streams receiving discharges may be sporadic but are often long lasting. Without containment of runoff or reclamation of source areas, open pits often continue to produce chronic discharge of polluted water. The time required to reverse damages to water quality and aquatic life is likely a complex function of the combined effects of differing treatments used to control and ameliorate acid mine drainage (Johnson and Hallberg 2005; Akcil and Koldas 2006), and of local geochemical controls on pH and the metals involved. Recovery further depends on annual variability in stream flows and the availability of unpolluted upstream habitat and flows for dilution and biological recolonization.

Linking acid mine drainage chemistry to biological responses typically shows a correspondence of low pH and elevated metals to toxicity and depletion of the types, abundance, and survival of algae, benthic invertebrates, and fish present in streams (Hogsden and Harding 2012b). Many studies, however, have been limited to short-term observations of correspondence to the spatial extent of acid mine drainage pollution or laboratory toxicity bioassays. Because of their ubiquity, diversity, sensitivity, and established use in aquatic bioassessment monitoring, benthic macroinvertebrates are used in a variety of water pollution studies (Barbour et al. 2000) and provide a measure of biological integrity that has been applied in many settings to monitor acid mine drainage (e.g., Chadwick et al. 1986; Nelson and Roline 1996; Carlisle and Clements 1999; Clements et al. 2000; DeNicola and Stapleton 2002; Gray and Harding 2012; Underwood et al. 2014). Documenting long-term acid mine drainage recovery trends of benthic macroinvertebrates can provide insight to the progress and success of different water treatment alternatives in the control of toxic metals and low pH, showing when and where recovery is occurring. In the present study of acid mine drainage from the Leviathan Mine US Superfund Site in the central Sierra Nevada of California, data gathered over an 18-yr period compare changes in benthic macroinvertebrate community structure, water chemistry, and stream flow along a downstream series below the mine and relative to local reference streams. The goals of the present study have been to provide an appraisal of benthic invertebrate indicator responses that measure the progress of actions to control and alleviate acid mine drainage pollution in the Leviathan Creek watershed, and to evaluate thresholds of impairment by mixed metals to set standards for remediation. Our objectives have been to: 1) quantify the structure and diversity of benthic macroinvertebrate communities along a set of stream sites exposed to variations in metal contamination and in unaffected reference streams, 2) monitor changes in relation to the chronology of capture and treatment of acid mine drainage effluent, and 3) examine the relationship between benthic macroinvertebrate metrics and dissolved metal concentration expressed in terms of predicted chronic toxicity.

Leviathan Mine is an abandoned open-pit sulfur mine site located just north of Monitor Pass on Highway 89 in Alpine County, California (2100-m elevation; northern aspect of drainage). Subsurface mining occurred at Leviathan Mine from the 1860s until the 1940s. During that time, an extensive system of adits (horizontal tunnels) was constructed to facilitate the removal of copper sulfate, copper, and sulfur. From 1952 to 1962, Anaconda Copper operated an open-pit mine, later abandoned, at the site to extract sulfur. The mining disturbance scar covers an area of 253 acres (Figure 1). Within the area of waste rock of the mine site, Leviathan Creek is contained within a cement channel, exiting to a natural channel below the tailings. Acid mine drainage enters Leviathan Creek and Aspen Creek, flows 2.5 km from their confluence to become Bryant Creek where it joins with unpolluted Mountaineer Creek, flowing another 11 km where it enters the East Fork of the Carson River in Douglas County, Nevada. Acid drainage emanates from the following identified locations: the adit, the pit underdrain, the channel underdrain, the Delta Seep, and Aspen Seep. Together these discharges contribute acid drainage containing a mixture of dissolved and particulate toxic metals and orange ferric hydroxide precipitates ("yellow-boy") to Leviathan Creek. In May 2000, the US Environmental Protection Agency (USEPA) listed Leviathan Mine as a Superfund Site to facilitate further site remediation and coordinate planning activities.

Discharge from the exposed mine pit, tailings, adit, and pit underdrain is collected and contained in 4 ponds constructed from 1983 to 1985. These ponds overflowed during late winter and spring snowmelt periods for most years from 1985 until 1999. The channel underdrain and Delta Seep formerly discharged directly to Leviathan Creek but are now captured and pumped during late spring and summer to the holding ponds for treatment. Active seasonal chemical treatment of acid mine drainage sources began in earnest in the fall of 1999 and has continued, with the result that the ponds have seldom overflowed since the spring of 1999. Pond water is typically treated through a biphasic neutralization by addition of lime (in the form of calcium hydroxide) from May through October, settled to remove precipitates, and then discharged to Leviathan Creek. The treatment system was improved in 2009 with a high-density sludge process that provides more complete removal of metals and clarification of water. Since 2001, the channel underdrain has been intercepted and actively treated through lime addition from late spring into fall, depending on weather conditions. The Delta Seep was partly or completely captured during the summers of 2003 to 2004, and from 2007 on, with some new uncontained seep zones forming in 2008 and discharging to Leviathan Creek. Even though there was treatment from May through October of most years, capture and treatment of the channel underdrain and Delta Seep were discontinued, and discharges were returned to Leviathan Creek at the end of each treatment season (October or November) when winter conditions prohibit continued site access. Another acid mine drainage source at Aspen Seep, from overburden deposited into Aspen Creek, has been partially to completely captured and treated year-round since 1998 in a bioreactor that uses sulfate-reducing bacteria to reverse the oxidation that produces acid mine drainage. The Aspen Seep bioreactor was improved in 2003 to a more efficient compost-free system (US Environmental Protection Agency 2006). These actions collectively have substantially reduced, but not entirely eliminated, the discharge of acid mine drainage to Leviathan Creek. During 2005 and 2006, under high runoff, the most substantial changes in treatment regime were a shorter channel underdrain treatment period and loss of Delta Seep capture until 2007. Between the effluent discharge of the Aspen Seep bioreactor and the downstream sample station on Aspen Creek, there are other uncontained seepage sources of acid mine drainage coming from mine tailings and overburden.

There is a core group of 6 sample stations that have been surveyed consistently from 1998 to 2015 (Table 1). Nearest the mine are stations on first-order Aspen below mine and Leviathan below mine just above their confluence; on second-order lower Leviathan Creek (Leviathan above Mountaineer) and at the main reference site on lower Mountaineer Creek (Mountaineer

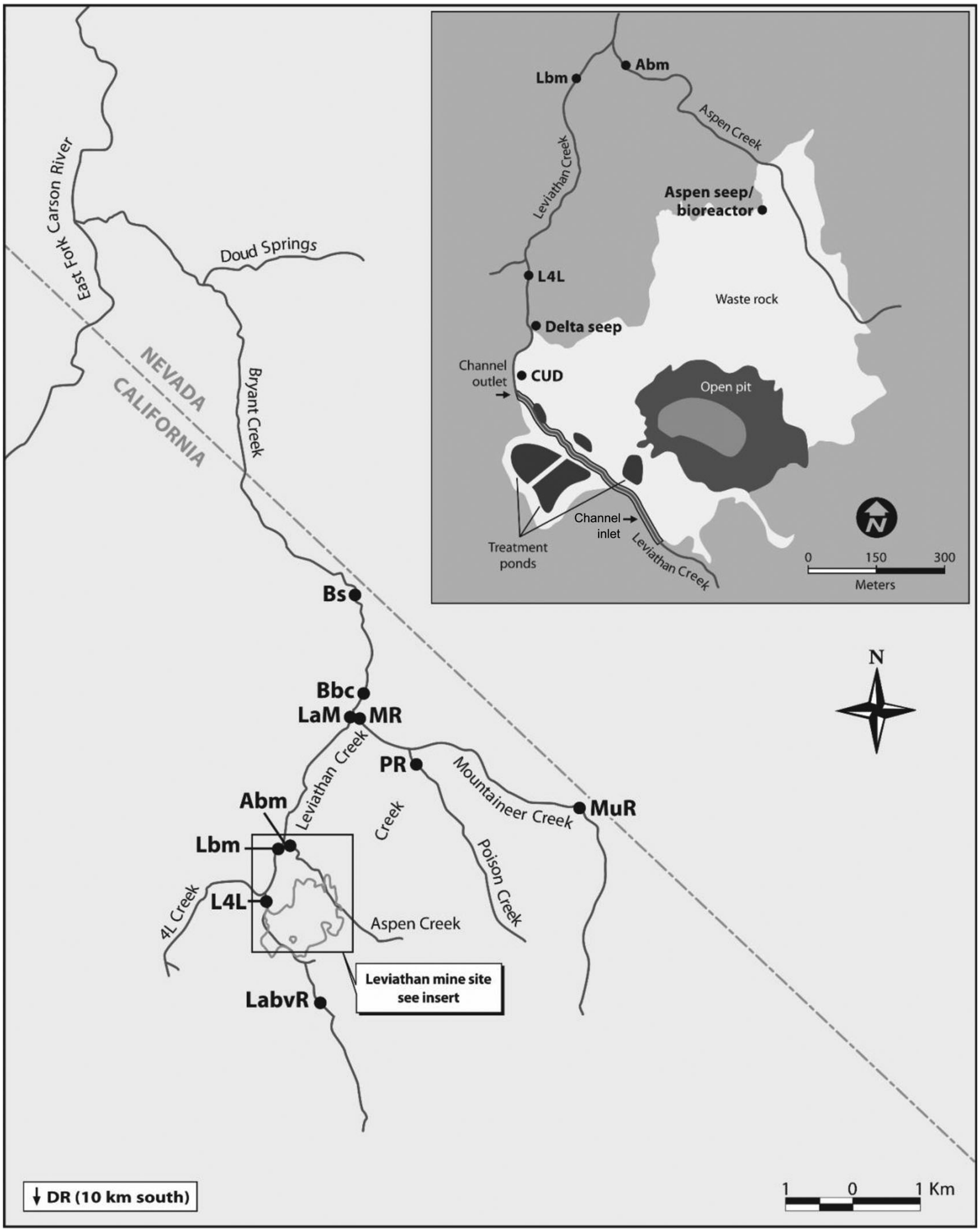

FIGURE 1: Map of Leviathan Mine watershed, with code names for all sites sampled and insert showing the abandoned mine site layout of sources and treatment ponds. Abm = Aspen below mine; Lbm = Leviathan below mine; L4L = Leviathan 4L; CUD = channel underdrain; Bs = Bryant state line; Bbc = Bryant below confluence; LaM = Leviathan above Mountaineer; MR = Mountaineer Reference; MuR = Mountaineer upper Reference.

Reference) just above their confluence; below the confluence where the stream becomes third-order Bryant Creek (Bryant below confluence); and on Bryant Creek near the boundary of California and Nevada state lines (Bryant state line). An additional site just upstream of Leviathan below mine, Leviathan 4L, was sampled in later drought years to substitute data when Leviathan below mine was dry. All study reaches were dominated by gravel and cobble substrata, with gradients of 2 to 4%, width of 2 to 4 m, and depth of 10 to 20 cm. Extensive yellow-boy deposition has occurred at times throughout Leviathan and Bryant Creeks, producing a coating and cementing of rock substratum with iron oxide deposits, especially early during the project; this is chronic at Leviathan below mine and still occurs sporadically at Leviathan above Mountaineer. Although the lime treatment removes metals and neutralizes pH, the specific conductivity of treated effluent entering Leviathan Creek remains well above background levels (ranging from an average of > 300 on Bryant Creek to > 1000 at

TABLE 1: Study site locations comparing average metal (µg/L[a]; and critical concentration units [CCUs[b]]) and water quality 1998–2015

Stream site (site code; no.)	Latitude/longitude	Elevation (m)	Al µg/L CCU	As µg/L CCU	Cu µg/L CCU	Fe µg/L CCU	Mn µg/L CCU	Ni µg/L CCU	Se µg/L CCU	Zn µg/L CCU	Hardness CaCO$_3$ mg/L	SO$_4$ mg/L	EC µS/cm
Leviathan below mine (Lbm; 29)	38.7175	2021	2886[a]	9.7[a]	4.9[a]	11 827[a]	3865[a]	223.5[a]	0.6[a]	40.0[a]	728	811	1310
	−119.6598		33.18[b]	0.06[b]	0.14[b]	11.83[b]	1.29[b]	0.97[b]	0.13[b]	0.08[b]			
Aspen below mine (Abm; 32)	38.7174	2020	34[a]	22.0[a]	4.1[a]	26[a]	124[a]	6.6[a]	5.1[a]	6.3[a]	170	136	406
	−119.6594		0.39[b]	0.15[b]	0.26[b]	0.03[b]	0.06[b]	0.07[b]	1.02[b]	0.03[b]			
Leviathan above Mountaineer (LaM; 33)	38.7351	1913	562[a]	4.2[a]	2.9[a]	1978[a]	792[a]	67.6[a]	2.3[a]	15.0[a]	345	336	677
	−119.6459		6.45[b]	0.03[b]	0.12[b]	1.98[b]	0.32[b]	0.45[b]	0.46[b]	0.05[b]			
Bryant below confluence (Bbc; 33)	38.7375	1907	51[a]	2.2[a]	0.4[a]	466[a]	216[a]	19.1[a]	0.4[a]	4.3[a]	143	96	323
	−119.6440		0.58[b]	0.01[b]	0.04[b]	0.47[b]	0.11[b]	0.24[b]	0.08[b]	0.03[b]			
Bryant state line (Bs; 29)	38.7513	1832	63[a]	2.9[a]	0.5[a]	181[a]	143[a]	9.8[a]	0.4[a]	2.8[a]	145	90	330
	−119.6454		0.73[b]	0.02[b]	0.05[b]	0.18[b]	0.07[b]	0.13[b]	0.08[b]	0.02[b]			
Mountaineer reference (MR; 33)	38.7353	1912	7[a]	1.2[a]	0.2[a]	24[a]	4[a]	0.5[a]	0.2[a]	3.4[a]	66	1	163
	−119.6455		0.08[b]	0.01[b]	0.02[b]	0.02[b]	0.003[b]	0.01[b]	0.04[b]	0.04[b]			
Other reference streams													
Mountaineer upper reference (MuR; 11)	38.7200	2178	11[a]	1.2[a]	1.0[a]	5.5[a]	5.6[a]	0.3[a]	0[a]	2.8[a]	54	0.8	129
	−119.6003		0.13[b]	0.01[b]	0.06[b]	0.006[b]	0.004[b]	0.006[b]	0[b]	0.05[b]			
Poison lower (PR; 2)	38.7295	1987	0[a]	0.7[a]	0[a]	0[a]	3.3[a]	1.2[a]	0[a]	4[a]	69	1.4	181
	−119.634l8		0[b]	0.004[b]	0[b]	0[b]	0.002[b]	0.02[b]	0[b]	0.05[b]			
Leviathan above mine (LabvR; 5)	38.6997	2191	0[a]	10[a]	0[a]	90[a]	8[a]	25[a]	0[a]	10[a]	46	5	144
	−119.6547		0[b]	0.07[b]	0[b]	0.09[b]	0.006[b]	0.33[b]	0[b]	0.16[b]			
Dixon lower (DR; 2)	38.5619	2074	6[a]	0.5[a]	0[a]	0[a]	1[a]	0[a]	0[a]	2.8[a]	18	1.3	62
	−119.7087		0.07[b]	0.003[b]	0[b]	0[b]	0.001[b]	0[b]	0[b]	0.11[b]			
Chronic criteria concentration[c]			87	150	9[c]	1000	1650[c]	52[c]	5	120[c]			

[a] µg/L
[b] Critical concentration units (CCUs)
[c] Chronic criteria concentrations were taken preferentially from the California Toxic Rule, Table 23. If no California Toxic Rule value was available, we used the US Environmental Protection Agency's (USEPA's) National Recommended Water Quality Criteria values. The Colorado Department of Public Health developed a hardness-adjusted criterion for Mn using the methods described in the USEPA Water Quality Standards Handbook (2017), and this adjusted value was used to calculate CCUs. Chronic criteria concentration adjusted value assuming hardness = 100 mg/L.
Cu = exp(0.843 × ln[hardness] − 1.702); Ni = exp(0.846 × ln[hardness] + 0.0584); Zn = exp(0.8473 × ln[hardness] + 0.884); Mn = exp(0.3331 × ln[hardness] + 5.8743).
Al = aluminum; As = arsenic; Cu = copper; Fe = iron; Mn = manganese; Ni = nickel; Se = selenium; Zn = zinc; SO$_4$ = sulfate; EC = electrical conductivity.

Leviathan below mine, compared with < 200 μS/cm at the Mountaineer Creek Reference), and so constitutes a persistent potential chemical stressor. This is in large part caused by sulfate ion enrichment (from the oxidation of sulfur during production of acid mine drainage), calcium, and carbonates (from lime addition). Sulfate averaged approximately 100 mg/L on Bryant Creek to more than 500 mg/L on Leviathan Creek sites, compared with approximately 1 mg/L background at Mountaineer Creek, whereas $CaCO_3$ hardness averaged approximately 150 mg/L on Bryant, more than 500 mg/L on Leviathan, and 66 mg/L at Mountaineer. Along with the core sites, a set of reference sites within the Leviathan watershed or nearby (also draining to the East Carson River) share similar first- to third-order size, physical habitat features, and geological formations of intrusive granitic and extrusive volcanic rocks (basalt, tuff) with concomitant pedogenesis; however, these reference sites have dissolved metals far below the acid mine drainage-exposed streams (Table 1). Multiple reference sites provide an improved assessment of the spatial and temporal variability that represent the potential range of ambient biological integrity defining the recovery of acid mine drainage-exposed streams.

The US Geological Survey gauge (#10308794) located just below the confluence of Leviathan and Mountaineer Creeks, on upper Bryant Creek, shows a hydrograph dominated by snowmelt, with periods of extreme flows from high-runoff wet years in 2005 to 2006 and 2011 to low-runoff dry years in 2001 and 2007, and sustained drought in 2012 to 2015 (Figure 2A). High flows have resulted in treatment pond overflows, usually from early spring runoff not being captured because snow cover prevented access to operate treatment facilities. Low-flow drought periods have sometimes resulted in late summer drying on Leviathan Creek at Leviathan below mine.

MATERIALS AND METHODS

Water quality and metal chemistry

At each macroinvertebrate sampling location, we simultaneously measured water quality data (pH, conductivity, and temperature) using handheld Yellow Springs Instrument meters (models 550A and 556). Water samples for dissolved metal analysis were filtered in the field through 0.45-μm cellulose acetate filters, and then acidified with analytical grade nitric acid and stored in acid-rinsed, high-density polyethylene containers with Teflon-lined closures. Field blank filtered samples were prepared using deionized water. At each location, a composite of the finest grain surface sediments visible on the streambed was collected with a scoop from loose depositional areas without algae or organic matter present and placed in amber glass jars with Teflon-lined lids.

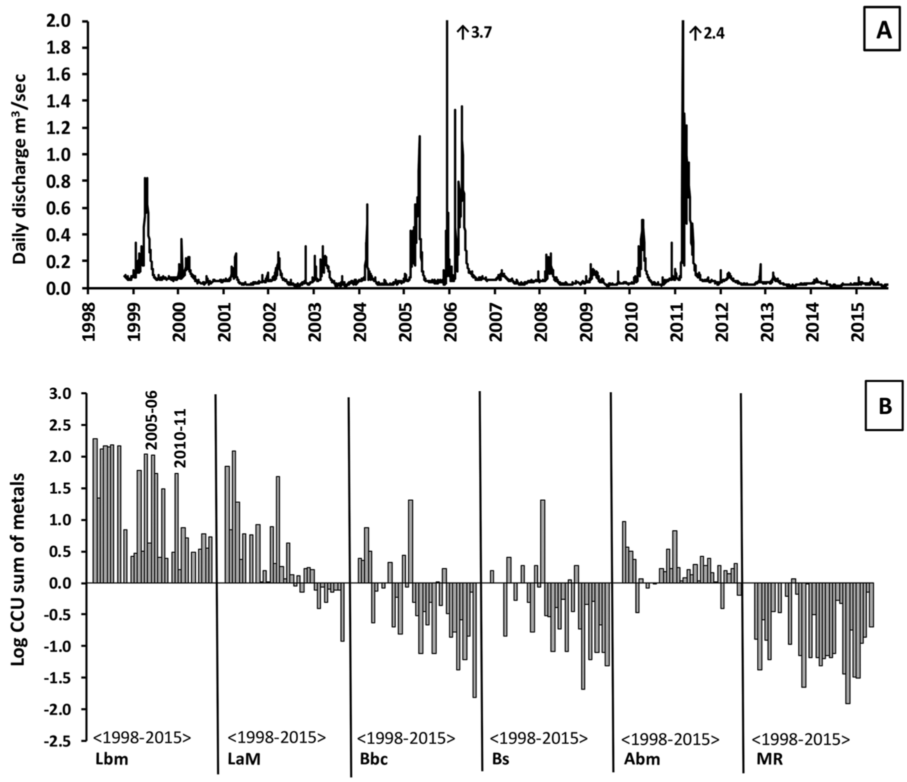

FIGURE 2: (A) Hydrograph of Bryant Creek below Mountaineer over the course of the present study (US Geological Survey gauge #10308794). (B) Trends in log cumulative criterion units (CCUs) total dissolved metals over time by site (codes are explained in Figure 1 caption), with levels initially high: dropping in 2003–2004, rising in 2005–2006, falling to variable in 2007–2009, rising in 2010–2011, falling in 2012–2013, then rising to stable in late drought of 2014–2015. Background of MR (reference site) mostly below log CCU of 0 (CCU = 1). When values fall below 0, nontoxic conditions are implied. Cumulative criterion units for 1999 and 2000 use estimates for selenium (Se), zinc (Zn), or manganese (Mn) based on regressions with related metals or electrical conductivity and/or interpolated trends.

Dissolved metals and cations (calcium [Ca] and magnesium [Mg], for hardness calculations) were analyzed per USEPA (1994) and Campisano et al. (2017) using USEPA Methods 200.7, 200.8, 200.9, and 245.1, and included aluminum (Al), arsenic (As), cadmium (Cd), chromium (Cr), copper (Cu), iron (Fe), lead (Pb), manganese (Mn), mercury (Hg), nickel (Ni), selenium (Se), and Zinc (Zn). Four of these, Cd, Cr, Pb, and Hg, were detected in 5 or fewer samples (of > 200), thus only the other 8 metals were examined in analysis of differences among sites. In 1999 and spring 2000, there were no data for some metals (Mn, Se, and Zn), hence values were estimated by regressions to other metals or interpolated from site trends for that interval. Sulfate in water was analyzed using USEPA Method 300.0. For the sediment samples, a subsample of the finest grain material (<2 mm grain size) was taken in the laboratory and acid-digested for metal analysis using USEPA Method 6010C. Dissolved metal concentrations were converted to cumulative criterion units (CCUs) utilizing the approach described by Clements et al. (2000). The CCU is the sum of concentration values of each metal divided by an associated chronic aquatic life criteria concentration taken from the California Toxic Rule or the USEPA National Recommended Water Quality Criteria. Criteria were adjusted for the hardness of each water sampling date and location (for Cu, Mn, Ni, and Zn) as described for the California Toxic Rule and the USEPA National Recommended Water Quality Criteria values. For Mn, neither the California Toxic Rule nor the USEPA National Recommended Water Quality Criteria includes criteria values; therefore, we used a value of 1000 μg/L in accord with Clements et al. (2000), not adjusted for hardness (per the existing Colorado Department of Public Health standards at the time of their publication). In addition, as indicated by Clements et al. (2000), and to account for some improvements in detection levels over the course of the present study, nondetects were treated as 0 value in the CCUs. The relative load exposure to metals over time and across sites was expressed simply as the CCU sum of the 8 target metals for each sample event (Figure 2B).

Bioassessment sampling

Benthic macroinvertebrate monitoring has been used to define the spatial extent and chronology of biological impacts and recovery in the Leviathan–Bryant Creek watershed from 1998 to 2015 (and is ongoing), with sampling typically conducted both in late spring and early fall of each year (early June and late September). These data have been collected along a downstream gradient of acid mine drainage-affected sites and in adjacent local reference streams of similar size and geology also tributary to the East Carson River but where acid mine drainage does not occur (Figure 1 and Table 1). Mountaineer Creek has served as the primary reference site for biomonitoring throughout the history of this survey program, supplemented by 4 additional sites sampled variously at 2 to 11 dates over the same years of monitoring acid mine drainage-exposed sites. The seasonal sampling times correspond to dates just after treatment operations begin, and just before they end. The time frame also represents changing hydrological conditions during spring runoff and fall base flow, and phenological changes in the development of insect populations. This dataset assesses improvements or relapses in biological health over time as well as metal response thresholds, and defines indicators for the re-establishment of aquatic life to a natural state.

Benthic macroinvertebrate sampling was conducted by collecting from riffle habitats in shallow, rocky stream sections within established survey reaches (100-m lengths). Standardized samples were taken within approximate square-foot areas (30×30 cm) directly upstream of a 30-cm wide, 250-μm mesh, D-frame net held vertically against the streambed. Streambed gravel and rocks within this area were sifted and rubbed by hand for approximately 1 min to dislodge benthic invertebrates that were then carried by water currents into the downstream net or swept by hand if water was too shallow or slow moving. Large wood, leaf, and rock debris were washed and removed from the net, and the sample procedure was repeated at 2 more locations within riffle zones (estimated area of each sample $= 3 \times 0.093$ for $0.28\,m^2$). This composite sample of 3 collections was then repeatedly washed and swirled in a bucket, pouring off lighter, suspended material into a fine mesh (100-μm mesh) collection net to separate invertebrates and small leaf/detritus litter from heavier sand/gravel. The gravel fraction remaining in buckets was then inspected in shallow, white trays; the remaining invertebrates were collected (e.g., cased caddis, snails), and the entire sample was preserved in 95% ethanol. This composite sampling was repeated 5 times at each site (moving upstream in randomly located riffle areas), for a combined total sample area coverage of $1.4\,m^2$. Composite collections were divided into subsample fractions using a rotating-drum splitter to obtain a minimum count of 300 organisms; others were counted in their entirety when holding fewer than 300 organisms. Subsample fractions and areas sampled were used to calculate density. Sample processing and sorting were performed using stereo and compound microscopes, and invertebrates were identified to the lowest verifiable taxonomical level (usually genus or species including midges and mites, except for oligochaetes, turbellarians, and ostracods) and cross-confirmed among laboratory taxonomists and reference specimens.

Data analysis

Compiled data from fall of 1998 to fall of 2015 were available for benthic macroinvertebrate taxa and metrics (densities and diversity), comprising nearly 400 000 specimens in 327 taxa over a series of 34 possible sample dates at acid mine drainage and reference sites. Data from the 5 sample replicates per site were combined to best represent entire study reaches; the density and relative abundance of taxa from each sample site and date were used for analysis of community structure, biotic index of community tolerance (sum product of relative abundance and tolerance value for each taxon), life history traits (voltinism), and functional feeding group composition. Tolerance values and functional feeding group assignments were based on standard listings (Appendix B, Barbour et al. 1999). Richness of total and Ephemeroptera, Plecoptera, and Tricoptera (EPT) taxa were estimated from rarefaction of combined replicate samples at a fixed count of 500, subsampled from larger counts using EcoSim (vegan package in R, Ver 2.4-2; Oksanen et al. 2017). Where sum sample counts were sometimes less than 500 (at Leviathan below mine), taxa richness was based on the entire sample. Recovery of trophic function was evaluated by comparing the significance of density increases across functional feeding groups from the early period of incomplete treatment and capture of acid mine drainage (1998–2006) with that of the later period, when treatments had become more efficient and complete (2007–2015) and coincided with reduced toxic metal loading (Figure 2B).

Nonmetric multidimensional scaling analysis, based on log $(x+1)$-transformed densities, was used to explore the changes in benthic macroinvertebrate community structure over time among sites, using PC-ORD, MjM Software, Ver 6.22 (McCune and Grace 2002). These values were not relativized, allowing data to represent spatial and temporal variations in both density and relative abundance. The data matrix consisted of 209 site-dates and included 168 taxa representing those present in $>5\%$ frequency of samples (rare taxa in fewer than 10 site-dates were removed to reduce noise and enhance the detection of relationships between community composition and environmental factors). Individual and combined metal CCUs, pH, conductivity, sulfate, and sample date were utilized in an environmental matrix to evaluate both when and what metals were related to differences within and among sites in composition of invertebrate taxa present. Dissimilarities among the data grouped by site and according to low, moderate, and high ranges of CCUs observed (<0.5, 0.5–2, and >2) were tested using multiple-response permutation procedures. Indicator species analysis (4999 Monte Carlo randomization runs) was also performed for taxa associated with these CCU ranges in the core group of 6 sites in the Leviathan/Bryant Creek watershed. From this we derived indicator values as the product of relative abundance and relative frequency of each taxon for significant associations with the low, moderate and high CCU ranges.

Ecological effect concentration ranges for metal CCUs were determined based on several approaches using 10 biological response metrics representing community structure, density, richness, and function (nonmetric multidimensional scaling axis 1 ordination scores, log total and EPT density, total and EPT richness, biotic index, log density of grazers, large predators, and semi-voltine taxa, and logit percentage of Chironomidae). Setting acceptable conditions at $>$ the 10th percentile of reference observations for each metric (consistent with California stream assessment standards; Mazor et al. 2016), the 90th percentile of the cumulative distribution of stream CCU measurements for all sites meeting this reference criterion was used to define a conservative estimate of the CCU threshold for effect level of impaired condition. In addition to this reference stream approach to defining effect levels, segmented regression was used to explore the range of variation in metric response. This technique has been applied in other evaluations of metal impact on benthic macroinvertebrate communities to provide another estimate of ecological thresholds (using R code; Sonderegger et al. 2009). Finally, nonparametric change point analysis was used on metric distributions partitioned into successive intervals (log CCU $<$ or >-1, –0.5, –0.25, 0, 0.25, 0.5, 1, 2, and >2) to find the minimum within-group deviations relative to overall variance (Qian et al. 2003).

RESULTS

Dissolved metal water chemistry

Remediation treatment activities have produced considerable declines in the estimated loading of toxic metals to Leviathan Creek (Figure 2B). After the onset of pond controls of overland runoff early in the present study, metals decreased progressively, and then increased again during periods of escalated runoff in 2005 to 2006 and 2010 to 2011. During the years between high flows, and in the recent drought, combined metal CCUs have continued to decline across all acid mine drainage-exposed sites.

Dissolved metals averaged over time at each site distinguished Aspen Creek from Leviathan and Bryant Creek stations by higher concentrations of As, Se, and Cu (Table 1). Leviathan and Bryant Creek sites were dominated by Al and Fe toxicity. Yellow-boy iron oxide precipitates were common in Leviathan and Bryant Creeks but nearly absent from Aspen Creek. These differences in chemistry were likely caused by practices used during mining when overburden (earth removed to access the ore body) was placed in the Aspen watershed, whereas tailings (low-grade ore body) were placed in Leviathan Creek. Downstream Leviathan and Bryant Creek sites showed diminishing CCU levels over time, consistently falling below CCU $=1$ (log CCU $=0$) after 2007 at Bryant below confluence and Bryant state line, and after 2011 at Leviathan above Mountaineer. The metal concentration exposure has not been uniformly distributed over the year, with poor water quality during the period of nontreatment (spring) alternating with improved conditions of pH, sulfate, and conductivity after treatment operations (fall) during average and wet years but not in dry years. Excluding Leviathan below mine and Aspen below mine, where metals have remained high (above CCU $=1$) and after 2000 when treatments were underway, seasonal patterns were evident, with spring metals higher than those in fall in 24 of 25 cases during average and wet years but with the opposite true in dry years of 2007 and 2012 through 2014 in 10 of 12 cases. An exception to this occurred in the severe drought of 2015 when spring water samples were taken during a storm spate that were consistent with high spring flows, producing elevated metals as in years with average or high flows. Continued elevated CCUs near the mine at Leviathan below mine came mostly from Al, and at Aspen below mine were the result of persistence of Cu, Se, and increasing As. Control of Al, Fe, Ni, and Mn over time in lower Leviathan and Bryant Creeks contributed to most of the reduction in overall CCUs at these sites. Selenium from Aspen continues to contaminate Leviathan above Mountaineer. Although the dissolved metal changes agreed closely with the history of treatments, flow, and spatial array of sites, there was no correlation between the sediment fractions of metals and dissolved concentrations (R^2 all less than 0.03, except Fe = 0.12).

Benthic macroinvertebrate trends and relation to metals

Consistent with the pattern of metals by site and over time, invertebrate community diversity increased over time and with distance downstream. Total and EPT richness of Bryant Creek sites have achieved reference condition over time but sites nearest the mine (Leviathan below mine and Aspen below mine) continue to lag (Figure 3A,B). The Leviathan above Mountaineer location is instructive of remediation efforts as an index site that integrates acid mine drainage from both Leviathan and Aspen sources, yet is undiluted by flows from uncontaminated Mountaineer Creek. In the present study, the progress of recovery can be tracked relative to a reference standard for acceptable conditions. Ephemeroptera, Plecoptera, and Tricoptera richness give a more stringent appraisal of recovery than total richness, taking longer to reach near-reference levels. Total and EPT richness usually showed a pattern of recovery from low values in spring to higher values in fall when metals had been contained (2001–2004 and 2007–2015), followed by reversal the following spring back to lower diversity. During the extended

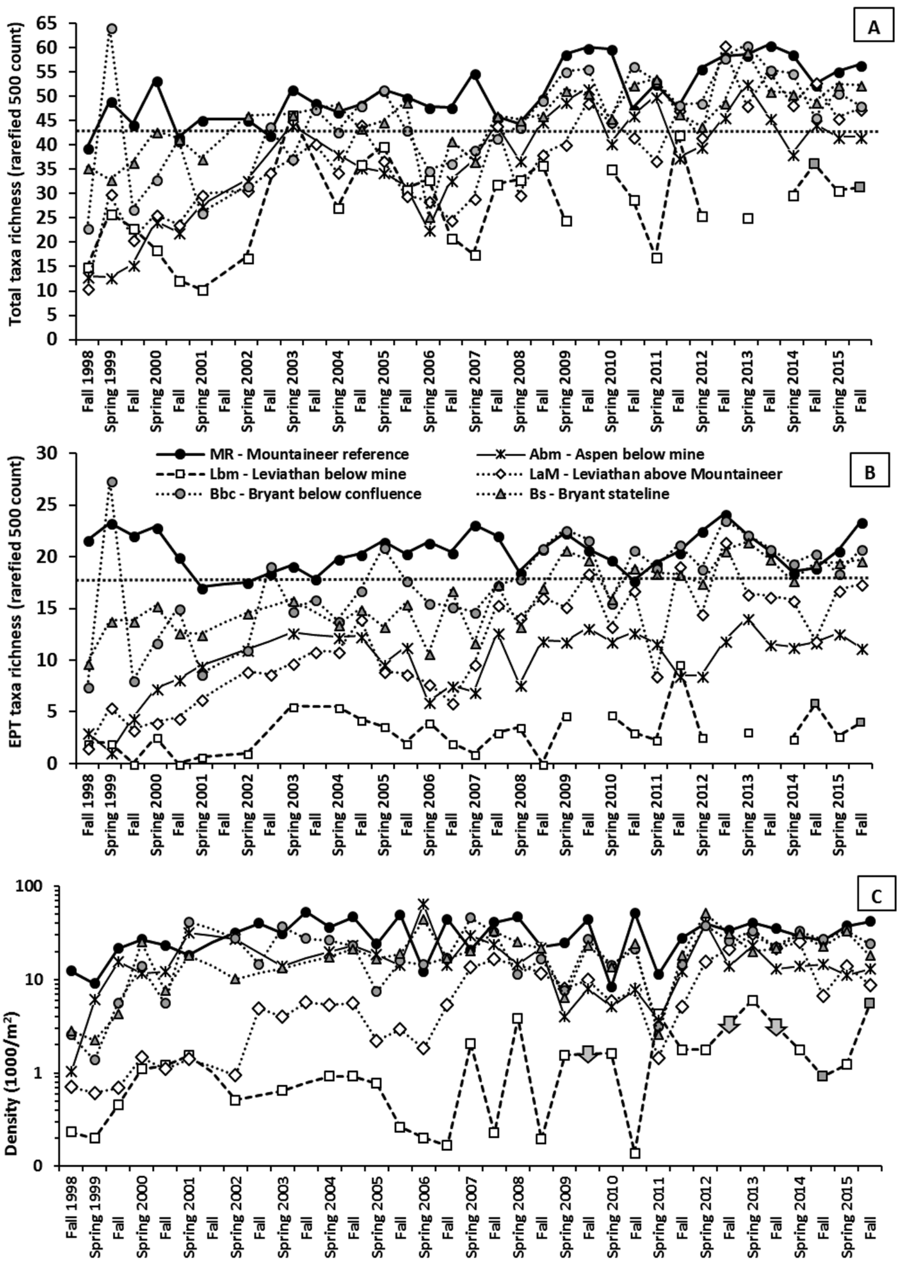

FIGURE 3: (A) Total rarefied richness of benthic macroinvertebrates over time (based on rarefaction to equalize sample counts to $N = 500$). (B) Ephemeroptera, Plecoptera, and Tricoptera (EPT) richness over time (from rarefaction to $N = 500$). Both graphs show the 10th percentile of reference richness for the time-averaged group of reference sites, including Mountaineer Creek as the standard for recovery. (C) Mean density. Codes for all sites are in (B), except gray squares are the substitution of Leviathan 4L for Leviathan below mine (Lbm) when this site was dry in fall of 2014 and 2015 (arrows in (C) show other dry periods at Lbm).

drought of 2012 to 2015, with more complete containment of metal load, richness levels at Leviathan above Mountaineer remained near reference conditions. Late in the drought when Leviathan below mine became dry by fall, data from Leviathan 4L where flows persisted were substituted (Figure 3, gray squares in fall of 2014 and 2015). This was verified as a reliable proxy site with data from spring of those years showing near-identical nonmetric multidimensional scaling community ordination.

Density has typically been lower at acid mine drainage sites than at the Mountaineer Reference (which averaged $> 30\,000\,m^{-2}$), indicating less benthic macroinvertebrate production (Figure 3C). Bryant Creek sites have recovered to the range found in the Mountaineer Reference but upstream sites have remained more sparsely populated. Whereas Leviathan below mine has often had less than $1000\,m^{-2}$, numbers even there have recently increased. The Mountaineer Reference density usually alternated between seasons, lower in the spring and higher in the fall, suggesting a natural pattern of recruitment and growth during summers. Whereas this was evident in years with average-to-high flow, this was not the case during drought years 2012 to 2015. Seasonal increase in density also occurred at times in Bryant Creek sites that improved with acid mine drainage control but again lost or even reversed this pattern during drought. Other acid mine drainage-exposed sites always had densities below those of reference sites.

Trophic structure, compared between the early and late periods of monitoring, showed response to improved treatment as increased densities over time of at least some groups at all sites (Table 2). Only micropredators escalated at Aspen below mine; however, these and shredders and collector–filterers also grew at Leviathan below mine, although still far below Mountaineer Reference densities. All groups increased at Leviathan above Mountaineer, with shredders approaching reference densities. Both Bryant sites showed enhancement in density of grazers, as well as all predators (and filterers at Bryant below confluence); other trophic groups were already abundant even during early years and had food web structure similar to the Mountaineer Reference.

Nonmetric multidimensional scaling ordinations of benthic macroinvertebrate communities show close clustering of reference site samples and more varied distribution of acid mine drainage-exposed sites as they have gone from impaired to a community composition comparable to reference condition (Figure 4A,B). Environmental chemistry vectors show early Leviathan/Bryant samples most dissimilar from references with positive nonmetric multidimensional scaling 1 values related to increased total CCUs, Zn, Fe, Al, Ni, and Mn metals; conductivity and sulfate; and negative values correlated with pH (Figure 4A). Aspen samples were distinguished by Se, As, and Cu; and negative nonmetric multidimensional scaling 1 values were associated with date and progressive recovery toward reference condition. Changes in the nonmetric multidimensional scaling 1 coordinates for exposed sites over time relative to the reference centroid coordinates further show recovery trends from the influence of acid mine drainage (Figure 4B). As with richness, the periods of failed acid mine drainage treatment are evident in departures from improvement in 2005 to 2006 and spring 2011, and later by rapid recovery in fall 2011. Based on the variability between the 10th through 90th percentiles of reference range of nonmetric multidimensional scaling axes, a distance within 0.5 units is indicative of recovery. Whereas nonmetric multidimensional scaling 1 showed metal influence, nonmetric

TABLE 2: Recovery or change in trophic function as density increases across sites from early years of incomplete acid mine drainage treatment (1998–2006) to later years of more complete treatment and reduced toxic metal concentrations (2007–2015)

Site code	No.	Shredders	Grazers	Medium–large predators	Micropredators	Collector– filterers	Collector– gatherers
Abm——Early	14	2997 (1014)	426 (156)	312 (60)	518 (139)	4393 (2173)	10 944 (1832)
Late	18	2311 (632)	377 (66)	403 (55)	2016 (296)	1963 (513)	8098 (1386)
		NS	NS	NS	***	NS	NS
Lbm——Early	14	11 (3)	8 (5)	23 (6)	31 (6)	6 (3)	572 (117)
Late	13	134 (45)	11 (3)	83 (33)	114 (28)	398 (200)	1293 (357)
		***	NS	NS	**	**	NS
LaM——Early	16	367 (121)	192 (76)	72 (13)	190 (49)	430 (176)	1547 (287)
Late	18	2680 (493)	1411 (304)	625 (106)	972 (202)	1400 (266)	6149 (1244)
		***	***	***	***	**	***
Bbc——Early	16	4684 (1206)	1657 (330)	451 (65)	645 (151)	1599 (560)	8566 (1607)
Late	18	4391 (974)	3121 (538)	787 (93)	2038 (359)	2671 (479)	10 865 (1640)
		NS	*	*	**	*	NS
Bs——Early	14	2948 (629)	1052 (221)	373 (50)	594 (111)	1373 (526)	9280 (2502)
Late	18	3620 (726)	3130 (509)	942 (142)	1410 (177)	2415 (464)	11 537 (1534)
		NS	**	**	**	NS	NS
MR——Early	16	2830 (538)	14 814 (2124)	795 (68)	1357 (231)	2668 (448)	7634 (876)
Late	18	3336 (460)	8741 (1140)	1123 (112)	3124 (404)	2294 (322)	13 743 (1699)
		NS	NS	NS	**	NS	**

Each contrast shows the number of observations, the mean values for each period (± SE), and the statistical significance. The Benjamini–Hochberg procedure is used with corrections for multiple comparisons and a false detection rate set at a probability value of 0.05.
* Significant at $p < 0.05$.
** Significant at $p < 0.01$.
*** Significant at $p < 0.001$.
Site codes are the same as in Table 1.
SE = standard error; NS = not significant.

multidimensional scaling axis 2 showed clear separation by season, with fall positive and spring negative in general (Figure 4C). The Mountaineer Reference and recovering Bryant state line sites showed least seasonal separation; the more acid mine drainage-affected sites were increasingly far apart, consistent with spring to fall difference in metal concentration exposure.

Indicator species analysis using taxa that were most common and abundant (indicator values > 20) showed that a large group of benthic macroinvertebrates were significantly associated with

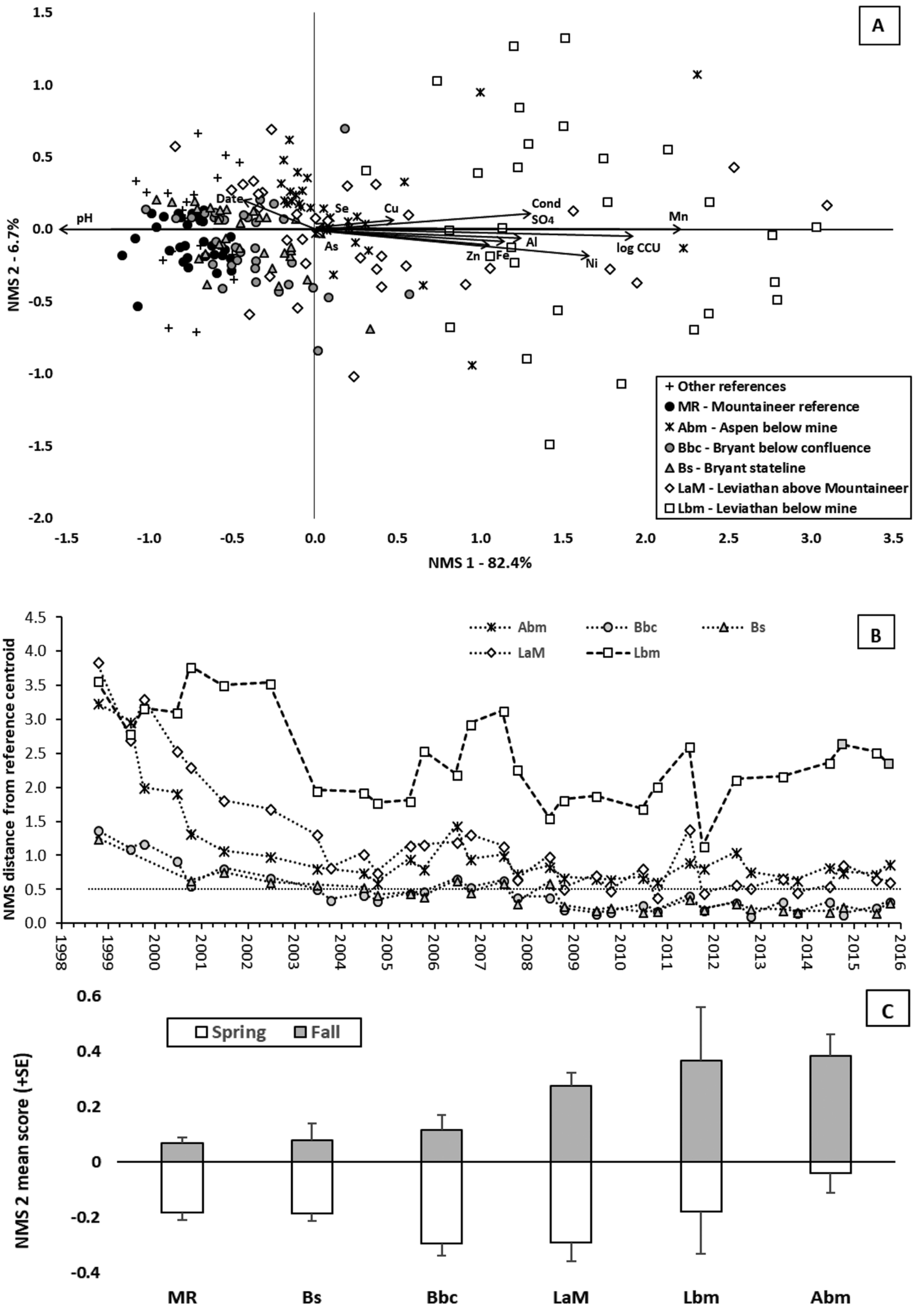

FIGURE 4: (A) Nonmetric multidimensional scaling (NMS) plot by sample and site from 1998–2015. (B) Distance from reference centroid. (C) Mean seasonal differences by site on nonmetric multidimensional scaling axis 2 and + standard error (SE). In (A) final stress = 14.05 for 2-dimensional solution; axis 1 explains 82.4% of variation; axis 2 explains 6.7% of variation; length of chemical and date vectors is proportional to R^2 values (highest = manganese [Mn] 0.73; log cumulative criterion units [CCUs] 0.65; pH 0.62). In (B) years are divided into seasonal quarters. Leviathan 4L substituted in fall 2014 and 2015 gray squares for Leviathan below mine (Lbm). Note that at Leviathan above Mountaineer (LaM) there is often a pattern of being farther from reference in spring and closer each fall, except during drought years. This coincides with metals being higher in spring and lower in fall of wet years. Se = selenium; Cu = copper; SO_4 = sulfate; Al = aluminum; Zn = zinc; Fe = iron; Ni = nickel; As = arsenic.

CCU values of less than 0.5, composed mostly of reference samples (Table 3). Ephemeroptera, Plecoptera, and Tricoptera comprised 27 of 70 (39%) taxa significantly associated with this group, along with many midges (21), elmid beetles (6), and mites (6). Taxa significantly associated with an intermediate transitional CCU range of 0.5 to 2 included ostracods, ceratopogonids, the midge *Parametriocnemus*, and the stonefly *Malenka*. At the high range of CCUs > 2, indicator taxa were all midges (6), except for the ceratopogonid *Monohelea* and the tipulid *Molophilus*.

Metal toxicity effect levels based on CCU values above the 90th percentile of samples that conformed to the reference condition for each metric, representing those with high risk of impairment, showed a metric range of CCU = 0.9 to 3.1 for grazer density to percentage of chironomids (log CCU of ~0–0.5 in Figure 5). Evaluation of the range of metric thresholds based on segmented regression (linear piecewise) showed lower CCU effect levels (0.1–0.6) than earlier, possibly related to the sensitivity of this method to more subtle change where linear slope first changes significantly to negative or positive (Table 4). Nonparametric change point analysis indicated that the threshold for transition between intact and impaired condition occurs in the range 1.3 to 2.5 CCUs, within the reference-defined assessment of thresholds. Using the metric means for these 3 approaches in order of decreasing CCU effect level suggests that recovery follows a progression: tolerant midges are reduced, densities increase, long-lived and sensitive taxa increase, overall diversity and community structure recover, followed by EPT richness and return of large predator and grazer trophic groups (Table 4).

The densities of selected prominent taxa were also evaluated as a measure of tolerance over the range of CCUs found in the core set of Leviathan watershed sites (Figure 6). Taxa with the greatest sensitivity to metals include mayflies of the family Heptageniidae and *Paraleptophlebia*, the caddisfly *Ceratopsyche*, and the large perlid stonefly predator *Doroneuria*. *Baetis* mayflies and the nemourid stoneflies *Malenka* and *Zapada* had intermediate tolerance of metals, whereas ceratopogonid biting midges and the midge *Eukiefferiella claripennis* grp. tended to proliferate at higher CCU levels.

Although the deposition of yellow-boy was often observed in poor water quality conditions and can cover and cement benthic habitat, compounding impacts on benthic macroinvertebrates, no relationship was found between biological indicators and metals in sediment deposits.

DISCUSSION

Multiple lines of evidence from this long-term monitoring of metals and stream macroinvertebrates show that capture of acid mine drainage and release of treated effluent are improving water quality and ecological integrity of streams below the abandoned pit of Leviathan Mine. Fall improvements after the summer season of treatment and later spring relapse after the overwinter period of uncontrolled acid mine drainage discharge suggest, however, that year-round treatment is needed to achieve lasting recovery. Sites nearest the mine have been slow to recover and thus remain degraded because of persistent high metal concentrations. Farther downstream, measures of taxa richness and community structure indicate that lower Leviathan Creek (Leviathan above Mountaineer) progressed with regularity into the range of reference conditions in the September samples but reversed again by the June sampling. This corresponds to the absence of neutralization over winter months and the release of treated effluent over the summer, as well as to natural spring–fall differences in discharge carrying toxic metal content proportional to flow (Clements et al. 2010). Water chemistry data support this in showing elevated metals and lower pH in Leviathan Creek when acid mine drainage capture and treatment are not operating; concentrations are higher in spring than in fall during average and wet years of runoff but not drought years. Early and peak runoff periods in snowmelt to the Arkansas River have also been shown to carry the highest concentrations of Zn, Cd, Cu, and Pb (Lewis and Clark 1997; Clements et al. 2010). Because dilute snowmelt runoff waters carry little Ca/Mg hardness, there is hardly any buffering of metal toxicity. This high flux of metals during spring pulse must be controlled so that the toxic effects are reduced or eliminated. Sensitivity and impact to insects may also be greater in spring when there is a greater proportion of smaller early-life stages, whereas by fall more tolerant, mature larvae and nymphs are present (Kiffney and Clements 1996). High flows have also been shown to be associated with an increase in tissue body burden of metals (Hornberger et al. 2009). At stations below the inflow of Mountaineer Creek Reference, Bryant Creek monitoring revealed earlier and prolonged recovery of the benthic invertebrate biota to conditions of similar density, community composition, trophic structure, and diversity as the Mountaineer Reference (Figures 4, 5, and 6). Spatial location may constrain the recovery of streams from acid mine drainage because of the limited availability of drifting colonists, especially in Leviathan Creek where there is little or no source area of uncontaminated water for invertebrates to originate (Kitto et al. 2015). Recovery of Bryant Creek sites benefits from the proximity to Mountaineer Creek as a recolonization source, whereas drift colonization of sites upstream of this confluence must run the gauntlet of acid mine drainage-contaminated channels. Leviathan and Aspen Creek sites may be slower to recover because of greater dependence on aerial colonization and upstream migration. Despite these limitations, contrast of the early period of incomplete treatments with later more effective remedial practices showed that densities of some or all functional feeding groups increased across sites, indicating progress in the recovery of food-web function.

Although advances have been made in biological recovery of streams in the Leviathan mine drainage, impediments to full recovery remain. High conductivity and elevated sulfate levels remaining in treated effluent are an uncontrolled source of chemical stress that may also prevent full recovery of Leviathan streams (Table 1). Sulfate along with other major anions have been shown to reduce mayfly abundance in mesocosm experiments (Clements and Kotalik 2016). Bound metals in yellow-boy deposits (in Leviathan Creek) could also be a chronic source of pollutants if they become dissolved and released. DeNicola and

TABLE 3: Indicator species analysis for cumulative criterion unit (CCU) groups ranked by indicator value[a]

<0.5 CCUs	0.5–2.0 CCUs	>2.0 CCUs
Epeorus——62.7***	Ostracoda——43.1***	*Pseudorthocladius*——44.5***
Ironodes——62.5***	*Bezzia-Palpomyia*——41.7***	*Monohelea*——37.0***
Micrasema——61.5***	*Malenka*——36.7**	*Eukiefferiella claripennis*——33.6*
Doroneuria baumanni——59.4***	*Parametriocnemus*——34.2**	*Limnophyes*——32.9**
Paraleptophlebia——57.2***	*Culicoides*——27.4**	*Apedilum*——28.8***
Testudacarus——54.9***	*Ceratopogon*——25.2*	*Molophilus*——28.4***
Rheotanytarsus——54.5***	*Dixa*——24.5***	*Polypedilum laetum*——27.5***
Diphetor hageni——54.2***	*Isoperla*——24.0**	*Pseudosmittia*——21.3**
Aturus——52.2***	*Cheiroseius*——23.1***	
Cinygmula——51.4***		
Feltria——50.3***		
Turbellaria——49.0***		
Lepidostoma——48.9***		
Rhyacophila brunnea——48.7***		
Serratella——48.5***		
Rhyacophila betteni——48.4***		
Ceratopsyche——47.0***		
Ameletus——46.9***		
Optioservus quadrimaculatus——46.7***		
Optioservus divergens——46.5***		
Tvetenia bavarica——45.3***		
Lebertia——43.9***		
Sweltsa——43.3***		
Baetis——43.1***		
Eukiefferiella brehmi——42.8***		
Eubrianax edwardsi——42.6***		
Atractides——42.6***		
Eukiefferiella gracei——42.5***		
Dicranota——41.9***		
Capniidae mixed spp.——41.7***		
Yoraperla——40.7***		
Tanytarsus——40.4***		
Sperchon——39.7***		
Simulium——39.7***		
Neoplasta——39.6***		
Micropsectra——39.3***		
Eukiefferiella devonica——39.1***		
Lara avara——38.8***		
Corynoneura——38.6***		
Pericoma——38.4***		
Orthocladius complex——38.4***		
Stempellinella——37.1***		
Heleniella——37.0***		
Thienemannimyia——36.6***		
Pisidium——36.5***		
Pagastia——36.5**		
Rheocricotopus——36.0***		
Diamesa——35.7***		
Zaitzevia——35.1***		
Larsia——32.1***		
Brillia——32.1**		
Narpus concolor——31.9***		
Cleptelmis addenda——30.8***		
Glossosoma——30.0***		
Virgatanytarsus——27.9***		
Protzia——27.4***		
Neophylax——26.4***		
Tabanus——24.3***		
Pentaneura——24.2*		
Brachycentrus americanus——23.8***		

[a] Within-group taxon abundance and frequency for values > 20.
* Significant at $p < 0.05$.
** Significant at $p < 0.01$.
*** Significant at $p < 0.001$.

Stapleton (2002) concluded that aqueous exposure may be more toxic than yellow-boy deposits; nevertheless, Macintosh and Griffiths (2015) showed clear inhibition associated with iron hydroxide deposits alone. Enduring resolution of open-pit source contamination may come only through re-establishing soil and vegetation capacity to retain and capture water and minerals and control erosion (Wong 2003).

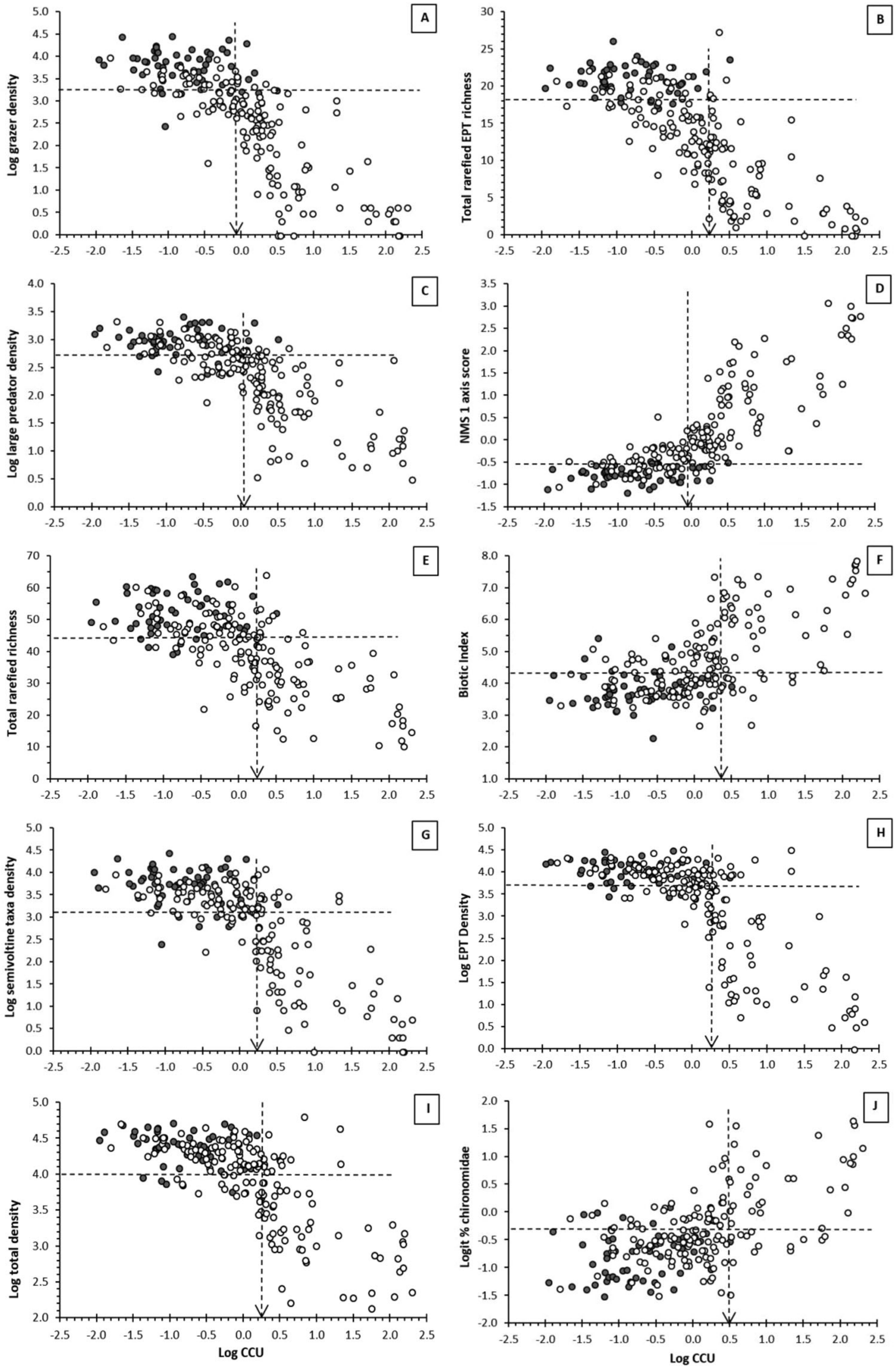

FIGURE 5: Biological metric responses to *x* axis log cumulative criterion units (CCUs = 1, at 0) in reference and acid mine drainage-exposed streams. The CCU effect level is shown as the 90th percentile (vertical dashed line with arrow) of all sites meeting the reference condition for that metric (defined as > 10th percentile of reference values, horizontal dashed line). (A) Log grazer density. (B) Total rarefied Ephemeroptera, Plecoptera, and Tricoptera (EPT) richness. (C) Log large predator density. (D) Nonmetric multidimensional scaling (NMS) axis 1 score. (E) Total rarefied richness. (F) Biotic index (sum of product relative abundances and tolerance values by taxon). (G) Log semi-voltine taxa density. (H) Log EPT density. (I) Log total density. (J) Logit % Chironomidae.

Monitoring trends indicate that the timing of recovery of stream biota coincides with levels of cumulative metal toxicity. Increases in overall density and a decrease in the proportion of tolerant midges occurred in the early stages of recovery at CCUs = 2 to 3 (Figures 3, 4, and 5; Table 4). This was followed by a recovery of long-lived taxa, increased proportion of sensitive relative to tolerant taxa, and an overall community composition and diversity that approached that of the reference at CCUs = 1 to 2. Finally, large predators, grazers, and EPT richness reached reference levels at approximately CCU = 1. The variable range over which the threshold effect of metals was observed is in part caused by differing metric responses and also by the varied mixtures of metals over the sites, and times of sampling and differences in flow regime and season among samples. In addition, the methods used to identify thresholds employed differing assumptions to infer change. The traditional regulatory approach using the reference distribution was based on local reference streams of similar type and biogeographic potential; change-point analysis showed the CCU value where the variance of the divided distribution was minimized within groups; and segmented regression used initial slope change to indicate where metals start to degrade the benthic invertebrate community. Expressing the threshold as a mean range for metrics from CCUs of 0.8 to 2.0 (median 1.3; Table 4) therefore seems most justifiable. This is similar to findings reported from streams in the mining belt of the Colorado Rocky Mountains based on the chronic criterion accumulation ratio (a refinement of the CCU concept) and also showed metrics bracketing a chronic criterion accumulation ratio of 1.0 (Schmidt et al. 2010).

In experimental studies of benthic communities from the Arkansas River in Colorado, Clements (2004) suggested that CCUs < 5.0 would be protective of biological communities here and at other streams of the region where Zn is the predominant heavy metal in mining discharges. In contrast, our long-term, multi-site field data from the Leviathan mine watershed suggest that a CCU near 1 is more completely protective of biological integrity; and although some community components may be in recovery at CCUs of 2 to 3, early signs of detrimental response may occur at CCU levels below 1. This greater apparent sensitivity may stem from several factors. The metals at Leviathan may be more toxic and include a more complex potent mix than the Colorado streams, with predominance of Al, Se, Fe, Ni, As, and Mn making greater contributions to toxicity than Zn (Table 1). In addition, small streams at higher elevations may have a more sensitive fauna because of greater metal toxicity in species having either smaller size (slower growth, earlier phenology), less adaptive variability present in populations in more constant environments, or resident communities comprised of a greater proportion of intolerant taxa (Kiffney and Clements 1996). These high elevation streams also have less acid neutralizing capacity and lower hardness levels that can offset metal toxicity. Setting metals limits based on short-term bioassays could also underestimate sublethal effects on population dynamics. Alleviating acid mine drainage effects on small streams such as those in Leviathan may therefore require standards to be more protective from metal toxicity, and the evidence presented in the present study is consistent with use of a standard CCU = 1. This supports the use of chronic exposure bioassay-based responses of test organisms (chronic criteria concentrations), indicating that even sublethal effects exhibited in community structure changes and altered species distributions occur within a range that encompasses this CCU standard.

Dissolved metal concentrations are likely to be correlated with metal in biofilms on rock surfaces where sensitive invertebrate grazers such as heptageniid mayflies feed, resulting in impaired growth and reduced abundance (Courtney and Clements 2002). Dietary exposure, particularly through periphyton, appears to be much more important as a direct source of body burden than what comes from dissolved metal (Poteat and Buchwalter 2014). Metal in sediment deposits, where samples were collected in the present study, did not bear any significant relation to dissolved metal or benthic macroinvertebrate metrics. Without an association with biofilms or periphyton, these sediment samples may simply have had no relation to dietary resources. Dissolved metal differed between Leviathan and Bryant versus Aspen Creeks, and CCUs showed that whereas Al was an important contributor to toxicity in all streams, Fe, Mn, and Ni were substantial in Leviathan (Zn less so)

TABLE 4: Threshold analyses of ranked mean cumulative criterion unit (CCU) effect level for selected metrics

Overall metric order for combined rank thresholds[b]	CCU values for effect thresholds by metric and analysis methods[a]			
	Segmented regression	Nonparametric change-point analysis	90th Percentile meeting reference	Mean CCU effect level
Grazer density	0.26	1.33	0.88	0.82
Rarefied EPT richness	0.12	1.33	1.06	0.84
Large predator density	0.26	1.78	1.13	1.06
NMS axis 1 score	0.55	2.47	0.93	1.28
Rarefied total richness	0.43	1.78	1.77	1.32
Biotic index	0.44	1.78	2.36	1.53
Semi-voltine density	0.60	2.47	1.69	1.55
EPT density	0.61	2.47	1.86	1.64
Total density	0.60	2.47	1.87	1.64
Percent Chironomidae	0.32	2.47	3.09	1.96

[a] Median value of all cumulative criterion units among all metrics and methods of threshold analysis = 1.33.
[b] Overall range = 0.1–3.1.
EPT = Ephemeroptera, Plecoptera, and Tricoptera; NMS = nonmetric multidimensional scaling.

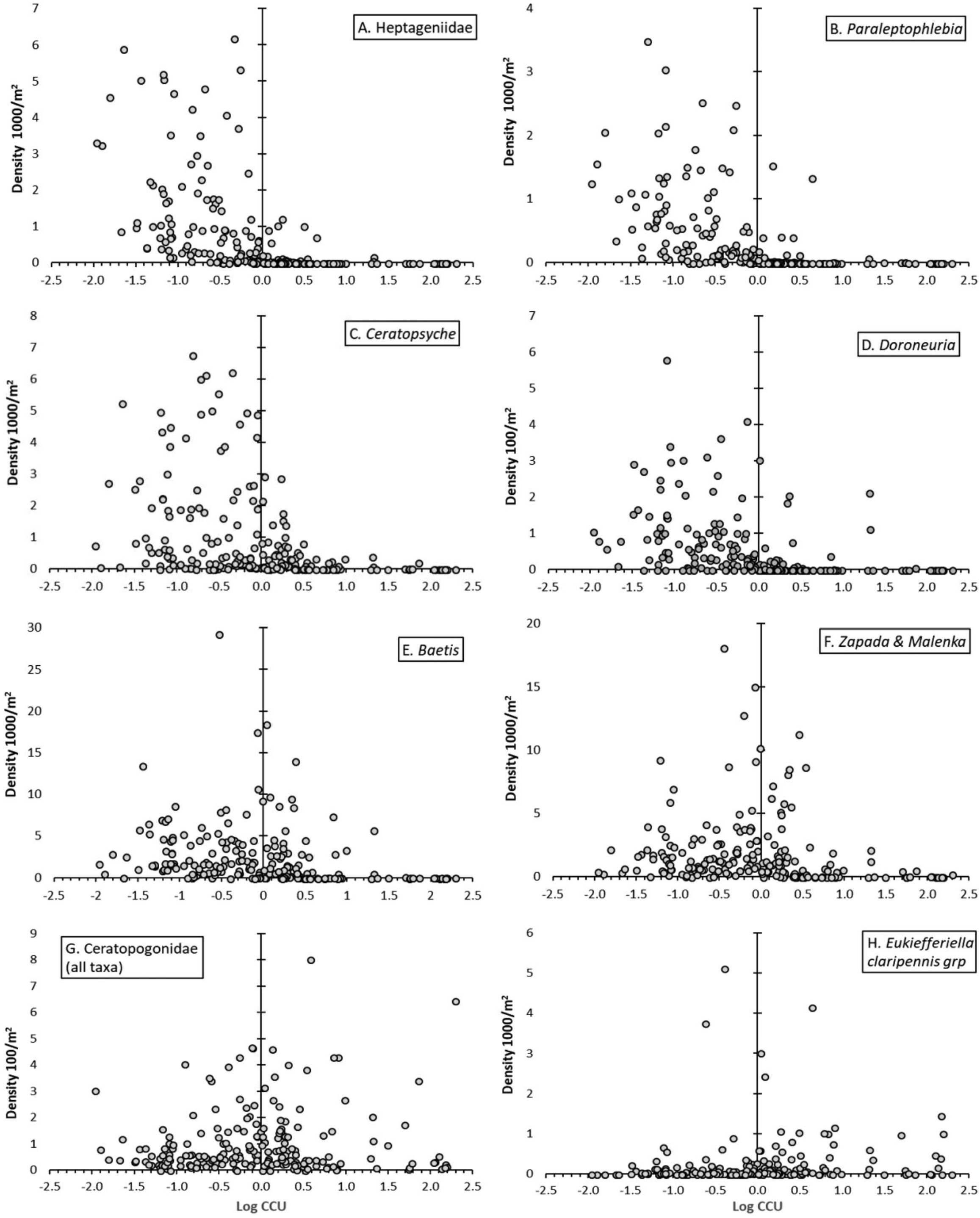

FIGURE 6: Selected taxa log abundance relation to log cumulative criterion units (CCUs) metal (*x* axis) variation over time among reference and acid mine drainage-exposed streams. (A) Heptageniidae. (B) *Paraleptophlebia*. (C) *Ceratopsyche*. (D) *Doroneuria*. (E) *Baetis*. (F) *Zapada* and *Malenka*. (G) Ceratopogonidae. (H) *Eukiefferiella claripennis* grp.

but Se, Cu, and As seemed to originate mostly in Aspen. Dissolved Mn and Fe are unlikely to produce direct toxicity to invertebrates but may indicate the presence of metal hydroxide deposits on body surfaces or substrata present in Leviathan Creek sites. How these mixed metals affect benthic food resource quality and availability operating through periphyton and detrital pathways, may provide further insight to the exposures and concentrations at which invertebrate communities are impaired.

The decreased abundance and diversity of benthic macroinvertebrates in acid mine drainage-affected streams is a well-known phenomenon (reviewed by Hogsden and Harding 2012a); however, there are few examples of how biological recovery proceeds over time with treatment of effluent, and with natural seasonal and inter-annual environmental variation (Clements et al. 2010; Mebane et al. 2015). In this regard, the Leviathan dataset provides an important case history in establishing the success of acid mine drainage remediation activities. The use of biomonitoring as an indicator of ecological toxicity and mining-related pollution impacts and improvements has been substantiated through studies that show close correlation of bioassessment metrics with the standard bioassays using specific test organisms, and with dissolved metal contaminant concentrations (Schmidt et al. 2002; Griffith et al. 2004). Studies combining controlled exposure of natural assemblages to metals and field manipulations of metal oxides on substrates have been used to detail recovery potential of different taxonomic groups (Cadmus et al. 2016). The well-documented trends of benthic macroinvertebrate community responses to improving water quality on the Arkansas River offer a strikingly similar example to Leviathan Creek of seasonal and flow-related effects, as well as downstream spatial variation in restoration success (Clements et al. 2010). Other field studies on streams in the mining district of the upper Arkansas River in Colorado showed that within 2 yr after water treatment that removed metals from contaminated inflows, EPT taxa increased and bioassessment metrics achieved upstream reference condition (Nelson and Roline 1996). Similar treatments on the Clark Fork in Montana required much longer periods for aquatic invertebrate recovery to occur (Chadwick et al. 1986) but were complicated by flows redistributing metal-contaminated sediments (Hornberger et al. 2009). Bioassessment monitoring of the Leviathan Creek watershed also revealed mixed results, with recovery occurring during periods of effluent control to the stream but reversion to degraded conditions when control of acid mine drainage pollution had partially failed (2005–2006). Although drought can have severe consequences to stream invertebrate communities, improved water quality was enabled by complete and earlier capture and treatment of acid mine drainage sources in years of reduced snowpack (2012–2015), appearing to offset and sustain recovery even in the face of low flows. This was evident in the near-reference taxonomic richness attained at Leviathan above Mountaineer during the drought. Conversely, the poor performance seen in metrics under high flows at AMD sites were not exhibited at Mountaineer reference. This implies that metals released under high runoff rather than scouring flows are the cause of degradation when discharge is high. The lower spring densities (Figure 4C) and seasonal shift in nonmetric multidimensional scaling axis 2 (Figure 6C) at Mountaineer Creek displayed a natural pattern of community turnover, whereas acid mine drainage-exposed sites had consistently lower densities and altered communities except as Bryant Creek sites recovered.

The algae and organic matter food resources of benthic invertebrates may become reduced in streams exposed to acid mine drainage. Growth of most algae on streambed surfaces is severely decreased under lower pH, elevated metal concentrations, and when metal hydroxides such as yellow-boy coat and cover substrata (Niyogi et al. 1999; Verb and Vis 2000). Microbial decomposition of leaf litter and wood that fall into streams is an integral trophic resource in forested watersheds, and the bacteria and fungi that mediate this process may be impaired by acid mine drainage (Niyogi et al. 2002; Schlief 2004). These results show that acid mine drainage may alter ecosystem processes of primary production and decomposition, changing food resource availability and distribution, and forcing food webs into simpler and less productive pathways. These types of changes in organization of Leviathan stream communities indicate that grazers, especially heptageniid mayflies (comprising $\sim$ 25% of grazers, overall), were slow to recover, possibly because of a combination of food limitation or contamination and their physiological susceptibility to low pH and metals (Figure 6A, Tables 2 and 4). The sensitivity and delayed recovery of heptageniids and grazers have previously been noted in monitoring on the Arkansas River (Clements et al. 2010). Predators such as large perlid stoneflies (Figure 6D) also recover more slowly; perhaps this relates to depletion of prey, poor recolonizing ability, and long life history. Rock surfaces where algae scrapers feed and active predators roam are also the microhabitats most affected by deposition of yellow-boy. Shredder taxa in contrast (dominated numerically by the nemourid stoneflies *Malenka* and *Zapada* that are less sensitive to metals, Figure 6F), may still be able to utilize the coarse particulate organic matter found at moderate levels of metal contamination. Reduced uptake of toxic Zn and Cu by *Zapada* has been found to occur at lower pH enabled apparently through competitive binding by H^+ ions and adaptive capacity of populations exposed to higher metal concentrations (D. Cain, US Geological Survey, Menlo Park, CA, USA, personal communication). Physiological resistance to metal stress appears to have a strong phylogenetic basis, with tolerance related to more effective capacity to eliminate metals in caddis and stoneflies than in mayflies (Buchwalter et al. 2008).

The mechanisms of alteration to benthic invertebrate communities by acid mine drainage are likely related to a mixture of factors. Direct mortality caused by high concentrations of toxic metals and low pH, along with exclusion from rock surfaces and interstices by yellow-boy deposits, may be most common where pollution is severe. Mild acidification from neutral pHs of 7 to 5.9 has been shown in experimental treatments of a stream to increase drift of mayflies, midges, and caddisflies; so, even without causing direct mortality, modest acidification can change the composition of stream benthos (Bernard et al. 1990; Clements 2004). Sublethal effects of metals deserve more attention to determine how individual and population growth rates may be diminished, leading to changes in community structure and function at relatively low metal concentrations (as observed using the segmented regression thresholds). Metal concentration was strongly correlated with acidic pH below 7; thus, the most severe biological impacts of these factors are intertwined. At neutral pH 7 and above, however, there was still a marked negative effect on metrics by CCUs ranging above and below a value of 1, suggesting that metal toxicity is the dominant stressor in acid mine drainage.

Mountaineer and other references showed a high degree of stability and constancy in community composition, whereas the acid mine drainage-exposed sites were highly variable——both in the sites that showed improvement toward the references and those that remained impaired. Variations in all metrics as CCUs increased were far greater than in the references and low CCU ranges, suggesting not only degradation of biological integrity but far less stability in community structure and function (Figure 5). This argues that transitions in these stream assemblages are not from one stable state to another but from a stable low-disturbance state to an instable fluctuating and stressed environment with a disrupted biotic structure. Erratic variations are inherent even in the simplified low-diversity communities of streams contaminated by acid mine drainage, with less coherence in taxonomic composition or functional traits as metals increase above a CCU of near 1.

Acknowledgment—We are indebted to a collaborative partnership and the on-site project staff for success in remediation efforts and monitoring. We thank K. Mayer, L. Deschambault, P. Husby, J. Sullivan, C. Lee, and A. Wagner of the US Environmental Protection Agency; T. Suk, D. Carey, and H. Schembri of the State of California's Lahontan Regional Water Quality Control Board; D. McMindes and C. Koger of the US Army Corps of Engineers; and A. Brown of Atlantic Richfield. We thank M. Bogan and S. Roberts for assisting with laboratory work. Comments from our colleagues W. Clements, D. Buchwalter, T. Short, G. Mancini, and G. Reller improved the writing and presentation of the present study. Funding for the present study was provided by the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (known also as Superfund), the US Environmental Protection Agency, the US Army Corps of Engineers, Atlantic Richfield, and the State of California's Lahontan Regional Water Quality Control Board.

Disclaimer—The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Data Accessibility—Data, associated metadata, and calculation tools are publicly accessible via Figshare or from the corresponding author (herbst@lifesci.ucsb.edu).

REFERENCES

Akcil A, Koldas S. 2006. Acid mine drainage (AMD): Causes, treatment and case studies. *J Clean Prod* 14:1139–1145.

Barbour MT, Gerritsen J, Snyder BD, Stribling JB. 1999. *Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates, and Fish*, 2nd ed. EPA 841-B-99-002. US Environmental Protection Agency, Office of Water, Washington, DC.

Barbour MT, Swietlik WF, Jackson SK, Courtemanch DL, Davies SP, Yoder CO. 2000. Measuring the attainment of biological integrity in the USA: A critical element of ecological integrity. *Hydrobiologia* 422–423:453–464.

Bernard DP, Neill WE, Rowe L. 1990. Impact of mild experimental acidification on short-term invertebrate drift in a sensitive British Columbia stream. *Hydrobiologia* 203:63–72.

Buchwalter DB, Cain DJ, Martin CA, Xie L, Luoma SN, Garland T. 2008. Aquatic insect ecophysiological traits reveal phylogenetically based differences in dissolved cadmium susceptibility. *Proc Natl Acad Sci U S A* 24:8321–8326.

Cadmus P, Clements WH, Williamson JL, Ranville JF, Meyer JS, Ginés MJG. 2016. The use of field and mesocosm experiments to quantify effects of physical and chemical stressors in mining-contaminated streams. *Environ Sci Technol* 50:7825–7833.

Campisano R, Hall K, Griggs J, Willison S, Reimer S, Mash H, Magnuson M, Boczek L, Rhodes E. 2017. Selected analytical methods for environmental remediation and recovery. EPA/600/R-17/356. US Environmental Protection Agency, Washington, DC.

Carlisle DM, Clements W. 1999. Sensitivity and variability of metrics used in biological assessments of running waters. *Environ Toxicol Chem* 18:285–291.

Chadwick JW, Canton SP, Dent RL. 1986. Recovery of benthic invertebrate communities in Silver Bow Creek, Montana, following improved metal mine wastewater treatment. *Water Air Soil Pollut* 28:427–438.

Clements WH. 2004. Small-scale experiments support causal relationships between metal contamination and macroinvertebrate community responses. *Ecol Appl* 14:954–967.

Clements WH, Carlisle DM, Lazorchak JM, Johnson PC. 2000. Heavy metals structure benthic communities in Colorado mountain streams. *Ecol Appl* 10:626–638.

Clements WH, Kotalik C. 2016. Effects of major ions on natural benthic communities: An experimental assessment of the US Environmental Protection Agency aquatic life benchmark for conductivity. *Freshwater Sci* 35:126–138.

Clements WH, Vieira NKM, Church SE. 2010. Quantifying restoration success and recovery in a metal-polluted stream: A 17-year assessment of physicochemical and biological responses. *J Appl Ecol* 47:899–910.

Courtney LA, Clements WH. 2002. Assessing the influence of water and substratum quality on benthic macroinvertebrate communities in a metal-polluted stream: An experimental approach. *Freshwater Biol* 47: 1766–1778.

DeNicola DM, Stapleton MG. 2002. Impact of acid mine drainage on benthic communities in streams: The relative roles of substratum vs aqueous effects. *Environ Pollut* 119:303–315.

Gray DP, Harding JS. 2012. Acid mine drainage index (AMDI): A benthic invertebrate biotic index for assessing coal mining impacts in New Zealand streams. *New Zealand J Mar Freshwater Res* 46:335–352.

Griffith MB, Lazorchak JM, Herlihy AT. 2004. Relationships among exceedances of metals criteria, the results of ambient bioassays, and community metrics in mining-impacted streams. *Environ Toxicol Chem* 23:1786–1795.

Hogsden KL, Harding JS. 2012a. Consequences of acid mine drainage for the structure and function of stream benthic communities: A review. *Freshwater Sci* 31:108–120.

Hogsden KL, Harding JS. 2012b. Anthropogenic and natural sources of acidity and metals and their influence on the structure of stream food webs. *Environ Pollut* 162:466–474.

Hornberger MI, Luoma SN, Johnson ML, Holyoak M. 2009. Influence of remediation in a mine-impacted river: Metal trends over large spatial and temporal scales. *Ecol Appl* 19:1522–1535.

Johnson DB. 2003. Chemical and microbiological characteristics of mineral spoils and drainage waters at abandoned coal and metal mines. *Water Air Soil Pollut* 3:47–66.

Johnson DB, Hallberg KB. 2005. Acid mine drainage remediation options: A review. *Sci Total Environ* 338:3–14.

Kiffney PM, Clements WH. 1996. Effects of metals on stream macroinvertebrate assemblages from different altitudes. *Ecol Appl* 6: 472–481.

Kitto JAJ, Gray DP, Greig HS, Niyogi DK, Harding JS. 2015. Meta-community theory and stream restoration: Evidence that spatial position constrains stream invertebrate communities in a mine-impacted landscape. *Restor Ecol* 23:284–291.

Lewis ME, Clark ML. 1997. How does streamflow affect metals in the upper Arkansas River? US Department of the Interior, US Geological Survey. FS-226-96. Report Series. Denver, CO.

Macintosh KA, Griffiths D. 2015. Changes in epilithic biomasses and invertebrate community structure over a deposit metal concentration gradient in upland headwater streams. *Hydrobiologia* 760:159–169.

Mazor RD, Rehn AC, Ode PR, Engeln M, Schiff KC, Stein ED, Gillett DJ, Herbst DB, Hawkins CP. 2016. Bioassessment in complex environments: Designing an index for consistent meaning in different settings. *Freshwater Sci* 35:249–271.

McCune B, Grace JB. 2002. *Analysis of Ecological Communities*, Ver 6.22, PC-ORD. MjM Software, Gleneden Beach, OR, USA.

Mebane CA, Eakins RJ, Fraser BG, Adams WJ. 2015. Recovery of a mining-damaged stream ecosystem. *Elementa*. DOI: 10.12952/journal.elementa.000024.

Nelson SM, Roline RA. 1996. Recovery of a stream macroinvertebrate community from mine drainage disturbance. *Hydrobiologia* 339:73–84.

Niyogi DK, McKnight DM, Lewis WM Jr. 1999. Influences of water and substrate quality for periphyton in a montane stream affected by acid mine drainage. *Limnol Oceanogr* 44:804–809.

Niyogi DK, McKnight DM, Lewis WM Jr. 2002. Fungal communities and biomass in mountain streams affected by mine drainage. *Archiv Hydrobiol* 155:255–271.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E, Wagner H. 2017. *Community Ecology Package*, R's vegan package, Ver 2.4.2. [cited 2017 March 21]. Available from: https://CRAN.R-project.org/package=vegan

Poteat MD, Buchwalter DB. 2014. Four reasons why traditional metal toxicity testing with aquatic insects is irrelevant. *Environ Sci Technol* 48:887–888.

Qian SS, King RS, Richardson CJ. 2003. Two statistical methods for the detection of environmental thresholds. *Ecol Model* 166:87–97.

Schlief J. 2004. Leaf-associated microbial activities in a stream affected by acid mine drainage. *Int Rev Hydrobiol* 89:467–475.

Schmidt TS, Clements WH, Mitchell KA, Church SE, Wanty RB, Fey DL, Verplank PL, San Juan CA. 2010. Development of a new toxic-unit model for the bioassessment of metals in streams. *Environ Toxicol Chem* 29:2432–2442.

Schmidt TS, Soucek DJ, Cherry DS. 2002. Modification of an ecotoxicological rating to bioassess small acid mine drainage-impacted watersheds exclusive of benthic macroinvertebrate analysis. *Environ Toxicol Chem* 21:1091–1097.

Sonderegger DL, Wang H, Clements WH, Noon BR. 2009. Using SiZer to detect thresholds in ecological data. *Front Ecol Environ* 7: 190–195.

Underwood BE, Kruse NA, Bowman JB. 2014. Long-term chemical and biological improvement in an acid mine drainage-impacted watershed. *Environ Monit Assess* 186:7539–7553.

US Environmental Protection Agency. 1994. Method No. 200.9, Revision 2.2: Determination of trace elements by stabilized temperature graphic furnace atomic absorption. Cincinnati, OH.

US Environmental Protection Agency. 2006. Compost-free bioreactor treatment of acid rock drainage, Leviathan Mine, California. EPA/540/R-06/009. Technical Report. National Risk Management Laboratory, Office of Research and Development. Cincinnati, OH.

US Environmental Protection Agency. 2017. *Water Quality Standards Handbook*. Washington, DC. [cited 2017 September 5]. Available from: https://www.epa.gov/wqs-tech/water-quality-standards-handbook

Verb RG, Vis ML. 2000. Comparison of benthic diatom assemblages from streams draining abandoned and reclaimed coal mines and nonimpacted sites. *J North Am Benthol Soc* 19:274–288.

Wong MH. 2003. Ecological restoration of mine degraded soils, with emphasis on metal contaminated soils. *Chemosphere* 50:775–780.

