DISSERTATION

NOT ALL QUESTIONS FIT IN BEAKERS - DIRECT AND INDIRECT TOXIC EFFECTS OF METAL MIXTURES AND THE APPLICATION OF ECOTOXICOLOGICAL EXPERIMENTS TO DERIVE BETTER WATER QUALITY STANDARDS AND PREDICT RECOVERY AFTER ABANDONED MINE RECLAMATION

Submitted by

Pete Cadmus

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2019

Doctoral Committee:

Advisor: William H. Clements

Boris C. Kondratieff
Dana L. Winkelman
James F. Ranville

ABSTRACT

NOT ALL QUESTIONS FIT IN BEAKERS - DIRECT AND INDIRECT TOXIC EFFECTS OF METAL MIXTURES AND THE APPLICATION OF ECOTOXICOLOGICAL EXPERIMENTS TO DERIVE BETTER WATER QUALITY STANDARDS AND PREDICT RECOVERY AFTER ABANDONED MINE RECLAMATION

Aqueous discharges from abandoned metal mines include complex mixtures of physical and chemical stressors. Consequently, identifying mechanisms and causal relationships between acid mine drainage (AMD) and community responses in the field is challenging. In addition to the direct toxicological effects associated with elevated concentrations of metals and reduced pH, mining activities influence aquatic organisms indirectly through physical alterations of habitat, including increased sedimentation, turbidity and substrate embeddedness. Although direct toxicity can sometimes be effectively studied in the laboratory, the indirect toxicity of toxicants rarely manifests into a measurable endpoint in the short duration and limited ecological realism of traditional laboratory toxicity experiments. The installation of a mine effluent treatment plant near Blackhawk Colorado (USA), had potential to remove the majority of aqueous metals from a mountain stream heavily degraded by Acid Mine Drainage (AMD). To investigate direct and indirect effects of Acid Mine Drainage (AMD) a series of field biomonitoring, field experiments, and mesocosm experiments were conducted. These studies quantified the relative importance of chemical (direct) and physical (indirect) stressors associated with AMD discharges and predicted recovery potential for dominant macroinvertebrate taxa.

Ferric Fe is often a dominant toxicant present in AMD but is largely believed to be non-toxic to aquatic life. Results of toxicity tests reported here suggest that the current USEPA chronic Fe criterion is underprotective and that the current criterion should be reduced to 25% of its current level (251 µg/L). These studies demonstrated additional risk to aquatic insects and periphyton in metal mixtures that included ferric Fe. Responses were primarily a result of indirect physical effects associated with Fe oxide deposition rather than direct toxicity. All aquatic insects hatch as nearly microscopic organisms and small size classes were consistently the most sensitive in numerous experiments. Sampling small age classes in nature and conducting toxicity trials with small age classes is difficult and therefore these studies are lacking from the scientific literature. Failure to characterize sensitivity of early size classes may lead to gross overestimation of tolerance. Mesocosm experiments conducted using natural benthic communities provide a unique opportunity to quantify the relative importance of these indirect physical effects.

DEDICATION

To my son, Darwin Cadmus. Thank you for your patience while I tried to finish this degree. I acknowledge that I was not there for you as much as I should have been. I cannot begin to describe how hard it was finding time for family, work, and school. So often I wished to take you camping or swimming or climbing or star gazing. But too often I was behind on number crunching or experiments or writing. I hope to be a better dad in the years to come, to show you the mountains, streams and oceans that inspire me. I hope to show you the glaciers in Alaska and Chile before they melt, because all the local ones are gone. I hope you get to see rivers that are not dammed and diverted to feed cows, because the local ones are. I hope you get to see the Milky Way, because no matter where I have taken you, it is no longer visible. I hope to teach you about the ecosystem services that keep you and all other humans alive on this planet and that you show others how imperative to humanity it is to maintain these functions. As humans, in pursuit of short term profit, quickly approach irreversible thresholds, I hope that my generation can internalize the negative externalities of their economies. Informed by science, I hope that we make wise decisions that will make our lives sustainable for your generation. If we fail, let me apologize to you, your brothers and sisters, your children… as I predict, in the light of today's data, that you will all experience a loss of health and quality of life that may be unthinkable. It is this scary uncertainty that compels me to work so many hours; the fear of my child being exiled from the multidimensional niche space in which humans evolved. You are a great son and a great friend and I hope the best for you and the earth that supports your life.

TABLE OF CONTENTS

ABSTRACT ii

DEDICATION iv

CHAPTER 1 - THE USE OF MESOCOSM AND FIELD EXPERIMENTS TO PREDICT RESTORATION EFFECTIVENESS AND POTENTIAL FOR RECOVERY IN MINING-CONTAMINATED STREAMS 1

CHAPTER 2 - STRUCTURAL AND FUNCTIONAL RESPONSES OF PERIPHYTON AND MACROINVERTEBRATE COMMUNITIES TO FERRIC FE, CU AND ZN IN STREAM MESOCOSMS 36

CHAPTER 3 - CHRONIC TOXICITY OF FERRIC IRON FOR NORTH AMERICAN AQUATIC ORGANISMS: DERIVATION OF A CHRONIC WATER QUALITY CRITERION USING SINGLE SPECIES AND MESOCOSM DATA 68

CHAPTER 4 - EXPERIMENTAL ASSESSMENT OF SIZE-DEPENDENT SENSITIVITY OF AQUATIC INSECTS TO METALS – HOMAGE TO HORTON HEARS A WHO! 100

CHAPTER 5 - SYNTHESIS AND SUMMERY OF RESULTS 131

APPENDIX A - SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 1. 138

APPENDIX B - SUPPLEMENTAL FIGURES, TABLES AND NARRATIVE FOR CHAPTER 2 148

APPENDIX C - SUPPLEMENTAL TABLES AND NARRATIVE FOR CHAPTER 3 159

APPENDIX D - SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 4 173

CHAPTER 1

THE USE OF MESOCOSM AND FIELD EXPERIMENTS TO PREDICT RESTORATION EFFECTIVENESS AND POTENTIAL FOR RECOVERY IN MINING-CONTAMINATED STREAMS[1]

Introduction

Responses of stream benthic communities to the impacts of metals and other stressors associated with mining operations have been well characterized in the literature (Clements et al. 2000; Niyogi et al. 2002; Maret et al. 2003; Hornberger et al. 2009; Schmidt et al. 2010). Reduced abundance, loss of species and shifts in community composition from sensitive to tolerant taxa routinely occur downstream from current and historical mining operations. Although less commonly measured, alterations in ecosystem processes, including reduced primary productivity, lower rates of litter decomposition and altered trophic relationships, have also been reported (Niyogi et al. 2001; Carlisle and Clements 2003). Because mining discharges often represent a complex mixture of physical and chemical stressors, identifying mechanisms and causal relationships between mining disturbance and community responses in the field is challenging (DeNicola and Stapleton 2002; Battaglia et al. 2005). In addition to the direct toxicological effects associated with elevated concentrations of metals and reduced pH, mining activity can influence aquatic organisms indirectly through physical alterations of habitat, including increased sedimentation and substrate embeddedness. At circumneutral pH, metals such as Fe, Al and Mn precipitate and coat substrate, fill interstitial spaces and can smother

1 After acceptance by my committee a revised version this chapter were published by Environmental Science and Technology as: Cadmus, P., W. H. Clements, J. L. Williamson, J. F. Ranville, J. S. Meyer, and M. J. Gutiérrez-Ginés. The Use of Field and Mesocosm Experiments to Quantify Effects of Physical and Chemical Stressors in Mining-Contaminated Streams. *Environmental Science & Technology* **2016** 50 (14), 7825-7833. DOI: 10.1021/acs.est.6b01911

periphyton and benthic macroinvertebrate consumers (Vuori 1995). Although water quality criteria have been developed for some metals associated with mining activities (e.g., Zn, Cu, Cd), effects of metal-oxide precipitates are generally not considered (Linton et al. 2007). Furthermore, there is evidence that aquatic insects, particularly early instars which are underrepresented in the databases used to develop these criteria (Brix et al. 2005), are considerably more sensitive to metals compared to traditional test species (Clements et al. 2013).

Results of studies attempting to quantify the relative importance of chemical and physical stressors on benthic communities in streams have been somewhat equivocal. In an assessment of acid mine drainage (AMD) impacts on macroinvertebrates and periphyton, McKnight and Feder (1984) concluded that physical effects of metal oxides were greater than the direct toxic effects of low pH or high concentrations of dissolved metals. In contrast, Battaglia et al. (2005) reported relatively little effect of metal-oxide-coated substrate on macroinvertebrates in the field, attributing most effects to poor water quality. Relatively low toxicity of metal-oxide precipitates has also been reported in field (DeNicola and Stapleton 2002) and in laboratory experiments (Dsa et al. 2008). However, DeNicola and Stapleton (2002) cautioned that metal-oxide precipitates could become a more significant source of metal contamination following remediation treatments that are designed to reduce metal loading at the source.

An understanding of the relative contributions of chemical and physical stressors in mining-impacted streams is crucial for designing stream restoration treatments and predicting effectiveness of these programs. Similar to the development of water quality standards, restoration efforts in mining-polluted streams generally focus on improving water quality, with the implicit assumption that reduced loading of dissolved metals will also result in improvements in the physical habitat. Although sophisticated mechanistic models have been developed to

predict downstream improvements in water quality following treatment of mining discharges (Walton-Day et al. 2012), it is unlikely these models can predict biological recovery in systems in which metal-oxide deposition has occurred (Niyogi et al. 2001). These indirect effects on benthic habitat will likely delay biological recovery of streams following improvements in water quality (Walter et al. 2012).

Predicting responses of benthic communities to chemical and physical stressors associated with AMD is also complicated by significant variation among taxa. Field experiments conducted in a mining-impacted Colorado stream demonstrated that some metal-sensitive taxa readily colonized metal-oxide-coated substrate, whereas other known metal-tolerant taxa avoided these materials (Courtney and Clements 2002). Sasaki et al. (2005) hypothesized that the physical effects of metal-oxide deposition will be greater for erosional species (e.g., mayflies, stoneflies and caddisflies) than for depositional species (e.g., chironomids), and showed that the sequence of recolonization of a restored coal mining stream was consistent with these predictions (Walter et al. 2012). A better understanding of differences in sensitivity to physical and chemical stressors among taxa may help explain the observed variation in ecological effects of mine discharges and improve the ability to predict responses to restoration treatments (Niyogi et al. 1999).

In addition to variation in susceptibility of macroinvertebrates to dissolved metals and metal-oxide-coated substrates, recovery of AMD-contaminated streams following improvements in habitat and water quality is highly dependent recolonization from upstream. Because macroinvertebrate colonization in streams occurs primarily by downstream drift (Waters 1972), predicting recovery requires an assessment of drift rates of individual taxa. Thus, estimating

recovery potential of individual taxa requires an understanding of three factors: tolerance of dissolved metals, tolerance of metal-oxide-coated substrates, and drift propensity.

The goal of this research was to quantify the relative importance of physical (metal-oxide deposition) and chemical (elevated dissolved metal concentrations) stressors on benthic communities. To predict colonization rate and recovery potential for dominant macroinvertebrate taxa, this study integrated experimentally derived estimates of tolerance to dissolved metals and contaminated substrate with measures of drift propensity obtained from the literature. My estimates of recolonization potential were then compared to benthic macroinvertebrate data collected at recovery sites downstream from a source of metal inputs.

Methods

Study Site

The field study was conducted in the North Fork of Clear Creek (NFCC), a 3rd-order tributary to the Clear Creek watershed located near Blackhawk, Colorado (Figure 1.1). NFCC was placed on the National Priorities ("Superfund") List in September 1983 because of elevated concentrations of metals in Clear Creek, which supplies drinking water to over 500,000 residents in metropolitan Denver. North Fork of Clear Creek is characterized as a high-gradient stream with steep canyon walls and strong seasonal variability in stream discharge. Two point sources of metals (National Tunnel and the Gregory Incline Tunnel) account for most of the metal loadings to NFCC (Butler et al. 2009). Water quality in NFCC downstream from these inputs is characterized as circumneutral (pH 6.5-7.5), with low alkalinity (< 15 mg/L as $CaCO_3$), low concentrations of dissolved organic carbon (DOC; 1-3 mg C/L) and high concentrations of Fe (2-20 mg/L), Zn (0.5-1.5 mg/L), and copper (0.01-0.05 mg/L). Water hardness is low upstream of

the AMD inputs (<50 mg/L as $CaCO_3$), but increases downstream to 200-300 mg/L as $CaCO_3$ as a result of those inputs.

A water treatment facility proposed for 2016 is expected to significantly reduce metal loadings to NFCC, providing a unique opportunity to test predictions about recovery of this system following improvements in water quality. Because the existing discharges will be diverted away from the stream, aqueous metal concentrations are expected to decrease immediately. However, elevated metal concentrations in the sediment and effects of metal-oxide deposition will continue to impact this system and may lead to a more complex response than would be expected based on water-column concentrations alone.

Field Biomonitoring

To assess effects of mining discharges in NFCC, quantitative benthic samples were collected in spring 2012 and 2013 along a longitudinal gradient of metal contamination. Reference, impacted and recovery reaches (n = 2 sites per reach) were located upstream, immediately downstream and several km downstream from the sources of metal inputs in the NFCC, respectively. Replicate (n = 5) samples were collected from riffle areas using a 0.1 m^2 Hess sampler. Samples were washed through a 350-µm sieve in the field and organisms were preserved in 80% ethanol. In the laboratory, all organisms were identified to the lowest practical level of taxonomic resolution (genus or species for most aquatic insects; subfamily for chironomids) and enumerated (Merritt et al. 2008).

Mesocosm Experiments

To quantify direct effects of metals on benthic communities, 2 mesocosm experiments were conducted using natural communities collected from an upstream reference site on the NFCC. Mesocosm experiments were conducted at the Stream Research Laboratory (SRL)

located at the Colorado State University Foothills Campus, Fort Collins, CO. Details of the experimental facility and the process used to obtain benthic communities have been described previously (Clements et al. 2013). Briefly, 10 x 10 x 10-cm plastic trays filled with natural pebble and cobble substrate were placed at a NFCC reference site for 30 d. Three holes (2.5 cm diameter) drilled in each side of the trays increased water flow and facilitated macroinvertebrate colonization. Trays were colonized by a diverse assemblage of aquatic insects that are representative of communities in the natural substrate. After colonization, trays were removed from the stream, placed in 4-L insulated containers filled with stream water (4 trays per container) and transferred to the SRL. The contents of each container were randomly assigned to one of 18 stream mesocosms. Diluent water for the mesocosms originates from a deep mesotrophic reservoir and was delivered to each 20-L stream at a rate of 1.0 L/min. Water quality characteristics (pH, conductivity, temperature, dissolved oxygen) were typical of mountain streams in Colorado. Paddlewheels maintained a constant current velocity of 0.35 m/s in the mesocosms.

Mesocosm experiment I (July 2012) exposed benthic communities to a mixture of metals (Cu, Fe, Mn, and Zn) at concentrations representative of those measured at the reference, impacted and recovery reaches in NFCC. Stock solutions of metals for this experiment were prepared using analytical grade $ZnSO_4 \cdot 7H_2O$, $CuSO_4 \cdot 5H_2O$, $FeCl_3 \cdot 6H_2O$ and $MnCl_2 \cdot 4H_2O$ dissolved in stream water collected from the reference reach. Streams (n = 2 per treatment) were dosed for 10 d with a mixture of metals at 9 concentrations: 0X (reference), 0.015X, 0.030X, 0.060X, 0.125X, 0.25X, 0.375X, 0.50X and 1.0X, where X = the mean metal concentration previously measured at the impacted sites in the NFCC (13 µg/L Cu; 6,580 µg/L Fe; 2,326 µg/L Mn; and 776 µg/L Zn). Stock solutions were brought to pH 7 using sodium hydroxide and were

aerated to ensure precipitates remained in suspension. Peristaltic pumps delivered stock solutions to treated streams from 20-L carboys at a rate of 10 ml/min.

We used this same experimental design in mesocosm experiment II (August 2013) except instead of using metal salts the streams were dosed with effluent collected directly from Gregory Incline, the primary source of metals to the NFCC. Effluent (~1200 L) was collected every 2 days for these experiments. To achieve a range of metal concentrations similar to that in mesocosm experiment I, peristaltic pumps delivered effluent to each treated stream at a rate of 1.3-86.9 ml/min.

Whole community metabolism was estimated in each stream by measuring changes in dissolved oxygen (DO) concentrations of trays placed in clear, sealed polycarbonate containers (Appendix A, Figure 1.S1). Peristaltic pumps circulated water in the containers over a DO probe (YSI Incorporated, Yellow Springs, OH, USA) and concentrations (mg O_2/L) were measured every 1-4 minutes for 30 minutes in light and dark conditions. Gross primary productivity (mg O_2 per hour) was estimated by summing oxygen production during light conditions (net community production) and oxygen consumption during dark conditions (community respiration). To measure the effects of metals on periphyton we estimated net primary productivity on 6.25 cm^2 unglazed porcelain tiles (Cinca Tile Co., Fiães, Portugal) placed in each mesocosm at the start of mesocosm experiment II. After 8 days 2 colonized tiles were removed and placed in a 24 ml respiration chamber (Appendix A, Figure 1.S2). Assessments were conducted under uniform light in a chilled water bath that matched the temperature of mesocosm streams. Changes in DO were calculated as described above.

Field Experiment

To assess the effect of metal-contaminated sediment on macroinvertebrates in the field, we conducted a substrate colonization experiment in NFCC during May 2013. Trays (n = 32) filled with cobble substrate were placed at an impacted site for 31 d. During this period the substrate and interstitial spaces in the trays accumulated a coating of metal oxides. These metal-contaminated trays and an additional 32 control trays containing clean substrate were then deployed at an upstream reference site and colonized for 31 d. To determine the influence of water flow on macroinvertebrate colonization and metal loss, 16 trays of each type were nested within an empty tray that lacked holes. After 31 d a randomly selected set of 4 trays from the same treatment were removed from the stream, composited and treated as a single experimental unit. Benthic macroinvertebrates in the trays were processed as described above. The final experimental design was a 2 x 2 factorial (n = 4), with 2 levels of substrate quality (control versus metal-contaminated) and 2 levels of water flow (low versus high). Additional trays of each type were used throughout the experiment to measure metal deposition and loss during the colonization period.

Estimating Colonization and Recovery Potential

To estimate recolonization and potential for recovery in NFCC, I integrated estimates of tolerance to dissolved metals and metal-oxide-coated substrate with published values of macroinvertebrate drift propensity. I defined tolerance to dissolved metals in the mesocosm experiments as the percent survival at the 2 highest metal treatments (relative to controls), averaged across both experiments. Tolerance to metal-oxide deposition was defined as the percent of organisms on contaminated substrate relative to control substrate. To estimate colonization potential for dominant macroinvertebrates in NFCC we used Rader's (1997)

assessment of drift propensity for 95 species based on 6 traits (intentional drift, habitat preferences, flow exposure, mobility, an index of drag and drift distance). Recovery potential for dominant NFCC macroinvertebrates was estimated based on their tolerance to dissolved metals, tolerance to contaminated substrate and relative drift propensity.

Chemical Analyses

Routine water quality characteristics (dissolved oxygen, pH, conductivity, temperature) were measured on days 2, 4 and 7 of each mesocosm experiment using hand held meters (models 550A and 63; YSI Incorporated, Yellow Springs, OH). Water samples (0.5 L) were collected to determine hardness and alkalinity in the laboratory using standard titration procedures. Concentrations of Cu, Fe, Mn and Zn, the primary metals of concern in NFCC, were measured in the field and in stream mesocosms. Water samples (15 ml) for analysis of dissolved metals were filtered through a 0.45-µm pre-rinsed glass fiber filter and acidified to a pH of < 2.0 with analytical grade nitric acid. Metal concentrations were analyzed using inductively coupled plasma-optical emission spectrometry (ICP-OES) with a Perkin Elmer Optima 5300 DV (Perkin Elmer, Waltham, Massachusetts). Quality assurance/quality control (QA/QC) samples included deionized (DI) water blanks (Barnstead Nanopure system, Thermo Fisher Scientific) that contained trace-metal-grade concentrated HNO_3 (Thermo Fisher Scientific) and certified continuing calibration verification (CCV) standards. The QA samples were analyzed immediately after instrument calibration, after every 10 samples, and at the end of each set of samples. Additionally, certified standard reference materials from the National Institute of Standards and Technology were analyzed for trace elements at the beginning and end of each analytical run.

We determined metal concentrations associated with the metal-oxide-contaminated substrate in the field experiment using a modification of EPA Method 3050B. Sediment was transferred from exposure trays to a 2-L polypropylene container, 100 ml of 50% trace-metal-grade nitric acid (diluted with Milli-Q water) was added to the container and the opening was covered with a Teflon watch-glass. The container was placed in a hot-water bath (~90°C) for ~15 min, then removed from the bath and allowed to cool. Concentrated trace-metal-grade nitric acid (50 ml) was added to the digestion container and this mixture was allowed to sit at room temperature for 30 min. The container was returned to the hot-water bath (~90°C) for 2 h, then removed and allowed to cool. Milli-Q water (50 ml) and 75 ml of hydrogen peroxide solution (Macron, 30% reagent grade) were added to the mixture. The container was returned to the hot water bath (~90°C) for 15 min, allowed to cool and the total mass was measured. The acid solution remaining in the container was sampled with a disposable polyethylene pipette into a conical tube (Falcon, 50 ml polypropylene), diluted 20X with Milli-Q water, and analyzed for dissolved metal concentrations as described above. The sediment mass, volume of extraction solution and measured metal concentrations were used to calculate mg metal/kg sediment.

Statistical Analyses

Differences in macroinvertebrate abundance and number of taxa among sites (reference, impacted and recovery) and between years (2012 and 2013) were analyzed using two-way ANOVA (PROC ANOVA; SAS Institute, Version 9.3, Cary, NC). If the overall effects were significant, we used multiple comparisons (REGWQ) to test for differences among means. We developed concentration-response relationships for community metrics (total macroinvertebrate abundance, taxonomic richness, abundance of major macroinvertebrate orders) and abundance of dominant taxa using linear regression (PROC REG) on log10 (x+1) transformed values. We also

used general linear models (PROC GLM) to test for differences in responses to metals between the 2 experiments (year x treatment interaction). We used 2-way ANOVA (PROC ANOVA) to determine effects of substrate quality, flow conditions and their interaction in the field experiment. Where necessary, abundance measures were log10 (x+1) transformed to satisfy assumptions of parametric statistics.

Results

Field Biomonitoring of Benthic Macroinvertebrates

Metal concentrations in the NFCC were generally low or below detection at the reference reach, increased 20X-150X below the mine adits and then decreased downstream below the confluence with Clear Creek (Table 1.S1, Supporting Information). Benthic macroinvertebrate communities were effectively eliminated at the metal-impacted sites, but showed significant improvement at the downstream recovery sites (Figure 1.2). Although communities at the reference sites consisted of a diverse assemblage of between 20-23 species of mayflies, stoneflies, caddisflies and dipterans, 3 dominant taxa (*Baetis* sp., *Hydropsyche* sp., and orthoclad chironomids) accounted for 74% of total macroinvertebrate abundance at the downstream recovery sites.

Mesocosm Experiments

With the exception of variables that were directly affected by metal treatments (e.g., conductivity), routine water quality characteristics in mesocosms were generally similar among treatments and between the 2 experiments (Table 1.S2, Supporting Information). Conductivity and water hardness increased as a function of metal treatment in both experiments, but this trend was more apparent in 2013. Circumneutral pH values were observed across all treatments in both

experiments, although pH decreased with metal treatment in the 2013 experiment. Measured concentrations of Cu, Fe, Mn and Zn increased with treatment levels (Table 1.S3, Supporting Information) and were generally consistent between the two experiments.

After 10 days of exposure to metals we observed highly significant concentration-response relationships in both mesocosm experiments; however, the strength of these relationships varied among taxa and macroinvertebrate community metrics. Compared to controls, total macroinvertebrate abundance was reduced by 53-58% and number of taxa was reduced by 25-30% at the highest treatment concentrations (Appendix A, Figure 1.S3). Total abundance of mayflies (Ephemeroptera), stoneflies (Plecoptera) and dipterans were significantly reduced by metals in both mesocosm experiments (Appendix A, Figure 1.S4). The greatest effects were observed on abundance of mayflies, which was reduced by > 90% at metal concentrations approximating those measured at impacted sites in NFCC. In contrast, caddisflies (Trichoptera) were relatively tolerant of metals and their abundance was not significantly related to metal concentration in either experiment.

Eight macroinvertebrate families (Baetidae, Ephemerellidae, Heptageniidae, Nemouridae, Chloroperlidae, Rhyacophilidae, Chironomidae, and Elmidae) were dominant in stream mesocosms and accounted for over 97% of total macroinvertebrate abundance in the 2 experiments. As indicated by the slopes and the strength (r^2 values) of the concentration-response relationships, considerable variation among these groups was observed (Figure 1.3). Some taxa were reduced at low to moderate metal concentrations, whereas others were unaffected even at the highest concentrations. Effects of metals were greatest on the 3 mayfly families (Baetidae, Ephemerellidae and Heptageniidae). Although I observed significant concentration-response relationships for other taxa (chloroperlid stoneflies in mesocosm experiment II; chironomids in

both experiments), these relationship were considerably weaker. Rhyacophilidae, the dominant caddisflies in mesocosms, and elmid beetles were unaffected by metals in either experiment.

The responses of macroinvertebrate communities to metals in the 2 mesocosm experiments were quite similar, despite the very different dosing procedures. Results of general linear models analysis showed no significant year x treatment interaction for most macroinvertebrate groups. These results indicate that regardless of whether metals were delivered as a mixture of laboratory salts or as whole effluent collected from Gregory Incline, effects were similar across macroinvertebrate taxa. The only exception to this observation was for abundance of chloroperlid stoneflies (Figure 1.3e), which showed significantly greater effects in the whole effluent experiment (experiment II) compared to the experiment using metal salts (year x treatment interaction, $p < 0.0001$).

Effects of effluent from Gregory Incline significantly reduced whole community metabolism and net primary productivity in mesocosm experiment II. Community respiration in trays exceeded photosynthesis in all treatments above 0.125X (Figure 1.4a). These effects were more extreme on periphyton that colonized tiles during the 10 d experiment (Figure 1.4b). All treatment groups >0.03X produced little or no oxygen, suggesting no colonization occurred on these tiles.

Field Experiments

Substrate deployed at the AMD-impacted site in NFCC accumulated high concentrations of metals (Table 1.S4, Supporting Information). Although metal concentrations decreased when trays were transferred to the reference site, they remained elevated compared to control (washed) substrate. Concentrations of Cu, Fe and Zn were 1.1-4.0 times greater on metal-oxide-contaminated substrate compared to control substrate at the end of the field experiment. These

differences were more apparent under low-flow conditions. In contrast to other metals, Mn concentrations increased when trays were placed at the reference site and did not differ between control and metal-oxide-contaminated substrate at the end of the experiment.

Patterns of macroinvertebrate colonization on trays in NFCC reflected differences in metal concentrations among treatments. Total macroinvertebrate abundance was 47% lower on trays containing metal-contaminated substrate compared to controls (Appendix A, Figure 1.S5). Effects of metal-contaminated substrate on macroinvertebrate colonization were highly significant ($p = 0.001$), but there was no effect of water flow on total abundance ($p = 0.126$). Total number of taxa was also significantly ($p = 0.001$) lower in trays containing contaminated substrate; however, this effect was generally limited to trays under low flow conditions, as indicated by the significant substrate x flow interaction term ($p = 0.0073$).

Eight macroinvertebrate taxa (Baetidae, Heptageniidae, Nemouridae, Chloroperlidae, Lepidostoma, Pericoma, Chironomidae, and Elmidae) accounted for over 88% of total abundance in the field experiment. Effects of substrate quality and flow conditions on these dominant groups varied among taxa (Figure 1.5). Although abundance of most taxa was reduced in trays containing contaminated substrate, baetid mayflies ($p = 0.1611$) and the stoneflies Nemouridae ($p = 0.3138$) and Chloroperlidae ($p = 0.3808$) were unaffected by substrate quality. Abundance of Baetidae ($p = 0.0008$), Heptageniidae ($p = 0.0001$) and Nemouridae ($p = 0.0463$) was significantly lower in trays with reduced water flow. Heptageniidae was the only group to show a significant substrate x flow interaction ($p = 0.0149$), indicating that effects of contaminated substrate on colonization were greater at low flow.

Colonization Potential of Benthic Macroinvertebrates

Estimates of colonization potential based on tolerance to dissolved metals, avoidance of contaminated substrate and drift propensity showed considerable variation among macroinvertebrate families (Figure 1.6). Because of their tolerance to metals and moderate drift propensity, I predicted that some caddisflies (Hydropsychidae, Rhyacophilidae) and stoneflies (Chloroperlidae) should recover rapidly following improvements in water quality. Despite their extreme sensitivity to dissolved metals, baetid mayflies are also expected to recover quickly because of their exceptionally high drift propensity and tolerance for metal-oxide-contaminated substrate. In contrast, the mayflies Ephemerellidae and Heptageniidae have moderate drift propensity, but were sensitive to both dissolved metals and contaminated substrate. These groups would be expected to recover slowly. Finally, the dipteran Psychodidae was highly tolerant to dissolved metals in stream mesocosms; however, these organisms are expected to recover very slowly because of their low tolerance to metal-oxide-contaminated substrate and very low drift propensity.

Discussion

Predicting effects of AMD discharges on streams and the effectiveness of restoration treatments in these systems is challenging because of the complex mixture of physical and chemical stressors often present (DeNicola and Stapleton 2002; Battaglia et al. 2005; Dsa et al. 2008; Hogsden and Harding 2012). Although the focus of most restoration efforts in AMD-contaminated streams is on improving water quality, it is generally acknowledged that removing this single stressor may not be sufficient for restoring structural and functional integrity of these systems (Niyogi et al. 2001). The frequently observed lag in recovery of benthic communities following improvements in water quality has been attributed to either the residual impacts of Fe

hydroxide deposition (Walter et al. 2012) or the contribution of AMD precipitates to aqueous metal loading (DeNicola and Stapleton 2002).

Because remediation of contaminated sediments is considerably more expensive than improving water quality, understanding the relative influence of metal-oxide deposition versus direct toxicity of AMD discharges has important practical applications. These mesocosm and field experiments provide insights into the relative importance of these physical and chemical stressors. Because these experiments were conducted using natural communities collected from NFCC, results are directly applicable to the proposed restoration treatments for this watershed.

I observed that several macroinvertebrate groups, particularly the mayflies Baetidae, Ephemerellidae and Heptageniidae, were highly sensitive to dissolved metal exposure. These findings are consistent with previous field (Clements et al. 2000; Schmidt et al. 2012) and mesocosm (Clements et al. 2013) studies that reported low tolerance of mayflies to dissolved metals. In the present study, exposure to only 12.5% of the metal concentrations measured at the most contaminated site in the NFCC reduced abundance of mayflies by 31-35% after only 10 d. Despite a 4-6 X increase in water hardness at higher metal concentrations in mesocosm experiment II, responses of mayflies and other sensitive taxa were similar in the two experiments. These results suggest that water hardness alone may not be a strong determinant of metal toxicity to aquatic insects, at least in this metal mixture that was dominated by Cu and Zn. These results also indicate that most mayflies in NFCC would respond to improvements in water quality, but large reductions in dissolved metal concentrations would be necessary. Other macroinvertebrate groups (e.g., chironomids, chloroperlid stoneflies) were also affected by metals in these mesocosm experiments; however, these taxa were considerably more tolerant of AMD stressors and would likely recover following modest improvements in water quality.

Macroinvertebrate abundance was approximately 50% lower on metal-oxide-contaminated substrate compared to controls in field experiments. These effects resulted from either avoidance of metal-oxide-coated substrate or direct mortality of organisms that colonized trays. Although my experimental design did not allow us to directly test these two hypotheses, I do not believe that exposure to metal oxides resulted in significant mortality. Previous mesocosm experiments have shown little evidence of direct toxic effects of dissolved Fe or metal-oxide-coated substrates (Cadmus, unpublished results). Furthermore, measured concentrations of Cd, Cu, and Zn associated with NFCC substrate were typically an order of magnitude less than the probable effect concentrations (PECs) for these metals (MacDonald et al. 2000). Although we recognize that these guidelines were based on sediment toxicity tests conducted using much finer grain sizes, they should provide a reasonable estimate of potential toxicity. Similar sediment quality guidelines have not been developed for Fe; however, it is well established that ferric Fe is relatively non-toxic to macroinvertebrates (Gerhardt 1992) and that effects in Fe-contaminated streams result primarily from deposition of Fe hydroxides either directly on organisms (Gerhardt 1992; Soucek et al. 2000) or in benthic habitats (Vuori 1995; Linton et al. 2007). I hypothesize that reduced colonization of contaminated substrate in these experiments was most likely a result of behavioral avoidance by macroinvertebrates. Given the low periphyton biomass observed at NFCC and the sensitivity of periphyton to metals in the mesocosm experiment, avoidance of contaminated substrate may be partially attributed to reduced food quality for grazing insects.

Previous field studies have attributed reduced abundance of macroinvertebrates and alterations in ecosystem processes to the deposition of Fe hydroxides (Niyogi et al. 2001; McKnight and Feder 1984; Courtney and Clements 2002; Sasaki et al. 2005). In contrast, others have shown relatively modest responses to AMD precipitates, attributing most of the observed

effects to either low pH or elevated concentrations of dissolved metals in water (DeNicola and Stapleton 2002; Battaglia et al. 2005; Dsa et al. 2008). Results of this study indicate that macroinvertebrates responded to both the direct toxicological effects of dissolved metals and the physical effects of metal-oxide deposition; however, the relative importance of these 2 stressors varied among macroinvertebrate groups. For example, heptageniid mayflies were eliminated at low dissolved metal concentrations in the mesocosm experiments and showed a strong avoidance response to metal-oxide-contaminated substrate in the field. In contrast, some taxa that were highly sensitive to aqueous metals in stream mesocosms (e.g., Baetidae) showed only modest responses to metal-oxide-contaminated substrate in the field.

In addition to quantifying tolerance to dissolved metals and contaminated substrate, this study used drift propensity to estimate recolonization potential for dominant macroinvertebrates in the NFCC. Species that are relatively tolerant to AMD discharges (e.g., caddisflies and stoneflies) and/or have high drift propensity (baetid mayflies) are expected to recolonize and recover faster than sensitive species or those that are poorly represented in the drift. The dominance of hydropsychid caddisflies and baetid mayflies at the downstream recovery reach in the NFCC also suggests that these taxa would recover rapidly following AMD treatments. The extreme sensitivity of grazing mayflies (e.g., Heptageniidae) to both dissolved metals (Clements et al. 2013) and Fe-contaminated substrate (Linton et al. 2007) is well established in the literature. Although these organisms have a relatively high drift propensity, one would expect slower recovery because of their sensitivity to AMD stressors. Finally, despite their high tolerance of dissolved metals, the dipteran *Pericoma* (Psychodidae) is expected to recover very slowly because of their extremely low drift propensity.

My estimates of recovery potential generally agree with results of other spatially extensive (Clements et al. 2000; Schmidt et al. 2012) and longitudinal (Clements et al. 2010) surveys conducted along gradients of metal contamination; however, there were several notable differences. In particular, orthoclad chironomids, organisms that are generally characterized as metal-tolerant based on field surveys (Chadwick et al. 1986; Iwasaki et al. 2009; Clements et al. 2010), are expected to recolonize slowly. It is important to note that these estimates of recovery potential for dominant macroinvertebrates only account for relative sensitivity to AMD stressors and drift propensity. Consequently, responses to improvements in water quality may differ from those observed in field surveys. In addition, recovery potential will vary among streams and is dependent on the restoration process. For example, if metal concentrations in water decrease rapidly but remain elevated in substrate, species with high drift propensity that are also tolerant of metal-oxide-contaminated substrate (e.g., Baetidae) should recover quickly.

The responses of benthic communities to AMD discharges observed in the field were generally much greater than those observed in mesocosm experiments. Most macroinvertebrates, including several metal-tolerant taxa, were completely eliminated from the impacted reach and showed little recovery downstream. Given their sensitivity to dissolved metals and/or intolerance to contaminated substrate, one would have expected very low densities of mayflies in the downstream reach. However, the absence of other taxa, such as the predatory caddisflies Rhyacophilidae, was not expected based on their observed tolerance to metals, moderate drift propensity and high density in the reference reach (approximately 400/m^2). It is likely that other factors, such as the absence of suitable prey species played an important role in limiting abundance of these organisms in the field.

Because the downstream transport of metals and the deposition of metal oxides in benthic habitats are strongly influenced by stream flow (Butler et al. 2008), hydrologic characteristics of a watershed must also be considered when estimating potential recovery rates. Benthic communities in flashy, high-gradient streams (in which interstitial spaces are frequently scoured) may recover faster from the physical effects of metal-oxide deposition compared to those in low-velocity streams. Recovery of aquatic life could be delayed in rivers systems that are regulated for hydroelectric, agriculture or municipal water use if spring floods and natural flows are not maintained.

The development of water quality criteria using traditional test species (e.g., *Ceriodaphnia dubia*, *Pimephales promelas*) has been criticized, largely because of the inability of these experiments to account for the indirect effects of contaminants (Cairns 1986; Clements and Rohr 2009). Because aquatic insects are poorly represented in the database used to establish these criteria (Brix et al. 2005), there is the potential that these organisms will not be protected by current water quality standards (Schmidt et al. 2010). Additionally, laboratory toxicity tests conducted with aquatic insects, particularly those using later instars, show much greater tolerance to metals compared to patterns observed in the field (Brinkman and Johnston 2008; Brinkman and Johnston 2012). For contaminants that do not show significant acute toxicity (e.g., total dissolved solids, suspended sediments, nutrients, and deposition of metal oxides), traditional laboratory-based toxicity tests may be inappropriate and should be supported by alternative approaches (Pacheco et al. 2005). For example, using a field-based assessment, Linton (Linton et al. 2007) concluded that a more stringent benchmark for Fe would be necessary to protect mayflies and other grazers from the indirect effects of metal-oxide deposition. Similarly, benchmarks for total dissolved solids based on field surveys of macroinvertebrate

communities were much lower than those based on traditional laboratory toxicity tests (Cormier and Suter 2013). A crucial challenge with field-based approaches is that other confounding factors will influence patterns of abundance and distribution and complicate the ability to determine causal relationships between stressors and responses.

Mesocosm experiments represent an important middle ground between laboratory toxicity tests and field-based assessments (Clements and Kotalik 2016). Although small-scale experiments have been criticized in the literature (Carpenter 1996; Schindler 1998), I believe that mesocosm experiments, particularly those conducted with natural communities and an obvious connection to natural ecosystems, provide an ecologically realistic complement to laboratory toxicity tests (Clements 2004). These experiments also control for the confounding variables associated with field-based approaches, thereby supporting causal relationships between stressors and responses. The similarity of responses between my mesocosm experiments, which were conducted in 2 different years and using very different exposure regimes, demonstrates the reproducibility of this approach. This study quantified the relative importance of chemical and physical stressors associated with AMD discharges and predicted recovery potential for dominant macroinvertebrate taxa. A better understanding of the mechanisms responsible for variation among macroinvertebrate taxa should improve the ability to predict restoration effectiveness in AMD-contaminated streams.

Figures and Tables from Chapter 1

Contains:

Figures 1.1 to 1.6

See Appendix A for Supplemental Figures 1.S1 to 1.S5 and Supplemental Tables 2.S1 to 2.S4.

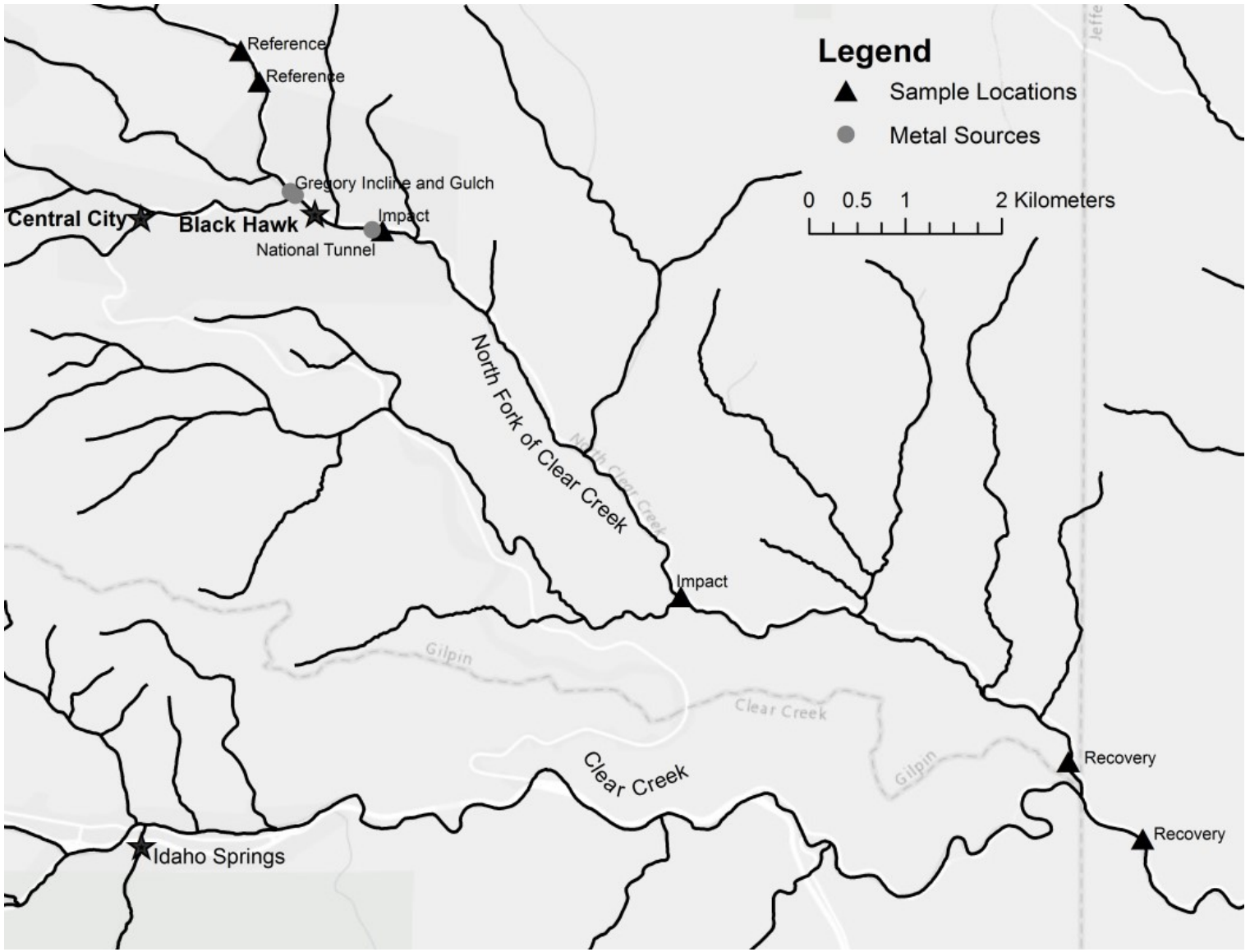


Figure 1.1 - Map showing location of upstream reference sites, impacted sites and downstream recovery sites in the NFCC, Colorado.

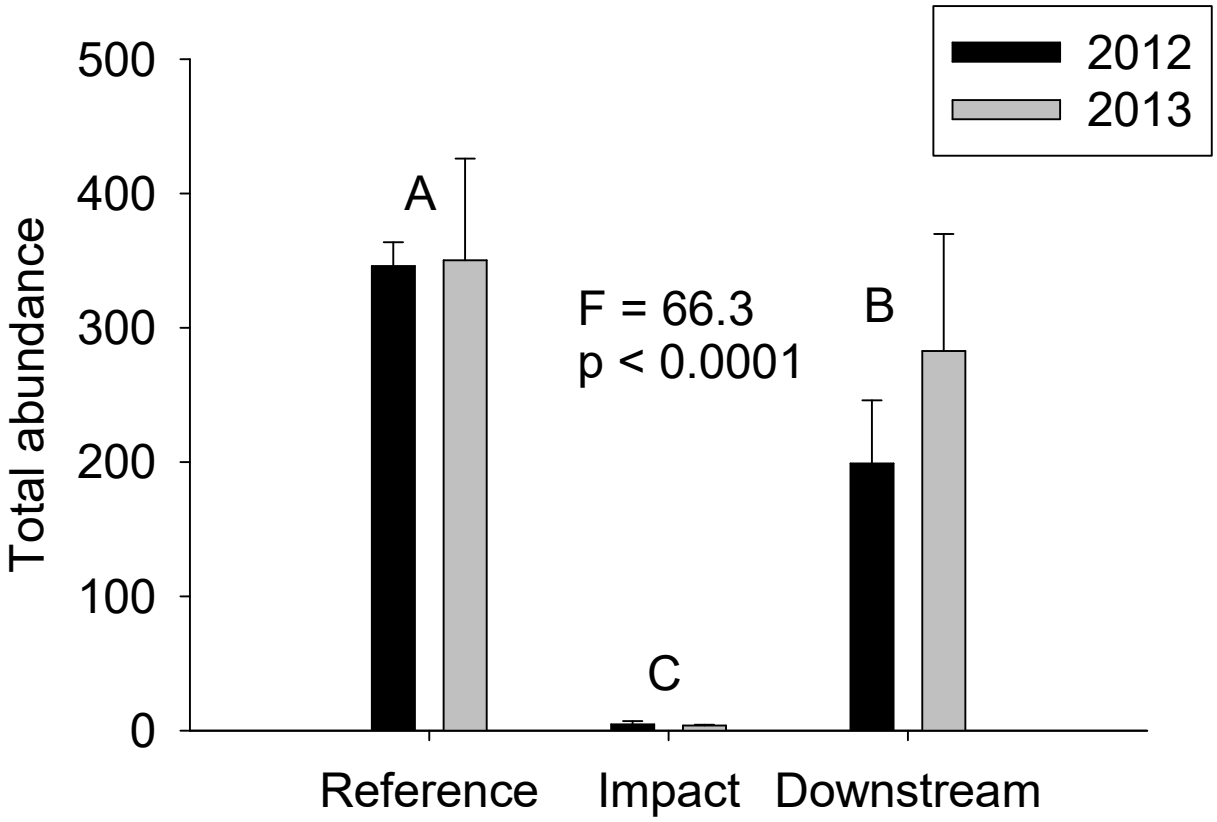


Figure 1.2 - Mean (± s.e.) macroinvertebrate abundance in Hess samples at reference, impacted and recovery stations in the North Fork of Clear Creek. Results of ANOVA testing for differences among stations are also shown (differences between years were not significant). Stations with the same letter were not significantly different.

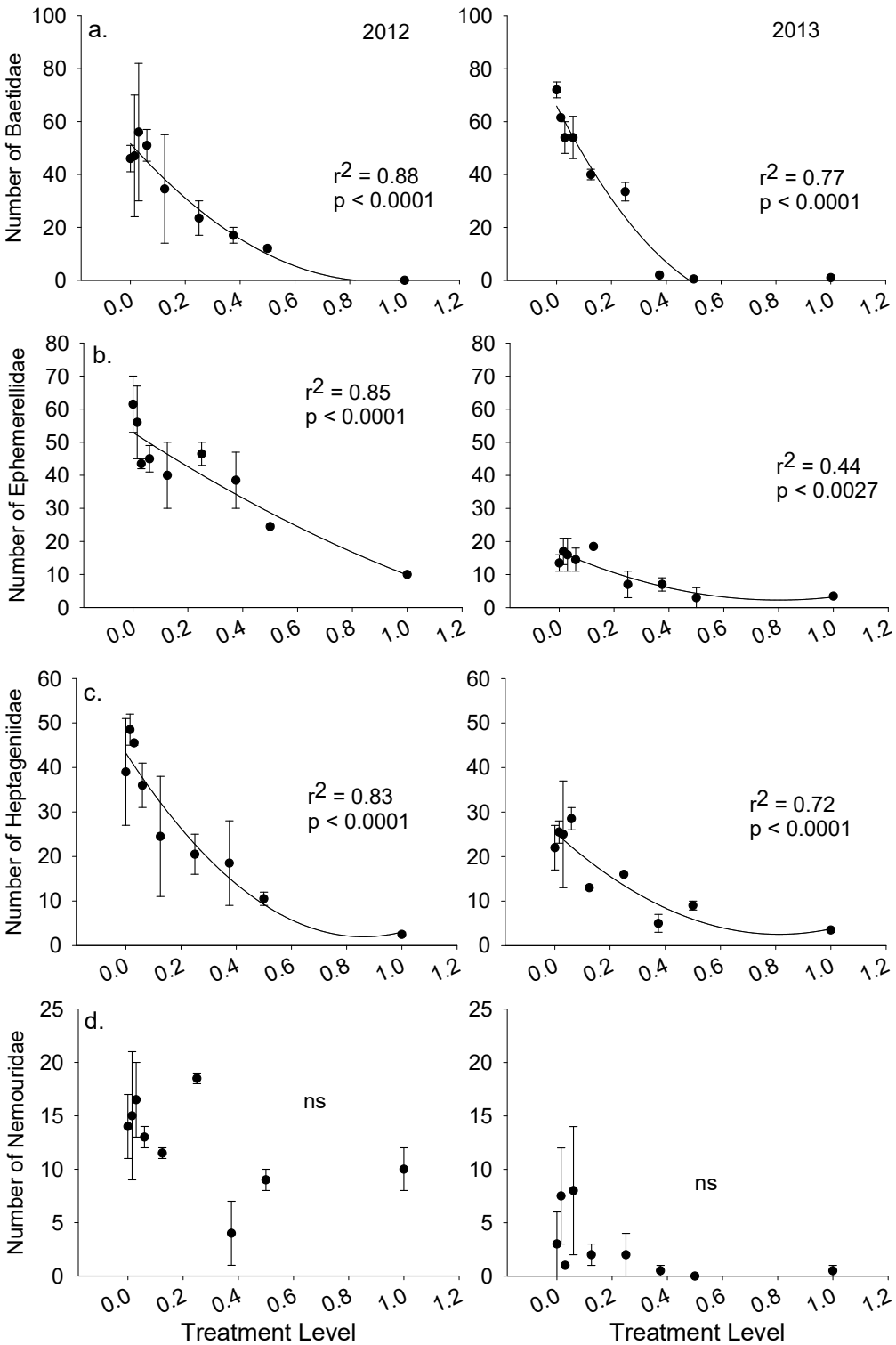


Figure 1.3 - Mean (+ s.e.) abundance of dominant macroinvertebrate taxa as a function of treatment level in stream mesocosm experiments conducted in 2012 and 2013. Results of statistical analyses testing for significant concentration-response relationships are shown for each dominant taxa.

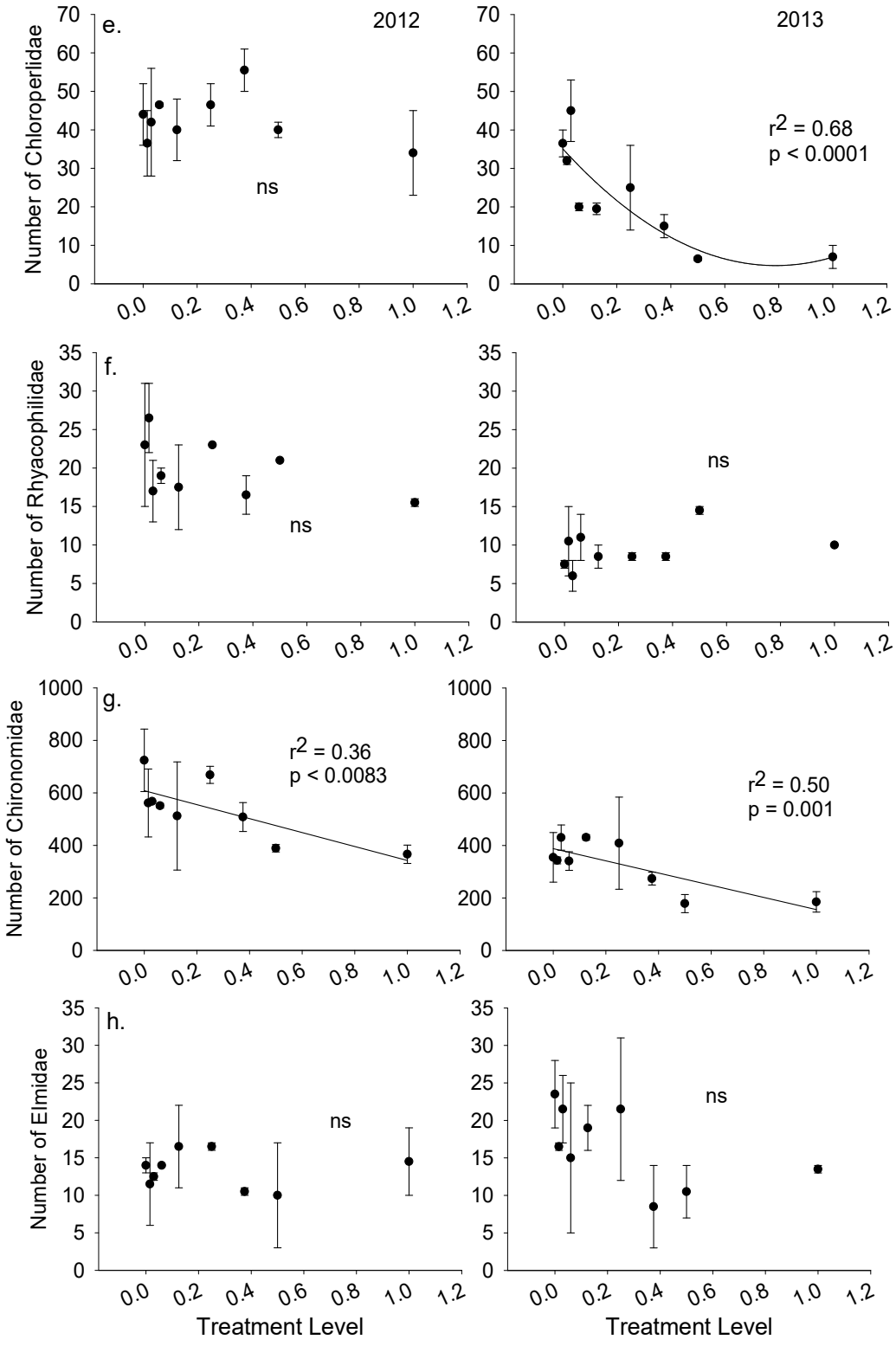


Figure 1.3 Continued

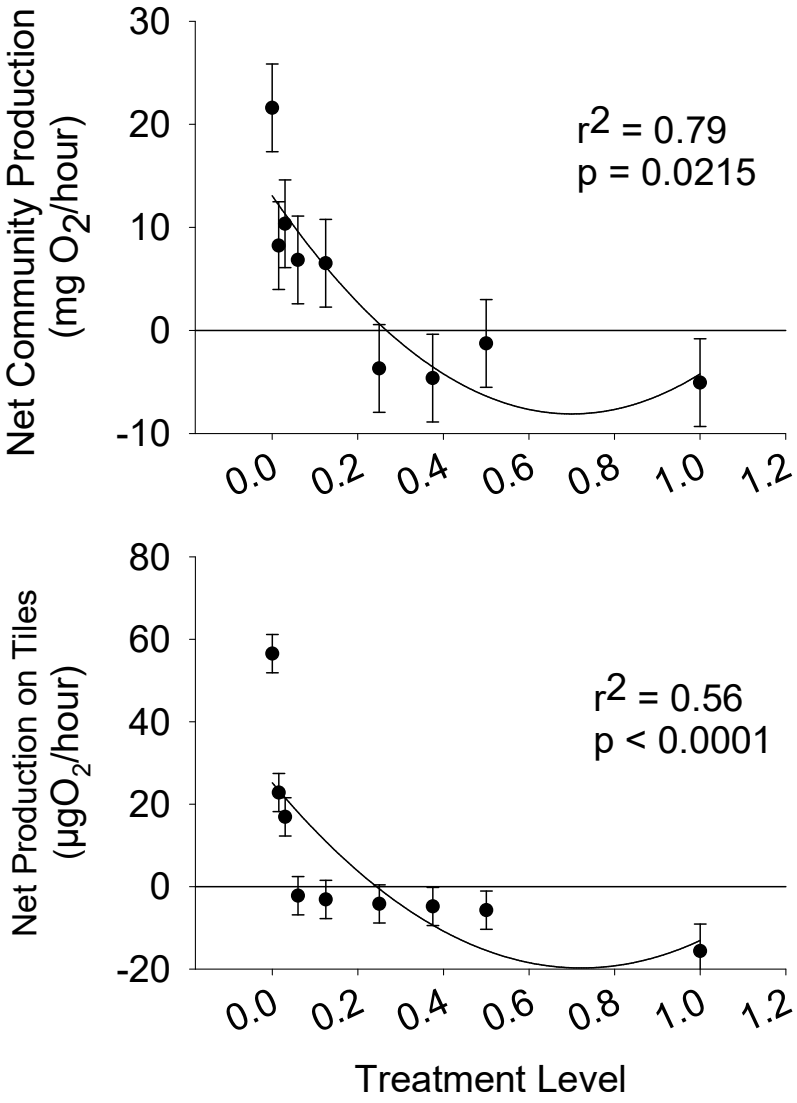


Figure 1.4 - Mean (± s.e.) whole community metabolism (a) and net primary production (b) as a function of treatment level in stream mesocosm experiments conducted in 2013. Results of statistical analyses testing for significant concentration-response relationships also are shown.

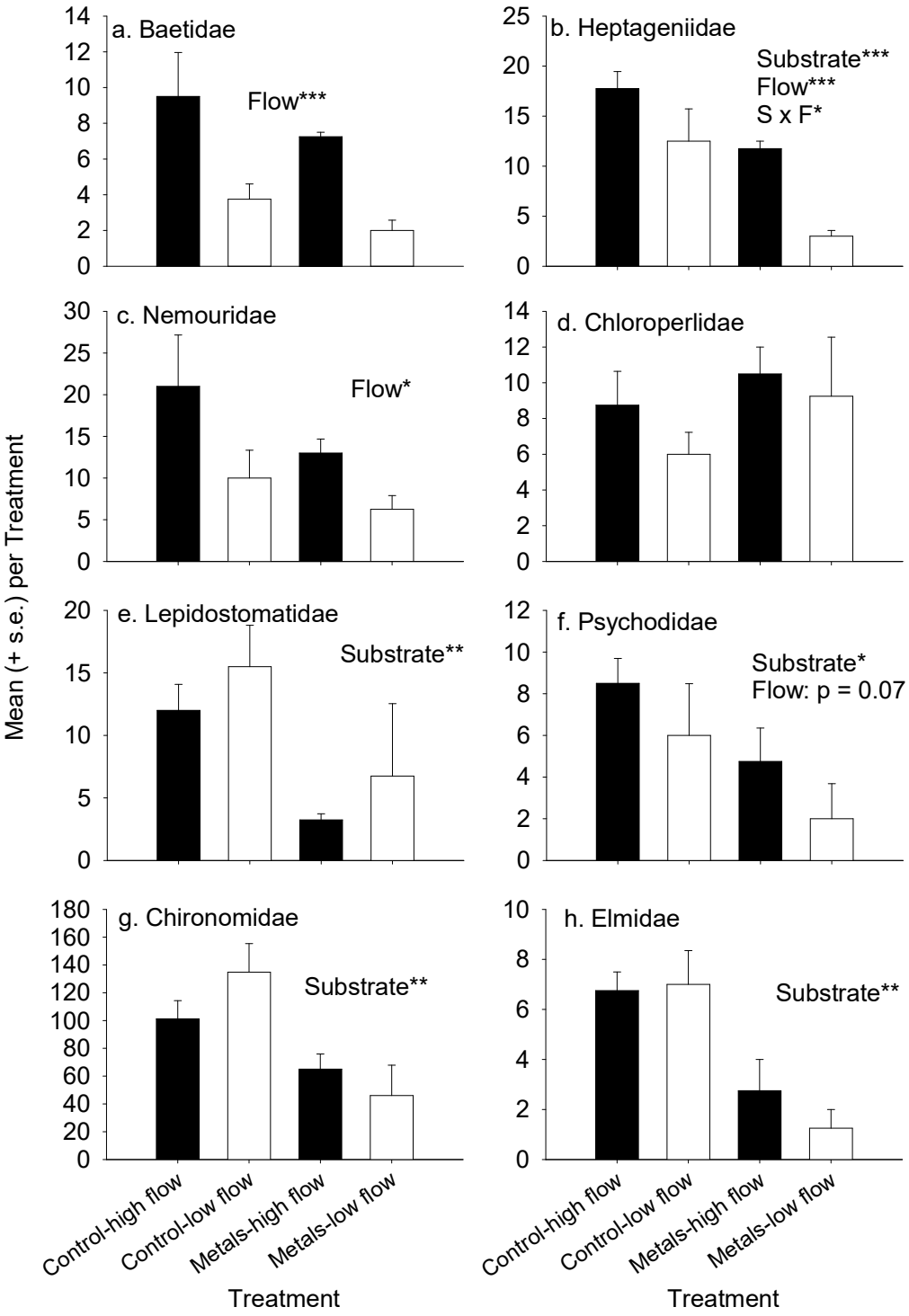


Figure 1.5 - Mean abundance of dominant macroinvertebrate taxa in colonization trays containing clean or metal-contaminated substrate placed at an upstream reference site. Results of 2-way ANOVA testing for effects of substrate (control versus metal-contaminated), flow (low versus high) and the substrate x flow (S x F) interaction are also shown. *$p < 0.05$; **$p < 0.01$; ***$p < 0.0001$.

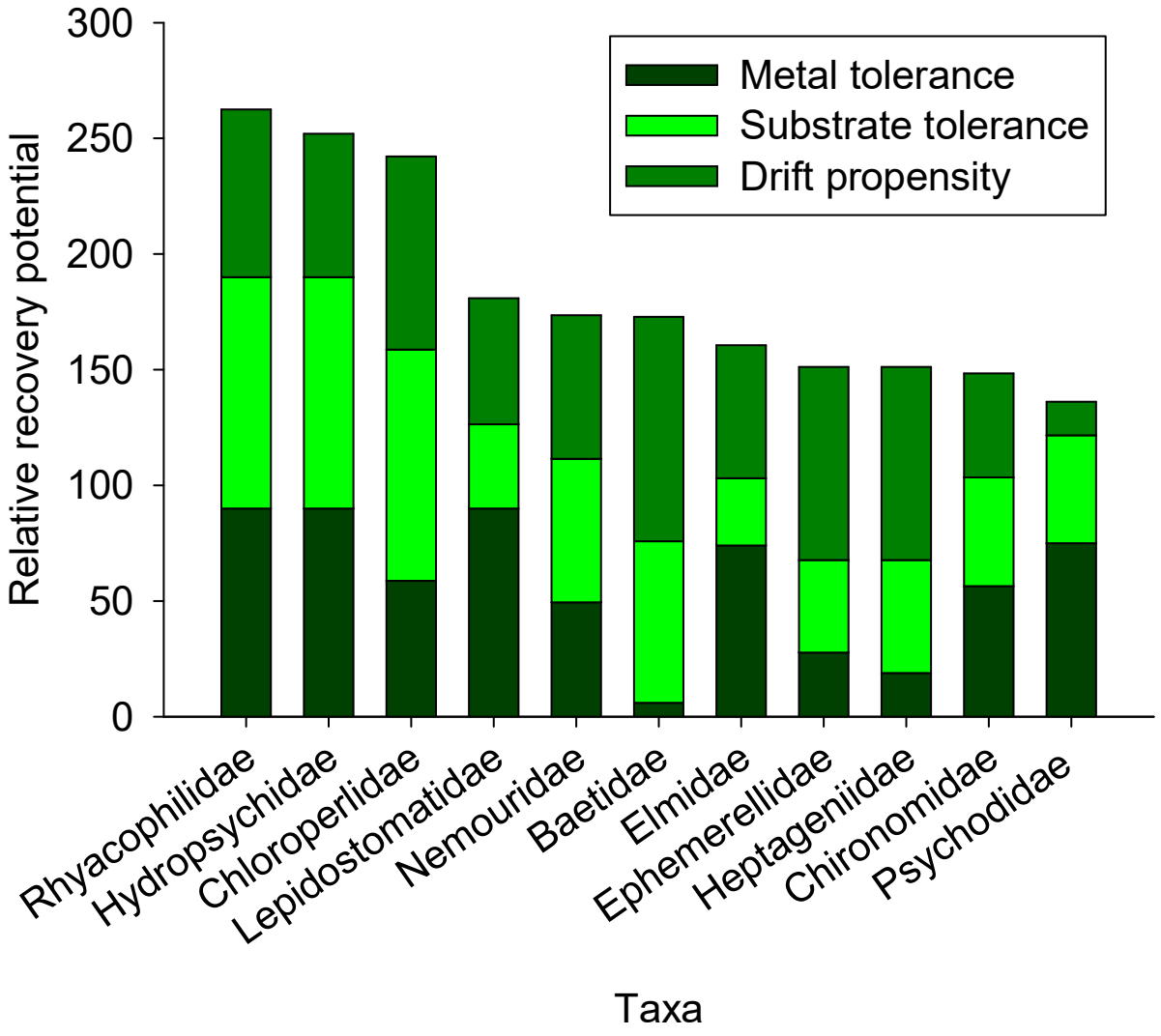


Figure 1.6 - Estimated recovery potential for dominant macroinvertebrate taxa in North Fork of Clear Creek, Colorado. Tolerances to metals and contaminated substrate were based on mesocosm and field experiments. Drift propensity was determined from Rader (1997).

References

Battaglia, M., G.C. Hose, E. Turak, and B. Warden, Depauperate macroinvertebrates in a mine affected stream: Clean water may be the key to recovery. *Environmental Pollution* **2005**, 138 (1), 132-141.

Brinkman, S.F. and W.D. Johnston, Acute toxicity of aqueous copper, cadmium, and zinc to the mayfly Rhithrogena hageni. *Archives of Environmental Contamination and Toxicology* **2008**, 54 (3), 466-472.

Brinkman, S.F. and W.D. Johnston, Acute toxicity of zinc to several aquatic species native to the Rocky Mountains. *Archives of Environmental Contamination and Toxicology* **2012**, 62 272-281.

Brix, K.V., D.K. DeForest, M. Burger, and W.J. Adams, Assessing the relative sensitivity of aquatic organisms to divalent metals and their representation in toxicity datasets compared to natural aquatic communities. *Human and Ecological Risk Assessment* **2005**, 11 1139-1156.

Butler, B.A., J.F. Ranville, and P.E. Ross, Observed and modeled seasonal trends in dissolved and particulate Cu, Fe, Mn, and Zn in a mining-impacted stream. *Water Research* **2008**, 42 (12), 3135-3145.

Butler, B.A., J.F. Ranville, and P.E. Ross, Spatial variations in the fate and transport of metals in a mining-influenced stream, North Fork Clear Creek, Colorado. *Science of the Total Environment* **2009**, 407 (24), 6223-6234.

Cairns, J.J., The myth of the most sensitive species. *BioScience* **1986**, 36 670-672.

Carlisle, D.M. and W.H. Clements, Growth and secondary production of aquatic insects along a gradient of Zn contamination in Rocky Mountain streams. *J. N. Am. Benthol. Soc.* **2003**, 22 (4), 582-597.

Carpenter, S.R., Microcosm experiments have limited relevance for community and ecosystem ecology. *Ecology* **1996**, 77 (3), 677-680.

Chadwick, J.W., S.P. Canton, and R.L. Dent, Recovery of benthic invertebrate communities in Silver Bow Creek, Montana, following improved metal mine wastewater treatment. *Water, Air, and Soil Pollution* **1986**, 28 427-438.

Clements, W.H., Small-scale experiments support causal relationships between metal contamination and macroinvertebrate community responses. *Ecological Applications* **2004**, 14 954-967.

Clements, W.H., P. Cadmus, and S.F. Brinkman, Responses of aquatic Insects to Cu and Zn in stream microcosms: understanding differences between single species tests and field responses. *Environmental Science & Technology* **2013**, 47 (13), 7506−7513.

Clements, W.H., D.M. Carlisle, J.M. Lazorchak, and P.C. Johnson, Heavy metals structure benthic communities in Colorado mountain streams. *Ecological Applications* **2000**, 10 626-638.

Clements, W. H. and C. Kotalik, Effects of major ions on natural benthic communities: an experimental assessment of the US Environmental Protection Agency aquatic life benchmark for conductivity. *Freshwater Science* **2016**, 35 (1), 126-138.

Clements, W.H. and J.R. Rohr, Community responses to contaminants: using basic ecological principles to predict ecotoxicological effects. *Environ. Toxicol. Chem.* **2009**, 28 (9), 1789-1800.

Clements, W.H., N.K.M. Vieira, and S.E. Church, Quantifying restoration success and recovery in a metal-polluted stream: a 17-year assessment of physicochemical and biological responses. *Journal of Applied Ecology* **2010**, 47 (4), 899-910.

Cormier, S. M. and G. W. Suter II, A method for deriving water-quality benchmarks using field data. *Environmental Toxicology and Chemistry* **2013**, 32 (2), 255-262.

Courtney, L.A. and W.H. Clements, Assessing the influence of water and substratum quality on benthic macroinvertebrate communities in a metal-polluted stream: an experimental approach. *Freshwater Biology* **2002**, 47 (9), 1766-1778.

DeNicola, D. M. and M. G. Stapleton, Impact of acid mine drainage on benthic communities in streams: the relative roles of substratum vs. aqueous effects. *Environmental Pollution* **2002**, 119 (3), 303-315.

Dsa, J.V., K.S. Johnson, D. Lopez, C. Kanuckel, and J. Tumlinson, Residual toxicity of acid mine drainage-contaminated sediment to stream macroinvertebrates: relative contribution of acidity vs. metals. *Water Air and Soil Pollution* **2008**, 194 (1-4), 185-197.

Gerhardt, A., Effects of subacute doses of iron (Fe) on Leptophlebia marginata (Insecta, Ephemeroptera). *Freshwater Biology* **1992**, 27 (1), 79-84.

Hogsden, K.L. and J.S. Harding, Consequences of acid mine drainage for the structure and function of benthic stream communities: a review. *Freshwater Science* **2012**, 31 (1), 108-120.

Hornberger, M.I., S.N. Luoma, M.L. Johnson, and M. Holyoak, Influence of remediation in a mine-impacted river: metal trends over large spatial and temporal scales. *Ecological Applications* **2009**, 19 (6), 1522-1535.

Iwasaki, Y., T. Kagaya, K. Miyamoto, and H. Matsuda, Effects of heavy metals on riverine benthic macroinvertebrate assemblages with reference to potential food availability for drift-feeding fishes. *Environ. Toxicol. Chem.* **2009**, 28 (2), 354-363.

Linton, T.K., M.A.W. Pacheco, D.O. McIntyre, W.H. Clement, and J. Goodrich-Mahoney, Development of bioassessment-based benchmarks for iron. *Environ. Toxicol. Chem.* **2007**, 26 (6), 1291-1298.

MacDonald, D.D., C.G. Ingersoll, and T.A. Berger, Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. *Archives of Environmental Contamination and Toxicology* **2000**, 39 20-31.

Maret, T.R., D.J. Cain, D.E. Maccoy, and T.M. Short, Response of benthic invertebrate assemblages to metal exposure and bioaccumulation associated with hard-rock mining in northwestern streams, USA. *J. N. Am. Benthol. Soc.* **2003**, 22 (4), 598-620.

McKnight, D.M. and G.L. Feder, The ecological effect of acid conditions and precipitation of hydrous metal-oxides in a rocky-mountain stream. *Hydrobiologia* **1984**, 119 (2), 129-138.

Merritt, R.W., K.W. Cummins, and M.B. Berg, eds. *An introduction to the aquatic insects of North America*. 4th ed. 2008, Kendall Hunt Publishing: Debuque, Iowa. 1158.

Niyogi, D.K., W.M. Lewis, and D.M. McKnight, Litter breakdown in mountain streams affected by mine drainage: Biotic mediation of abiotic controls. *Ecological Applications* **2001**, 11 (2), 506-516.

Niyogi, D.K., W.M. Lewis, and D.M. McKnight, Effects of stress from mine drainage on diversity, biomass, and function of primary producers in mountain streams. *Ecosystems* **2002**, 5 (6), 554-567.

Niyogi, D.K., D.M. McKnight, and W.M. Lewis, Influences of water and substrate quality for periphyton in a montane stream affected by acid mine drainage. *Limnology and Oceanography* **1999**, 44 (3), 804-809.

Pacheco, M.A.W., D.O. McIntyre, and T.K. Linton, Integrating chemical and biological criteria. *Environ. Toxicol. Chem.* **2005**, 24 (11), 2983-2991.

Rader, R.B., A functional classification of the drift: traits that influence invertebrate availability to salmonids. *Canadian Journal of Fisheries and Aquatic Sciences* **1997**, 54 (6), 1211-1234.

Sasaki, A., A. Ito, J. Aizawa, and T. Umita, Influence of water and sediment quality on benthic biota in an acidified river. *Water Research* **2005**, 39 (12), 2517-2526.

Schindler, D.W., Replication versus realism: The need for ecosystem-scale experiments. *Ecosystems* **1998**, 1 (4), 323-334.

Schmidt, T.S., W.H. Clements, K.A. Mitchell, S.E. Church, R.B. Wanty, D.L. Fey, P.L. Verplanck, and C.A. San Juan, Development of a new toxic-unit model for the bioassessment of metals in streams. *Environmental toxicology and chemistry* **2010**, 29 (11), 2432-2442.

Schmidt, T.S., W.H. Clements, R.B. Wanty, P.L. Verplanck, S.E. Church, C.A. San Juan, D.L. Fey, B.W. Rockwell, E.H. DeWitt, and T.L. Klein, Geologic processes influence the effects of mining on aquatic ecosystems. *Ecological Applications* **2012**, 22 (3), 870-879.

Soucek, D.J., D.S. Cherry, and G.C. Trent, Relative acute toxicity of acid mine drainage water column and sediments to Daphnia magna in the Puckett's Creek watershed, Virginia, USA. *Archives of Environmental Contamination and Toxicology* **2000**, 38 (3), 305-310.

Vuori, K.M., Direct and indirect effects of iron on river ecosystems. *Annales Zoologici Fennici* **1995**, 32 (3), 317-329.

Walter, C.A., D. Nelson, and J.I. Earle, Assessment of stream restoration: sources of variation in macroinvertebrate recovery throughout an 11-year study of coal mine drainage treatment. *Restoration Ecology* **2012**, 20 (4), 431-440.

Walton-Day, K., R.L. Runkel, and B.A. Kimball, Using spatially detailed water-quality data and solute-transport modeling to support total maximum daily load development. *Journal of the American Water Resources Association* **2012**, 48 (5), 949-969.

Waters, T.F., The drift of stream insects. *Annual Review of Entomology* **1972**, 17 (1), 253-272.

CHAPTER 2

STRUCTURAL AND FUNCTIONAL RESPONSES OF PERIPHYTON AND MACROINVERTEBRATE COMMUNITIES TO FERRIC FE, CU AND ZN IN STREAM MESOCOSMS[2]

Introduction

Iron (III) oxides often occur in streams associated with historical mining activity under pH neutral and well-oxygenated conditions. Unlike the more directly toxic ferrous (Fe^{2+}) forms of iron (Fe) typical of low pH river systems, traditional laboratory toxicity tests have found ferric iron (Fe^{3+}) to be relatively harmless to aquatic organisms (Boutet and Chaisemartin 1973, Smith and Sykora 1976, Randall et al. 1992, Van Dam et al. 1998, Biesinger and Christensen 1972). However, river systems polluted with even low concentrations of Fe (III) precipitates are often characterized by reduced biomass and production of periphyton as well as decreased benthic invertebrate richness and abundance (Sode 1983, McKnight and Fender 1984, Rasmussen and Lindengaard 1988, Vuiri 1995, Wellnitz and Sheldon 1995, DeNicola and Stapleton 2002, Guasch et al. 2012). Field-based derivations of regional benchmarks using quantile regression found total Fe concentrations of 0.21 mg/L would be protective of all taxonomic groups considered, while 1.74 mg/L would result in the loss of sensitive species but should maintain ecosystem functions (Linton et al. 2007). These values bracket the USEPA national criterion value of 1.0 mg/L (total recoverable), which was based largely on the presence or absence of fish across a gradient of ferric Fe in a small Colorado (USA) stream (USEPA 1976). Although more

2 After approval from my committee a revised version of this chapter was published by Environmental Toxicology and Chemistry as: Cadmus P, Guasch H, Herdrich AT, Bonet B, Urrea G, Clements WH. (2018) Structural and functional responses of periphyton and macroinvertebrate communities to ferric Fe, Cu, and Zn in stream mesocosms. *Environ Toxicol Chem.* 37:1320-1329. doi: 10.1002/etc.4070.

ecologically realistic, the development of water quality criteria based exclusively on field studies is challenging because Fe commonly co-occurs with other toxic metals and interaction effects are difficult to predict. Experiments that examine potential interactions between ferric Fe and other toxic metals are limited.

Single species tests that measure the toxicity of ferric Fe have been conducted almost exclusively using organisms that are either well adapted to fine sediments (e.g., chironomids) or free swimming organisms such as fish (Smith and Sykora 1976) and cladocerans (Randall et al. 1992, Van Dam et al. 1998, Biesinger and Christensen 1972). Few studies have investigated effects on benthic organisms. In neutral pH waters Fe precipitates as iron oxides, which clogs interstitial spaces of benthic habitat, smothers periphyton and reduces available light for photosynthesis (Vuori 1995). Alterations in the structure and function of benthic communities as a result of these indirect effects are also not considered in single species toxicity tests (Cairns 1983). If indirect effects of Fe induce loss of primary producers or consumers, alterations in community structure might eventually impact pelagic organisms that were thought to be tolerant. Finally, the limited duration of many laboratory toxicity tests is believed to be a key reason why these experiments often underestimate the sensitivity of aquatic insects (Buchwalter et al. 2007, Clements et al. 2013). Unlike aqueous toxicants, the indirect effects of Fe are observed only after deposition has embedded substrate, so even chronic tests (~30 d) may not be of a sufficient duration for these physical effects to materialize.

Mesocosm experiments using natural macroinvertebrate and periphyton communities provide an opportunity to assess the direct and indirect effects of contaminants across levels of biological organization. Previous studies have shown that natural communities of macroinvertebrates respond to metals and other toxicants at lower concentrations than single

species trials (Clements et al. 2013, Clements 2004, Clements and Kotalik 2016). Mesocosm experiments conducted with natural benthic communities include rare taxa, sensitive taxa and early life stages which are difficult to obtain or culture for single species trials. Grazing insects are especially sensitive to metals because they are dependent on healthy periphyton for food and have been found to accumulate high concentrations of metals through their diet (Irving et al. 2003, Xie and Buchwalter 2011, Xie et al. 2010). Dietary exposure to metals is often ignored in laboratory toxicity tests, but routinely occurs in community mesocosm experiments.

To examine the effects of Fe-oxide precipitates I exposed natural benthic invertebrate communities to a gradient of ferric Fe using experimental stream mesocosms under flow through conditions (Clements 2004). I predicted this exposure system would better simulate natural conditions in which Fe concentrations in the water column remain constant while interstitial spaces of benthic habitat become clogged with Fe floc. The primary goal of this experiment was to determine levels of Fe that are protective of benthic invertebrates common in highly oxygenated mountain streams with circumneutral pH. I assessed Fe deposition, community respiration/metabolism, abundance, diversity and drift of invertebrates.

The discrepancy between field and laboratory characterizations of Fe toxicity could be explained by interactions between ferric Fe and other metals co-occurring at mine impacted sites. However, most laboratory experiments have found Fe to have an antagonistic effect on aqueous metal toxicity (Sunda et al. 1981, Foster and Morel 1982, Stauber and Florence 1985, Sunda et al. 1987, Hare et al. 1991, Gerhardt 1994, Gerhardt and Westermann 1995). Because Fe (III) oxides readily adsorb to metals such as Cu, Cd and Zn (Davis et al. 1978, Millward 1980, Millward and Moore 1982), it is hypothesized that Fe precipitates have a protective effect against metal toxicity by reducing bioavailability. I tested this hypothesis in a second mesocosm

experiment by measuring the combined effects of ferric Fe and metals on structural and functional endpoints including antioxidant enzyme activities (AEA), algal growth, protein content, biomass, community composition and photosynthesis. The goal of this study was to assess the toxicity of ferric Fe below the USEPA benchmark of 1.0 mg/L total recoverable Fe and to investigate the influence of aqueous metals that often co-occur with ferric Fe. Because water quality standards generally assume an additive relationship among contaminants, I also tested for an interaction between ferric Fe and other metals.

Materials and Methods

Mesocosm experiment 1: responses to Fe

The goal of mesocosm experiment 1 was to quantify the direct and indirect effects of ferric Fe on stream benthic communities. Uniform amounts of small cobble and pebbles were placed in 10 x 10 x 6 cm colonization trays that were deployed for 32 d in the South Fork of the Michigan River, a stream originating from a wilderness area in the Routt National Forest (Colorado, USA). Trays were randomly assigned to 18 coolers and transported to experimental streams at the Colorado State University Stream Research Laboratory (SRL) in Fort Collins, Colorado, USA (Clements 2004). The experimental streams receive untreated water from a mountain reservoir with water quality and temperature typical of regional reference streams in Colorado. Water was circulated by paddle wheels to maintain current velocities similar to high gradient mountain streams. On 17 Oct. 2010, 18 stream mesocosms were randomly assigned in triplicate to one of 6 ferric Fe treatments (0, 0.4, 1.0, 2.5, 6.2 and 15.6 mg/L). Peristaltic pumps were calibrated to deliver 10 ml/min of stock solution to streams receiving 1 L/min flow. To convert all Fe (II) to Fe (III), stock solutions in 20 L carboys for each stream were thoroughly

stirred and aerated, while NaOH was added until a pH > 6.5 was observed. After 60 minutes of stirring, pH was reassessed before using stock solutions. To maintain a homogenous suspension of Fe (III) oxide precipitates in the carboys, stock solutions were vigorously aerated throughout the experiment. On Day 2 of exposure, nets (100 μm mesh) were placed in each mesocosm to capture drifting organisms as previously described (Clements 2004). After 24 h the nets were removed and organisms were preserved in 80% ethanol. To assess deposition of Fe oxides, 50 ml beakers containing 20 ml of glass beads (6 mm diameter) were placed in each stream mesocosm (Appendix B, Figure 2.S1). Fe was allowed to settle in the interstitial spaces of the beads for 96 h, after which the contents were transferred to polypropylene jars and rinsed with ultrapure water. Samples were passed through a 5 mm sieve (to remove beads) and filtered through a vacuum funnel containing a pre-weighed 10 μm glass fiber filter (GC-50, Sterlitech Corp. Kent, WA, USA). The filter and Fe oxide precipitates were then dried and weighed.

After 10 d I measured effects of Fe on community metabolism by transferring 2 of the 4 trays from each experimental stream to transparent and opaque airtight chambers that received recirculating flow. The chambers remained immersed within the experimental streams to maintain a constant temperature. Measures of dissolved oxygen (DO) in transparent chambers reflect respiration and photosynthesis, whereas measures of DO in the opaque chambers reflect autotrophic and heterotrophic respiration. I assumed that differences between light and dark measures would provide an estimate of whole community metabolism (Hall et al. 1987, Kashian et al. 2004). Respiration measurements were taken after all streams had equilibrated to ensure turbidity did not affect photosynthetic rates. Following assessment of community metabolism, trays were removed from the mesocosms and substrate and experimental streams were gently scrubbed to remove surviving organisms. All macroinvertebrates retained in a 355 μm sieve

were preserved in 80% ethanol. Organisms were enumerated and most were identified to genus (Merritt et al. 2008, Ward et al. 2002). Chironomids were identified to tribe (Wiederholm 1983).

Mesocosm experiment 2: responses to Fe and other metals

The goal of mesocosm experiment 2 was to quantify the effects of Fe on the toxicity of other metals (Cu, Zn) to periphyton and benthic macroinvertebrate communities. Periphyton communities were naturally colonized on 5 cm^2 unglazed porcelain tiles placed in each of the 18 flow-through stream mesocosms. After 23 d (July 2011), natural benthic macroinvertebrate communities for each stream mesocosm were obtained by combining two core samples (10 cm diameter x 8 cm depth) and four Surber samples (0.1 m^2; 400 µm mesh) collected from Joe Wright Creek, a tributary of the Cache La Poudre River, Colorado, USA. Benthic communities were transported and randomly assigned to the 18 stream mesocosms. Peristaltic pumps were calibrated to deliver 10 ml/min of stock solution to streams as described for mesocosm experiment 1. Using a 2 x 3 factorial design (n=3) the presence (Fe) or absence (NoFe) of Fe at a target concentration of 0.6 mg/L was crossed with 3 levels of Cu and Zn: control, low metals (LM: 0.01 mg/L Cu and 0.1 mg/L Zn) and high metals (HM: 0.05 mg/L Cu and 0.5 mg/L Zn). Target concentrations in LM streams bracketed hardness adjusted acute and chronic criteria values for Zn (acute and chronic: 0.049 mg/L; USEPA 1996) and Cu (acute: 0.016 mg/L, chronic: 0.005 mg/L; USEPA 2007). HM treatment levels were within the range observed in Colorado drainages altered by mining activity (Schmidt et al. 2010, Clements 2000). Stock solutions were prepared by adding, $FeCl_2$ x 6 H_2O and concentrated solutions of $ZnSO_4$ x 7 H_2O and $CuSO_4$ x 6 H_2O to 18 L carboys. Stock solutions were neutralized as described above.

Algal biomass (as chlorophyll-*a* fluorescence, F_o) and maximal quantum yield (Y_{max}), the fraction of absorbed quanta used for photosystem II photochemistry, were measured before (0, 2, 4, 6, 12 and 24 d before) and during (1, 3, 5, 7 and 9 d) exposure to metals as indicators of photosynthetic performance (Corcoll et al. 2011). Three colonized tiles from each experimental stream were removed and analyzed for chlorophyll-*a* in vivo fluorescence using a Pulse Amplitude Modulated (PAM) fluorometer (Heinz Walz GmbH, Effeltrich, Germany). Y_{max} was calculated from the fluorescence signal recorded at 665 nm and reported as relative units of fluorescence (Genty et al. 1989).

Antioxidant enzyme activities (AEA) and protein content of periphyton, considered indicators of oxidative stress and an early warning system for detecting effects on autotrophic communities, were measured after 7 and 9 days of exposure. Because samples taken on days 7 and 9 did not differ statistically, these measures were pooled and used to represent effects after 1 week of exposure. Periphyton was scraped from tiles and placed in cryogenic tubes, flash frozen in liquid nitrogen and stored at -80°C until analysis. Prior to protein extraction samples were thawed, centrifuged at 2,300 *g* for 5 min at 10°C to remove excess water and then weighed (wet weight). Protein extraction and AEA measurements were performed using homogenization and disruption with glass beads Bonnineau et al. 2011). The optimal concentration of substrate or cofactor was determined by testing the concentration of H_2O_2 for the catalase (CAT) and ascorbate peroxidase (APX) assays and nicotinamide adenine dinucleotide phosphate (NADPH) was used for the glutathione reductase (GR) assay. CAT (Aebi 1984), APX (Nakano and Asada 1981) and GR (Schaedle and Bassham 1977) were assessed spectrophotometrically and were calculated as μmol H_2O_2 mg protein^{-1} min^{-1} (extinction coefficient, ε: 0.039 M cm^{-1}), μmol ascorbate mg protein^{-1} min^{-1} (extinction coefficient, ε: 2.8 M cm^{-1}), and μmol NADPH min^{-1} μg^{-1}

of protein, respectively (See Appendix C Supplemental Methods for a detailed narrative of AEA extraction and assessment).

To assess effects of Fe and metals on the structure and composition of periphyton communities, an algal suspension (4 ml) from each colonized tile was collected for taxonomic analysis of diatoms. Organic matter was removed using 0.5 ml hydrochloric acid and 20 ml of H_2O_2 (40%) to obtain clean frustules. Frustules were then washed, desiccated, and mounted in Naphrax. Algal cells were subsampled, identified and enumerated. A subsample of periphyton from each mesocosm was dried, weighed and digested in HNO_3 and H_2O_2 for metals analysis (see details below).

After a 10 d exposure to metals all trays from each experimental stream were placed in transparent respiration chambers to measure community metabolism as described above. Dark conditions were created by covering chambers with opaque plastic sheeting to assess respiration. Macroinvertebrate communities on trays were sampled and processed as described in mesocosm experiment 1.

For both mesocosm experiments temperature, pH, dissolved oxygen and conductivity were measured every 2 days using YSI 63 and 550a hand-held meters (YSI Inc. Yellow Springs, Ohio, USA). Alkalinity and hardness were assessed using EPA methods 200.7 and 310.7. Additionally, pH was assessed immediately before and after changing stock solutions to ensure all streams maintained similar neutral values. Total (unfiltered) and dissolved (0.45 µm filter) metal concentrations were analyzed by atomic absorption using a Varian SpectrAA 110 FS furnace with deuterium background correction for Cu and an Instrumentation Laboratory Video 22 with Smith-Hieftje background correction for Zn and Fe. Phosphate concentrations were assessed per Murphy and Riley (1962).

Data analysis

Effects of Fe on macroinvertebrate communities, macroinvertebrate drift and community metabolism in mesocosm experiment 1 were determined using one-way ANOVA (Proc ANOVA, SAS 9.3). Data were log transformed to achieve normality and homogeneity of variance when required. If the overall F-statistic was significant ($p < 0.10$) I used Duncan's multiple range test to determine differences among treatments. I used an alpha-value of 0.10 to protect against type II errors because of the limited number of replicate mesocosms and expected high within treatment variation. Treatment effects of Fe and metals in mesocosm experiment 2 were analyzed using two-way ANOVA. Interaction terms were analyzed only when the overall model and the main effects were significant. Community metabolism in experiment 2 was analyzed using general linear models (Proc GLM) because sample sizes were unbalanced due to an equipment failure in one of the 18 stream mesocosms.

Changes in periphyton biomass during initial colonization (22 d) and metals exposure (9 d) were fit to a 3-parameter log-normal curve using SigmaPlot v11 (Appendix B, Figure 2.S2). I tested for differences between the estimated parameters using the following equation,

$$y = \left(\frac{a}{x}\right)^{-0.5\left[\frac{\ln(x/x_o)}{b}\right]^2}$$

where y = periphyton biomass, $\boldsymbol{a}$ is associated with the peak of the curve, **b** is the rate of inhibition after the peak, $\mathbf{X_o}$ is the time required to reach the maximum or peak value and **X** is time in days.

Results

Mesocosm experiment 1: Effects of ferric Fe

Except for specific conductance, which increased with metal concentration, water quality characteristics in the 2 mesocosm experiments were consistent across treatments and typical of conditions observed in low order mountain streams throughout Colorado (Supplemental Table 2.S1). Phosphate levels were below the detection limit (0.01 mg/L) and previous mesocosm studies conducted at the SRL showed that concentrations of dissolved organic carbon (2.5 mg/L), sulfate (5.6 mg/L), potassium (0.7 mg/L), chloride (1.9 mg/L) and sulfide (0.1 mg/L) were consistently low in diluent water Naddy et al. 2007). Total Fe in mesocosm experiment 1 increased with the treatment levels and approximated target concentrations (Table 2.1), while dissolved Fe remained below detection limits (0.1 mg/L). Fe oxide deposition was observed in all treated mesocosms and was significantly correlated with total Fe measured in the water column (R^2=0.81; Appendix B, Figure 2.S1).

Nine dominant taxa accounted for 81% of the total macroinvertebrate community in stream mesocosms (Figure 2.1). Exposure to ferric Fe altered the composition of benthic communities and reduced abundance of most taxa. Although there was considerable variation in sensitivity among groups, some taxa were affected by Fe at relatively low concentrations. In particular, the mayfly *Epeorus* sp. and the chironomid tribe Tanytarsini were highly sensitive to Fe and significantly reduced at 0.4 and 1.0 mg Fe/L, respectively. Total abundance of the 4 major macroinvertebrate orders (Ephemeroptera, Plecoptera, Trichoptera and Diptera) followed similar trends and showed strong concentration-response relationships with Fe, but effects differed among groups (Appendix B, Figure 2.S3). In particular, total abundance of Diptera was highly sensitive to Fe exposure and was significantly reduced at the lowest Fe concentration.

Community metabolism was significantly reduced in all treatments greater than 0.4 mg/L Fe, but macroinvertebrate drift increased only at the highest Fe level (Figure 2.2).

Mesocosm Experiment 2: Interactions between Fe and other metals

Mean concentrations of total Fe measured in treated streams during experiment 2 ranged from 0.60-0.69 mg/L. Measured concentrations of dissolved Cu (LM = 0.01 mg/L; HM = 0.03 mg/L) and Zn (LM = 0.08 mg/L; HM = 0.52 mg/L) approximated target concentrations and were similar between Fe and NoFe treatments (Table 2.2).

Concentrations of Cu and Zn in periphyton increased with exposure level ($p<0.001$), but metal concentrations in biofilm were also increased by the presence of Fe ($p<0.001$; Figure 2.3). The significant Fe x metal interaction term for Cu indicates that these effects of Fe on metal uptake increased with metal treatment. Community metabolism was also significantly reduced by exposure to Fe ($p= 0.0004$) but not metals (Supplemental Table 2.S2). Across all metal treatments, community metabolism was approximately 40% lower in Fe-treated streams compared to NoFe streams.

Dissolved metals and Fe significantly affected periphyton biomass, growth, photosynthetic activity, and community metabolism, but responses differed between the Fe and metals treatments. The number of days required for periphyton to reach the maximum biomass (X_o) was significantly greater for the metals ($p=0.019$) and Fe ($p=0.048$) treatments compared to controls (Appendix B, Figure 2.S4). PAM fluorimetry measurements revealed negative effects of metal exposure on the photosynthetic performance of periphyton (Y_{max}) and on periphyton biomass (F_o) within 2-4 days of exposure (Figure 2.4). Effects of metals on biomass were generally greater than effects on photosynthetic inhibition. Periphyton biomass was 61% lower

in HM treatments compared to controls after 7 d. In contrast to Cu and Zn treatments, Fe had relatively little effect on periphyton biomass or photosynthetic performance.

Exposure to metals and Fe significantly reduced the protein content of periphyton, but measures of antioxidant enzyme activity (AEA) showed relatively little variation among treatments (Figure 2.5; Appendix B, Table 2.S2). Effects of metals on protein content were strongly influenced by the presence of Fe, as indicated by a highly significant Fe x treatment interaction. The presence of Fe increased effects of metals in the LM treatments, but reduced effects in the HM treatment. Although glutathione reductase increased in response to Fe and metals, these differences were not statistically significant. Unlike these physiological and functional measures, Fe and metal treatments had no effect on diversity or species richness of periphyton (Appendix B, Figure 2.S5). However, low abundance and high variability of most species across experimental units made discerning treatment effects difficult.

Effects on macroinvertebrates

Of the six dominant macroinvertebrate taxa in mesocosm experiment 2, the heptageniid mayfly *Cinygmula* ($p<0.0001$) and the chironomid tribes Tanytarsini ($p<0.0001$) and Orthocladiini ($p=0.0422$) were especially sensitive to metals at both the low and high metal treatment levels (Figure 2.6). However, *Cinygmula* was the only taxa significantly reduced by Fe exposure. Exposure to Zn and Cu significantly reduced total macroinvertebrate abundance ($p<0.0001$) and abundance of mayflies (Ephemeroptera; $p<0.0001$) and dipterans ($p<0.0001$); however, Ephemeroptera was only group significantly reduced by Fe ($p = 0.0992$; Appendix B, Figure 2.S5).

Discussion

I observed significant alterations in the structure and function of benthic communities after 10 d of exposure to ferric Fe in stream mesocosms. Effects of Fe on some endpoints (e.g., abundance of grazing mayflies and chironomids, physiological responses, community metabolism) were observed at concentrations either at or below the existing chronic water quality criterion for Fe (1.0 mg/L). These results suggest that the current criterion value for Fe may be under-protective for some components of stream benthic communities. I hypothesize that the effects of Fe on primary producers and grazers were a result of increased deposition of Fe oxides and the loss of interstitial space and not due to direct Fe toxicity. The highly oxygenated and neutral pH conditions of stream mesocosms ensured that dissolved Fe was negligible and similar in control and treated mesocosms. The negative effects of Fe oxides, which are generally not considered in laboratory toxicity tests, were consistent with field observations showing loss of interstitial space, lowered algal biomass, reduced primary production and altered macroinvertebrate community structure in streams polluted with Fe (III) precipitates (Sode 1983, Guasch et al. 2012). A better understanding of these indirect physical effects of Fe oxides would improve the ability to derive ecologically relevant and protective water quality standards for Fe.

Previous single species toxicity tests suggest that ferric Fe is either non-toxic Boutet and Chaisemartin 1973, Smith and Sykora 1976, Randall et al. 1992, Van Dam et al. 1998, Biesinger and Christensen 1972 or has a protective (antagonistic) effect on the toxicity of other metals (Sunda et al. 1981, Foster and Morel 1982, Stauber and Florence 1985, Sunda et al. 1987, Hare et al. 1991, Gerhardt 1994, Gerhardt and Westermann 1995). Herewithin, we observed only weak evidence for interactive effects between aqueous metals and ferric Fe on benthic stream communities. An important exception was for Cu and Zn accumulation in periphyton, which

increased significantly in the presence of ferric Fe. Greater concentrations of metals may have resulted from adsorption to Fe floc that was later accumulated by periphyton; however, we did not find evidence that Cu or Zn were bound to Fe when we compared concentrations of dissolved and total metals.

Macroinvertebrate drift was relatively insensitive to Fe exposure. Drift behavior was measured only on the second day of exposure, before interstitial spaces of substrate became clogged. This is likely why effects were only observed in the highest treatment level. The dramatic effects of Fe on heptageniid mayflies were expected, as these organisms are obligate grazers and therefore would likely be the first organisms to respond to reduced primary production. Compared to other grazing mayflies (e.g., Baetidae*)*, heptageniids are also considerably less mobile (Rader 1997), which reduces their ability to avoid habitats impacted by Fe precipitates. In contrast, some of the effects of Cu and Zn on macroinvertebrates observed in this study were likely a result of dietary exposure. Diet is known to be an important pathway of metal accumulation (Irving et al. 2003, Xie and Buchwalter 2011, Xie et al. 2010) and feeding rates of grazing mayflies are often reduced when fed algae that has accumulated trace metals (Irving et al. 2003). Even if Fe oxide reduces bioavailability of Cu or Zn to gill structures as laboratory experiments suggest, grazing mayflies would likely experience increased accumulation in proportion to metal concentrations in periphyton. Although dietary exposure to metals is rarely considered in derivation of water quality standards, these results suggest it should be included in the development of predictive models.

Although Fe alone had a significant effect on several macroinvertebrate taxa in both experiments, there were important differences between the 2 mesocosm experiments. These differences may be partially explained by the differences in experimental designs and the low Fe

concentrations used in experiment 2, which were at the lower end of the concentration gradient used in experiment 1. However, some taxa showed different responses between the experiments even at similar Fe concentrations. For example, the mayfly *Cinygmula* was unaffected by high levels of Fe (> 2.5 mg/L) in experiment 1, but was reduced by 24% at much lower concentrations (0.6 mg/L) in experiment 2. Conversely, Tanytarsini chironomids were highly sensitive to low concentrations of Fe in experiment 1, but were unaffected in experiment 2 (although they were highly sensitive to Cu and Zn). One unique advantage of the approach used in these mesocosm experiments is that it allows for inclusion of a wide range of community types, species and life stages. By comparing responses of communities collected from different locations or at different times of year, we can begin to explain some of the variation often observed in field studies. Because communities used in these mesocosm experiments were collected from 2 different streams and at different times of year (October versus July), we expected to see differences in responses based on insect phenology. Previous mesocosm experiments ((Kiffney and Clements 1996, Clements et al. 2013) , laboratory toxicity tests (Clements et al. 2012) and field studies (Clark and Clements 2006) have shown that early life stages of aquatic insects are especially sensitive to metals and other contaminants (see Chapter 3). Failure to account for the extreme sensitivity of these early instars may help explain the discrepancies between results of laboratory toxicity tests and field surveys of aquatic insects (Clements et al. 2013). Because of differences in community composition among sites and interspecific differences in their development and phenology, different species and life stages of aquatic insects will be present at a site at any particular point in time, resulting in significant differences in observed community-level effects. An understanding of these context-dependent responses to contaminants (Clements et al. 2016,

Clements et al. 2012), which could significantly improve the ability to predict effects in the field, can only be obtained by using targeted and controlled experiments.

Previous investigators have hypothesized that because of ecological redundancy, a large proportion of biodiversity must be lost before functional changes in ecosystems are observed (Ehrlich and Ehrlich 1981, Walker 1992), and meta-analyses conducted with periphyton communities support this hypothesis (Cardinale et al. 2011). This study found highly significant effects of Cu and Zn on physiological and functional endpoints before structural changes in periphyton communities, which showed little response to Fe or metals. In agreement with previous investigations (Cadmus et al. 2016), adverse effects of metals on photosynthetic performance were observed after two days of exposure, followed by biomass reduction after four days. Sublethal responses such as growth rate, photosynthetic activity, protein content, or community metabolism may be more meaningful measures of Fe effects on periphyton than alterations in community structure. We expected to see greater effects of metals in the LM and FeLM treatments on periphyton community composition as the measured Cu concentration exceeded the toxicity threshold for periphyton (0.009 mg/L) observed in previous studies Guasch et al. 2004). We have observed that the presence of Zn can reduce the toxicity of Cu on macroinvertebrate communities (Clements 2004), and it is possible that a similar protective effect of Zn occurred in the present study. Interestingly, my inability to detect an effect of metals on periphyton community structure was in contrast to the responses of macroinvertebrate communities, which were highly sensitive to both metals and ferric Fe.

Experiments investigating effects of metal mixtures on benthic communities are surprisingly limited in the literature (Clements et al. 2013, Cadmus et al. 2016), despite the fact that aquatic insects often dominate stream ecosystems and most mining polluted streams are

impacted by multiple metals. This study demonstrated additional risk to aquatic insects and periphyton in metal mixtures that included ferric Fe. Although significant deposition of Fe oxides was observed in these mesocosm studies, it is unlikely that a 10 d experiments adequately reproduced the extreme effects observed at mining-impacted sites in the field. Mesocosm or field experiments conducted for longer duration would likely result in more severe alterations in benthic habitat. Regardless, we observed significant alterations in the structure and function of stream benthic communities at Fe concentrations near or below the national standard of 1.0 mg/L. I am confident that these responses were primarily a result of indirect physical effects associated with Fe oxide deposition rather than direct toxicity. Mesocosm experiments conducted using natural benthic communities provide a unique opportunity to quantify the relative importance of these indirect physical effects and to develop a better understanding of the relationship between basal resources and consumers in natural stream ecosystems. An approach that integrates mesocosm experiments, field assessments and standard laboratory toxicity tests could increase the reliability of water quality criteria and may help explain the discrepancies between laboratory and field responses to contaminants often reported in the literature.

Figures and Tables from Chapter 2

Contains:

Figures 2.1 to 2.6

Tables 2.1 to 2.2

See Appendix B for Supplemental Figures 2.S1 to 2.S5, Supplemental Tables 2.S1 to 2.S2 and Supplemental Methods Narrative for Chapter 2.

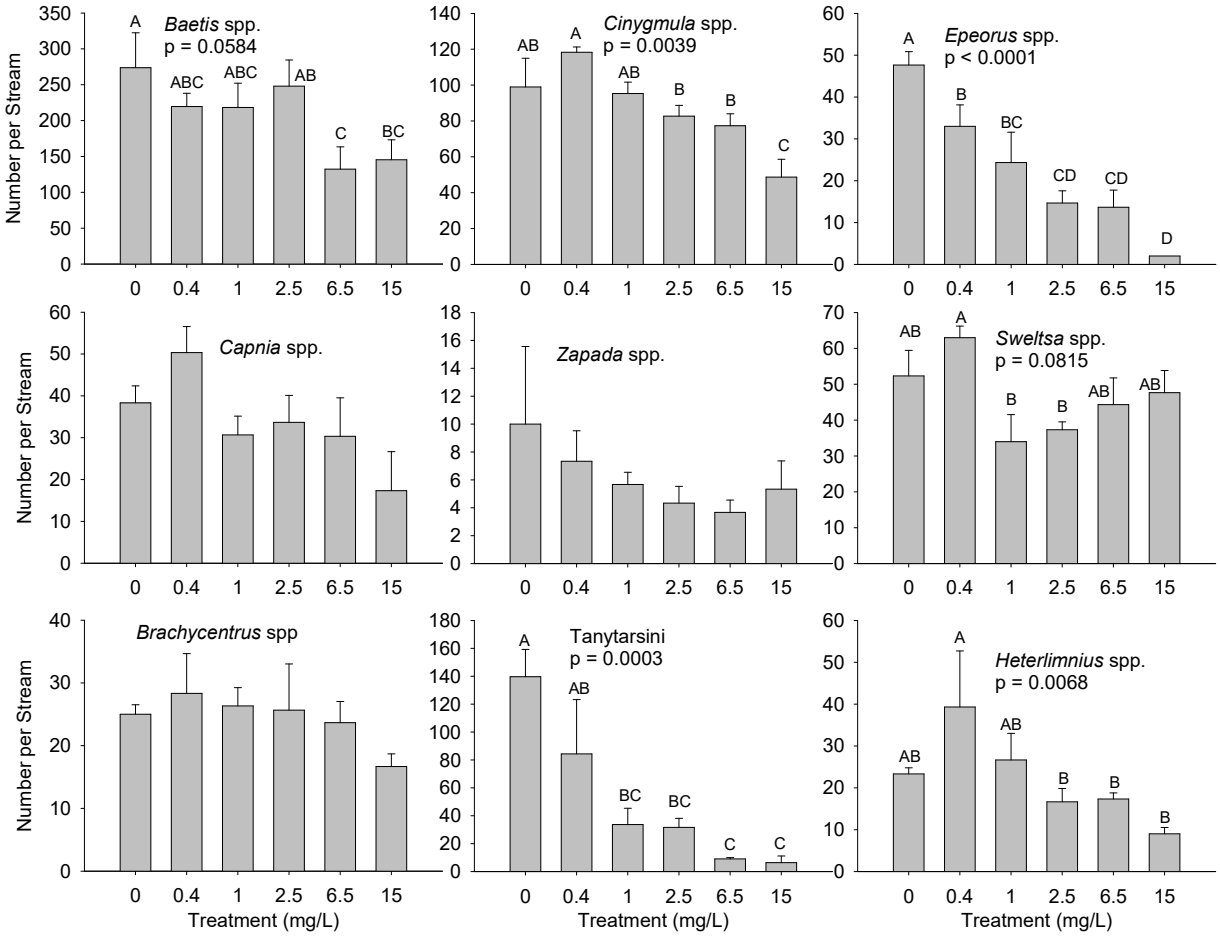


Figure 2.1 – Mean (+ s.e.) abundance of the 9 dominant macroinvertebrate taxa from mesocosm experiment 1 after 10 days of exposure to ferric Fe. Results of Duncan's multiple range test are shown for all significant responses ($p < 0.10$). Treatments with the same letter were not significantly different from each other.

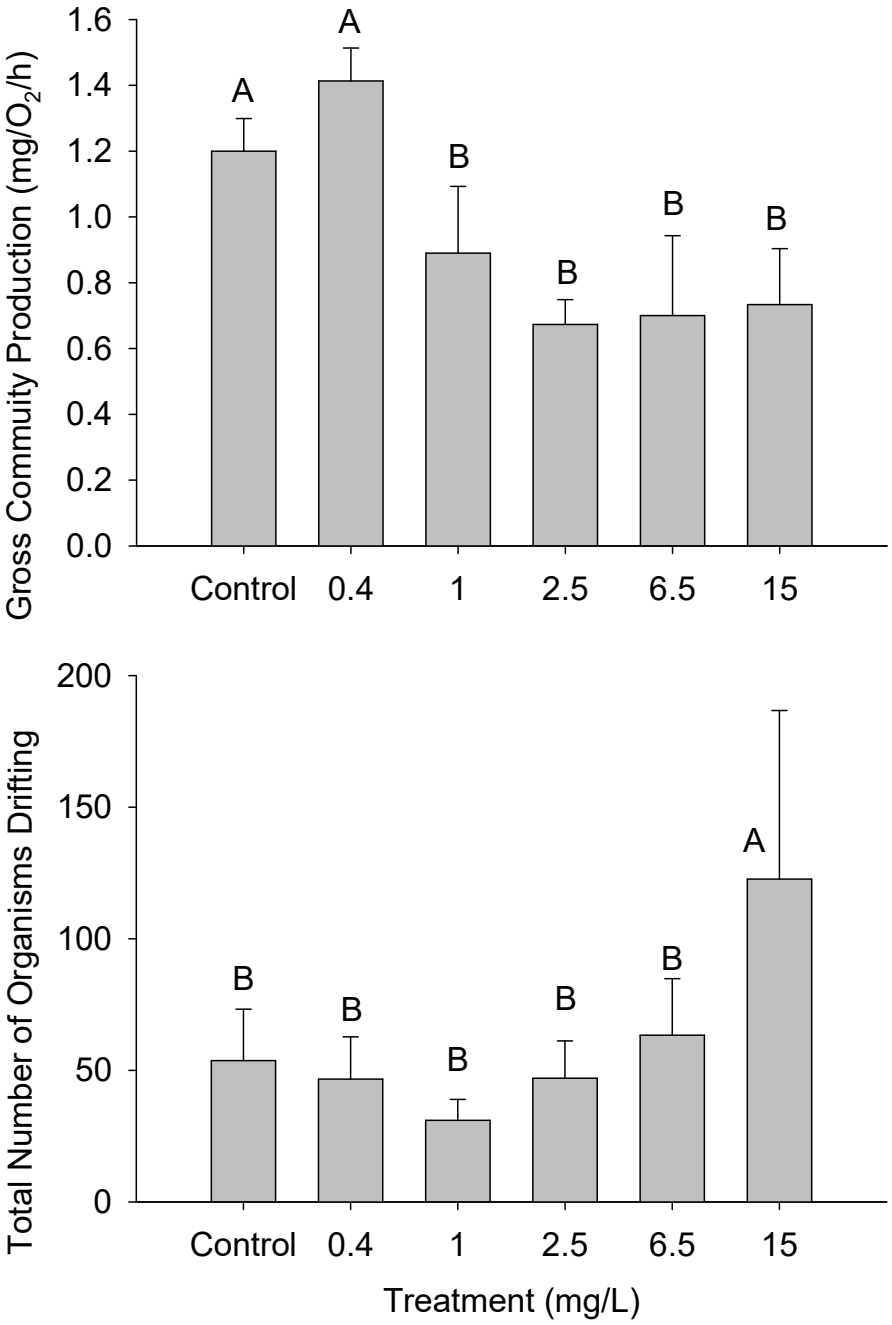


Figure 2.2 - Mean (+s.e.) community metabolism (A) and macroinvertebrate drift (B) measured in mesocosm experiment 1. Treatments with the same letter were not significantly different from each other based on results of Duncan's multiple range test.

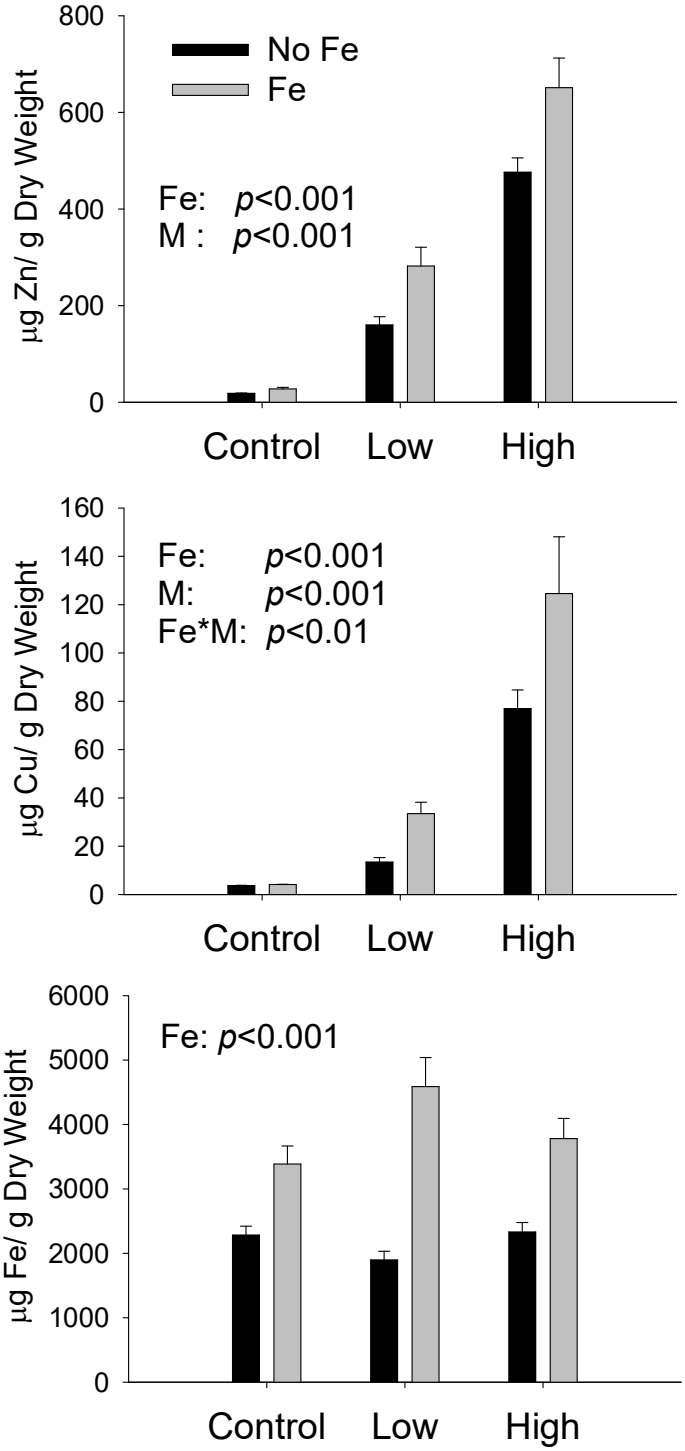


Figure 2.3 - Mean (+ s.e.) concentrations of Zn, Cu and Fe in periphyton after 9 days of exposure in mesocosm experiment 2. Results of 2-way ANOVA showing effects of Fe, metals and the Fe x metal interaction are shown for all significant (p<0.10) responses.

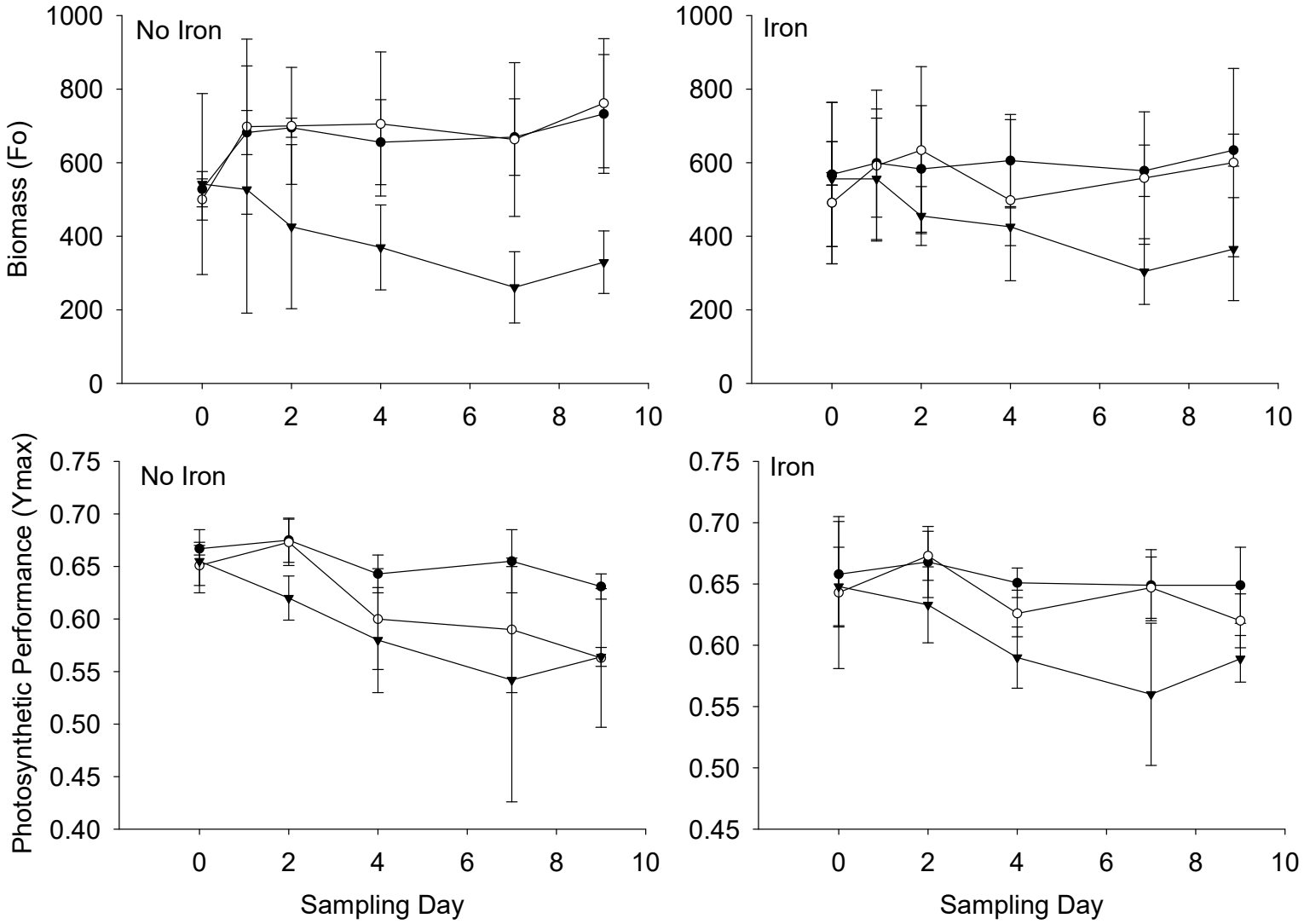


Figure 2.4 - Mean (± s.e.) biomass (F_o) and photosynthetic performance (Y_{max}) of periphyton after 9 days of exposure to Fe and metals in mesocosm experiment 2. Details of statistical analyses are shown in Appendix B, Table 2.S2.

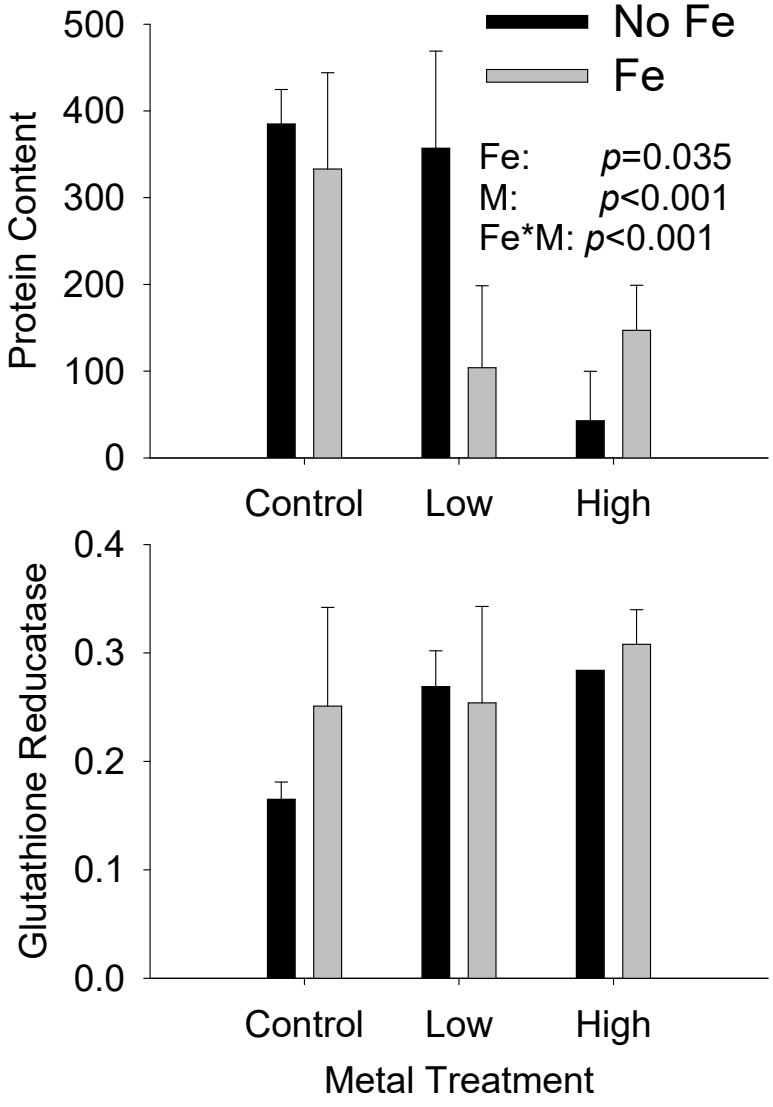


Figure 2.5 - Mean (+ s.e.) protein content (µg/ml) and glutathione reductase (µmol NADMH/min/µg protein) after 9 days of exposure to Fe and metals in mesocosm experiment 2. Results of 2-way ANOVA showing effects of Fe, metals and the Fe x metal interaction are included for all significant (p<0.10) responses. Details of statistical analyses are shown in Appendix B Table 2.4.

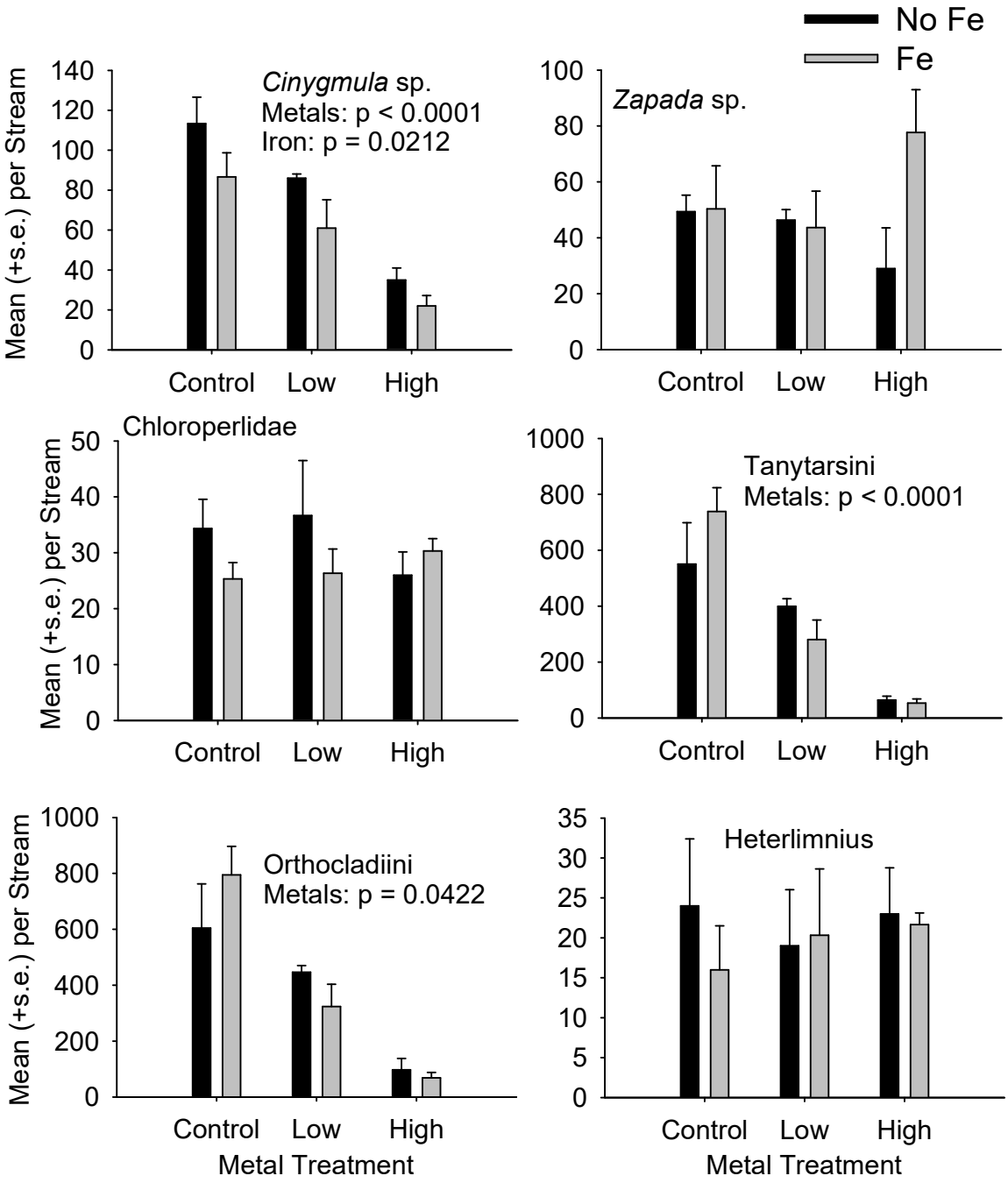


Figure 2.6 - Mean (+s.e.) abundance of the 6 dominant macroinvertebrate taxa from mesocosm experiment 2 after 10 days of exposure to metals and ferric Fe. Results of 2-way ANOVA showing effects of Fe, metals and the Fe x metal interaction are included for all significant (p< 0.10) responses. Details of statistical analyses are shown in Supplemental Table 2.

Table 2.1 - Mean (± s.e.) concentrations of Fe (mg/L; total, unfiltered) measured in stream mesocosms during experiment 1. Dissolved Fe was near the instrument detection limit (0.10 mg/L).

Target Fe	Measured Fe
0	0.12 (0.006)
0.40	0.46 (0.039)
1.0	0.94 (0.073)
2.5	2.42 (0.083)
6.3	5.24 (0.354)
15.6	14.1 (0.450)

Table 2.2 - Mean (± s.e.) concentrations of Fe (total, unfiltered), Cu and Zn (dissolved) measured in stream mesocosms during experiment 2. All values reported as mg/L.

Iron treatment	Metals treatment	Cu	Zn	Fe
No Fe	Control	0.002 (0.007)	0.004 (0.003)	0.185 (0.012)
	Low	0.010 (0.001)	0.089 (0.009)	0.142 (0.024)
	High	0.037 (0.007)	0.546 (0.017)	0.190 (0.014)
Fe	Control	0.003 (0.001)	0.003 (0.001)	0.603 (0.038)
	Low	0.011 (0.002)	0.075 (0.003)	0.622 (0.040)
	High	0.029 (0.003)	0.508 (0.016)	0.687 (0.050)

References

Aebi H. 1984. Catalase in vitro. *Methods in enzymology* 105:121-126.

Biesinger, KE, Christensen,GM. 1972. Effects of various metals on survival, growth, reproduction, and metabolism of *Daphnia magna*. *J Fish Res Bd Can* 29:1691- 1700.

Bonnineau C, Bonet B, Corcoll N, Guasch H. 2011. Catalase in fluvial biofilms: a comparison between different extraction methods and example of application in a metal-polluted river. *Ecotoxicology* 20:293-303.

Boutet C, Chaisemartin C. 1973. Specific toxic properties of metallic salts in *Austropotamobius pallipes pallipes* and *Orconectes limosus*. *CR Soc Biol* 167:1933-1938.

Buchwalter DB, Cain DJ, Clements WH, Lucima SN. 2007. Using biodynammic models to reconcile differences between laboratory toxicity tests and field biomonitoring with aquatic insects. *Environ Sci Technol* 41:4821-4828.

Cadmus P, Clements WH, Williamson JL, Ranville JF, Meyer JS, Gutiérrez MJ. 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.

Cairns JJ. 1983. Are single species toxicity tests alone adequate for estimating environmental hazard? *Hydrobiologia* 100:47-57.

Cardinale BJ, Matulich KL, Hooper DU, Byrnes JE, Duffy E, Gamfeldt L, Balvanera P, O'Connor MI, Gonzalez A. 2011. The functional role of producer diversity in ecosystems. *Am J Bot* 98:572-592.

Clark JL, Clements WH. 2006. The use of in situ and stream microcosm experiments to assess population- and community-level responses to metals. *Environ Toxicol Chem* 25:2306-2312.

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

Clements WH, Cadmus P, Brinkman SF. 2013. Responses of aquatic insects to cu and zn in stream microcosms: understanding differences between single species tests and field responses. *Environ Sci Technol* 47:7506-7513.

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, Hickey CW, Kidd KA. 2012. How do aquatic communities respond to contaminants? It depends on the ecological context. *Environ Toxicol Chem* 31:1932-1940.

Clements WH, Kashian DR, Kiffney PM, Zuellig RE. 2016. Perspectives on the context-dependency of stream community responses to contaminants. *Freshw Biol* 61:2162-2170.

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. *Freshw Sci* 35:126-138.

Corcoll N, Bonet B, Leira M, Guasch H. 2011. Chl-a fluorescence parameters as biomarkers of metal toxicity in fluvial biofilms: an experimental study. *Hydrobiologia* 673:119-136.

van Dam RA, Barry MJ, Ahokas JT, Holdway DA. 1998. Effects of water-borne iron and calcium on the toxicity of diethylenetriamine pentaacetic acid (DTPA) to *Daphnia carinata*. *Aquat Toxicol* 42:49-66.

Davis JA, James RO, Leckie JO. 1978. Surface ionization and complexation at oxide-water interface .1. Computation of electrical double-layer properties in simple electrolytes. *J Colloid Interface Sci* 63:480-499.

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.

Ehrlich P, Ehrlich A. 1981. *Extinction: The Causes and Consequences of the Disappearance of Species*. Random House, New York.

Foster PL, Morel FMM. 1982. Reversal of cadmium toxicity in a diatom - an interaction between cadmium activity and iron. *Limnol Oceanogr* 27:745-752.

Genty B, Briantais JM, Baker NR. 1989. The relationship between the quantum yield of photosynthetic electron-transport and quenching of chlorophyll fluorescence. *Biochimica Et Biophysica Acta* 990:87-92.

Gerhardt A. 1994. Short term toxicity of iron (Fe) and lead (Pb) to the mayfly *Leptophlebia marginata* (L.) (Insecta) in relation to freshwater acidification. *Hydrobiologia* 284:157-168.

Gerhardt A, Westermann F 1995. Effects of precipitations of Iron Hydroxides on Leptophlebia-marginata (l) (Insecta, Ephemeroptera) in the field. *Archiv Fur Hydrobiologie* 133:81-93.

Guasch H, Acosta XG, Urrea G, Baneras L. 2012. Changes in the microbial communities along the environmental gradient created by a small Fe spring. *Freshw Sci* 31:599-609.

Guasch H, Bonet B, Bonnineau C, Corcoll N, López-Doval J, Muñoz I, Ricart M, Serra A, Clements W. 2012. How to link field observations with causality? Field and experimental approaches linking chemical pollution with ecological alterations. In Guasch H, Ginebreda A, Geiszinger A, eds, *Emerging and Priority Pollutants in Rivers*. -The Handbook of Environmental Chemistry. Springer Berlin Heidelberg, pp 181-218.

Guasch H, Navarro E, Serra A, Sabater S. 2004. Phosphate limitation influences the sensitivity to copper in periphytic algae. *Freshw Biol* 49:463-473.

Hall RJ, Driscoll CT, Likens GE. 1987. Importance of hydrogen-ions and aluminum in regulating the structure and function of stream ecosystems - an experimental test. *Freshw Biol* 18:17-43.

Hare L, E. Saouter, PGC. Campbell, A Tessier, F Ribeyre, and A Boudou. 1991. Dynamics of cadmium, lead, and zinc exchange between nymphs of the burrowing mayfly *Hexagenia rigida* (Ephemeroptera) and the environment. *Can J Fish Aquat Sci* 48:39-47.

Irving EC, Baird DJ, Culp JM. 2003. Ecotoxicological responses of the mayfly *Baetis tricaudatus* to dietary and waterborne cadmium: implications for toxicity testing. *Environ Toxicol Chem* 22:1058-1064.

Kashian DR, Prusha BA, Clements WH. 2004. Influence of total organic carbon and UV-B radiation on zinc toxicity and bioaccumulation in aquatic communities. *Environ Sci Technol* 38:6371-6376.

Kiffney PM, Clements WH. 1996. Size-dependent response of macroinvertebrates to metals in experimental streams. *Environ Toxicol Chem* 15:1352-1356.

Linton TK, Pacheco MAW, McIntyre DO, Clement WH, Goodrich-Mahoney J. 2007. Development of bioassessment-based benchmarks for iron. *Environ Toxicol Chem* 26:1291-1298.

McKnight DM, Feder GL. 1984. The ecological effect of acid conditions and precipitation of hydrous metal-oxides in a rocky-mountain stream. *Hydrobiologia* 119:129-138.

Merritt RW, Cummins KW, Berg MB eds. 2008. An introduction to the aquatic insects of North America, fourth edition. Kendall/Hunt Publ. Co., Dubuque, Iowa.

Millward GE. 1980. The adsorption of cadmium by iron(III) precipitates in model estuarine solutions. *Environ Techn Lett* 1:394-399.

Millward GE, Moore RM. 1982. The adsorption of Cu, Mn and Zn by iron oxyhydroxide in model estuarine solutions. *Water Res* 16:981-985.

Murphy J, Riley JP. 1962. A modified single solution method for determination of phosphate in natural waters. *Analytica Chimica Acta* 26:31-&.

Naddy RB, Rehner AB, McNerney GR, Gorsuch JW, Kramer JR, Wood CM, Paquin PR, Stubblefield WA. 2007. Comparison of short-term chronic and chronic silver toxicity to fathead minnows in unamended and sodium chloride-amended waters. *Environ Toxicol Chem* 26:1922-1930.

Nakano Y, Asada K. 1981. Hydrogen peroxide is scavenged by ascorbate-specific peroxidase in spinach chloroplasts. *Plant Cell Physiol* 22:867-880.

Rader RB. 1997. A functional classification of the drift: traits that influence invertebrate availability to salmonids. *Can J Fish Aquat Sci* 54:1211-1234.

Randall S, Harper D, Brierley B. 1999. Ecological and ecophysiological impacts of ferric dosing in reservoirs. *Hydrobiologia* 395:355-364.

Rasmussen K, Lindegaard C. 1988. Effects of iron compounds on macroinvertebrate communities in a danish lowland river system. *Water Res* 22:1101-1108.

Schaedle M, Bassham JA. 1977. Chloroplast glutathione reductase. *Plant Physiol* 59:1011-1012.

Schmidt TS, Clements WH, Mitchell KA, Church SE, Wanty RB, Fey DL, Verplanck 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.

Smith EJ, Sykora JL. 1976. Early developmental effects of lime-neutralized Iron Hydroxide suspensions on Brook Trout and Coho Salmon. *Trans Amer Fish Soc* 105:308-312.

Sode A. 1983. Effect of ferric hydroxide on algae and oxygen consumption by sediment in a Danish stream. *Archs Hydrobiol Suppl* 65:134-162.

Stauber JL, Florence TM. 1985. The influence of iron on copper toxicity to the marine diatom, *Nitzschia-closterium* (Ehrenberg) Smith,W. *Aquat Toxicol* 6:297-305.

Sunda WG, Barber RT, Huntsman SA. 1981. Phytoplankton growth in nutrient rich seawater - importance of copper-manganese cellular interactions. *J Mar Res* 39:567-586.

Sunda WG, Tester PA, Huntsman SA. 1987. Effects of cupric and zinc ion activities on the survival and reproduction of marine copepods. *Mar Biol* 94:203-210.

Ward JV, Kondratieff BC, Zuellig RE eds. 2002. An Illustrated Guide to the Mountain Streams Insects of Colorado, Second Edition, 2nd edition ed. University Press of Colorado, Boulder, Colorado.

Wiederholm T ed. 1983. Chironomidae of the Holarctic region. Keys and diagnoses. Vol. 1. Larvae. Entomological Society of Lund, Sweden.

US Environmental Protection Agency. 1976. Quality Critera for Water. Washington D.C.

Vuori KM. 1995. Direct and indirect effects of iron on river ecosystems. *Annal Zool Fenn* 32:317-329.

Walker BH. 1992. Biodiversity and ecological redundancy. *Conservation Biology* 6:18-23.

Wellnitz TA, Sheldon SP. 1995. The effects of iron and manganese on diatom colonization in a Vermont stream. *Freshw Biol* 34:465-470.

Xie LT, Buchwalter DB. 2011. Cadmium exposure route affects antioxidant responses in the mayfly Centroptilum triangulifer. *Aquat Toxicol* 105:199-205.

Xie LT, Funk DH, Buchwalter DB. 2010. Trophic transfer of Cd from natural periphyton to the grazing mayfly Centroptilum triangulifer in a life cycle test. *Environ Pollut* 158:272-277.

CHAPTER 3

CHRONIC TOXICITY OF FERRIC IRON FOR NORTH AMERICAN AQUATIC ORGANISMS: DERIVATION OF A CHRONIC WATER QUALITY CRITERION USING SINGLE SPECIES AND MESOCOSM DATA[3]

Introduction

While iron is abundant in the earth's crust and occurs naturally in the aquatic environment, concentrations can be elevated due to human activities. Mining activities that expose pyrite and other sulfidic minerals to air and water lead to oxidation and release of iron and sulfuric acid in a process known as acid mine drainage (AMD). An estimated 20,000 to 50,000 mines in the western United States produce AMD which "seriously" affects 8,000-15,000 km of streams (USDA 1993) and is considered the greatest water quality problem in the Rocky Mountain region (Mineral Policy Center 1997). In the eastern United States, acid drainage from coal mines affects more than 7,000 km of streams (Kim 1982). Despite the widespread and harmful effects of iron, fewer than half of states in the USA have adopted a numeric chronic iron standard to protect aquatic life, and several states have deleted iron standards. The current USEPA chronic iron criterion of 1,000 µg/L (total recoverable) for protection of aquatic life was adopted in 1976, and is largely based on field observations of a single iron-polluted Colorado stream in which trout and other fishes were absent at iron concentrations > 1,000 µg/L (USEPA 1976). A field study conducted in Kentucky supported the 1,000 µg/L criterion (Birge et al. 1985). Nevertheless, the basis for this criterion is generally regarded to be insufficient (Thurston

3 After acceptance by my committee, a revised version of this chapter was published by Archives of Environmental Contamination and Toxicology as: Cadmus, P, Brinkman, SF and May, MK (2018) Chronic Toxicity of Ferric Iron for North American Aquatic Organisms: Derivation of a Chronic Water Quality Criterion Using Single Species and Mesocosm Data. Arch Environ Contam Toxicol 74: 605-615. https://doi.org/10.1007/s00244-018-0505-2

et al. 1979; Ohio EPA 1998). Development of a more scientifically rigorous iron criterion has been challenging because of its complex speciation, which is influenced by redox, dissolved oxygen, light, pH, and organic matter (Vuori 1995). In the aqueous environment, iron exists in two oxidation states: reduced ferrous ion (Fe II) and oxidized ferric ion (Fe III). In oxygenated waters, soluble ferrous ions (Fe II) oxidize to ferric ions (Fe III; Hem 1985). In circumneutral waters (pH > 6.5), ferric ions are insoluble and rapidly precipitate as hydroxides and oxyhydroxides (Hem 1985, Kimball et al. 2007). While iron speciation is indeed complex, ferric precipitates are the predominant form in waters capable of supporting aquatic life (i.e. oxygenated and circumneutral pH). Thus, ferric precipitates are the most relevant form of iron to consider for the development of a criterion for the protection of aquatic life.

U.S. water quality criteria are usually derived using the methodology outlined by Stephan et al. (1985). Briefly, a chronic criterion is intended to be protective of 95% of genera estimated from a dataset of toxicity values consisting of minimum of eight families that includes Salmonidae, another fish family in class Osteichthyes, a third family in Chordata, a planktonic crustacean, a benthic crustacean, an insect, a family in a phylum other than Arthropoda or Chordata, and finally a family in any order of insect or phylum not already represented. For this study, a literature review was conducted to identify chronic iron toxicity tests that met the following four criteria: 1. The species of test organisms used must exist in freshwater systems in North America; 2. The duration of the test was sufficiently long to detect sublethal effects (≥25 days or ≥7 days for Daphnids); 3. Ferric iron was used as the toxicant, because precipitates are the overwhelmingly predominant form of iron in circumneutral oxygenated waters; and 4. Toxicity tests were conducted at pH between 6.5 and 9.0 in order to minimize confounding effects of pH on results (see e.g. Radford 1997). Suitable tests existed for genera from

Salmonidae (*Oncorhynchus, Salvelinus),* another fish from class Osteichthyes (*Pimephales*), a planktonic crustacean (*Daphnia*), a benthic crustacean (*Orconectes*), and an insect (*Chironomus*) (Table 3.1). To add to this existing dataset, we conducted chronic single species laboratory toxicity tests on a Chordate (*Bufo*), an insect (*Hexagenia*), a non-arthropod invertebrate (*Lumbriculus*), additional members of Salmonidae (*Salmo, Prosopium)* and a family in another insect order or a phylum not otherwise represented (*Dugesia*). With these toxicity test results, iron toxicity data are available for a sufficiently diverse array of organisms to meet USEPA's recommended methods to calculate a chronic final value (FCV) for total Fe.

Traditionally, results from single species toxicity tests are used exclusively to derive US water quality criteria. While single species tests offer a greater degree of control and evidence for causation, they cannot evaluate interspecific interactions such as increased susceptibility to predation and ecologically relevant endpoints such as drift, which mesocosms are able to provide. A recent paper (Buchwalter et al. 2017) suggested that ecologically-relevant lines of evidence be used in the creation of water quality criteria. Specifically, the authors recommended the inclusion of mesocosm data in criteria development and the immediate use of mesocosm studies to test the hypothesis that a criteria is protective. To test the hypothesis that the single-species derived FCV was protective, I conducted a ten day mesocosm experiment using naturally colonized communities of benthic macroinvertebrates. The results from this experiment were compared to the Fe criterion calculated from the single species experiments. Lastly, effect concentration (EC_{20}) values for species in the mesocosm experiments were incorporated into the species sensitivity distributions from single species tests to derive a FCV using both single species and mesocosm data.

Methods

Single species test methods followed ASTM method E1241, *Standard Guide for Conducting Early Life-Stage Toxicity Tests with Fishes* (ASTM 1997) using ferric chloride as the toxicant. Dissolution of ferric chloride and the subsequent precipitation of ferric hydroxide release acidic protons according to the reaction:

$$FeCl_3 + 6H_2O \rightarrow Fe(OH)_3 + 3H^+ + Cl^-$$

As a result, adding a concentrated stock solution of ferric chloride to dilution water would lower pH and alkalinity and confound interpretation of toxicity results (see *e.g.* Radford 1997). To prevent changes in pH and alkalinity among iron exposure levels, sodium hydroxide was added to the stock solution in a 3:1 stoichiometric ratio to neutralize the acid formed by the precipitation of ferric hydroxide. Stock solutions were >6.5 pH before use. Measured alkalinity and pH were similar among the iron exposure levels for all studies. In the flow-through experiments, aeration of stock solutions, diluter compartments and exposure chambers were used to minimize settling of ferric precipitates.

Brown Trout and Mountain Whitefish

Freshly fertilized eggs were collected from wild spawning adults. Brown trout (*Salmo trutta,* Linnaeus) eggs and milt were collected as part of the annual Colorado Parks and Wildlife spawning operations (North Delaney Buttes Reservoir, Jackson County, Colorado, USA). Mountain whitefish (*Prosopium williamsoni*, Girard) eggs and milt were collect from adults in

spawning condition (Mad Creek, Routt County, Colorado, USA). Eggs were stripped, fertilized and water-hardened in the field and transported in coolers to the Colorado Parks and Wildlife (CPW) Aquatic Toxicology Laboratory in Fort Collins, Colorado, USA. Upon arrival, eggs were treated with 1600 ppm formalin for 15 minutes to control fungus (Piper et al. 1986).

A continuous-flow diluter (Benoit et al. 1982) constructed of Teflon, polyethylene and polypropylene components delivered five exposure levels of iron hydroxide and an exposure control. Source water was dechlorinated municipal tap water (Fort Collins, Colorado, USA). Target Fe concentrations were 0, 625, 1,250, 2,500, and 5,000 μg/L total iron. In order to accommodate the additional aquaria needed to test two species simultaneously the number of exposure levels was reduced from five to four. Iron stock solution was prepared by dissolving ferric chloride hexahydrate ($FeCl_3 \cdot 6H_2O$, Mallinckrodt analytical reagent grade) with sufficient NaOH (1:3 stoichiometry) to neutralize acidic conditions caused by precipitation of ferric hydroxide. The stock solution was pumped to the diluter with a peristaltic pump at a rate of 2 mL/min. A flow splitter equally allocated each iron concentration to each of six replicate 7.5 L glass aquaria at 30 mL/min. Aquaria, stock solutions and diluter compartments were aerated to keep iron precipitates suspended in the water column. Exposure solutions were delivered via food-grade vinyl tubing to egg incubation cups constructed of 1,000 μm nylon screen affixed to PVC pipe segments (53 mm I.D. X 75 mm) with aquarium-grade silicone adhesive. Each incubation cup was suspended in a 7 L glass aquarium with a standpipe that allowed the exposure solution to overflow into a temperature-controlled water bath. Thirty eggs were distributed to each incubation cup. Treatments were arranged so that each species was exposed to three replicates of each iron concentration. Treatments were randomized in complete blocks. Ambient fluorescent light (16h:8h photoperiod) provided illumination. Temperature of dilution

water and water bath was initially 7 °C and then increased to 12 °C after hatch of whitefish (Appendix C Table 3.S1). Mountain whitefish temperatures were assigned a low temperature during egg incubation and a higher temperature after hatch for three reasons: 1. Mountain whitefish eggs do not survive temperatures >8 °C (Rajagopal 1979; Brinkman et al. 2013), 2. The egg stage was expected to be a sensitive life stage and lower incubation temperatures would extend exposure times, 3. Lower temperature during egg incubation is a more natural temperature regime for fall spawning species such as mountain whitefish and brown trout.

Incubation cups were inspected daily for egg mortality and hatch. The first twelve brown trout eggs and first fifteen whitefish eggs to hatch were carefully transferred from the incubation cup to the aquarium using a glass tube and pipette pump. Remaining eggs in the incubation cups were monitored for hatching and removed once hatching was completed. Thus, hatching success for each species was based on 30 embryos in each incubation cup while fry survival and growth were based on 12 and 15 fry transferred to the aquaria for brown trout and mountain whitefish, respectively. After absorption of the yolk-sac, brown trout fry were fed starter trout chow (Rangen soft-moist) five times per day with an automatic feeder at a rate of 5% body weight (BW)/day. Whitefish fry were fed <24 hour brine shrimp nauplii three times per day (one to two times per day on weekends and holidays) at a rate of 5% BW/day. Tests were ended 30 days post-swimup and fry were terminally anesthetized with MS-222 and weighed (g). Total duration of exposure including embryo and larval stages was 79 days for brown trout and 78 days for mountain whitefish.

Boreal Toad Tadpoles

Fertilized boreal toad eggs (*Bufo boreas,* Baird and Girard) were obtained from Trout Lake (Larimer County, Colorado, USA). A continuous-flow diluter (described above) delivered five exposure levels of iron and an exposure control. Target concentrations of 0, 500, 1,000, 2,000, 4,000, and 8,000 µg/L total iron were delivered at a rate of 40 mL/min to 2.8 L polypropylene tanks at 20 °C using methods described above. Five tadpoles (c.a. stage 18; Gosner 1960) were carefully distributed into each tank (n= 4 replicate tanks) using a glass pipette. Tadpoles were fed (*ad libitum*) a mixture of Mazuri amphibian feed and powdered algae wafers (1:1) and a processed slurry of kale, mustard greens and squash. Tanks were cleaned to remove feces and excess food every 2 days. Tanks were monitored daily for mortality. After 35 days of exposure, tadpoles were terminally anesthetized with MS-222 and lengths (mm), weights (g) and developmental stage (Gosner 1960) measured for each tadpole.

Lumbriculus

Toxicity trials of *Lumbriculus variegatus* (Müller) were conducted using organisms from an onsite culture obtained from the USEPA laboratory in Duluth Minnesota. The onsite culture was maintained in a 39 L glass aquarium with washed coarse sand as a substrate and fed a slurry of trout starter feed. At the start of the experiment, 15 individuals were weighed and placed into each 2.8 L polypropylene exposure chamber. Each contained 150 mL of coarse washed sand and was maintained at 21ºC. Target concentrations of 0, 1,000, 2,000, 4,000, and 8,000 µg/L total Fe were delivered to five replicate treatment tanks per exposure level as described above. After 35 days of exposure, individuals in each tank were enumerated and weighed.

Hexagenia

Hexagenia limbata (Serville) nymphs were obtained from Aquatic Research Organisms (Hampton, NH, USA). At the start of the experiment, 10 individuals were placed into each 2.7 L polypropylene exposure chamber maintained at 17ºC. Glass tubes 5 cm long and of varying inside diameters (4.9, 6, 7, 9 mm) were provided as artificial burrows (Fremling and Mauck 1980) which the nymphs readily adopted. Nymphs were fed 2.0 mL of a slurry consisting of 500 mL yeast-trout chow-Cerophyl (YTC), 20 g Tetramin fish food and 5 g wheatgrass. Target concentrations of 0, 500, 1,000, 2,000, 4,000, and 8,000 µg/L total iron were delivered to four replicate 2.8 L treatment tanks per treatment level as described above. After 30 days of exposure, individuals in each tank were enumerated and weighed.

Dugesia

Planarian worms (*Dugesia dorotocephala*, Girard) were field-collected near the outflow of the Colorado Parks and Wildlife Bellvue-Watson Rearing Unit (Larimer County, Colorado, USA). Six individuals were randomly placed in each of 24 polystyrene petri dishes (145 x 20 mm) each containing 100mL of exposure solution. Preliminary studies showed *Dugesia dorotocephala* to be tolerant. Target concentrations of 0, 2,500, 5,000, 10,000, 20,000, and 40,000 µg/L total iron were made using dechlorinated municipal tap water (Fort Collins, Colorado, USA) and an iron stock solution of ferric chloride hexahydrate ($FeCl_3 \cdot 6H_2O$, Mallinkrodt analytical reagent grade) with sufficient NaOH (1:3 stoichiometry) to neutralize acidic conditions caused by precipitation of ferric hydroxide. Petri dishes were renewed twice weekly with freshly prepared exposure solutions. Dishes were monitored daily for mortality and fissioning (asexual reproduction in which a single organisms physically splits into two

organisms). After 30 days, the test was terminated and all remaining planarians were enumerated and weighed.

Aquatic Macroinvertebrate Community

In October 2010 effects of ferric Fe on communities of aquatic macroinvertebrates were measured using stream mesocosms containing naturally colonized substrate (Clements 2004, Cadmus et al. 2018). Colonization trays (10 x 10 x 6 cm) filled with cobble were allowed to colonize for 32 d in the South Fork of the Michigan River, a stream originating from a wilderness area in the Routt National Forest (Colorado, USA). Trays were randomly assigned to coolers (4 trays in each) and were transferred to the 18 experimental streams at the Colorado State University Stream Research Laboratory (SRL) in Fort Collins, Colorado, USA. Flow-through conditions were maintained at 1.0 L/min of untreated water from a reservoir fed by mountain streams (Horsetooth Reservoir, Fort Collins, Colorado, USA). Peristaltic pumps delivered stock solution to experimental streams to create six ferric Fe treatments (n = 3; 0, 464, 944, 2,425, 5,238 and 14,073 µg/L) for 10 d. Stock solutions of Fe were neutralized with NaOH as described above. Reagents were mixed and vigorously aerated for 1 h. Prior to use, all stock solutions were tested for a circumneutral pH (6.5 to 7.5). After 10 d all organisms retained in a 355 µm sieve were preserved in ethanol for identification to genus (tribe or subfamily for chironomids) and enumeration.

Water Quality

For all single species experiments, unfiltered (total) samples for total iron were collected weekly from each exposure level in 60 mL high density polyethylene bottles (Nalgene). During

the mesocosm experiment total iron was sampled from each experimental unit every other day in 15 ml polypropylene centrifuge tubes (Falcon). Samples were immediately preserved with high-purity nitric acid (JT Baker) to pH <2. Iron concentrations were measured using an Instrumentation Laboratory Video 22 (Allied Analytical Systems, Franklin, Massachusetts, USA) atomic absorption spectrometer with air-acetylene flame and Smith-Hieftje background correction. The spectrometer was calibrated prior to each use and the calibration curve verified through analyses of external quality assurance samples (High Purity Standards, Charleston, South Carolina USA). Sample splits and spikes were collected at each sampling event to verify analytical reproducibility and recovery.

Water quality (pH, temperature, dissolved oxygen, hardness, alkalinity and conductivity) was assessed every other day from each experimental unit in the mesocosm trial. During single species trials water quality characteristics were measured weekly in all aquaria within a block. A different replicate was selected rotationally each week. Alkalinity was determined according to standard methods (APHA 1998). Dissolved oxygen and pH were measured with an electronic meter (Oakton Model 300 or YSI model 550a and 63) calibrated prior to each use. Conductivity was measured with an YSI model 35 or 63 conductance meter.

Statistical Analyses

The maximum allowable toxicant concentration (MATC) was calculated as the geometric mean of the no-observed-effect concentration (NOEC) and lowest-observed-effect concentration (LOEC) of the most sensitive endpoint. These were determined using Analysis of Variance (ANOVA) using Toxstat version 3.5 software (West Inc. 1996, Cheyenne, WY, USA). Hatching success and survival data were arcsine square root transformed prior to ANOVA. Normality and

homogeneity of variances were tested using Chi-square and Bartlett's test, respectively. Treatment means were compared to the control using William's one-tailed test (Williams 1971; Williams 1972). The highest measured iron concentration not associated with a treatment effect (e.g. decreased survival, decreased body weight) was designated as the NOEC. The lowest measured iron concentration associated with a statistically significant treatment effect was designated as the LOEC. Use of NOEC and LOEC values has fallen out of favor with the scientific community (Warne and van Dam 2008, USEPA 1999). For this reason we used the concentration predicted to cause a 20% reduction in survival or performance (EC_{20}) when data were available. EC_{20}s were calculated using USEPA's Toxicity Relationship Analysis Program (TRAP version 1.30a; USEPA 2015a). Three parameter piecewise linear estimates of log transformed abundance data were used to calculate EC_{20} values of mesocosm results. Linear analysis of log transformed abundance data has been successfully employed in community mesocosm experiments exposing insect communities to toxicants (Clements et al. 2013). A threshold sigmoidal model was used to model single species results.

Derivation of Final Chronic Value

USEPA's ECOTOX database (USEPA 2015b) and science literature databases were used to identify iron toxicity tests that were of sufficient duration (≥25d, Daphnids ≥7d) to detect sublethal effects such as reduced growth or reproduction. Only studies that used the ferric iron and conducted at circumneutral pH (6.5 to 9.0) were included. If sufficient data were reported, regression analysis was used to calculate chronic values. USEPA's Toxicity Relationship Analysis Program version 1.30a (TRAP), was used to determine EC_{20} for both single species laboratory tests and mesocosm tests. If insufficient partial effects were observed to produce a

reliable estimate or if insufficient data were reported to run TRAP, MATCs or other effect concentrations reported by the authors were for chronic values. Chronic values of toxicity tests that met the screening requirements are reported in Table 3.1. Details of TRAP results and chronic values are reported in the Methods Narrative in Appendix C. Species Mean Chronic Values (SMCV) were calculated as the geometric mean of chronic values in the limited instances where multiple chronic values were available for the same species. Genus Mean Chronic Values (GMCV) were calculated as the geometric mean of relevant SMCVs. GMCVs were ranked and a Final Chronic Value (FCV) was calculated using methods described by Stephan et. al (1985).

To test the hypothesis that the FCV derived from single species tests is protective of natural benthic communities, EC_{20} values from the mesocosm study were compared to the FCV. Only fourteen of the forty taxa present occurred in sufficient abundance to calculate an EC_{20} or to demonstrate there was no significant effect at my highest concentration. Given the short duration of the mesocosm experiment and the inclusion of two insect families in the FCV, we predicted that the FCV was sufficiently below the EC_{20} of any macroinvertebrate taxa in our mesocosm. Finally, a second FCV was calculated by adding the 14 genera from the mesocosm study to the species sensitivity distribution that included GMCVs from single species trials.

Results

Single Species Toxicity Tests

Details on water quality measurements and toxicity endpoints can be found in Appendix C, Tables 3.S1-S6. Single species toxicity tests were deemed acceptable based on ASTM criteria

(1997). Hatch success of brown trout and mountain whitefish exceeded 80% in the control treatments. Posthatch survival was 91% and 84% in brown trout and mountain whitefish controls, respectively. Control survival was 100% for boreal toad tadpoles and 84% for *Hexagenia* nymphs. The number of individuals in control treatments increased by a factor of 7.5 and 1.4 for *Lumbriculus* and *Dugesia*, respectively. Measured dissolved oxygen concentrations were near saturation (mean 97%, range 82%-105%) and biomass loading never exceeded 3.1 g/L or 0.15 g/L/24h. Temperatures within each experiment were consistent among exposure levels. Alkalinity and pH measurements within each experiment were consistent throughout the duration of each test and also consistent among exposure concentrations demonstrating that ferric chloride was neutralized by the addition of sodium hydroxide.

Iron was not lethal to any of the organisms tested except for boreal toad tadpoles. Significant sublethal effects were detected including reduced growth for boreal toad tadpoles and mountain whitefish, reduced development for boreal toad tadpoles and reduced reproduction for *Lumbriculus* (Figure 3.1 and Figure 3.2). EC_{20}s were 1318 µg/L for mountain whitefish based on biomass, 3145 µg/L for boreal toad tapoles based on biomass and 870 µg/L for *Lumbriculus* based on number of organisms at the end of the test. No significant effects were detected for brown trout, *Hexagenia* or *Dugesia* at exposure concentrations used in the tests.

Final Chronic Value

Chronic toxicity data reported in the scientific literature were combined with results from this study (Table 3.1). Details on screening, acceptability and treatment of toxicity data from the literature are available in the Supplemental Narrative in Appendix C. Only 12 genera are represented, six of which are from this study. Though limited in taxonomic diversity and number

of genera, the dataset met the eight family minimum requirement needed to calculate a criterion (Stephan et al. 1985). Using the four most sensitive genera (*Pimephales, Lumbriculus, Prosopium* and *Daphnia*) and 12 as the number of genera, a Final Chronic Value (FCV) of 499 µg/L total Fe was calculated.

Mesocosm Toxicity Test

Physical and chemical assessments of experimental streams showed Fe concentrations near target levels and water quality similar to that of a high mountain stream (Appendix C, Table 3.S6). Despite the high-flow, turbulent environment produced in the experimental streams we observed Fe precipitates clogging interstitial space and covering substrate and organisms. The Fe oxides appeared to indirectly affect benthos by reducing availability of benthic habitat, increasing turbidity and reducing periphyton quality (Cadmus et al. 2018). Fourteen of the forty aquatic insect taxa found in community mesocosms were sufficiently abundant (> 20 individuals in control streams) to calculate an EC_{20} value using TRAP (Table 3.2). EC_{20} values of the mayfly *Epeorus* sp. (335 µg/L), the caddisfly *Micrasema* sp. (356 µg/L), and the chironomid tribe Tanytarsini (234 µg/L) were below the FCV of 499 µg/L (Table 3.2). Additionally EC_{20} values of the chironomid subfamily Orthocladiinae (776 µg/L) were below the current national criterion of 1,000 µg/L. If taxa from the mesocosm experiment are included a FCV of 251 µg/L total Fe is calculated (Appendix C Table 3.S9).

Discussion

Iron was not lethal in the single species toxicity tests except for boreal toad tadpoles. Instead, iron toxicity effects were sublethal which included reduced growth for boreal toad tadpoles and mountain whitefish, reduced development for boreal toad tadpoles and reduced reproduction for *Lumbriculus* (Figure 3.1 and Figure 3.2). *Lumbriculus* are generally regarded as tolerant to dissolved metal exposure. However, iron precipitates which accumulated on the substrate may have interfered with feeding. This would be consistent with the notion that iron precipitates act as an indirect or physical stress on organisms and ecosystems rather than direct chemical toxicity. In neutral waters Fe has been found to increase turbidity, reduce primary production and reduce interstitial space in the benthic zones which smothers invertebrates, periphyton and eggs (USEPA 1976; Goettl and Davies 1977; DeNicola et al. 2002; McKnight et al. 1984; Vuori 1995; Linton et al. 2007; Hayer et al. 2013). Iron precipitates also physically clog and damage gills causing respiratory impairment (Peuranen et al. 1994, Dalzell et al. 1999).

The single species FCV was calculated using EC_{20}s as chronic values, in instances where EC_{20}s could be reliably estimated. Otherwise, MATCs or other chronic toxicity values reported by authors were used. Chronic values based on regression analysis enables a uniform level of effect among different tests. In contrast, chronic values based on hypothesis testing to determine LOEC and NOEC is not based on magnitude of an effect, and is sensitive to sample size, number of replicates and variability of endpoints. Use of EC_{20}s to derive FCVs was a risk management decision made by USEPA (1999), reflecting a compromise between a low level of effect such as EC_{10} which is rarely significantly different from a control, and an EC_{50} which can be estimated with greatest precision but is clearly too large of an adverse effect for adequate protection. For

the Fe chronic toxicity dataset, EC_{20}s were often close to MATCs and nearly always between the NOECs and LOECs (Appendix C Table 3.S6). Nevertheless, a FCV calculated using MATCs would have increased to 628 µg/L from the 499 µg/L calculated using EC_{20}s.

The current USEPA chronic Fe criterion for protection of aquatic life is 1000 µg/L total Fe, a value based principally on limited field observations which has not been updated since 1976. Using single species toxicity data in Table 3.1, a Final Chronic Value of 499 µg/L total Fe was calculated by applying USEPA methodology (Stephan et al. 1985). We believe this FCV is more rigorous and has a stronger scientific basis than the current criterion. This methodology uses chronic values of the four most sensitive genera to estimate a concentration that would protect 95% of taxa (Stephan et al. 1985). Interestingly, extrapolation of the trendline of the percentile versus genus mean chronic values of all genera to the 0.05 percentile yielded a concentration of 439 µg/L (Figure 3.3), in good agreement with 499 µg/L derived using USEPA methodology.

Results from single species toxicity tests are currently the preferred data for deriving USEPA water quality criteria. Such laboratory tests provide a high degree of control, standardization and reproducibility. However, restricting water quality criteria to single species data clearly has its limitations. Single species tests lack environmental realism, rely on a limited number of easy to culture organisms and do not consider interactions at higher levels of biological organization. Buchwalter et al. (2017) argue that water quality criteria should incorporate more ecologically relevant data. One recommendation is to include results of mesocosm experiments. Indeed, results of the mesocosm identified three taxa that would not be protected by the FCV calculated from single species tests. EC_{20} values of the mayfly *Epeorus* sp.

(335 µg/L), the caddisfly *Micrasema* sp. (356 µg/L), and the chironomid tribe Tanytarsini (234 µg/L) were below the FCV of 499 µg/L derived using single species toxicity tests (Table 3.2). These findings highlight the limitations of single species based criterion and strongly support inclusion of mesocosm test results for derivation of water quality criteria. Including mesocosm data would lower the FCV to 251 µg/L total Fe. The FCV with the mesocosm data included is supported by an assessment of field data that found iron as low as 210 µg/L may be necessary to protect sensitive insect species (Linton et al. 2007).

Inclusion of benthic mesocosm results clearly adds environmental relevance and reduced the risk of calculating an underprotective FCV for total Fe. However, mesocosm experiments are not a panacea for improving water quality standards, particularly if they are poorly designed or are unrepresentative of natural biota. Improperly designed mesocosm experiments can overestimate safe pollution levels in the same ways as single species experiments. During this exercise we observed several qualities of mesocosm data that should be considered when deriving a standard with this approach. A single mesocosm experiment has potential to drastically influence a water quality standard. The derivation of water quality criteria described by Stephan et al. (1985) considers the four most sensitive species, while the addition of more tolerant species (Number of Genera) increases the FCV. In one single mesocosm experiment, we considered forty macroinvertebrate taxa. Fourteen of these taxa were abundant enough to safely calculate an EC_{20}. The original FCV would have increased 250% or 148%, respectively, if the species considered were more tolerant than the fourth most sensitive genus (*Daphnia*) in the species sensitivity distribution. Standards are best informed by mesocosm experiments when those experiments represent the natural community and thus avoid consideration of an un-natural number of tolerant species or tolerant age classes. For this reason we used naturally colonized

communities from a pristine stream. Sampling disturbed sites or artificially building communities from tolerant species might have artificially inflated the FCV. This experiment considered 40 species but only 12 were abundant enough to derive an EC_{20} value using TRAP. A larger sample size (more naturally colonized substrate) in each experimental unit would have allowed for inclusion of the more rare and sensitive species, better characterizing the entire community. In this experiment we collected organisms retained in a 355 micron sieve. Use of a larger mesh size when sampling aquatic macroinvertebrates might underrepresent smaller, more sensitive, age classes. Additionally, community structure of benthos and organism size changes seasonally. To best characterize sensitivity of each taxa, mesocosm experiments might need to be repeated through the seasons. Improper design or methods can easily lead to underprotective FCV. However, we struggle to envision a situation in which similar mesocosm experiments could lead to an overprotective FCV when species composition of experiments is created using natural colonization from environmentally relevant locations.

The 10-day duration of my mesocosm experiment likely underrepresented the toxic effects of Fe that could be possible after 30 d, or after a complete life cycle for these organisms. Shorter exposure durations typically increase the concentration at which a response is detectable. Mesocosm techniques using naturally colonized benthos have been conducted upwards of 30 d durations (Mebane et al. 2016). It is likely that increased exposure durations would have produced much lower FCV than we report here.

Results of toxicity tests reported here suggest that the current USEPA chronic Fe criterion of 1000 µg/L is underprotective of sensitive aquatic life. A chronic criterion calculated from single species toxicity tests suggest that the current criterion should be reduced by half to 499 µg/L. Mesocosm results and field data (Linton et al. 2007) suggest that sensitive species may

require the Fe criterion to be reduced by half again to 251 µg/L. Some field studies observed that aquatic life appear unaffected at iron concentrations that exceed the water quality criterion of 1,000 µg/L (Ohio EPA 1998; Loeffelman et al. 1985). Water quality criteria are intended to protect 95% of species, and as such, may appear overly protective in circumstances where more tolerant organisms are present or in communities where sensitive species have been extirpated. Field studies that fail to detect adverse effects to aquatic life at concentrations above a criterion value should not necessarily be interpreted as demonstrating an overprotective criterion.

Figures and Tables from Chapter 3

Contains:

Figures 3.1 to 3.3

Tables 3.1 to 3.2

See Appendix C for supplimental materials including Tables 3.S1 to 3.S9 and the supplemental narrative of methods for Chapter 3.

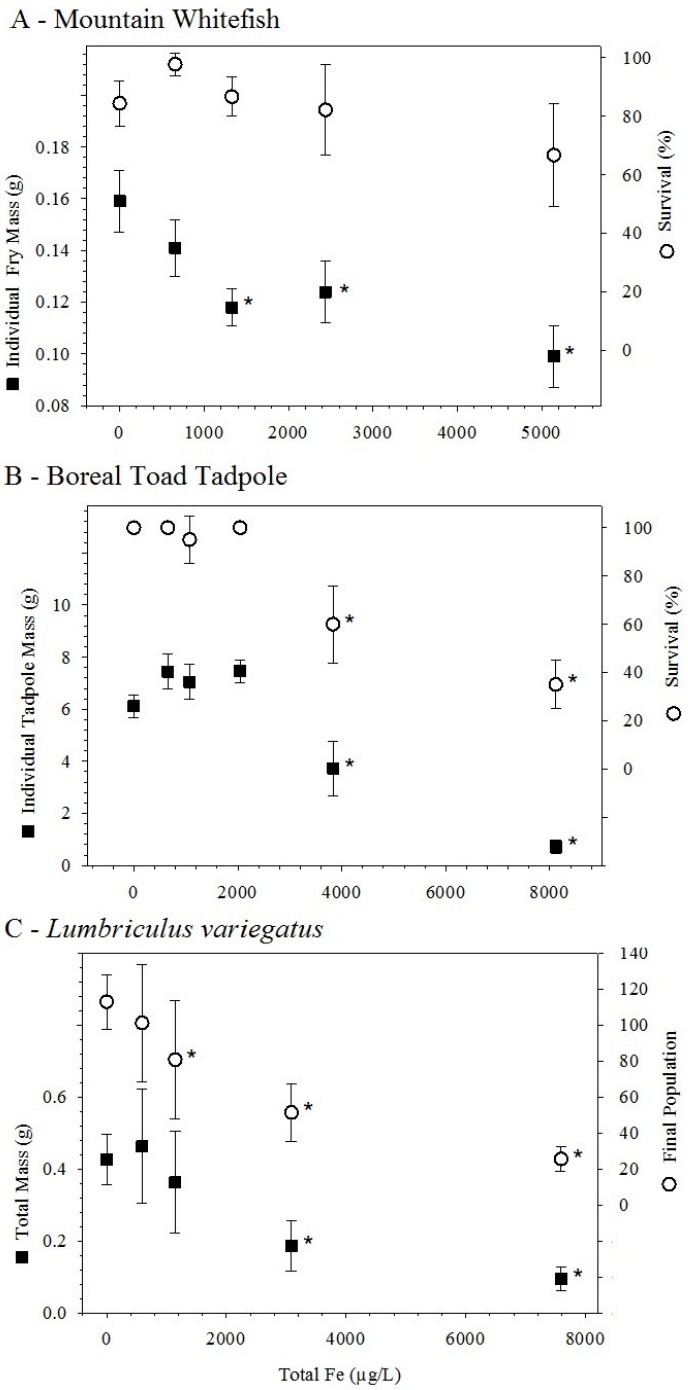


Figure 3.1 - Response of organisms to total Fe. A) Survival of Mountain White Fish was a less sensitive endpoint than growth measured in individual fry mass. B) Reduced survival and mass of Boreal Toad tadpoles were observed at 3831 µg/L total iron. C) Population size of *Lumbriculus* worms was reduced in the 1145 µg/L treatment levels when compared to controls. All experimental units started at 16 individuals. Asterisks denote treatment means significantly less than control ($p < 0.05$).

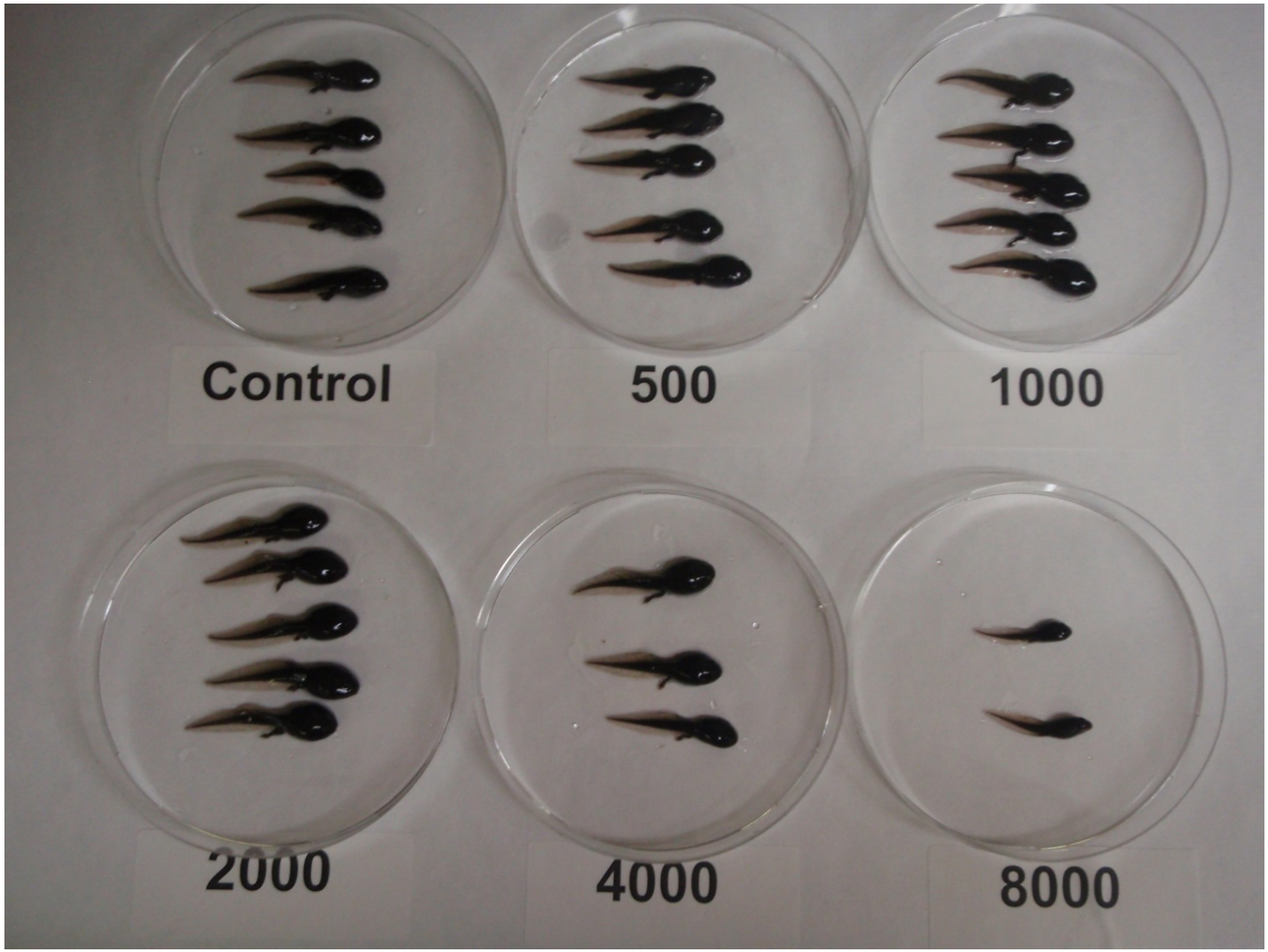


Figure 3.2 - Boreal toad (*Bufo boreas)* tadpoles exposed to 8,000, 4000, 2000, 1000, 500 µg/L iron at termination of 35 d toxicity test. Tadpoles exposed to 8000 µg/L were significantly smaller and less developed than controls ($p < 0.05$).

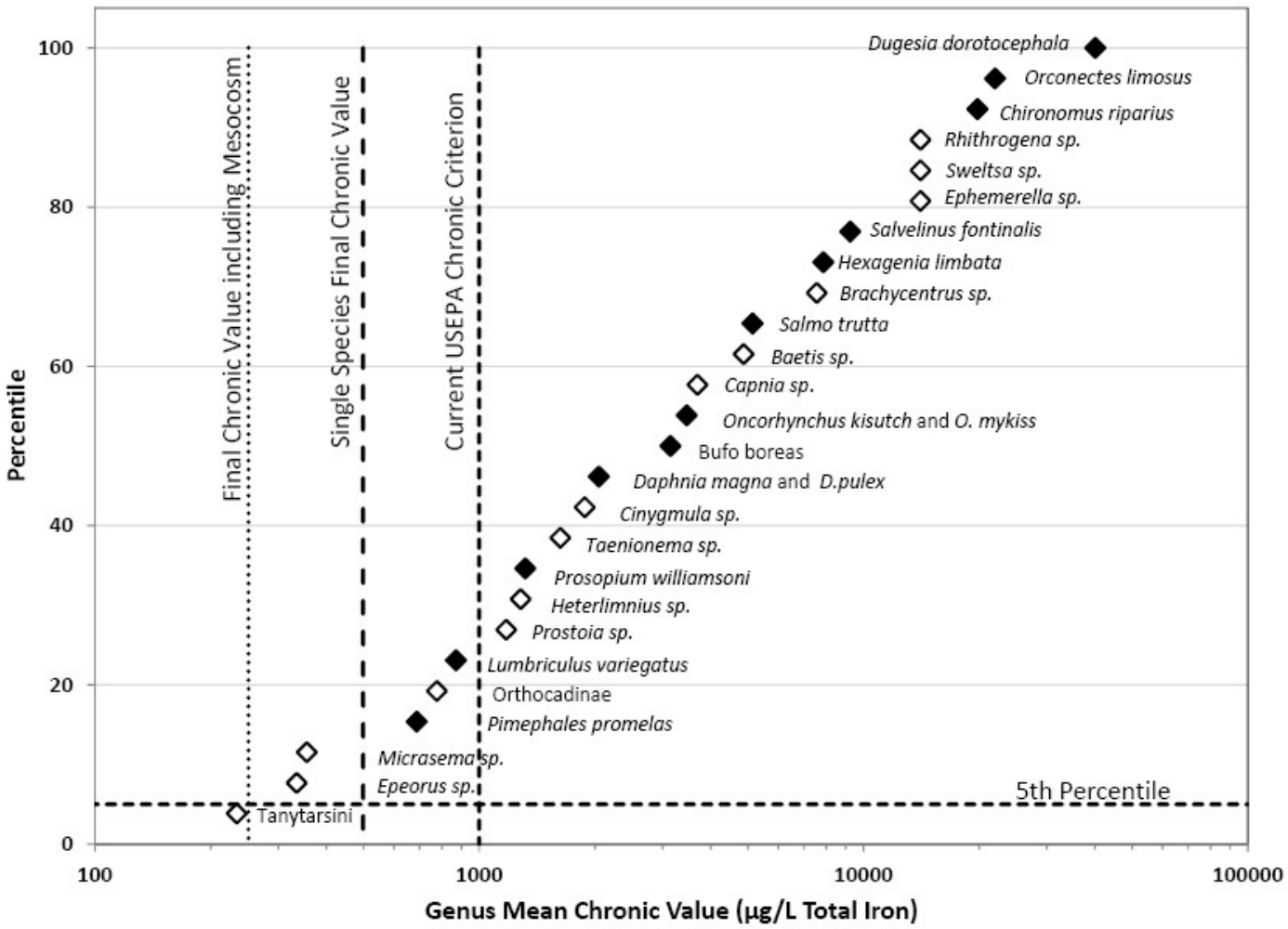


Figure 3.3 - Genus Sensitivity Distribution Sensitivity distribution of genera to iron. ♦ = chronic single species experiments. ◊ = 10 d mesocosm experiment.

Table 3.1 - Chronic Values from experiments included in calculation of Species Mean Chronic Values (SMCV) and Genus Mean Chronic Values (GMCV). Fe concentrations in µg/L total or total recoverable Fe. Excludes mesocosm experiments.

Rank	Scientific name	Common name	Chronic Value (µg/L)	SMCV (µg/L)	GMCV (µg/L)	Reference
12	*Dugesia dorotocephala*	Planarian	40134	40134	40134	This study
11	*Orconectes limosus*	Crayfish	22000	22000	22000	Boutet and Chaisemartin 1973
10	*Chironomus riparius*	Midge	19818	19811	19811	Radford 1997
9	*Salvelinus fontinalis*	Brook trout	9237	9237	9237	Sykora et al. 1975
8	*Hexagenia limbata*	Mayfly	7863	7863	7863	This study
7	*Salmo trutta*	Brown trout	5146	5146	5146	This study
6	*Oncorhynchus kisutch*	Coho salmon	4870	3605	3467	Smith and Sykora 1976
	Oncorhynchus kisutch	Coho salmon	3300			Brenner and Cooper 1978
	Oncorhynchus kisutch	Coho salmon	2915			Updegraff and Sykora 1976
	Oncorhynchus mykiss	Rainbow trout	1483	3335		Goettl and Davies 1977
	Oncorhynchus mykiss	Rainbow trout	7500			Steffens et al. 1993

Table 3.1 Continued

Rank	Scientific name	Common name	Chronic Value (µg/L)	SMCV (µg/L)	GMCV (µg/L)	Reference
5	*Bufo boreas*	Boreal toad (tadpole)	3145	3145	3145	This study
4	*Daphnia magna*	Cladoceran	4380	4380	2048	Biesinger and Christensen 1972
	Daphnia pulex	Cladoceran	958	958		Birge et al. 1985
3	*Prosopium williamsoni*	Mountain whitefish	1318	1318	1318	This study
2	*Lumbriculus variegatus*	Worm	870	870	870	This study
1	*Pimephales promelas*	Fathead minnow	910	688	688	Birge et al. 1985
	Pimephales promelas	Fathead minnow	520			Smith et al. 1973

Table 3.2 - EC_{20} values for genera (tribe for Chironomids) present in a 10 d mesocosm experiment exposing ferric iron to naturally colonized communities of benthic invertebrates.

Genus (or tribe/subfamily)	EC_{20} (µg/L)
Rhithrogena sp.	> 14073
Ephemerella sp.	> 14073
Sweltsa sp.	> 14073
Brachycentrus sp.	7558
Baetis sp.	4870
Capnia sp.	3697
Cinygmula sp.	1882
Taenionema sp.	1626
Heterlimnius sp.	1282
Prostoia sp.	1176
Orthocladinae	776
Micrasema sp.	356
Epeorus sp.	335
Tanytarsini	234

References

ASTM (1997) Standard guide for conducting early life-stage toxicity tests with fishes. Standard E 1241-05. American Society for Testing and Materials: West Conshohocken, PA pp 585-612

Benoit DA, Mattson VR, Olsen DC (1982) A continuous flow mini-diluter system for toxicity testing. Water Res 16:457-464

Biesinger KE, Christensen GM (1972) Effects of various metals on survival, growth, reproduction and metabolism of *Daphnia magna*. J Fish Res Board of Can 29:1691-1700

Birge WJ, Black JA, Westerman, AG, Short TM, Taylor SB, Bruser DM, Wallingford ED (1985) Recommendations on numerical values for regulating iron and chloride concentrations for the purpose of protecting warmwater species of aquatic life in the commonwealth of Kentucky. Memorandum of Agreement No. 5429

Boutet C, Chasemartin C (1973) Proprietes toxiques specifiques des sels metalliques chez Austropotamobius pallipes pallipes et Orconectes limosus. Comptes Rendus des Seances de la Societe de Biologie et de ses Filiales 167:1933-1938

Brenner FJ, Cooper WL (1978) Effect of suspended iron hydroxide on the hatchability and embryonic development of coho salmon. Ohio J Sci 78:34-38

Brinkman SF, Crockett HJ, Rogers KB (2013) Upper thermal tolerance of mountain whitefish eggs and fry. Transactions Am Fish Soc 142:834-832

Buchwalter DB, Clements WH, Luoma SN (2017) Modernizing water quality criteria in the United States: a need to expand the definition of acceptable data. Environ Toxicol Chem 36:285-291

Cadmus P, Guasch H, Herdrich AT, Bonet B, Urrea G, Clements WH (2018) Structural and functional responses of periphyton and macroinvertebrate communities to ferric Fe, Cu and Zn in stream mesocosms. Environ Toxicol Chem (*In Press*) doi:10.1002/etc.4070

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

Clements WH, Cadmus P, Brinkman SF (2013) Responses of aquatic insects to Cu and Zn in stream microcosms: understanding differences between single species tests and field responses. Environ Sci Technol 47:7506-7513

Dalzell DJB, Macfarlane NAA (1999) The toxicity of iron to brown trout and effects on the gills: a comparison of two grades of iron sulphate. J Fish Biol 55:301-315

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

Fremling CR, Mauck WL (1980) Methods for using nymphs of burrowing mayflies (Ephemeroptera, *Hexagenia)* as toxicity test organisms. In: Buikema AL Jr, Cairns J Jr (Eds.) Aquatic Invertebrate Bioassays. ASTM STP 715. ASTM, Philadelphia, PA pp 81-97

Goettl JP, Davies PH (1977) Study of the effects of metallic ions on fish and aquatic organisms. Water pollution studies, federal aid in fish and wildlife restoration project F-33-R12. CPW, Fort Collins, CO. pp 37

Gosner KL (1960) A simplified table for staging anuran embryos and larvae with notes on identification. Herpetologica 16:183-190

Hayer CA, Holcomb BM, Chipps SR (2013) Associations between iron concentration and productivity in montane streams of the Black Hills, South Dakota. Prairie Nat 45:68-76

Hem JD (1985) Study and interpretation of the chemical characteristics of natural water. USGS. Water Supply Paper 2254. USGS. Alexandria, VA

Kim AG, Heisey B, Kleinmann R, Duel M (1982) Acid mine drainage: control and abatement research. US Department of Interior, Bureau of Mines information circular 8905 pp 21-22

Kimball BA, Walton-Day K, Runkel RL (2007) Quantification of metal loading by tracer injection and synoptic sampling 1996-2000. In: Church SE, von Guerard P, Finger SE (Eds) Integrated Investigations of Environmental Effects of Historical Mining in the Animas River Watershed, San Juan County, Colorado. USGS, Reston, VA pp 423-495

Linton TK, Pachego MA, McIntyre DO, Clement WH, Goodrich-Mahoney J (2007) Development of bioassessment-based benchmarks for iron. Environ Toxicol Chem 26:1291-1298

Loeffelman PH, Van Hassel JH, Arnold TE, Hendricks JC (1985) A new approach for regulating iron in water quality standards. In: Bahner RC (Ed) Aquatic Toxicology and Hazard Assessment, Eighth Symposium. ASTM STP 891. ASTM, Philadelphia, PA pp 137-152

McKnight DM, Feder GL (1984) The ecological effect of acid conditions and precipitation of hydrous metal oxides in a Rocky Mountain stream. Hydrobiologia 119:129-138

Mebane CA, Schmidt TS, Balistrieri LS (2016) Larval aquatic insect responses to cadmium and zinc in experimental streams. Environ Toxicol Chem 36:749-762

Mineral Policy Center (1997) Golden dreams, poisoned streams. MPC, Washington, D.C.

Ohio EPA (1998) Empirically derived guidelines for determining water quality standards and NPDES discharge limits for iron protective of aquatic life in Ohio rivers and streams. Technical Bulletin *MAS/1998-9-1.* Ohio EPA, Division of Surface Water, Columbus OH

Peuranen S, Vuorinen PJ, Vuorinen M, A. Hollender (1994) The effects of iron, humic acids and low pH on the gills and physiology of Brown Trout (Salmo trutta). Ann Zool Fenn 31:389-396

Piper RG, McElwain IB, McCraren JP, Fowler LG, Leonard JR (1986) Fish Hatchery Management, 2nd edn. US Department of Interior, Fish and Wildlife Service, Washington, D.C.

Radford NP (1997) Ecotoxicological impact of iron III sulphate dosing on chironomid cultures and profundal reservoir communities. Doctoral Thesis, University of Leicester

Rajagopal PK (1979) The embryonic development and the thermal effects on the development of the Mountain Whitefish, *Prosopium williamsoni* (Girard). J Fish Biol, 15:153-158

Smith EJ, Sykora JL (1976) Early developmental effects of lime-neutralized iron hydroxide suspensions on brook trout and coho salmon. Trans Am Fish Soc 105:308-312

Smith EJ, Sykora JL, Shapiro MA (1973) Effect of neutralized iron hydroxide suspensions on survival, growth and reproduction of the fathead minnow (*Pimephales promelas*). J Fish Res Board Can 30:1147-1160

Stephan CE, Mount DI, Hansen DJ, Gentile JR, Chapman GA, Brungs WA (1985) Guidelines for deriving numerical standards for the protection of aquatic organisms and their uses. PB85-227049. USEPA, Springfield, VA

Steffens W, Mattheis TH, Riedel M (1993) Field observations on the production of Rainbow Trout (*Oncorhynchus mykiss*) under high concentrations of water-borne iron. Aquat Sci 55:173-178

Sykora JL, Smith EJ, Synak M, Shapiro MA (1975) Some observations on spawning brook trout (*Salvalinus fontinalis* Mitchell) in lime neutralized iron hydroxide suspensions. Water Res 9:451-458

Thurston RV, Russo RC, Fetterolf Jr. CM, Edsall TA, Barber Jr. YM (1979) A Review of the EPA Red Book: Quality Criteria for Water. AFS, Bethesda, MD

Updegraff KF, Sykora JL (1976) Avoidance of lime-neutralized iron hydroxide solutions by coho salmon in the laboratory. Environ Sci Technol 10:51-54

USDA Forest Service (1993) Acid mine drainage from mines on the National Forests: A management challenge. USDA, Washington, D.C.

USEPA (1976) Quality Criteria for Water. PB263-943. USEPA, Washington, D.C.

USEPA (1999) 1999 update of aquatic life ambient water quality criteria for ammonia. EPA 822-R-99-014. USEPA, Washington, D.C.

USEPA (2015a) Toxicity Relationship Analysis Program (TRAP Version 1.30a). USEPA. Duluth, MN

USEPA (2015b) ECOTOX User Guide: ECOTOXicology Knowledgebase System. Version 4.0. https://cfpub.epa.gov/ecotox/blackbox/help/userhelp4.pdf. Accessed 28 April 2017

Vuori KM (1995) Direct and indirect effects of iron on river ecosystems. Ann Zool Fenn 32:317-329

Warne, M.S.J. and R. van Dam. 2008. NOEC and LOEC data should no longer be generated or used. Autralasian Journal of Ecotoxicology 14: 1-5

West Inc (1996) Toxstat Version 3.5. Western EcoSytems Technology, Cheyenne, WY

Williams DA (1971) A test for differences between treatment means when several dose levels are compared with a zero dose control. Biometrics 27:103-117

Williams DA (1972) The comparison of several dose levels with a zero dose control. Biometrics 28:519-531

CHAPTER 4

EXPERIMENTAL ASSESSMENT OF SIZE-DEPENDENT SENSITIVITY OF AQUATIC INSECTS TO METALS – HOMAGE TO HORTON HEARS A WHO!

Introduction

In the classic children's book "Horton Hears a Who!" (Seuss 1954) Horton the elephant discovers a population of microscopic organisms named "Whos." Other large vertebrates in the Jungle of Nool were oblivious to the existence of such small organisms. After further study and debate these megafauna instated policies protective of small organisms. This is not unlike the current understanding of aquatic communities. Ecological studies are often limited to macrofauna because of the difficulty in sampling meiofauna and microorganisms. Species sensitivity distributions rarely include microorganisms. Although all aquatic insects start life as nearly microscopic size classes, most toxicology studies use only larger or older age classes. Policies and numeric standards based on traditional toxicity experiments using mature aquatic insects may be underprotective if smaller age classes are more sensitive to pollution.

Links between metal pollution and degradation of aquatic communities in streams are well established in the literature (Clements et al. 2000; Mebane 2001; Maret et al. 2003; Cain et al. 2004; Brix et al. 2005; Herbst et al. 2018; Hornberger et al. 2009). Laboratory experiments have routinely demonstrated that aquatic insects are tolerant to trace metals (Brix et al. 2005; Brinkman and Johnston 2008; Brinkman and Johnston 2012; Mebane et al. 2012); however, biomonitoring studies often indicate that aquatic insects are sensitive to metals at relatively low concentrations (Clements et al. 2000; Clements 2004; Buchwalter et al. 2007; Schmidt et al. 2010; Clements et al. 2013). This discrepancy in reported metal tolerance may be the result of

invertebrate assemblage size structure (Kiffney and Clements 1996; Clark and Clements 2006; Clements and Kotalik 2019, *in review*). Natural benthic communities contain a diversity of taxa that can widely differ in their rates of development due mainly to phenology (i.e., seasonal environemental cues for development) and voltinism (i.e., number of life cycles per year). These life history traits are spatially and temporally variable, resulting in a diversity of invertebrate developmental sizes within and among different species that differ in response to metal expsoure. All aquatic insects hatch from eggs as nearly microscopic first instar larva. However, the smaller developmental sizes are seldom used in laboratory toxicity tests.

Laboratory and mesocosm experiments that have compared early and late life stages of aquatic invertebrates have reported greater sensitivity of smaller size classes (Powlesland and George 1986; McCahon et al. 1989; Diamond et al. 1992; Stuhlbacher et al. 1993; Kiffney and Clements 1996; Clark and Clements 2006; Soucek and Dickinson 2015). In their seminal study of phylogenetic influences on metal sensitivity in aquatic insects, Buchwalter et al. (2008) controlled for the potential confounding effect of body size on species sensitivity to metals. These differences may in part result from the influence of surface area to volume ratios and the related size-dependent uptake and turnover rates of major body ions (Grosell et al. 2002). Additionally, lower fat to protein ratios, more rapid accumulation of toxicants in organs, less developed antioxidant systems and less developed physical structures may contribute to the increased sensitivity of smaller organisms (Mohammed 2013). Potential or maximum body size is also considered an important species trait in predicting aquatic macroinvertebrate colonization and occupancy in disturbed habitats, and it has been used as a biotic indicator of metal pollution (Archaimbault et al. 2005; Statzner et al. 2005; Doledec and Statzner 2008; Statzner et al. 2001; Statzner and Beche 2010; Pomeranz et al. 2018).

Early instar aquatic insects are typically too small to collect in the field or manipulate in the laboratory (Figure 4.1). Because of this the current understanding of aquatic insect metal sensitivity is based predominately on larger instars. Mesocosm studies have improved predictions of metal sensitivity in the field by integrating naturally colonized communities that contain numerous taxa at differing stages of development, including early instars (Kiffney and Clements 1996; Clark and Clements 2006; Clements et al. 2013). Despite this, few experimental studies directly address the relationship between aquatic insect size and metal sensitivity. This distinction is important because differences in metal sensitivity among aquatic invertebrates are used to generate species sensitivity distributions (SSD) that serve as the basis for deriving water quality standards (Stephan et al. 1985; Von der Ohe et al. 2004), but logistic challenges in obtaining, culturing, and/or testing early instars may bias these SSDs. Although many standardized testing procedures encourage the use of early life stages or full life cycle trials for vertebrates (e.g., fish), similar experiments are rarely conducted for aquatic insects.

Herein, we report the results of a series of mesocosm and laboratory experiments that test the hypothesis that early life stages of aquatic insects are more sensitive to metals than mature, later instars. We tested the following specific hypotheses: 1) metal sensitivity increases as body mass decreases for Ephemeroptera, Plecoptera and Trichoptera (EPT) species; 2) head capsule width (i.e., body size) and metal concentration is a better predictor of aquatic insect mortality than metal concentration alone; 3) smaller size classes of 4 common aquatic insect species are more susceptible to metal mixtures than larger size classes; and 4) acute median lethal concentrations (LC_{50} values) for three age classes of *Baetis tricaudatus* (Dodds) exposed to Zn increase as age class increases (i.e., older age class are less sensitive because they are larger in size).

Methods

Overview

We examined aquatic insect size distributions from mesocosm studies that exposed natural benthic macroinvertebrate communities to different metal combinations (Cu, Zn, Cd). Macroinvertebrate head capsule widths and body mass are commonly used to estimate invertebrate size (Benke et al. 1999). We measured the head capsule width of the mayfly *Baetis* spp., as well as taxa from three other dominant aquatic insect orders (*Isoperla* spp., Plecoptera; *Hydropsyche* sp., Trichoptera; and Orthocladiinae, Diptera). A diversity of taxa and body sizes were used to evaluate inter- and intraspecific metal sensitivity. Similarly, average mass of each taxon from each mesocosm experiment was used to estimate sensitivity across metal concentrations. We hypothesized that taxa with lower mass would exhibit a wide range of sensitivity to metals, whereas larger taxa would be consistently tolerant. Lastly, acute Zn toxicity tests were conducted using first instar (< 24 h post-hatch, originating from field collected eggs) and mid-instar mayflies (~ 1 mo post-hatch, field collected). We then compared results from these early life stages to results from late instars obtained under identical laboratory conditions by Brinkman and Johnston (2012).

Mesocosm experiments

In four previous mesocosm experiments, naturally colonized benthic communities were exposed to different combinations of metals (Cu alone, September 2007; Cu and Zn, October 2007; Cu, Cd and Zn, August 2010; Cu and Zn, September 2015) at the Colorado State University Stream Research Laboratory (SRL; Fort Collins, Colorado, USA) . Details of the SRL design and water chemistry have been described previously (Clements et al. 2013). The 2007-2010

experiments were 10 d exposures, and the 2015 experiment was a 14 d exposure. Mesocosm experiments conducted in 2007-2010 exposed benthic communities from the South Fork of the Michigan River (Gould, Colorado, USA); the experiment conducted in 2015 exposed benthic communities from the Arkansas River (Leadville, Colorado, USA). At the end of each experiment, benthic organisms retained in a 355 µm sieve were preserved in ethanol (80%), and individuals were enumerated and identified to the lowest practical level of taxonomic resolution. Because these experiments used different combinations of metals, cumulative criterion units (CCUs) based on the U.S. Environmental Protection Agency's hardness-adjusted criteria were used to quantify metal concentrations in the mesocosms (Appendix D, Table 4.S1) (Clements et al. 2013). Other factors that influence metal toxicity and bioavailability (e.g., pH, dissolved organic carbon) were consistent among treatments and experiments. Because models predicting bioavailability of these metal mixtures were unavailable, hardness-adjusted criteria were used for these analyses. Detailed water chemistry measured in these experiments is listed in supplemental information (Appendix D, Table 4.S3).

Head capsule widths of *Baetis* spp. from 2007-10 experiments were measured using a stereo microscope (Meji EMZ-TR) with a reticle SFW20x eyepiece that provided 0.1 mm resolution. Greater measurement resolution was achieved with the 2015 experiment, which used a high definition microscopy camera (ACCU-SCOPE® *Excelis* Camera AU-600-HD) attached to a stereoscope (Meji EMZ-TR). A stage micrometer (0.01 mm precision) was used to calibrate measurements, and three observations were taken on each individual.

To quantify invertebrate body mass, the wet (preserved in 80% ethanol) mass of every organism from controls of the 2007-2010 experiments was measured. Preserved organisms were placed on dry filter paper on a Buchner funnel for 30 s and weighed (O'Hause GS200D balance;

0.00001 g resolution). Average organism mass of each EPT taxon in controls was calculated and log transformed (ln(mg+1)). Relative abundance after 10 d of exposure (expressed as a proportion of the mean abundance in controls; n=2) for each taxon in each experiment was log transformed (ln([Abundance]/[Average Abundance In Controls]+1)) and regressed on log-transformed CCUs (ln(CCU+1)). The "LM" function in package 'car' in R (R Core Team v 3.5.1) was used to estimate slope (Fox and Weisberg 2011). The reverse sign of each respective slope estimate was used as a measure of sensitivity for each taxon in each experiment. Lastly, weighted regression ('weights=') was used to regress sensitivity values from all three mesocosm experiments across body mass. Abundance of each taxon in the controls was used as the weight in the regression analyses to ensure poorly represented taxa did not have a disproportionate influence on the relationship between mass and metal tolerance. Dipterans (true flies), coleopterans (beetles) and non-insect taxa from mesocosms were not consistently represented among experiments and were not included in the analysis.

Because the 2007-2010 mesocosm experiments employed a regression experimental design with low replication (n=2), analysis of covariance (ANCOVA) was used to estimate the slope of the relationships between abundance of surviving organisms and metal concentration for different size classes of *Baetis* spp. After combining control head capsule distributions among the three experiments, size classes were determined using the "split" function in package 'Hmisc' (Harrell and Dupont 2006). Size class distributions varied among these three experiments due to their different colonization periods and phenological differences; therefore, uneven size class groupings were chosen to allow for absolute size comparisons of *Baetis* sp. among the three experiments. The number of surviving organisms and the metal concentrations were transformed (ln + 1) to satisfy assumptions of parametric statistics. Using the "LM" function, survival was

regressed on metal concentration. Akaike Information Criterion (AIC) (Burnham and Anderson 2004) was used to select the model that best predicted mortality based on insect size, metal concentration, and (or) the size x metal interaction. To identify differences in responses to metals among size classes, we used the "estimated marginal means of linear trends" (emtrends) function in package 'emmeans' (Lenth 2018), and the 'multcompLetters' package (Graves et al. 2012) with a Tukey HSD multiple-comparison adjustment.

The 2015 mesocosm study was designed with greater replication (n=3) to allow use of two-factor ANOVA (package 'car') to test the hypothesis that differences in mortality across metal treatments were determined by insect body size (i.e., head capsule width). To separate size classes for each aquatic insect order, the "split" function was used to fit either 6 or 7 size class groupings. Evenly separated size class groupings were used because we wanted to compare size gradient responses to metals among taxa. Because of differences in abundance among head-capsule size groupings, abundance data were normalized to proportion mortality relative to mean control abundance (n=3) for each size grouping.

Effects on early instar Baetis tricaudatus in the laboratory

Early life stages of the mayfly *B. tricaudatus* were exposed to a gradient of Zn concentrations for 96 h. Early instar organisms (mean head capsule width 113.5 µm SD=10 n=7) were obtained by rearing eggs. Egg masses were collected from the Cache la Poudre River (Colorado, USA) in September 2014 substrate (Appendix D, Figure 4.S1). Mid-instar organisms (mean head capsule width 260.1 µm SD=25 n=7) of ~1 month age, were collected from cobble at the same location using 7.5 ml transfer pipettes (16 November 2014). Early and mid-instar baetids were nearly microscopic and were contained and enumerated using a novel toxicant exposure system that reproduced the natural flows of benthic habitats in high-gradient streams without losing

organisms (Appendix D, Figures 4.S2-4.S5). Importantly, this acute exposure to Zn used the same exposure methodology and dilution water supply as described by Brinkman and Johnston (2012) for large instars. After initial range-finding experiment (s) for each size, first instars were exposed to 0, 133, 300, 642, 1433, and 3263 µg/L Zn (26 Oct 2014; Appendix D, Table 4.S6). Mid-instars were exposed to 0, 4600, 9380, 20450, 46550, 84800 µg/L Zn (16 Nov 2014). Because phenotypic characteristics used to identify *Baetis* spp. are not developed until organisms are more mature (late instar), a subsample of surviving organisms from each experiment was preserved for genetic analysis. Ninety-six h LC_{50} values for first and mid-instar size classes were calculated using the dose response model function ("drm") in package 'drc' (Ritz and Streibig 2016).

Results

Routine water quality characteristics (pH, hardness, conductivity, temperature) measured in stream mesocosms were similar among the 4 experiments and showed relatively little variation among treatments. Water chemistry in the 2010 and 2015 mesocosm experiments (Cu+Zn+Cd and Cu+Zn) were very similar to experiments conducted in 2007 (Cu+Zn and Cu; Clements et al. 2013), with sourced water representative of oligotrophic headwater streams **(Appendix D, Tables 4.S2-4.S4)**.

Body mass of aquatic insects in the 2007-2010 mesocosm studies ranged from 0.013 to 36.8 milligrams. Sensitivity to metals significantly decreased as body size increased across the dominant EPT taxa (Slope: -0.0806, $p < 0.0001$; Figure 4.2). As predicted, smaller taxa had the greatest range in sensitivity to metals, whereas larger taxa were represented only by metal-

tolerant organisms. This wedge shaped response distribution contributed to the relatively low r^2 (0.31) for this regression.

Head capsule widths for *Baetis* spp., the dominant mayfly in the 2007-2010 mesocosm experiments, was an important addition to CCU in predicting mortality (Appendix D, Table 4.S5). Instar size x metal concentration interaction terms were significant for Cu (p=0.0082) and Cu+Zn+Cd (p=0.0171), but not for Cu+Zn (p=0.2395). AIC results support the addition of the interaction term for all ANCOVA models, indicating that size and CCU better explained mortality across treatments compared to CCU or instar size alone (Table 4.1).

In the 2015 mesocosm experiment, survival of *Baetis* spp. (Ephemeroptera), Orthocladiinae (Diptera), *Isoperla* (Plecoptera), and *Hydropsyche* sp. (Trichoptera) decreased with CCU but increased with instar size ($p < 0.05$, Table 4.2 and Figure 4.3). Therefore, including body size in the regression improves model predictions. Additionally, interaction terms (CCU x Instar Size) for *Baetis* spp., Orthocladiinae, and *Hydropsyche* spp. were statistically significant. Body size of the stonefly *Isoperla* spp. seemed to influence responses to metals, but the interaction term was not significant (p=0.0762) likely due to high variability among treatments. In general, the greatest mortality was observed for smaller instars (i.e., lower mortality as organisms become larger; Figure 4.3). Treatment effects for Trichoptera were highly size-dependent, with less than 5% mortality at 53 CCU for the largest instars (>1.05 mm), while the smallest instars (< 0.30 mm) had greater than 50% mortality even in the lowest treatment (4 CCU). The slopes describing the body size-survival relationship of Orthocladiinae were similar across treatments, whereas the influence of size for *Baetis* and *Isoperla* was more pronounced at the lower metal concentrations due to high or complete mortality in the higher treatments.

Acute toxicity of Zn to the mayfly *Baetis tricaudatus* in the single-species experiment decreased as organism size increased (Figure 4.4; Appendix D, Table 4.S6). LC_{50} values for first and mid-instar *B. tricaudatus* were 600.1 (±460.5-782.1) µg Zn/L and 6094.3 (±4946.2-7509.1) µg Zn/L, respectively. These experiments were conducted in the same laboratory and used the same water sources as *Baetis tricaudatus* experiments described by Brinkman and Johnston (2012) who reported LC_{50} values of 10,020 µg/l. Water quality (Appendix D, Table 4.S3 and 4.S6) did not differ between these studies.

Discussion

We present several lines of evidence that body size of aquatic insects is a strong predictor of metal sensitivity, with greater sensitivity observed in smaller individuals than in larger individuals. The naturally colonized benthic communities used in the mesocosm studies incorporated a diverse size structure within and among taxa. This enabled us to evaluate aquatic insect responses to metals across numerous taxonomic groups and developmental size classes. At metals concentrations in which partial mortality occurred, smaller organisms were consistently more sensitive than larger organisms. Size-dependent responses of *Baetis* spp., the dominant mayfly in many western streams (Ward et al. 2002; Merritt et al. 2008, McCafferty et al. 2012), occurred in the four mesocosm experiments and in the single-species toxicity tests. Across all taxa, metal sensitivity was inversely correlated with body mass, with small organisms displaying a wide range of sensitivity to aqueous metals, but large organisms, regardless of species, displaying greater tolerance. Importantly, size-dependent sensitivity occurred even in taxa that are generally considered tolerant to metal exposure. For example, laboratory and field studies have demonstrated that hydropsychid caddisflies are highly tolerant to metals (Cain and Luoma

1998; Clements et al. 2000; Mebane et al. 2012), but in my study *Hydropsyche* spp. had the most pronounced size-dependent treatment effects, with greater than 50% mortality of early instars (< 0.3 mm) in the lowest metal concentrations (4 CCU).

Consistent with my hypothesis, small size classes had a range of sensitivities to metals, but taxa represented primarily by large size classes (e.g. *Drunella* spp., *Arctopsyche* sp., *Brachycentrus* sp.) were only tolerant. All aquatic macroinvertebrates hatch as small-bodied individuals, and selection against sensitive taxa likely occurs during these early stages of development. Phylogenetic differences in acclimating to stressors is perhaps of greatest importance for early instars. Observational studies have demonstrated that maximal body size is a trait commonly associated with taxa at contaminated sites (Statzner et al. 2001; Archaimbault et al. 2005; Statzner et al. 2005; Doledec and Statzner 2008; Statzner and Beche 2010; Pomeranz et al. 2018). However, observational studies are limited in addressing these relationships because immigration and emigration are not controlled, whereas my experiments measured the direct toxicological effects experimentally. It is possible that maximal body size predicts which taxa can immigrate and survive at a site, but minimal body size at a site might better explain which species can actually complete their full life cycle.

Single-species laboratory studies with aquatic insects routinely suggest that these organisms are highly tolerant to metals (Brix et al. 2005; Brinkman and Johnston 2008; Brinkman and Johnston 2012; Mebane et al. 2012). Laboratory experiments using field-collected aquatic insects (i.e., *Drunella doddsii*, *Ephemerella* sp., *Cinygmula* sp., *Lepidostoma* sp. and Chloroperlidae) report LC_{50} ranging from 32,000 to 64,000 µg Zn/L (Brinkman and Johnston 2012). In these studies, larvae were large enough to be collected by hand, and survival was easily assessed without magnification. These LC_{50} values are orders of magnitude higher than

thresholds reported in mesocosm experiments and field studies (Clements et al. 2000; Clements et al. 2004; Schmidt et al. 2010). Although other environmental factors such as colonization dynamics, drift and emergence propensity, and duration of life cycle likely contribute to laboratory and field discrepancies, my results strongly suggest that the developmental size progression of aquatic insects influences metal sensitivity. These results may also explain, in part, why laboratory experiments typically demonstrate aquatic insects are tolerant to metals, while mesocosm and field studies in contrast indicate they are quite sensitive.

The physical and chemical cues that influence the phenology of macroinvertebrates in the field likely affect their spatiotemporal sensitivity to contaminants. For example, environmental cues such as degree days, stream flow, and day length influence hatching, adult aquatic insect emergence, diapause, and secondary production (Benke 1979; Vannote and Sweeney 1980; Peckarsky 2000). Seasonal fluctuations in metal concentrations may co-occur with the presence of sensitive or tolerant life stages, and changes in water chemistry may affect certain life stages of some taxa but not others based on their timing of development.

Benthic survey comparisons in the Rocky Mountains have demonstrated the influence of insect phenology on metal sensitivity along elevation gradients and among seasons (Kiffney and Clements 1996; Clark and Clements 2006). Although we generally observed greater mortality in less mature instars, the results were complicated by the concurrent emergence of larger organisms during my experiments. For example, *Baetis* spp. in the October 2007 experiment was dominated by late instars. It is possible that some of the lower abundances in larger size classes that we attributed to larval mortality were at least partly the result of adult emergence, which was not quantified in these experiments. Toxicity models need to better incorporate early instar sizes and differentiate sensitivity throughout an organism's life cycle. Moreover, linking invertebrate

phenology to the temporal changes of contaminant concentrations in the field will better characterize exposure outcomes.

Standard testing guidelines (e.g., Stephan et al. 1985; American Society for Testing and Materials 1993; USEPA Office of Water 2002) have long noted the importance of using early life stages in toxicity tests. These same policies limit "acceptable" mortality in controls to 5-10%, a requirement likely intended to limit the risk of erroneously determining a toxic effect when none exists. Starting in the early twentieth century, ecologists have used the concept of survivorship curves (Figure 4.5) to describe the natural rates of mortality throughout an organism's lifespan (Deevey 1947; Pearl and Miner 1935). Fish and aquatic insects generally display a Type III survivorship curve, with high mortality in early life stages (dashed box in Figure 4.5) and a lower mortality in later life stages (solid box in Figure 4.5). High mortality in early life stages can be attributed to predation, limited resources, competition, and the stochastic mortality commonly observed in r-selected species. For example, Willis and Hendricks (1992) conducted a comprehensive study of the population dynamics of the caddisfly *Hydropysche slossonae* in an undisturbed river and observed first instar mortality approaching 93%. These high rates of natural mortality would be unacceptable in the current testing guidance (Stephan et al. 1985; American Society for Testing and Materials 1993; USEPA Office of Water 2002) This, illustrates the challenges associated with developing test protocols for aquatic insects that balance environmental realism and laboritory control. Early instar toxicity tests are rarely attempted or the results are excluded from criteria/guideline derivation datasets. More research is needed to characterize background mortality of early instars of aquatic insects, so benchmark "acceptable" control mortality can be established for early life stages.

The novel single-species toxicity test methodology presented in this paper, along with the ability to genetically identify species before diagnostic morphological characteristics develop, improves the ability to test the responses of early instars of aquatic insects to contaminants. The toxicity test method incorporates flow in a way that better simulates hyporheic hydrologic processes (e.g., exchange of dissolved oxygen and water, and toxicant replenishment) and enables handling and enumeration of small early instars. Although this method routinely produced acceptable control survival (94-100%), success may be limited to species that oviposit in clusters (pads) and have higher rates of survival in early age classes. This experiment was only possible after a decade of efforts to culture numerous mayfly species, in which *Baetis* was found to be the most tolerant of laboratory conditions. Although these methods produced an acute LC_{50} value for early instars at 6% of the value obtained from late instars, even surrogate test species like this might routinely fail to represent the sensitivity of aquatic invertebrates found in natural communities (Cairns 1986). This stark limitation implies the need to develop more innovative testing methods and/or ways to incorporate streammesocosm results into the development of water quality guidelines and criteria (Buchwalter et al. 2017).

My results demonstrate that aquatic insect body size is a strong predictor of susceptibility to metals. Size-dependent responses occurred among multiple aquatic insect orders, with smaller invertebrates generally displaying greater susceptibility to metals than in larger, mature invertebrates. The addition of body size improved toxicity model fit compared to metal concentration alone. Testing methodologies used to establish water quality criteria would benefit by mandating early life stage testing of aquatic insects. Additionally, improved field sampling methods that target these small but particularly sensitive life stages would improve the ability to characterize effects in the field. Toxicity models that account for the sensitive life stages of

aquatic insects have the potential to improve the accuracy in predicting effects of contaminants in the field.

All aquatic insects hatch as nearly microscopic organisms and small size classes were consistently the most sensitive in my experiments. These findings have important implications for biomonitoring studies designed to assess effects of contaminants. Field studies typically use sampling procedures that retain only large benthic organisms (e.g. 500 or 350 μm mesh). Early instars are not retained in these samples so effects of metals and other stressors may be underestimated in the field. Sampling procedures that collect early instars (e.g., smaller mesh sizes) have the potential to improve ecotoxicological studies. Sampling small age classes in nature and conducting toxicity trials with small age classes is difficult and therefore these studies are lacking from the scientific literature. Failure to characterize sensitivity of early size classes may lead to gross overestimation of tolerance. To paraphrase Horton in reference to Who-ville (Seuss 1954), “an [insect’s] an [insect] no matter how small.”

Figures and Tables from Chapter 4

Contains:

Figures 4.1 to 4.5

Tables 4.1 and 4.2

See Appendix D for Figures 4.S1 to 4.S5 and Tables 4.S to 4.S6.

Figure 4.1 - *Baetis* spp. A): first or early instar 96 to 108 h post-hatch from single-species experiments; B) and C): mid-instars from single-species experiments that were field collected 30 d after egg masses were observed hatching; D): late instar typical of field collected organisms in mesocosm communities. Height of T in "TRUST" is ~800 µm.

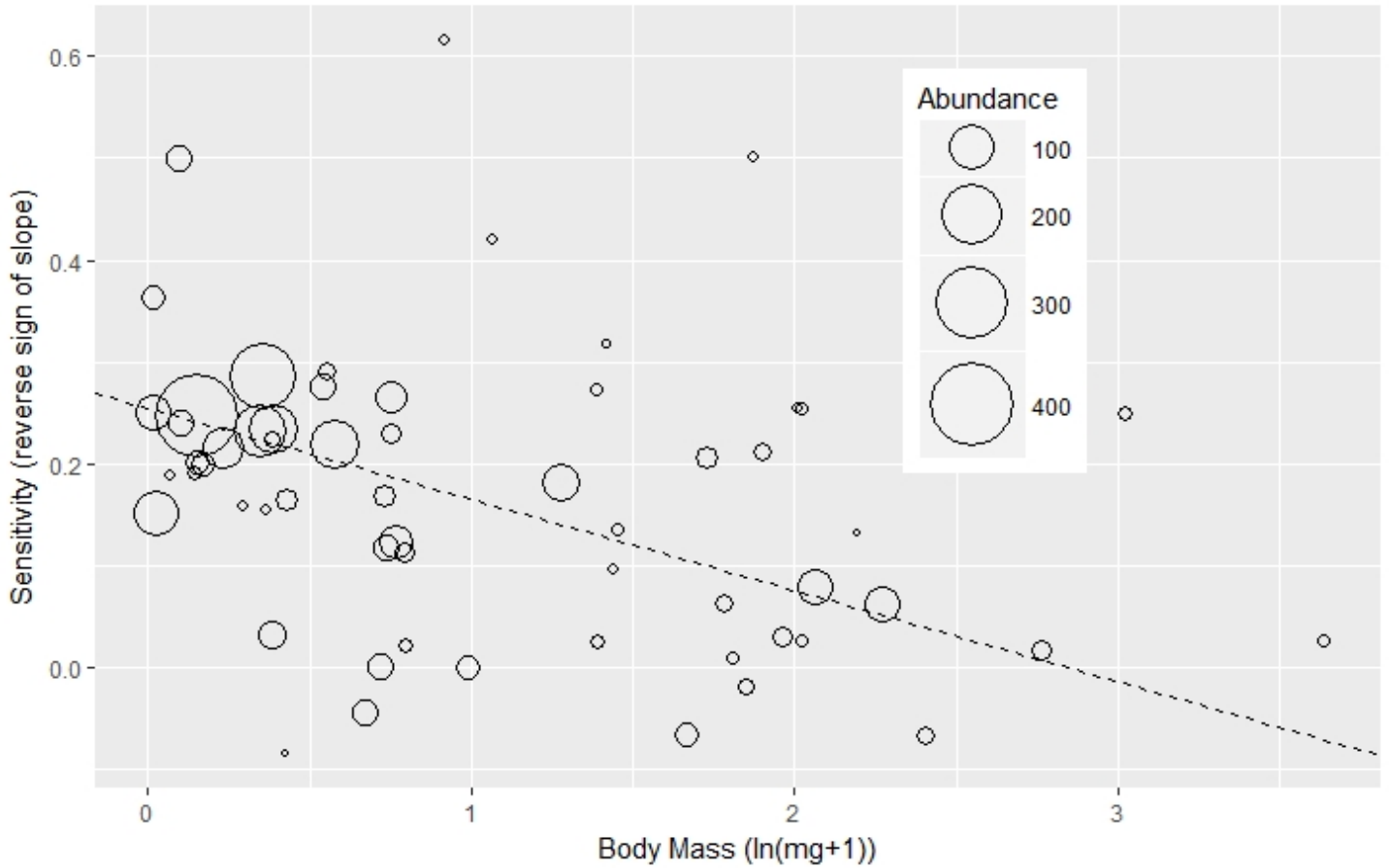


Figure 4.2 - Relationship of sensitivity index to body mass of EPT (Ephemeroptera Plecoptera and Trichoptera) larvae in the 2007 (Cu and Cu+Zn) and 2010 (Cu+Zn+ Cd) mesocosm experiments combined. Sensitivity index equals the reverse sign of the slope of ln([Abundance]/[Average Abundance In Controls]+1) regressed on ln(CCU+1), where CCU=chronic criterion units for the metal(s). Diameter of the points reflects average abundance in controls for each taxa at the end of the experiment. Dashed regression line was weighted for average abundance in controls (Slope:-0.08052 (±0.01606) $p < 0.0001$. Intercept: 0.24119 (±0.01399) $p < 0.0001$.)

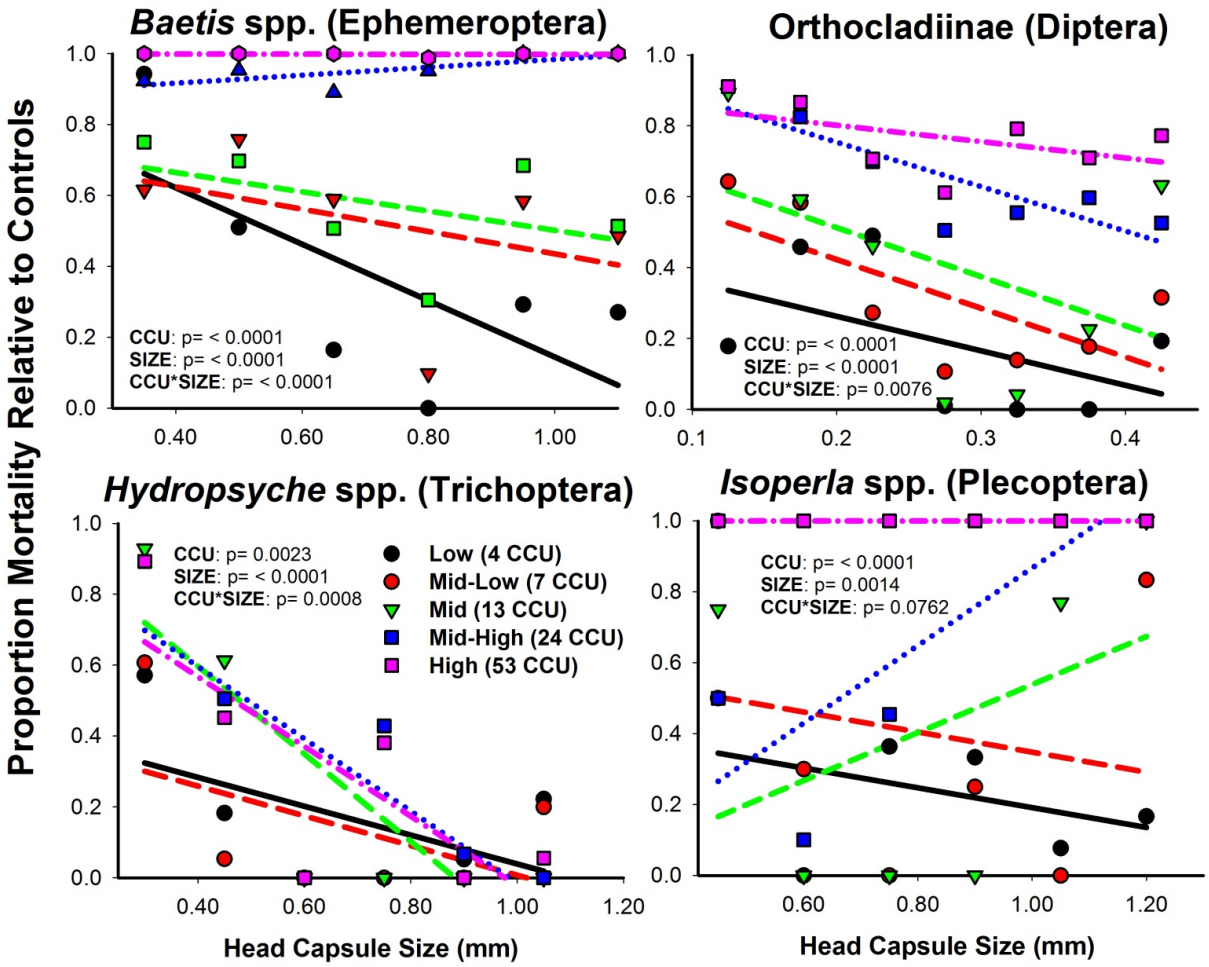


Figure 4.3 - Relationships between mortality and head-capsule size in mesocosm experiments (Cu + Zn) from 30 August to 12 September of 2015. Each symbol represents the average proportional mortality of three replicates in each treatment level.

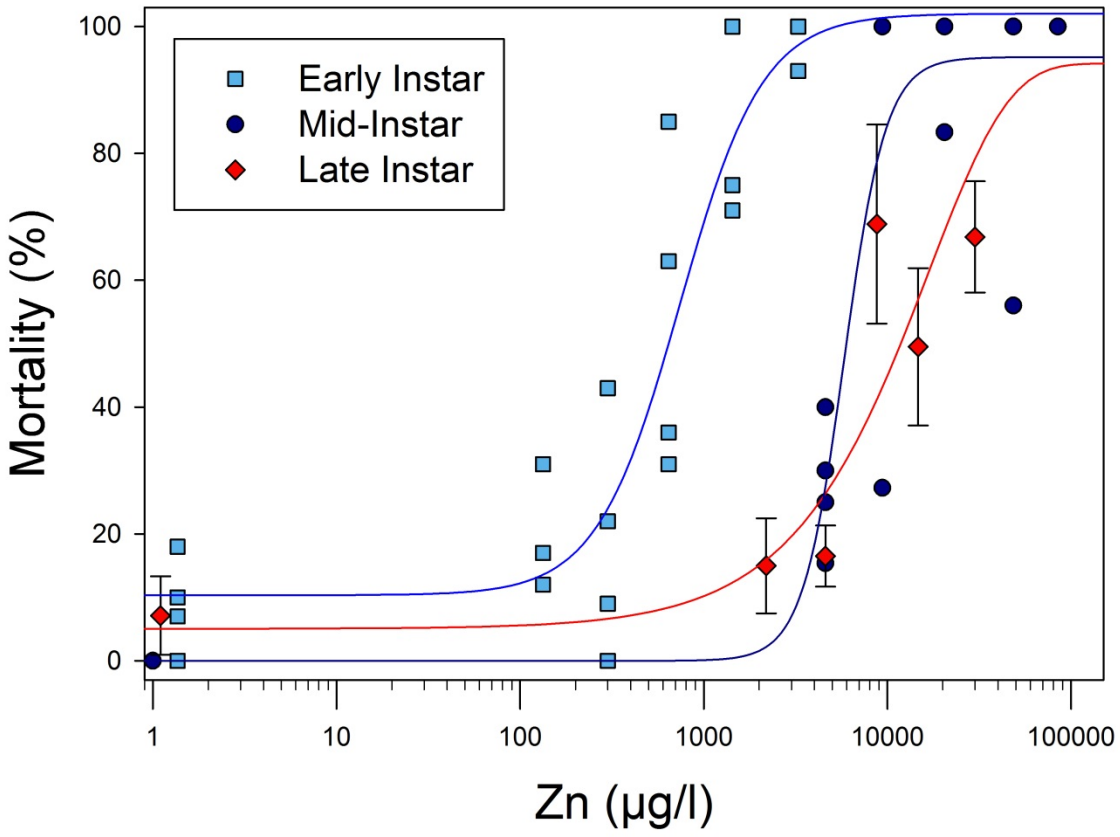


Figure 4.4 - Mortality of early instars (light blue squares) and mid-instars (dark blue circles) of *Baetis tricaudatus* after 96 h exposure to Zn. Results from Brinkman and Johnston (2012; red diamonds) are included for a comparison to late instars. (± s.e.; n=4).

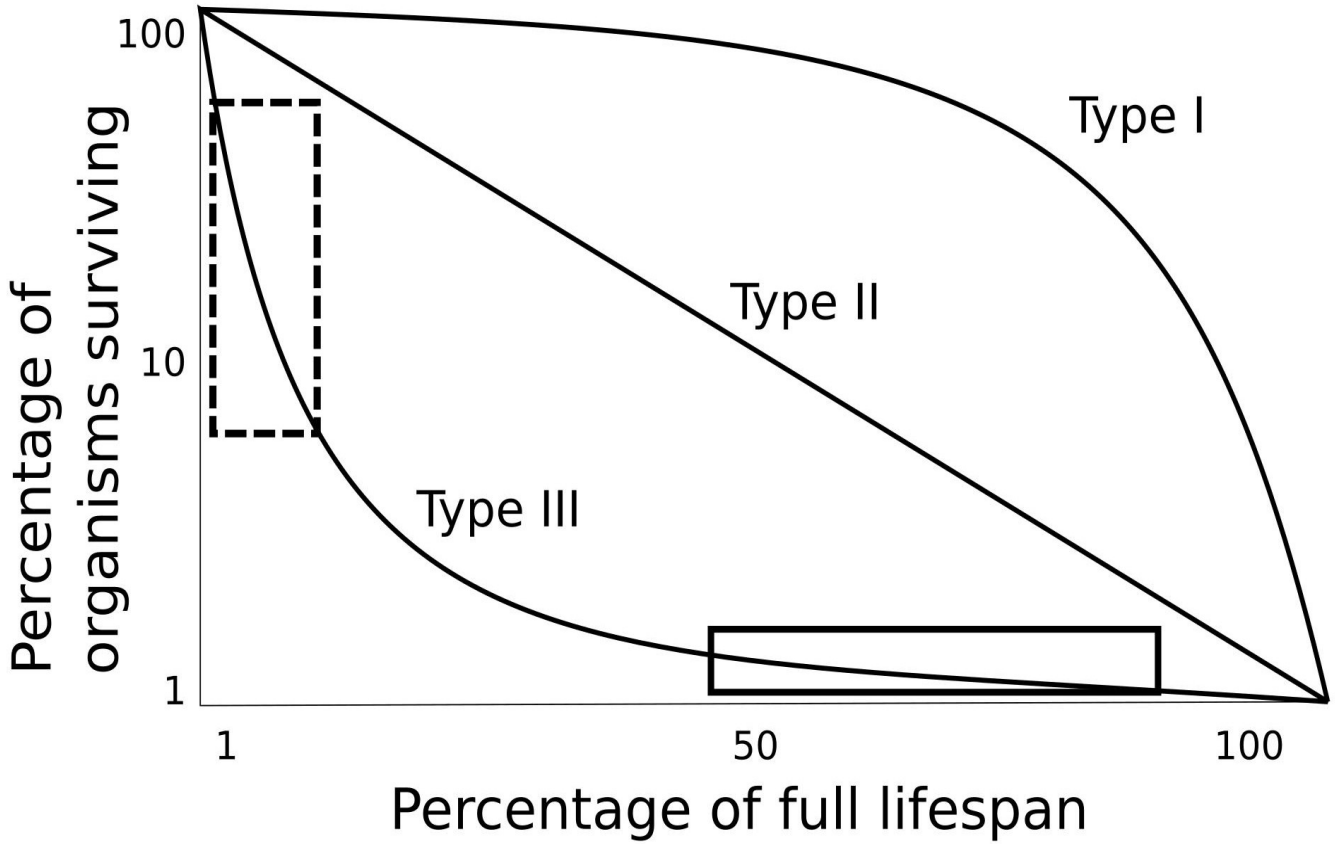


Figure 4.5 - Conceptual survivorship curves are commonly used by ecologists to characterize life history traits. Fish and insects generally occupy a Type III curve, whereas longer lived species such as large mammals typically occupy a Type I curve. The dashed box includes early, more sensitive life stages; the solid box represents larger, more tolerant age classes. Mortality in controls similar to that in the dashed box would be deemed unacceptable in standardized testing guidelines, but it is common in natural aquatic communities.

Table 4.1 - Akaike Information Criterion (AIC) model selection results for model terms in 2007-2010 mesocosm experiments. ANCOVA was used to test for the responses of *Baetis* spp. abundances given the model predictors of metals concentration(s) (CCU), Instar Size, and the CCU x Instar Size interaction.

Treatment	Model Term	AIC	Delta-AIC
Cu	CCU x Instar Size	234.64	0
0 - 5.1 CCU	CCU, Instar Size	244.03	9.39
	Instar Size	250.36	15.72
	CCU	280.45	45.81
Cu+Zn	CCU, Instar Size	247.88	0
0 - 7.0 CCU	CCU x Instar Size	249.05	1.17
	Instar Size	252.96	5.08
	CCU	265.45	16.40
Cu+Zn+Cd	CCU x Instar Size	312.89	0
0 - 12.9 CCU	CCU, Instar Size	319.18	6.29
	Instar Size	346.31	33.42
	CCU	368.07	55.18

Table 4.2 - Two-factor ANOVA results from the 2015 experiments in which metal concentration(s) (CCU), Instar Size, and CCU x Instar Size interaction were used to predict mortality of the four dominant taxa: *Baetis* spp. (Ephemeroptera), Orthocladiinae (Diptera), *Hydropsyche* spp. (Trichoptera), and *Isoperla* spp. (Plecoptera).

Taxa	Model Term	F-Value	P-value
Baetis spp.	CCU Treatment	100.12	<0.0001
	Instar Size	14.68	<0.0001
	CCU Treatment*Instar Size	4.50	<0.0001
Orthocladiinae	CCU Treatment	53.43	<0.0001
	Instar Size	17.99	<0.0001
	CCU Treatment*Instar Size	2.13	0.0076
Hydropsyche spp.	CCU Treatment	4.56	0.0023
	Instar Size	53.73	<0.0001
	CCU Treatment*Instar Size	2.75	0.0008
Isoperla spp.	CCU Treatment	14.75	<0.0001
	Instar Size	4.54	0.0014
	CCU Treatment*Instar Size	1.62	0.0762

References

American Society for Testing and Materials (ASTM). *Standards on aquatic toxicology and hazard evaluation* 1993, ASTM Publication Code 03-547093-16.

Archaimbault, V.; Usseglio-Polatera, P.; Vanden Bossche, J. P. Functional differences among benthic macroinvertebrate communities in reference streams of same order in a given biogeographic area. *Hydrobiologia* 2005, 551 (1), 171-182.

Benke, A. C. A modification of the Hynes method for estimating secondary production with particular significance for multivoltine populations. *Limnology and Oceanography* 1979, 24 (1), 168-171.

Benke, A. C.; Huryn, A. D.; Smock, L. A.; Wallace, J. B. Length-mass relationships for freshwater macroinvertebrates in North America with particular reference to the southeastern United States. *Journal of the North American Benthological Society* 1999, 18 (3), 308-343.

Brinkman, S. F.; Johnston, W. D. Acute toxicity of aqueous copper, cadmium, and zinc to the mayfly *Rhithrogena hageni*. *Archives of Environmental Contamination and Toxicology* 2008, 54 (3), 466-472.

Brinkman, S. F.; Johnston, W. D. Acute toxicity of zinc to several aquatic species native to the Rocky Mountains. *Archives of Environmental Contamination and Toxicology* 2012, 62 (2); 272-281.

Brix, K.V.; DeForest, D.K.; Adams, W.J. Assessing acute and chronic copper risks to freshwater aquatic life using species sensitivity distributions for different taxonomic groups. *Environmental Toxicology and Chemistry* 2001, 20 (8), 1846-56.

Brix, K. V.; D. K. Deforest, Burger, M.; Adams, W. J. Assessing the relative sensitivity of aquatic organisms to divalent metals and their representation in toxicity datasets compared to natural aquatic communities. *Human and Ecological Risk Assessment* 2005, 11 (6), 1139-1156.

Buchwalter, D. B.; Cain, D. J.; Clements, W. H.; Luoma, S. N. Using biodynamic models to reconcile differences between laboratory toxicity tests and field biomonitoring with aquatic insects. *Environmental Science & Technology* 2007, 41 (13), 4821-4828.

Buchwalter, D. B.; Clements, W. H.; Luoma, S. N., Modernizing water quality criteria in the United States: a need to expand the definition of acceptable data. *Environ. Toxicol. Chem.* 2017**,** 36 (2), 285-291.

Buchwalter, D. B.; Cain, D. J.; Martin, C. A.; Xie, L; Luoma, S. N.; Garland Jr., T. Aquatic insect ecophysiological traits reveal phylogenetically based differences in dissolved cadmium susceptibility. *Proceedings of the National Academy of Sciences of the United States of America* 2008, 105 (24), 8321-8326.

Burnham, K. P.; Anderson, D. R.; Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research* 2004, 33 (2), 261-304.

Cain, D. J.; Luoma, S. N. Metal exposures to native populations of the caddisfly *Hydropsyche* (Trichoptera : Hydropsychidae) determined from cytosolic and whole body metal concentrations. *Hydrobiologia* 1998, 386 (1-3), 103-117.

Cain, D. J.; Luoma, S. N.; Wallace, W. G. Linking metal bioaccumulation of aquatic insects to their distribution patterns in a mining-impacted river. *Environmental Toxicology and Chemistry* 2004, 23 (6), 1463-1473.

Cairns, J. The myth of the most sensitive species. *Bioscience* 1986, 36 (10), 670-672.

Clark, J. L.; Clements, W. H. The use of in situ and stream microcosm experiments to assess population- and community-level responses to metals. *Environmental Toxicology and Chemistry* 2006, 25 (9), 2306-2312.

Clements, W. H. Small-scale experiments support causal relationships between metal contamination and macroinvertebrate community responses. *Ecological Applications* 2004, 14 (3):954-967.

Clements, W. H., P. Cadmus, and S. F. Brinkman. Responses of aquatic insects to Cu and Zn in stream microcosms: understanding differences between single species tests and field responses. *Environmental Science & Technology* 2013, 47 (13), 7506-7513.

Clements, W. H., D. M. Carlisle, J. M. Lazorchak, and P. C. Johnson. Heavy metals structure benthic communities in Colorado mountain streams. *Ecological Applications* 2000, 10 (2):626-638.'

Deevey, E. S. Life tables for natural populations of animals. *The Quarterly Review of Biology* 1947, 22 (4), 283-314.

Diamond, J. M.; Winchester, E. L.; Mackler, D. G.; Gruber, D. Use of the mayfly *Stenonema modestum* (Heptageniidae) in subacute toxicity assessments. *Environmental Toxicology and Chemistry* 1992, 11 (3), 415-425.

Doledec, S.; Statzner, B. Invertebrate traits for the biomonitoring of large European rivers: an assessment of specific types of human impact. *Freshwater Biology* 2008, 53 (3), 617-634.

Fox, J.; Weisberg, S. 2011. An R companion to applied regression. SAGE Publications, Thousand Oaks, California, 2011.

Graves, S.; Piepho, H. P.; Selzer, L.; Dorai-Jai, S. 2012. Multcomp View: visualizations of paired comparisons. R Package Version 0.1-5.

Grosell, M.; Nielsen, C.; Bianchini, A. Sodium turnover rate determines sensitivity to acute copper and silver exposure in freshwater animals. *Comparative Biochemistry and Physiology: Part C* 2002. 133:287-303.

Harrell Jr., F. E.; Dupont, M. C. The Hmisc Package. R Package, version 2006: 2-0.

Herbst, D. B.; Medhurst, R. B.; Black, N. J. Long-term effects and recovery of streams from acid mine drainage and evaluation of toxic metal threshold ranges for macroinvertebrate community reassembly. *Environmental Toxicology and Chemistry* 2018, 37 (10): 2575-2592.

Hornberger, M. I.; Luoma, S. N.; Johnson, M. L.; Holyoak, M. Influence of remediation in a mine-impacted river: metal trends over large spatial and temporal scales. *Ecological Applications* 2009, 19 (6): 1522-1535.

Kiffney, P. M.; Clements, W. H. Size-dependent response of macroinvertebrates to metals in experimental streams. *Environmental Toxicology and Chemistry* 1996, 15 (8), 1352-1356.

Lenth, R. 2018. Estimated marginal means, aka least-squares means. URL: https://github.com/rvlenth/emmeans.

Maret, T.R.; Cain, D.J.; MacCoy, D.E.; Short, T.M. Response of benthic invertebrate assemblages to metal exposure and bioaccumulation associated with hard-rock mining in northwestern streams, USA. *Journal of the North American Benthological Society* 2003, 22 (4), 598-620.

McCafferty, W.P.; Randolph, R.P; Jacobus, L.M. Mayflies of the Intermountain West. *Memoirs of the American Entomological Institute* 2012, 85, 1-317.

McCahon, C. P.; Whiles, A. J.; Pascoe, D. The toxicity of cadmium to different larval instars of the trichopteran larvae *Agapetus fuscipes* (Curtis) and the importance of life-cycle information to the design of toxicity tests. *Hydrobiologia* 1989, 185 (2), 153-162.

Mebane, C. A. Testing Bioassessment Metrics: Macroinvertebrate, Sculpin, and Salmonid Responses to Stream Habitat, Sediment, and Metals. *Environmental Monitoring and Assessment* 2001, 67 (3), 293-322.

Mebane, C. A.; Dillon, F. S.; Hennessy, D. P. Acute toxicity of cadmium, lead, zinc, and their mixtures to stream-resident fish and invertebrates. *Environmental Toxicology and Chemistry* 2012, 31 (6), 1334-1348.

Mohammed, A. Why are early life stages of aquatic organisms more sensitive to toxicants than adults? New insights into toxicity and drug testing. *IntechOpen* 2013, 49-62.

Pearl, R.; Miner, J. R. Experimental studies on the duration of life XIV The comparative mortality of certain lower organisms. *Quarterly Review of Biology* 1935, 10 (1), 60-79.

Peckarsky, B. L.; Taylor, B. W.; Caudill, C. C. Hydrologic and behavioral constraints on oviposition of stream insects: implications for adult dispersal *Oecologia* 2000, 125 (2), 186-200.

Pomeranz J. P. F.; Warburton, H. J.; Harding, J. S. Anthropogenic mining alters macroinvertebrate size spectra in streams. *Freshwater Biology* 2018, 64 (1), 81-92.

Powlesland, C.; George, J. Acute and chronic toxicity of nickel to larvae of *Chironomus riparis* (Meigen). *Environmental Pollution Series A-Ecological and Biological* 1986, 42 (1), 47-64.

Schmidt, T. S.; Clements, W.H.; Mitchell, K. A.; Church, S. E.; Wanty, R. B.; Fey, D. L.; Verplanck, P. L.; San Juan, C. A. Development of a new toxic-unit model for the bioassessment of metals in streams. *Environmental Toxicology and Chemistry* 2010, 29 (11), 2432-2442.

Seuss, Dr. *Horton Hears a Who!* Random House, New York, NY, USA. 1954.

Soucek, David J., and Amy Dickinson. "Full-life chronic toxicity of sodium salts to the mayfly *Neocloeon triangulifer* in tests with laboratory cultured food. *Environmental Toxicology and Chemistry* 2015, 34 (9), 2126-2137.

Statzner, B., P. Bady, S. Doledec, and F. Scholl. Invertebrate traits for the biomonitoring of large European rivers: an initial assessment of trait patterns in least impacted river reaches. *Freshwater Biology* 2005, 50 (12), 2136-2161.

Statzner, B., and L. A. Beche. Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? *Freshwater Biology* 2010, 55 (1):80-119.

Statzner, B., B. Bis, S. Doledec, and P. Usseglio-Polatera. Perspectives for biomonitoring at large spatial scales: a unified measure for the functional composition on invertebrate communities in European running waters. *Basic and Applied Ecology* 2001, 2 (1):73-85.

Stephan, C. E.; Mount, D. I.; Hanson, D. J.; Gentile, J. R.; Chapman, G. A.; Brungs, W. A. *Guidelines for deriving numerical standards for the protection of aquatic organisms and their uses (PB85-227049)*. U.S. Environmental Protection Agency - USEPA Office of Water. Washington, D.C. 1985.

Stuhlbacher, A.; Bradley, M. C.; Naylor, C.; Calow, P. Variation in the development of cadmium resistance in *Daphnia magna* (Straus) effect of temperature, nutrition, age and genotype. *Environmental Pollution* 1993, 80 (2), 153-158.

U.S. Environmental Protection Agency - USEPA Office of Water. *Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms (4303T).* U.S. Environmental Protection Agency, 1200 Pennsylvania Avenue, NW. Washington, DC 2002.

Vannote, R L.; Sweeney, B. W. Geographic analysis of thermal equilibria: a conceptual model for evaluating the effect of natural and modified thermal regimes on aquatic insect communities. *The American Naturalist* 1980, 115 (5), 667-695.

Von der Ohe, P. C.; Liess, M. Relative sensitivity distribution of aquatic invertebrates to organic and metal compounds. *Environmental Toxicology and Chemistry* 2004, 23 (1), 150-156.

Ward J. V.; Kondratieff , B. C.; Zuellig R. E. *An Illustrated Guide to the Mountain Stream Insects of Colorado*. University Press of Colorado, Boulder, Colorado, USA. 2002

2 edition (April 15, 2002).

Willis Jr, L. D.; Hendricks, A. C. Life history, growth, survivorship, and production of Hydropsyche slossonae in Mill Creek, Virginia. *Journal of the North American Benthological Society* 1992, 11 (3), 290-303.

Xie, L.; Buchwalter, D. B. Cadmium exposure route affects antioxidant responses in the mayfly *Centroptilum triangulifer*. *Aquatic Toxicology* 2011, 105 (3-4), 199-200.

Xie, L.; Funk, D. H.; Buchwalter, D. B. Trophic transfer of Cd from natural periphyton to the grazing mayfly *Centroptilum triangulifer* in a life cycle test. *Environmental Pollution* 2010, 158 (1), 272-277.

CHAPTER 5

SYNTHESIS AND SUMMERY OF RESULTS

The field of aquatic toxicology functions across a wide spectrum from basic science to applied science. Research often addresses very scientifically basic questions not limited to the fields of ecology, physiology, physics and chemistry. Results from these sciences are applied to the creation of policy such as risk assessment, the establishment of safe pollution limits and the restoration of damaged ecosystems. Internalization of negative externalities, or regulation of pollution, associated with mining, chemical production, oil extraction and refinement is often contentious. This, in part, has led to standardization of methods to evaluate toxicity and derive toxicity thresholds. Such guidelines for toxicology experiments reduced the risk of Type I errors, the situation in which an experiment finds a toxic effect erroneously. Simultaneously, such experimental guidelines increased the risk of a Type II error in which an experiment finds no effect of toxicants when in nature toxic effects do exist. Starting in the 1960's and 1970's reductionist experiments exposing a single species to a single toxicant in laboratory became the primary source of data. Inexpensive and well suited to rigid experimental guidelines, this type of toxicity data became abundant, often drowning out the inference from the field of ecotoxicology. The terms "ecotoxicology" and "ecotoxicological" within this dissertation refers to the term originally coined in North America in the 1960s and 1970s referring to ecology and ecological research that is concerned with toxic effects. It does not use the term adopted in European policy that refers to traditional toxicology research that is concerned with any non-human species including cladocerans in a beaker. Throughout this dissertation, I present experiments that were designed to increase environmental realism. These studies focus on endpoints at higher levels of biological organization than traditional tests and are better suited to capture species interaction,

examine functional responses, include multiple stressors (as well as both direct and indirect toxicity) and consider conditional or phenological differences. These studies attempt to simulate the complexities of nature or use traditional experimental design that allows testing of more sensitive species and age classes. These improved studies not only advance ecological understanding and create new questions on the basic side of the applied-basic spectrum, but they also are used within to propose new water quality criteria, predict ecosystem response after a restoration, and critique the current understanding of aquatic insect sensitivity to aqueous metals.

The title 'Not all questions fit in beakers' was inspired from the wealth of data that suggested iron, iron floc or iron rich particles are non-toxic. This was based largely on many well replicated EPA/ASTM compliant experiments in which cladocerans, or sometimes larger pelagic species, were placed in small beakers with Fe floc or Fe rich particles. Beakers were not under flow-through conditions but instead static renewal. Iron oxides settled to the sole of the beakers. No effort was made to simulate a natural lotic environment in which Fe would remain homogenous in the water column. The duration of these experiments was too short to capture the loss of ecosystem function observed when Fe oxides embed in the interstitial spaces of the benthic zone and retard primary productivity. The majority of these experiments assessed only the response variable of mortality, missing more sensitive sub-lethal effects. Based on such tests consulting firms in Colorado were advocating that Fe standard was too stringent or that iron rich particles should be filtered from metal samples used to assess compliance (now coined "pre-filtered") because Fe is not toxic and is not bioavailable. Yet in one single mesocosm experiment Colorado State University's Ecotoxicology Laboratory found significant evidence that the Iron standard should be one-fourth the current level (Chapters 2 and 3). Under static renewal regimes, cladocerans and other pelagic species may survive well when Iron floc settles to the floor of

beakers. However, Colorado School of Mines observed extensive mortality of cladocerans in any treatment level above controls when caged in the mesocosm experiment reported in Chapter 1 of this dissertation. This system created a far more natural environment with more consistent homogeneity of Fe in the water column. The reason for mortality was believed to be that the iron floc was clinging to their swimming seta (M. Ramiro Pastorinho. Department of Biology at the Universidade de Évora. Personal Communication). The experience showcased that a large amount of data can be produced in a short amount of time when a scientist employs short single species experiments in small glass tanks or beakers in a laboratory. If the ecological relevance is poor or biased, those many inexpensive studies hold a disproportionate weight in the scientific literature and in derivation of standards. Philosophy of Science surrounding Abraham Maslow's Law of the Instrument (A.K.A. Law of the Hammer) warned against repeatedly using the same experimental tests to address all questions. Maslow (1966) wrote "I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail." Within this dissertation I attempt to showcase advantages of tailoring experimental designs to the question being asked and also give examples of how policies enshrined the use of inappropriate standardized methods.

Aqueous discharges from abandoned metal mines include complex mixtures of physical and chemical stressors. Consequently, identifying mechanisms and causal relationships between acid mine drainage (AMD) and community responses in the field is challenging. In addition to the direct toxicological effects associated with elevated concentrations of metals and reduced pH, mining activities influence aquatic organisms indirectly through physical alterations of habitat, including increased sedimentation, turbidity and substrate embeddedness. Although the focus of most restoration efforts in AMD-contaminated streams is on improving water quality, it is generally acknowledged that removing this single stressor may not be sufficient for restoring

structural and functional integrity of these systems. Therefore, an understanding of the relative contributions of chemical and physical stressors is crucial for designing effective restoration treatments in streams impacted by AMD. The goal of this research, in part, was to quantify the relative importance of physical (metal-oxide deposition) and chemical (elevated dissolved metal concentrations) stressors on benthic macroinvertebrate communities in the North Fork of Clear Creek (NFCC) near Blackhawk, Colorado, USA. North Fork of Clear Creek is a U.S. EPA Superfund site that is highly degraded by AMD below abandoned mine adits but rather pristine above these pollution sources. Field surveys of benthic macroinvertebrates conducted in NFCC showed highly significant alterations in abundance and community composition. Mesocosm experiments conducted with natural assemblages of benthic macroinvertebrates were used to establish concentration-response relationships between metals and community structure and to identify metal-sensitive taxa. The greatest effects of metals were observed on mayflies, whose abundances were reduced by >90% at metal concentrations similar to those measured in the field. Field experiments in NFCC were used to quantify avoidance of metal-contaminated substrate by macroinvertebrates. To predict the recovery potential of dominant taxa in this system, measures of metal tolerance and substrate tolerance were integrated with estimates of drift propensity obtained from the literature. Estimates of recovery potential were consistent with patterns observed at downstream recovery sites in the NFCC, which were dominated by caddisflies and baetid mayflies. Mesocosm and small-scale field experiments, particularly those conducted with natural communities, provide an ecologically realistic complement to laboratory toxicity tests. These experiments also control for the confounding variables associated with field-based approaches, thereby supporting causal relationships between stressors and responses.

Two mesocosm experiments were conducted to examine effects of ferric iron (Fe) and mixtures of ferric Fe with aqueous metals (Cu, Zn) on the structure and function of stream benthic communities. In 2010, naturally colonized benthic communities were exposed to a gradient of ferric Fe (0, 0.4, 1.0, 2.5, 6.2 and 15.6 mg/L) that bracketed the current U.S. EPA water quality criterion value (1.0 mg/L). After 10 d of exposure, total macroinvertebrate abundance, number of taxa and the abundance of all major aquatic insect groups (Ephemeroptera, Plecoptera, Trichoptera, and Diptera) were significantly reduced. Heptageniid mayflies (*Epeorus* sp.) and chironomids (Tanytarsini) were especially sensitive to the physical effects of Fe oxide deposition and were significantly reduced at 0.4 and 1.0 mg/L total Fe, respectively. In a second mesocosm experiment periphyton and macroinvertebrate communities were exposed to ferric Fe (0.60 mg/L) with or without aqueous copper (Cu) and zinc (Zn) at to treatment levels: low (0.01 mg/L Cu + 0.1 mg/L Zn) and high (0.05 mg/L Cu + 0.5 mg/L Zn). Contrary to previous research, no evidence of a protective effect of Fe was observed. Uptake of Cu and Zn by periphyton significantly increased in the presence of ferric Fe. Growth rates and protein content of periphyton were significantly reduced by both ferric Fe and aqueous metals, whereas abundance of a heptageniid mayfly (*Cinygmula* sp.) and whole community metabolism were significantly reduced by ferric Fe alone. Fe oxides inhibited algal growth and enhanced aqueous dissolved metal accumulation leading to a reduction in the quantity (lower biomass) and quality (lower protein content) of food resources for grazers. Mesocosm experiments conducted using natural benthic communities provide a unique opportunity to quantify the relative importance of indirect physical effects and to develop a better understanding of the relationship between basal food resources and consumers in natural stream ecosystems.

Iron is a common pollutant in waters near coal and hard rock mine disturbances. The current 1000 µg/L total recoverable chronic criterion for iron (Fe) for protection of aquatic life in the United States was developed using limited data in 1976 and has not been revised since. To develop a more scientifically-based criterion, several chronic laboratory toxicity experiments (>30 d) were conducted with ferric Fe at circumneutral pH on a taxonomically diverse group of organisms including brown trout (*Salmo trutta*), mountain whitefish (*Prosopium williamsoni*), boreal toad tadpoles (*Bufo boreas*), the oligochaete worm *Lumbriculus variegatus*, the mayfly *Hexagenia limbata* and the planarian *Dugesia dorotocephala*. Results of these tests and those of previously published toxicity data were used to derive a Final Chronic Value (FCV) of 499 µg/L using the US Environmental Protection Agency's (USEPA) recommended methods based on single species toxicity tests. In addition to single species toxicity tests, ferric Fe toxicity experiments (10 d) were performed on mesocosms containing naturally colonized communities of benthic macroinvertebrates. Fourteen genera in the mesocosms occurred at sufficient densities to estimate an iron concentration resulting in 20% reduction in abundance (EC_{20}). Three of these taxa had EC_{20}s less than the FCV of 499 µg/L derived from single species tests: the mayfly *Epeorus* sp. (335 µg/L), the caddisfly *Micrasema* sp. (356 µg/L), and the midge Tanytarsini (234 µg/L). When mesocosm results were included, the FCV was lowered to 251 µg/L. These findings support the suggestion that modernization of water quality criteria should include data generated from mesocosm experiments and other lines of evidence.

Laboratory assessments of trace metal toxicity generally demonstrate aquatic insects tolerate relatively high concentrations of metals in aqueous exposures; however, mesocosm experiments and field biomonitoring often indicate effects at relatively low metal concentrations. Several hypotheses have been proposed to reconcile these discrepancies, yet minimal research

has addressed how the size of aquatic insects influences their responses to metals. Field colonized benthic communities were exposed to trace metals in a series of mesocosm experiments. In addition, a novel single-species test system was used to expose first instar, mid-instar, and late instar mayfly (*Baetis tricaudatus*, Dodds) to Zn. These experimental approaches tested the hypothesis that small invertebrate size classes are more sensitive than large, mature size classes. Mesocosm results demonstrated strong size-dependent responses of aquatic insects to metals. Smaller organisms generally displayed greater mortality than large, mature individuals, and models were improved when size was included as a predictor of mortality. Size-dependent responses of *Baetis* spp. occurred in mesocosm experiments and in my single-species test system. The median lethal concentration (LC_{50}) for early instar *B. tricaudatus* was less than 6% of the previously reported LC_{50} for late instars. Together, these results suggest that aquatic insect body size is an important predictor of susceptibility to aqueous metals. Toxicity models that account for insect phenology by integrating the natural size progression of organisms have the potential to improve accuracy in predicting effects of metals in the field.

References

Abraham H. Maslow (1966). The Psychology of Science. New York: Harper & Row. p.15

APPENDIX A

SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 1.

Contains:

Figures 1.S1 to 1.S5

Tables 1.S1 to 1.S4

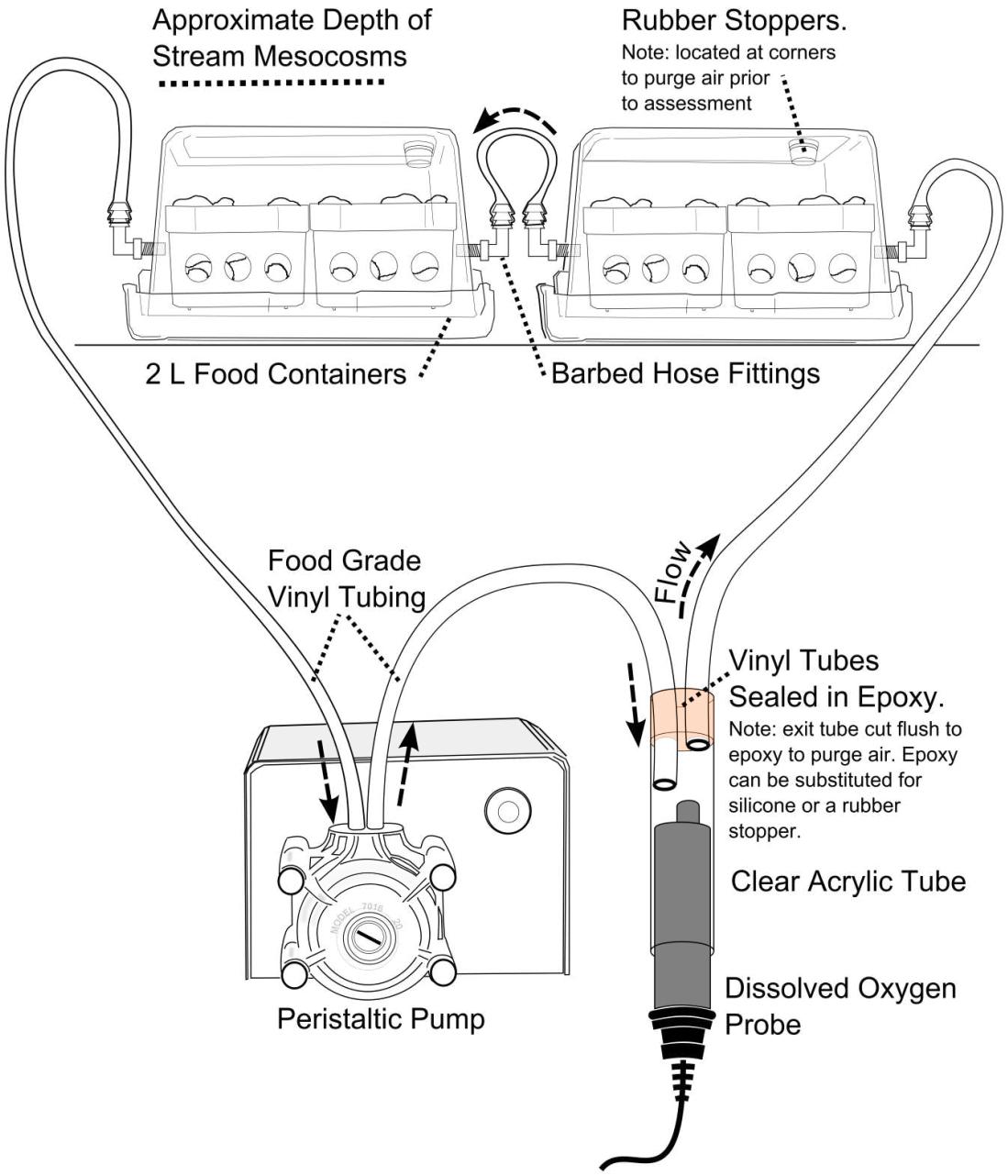


Figure 1.S1 - Schematic of community metabolism chambers. Trays were placed in sealed food containers which were purged of air bubbles. Peristaltic pumps circulated water within the containers through vinyl tubes over a dissolved oxygen probe (YSI Incorporated, Yellow Springs, OH, USA). The tubing and probe were insulated with foam pipe insulation. Dark conditions were created by blocking natural sunlight with opaque black plastic sheeting.

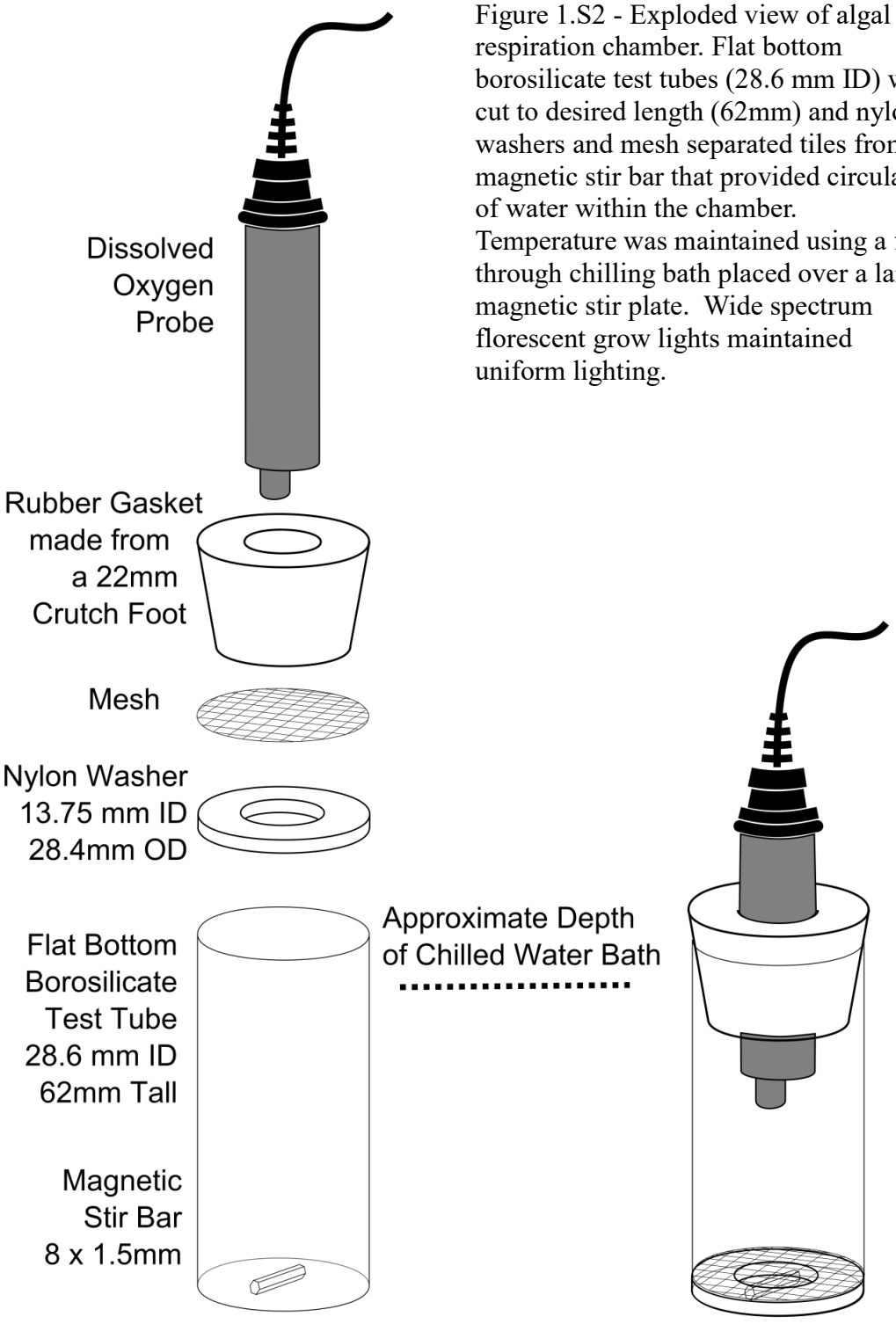


Figure 1.S2 - Exploded view of algal tile respiration chamber. Flat bottom borosilicate test tubes (28.6 mm ID) were cut to desired length (62mm) and nylon washers and mesh separated tiles from a magnetic stir bar that provided circulation of water within the chamber. Temperature was maintained using a flow through chilling bath placed over a large magnetic stir plate. Wide spectrum florescent grow lights maintained uniform lighting.

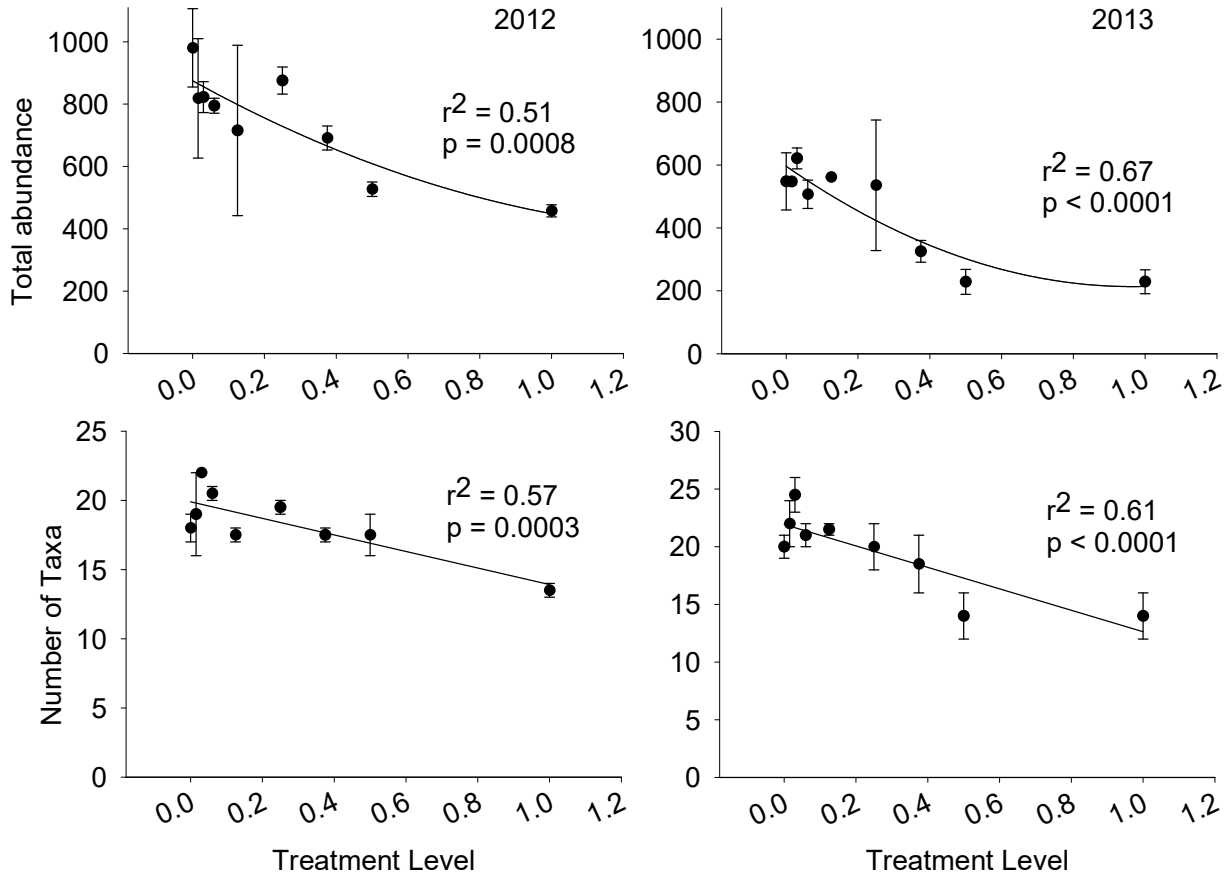


Figure 1.S3 - Mean (± s.e.) abundance (upper panels) and number of taxa (lower panels) as a function of treatment level in stream mesocosm experiments conducted in 2012 and 2013. Results of statistical analyses testing for significant concentration-response relationships are shown for each metric.

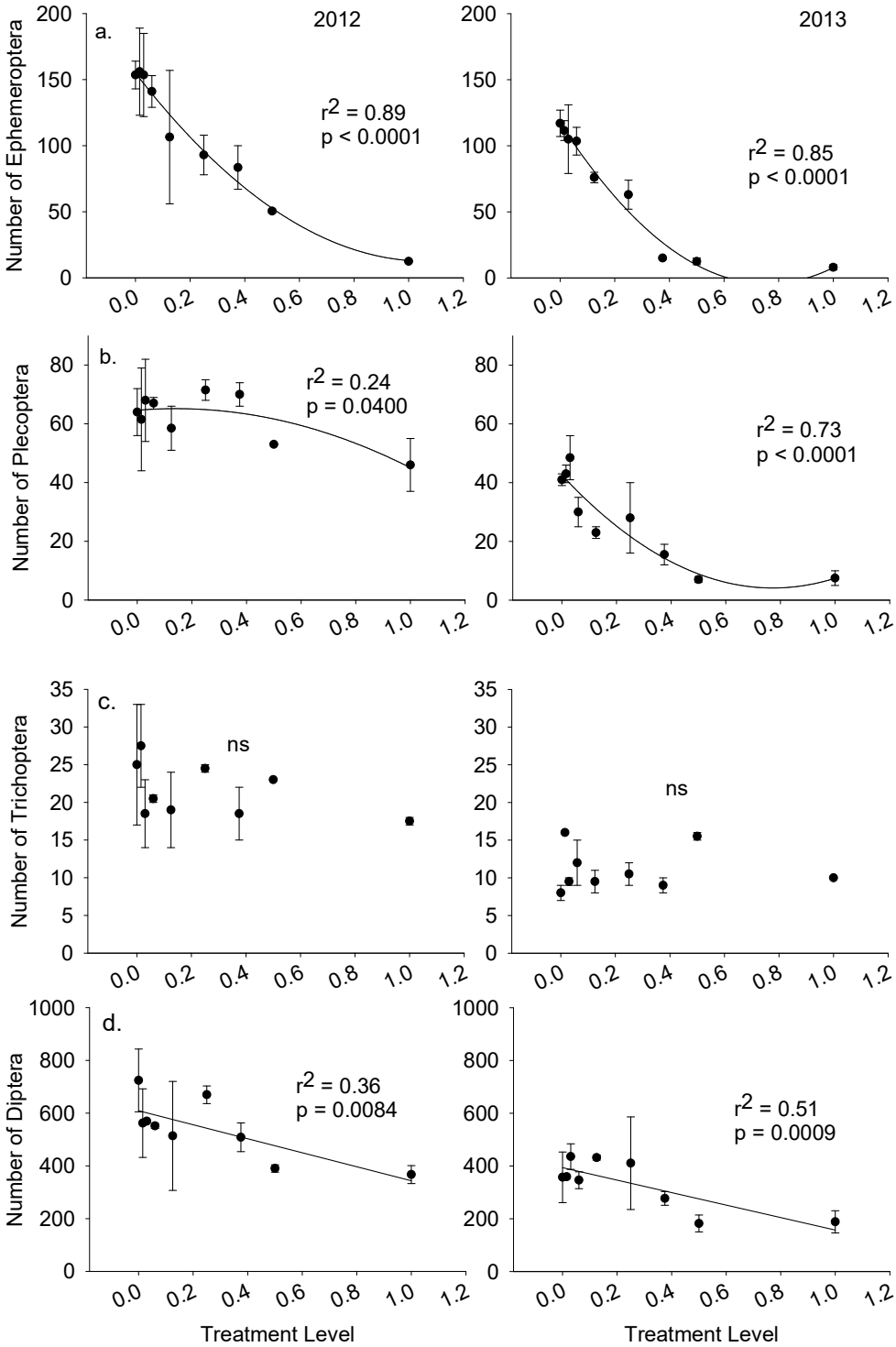


Figure 1.S4 - Mean (± s.e.) abundance of major macroinvertebrate groups as a function of treatment level in stream mesocosm experiments conducted in 2012 and 2013. Results of statistical analyses testing for significant concentration-response relationships are shown for each metric.

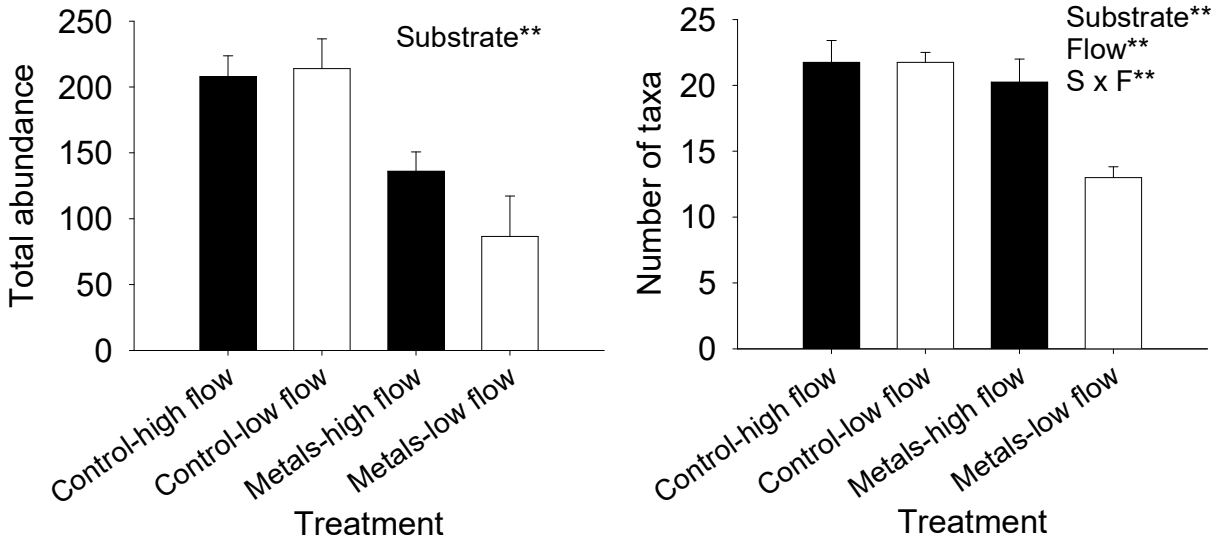


Figure 1.S5 - Mean (± s.e.) macroinvertebrate abundance and number of taxa in colonization trays containing clean or metal-contaminated substrate placed at an upstream reference site. Results of 2-way ANOVA testing for effects of substrate (control versus metal-contaminated), flow (low versus high) and the substrate x flow (S x F) interaction are also shown. *$p < 0.05$; **$p < 0.01$; ***$p < 0.0001$.

Table 1.S1 - Mean (± s.e.) dissolved metal concentrations (µg/L) measured at reference, impacted and recovery stations in the North Fork of Clear Creek, 2012-2013.

Reach	Cu		Zn		Fe		Mn	
Reference	2.4	(0.4)	37.7	(14.7)	89.2	(64.1)	16.0	(11.9)
Impacted	47.8	(15.6)	1133.4	(112.2)	6831.1	(1950.8)	2407.9	(439.6)
Recovery	10.3	(2.1)	440.4	(109.7)	24.9	(2.8)	714.5	(242.7)

Table 1.S2 - Routine water quality characteristics (mean ± s.e.) measured in stream mesocosms, 2012 and 2013. Mesocosm experiment I exposed communities to a mixture of metal salts. Mesocosm experiment II exposed communities to effluent collected from Gregory Incline, the primary source of metals to the NFCC. X = target metal treatment based on at the measured metal concentration at downstream contaminated stations.

Meso-cosm Exper-iment	Treatment	Temp. (C°)		D.O. (mg/l)		pH		Conduct-ivity (µS/cm)		Hard-ness mg/L $CaCO_3$		Alkal-inity mg/L $CaCO_3$	
I (2012)	0.0X	13.8	(0.2)	7.5	(0.4)	7.3	(0.1)	62.3	(0.7)	22.2	(4.9)	27.5	(1.5)
	0.015X	13.6	(0.2)	7.6	(0.1)	7.3	(0.1)	63.4	(0.4)	22.7	(5.0)	26.2	(1.3)
	0.03X	13.4	(0.2)	7.6	(0.1)	7.2	(0.1)	64.5	(0.7)	22.5	(4.9)	26.8	(1.6)
	0.06X	13.6	(0.2)	7.5	(0.4)	7.2	(0.1)	64.9	(0.4)	21.9	(4.8)	25.7	(1.0)
	0.125X	13.5	(0.2)	7.6	(0.0)	7.3	(0.1)	67.3	(0.8)	23.5	(5.2)	28.5	(1.9)
	0.25X	13.5	(0.2)	7.6	(0.1)	7.2	(0.0)	73.3	(0.9)	24.7	(5.5)	31.8	(1.7)
	0.375X	13.8	(0.2)	7.6	(0.1)	7.2	(0.1)	78.9	(1.9)	33.1	(0.9)	29.0	(1.2)
	0.50X	13.2	(0.3)	7.6	(0.1)	7.2	(0.1)	83.3	(2.2)	22.5	(5.0)	27.3	(1.4)
	1.0X	14.0	(0.2)	7.6	(0.1)	7.0	(0.1)	112.0	(4.1)	26.0	(5.9)	28.2	(2.6)
II (2013)	0.0X	11.1	(3.5)	8.1	(0.1)	7.6	(0.0)	64.7	(3.0)	30.3	(0.4)	22.6	(0.9)
	0.015X	11.0	(3.5)	8.1	(0.0)	7.6	(0.0)	71.3	(2.0)	32.5	(0.5)	21.6	(1.2)
	0.03X	11.1	(3.5)	8.0	(0.0)	7.5	(0.0)	78.1	(0.5)	35.1	(0.6)	23.2	(0.6)
	0.06X	11.0	(3.5)	7.9	(0.0)	7.4	(0.0)	87.3	(0.4)	37.3	(0.7)	19.9	(0.7)
	0.125X	10.9	(3.4)	7.9	(0.0)	7.4	(0.0)	100.4	(1.4)	44.2	(0.7)	22.4	(0.9)
	0.25X	11.2	(3.5)	7.8	(0.0)	7.3	(0.0)	114.6	(1.6)	50.8	(0.7)	20.9	(0.9)
	0.375X	11.3	(3.6)	7.7	(0.0)	7.1	(0.0)	167.6	(5.5)	76.9	(1.6)	18.8	(1.1)
	0.50X	11.5	(3.6)	7.7	(0.0)	7.0	(0.0)	204.9	(1.8)	92.8	(2.4)	17.8	(1.0)
	1.0X	11.9	(3.7)	7.5	(0.0)	6.8	(0.0)	309.0	(4.9)	154.0	(5.0)	16.4	(1.0)

Table 1.S3 - Mean (± s.e.) concentrations (µg/L) of dissolved (Cu, Zn) and total (Fe, Mn) metals measured in stream mesocosms experiments, 2012 and 2013. Mesocosm experiment I exposed communities to a mixture of metal salts. Mesocosm experiment II exposed communities to effluent collected from Gregory Incline. X = target metal treatment based on at the measured metal concentration at downstream contaminated stations.

Mesocosm Experiment	Treatment	Cu		Zn		Fe		Mn	
I (2012)	0.0X	4.0	(0.6)	12.2	(2.2)	145.8	(8.2)	60.0	(4.7)
	0.015X	6.2	(0.8)	22.1	(1.7)	330.1	(46.9)	96.1	(10.1)
	0.03X	7.2	(1.3)	30.1	(5.0)	437.7	(30.4)	110.4	(9.9)
	0.06X	9.1	(1.0)	48.4	(2.2)	881.3	(48.9)	192.5	(12.3)
	0.125X	9.1	(0.6)	73.0	(3.5)	1583.6	(140.1)	320.4	(27.0)
	0.25X	11.0	(1.2)	120.3	(9.7)	2747.4	(202.7)	544.2	(29.0)
	0.375X	12.9	(2.2)	166.1	(8.8)	4119.1	(343.0)	779.8	(53.6)
	0.5X	6.7	(0.9)	207.4	(16.0)	5385.2	(458.8)	1072.6	(53.6)
	1.0X	16.0	(5.7)	576.2	(64.8)	12527.9	(541.4)	2201.2	(107.0)
II (2013)	0.0X	2.0	(0.8)	20.5	(4.5)	144.6	(3.1)	12.6	(0.9)
	0.015X	2.5	(0.9)	29.1	(4.5)	301.7	(10.8)	59.0	(1.8)
	0.03X	1.6	(0.3)	31.0	(4.7)	478.9	(36.9)	104.0	(3.2)
	0.06X	2.1	(0.4)	45.9	(5.2)	792.4	(44.1)	170.1	(3.1)
	0.125X	2.3	(0.6)	67.8	(5.4)	1459.2	(97.6)	318.1	(15.8)
	0.25X	2.1	(0.3)	93.1	(5.3)	2364.6	(162.7)	476.5	(12.1)
	0.375X	4.1	(0.7)	192.3	(8.7)	4637.7	(425.2)	1010.8	(28.6)
	0.5X	4.9	(0.8)	245.2	(6.9)	6024.1	(284.7)	1295.0	(14.8)
	1.0X	5.1	(0.7)	472.9	(20.8)	16875.5	(1700.7)	2509.1	(38.5)

Table 1.S4 - Concentrations (µg/g) of Cu, Fe Mn and Zn (mean ± s.e.) on rock substrate before trays were deployed at the reference site (initial metal concentrations) and at the end of the field colonization experiment, May 2013. Substrate in control trays was washed prior to deployment in the reference stream to remove metal-oxide deposits. Metal-contaminated substrate was not washed. Flow treatments (low versus high) were hypothesized to influence the loss of metals during colonization in the reference stream.

Treatment	Cu		Fe		Mn		Zn	
Initial concentration	22.6	(2.5)	3094.2	(224.9)	28.0	(2.0)	29.1	(3.4)
Control								
Low flow	4.4	(1.2)	1299.6	(153.5)	45.6	(7.0)	10.9	(1.7)
High flow	4.1	(0.5)	1216.0	(30.5)	39.3	(7.1)	9.0	(0.6)
Metals								
Low flow	16.2	(2.3)	2570.3	(153.5)	49.6	(22.1)	25.0	(1.1)
High flow	6.0	(1.7)	1463.9	(282.0)	27.2	(4.5)	10.1	(2.3)

APPENDIX B

SUPPLEMENTAL FIGURES, TABLES AND NARRATIVE FOR CHAPTER 2

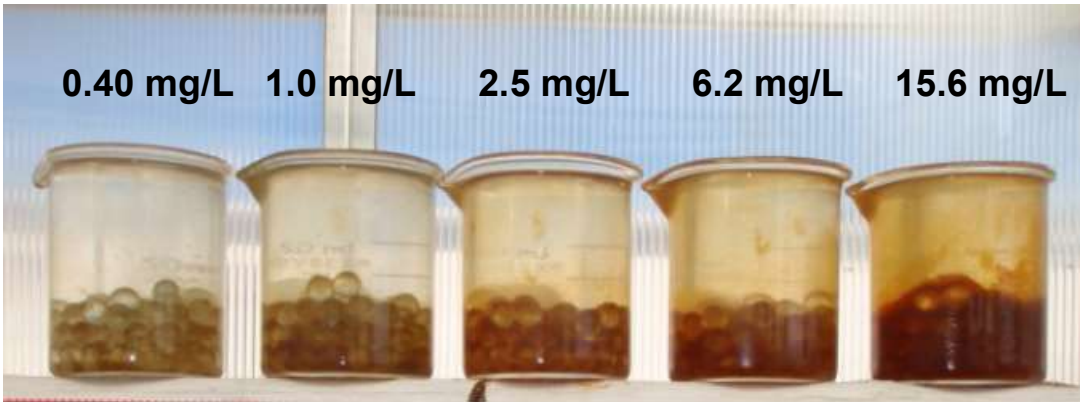
Contains:

Figures 2.S1 to 2.S5

Supplemental tables 2.S1 to 2.S2

Supplemental Methods Narrative for Chapter 2.

A.



B.

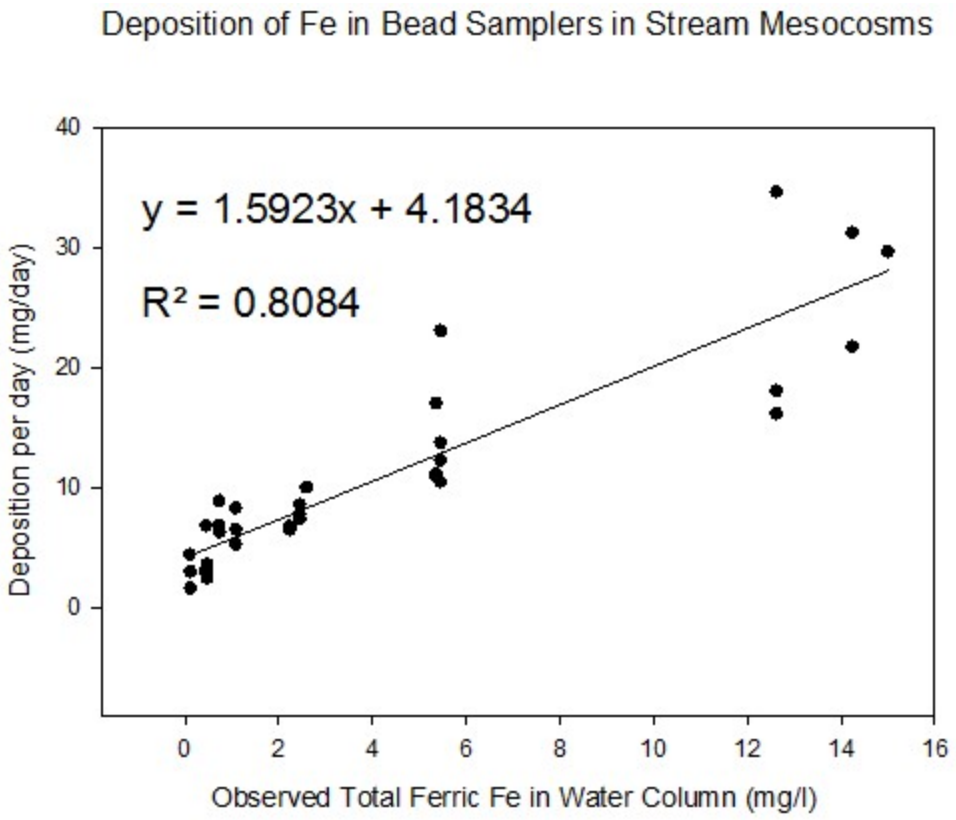


Figure 2.S1 - Deposition of Fe precipitates in Bead Samplers. A: Deposition in bead samplers after only 4 days. Interstitial space was lost in highest 3 treatment levels. B: Dry mass of Fe precipitates (mg/day) was well correlated with Fe observed in the water column (R^2=0.8084).

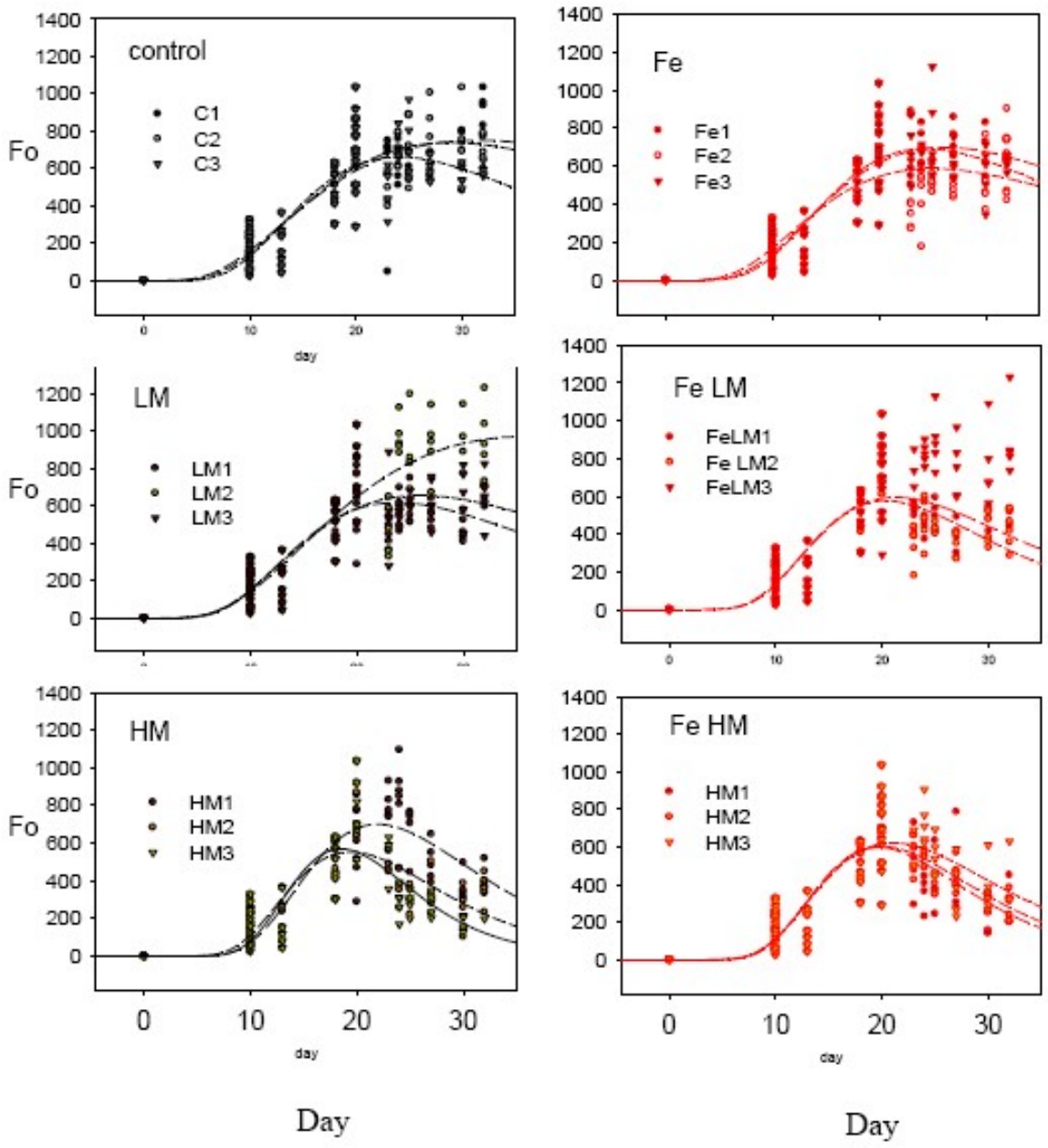


Figure 2.S2 - Biomass over time. Use of a three-parameters peak log normal curve to model algal biomass (Fo) throughout the experiments.

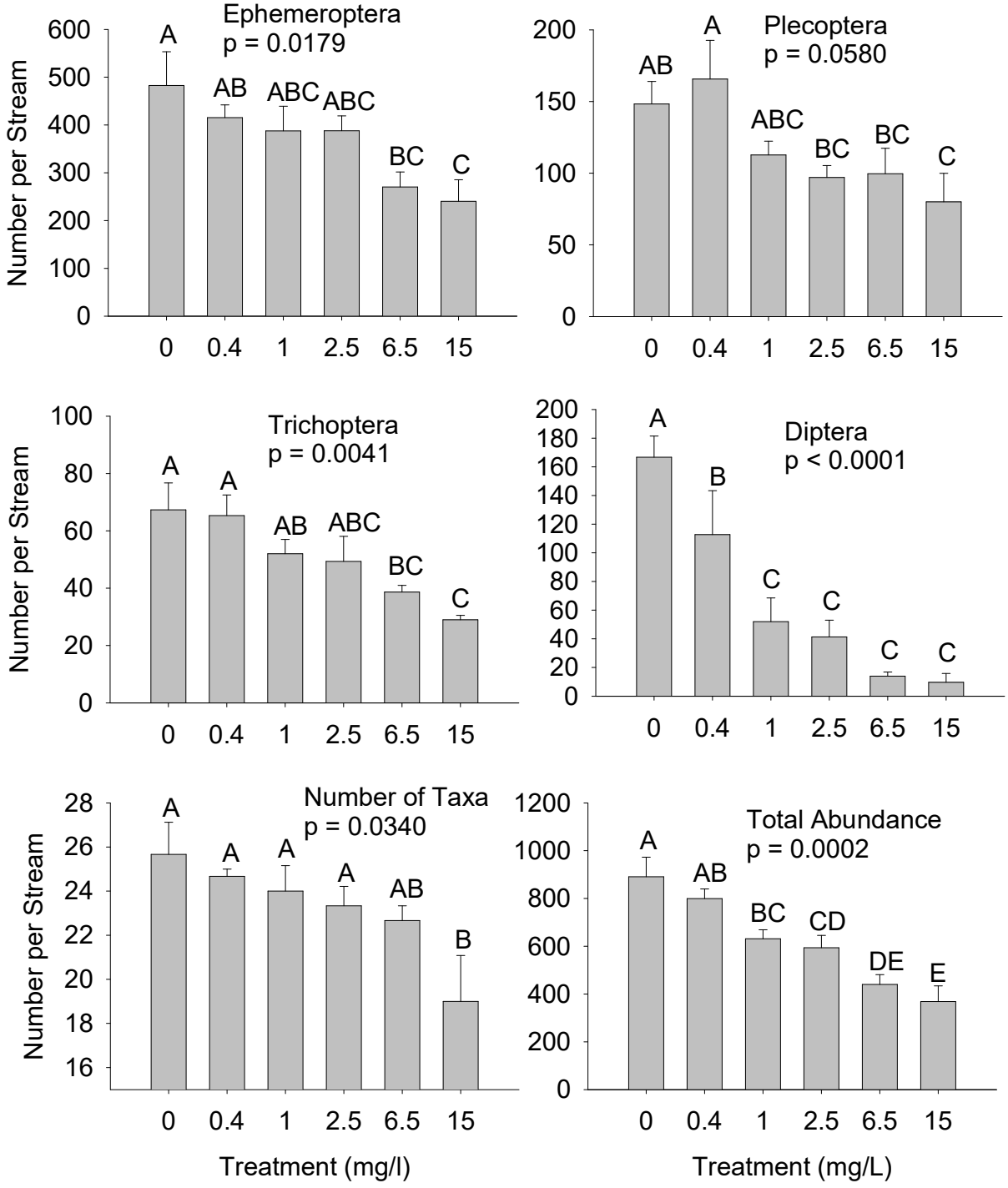


Figure 2.S3 - Abundance and diversity of macroinvertebrate taxa after 10 days of exposure to ferric Fe. Error bars denote standard error. Means with the same letter were not significantly different based on Duncan's Multiple Range test.

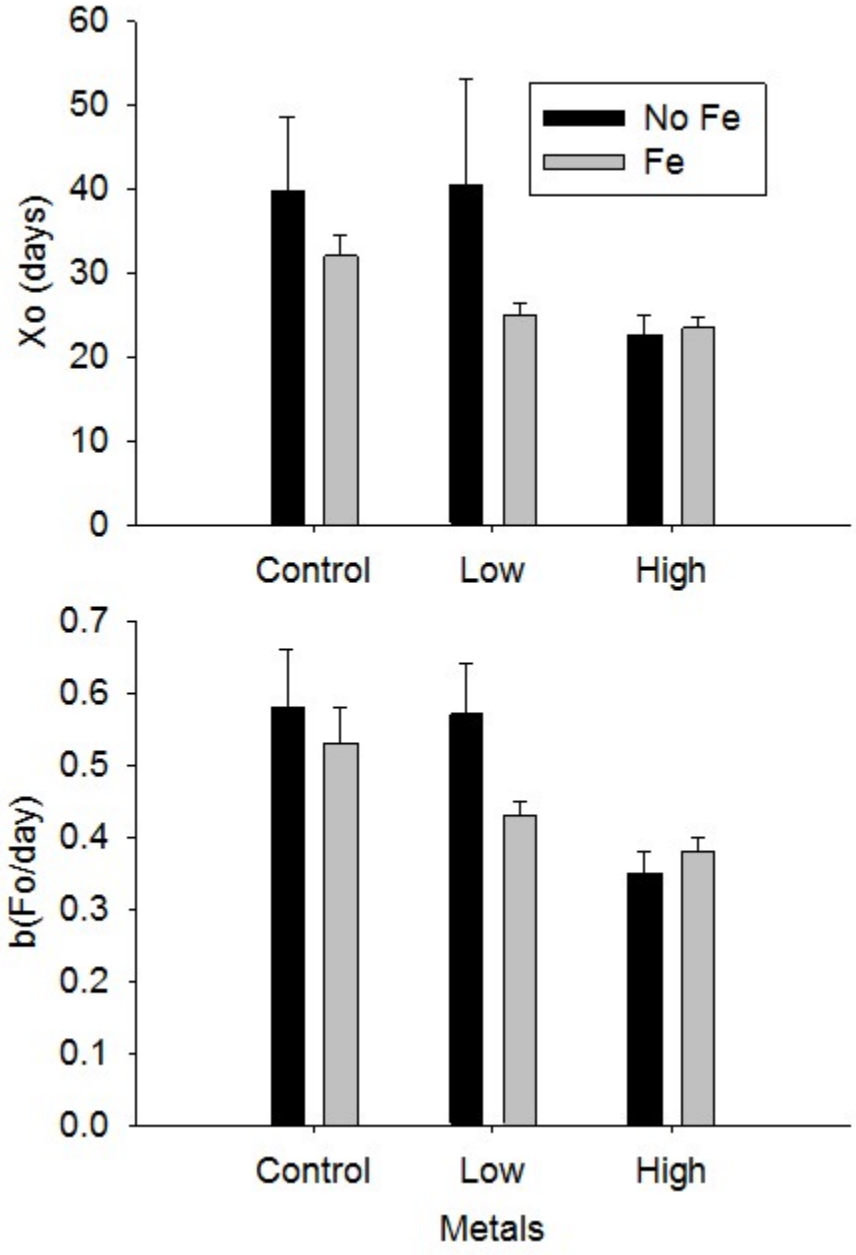


Figure 2.S4 - Influence of metals and iron on periphyton growth. Average and standard deviation of the parameters b and Xo. Differences between groups using multiple comparisons within ANOVA are denoted with different letters. The parameter representing the time required to reach the maximum biomass, Xo, was significant for both the metal ($p=0.019$) and Fe ($p=0.048$) factors. This implies that biomass would ultimately be reduced in treatments exposed to Fe or aqueous metals. Metal ($p<0.001$) exposure was found to reduce b, the parameter representing growth or Fo/day.

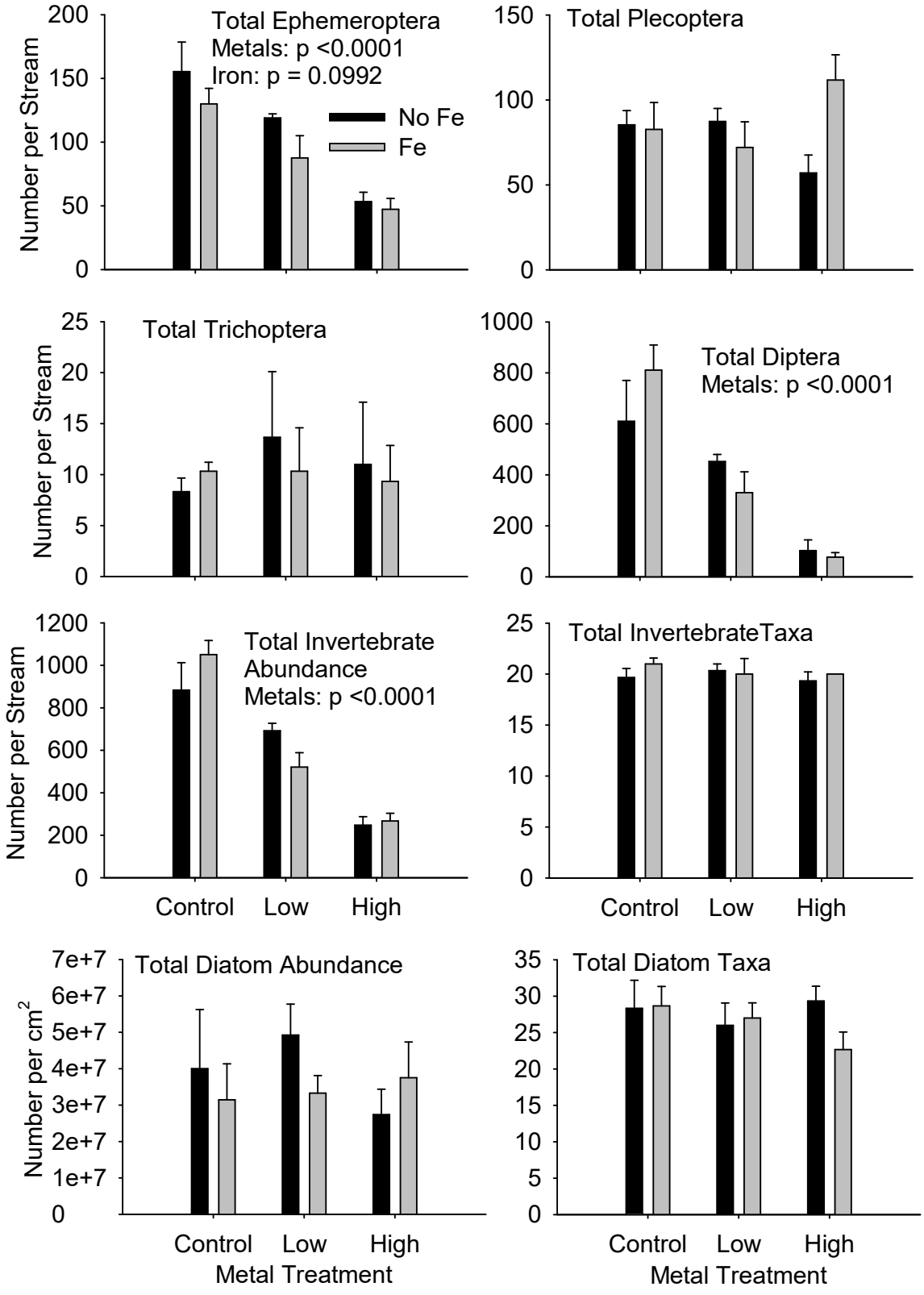


Figure 2.S5 - Influence of metals and iron on macroinvertebrate and diatom abundance and diversity from Experiment 2. Error bars denote standard error.

Table 2.S1 –Mean (s.e.) water quality characteristics measured in stream mesocosm experiments 1 and 2.

Experiment 1

Treatment (mg/L)	pH	Temp. (°C)	Spec. Cond (µS/cm)	Dis. O2 (mg/L)	Hardness (mg/L)	Alkalinity (mg/L)
0	7.2 (0.1)	11.7 (0.3)	90.4 (9.3)	7.7 (0.1)	35.0 (0.6)	27.3 (2.1)
0.4	7.2 (0.1)	11.8 (0.3)	95.4 (9.2)	7.5 (0.1)	35.7 (0.7)	29.0 (0.6)
1	7.3 (0.1)	11.6 (0.3)	99.9 (9.2)	7.6 (0.1)	32.7 (0.9)	28.0 (1.5)
2.5	7.2 (0.1)	11.8 (0.3)	110.8 (9.7)	7.4 (0.1)	32.7 (2.8)	32.7 (2.8)
6.25	7.3 (0.1)	11.9 (0.3)	128.1 (9.2)	7.2 (0.05)	33.8 (0.9)	25.7 (1.2)
15.6	7.2 (0.1)	11.8 (0.3)	204.8 (10.7)	7.2 (0.1)	34.3 (1.2)	24.2 (1.6)

Mesocosm Experiment 2

Fe Treatment	Metals	pH	Temp. (°C)	Spec. Cond (µS/cm)	Dis. O2 (mg/L)	Hardness (mg/L)	Alkalinity (mg/L)
None	Control	7.5 (0.05)	12.5 (0.1)	73.9 (0.3)	9.8 (0.1)	35.0 (1.0)	30.4 (0.6)
	Low	7.4 (0.02)	12.4 (0.1)	74.8 (0.2)	8.7 (0.1)	34.8 (0.6)	31.3 (0.5)
	High	7.2 (0.03)	12.3 (5.0)	76.6 (0.3)	8.4 (0.05)	35.0 (0)	30.0 (0.6)
Fe	Control	7.5 (0.05)	12.4 (0.1)	76.2 (0.8)	9.5 (0.1)	34.0 (0.4)	31.0 (0.9)
	Low	7.4 (0.1)	12.5 (0.1)	77.5 (0.3)	9.0 (0.1)	35.8 (0.5)	32.0 (0.7)
	High	7.3 (0.1)	12.5 (5.1)	79.4 (0.3)	8.4 (0.1)	35.8 (1.0)	30.0 (0.7)

Table 2.S2 - Average and confidence intervals of protein content (in µg protein/ml); AEA catalase (CAT) in µmol H_2O_2 mg protein-1 min-1 ; ascorbate peroxidase (APX) in µmol ascorbate mg protein-1 min-1, glutathione reductase (GR) in µmol NADPH min-1 µg-1 of protein obtained from biofilm samples after exposure and community metabolism. Results of the two-way analysis of variance are also included. The term n.s. denotes a *p*-value > 0.10. and n.a. denotes that the interaction term was not included in the model.

		Treatment						**TWO-WAY ANOVA**			
	Day	**Control**	**ML**	**MH**	**Fe**	**Fe ML**	**Fe MH**	Model	Fe	Metals	Fe*M
Proteins	1	393	276	32.6	262	278	207		n.s.	0.017	n.a.
		±50.2	±41.2	±31.0	±100	±57.3	±74				
	7-9	385	357	42.8	333	104	147		0.035	<0.001	<0.001
		±39.8	±112	±57.1	±111	±94.5	±52				
CAT	1	0.108	0.120	0.084	0.077	0.071	0.071		n.s.	n.s.	n.a.
		±0.042	±0.045		±0.012	±0.033	±0.027				
	7-9	0.079	0.092	0.051	0.059	0.077	0.076		n.s.	n.s.	n.a.
		±0.007	±0.022	±0.001	±0.012	±0.011	±0.010				
APX * 10^3	1	0.411	0.424	0.504	0.693	0.734	0.392		n.s.	n.s.	n.a.
		±0.030	±0.008		±0.228	±0.259	±0.031				
	7-9	0.403	0.455	0.431	0.434	0.514	0.336		n.s.	n.s.	n.a.
		±0.020	±0.023	±0.064	±0.040	±0.047	±0.061				
GR * 10^3	1	0.141	0.256	0.606	0.285	0.249	0.355		n.s.	0.004	0.045
		±0.038	±0.035		±0.085	±0.019	±0.034				
	7-9	0.165	0.269	0.284	0.251	0.254	0.308		n.s.	n.s.	n.a.
		±0.016	±0.033		±0.091	±0.089	±0.032				
Community Metabolism	10	1.2	1.4	0.89	0.67	0.7	0.73	0.0018	0.0004	n.s.	n.a.
		±0.10	±0.07	±0.14	±0.05	±0.17	±0.12				

Supplemental Methods Narrative for Extraction and Analysis of Protein and Antioxidants from Algae

Sampling, protein extraction and AEA measurements were performed per Bonnineau et al. (2011). The minimum required biomass for AEA analysis was obtained from 10 to 16 cm^2 of surface area which differed among experimental units. Biofilm was scraped from tiles to cryogenic tubes, flash frozen in liquid nitrogen and stored in -80°C conditions until analysis. Prior to protein extraction, samples were defrosted, centrifuged at 2,300 *g* for 5 min at 10°C to remove excess water and then weighed (wet weight). 200 µl of extraction buffer (100 mM Na_2HPO_4/KH_2PO_4, pH 7.4, 100 mM KCl, 1 mM EDTA) was added for each 100 mg of wet weight. Using a Heidolph DIAX900 homogenizer (Schwabach, Germany), samples were homogenized in two 30 s pulses. After adding 100 mg of glass beads (500 µm diameter) for each 100 mg of wet weight, cells were disrupted by performing three 30 s pulses of beadbeating using an MP FastPrep-24 ($v=4$ m s^{-1}) with 2 min intervals on ice. After cell disruption, homogenates were centrifuged at 10,000 *g* for 30 min at 4°C. The supernatant of each sample was retained as the enzyme source (Bonnineau et al. 2011).

The supernatant of each sample was split in triplicate and protein concentration was measured per Bradford (1976) using Coomassie Brilliant Blue G-250 dye reagent concentrate (Bio-Rad Laboratories GmbH, Munich, Germany) and bovine serum albumin as a standard. AEA measurements were performed in microtiter plates (UV-Star 96 well plate, Greiner®, Frickenhausen, Germany), changes in absorbance were followed using a Synergy4 microtiter plate reader (BioTek®, Winooski, VT, USA). For all assays the optimal protein concentration was 4 µg of protein. All AEA were calculated as specific activities (i.e. per µg of proteins). For each assay the optimal concentration of substrate or cofactor was determined by testing the

concentration in mM of H_2O_2 for the CAT and APX assay and nicotinamide adenine dinucleotide phosphate (NADPH) for the glutathione reductase (GR) assay. CAT activity was measured spectrophotometrically at 240 nm per Aebi (1984). 250 µl reaction mixtures were obtained adding by potassium phosphate buffer (pH 7.0; 80 mM final concentration), H_2O_2 (30 mM final concentration) and the enzyme extract (4 µg protein). The H_2O_2 consumption was determined by measuring the decrease in absorbance at 25°C for 3 min. CAT activity was calculated as µmol H_2O_2 mg protein^{-1} min^{-1} (extinction coefficient, ε: 0.039 M cm^{-1}). APX activity was assessed by monitoring the decrease in absorbance at 290 nm at 25°C and for 2 min, due to ascorbate oxidation, according to Nakano and Asada (1981). 250 µl reaction mixtures were obtained by adding potassium phosphate buffer (pH 7.0; 80 mM final concentration), H_2O_2 (4 mM final concentration), Na-Ascorbate (1.5 mM final concentration) and the enzyme extract (4 µg protein). APX activity was calculated as µmol Ascorbate mg protein^{-1} min^{-1} (extinction coefficient, ε: 2.8 M cm^{-1}). GR activity was assessed by monitoring the decrease in absorbance at 340 nm at 25°C and for 2 min (Schaedle and Bassham 1977). 200 µl reaction mixtures were obtained by adding Tris hydrochloride buffer (pH 7.5; 100 mM final concentration) and EDTA (1 mM), oxidized glutathione (GSSG; 1 mM final concentration), NADPH (0.25mM final concentration) and enzyme extract (4µg protein). GR activity was calculated as µmol NADPH min^{-1} µg^{-1} of protein.

References

Aebi H (1984) Catalase in vitro. Methods Enzymol 105:121-6.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254.

Bonnineau C, Bonet B, Corcoll N, Guasch H (2011) Catalase in fluvial biofilms: a comparison between different extraction methods and example of application in a metal-polluted river. Ecotoxicology 20:293-303.

Nakano Y, Asada K (1981) Hydrogen peroxide is scavenged by ascorbate-specific peroxidase in spinach chloroplasts. Plant and Cell Physiology 22:867-880.

Schaedle M, Bassham JA (1977) Chloroplast glutathione reductase. Plant Physiology 59:1011-1012.

APPENDIX C

SUPPLEMENTAL TABLES AND NARRATIVE FOR CHAPTER 3.

Contains:

Supplemental tables 3.S1 to 3.S9

Supplemental method narrative for Chapter 3.

Table 3.S1 - Water quality measurements and fry and egg survival for Brown Trout and Mountain Whitefish early-life stage Fe toxicity tests. Standard deviations are in parentheses. Asterisks denote treatment means significantly less than control ($p<0.05$). Fe reporting limits of <50 were reported for means of Fe between 0 and 49.

Target Fe concentration (µg/L)	0	625	1250	2500	5000
Total Fe concentration (unfiltered) (µg/L)	<50 (13)	658 (127)	1329 (169)	2438 (216)	5146 (835)
Dissolved Fe concentration (µg/L)	<50 (15)	<50 (34)	<50 (20)	<50 (18)	<50 (13)
Brown Trout hatch (%)	81.1 (7.9)	64.4 (6.8)	61.1 (10.3)	65.6 (7.9)	74.4 (6.8)
Brown Trout fry survival (%)	91.7 (6.8)	100 (0)	97.2 (3.9)	94.4 (7.7)	97.2 (3.9)
Brown Trout weight at termination (g)	0.295 (0.015)	0.277 (0.014)	0.271 (0.012)	0.285 (0.011)	0.248 (0.034)
Brown Trout biomass at termination (g)	3.241 (0.080)	3.327 (0.163)	3.161 (0.251)	3.229 (0.281)	2.894 (0.310)
Mountain Whitefish Hatch (%)	86.7 (9.4)	85.6 (1.6)	90.0 (4.7)	82.2 (4.2)	88.9 (1.6)
Mountain Whitefish fry survival (%)	84.4 (6.3)	97.8 (3.1)	86.7 (5.4)	82.2 (12.6)	66.7 (14.4)
Mountain Whitefish fry weight (g)	0.159 (0.010)	0.141 (0.009)	0.118* (0.005)	0.124 * (0.010)	0.099 * (0.010)
Mountain Whitefish biomass at termination (g)	2.005 (0.046)	2.067 (0.141)	1.532* (0.076)	1.544* (0.368)	0.980* (0.206)
Alkalinity (mg/L)	33.8 (1.1)	33.9 (0.8)	33.6 (1.2)	33.8 (1.1)	33.5 (1.3)
pH (SU)	7.47 (0.09)	7.50 (0.08)	7.47 (0.07)	7.49 (0.07)	7.49 (0.08)
Incubation temperature (°C)	7.6 (0.2)	7.5 (0.2)	7.6 (0.2)	7.6 (0.2)	7.5 (0.2)
Post-hatch temperature (°C)	12.3 (0.1)	12.4 (0.1)	12.3 (0.1)	12.3 (0.1)	12.2 (0.1)
Conductivity (µS/cm)	74.8 (1.2)	76.3 (1.4)	79.1 (1.1)	81.7 (2.4)	90.8 (1.7)
Dissolved oxygen (mg/L)	9.50 (0.41)	9.46 (0.50)	9.46 (0.46)	9.49 (0.47)	9.48 (0.47)

Table 3.S2 - Water quality measurements, survival, growth, development and mass of Boreal Toad tadpoles exposed to Fe for 35 days. Standard deviations are in parentheses. Asterisks denote treatment means significantly less than control ($p<0.05$). Fe reporting limits of <50 were reported for means of Fe between 0 and 49.

Target Fe concentration (µg/L)	0	500	1000	2000	4000	8000
Total Fe concentration (unfiltered) (µg/L)	<50 (13)	654 (89)	1073 (55)	2044 (89)	3831 (218)	8115 (701)
Dissolved Fe concentration (µg/L)	<50 (16)	<50 (41)	<50 (35)	<50 (21)	<50 (18)	<50 (13)
Survival (%)	100 (0)	100 (0)	95 (10)	100 (0)	60 (16)*	35 (10)*
Mean length (mm)	56 (1)	55 (2)	55 (1)	55 (1)	52 (1)	36 (2)*
Mean weight (g)	1.423 (0.052)	1.490 (0.159)	1.491 (0.100)	1.493 (0.075)	1.242 (0.100)	0.414 (0.061)*
Gosner Stage	39.4 (0.2)	39.5 (0.2)	39.7 (0.1)	39.6 (0.5)	39.0 (0.7)	35.8 (0.3)*
Biomass (g)	7.113 (0.262)	7.451 (0.794)	7.058 (0.591)	7.464 (0.377)	3.737 (0.907)*	0.728 (0.217)*
Alkalinity (mg/L)	34.9 (1.0)	33.2 (0.6)	33.9 (1.3)	34.5 (1.9)	35.4 (2.4)	33.4 (1.3)
pH (SU)	7.10 (0.18)	7.00 (0.07)	7.15 (0.12)	7.13 (0.15)	7.13 (0.14)	7.08 (0.14)
Conductivity (µS/cm)	109 (4)	110 (4)	113 (5)	121 (5)	131 (6)	148 (14)
Dissolved oxygen (mg/L)	7.19 (0.34)	7.49 (0.9)	7.60 (0.45)	7.36 (0.52)	7.45 (0.45)	7.65 (0.12)

Table 3.S3 - Water chemistry measurements, population growth (from 15 individuals) and biomass of *Lumbriculus variegatus* after 35 days exposure to ferric iron. Standard deviations are in parentheses. Asterisks denote treatment means significantly less than control ($p<0.05$). Fe reporting limits of <100 were reported for means of Fe between 0 and 99.

Target Fe concentration	0	1000	2000	4000	8000
Total Fe concentration (unfiltered) (µg/L)	<100 (13)	593 (84)	1145 (153)	3087 (308)	7592 (464)
Number of organisms	113.0 (13.6)	101.2 (29.1)	80.8* (29.4)	51.4* (14.4)	25.6* (6.2)
Biomass (g)	0.426 (0.062)	0.463 (0.143)	0.363 (0.126)	0.186* (0.062)	0.094* (0.029)
Hardness (mg/L)	44.6 (2.0)	44.5 (1.7)	44.1 (1.2)	44.4 (1.1)	44.2 (1.3)
Alkalinity (mg/L)	35.4 (2.2)	34.6 (1.4)	34.4 (1.3)	34.6 (1.8)	34.0 (2.6)
pH (SU)	7.86 (0.10)	7.91 (0.11)	7.88 (0.12)	7.88 (0.13)	7.85 (0.08)
Temperature (°C)	22.1 (0.1)	22.1 (0.1)	22.0 (0.3)	22.2 (0.1)	22.1 (0.1)
Conductivity (µS/cm)	114.0 (3.8)	116.4 (5.8)	121.2 (6.3)	134.2 (4.8)	165.7 (4.5)
Dissolved oxygen (mg/L)	7.55 (0.72)	7.51 (0.78)	7.57 (0.79)	7.54 (0.77)	7.62 (0.75)

Table 3.S4 - Mean water chemistry measurements, survival and mass of *Hexagenia limbata* after 30 days exposure to ferric Fe. Standard deviations are in parentheses. No significant differences from the control group were observed. Fe reporting limits of <100 were reported for means of Fe between 0 and 99.

Target Fe concentration	0	500	1000	2000	4000	8000
Total Fe concentration (unfiltered) (µg/L)	<100 (38)	464 (63)	903 (100)	1933 (171)	3829 (195)	7863 (244)
Survival (%)	86.1 (16.7)	86.1 (16.7)	83.3 (11.1)	83.3 (11.1)	91.7 (10.6)	83.3 (11.1)
Weight (mg)	256 (18)	219 (9)	211 (42)	254 (37)	255 (49)	224 (23)
Biomass (mg)	2050 (495)	1705 (384)	1600 (463)	1921 (457)	2124 (599)	1680 (350)
Alkalinity (mg/L)	33.1 (1.2)	33.2 (1.1)	33.4 (1.1)	33.0 (0.9)	32.9 (0.7)	33.4 (1.1)
pH (SU)	7.72 (0.04)	7.77 (0.06)	7.75 (0.07)	7.76 (0.08)	7.75 (0.04)	7.76 (0.10)
Temperature (°C)	17.0 (0.6)	17.0 (0.5)	17.1 (0.6)	17.0 (0.7)	17.0 (0.6)	17.0 (0.6)
Conductivity (µS/cm)	94.5 (9.5)	97.0 (3.5)	100.7 (4.1)	103.8 (4.3)	111.8 (7.7)	133.8 (3.8)
Dissolved oxygen (mg/L)	7.61 (0.17)	7.65 (0.17)	7.58 (0.14)	7.52 (0.18)	7.62 (0.20)	7.52 (0.17)

Table 3.S5 - Mean water chemistry measurements, fissions, population growth and mass of *Dugesia dorotocephala* after 30 days exposure to ferric Fe. Standard deviations are in parentheses. No significant differences from the control group were observed. Fe reporting limits of <100 were reported for means of Fe between 0 and 99.

Target Fe concentration	0	5000	1000	20000	30000	40000
Total Fe concentration (unfiltered) (µg/L)	<100 (38)	2502 (58)	5050 (183)	10214 (351)	20566 (666)	40134 (1106)
Population growth (%)	145.83 (25.00)	141.67 (21.52)	108.33 (16.67)	162.50 (34.36)	150.00 (13.61)	137.50 (15.96)
Fissions	2.75 (1.5)	2.50 (1.29)	0.50 (1.00)	3.75 (2.06)	3.00 (0.82)	2.25 (0.96)
Biomass (mg)	15.65 (0.72)	15.78 (2.29)	15.45 (1.51)	16.45 (1.88)	13.60 (1.88)	14.98 (1.18)
Alkalinity (mg/L)	75.13 (8.71)	76.80 (14.67)	75.33 (6.33)	74.07 (5.55)	74.07 (3.75)	72.60 (2.16)
pH (SU)	7.19 (0.14)	7.25 (0.17)	7.22 (0.07)	7.27 (0.05)	7.26 (0.15)	7.26 (0.14)
Temperature (°C)	22.63 (1.80)	22.67 (1.89)	22.80 (1.93)	22.80 (2.10)	22.77 (2.25)	22.97 (2.25)
Conductivity (µS/cm)	398.67 (44.60)	454.67 (59.34)	466.33 (37.58)	512.33 (40.53)	565.00 (48.12)	620.67 (47.25)
Dissolved oxygen (mg/L)	4.18 (3.05)	4.34 (2.82)	4.66 (2.63)	4.61 (2.61)	4.60 (2.64)	4.65 (2.57)

Table 3.S6 - Mean water chemistry measurements from mesocosm experiment exposing naturally colonized communities of benthic macroinvertebrates to ferric Fe for 10 days. * = below instrument detection limits.

Target Iron concentration	0	400	1000	2500	6250	15000
Total Fe concentration (unfiltered) (µg/L)	124 * (6)	446 (39)	944 (73)	2425 (83)	5238 (354)	14073 (450)
Hardness (mg/L)	35 (0.6)	35.7 (0.7)	32.7 (0.9)	32.7 (2.8)	33.75 (1.0)	34.33 (1.2)
Alkalinity (mg/L)	27.25 (2)	29 (0.5)	28 (1.5)	32.6 (2.9)	25.7 (1.2)	24.2 (1.6)
pH (SU)	7.24 (0.10)	7.22 (0.09)	7.27 (0.10)	7.23 (0.10)	7.25 (0.12)	7.23 (0.09)
Temperature (°C)	11.71 (0.26)	11.75 (0.28)	11.63 (0.26)	11.75 (0.26)	11.92 (0.33)	11.76 (0.28)
Conductivity (µS/cm)	90.4 (9)	95.4 (9)	100.0 (9)	101.0 (10)	128.1 (9)	204.8 (11)
Dissolved Oxygen (mg/L)	7.7 (0.1)	7.5 (0.1)	7.6 (0.1)	7.5 (0.1)	7.2 (0.1)	7.2 (0.1)

Table 3.S7 - No observed effect concentration (NOEC), lowest observed effect concentration (LOEC), maximum allowable toxicant concentration (MATC) and EC_{20} values. Underlined values used as chronic value for derivation of Final Chronic Value. Fe concentrations in µg/L total or total recoverable Fe.

Scientific name	Common name	NOEC	LOEC	MATC	EC_{20}	Reference
Dugesia dorotocephala	Planarian	>40134		>40134	<u>>40134</u>	This report
Orconectes limosus	Crayfish			22000[c]	[b]	Boutet and Chaisemartin 1973
Chironomus riparius	Midge	15000	30000	21213	<u>19818</u>	Radford 1997
Salvelinus fontinalis	Brook trout	7800	13420	10231	<u>9237</u>	Sykora et al. 1975
Hexagenia limbata	Mayfly	>7863		>7863	<u>>7863</u>	This report
Salmo trutta	Brown trout	>5146		>5146	<u>>5146</u>	This report
Oncorhynchus kisutch	Coho salmon	2830	4635	3621	<u>4870</u>	Smith and Sykora 1976
Oncorhynchus kisutch	Coho salmon	>3300		>3300	<u>>3300</u>	Brenner and Cooper 1978
Oncorhynchus kisutch	Coho salmon	2000	4250	<u>2915</u>	[b]	Updegraff and Sykora 1976
Oncorhynchus mykiss	Rainbow trout	1000	2200	<u>1483</u>	[b]	Goettl and Davies 1977
Oncorhynchus mykiss	Rainbow trout	>7500		>7500	<u>>7500</u>	Steffens et al. 1993
Bufo boreas	Boreal toad	2044	3831	2798	<u>3145</u>	This report
Daphnia magna	Cladoceran			<u>4380</u>[d]	[b]	Biesinger and Christensen 1972
Daphnia pulex	Cladoceran	700	1310	<u>958</u>	979[a]	Birge et al. 1985
Prosopium williamsoni	Mountain whitefish	658	1329	935	<u>1318</u>	This report
Lumbriculus variegates	Worm	593	1145	880	<u>870</u>	This report
Pimephales promelas	Fathead minnow	316	1008	569	<u>910</u>	Birge et al. 1985
Pimephales promelas	Fathead minnow		<2000	<2000	<u>520</u>	Smith et al. 1973

[a] Insufficient partial effects for reliable estimate of EC_{20}. Value reported for comparison purpose only.
[b] Insufficient data reported to run TRAP for EC_{20} value.
[c] 30d LC50 reported by authors used as MATC.
[d] 21d EC16 reported by authors used as MATC.

Table 3.S8 - Derivation of Final Chronic Value per Stephen et al. (EPA 1985) using single species experiments including spreadsheet equations in Microsoft Excel format.
N= Number of Genera = 12
R= Rank

RANK	GENUS	GMCV	Ln(GMCV)	Ln(GMCV)**2	P=R/N+1	P**0.5
4	*Daphnia*	2048	7.6246	58.1348	0.3077	0.5547
3	*Prosopium*	1318	7.1839	51.6080	0.2308	0.4804
2	*Lumbriculus*	870	6.7685	45.8125	0.1538	0.3922
1	*Pimephales*	688	6.5338	42.6904	0.0769	0.2774
	SUM		28.1108	198.2457	0.7692	1.7047
	SUM SQUARED		790.2155	39301.3616	0.5917	2.9059

S2 = SUM(LnGMCV)2 - (SUMLnGMCV)2/4/SUM(P)-SUM(P**0.5)**2/4	16.1802	4.022
L=(SUM(LnGMCV)-S*SUM(P**0.5))/4	5.3135	
A=S*SQRT(0.5) +L	6.2129	
FCV=EXP(A)	499	

References:

Stephan CE, Mount DI, Hansen DJ, Gentile JR, Chapman GA, Brungs WA (1985) Guidelines for deriving numerical standards for the protection of aquatic organisms and their uses. PB85-227049. USEPA, Springfield, VA

Table 3.S9 - Derivation of Final Chronic Value per Stephen et al. (EPA 1985) using single species experiments including spreadsheet equations in Microsoft Excel format.
N= Number of Genera = 26
R= Rank

RANK	GENUS	GMCV	Ln(GMCV)	Ln(GMCV)**2	P=R/N+1	P**0.5
4	*Pimephales*	688	6.5338	42.6904	0.1481	0.3849
3	*Microsema sp.*	356.29	5.8757	34.5244	0.1111	0.3333
2	*Epeorus sp.*	334.5	5.8126	33.7867	0.0741	0.2722
1	Tanytarsini (Tribe)	233.65	5.4538	29.7442	0.0370	0.1925
SUM			23.6760	140.7457	0.3704	1.1828
SUM SQUARED			560.5527	19809.3584	0.1372	1.3991
S2 = SUM(LnGMCV)2 - (SUMLnGMCV)2/4/SUM(P)-SUM(P**0.5)**2/4				29.5102		5.432
L=(SUM(LnGMCV)-S*SUM(P**0.5))/4				4.3126		
A=S*SQRT(0.5) +L				5.5273		
FCV=EXP(A)				251		

References:

Stephan CE, Mount DI, Hansen DJ, Gentile JR, Chapman GA, Brungs WA (1985) Guidelines for deriving numerical standards for the protection of aquatic organisms and their uses. PB85-227049. USEPA, Springfield, VA

Supplemental Methods Narrative for Literature Review and Creation of a Species Sensitivity Distribution from Chapter 3

Chronic Iron Toxicity Data for Derivation of Final Chronic Value using Single Species Trials

Chronic iron toxicity tests were identified using United States Environmental Protection Agency's Ecotox database (USEPA 2015a) and other electronic literature databases and then screened using the following criteria: 1. Species of test organism used must exist in freshwater systems in North America, 2. The duration of the test was sufficiently long to detect sublethal effects (≥25 days or ≥7days for Daphnids), 3. Ferrous iron was used as the toxicant. This third criterion was selected because ferrous iron and its precipitates are the overwhelming predominant form of iron in circumneutral oxygenated waters, 4. Toxicity tests were conducted at pH between 6.5 and 9.0 in order to minimize confounding effects of pH on toxicity results (see e.g. Radford 1997).

Toxicity tests meeting these criteria were relatively few but when added to tests reported here met the minimum of eight families needed to derive a chronic criterion for protection of aquatic life (Stephan et al. 1985). Test organisms included Salmonidae (*Oncorynchus kisutch, Oncorynchus mykiss, Prosopium williamsoni, Salvelinus fontinalis, Salmo trutta*), another fish family in class Osteichthyes (*Pimephales promelas*), a third family in Chordata (*Bufo boreas*), planktonic crustaceans (*Daphnia magna, Daphnia pulex*), benthic crustaceans (*Orconectes limosus*), Arthropoda (*Chironomus riparius, Hexagenia limbata*), a family in a phylum other than Arthropoda or Chordata (*Lumbriculus variegatus*), and finally a family in any order of insect or phylum not already represented (*Dugesia dorotocephala*; Appendix C, Table 3.S1). A study on the avoidance of *Oncorynchus kisutch* to ferric iron suspensions (Updegraff and Sykora

1976) did not meet the criterion for test duration but was included because avoidance was considered to be a relevant chronic endpoint.

Most studies reported maximum allowable toxicant concentrations (MATC), calculated as the geometric mean of a no observed effect concentration (NOEC) and lowest observed effect concentration (LOEC) (Table 3.6). Many studies reported data in sufficient detail to utilize USEPA's Toxicity Relationship Analysis Program version 1.30a (TRAP; USEPA 2015b) to perform regression analyses of results. Log transformed threshold sigmoid was routinely a good fit for data from single species trials. In such instances, iron concentrations estimated to cause a 20 percent reduction in response relative to control treatments (EC_{20}) were calculated to be consistent with current USEPA procedures for development of chronic criteria. EC_{20}s were generally in good agreement with MATCs and were always between NOECs and LOECs reported by the authors (Table 3.6).

Five iron toxicity tests did not detect an adverse effect at the highest iron concentration tested. In these instances, an EC_{20} or MATC could not be determined. If all of the iron exposure concentrations were very low, unbounded "greater than" toxicity thresholds would provide little information. However, if the highest exposure concentration was high, the "greater than" toxicity value indicates a species with high resistance to iron. Stephan et al. (1985) noted that unbounded "greater than" acute toxicity values should be used because excluding results from such resistant species would unnecessarily lower the Final Acute Value. This reasoning is equally applicable to calculating a Final Chronic Value. The five unbounded "greater than" toxicity values were deemed sufficiently high to indicate resistant species and were used as the chronic value for *Oncorhynchus mykiss* (>7500 µg/L; Steffens et al.1993), *Oncorhynchus kisutch* (>3300 µg/L;

Brenner and Cooper 1978), *Salmo trutta* (>5146 µg/L), *Hexagenia limbata* (>7863 µg/L) and *Dugesia dorotocephala* (>40,126 µg/L).

Chronic values used to calculate the Final Chronic Value (FCV) were based on EC_{20} values in instances where reliable estimates could be made. Otherwise, chronic values were MATCs or other toxicity threshold values reported by the authors (21d EC16 for *Daphnia magna;* Biesinger and Christensen 1972 and 30d LC_{50} for *Orconectes limosus*; Boutet and Chaisemartin 1973). Efforts were made to include data from as many experiments as possible given the limited number of relevant studies. After reviewing the toxicity reports, results of three studies were excluded from derivation of the Final Chronic Value (FCV). Dave (1984) reported a maximum allowable toxicant concentration (MATC) of 181 µg/L for *Daphnia magna* which was much lower than other chronic values and deemed inconsistent with other chronic values reported for Daphnids. Unpublished toxicity results for *Chematopsyche* and *Gammarus minus* (Sykora 1972) lacked a clear concentration-response and were also excluded.

Two other Daphnid studies were excluded because I could find no record of their presence in North America. A MATC for *Daphnia carinata* was 2419 µg/L (Van Dam et al. 1998). Randall et al. (1999) report a "safe limit" of particulate iron of 1690 µg/L for *Daphnia longispina*. While the MATCs for these Daphnids were excluded, it is worth noting that they fell within the range of 958 µg/L for *Daphnia pulex* and 4380 µg/L for *Daphnia magna*.

References

Boutet C, Chasemartin C (1973) Proprietes toxiques specifiques des sels metalliques chez Austropotamobius pallipes pallipes et Orconectes limosus. Comptes Rendus des Seances de la Societe de Biologie et de ses Filiales 167:1933-1938

Dave G (1984) Effects of waterborne iron on growth, reproduction, survival and haemoglobin in Daphnia magna. Comp Biochem Phys 78C:433-438

Radford NP (1997) Ecotoxicological impact of iron III sulphate dosing on chironomid cultures and profundal reservoir communities. Doctoral Thesis, University of Leicester

Randall S, Harper D, Brierley B (1999) Ecological and ecophysiological impacts of ferric dosing in reservoirs. Hydrobiologia 395/396: 355-364

Stephan CE, Mount DI, Hansen DJ, Gentile JR, Chapman GA, Brungs WA (1985) Guidelines for deriving numerical standards for the protection of aquatic organisms and their uses. PB85-227049. USEPA, Springfield, VA

Sykora JL, Smith EJ, Synak M, Shapiro MA (1975) Some observations on spawning brook trout (Salvalinus fontinalis Mitchell) in lime neutralized iron hydroxide suspensions. Water Res 9:451-458

Updegraff KF, Sykora JL (1976) Avoidance of lime-neutralized iron hydroxide solutions by coho salmon in the laboratory. Environ Sci Technol 10:51-54

USEPA (2015a) ECOTOX User Guide: ECOTOXicology Knowledgebase System. Version 4.0. https://cfpub.epa.gov/ecotox/blackbox/help/userhelp4.pdf. Accessed 28 April 2017

USEPA (2015b) Toxicity Relationship Analysis Program (TRAP Version 1.30a). USEPA. Duluth, MN

Van Dam RA, Barry MJ, Ahokas JT, Holdway DA (1998). Effects of water-borne iron and calcium on the toxicity of diethylenetriamine pentaacetic (DTPA) to Daphnia carinata. Aquat. Toxicol. 42:49-66

APPENDIX D

SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 4

Contains:

Figures 4.S1 to 4.S5

Tables 4.S to 4.S6

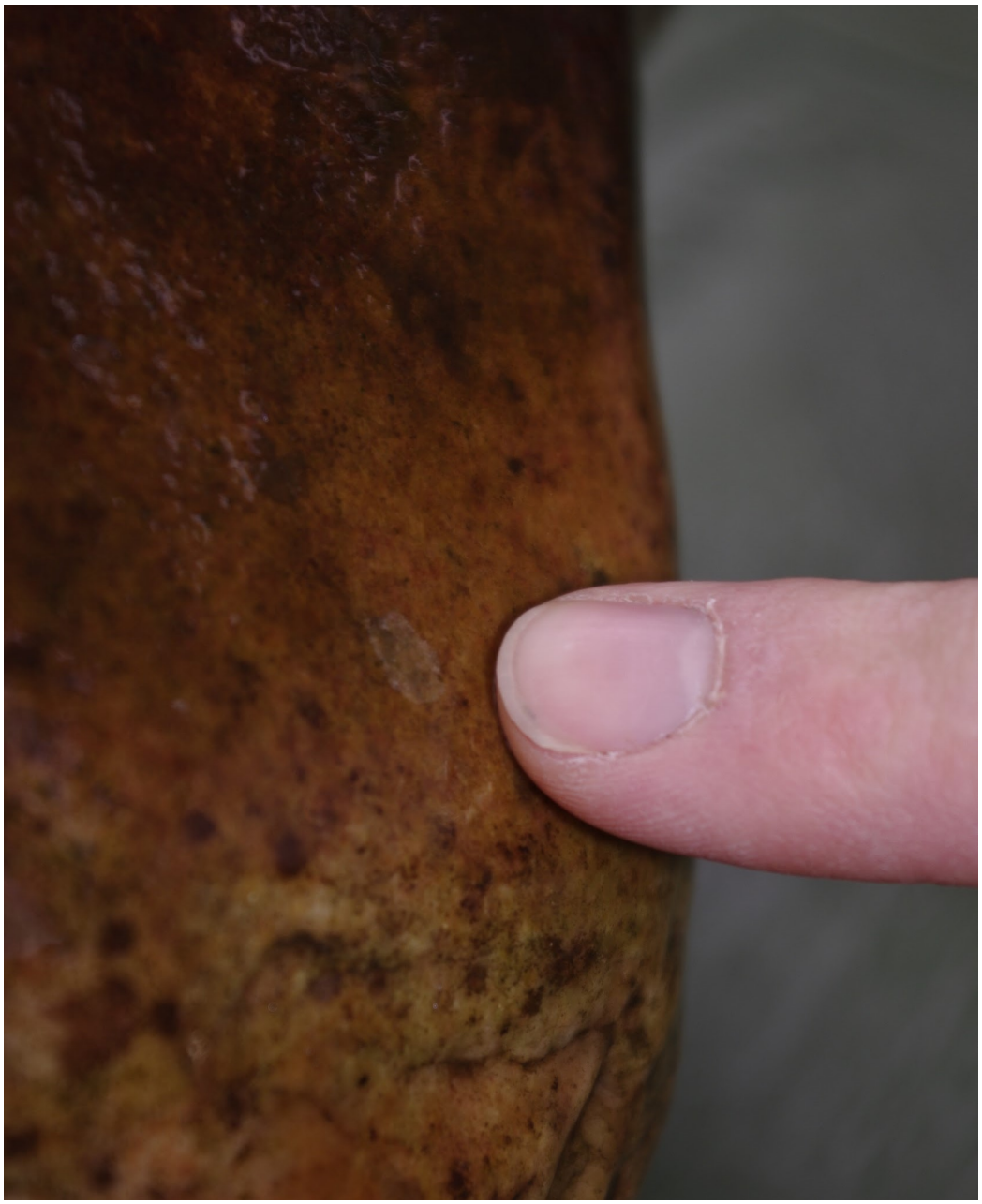


Figure 4.S1 - *Baetis tricaudatus* egg mass. Members of Baetidae (Ephemeroptera) often oviposit eggs on lenticular cobble 6 to 30 cm in diameter. Almost translucent, the bumpy white patch of eggs covers less than 0.5 cm^2. Egg masses were carefully removed using stainless steel razor blades and placed in Cache la Poudre River water in 35-mm polystyrene Petri dishes. Petri dishes were aerated using pasture pipettes and maintained at 12-15 °C. Eggs were assessed for hatching twice daily. Immediately upon hatch (after ~2 weeks), organisms were fed a 10 mL suspension of the diatoms *Navicula* sp. and *Synedra* sp. (Carolina Biological Supply, Burlington, NC, USA) at 0 and 2 d of exposure.

Figure 4.S2 - Exposure tubes. Aquarium-grade silicone was used to affix stainless steel mesh (105 x 125 micron pore size) to the beveled (30°) end of 24 borosilicate glass tubes (8 mm I.D., 12 mm O.D., 27 cm long). These tubes retained mid- and first-instar age classes during exposure and enumeration. Small strips of fine mesh (3 x 30 mm) were placed inside each tube as substrate for the organisms.

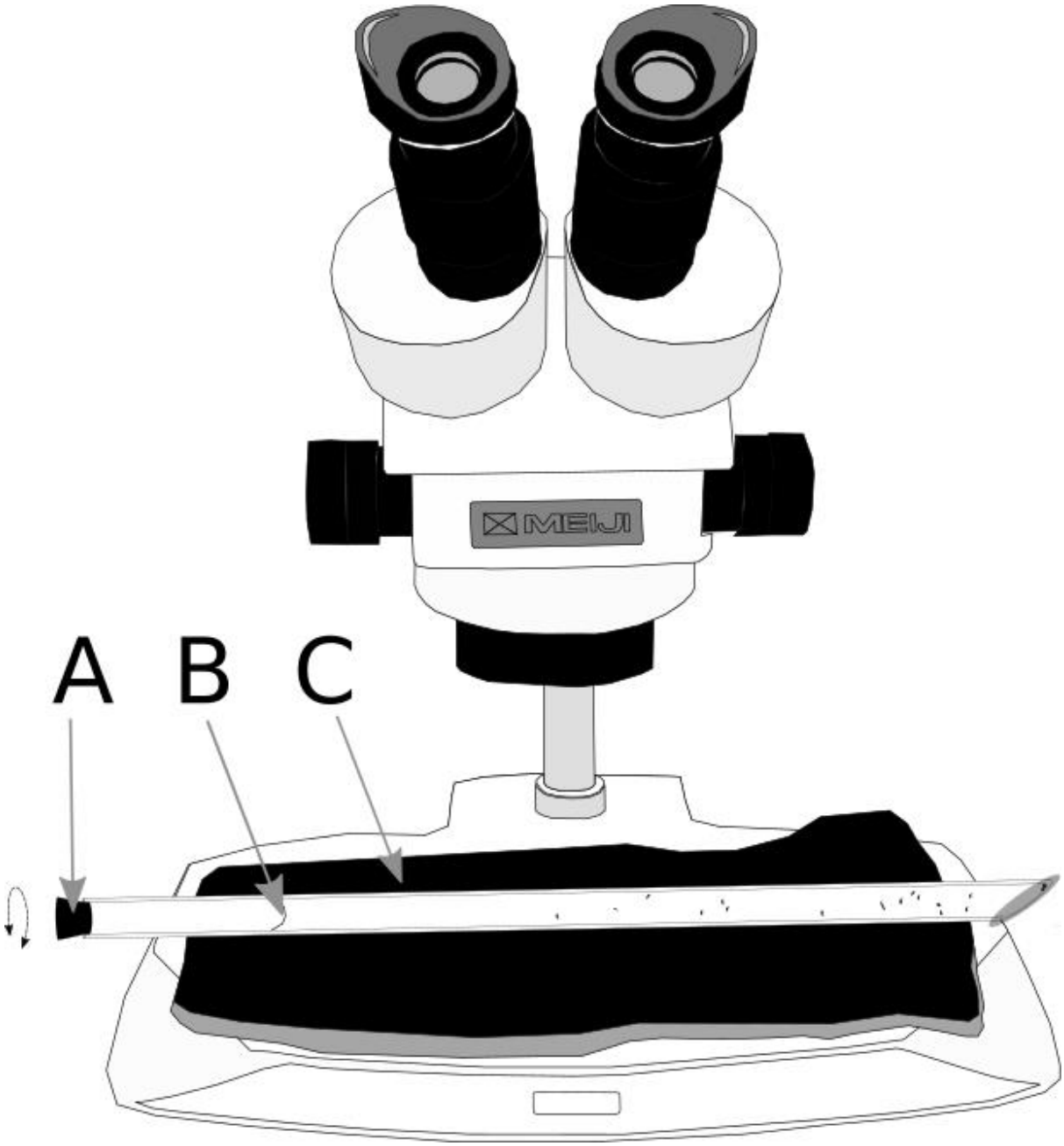


Figure 4.S3 - Exposure tubes were placed upright in a 500-ml beaker filled with chilled, highly-oxygenated water. Using a dissecting microscope (Meji EMZ-TR with 20x eyepieces), 10-20 organisms were transferred to each of the 24 exposure tubes using a 100-µl pipette. Rubber stoppers (A) were placed in exposure tube during enumeration and transport to and from the exposure system. This held the water level in the exposure tube at approximately B. Enumeration of surviving organisms was assessed at 0 (before exposure), 48 and 96 h. During enumeration, the exposure tube was rotated in the focal field of the microscope. Chilled slate tiles (C) were used to maintain temperature in exposure tubes during microscopy.

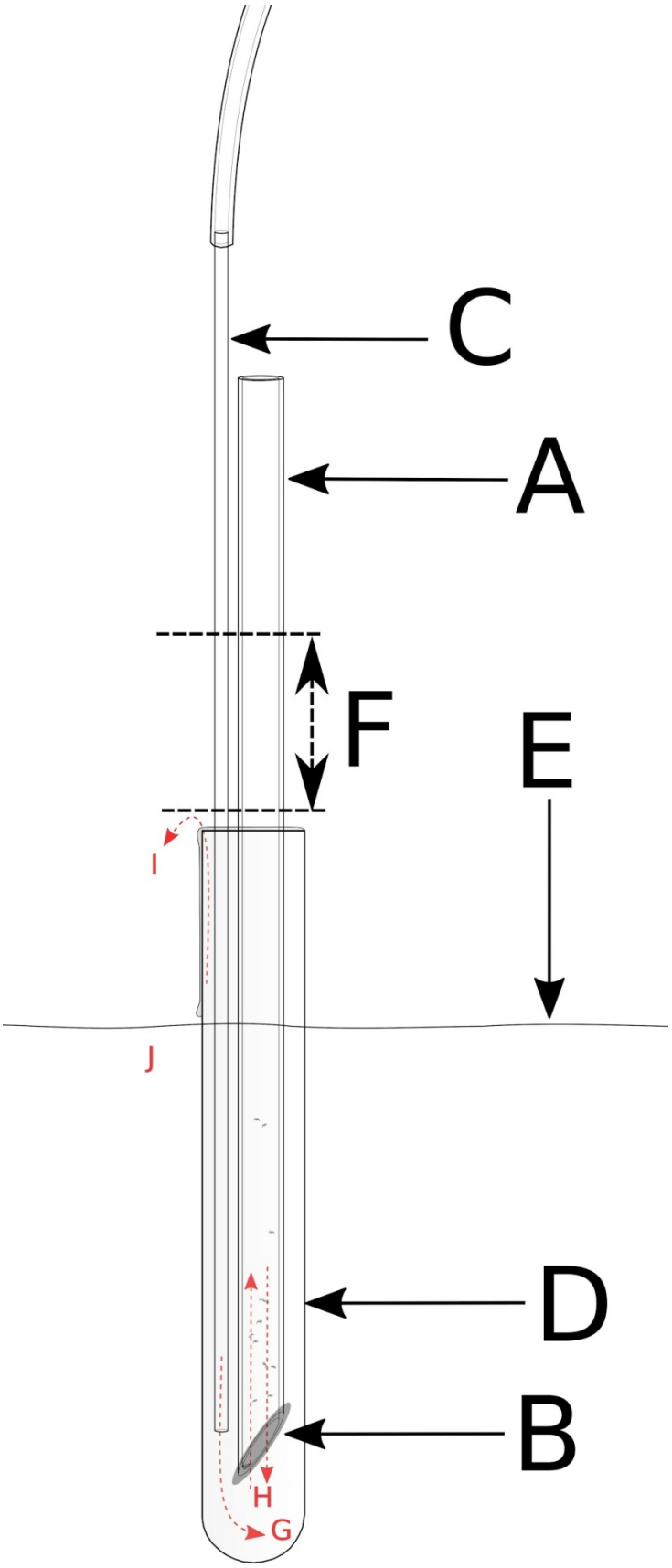


Figure 4.S4 - Organisms were contained in 24 exposure tubes (A). Fine stainless steel mesh (B) retained organisms during the experiment. Vinyl food-grade tubing and fine glass tubes (C; 3 mm ID, 5 mm OD) were used to transport toxicants and aerated water from a serial diluter to 24 test tubes (D; 23 mm ID x 15 cm tall). Test tubes were held in test tube racks placed in a chilled water bath with a standpipe holding the water level at E. Exposure tubes and toxicant delivery tubes were affixed by rubber band to a motorized teeter-totter (Figure 4.S5) that dipped and raised the tubes approximately 3cm (F) every 30 s. Highly-oxygenated water and toxicants from the serial diluter entered the system at G. The change in hydraulic head forced water in and out (H) of exposure tubes. Water and toxicants overflowed (I) into the chilled water bath (J). Flow-through conditions and the flood and ebb created by the mechanical teeter-totter simulated the hydraulic microhabitat of *Baetis* sp. and prevented debris from clogging the fine-mesh screens. The beveled end of exposure tubes allowed debris to collect at the tip without occluding the entire screen.

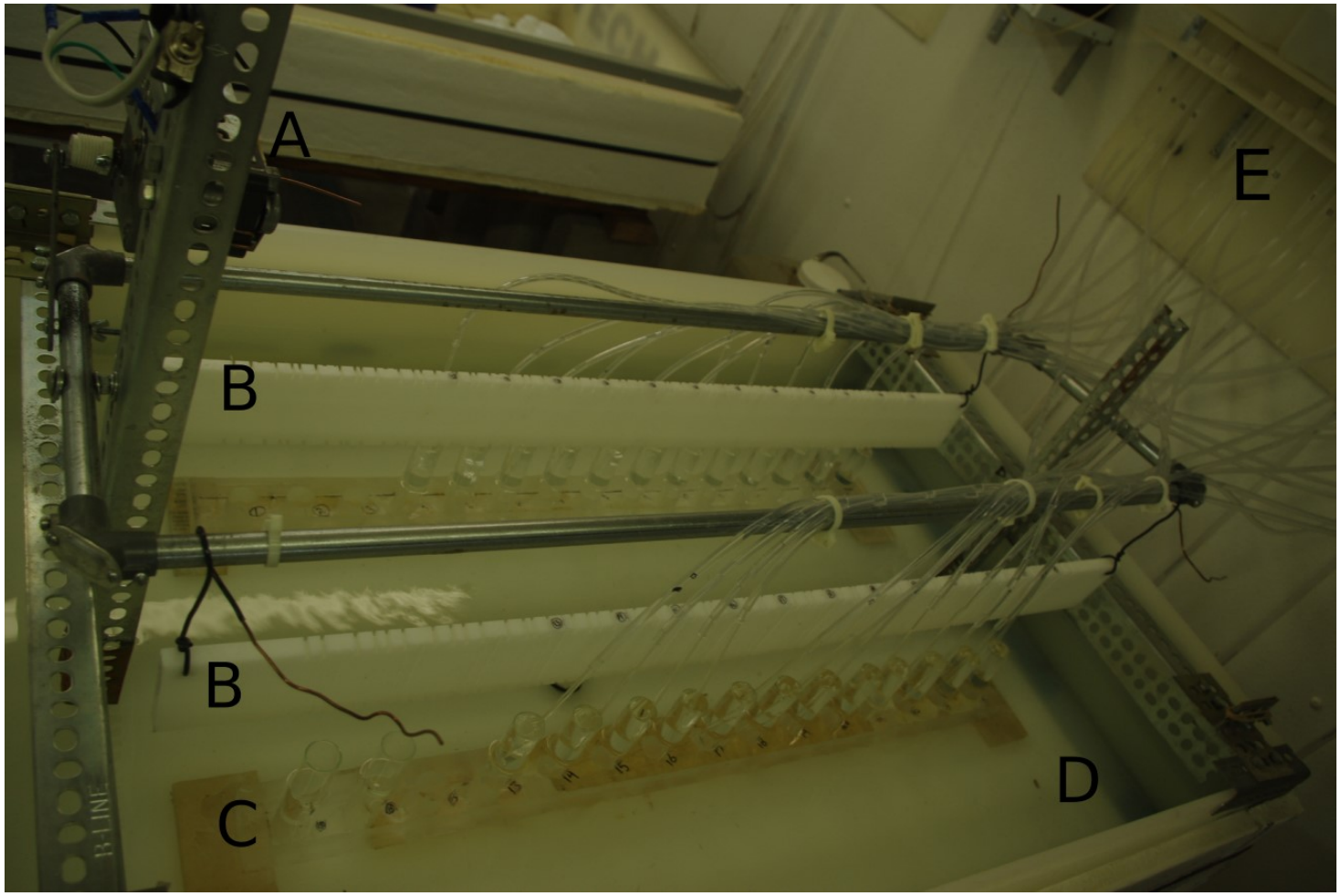


Figure 4.S5 - Exposure system. A motorized “teeter-totter” powered by a gear motor (A) was constructed of perforated angle iron and electrical conduit. White, high-density polyethylene boards (B) were suspended by wire on each arm of the teeter-totter. Dado cuts (grooves) accommodated rubber bands (not shown), which held exposure tubes and toxicant delivery tubes perfectly aligned with 24 test tubes. Test tubes were held in heavy, acrylic test-tube racks (C) in a water bath (D). Continuous-flow serial diluters (E; Benoit et al. 1982) delivered 40 ml/min of dechlorinated municipal tap water (Fort Collins, CO, USA) to test tubes. Both the water bath and flow from the continuous-flow serial diluters were chilled at 12 °C.

Table 4.S1 - Hardness-adjusted criteria for metals used in mesocosm studies. The hardness-adjusted equations were selected to be consistent with previous studies (Clements et al. 2013, Clements 2004) at the Stream Research Laboratory.

	2007 (Cu, Cu + Zn) 2010 (Cu + Zn + Cd)		**2015 (Cu + Zn)**		
	Hardness (mg/L $CaCO_3$)	Criterion (µg/L)	Hardness (mg/L $CaCO_3$)	Criterion (µg/L)	
Cu	35	5	30	4.3	EXP(0.9422*(LN(Hardness))-1.7)*0.96 **(US EPA, 1985)**
Zn	35	48.1	30	42.2	EXP(0.8473*(LN(Hardness))+0.884)*0.978 **(US EPA, 2002)**
Cd	35	0.7	30	0.6	EXP(1.0166*(LN(Hardness))-3.924)*(1.137-(LN(Hardness)*(0.041))) **(US EPA, 2001)**

Copper

U.S. EPA. *Ambient aquatic life criteria for copper, EPA 440-5-84-031*. U.S. Environmental Protection Agency, Office of Water Regulations and Standards. Washington, D.C. 1985

Zinc

U.S. EPA. *National Recommended Water Quality Criteria, EPA- 822-R-02-047*. U.S. Environmental Protection Agency, Office of Water: Washington, D.C. 2002.

Cadmium

U.S. EPA. *2001 Update of Ambient Water Quality Criteria for Cadmium (EPA-822-R-01-001).* USEPA Office of Water. Washington D.C.. 2001

Total to Dissolved Conversion (when applicable)

Stephan, C. E . *Derivation of conversion factors for the calculation of dissolved freshwater aquatic life criteria for metals (620A95001).* U.S. EPA Office of Research and Development. Duluth MN, USA. 1995

Table 4.S2 - Average (+/- s.d.) Cu, Zn and Cd concentrations in the 2010 and 2015 mesocosm experiments. Cumulative Criterion Units (CCUs) were developed using hardness-adjusted metal criteria listed in Table 4.S1.

	Nominal CCU	Observed CCU	Measured Cu (µg/L)	Measured Zn (µg/L)	Measured Cd (µg/L)
2010	0	BRL	BRL	9.6 (2.4) *	BRL
(Cu+Zn+Cd)	3	3.96 (0.35)	6.35 (0.21)	71.18 (10.24)	0.88 (0.08)
Mesocosm	6	6.85 (0.54)	9.85 (2.38)	133.9 (7.91)	1.52 (0.07)
	12	12.24 (0.40)	17.0 (4.96)	251.1 (15)	2.64 (0.23)
	25	23.71 (2.14)	25.37 (3.6)	565.7 (52.9)	5.0 (0.61)
	50	46.9 (2.11)	62.23 (8.5)	965.2 (44.3)	10.5 (0.61)
2015	Control	0.56 (0.09)	0.94 (0.36)	15.26 (1.24)	N/A
(Cu+Zn)	Low	3.94 (0.66)	7.31 (2.14)	95.56 (8.68)	N/A
Mesocosm	Mid-Low	6.95 (1.46)	13.07 (2.10)	167.03 (24.11)	N/A
	Mid	12.68 (0.75)	23.74 (2.46)	305.67 (9.10)	N/A
	Mid-High	24.05 (3.81)	37.56 (8.42)	653.42 (78.32)	N/A
	High	52.95 (2.00)	78.33 (6.48)	1481.6 (73.76)	N/A

Quality Assurance and Repeatability notes: A Varian Spectra AA 22ss atomic absorption spectrometer with deuterium background correction was used to analyze Zn and Cu (>25 µg/L) by flame and Cu and Cd by furnace. Matrix solution for blanks standards and dilutions were made of deionized water (Barnstead Nanopure system; Thermo Fisher Scientific or Milli-Q system, MilliporeSigma). Samples and solutions were preserved with ultra-pure (Ultrex®II, J.T. Baker or equivalent) nitric acid (1ml per L or one drop per 5 ml of sample). Five point calibration for each element was conducted prior to each batch of 20 samples and was analyzed after each batch to ensure no drift. Each batch was accompanied with one duplicate sample at the time of collection and one sample split just prior to analysis, each flagged if duplicate or split was >5 or 10% from original. Blanks were flagged if greater than 5% of detection limit suggested by manufacturer or detection limit calculated from previous batches. External quality assurance standards for each element were assessed every 10 samples and were flagged if >10% from nominal or more frequently at the analyst's discretion. External standards obtained from nationally certified firms were NIST -Traceable to the SRM 3100 Series. Standards had a certificate of analysis and SDS that guaranteed accuracy (99.999% certified accuracy to ±0.3%) and stability. If any QAQC flags were observed instrument was recalibrated and all samples of that batch were reanalyzed. Samples found above the highest standard in the calibration curve were diluted (1:2, 1:5 or 1:10) and reanalyzed in a subsequent batch.

Table 4.S3 - Average (+/- s.d.) and the range of hardness, alkalinity, pH, conductivity and temperature during the 2010 and 2015 mesocosm experiments and the acute single-species tests using *Baetis tricaudatus*. The 2005 Cu and Cu+Zn mesocosm results were published in Clements et. al (2013). Values for the late-instar *Baetis* acute test are available in Brinkman and Vieira (2007). Acute tests include only observations from the controls, because the $ZnSO_4$ added to all non-control waters increased the hardness and conductivity.

	2010 Cu+Zn+Cd	2015 Cu+Zn	2014 First & Mid instar acute
Hardness (mg/L $CaCO_3$)	Not Available	29.5 (0.8) 28-30	51.7 (2.4) 47.4-55.2
Alkalinity (mg/L $CaCO_3$)	Not Available	30.4 (0.5) 30-31	40.3 (2.1) 38.4-45
pH	7.2 (0.01) 7.08-7.26	7.25 (0.09) 7.04-7.51	7.3 (0.4) 6.5-7.7
Specific Conductance (µS)	57.7 (0.2) 56.4-59.8	80.85 (1.69) 78.5-85.1	Not Available
Temp (C°)	12.26 (0.06) 11.7-12.9	12.42 (0.6) 11-13.6	11.8 (0.4) 11-13.5

Brinkman, S. F.; Vieira, N. K. M. *Water Pollution Studies; Job Progress Report; Federal Aid Project F-243-R14*. Colorado Division of Wildlife: Fort Collins, CO, USA. 2007.

Table 4.S4 - Average (mg/L) sulfate (SO_4), total organic carbon, (TOC), calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), and chloride (Cl) concentrations during the calendar month(s) of each mesocosm (2007, 2010, 2015) and single-species (2014) experiment. A water-supply line delivers untreated water from Horsetooth Reservoir (Fort Collins, Colorado, USA) to the Fort Collins Municipal Water Treatment Facility and the Colorado State University Experimental Stream Research Facility, where all mesocosm studies were conducted. Single-species tests using first instar and mid-instar mayflies were conducted at the Colorado Parks and Wildlife Aquatic Toxicology Laboratory using dechlorinated municipal tap water. Data courtesy of Jeff Cannon and the City of Fort Collins Utilities' Water Quality Laboratory (4316 La Porte Ave., Fort Collins, Colorado, USA). BDL=below detection limit. SD= standard deviation.

	2007 Cu	2007 Cu+Zn	2010 Cu+Zn+Cd	2015 Cu+Zn	2014 First Instar	2014 Mid-instar
SO_4	BDL (<5) (n=1)	BDL (<5) (n=1)	BDL (< 5) (n=1)	BDL (< 5) (n=2)	12.8 (n=1)	12.2 (n=1)
TOC	3.08 (n=4, SD=0.1)	3.06 (n=4, SD=0.07)	3.7 (n=5, SD=0.2)	3.6 (n=9, SD=0.1)	1.9 (n=4, SD=0.13)	1.8 (n=4,se=0.05)
Ca	8.6 (n=1)	9.0 (n=1)	10.5 (n=1)	9.5 (n=6, SD=0.28)	16.0 (n=2 se=0.35)	15.7 (n=3 se=0.1)
Mg	1.7 (n=1)	1.6 (n=1)	1.3 (n=1)	1.8 (n=4, SD=0)	1.8 (n=1)	1.9 (n=1)
Na	2.7 (n=1)	2.6 (n=1)	2.9 (n=1)	3.0 (n=4, SD=0)	3.0 (n=1)	3.2 (n=2, SD=0.07)
K	0.8 (n=1)	0.8 (n=1)	0.8 (n=1)	0.8 (n=4, SD=0.05)	0.8 (n=1)	0.9 (n=1)
Cl	1.3 (n=1)	1.3 (n=1)	2.0 (n=1)	2.6 (n=2, SD=0.1)	3.2 (n=1)	3.2 (n=1)

Table 4.S5 - ANCOVA results from 2007-2010 experiments, EMTRENDS slope estimates and multiple comparisons of slopes.

ANCOVA				**EMTRENDS ESTIMATES & MULTIPLE COMPARISON OF SLOPES**			
Treatment	**Model Term**	**F-Value**	**P-value**	**Instar Size**	**Slope**	**Confidence Limits**	**Mult. Comp. of Slopes**
Baetis Cu (0 - 5.1 CCU)	CCU	11.53	0.0023	<0.7 mm	-3.928	(-0.38, -7.48)	A
	Instar Size	29.44	<0.0001	0.7-0.8 mm	-7.898	(-11.45, -4.35)	AB
	CCU*Instar Size	4.94	0.0082	0.8-1.0 mm	0.431	(-3.12, 3.98)	B
				> 1 mm	-0.298	(-3.85, 3.25)	B
Baetis Cu + Zn (0 - 7.0 CCU)	CCU	7.06	0.0137	<0.7 mm	-0.184	(-2.94, 2.57)	A
	Instar Size	7.10	0.0014	0.7-0.8 mm	-1.019	(-3.77, 1.75)	A
	CCU*Instar Size	1.50	0.2395	0.8-1.0 mm	-1.901	(-4.65, 0.85)	A
				> 1 mm	-3.989	(-6.74, -1.24)	A
Baetis Cu + Zn + Cd (0 - 12.9 CCU)	CCU	44.68	<0.0001	<0.7 mm	-3.266	(-4.33, -2.19)	A
	Instar Size	26.49	<0.0001	0.7-0.8 mm	-1.580	(-2.65, -0.51)	AB
	CCU*Instar Size	3.86	0.0171	0.8-1.0 mm	-1.329	(-2.41, -0.28)	AB
				> 1 mm	-0.891	(-1.96, 0.18)	B

Table 4.S6 - Average (+/- s.e.) Survival (n=4) and Zn concentrations (n=4) in the 2014 acute single-species trials using *Baetis tricaudatus*.

	Nominal CCU	96 Hr Survival (%)	Measured Zn (µg/L)
2014	Control	91.4 (3)	0.34 (0.03)
(Zn)	Low	80.3 (5*)	133.00 (2.2)
B. tricaudatus	Mid-Low	81.5 (8)	300.33 (3.9)
First Instar	Mid	46.5 (11)	642.67 (7.7)
	Mid-High	19.2 (6)	1433.33 (17.8)
	High	1.5 (1)	3263.33 (68.4)
2015	Control	100 (0)	2.78 (0.499)
(Zn)	Low	72.4 (4)	4600 (102)
B. tricaudatus	Mid-Low	18.2 (16)	9380 (147.5)
Mid-Instar	Mid	4.2 (4)	20450 (287)
	Mid-High	11.1 (10)	46550 (1041)
	High	0 (0)	84800 (1335)

*: n=3 because one outlier was removed from consideration.

Quality Assurance and Repeatability notes: A ThermoScientific iCAP 6000 ICP-OES was used to analyze Zn for single species trials. Matrix solution for blanks standards and dilutions were made of deionized water (Barnstead Nanopure system; Thermo Fisher Scientific). Samples and solutions were preserved with ultra-pure (Ultrex®II, J.T. Baker or equivalent) nitric acid (1ml per L or one drop per 5 ml of sample). Five point calibration for each element was conducted prior to each batch of 20 samples and was analyzed after each batch to ensure no drift. At a minimum the following quality assurance samples (QA) were analyzed. Each batch was accompanied with at least one duplicate sample at the time of collection and at least one sample split just prior to analysis, each flagged if duplicate or split was >5 or 10% from original. Blanks were flagged if greater than 5% of detection limit suggested by manufacturer or detection limit calculated from previous batches. External quality assurance standards for each element were assessed every 10 samples and were flagged if >5% from nominal or more frequently at the analyst's discretion. External standards obtained from nationally certified firms were NIST -Traceable to the SRM 3100 Series. Standards had a certificate of analysis and SDS that guaranteed accuracy (99.999% certified accuracy to ±0.3%) and stability. A yttrium internal calibration standard was continuously introduced into the plasma along with each sample. If any QAQC flags were observed instrument was recalibrated and all samples of that batch were reanalyzed. Samples found above the highest standard in the calibration curve were diluted (1:2, 1:5 or 1:10) and reanalyzed in a subsequent batch.