DISSERTATION

EVALUATING THE DIRECT AND INDIRECT EFFECTS OF ACID MINE DRAINAGE ON STREAM ECOSYSTEMS USING FIELD AND NOVEL EXPERIMENTAL APPROACHES

Submitted by

Christopher J. Kotalik

Department of Fish, Wildlife, and Conservation Biology

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
Joseph S. Meyer

ABSTRACT

EVALUATING THE DIRECT AND INDIRECT EFFECTS OF ACID MINE DRAINAGE ON STREAM ECOSYSTEMS USING FIELD AND NOVEL EXPERIMENTAL APPROACHES

Mineral extraction in the western United States has resulted in widespread stream impairment due to habitat degradation and water quality impacts from acid mine drainage (AMD). Stream ecosystems are exposed to differing mixtures and concentrations of dissolved metals and metal oxides, acidified waters, and fine sediments. The models used to predict the effects of these stressors on aquatic life are predominately-derived using laboratory toxicity tests with surrogate test organisms; however, these approaches often have difficulty predicting the responses we observe in nature. Novel experimental approaches such as stream mesocosm testing can assist researchers in identifying the causal mechanisms responsible for sensitivity of naturally occurring stream communities, and they provide an opportunity to reconcile differences between laboratory and field responses to AMD stressors.

Our understanding of the fate and effects of contaminants has evolved as researchers advance our knowledge of the ecological interdependence between aquatic and terrestrial environments. To account for linked aquatic-terrestrial (A-T) effects, the methods researchers use to develop aquatic life criteria should extend beyond the aquatic environment. Mesocosm experiments provide the spatial and temporal resolution required to describe A-T contaminant transfer and subsidy alterations. In Chapter 1 (Kotalik et al. 2019a), we focus on aquatic insect emergence evaluated with mesocosm approaches because of the dual utility that both mesocosms and emergence assessments provide in predicting aquatic contaminant effects on aquatic-

dependent terrestrial ecosystems. We will review literature of mesocosm studies that evaluate aquatic insect adult emergence endpoints, provide key considerations to conducting these experiments, and discuss integrating emergence into the aquatic life criteria development process.

In Chapter 2 (Kotalik and Clements 2019), we present the results of two stream mesocosm experiments that exposed aquatic insects to mixtures of Cu and Zn. We compared responses of larvae and emerging adults in a single-species experiment with the mayfly *Rhithrogena robusta* and a benthic community experiment. Results showed that *R. robusta* larvae and emerging adults were highly tolerant of metals. In the benthic community experiment, larval and emerging adult life stages of the mayfly Baetidae were highly sensitive to metals exposure, with significant alterations in adult sex ratios. In contrast, the emergence of Chironomidae (midge) was unaffected, but larval abundance strongly decreased. Timing of adult emergence was significantly different among treatments and varied among taxa, with emergence stimulation in Chironomidae and delays in emergence in *R. robusta* and Simuliidae. Our results demonstrate that metal tolerance in aquatic insects is life stage dependent and that taxa sensitivity is influenced by a combination of physiology and phylogeny. Regulatory frameworks would benefit by including test results that account for effects of contaminants on metamorphosis and adult insect emergence for the development of aquatic life standards.

Ferric iron (Fe(III)) oxyhydroxides commonly precipitate at neutral pH and in highly oxygenated conditions in waterways receiving acid mine drainage, degrading stream benthic communities by smothering habitat, primary producers, and aquatic invertebrates. In Chapter 3 (Kotalik et al. 2019b), stream mesocosms were used to expose naturally colonized benthic communities to a gradient of ferric Fe (0-15 mg/L) for 14 d to estimate the effects of Fe

precipitates on primary production, larval and emerging adult aquatic insects, and macroinvertebrate community structure. Community composition was significantly altered at concentrations near or below the USEPA chronic Fe criterion (1.0 mg/L). Iron exposure significantly decreased larval and emerging adult abundances of Baetidae (mayfly) and Chironomidae (Diptera); however, while Simuliidae (Diptera) larvae were not reduced by the Fe treatments, abundance of emerged adults significantly decreased. Iron substantially decreased the colonization biomass of green algae and diatoms, with estimated EC20 values well below the Fe criterion. In contrast, cyanobacteria were stimulated with increasing Fe concentration. By integrating environmentally realistic exposure conditions to native benthic communities that have complex structural and functional responses, the ability to predict the effects of Fe in the field is improved. Traditional toxicity testing methodologies were not developed to evaluate indirect effects of contaminants, and modernized approaches such as community mesocosm experiments better characterize and predict responses in aquatic ecosystems outside the laboratory. Therefore, the development of water quality standards would benefit by including mesocosm testing results.

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. In Chapter 4 (Cadmus et al. 2019), we exposed field colonized benthic communities 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 mayflies (Ephemeroptera: *Baetis tricaudatus*) 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 our 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.

The North Fork of Clear Creek (NFCC), CO, USA was historically impaired by AMD from two major point-source inputs, with certain stream segments devoid of aquatic life prior to remediation. In the summer of 2017, the North Clear Creek Water Treatment Plant (NCCWTP) began AMD treatment. In Chapter 5, to predict and characterize the biological recovery of NFCC to improvements in water quality, we used a Before-After Control-Impact biomonitoring study supplemented with stream mesocosm and field experimental approaches. The NFCC stream community responded rapidly to improved water quality. Benthic algal biomass increased at impacted sites, and colonizing and *in situ* benthic macroinvertebrate surveys showed significant increases in abundance, taxa richness, and emerging adult aquatic insects. However, the dominant taxa that we predicted to recover given stream mesocosm results strongly differed from the dominant taxa colonizing downstream. We hypothesize that this discrepancy is the result of differences in exposure regime (i.e., stable vs fluctuating), colonization attributes (i.e., closed vs open system), and temporal differences in sensitivity to physical and chemical stressors. The novel approaches used in this research improved our ability to predict, characterize, and validate effectiveness of AMD remediation. We expect continued biological recovery in NFCC, but habitat impairment and elevated concentrations of dissolved metals will continue to impair aquatic life. Applying both experimental and field approaches to predict and evaluate AMD-remediation projects in the future will improve our ability to understand the physical, chemical, and biological mechanisms influencing stream recovery.

Cadmus, P., Kotalik, C. J. et al. 2019. Experimental assessment of size-dependent sensitivity of aquatic insects to metals- homage to Horton Hears a Who! *Environmental Science & Technology (in review)*

Kotalik, C.J., P. Cadmus, W.H. Clements. 2019. Indirect effects of iron oxide on stream benthic communities: capturing ecological complexity with controlled mesocosm experiments. *Environmental Science & Technology* 53*:*11532-11540

Kotalik, C.J., W.H. Clements. 2019b. Stream mesocosm experiments show significant differences in sensitivity of larval and emerging adults to metals. *Environmental Science & Technology* 53:8362-8370

Kotalik, C.J., T.S. Schmidt, C.A. Mebane. 2019a. Mesocosms to evaluate aquatic-terrestrial contaminant linkages using aquatic insect emergence: Utility for aquatic life criteria development. Contaminants and Ecological Subsidies at the Land-Water Interface. Edited Volume. (*accepted)*

ACKNOWLEDGMENTS

This dissertation research would not have been possible without the support of my family, friends, and colleagues. I want to first thank my wife Alexandra for her endless support through this journey, she selflessly helped with much of this research and provided emotional support through the stress and demands of graduate school. I could not have done it without her. I want to thank my Mom, Dad, and my brothers for keeping me grounded, from weekend meals to adventures together camping, fishing, and hunting. David Rees for the mentorship, guidance, and friendship. Pete Cadmus— the majority of this research was generated from long discussions and collaboration over the last decade. Boris Kondratieff for teaching me the foundations of aquatic entomology and for providing support and advice dating back to my undergraduate degree. Numerous work-study students that have spent countless hours in the ecotoxicology laboratory making this research possible. Fellow Ecotox Lab members and FWCB graduate students for the support, comradery, and friendship, we were in it together. My graduate committee for providing outstanding help and constructive criticism. Lastly, Will Clements for inspiring all of this research, for providing countless opportunities, and for always being patient and present as a mentor.

TABLE OF CONTENTS

CHAPTER 1

MESOCOSMS TO EVALUATE AQUATIC-TERRESTRIAL CONTAMINANT LINKAGES USING AQUATIC INSECT EMERGENCE: UTILITY FOR AQUATIC LIFE CRITERIA DEVELOPMENT

Introduction

Aquatic and terrestrial environments interact through movements of resources, nutrients, and prey, and they are not functionally independent of one another (Hynes 1975, Vannote et al. 1980, Nakano and Murakami 2001, Baxter al. 2005). As a result, aquatic contaminants have the potential to impact terrestrial environments through alterations to aquatic subsidies and contaminant transfer to terrestrial ecosystems (Walters et al. 2008, Paetzold et al. 2011, Kraus et al. 2014a). Developing aquatic life criteria to account for aquatic contaminant effects on linked terrestrial ecosystems requires contaminant evaluation methodologies that have the experimental capacity to assess these relationships. Traditionally utilized single-species assessments fall short of integrating the level of biological complexity needed to evaluate these interactions (Cairns 1986, Clements and Kiffney 1994, Versteeg et al. 1999, Buchwalter 2017). Alternatively, mesocosm toxicity testing offers the spatiotemporal scale, control, and degree of replication required to assess aquatic-terrestrial (A-T) contaminant transfer and subsidy alteration to terrestrial ecosystems, and critically, the potential to contribute A-T data for aquatic life development.

In this chapter, we review aquatic community mesocosm studies that incorporate aquatic insect emergence, with the purpose of synthesizing this literature and providing guidance to researchers and regulators in selecting aquatic insect emergence endpoints to make A-T contaminant inferences. We provide key considerations to conducting mesocosm studies using aquatic insect emergence and suggest how these data may be integrated into the aquatic life criteria development process.

Traditional Single-Species Testing

Single-species toxicity testing is the main source of biological data used for aquatic life criteria development in the United States (US) (Stephan et al. 1985). These tests are acute or chronic exposures on surrogate organisms intended to represent the diversity of aquatic animals in cold and warm waters, and with benthic and pelagic life stages. For acute tests, dose-response toxicity tests (generally less than 5 d) estimate lethal concentrations that reduce organism abundance by 50% (LC50). For chronic tests (greater than 24 d), endpoints such as long-term mortality, growth, and reproduction are used to generate no and lowest observed effects concentrations (NOEC, LOEC), or EC20 values (effect concentration that reduced endpoint by 20% from controls). Variation in chronic and acute chemical sensitivity among species is described using an empirical distribution function, the species sensitivity distribution (SSD). The reliance on single-species testing is largely directed by the Environmental Protection Agency's (U.S. EPA) 1985 guidelines for "deriving numerical national water quality criteria for the protection of aquatic organisms and their uses" (Stephan et al. 1985). These guidelines assume that effects which occur in standardized single-species laboratory tests will generally occur for the same species in comparable field situations. If the specified minimum number of organisms are tested, the resultant criteria are expected to be sufficient to protect most aquatic species in the US (Stephan et al. 1985).

Standardized single-species tests are valuable in hypothesis generation, informing relative toxicities of contaminants, and testing contaminant mixtures; however, exclusively relying on these tests to develop criteria has been criticized due to lack of ecological realism and extrapolating results to higher levels of biological organization (Cairns 1986, Buchwalter et al. 2017). For example, single-species tests of dissolved metal toxicity often find aquatic insects highly tolerant, while effects in the field are observed at significantly lower concentrations (Clements et al. 2013, Mebane et al. 2013, Iwasaki et al. 2018). Other concerns with use of SSDs to establish water quality criteria are that they only require aqueous exposures, though diet is a critical route of exposure for contaminants (Xie et al. 2010, Kim et al. 2012, Poteat and Buchwalter 2014). Furthermore, commonly used surrogate test species (e.g., *Daphnia*) do not adequately represent responses of other taxonomic groups such as aquatic insects, which are generally underrepresented in toxicity tests and sensitive species distributions (Cairns 1986, Clements and Kiffney 1994, Pond et al. 2008). Finally, these tests by definition exclude interspecies relationships (e.g., competitive release, resource competition) and trophic interactions (e.g., A-T linkages), which drive indirect effects of contaminants on natural communities (Clements et al. 2013, Kraus et al. 2016, Rogers et al. 2016).

Regulatory Perspectives of Mesocosm Testing

One solution to address many of the issues associated with single-species testing for criteria development is to use a methodology that increases realism without sacrificing causality. Aquatic mesocosm studies offer environmentally realistic approaches to aquatic toxicity testing. Here, we define ‘mesocosms’ as experimental systems that integrate the abiotic and biotic components of natural aquatic communities under controlled conditions. We use the term broadly, without attempting to discriminate between typical associations of indoor ‘microcosms’ or outdoor ‘mesocosms.’ Mesocosms incorporate higher levels of biological organization than single-species tests (e.g., organisms interacting at multiple trophic levels), and have the potential to more accurately predict changes to actual communities in nature compared to laboratory testing alone (Clements 2004). Traditional endpoints used in single-species toxicity tests such as mortality and growth can be assessed with mesocosms; however, certain functional endpoints are only possible to assess with mesocosms. For example, rates of aquatic insect emergence and stream metabolism require biological complexity (i.e., natural occurring community assemblages) and realistic testing conditions (i.e., flow through conditions with natural habitat).

Similar to traditional aquatic toxicity testing methodology, mesocosms can generate concentration-response relationships, as well as commonly used metrics for populations (e.g., EC50: concentration reducing maximal response by 50% compared to the controls). These comparisons can integrate results obtained from both single-species testing and field responses. For example, mesocosm-derived EC50 values can be compared to laboratory LC50 results, and metrics characterizing mesocosm community responses can evaluate aquatic communities in the field. Furthermore, these experiments can apply contaminant-gradient relationships to communities and functional endpoints (Clements 2000, Hickey and Golding 2002), and analyze changes in community structure using multivariate approaches (Van den Brink 1998, Van den Brink 1999) and structural equation models (Grace 2006, Rosseel 2012, Rogers et al. 2016).

In the United States, mesocosm testing has historically received mixed reviews from researchers and regulators on their usefulness for regulatory purposes, but other countries including Canada and the European Union have been less divided in their opinion regarding the utility of toxicity data generated by mesocosms. For example, from 1988 through 1992, the U.S. EPA required higher tier risk assessments (e.g., mesocosms) to study secondary effects in non-target organisms for pesticide registration under the Federal Insecticide, Fungicide, and Rodenticide Act using simulated aquatic field studies (Touart 1988). However, the U.S. EPA rescinded mesocosm testing requirements for pesticides on the basis that “they do not provide substantial information for making risk decisions beyond that already revealed by lower tiered (e.g., single-species laboratory testing) studies,” with key criticisms regarding their high variability, lack of repeatability, and impractical experimental scale (U.S. EPA 1993). Instead, the U.S. EPA focused their efforts on improving laboratory-based risk assessments and mitigating and monitoring pesticides in the field, whereas the regulatory focus in Europe was to improve mesocsom practices by standardizing methods for evaluating population *and* community-level effects (Van den Brink 2006).

Presently, development of water quality guidelines in the U.S. allows for, but does not mandate, the use of mesocosm data in criteria development. Criteria to protect aquatic life can utilize “other data” from mesocosm testing in a secondary role as a validation step that can be used to lower criteria values if mesocosm data indicate that criteria calculations from specified single-species testing are not sufficiently protective (Stephan et al. 1985). In contrast, Canada’s updated aquatic life development process uses available mesocosm data in a primary role to derive criteria, as long as the data meet “quality acceptability” requirements. The Canadian guidelines state, *“Controlled microcosm and mesocosm studies are acceptable and are ranked according to the applicable categorization criteria. A clear dose-response relationship should be experimentally established and effects reasonably apportioned to the substance.”* (CCME 2007).

Mesocosms certainly have their limitations; for example, investigators have encountered enclosure bias (Bloesch et al. 1988, Carpenter 1996, Schmidt et al. 2018), high variability and low signal to noise ratio (Caquet et al. 2001, Van den Brink 2013), significant experimental costs (Perceval et al. 2009), and poor repeatability (Elberhardt and Thomas 1991). Still, comparisons of mesocosm-generated data to single-species and field data have shown valuable mechanistic insights (Clements 2004, Rogers et al. 2016, Mebane et al. 2017), and testing system designs have been developed to reduce variability to the point that concentration-response curves can be fit to larval community metrics (Clements et al. 2013, Schmidt et al 2018, Mebane et al 2017).

Considering Aquatic Insect Emergence

Adult aquatic insects export detrital material and nutrients from aquatic to terrestrial ecosystems (Nakano and Murakami 2001, Gratton and Vander Zanden 2009), and they provide a significant resource subsidy to linked consumers such as birds, bats, and spiders (Sabo and Power 2002, Echols et al. 2004, Baxter et al. 2005). Emerged adult aquatic insects are also critical vectors of contaminant transfer from aquatic to terrestrial environments (Walters et al. 2008, Kraus et al. 2014b). Thus, aquatic insect emergence can be used to characterize aquatic contaminant fate and aquatic subsidy alterations to linked terrestrial ecosystems.

Field observations have predominately driven our understanding of these A-T contaminant linkages (Walters et al. 2008, Paetzold et al. 2011, Schmidt et al. 2013, Kraus et al. 2014a, Kraus et al. 2016). For example, Schmidt et al. (2013) conducted an analysis of field survey data on macroinvertebrate larval and adult densities along a metals gradient in Colorado (USA) mountain streams. They compared larval and adult responses to metal concentration and observed differences in threshold responses between life stages. Specifically, larvae were reduced at or above the cumulative criterion accumulation ratio (CCAR; a toxic unit model modified from the biotic ligand metal bioavailability model; values at or below 1 are considered protective of aquatic life), but adult emergence declined at CCAR values less than or equal to 1. These results suggest uncertainty whether aquatic life criteria are protective of full aquatic insect life cycles, and thus protective of A-T linkages. However, our ability to develop aquatic life criteria using field surveys alone is limited by replication, inability to control for extraneous variables, and naturally high variability.

Furthermore, traditional single-species toxicity tests are too limited in predicting cross ecosystem effects. For example, aquatic invertebrate species such as *Daphnia* spp. are widely employed because they are relatively sensitive and used in many single-species tests, but zooplankton do not metamorphose into a terrestrial life stage; therefore, no equivalent endpoint exists to estimate these effects. Most natural assemblages of aquatic (e.g., lotic or lentic) communities contain aquatic insect taxa that complete their life cycle as winged-terrestrial adults (Merritt and Cummins 1996). Certainly, some surrogate test organisms employed with single-species testing do metamorphose to an imago (e.g., adult) life stage, such as the hemimetabolous mayfly *Neocloeon triangulifar* (Wesner et al. 2014, Soucek and Dickinson 2015), and the holometabolous midge *Chironomus* spp. (Sibley et al. 1997, Mebane et al. 2008). These metamorphosing invertebrates are important in evaluating contaminant effects through their emerged adult life stages, but using single-species testing alone to predict aquatic-derived terrestrial contaminant effects fails to incorporate the complexity of natural aquatic ecosystems that influence exposure outcomes.

Adult Insect Emergence Assessments with Mesocosm Studies

By quantifying adult emergence in aquatic mesocosm experiments, we can estimate potential effects to linked terrestrial ecosystems using experimentally robust methodology. Thus, cause and effect relationships relating aquatic contaminant exposure to terrestrial effects can be quantitatively supported. Importantly, most mesocosm facilities can practically quantify emergence by modifying netting and emergence traps for their respective test systems. **Table 1.1** summarizes aquatic insect adult emergence endpoints conducted with community mesocosm studies described in this chapter. Each mesocosm experiment has accompanying endpoint responses, contaminant and (or) stressor inferences, and potential A-T outcomes. The goal of detailing these mesocosm studies is to provide guidance to researchers and regulators in selecting aquatic insect emergence endpoints that they can apply to research and develop aquatic life criteria.

Life Stage Sensitivity

Larval abundances in aquatic ecosystems are often assumed to be linked to the abundances of emerging insect adults (Statzner and Resh 1993, Gratton and Vander Zanden 2009). However, metamorphosis of aquatic organisms from their larval to adult life stages is a biologically stressful process (Campero et al. 2008), and exposure to contaminants either before or during this transition may cause organisms to be even more sensitive when metamorphosing (**Figure 1.1**, Wesner et al. 2014). This pattern is seen in both single-species laboratory exposures and field surveys, with adult life stages displaying greater sensitivity compared to larvae (Palmquist et al. 2008, Schmidt et al. 2013, Wesner et al. 2014, Debecker et al. 2017). During the metamorphosis process, substantial reorganization of tissues occurs; and therefore, metamorphosing life stages of invertebrates may be more susceptible compared to earlier molting events (Liber et al. 1996, Wesner et al. 2014). Most acute single-species laboratory toxicity tests lack sufficient exposure duration and variation in size structure to evaluate effects among aquatic insect life stages. Furthermore, field surveys have confounding variables (e.g., seasonality, immigration, emigration) that affect the ability to characterize sensitivity of larvae and emerging adults. If invertebrates are unable to metamorphose and complete their life cycles as emerged adults, population persistence is at risk (Zwick 1990, Rasmussen and Rudolf 2015), and biomass available to terrestrial consumers is reduced (Kraus et al. 2014a).

Mesocosms are particularly useful for assessing contaminant sensitivity through metamorphosis because they can integrate and expose benthic communities through multiple developmental life stages, including the transition to emerged adults (Kotalik and Clements 2019). Differences in sensitivity among aquatic insect life stages, including larvae transitioning to adults, have been confirmed with numerous community mesocosm studies (Kiffney and Clements 1996, Clements and Kotalik 2016, Kotalik et al. 2017, Kotalik et al. 2019, Kotalik and Clements 2019). Community mesocosm exposures have shown greater relative sensitivity in emerging adults compared to larval populations in Ephemeroptera (Culp et al. 2003, Colombo et al. 2013), Chironomidae (O' Halloran et al. 1996), and other Diptera (Kotalik et al. 2019, Brock et al. 2009). Likewise, Mohr et al. (2011) did not observe a significant difference in larval community structure when exposed to pulses of imidacloprid but did observe significant alterations in emerging adult community structure. In contrast, other mesocosm studies have shown just the opposite relationship, with emerging adult aquatic insects more tolerant to exposure than early larval life stages. Instances of reduced larval abundances, but not emerging adult abundances, have been observed with pesticides (Arthur 1983, Beketov et al. 2008), household and personal care products (Belanger et al. 2004, Belanger et al. 2005), bacterial insecticides (Richardson and Perrin 1994), and metals (Kotalik and Clements 2019).

Critically, emerged adults must develop to a minimum weight before larvae can metamorphosis, which influences comparisons of life stage sensitivity. For example, metamorphosis may be delayed to facilitate growth (i.e., fecundity) if environmental conditions are suitable (Sibley et al. 1997, Peckarsky et al. 2001), or individuals may avoid exposure by emerging before maximum body size is achieved (see *Timing of Emergence*). Contaminant tolerance in developmentally mature individuals can be beneficial for population persistence because female aquatic insects can disperse to less contaminated aquatic environments to lay eggs, and because some aquatic insect egg life stages are relatively more tolerant to contaminants compared to their larvae life stages (Kefford et al. 2007, Schriever and Liess 2007).

Timing of Emergence

Altered rates of emergence can indicate stress related to direct (i.e. aqueous or dietary routes of exposure) or indirect (e.g., reduced food) toxicity to aquatic communities. For example, aquatic insects may speed up or delay their larvae development either to reduce their exposure, or as a direct consequence of toxic effects. Irregular timing of emergence may result in individuals encountering terrestrial conditions that are unsuitable for their adult survival (Nebeker 1971) and shift the timing of A-T resource subsidies that terrestrial consumers are temporally cued to encounter (Takimoto et al. 2002, Yang et al. 2010, Hastings 2012).

Increased rates of emergence in community mesocosm studies have been observed with trace metals (Joachim et al. 2017, Kotalik and Clements 2019), pesticides (Dewey 1986, Maund et al. 1992, Gruessner and Watzin 1995, Schulz and Liess 2001, Belanger et al. 2004, Farmer et al. 1995, Alexander et al. 2008, Rogers et al. 2016), illicit drugs and pharmaceuticals (Lee et al. 2016, Richmond et al. 2016). Increases in rates of emergence can be beneficial because emerged adults may reduce larval exposure and complete their life cycle as emerged adults; however, increased rates of development may also reduce size (i.e., fecundity) of emerging adults that have less time for growth. In contrast, delays in emergence have been reported with petroleum exposures (Pettigrove and Hoffman 2005) and reduced dissolved oxygen treatments (Nebeker 1972, Connolly et al. 2004). Delayed emergence may be the result of contaminant-induced stress that reduces rates of tissue reorganization required for the morphologically transition from aquatic to terrestrial life stages.

Contaminants can directly affect rates of emergence by altering neurological function (Bloomquist 1996), metabolic rates (Nebeker 1972), and hormone balance (Zhao et al. 2009, Nation 2015). In addition, altered rates of insect emergence have been associated with the effects of contaminants on food palatability (Postma et al. 1994), and reductions in overall food resources (Dewey 1986, Gruessner and Watzin 1995). Single-species tests are not able to account for consumer-resource interactions, while field surveys lack the experimental control to characterize timing of emergence due to numerous potential confounding effects (e.g., immigration, predation, climatic variability). Mesocosm community exposures are particularly useful to evaluate timing of emergence because indigenous taxa at various life stages of development and with complex physiologies respond directly (i.e., toxicant accumulation) and indirectly (i.e., impaired food resources) to contaminants.

Size and Biomass of Emerging Adults

Size and biomass of emerging adult aquatic insects provides a useful endpoint in addressing contaminant effects to A-T linkages. Reductions in emerging biomass can be related to direct toxicity to larvae, which obviously eliminates emergence (Belanger et al. 2002), or mortality of individuals during the metamorphosis process (Wesner et al. 2014, Debecker et al. 2017). Reductions in biomass and size of adult aquatic insects can be linked to altered food resources, with autotrophic stimulation or suppression resulting in changes to food availability in larvae that then affects emerging adult body size (Blumenshine et al. 1997, Sardina et al. 2017). Mesocosms are unique in identifying the underlying mechanisms influencing changes in emerging biomass because autotrophic functions (e.g., periphyton growth) are included, and changes in food resources can be in direct response to the contaminant. In contrast, single-species tests commonly feed organisms *ad libitum*, and the food resource does not directly respond to the stressor. Thus, single-species tests can only characterize direct toxicity on insect larvae leading to reduced emergence biomass or size.

Mesocosm studies are useful to decouple alterations in autotrophic production with emerging size and overall biomass. For example, mesocosm exposures have linked contaminant-induced autochthonous stimulation with increased body size and biomass of emerging insects (Mundie et al. 1991, Blumenshine et al. 1997, Perrin and Richardson 1997, Culp et al. 2003, Piggot et al. 2015, Elbrecht et al. 2016). Similarly, elevated water temperatures that increase primary production have been shown to increase the total biomass of emerging insects (Nordlie and Arthur 1981, Greig et al. 2012, Piggott et al. 2015, Sardina et al. 2017); however, aquatic insects in some of these thermal exposures actually emerged as smaller adults (Piggot et al. 2015, Sardina et al. 2017). While autochthonous subsidies contribute to food web stability (Nakano and Murakami 2001), unnatural increases in subsidy biomass may decouple consumer-resource interactions and reduce stability (Huxel and McCann 1998, Hocking and Reynolds 2011). For example, linked terrestrial consumers may shift to more abundant emerging aquatic insects that respond to greater food resources (e.g. strengthening these linkages), but reduce their top-down control of other terrestrial-derived consumer resources (e.g., weakening these linkages).

Aerial Recolonization Dynamics

Recovery of natural aquatic ecosystems following disturbances by natural and anthropogenic stressors is often facilitated by aerial recolonization by aquatic insect adults (Merritt and Cummins 1996). Long-duration community mesocosm studies have demonstrated the importance of understanding invertebrate responses throughout their life cycles for predicting contaminant effects and recovery dynamics (Boyle et al. 1996, Schumde et al. 1999, Caquet et al. 2007, Beketov et al. 2008, Mohr et al. 2012). Outdoor mesocosms that enable external colonization by winged-adults have confirmed the importance of egg laying by adults for benthic recovery following pulse pesticide exposures (Caquet et al. 2007, Beketov 2008). Results from these studies have demonstrated differences in recovery propensity in multi-generational (multivoltine) groups of organisms, compared to longer-lived (uni- or semivoltine) taxa (Caquet et al. 2007, Beketov et al. 2008, Mohr et al. 2012).

Given that multivoltine taxa have greater probability of recolonizing from nearby source populations as adults, mesocosm experiments that compare “open” versus “closed” systems can test recovery duration required to achieve original taxa density and distribution. Single-species exposures, even with longer, chronic exposure durations (e.g., greater than 30 days), lack the environmental complexity and colonization capacity required to examine such endpoints. For example, recolonization by aerial adults to small indoor test systems (i.e., beakers) is simply not possible. Mesocosm studies offer a suitable spatiotemporal scale to characterize and predict trajectories of benthic community recovery. Importantly, estimating rates of recovery following contaminant disturbances are important for predicting the responses of linked terrestrial ecosystems that receive aquatic subsidies.

Interspecies Interactions and Multiple Stressors

The effects of interspecies interactions and multiple stressors on contaminant and resource transfer by adult aquatic insects can be effectively separated using mesocosms. For example, mesocosm studies have shown that the presence of fish and nutrients can control methyl mercury (MeHg) flux from aquatic ecosystems via emergent aquatic insects (Jones et al. 2013, Tweedy et al. 2013). Greig et al. (2012) demonstrated enhanced cross-ecosystem subsidy exchanges via emerging insects with warming and increased nutrients, but predatory fish eliminated these effects by reducing emerging insect biomass and delaying insect phenology. Similarly, Alexander et al. (2016) used mayfly emergence and stream mesocosm exposures to decouple density-dependent relationships of imidacloprid, nutrient enrichment, and predation (stonefly *Agnetina capitata*) pressures on subsidy-stress responses. The use of adult insect emergence endpoints in these studies uniquely quantified these complex interactions of subsides and stressors on bottom-up and top-down effects.

Multiple stressors can destabilize food webs and decouple trophic linkages that are critical for the recovery of both aquatic and linked terrestrial ecosystems (MacNeale et al. 2010). Mesocosms that measure emergence have separated interactive effects among ubiquitous anthropogenic stressors such as warming temperatures, flow alteration, fine-sediment deposition, pesticides, and nutrients (Greig et al. 2012, Magbanua at al. 2016, Piggot et al. 2015, Alexander et al. 2016, Elbrecht et al. 2016). Magbanua et al. (2016) used mesocosms to demonstrate the additive effects of the insecticide glyphosate and fine sediment on emergence. Similarly, Piggot et al. (2015) observed greater total emergence in response to fine sediment, as well as significant 2- and 3-way interactions with nutrients and temperature, reflecting life-stage-dependent sensitivity. Most all impaired aquatic systems have more than one stressor contributing to disturbances, and novel experimental methodologies such as mesocosms and aquatic insect emergence provide pragmatic approaches to isolate stressor effects and to characterize the complex interactions of physical and chemical stressors.

Mesocosm Testing Considerations for Aquatic Insect Emergence

The scale, duration of exposure, and sources of aquatic communities used in mesocosm exposures will influence data obtained from aquatic insect emergence results. Several documents have been developed outlining "appropriate" mesocosm experimental design for research and regulatory applications (Pontasch and Cairns 1989, Graney et al. 1994, Cairns and Niederlehner 1995). The European Union (EU) has produced several guidance documents for micro- and mesocosm studies for pesticides (European Union 1997, Campbell et al. 1999, European Union 2002). For this section, our aim is to address key mesocosm testing considerations for quantifying aquatic contaminant effects to linked terrestrial ecosystems using aquatic insect communities and adult aquatic insect emergence. These considerations include duration and concentrations of exposure, traits of the sourced communities, size of mesocosms, and appropriate statistical analyses. We also provide basic guidance for incorporating these data into the criteria development process.

Exposure Duration and Concentration

One of the first decisions faced by researchers using mesocosms to test effects of a particular contaminant, or contaminants, on aquatic-terrestrial linkages is to choose the concentrations and duration of exposure. For example, mesocosm exposures can vary tremendously, from days (Clements 2004) to multiple years (Joachim et al. 2017). Because exposure lengths will influence biological outcomes, selecting the duration of exposure for mesocosm experiments is a critical decision. Longer duration exposures are more comparable to field conditions, while shorter durations may be more similar to single-species laboratory studies. Researchers must question if their exposure duration reflects the exposure regime observed in the field, affording adequate time to observe biological effects similar those in nature. Contaminant exposures in aquatic environments can have both short (acute) and long-term (chronic) effects; for example, acute aqueous exposure followed by chronic exposure from the release of contaminants bound in sediments (Liess and von der Ohe 2005). Chronic exposure scenarios, such as toxicant inputs from mining activities or point source effluents, may require a longer duration to reflect temporally dependent rates of accumulation and toxic effects (Poteat and Buchwalter 2014). Ultimately, logistical considerations (e.g., monetary costs, technicians, experimental setup) must be weighed with the ecotoxicological relevance from selected exposure regimes.

Longer exposure durations and (or) longer experiment lengths may be required for mesocosm studies aimed at quantifying aqueous contaminant effects on terrestrial ecosystems because most aquatic insects have complex life histories that require much greater time to complete compared to most cultured aquarium species. Therefore, longer-lived taxa such as uni- or semivoltine invertebrates will require longer exposure durations to capture all developmental instars (life stages) through emergence, whereas multivoltine taxa that have numerous life cycles per year may be sufficiently exposed with less time. Life-cycle traits should be considered because inadequate duration of larval exposure may underestimate effects to longer-lived species (e.g., uni- or semivoltine;Van den Brink et al. 1996, Caquet et al. 2007, Beketov et al. 2008). For example, Beketov et al. (2008) observed the fastest taxa recovery in exposed test systems by multivoltine Diptera (e.g., chironomids, simulids) following pulsed insecticide exposure, but much slower recovery in a univoltine stonefly. Most aquatic insects have well documented life histories (Merritt and Cummins 1996), and researchers have the opportunity to use this information to determine the appropriate exposure duration.

Source Community Traits

A second decision made by researchers in preparation for mesocosm experiments is to determine which communities to expose, and how to collect them from the field. Sourcing communities from reference sites that are devoid of unnatural stressors, and that have not been disturbed by natural stochastic events (e.g., floods, fire, drought), reduces confounding factors that may limit the ability to significantly detect a treatment effect. However, communities sourced from sites with previous exposure, or sites that are exposed to unfamiliar contaminants (i.e., novel stressors), can provide opportunities to examine context-dependent responses.

For researchers interested in measuring emergence endpoints, community assemblages selected for mesocosm experiments need to contain taxa that metamorphose to air-breathing terrestrial adults, which can be determined *a priori* and referenced using Merritt and Cummins (1996). For example, individuals must be developmentally mature to the point of metamorphosis to capture emergence. Either larvae are exposed as they develop and metamorphose to adults, or developmentally mature larvae are present in communities collected from the field and exposed at or close to metamorphosis. Timing of source community collection will also affect experimental outcomes. For example, communities collected in the winter may have very different A-T exposure outcomes compared to communities in mid-summer due to phenological differences in emergence and developmental life stages.

In terms of collecting source communities for mesocosm experiments, various methods exist, such as substrate filled colonization trays, or field collection using quantitative benthic sampling methods (e.g. surber, core sampling) (Pontasch and Cairns 1989, Clements et al. 1989, Pontasch and Cairns 1991, Cadmus et al. 2018). These methods can be used to target annual differences in developmental life stages (e.g., egg, early-mid-late instar, and pupa) among invertebrates in their selected community assemblage. Including a variety of life stages within an experiment is important because many aquatic insect larvae stop feeding as they near metamorphosis (Merritt and Cummins 1996), limiting dietary routes of exposure. Moreover, larval size alone can strongly influence sensitivity to contaminants (Kiffney and Clements 1996, Poteat and Buchwalter 2014), and therefore the ability of individuals to survive to metamorphosis.

Developmental and phenological changes within experimental communities can occur very rapidly during an experiment. Duration of emergence varies among insect taxa; for example, some aquatic insects such as *Pteronarcys* spp. or *Hexagenia* spp. emerge for only one week annually (Giberson and Rosenberg 1994, Sheldon 1999), while other taxa have emergence that extends for many months (e.g., midges, mayflies). Even though some insect taxa may only briefly emerge to the terrestrial environment, superabundant "pulsed" subsides can represent disproportionally large carbon and nutrient flux to terrestrial ecosystems compared to relatively constant "press" emergence (Yang et al. 2010, Walters et al. 2018). Therefore, timing of toxicant exposure to source communities should be carefully planned.

Mesocosm Scale

Mesocosm scale can dictate the community assemblage structure, functional processes, and possible duration of exposure. Small-scale mesocosms (e.g., buckets, small aquaria) typically have greater replication and statistical power and owing to the need for less infrastructure and fewer resources (less water, less chemical), are relatively more cost-effective. However, smaller scale mesocosms are generally constrained to shorter exposure periods (weeks to months, not years) due to difficulty maintaining natural functional processes such as delivery of organic matter and colonization of new organisms. Conversely, larger mesocosms (e.g., ponds, large raceways) generally exhibit greater structural and functional realism because functional processes such as immigration maintain natural population and community dynamics. For example, Joachim et al. (2017) allowed aquatic insect adults to seed and maintain larval populations in their mesocosms through aerial deposition of eggs over a two-year exposure. Larger mesocosms can be implemented for longer exposure assessments, they can be used to evaluate recovery of benthic communities via aerial colonization by adults, and incorporate higher trophic levels interactions (e.g., predatory fish). Yet, large mesocosms consume relatively more resources and take up more space, resulting in fewer replicates that reduce statistical power, and because they are often open to the environment are more susceptible to stochastic events (e.g., un-wanted predation or vandalism). Researchers do not always have a choice regarding scale of their experiments, which may be limited to available facilities. Although most A-T endpoints are possible to quantify with most mesocosm experimental scales, the relationship between scale and exposure duration is tightly coupled, with extended exposure regimes requiring larger mesocosms in order to maintain functional processes.

Statistical Considerations

Finally, researchers must choose statistical analyses that are appropriate to their research questions and experimental design. Statistical analyses of mesocosm treatment effects are often conducted at the population level, and most univariate techniques (e.g., ANOVA, regression, and GLM), as well as *post hoc* comparisons (e.g., Dunnets, Duncans, and Williams) employed to analyze larval treatment responses can be applied to aquatic insect emergence data. Temporally-dependent A-T endpoints can be tested with repeated-measures ANOVA (Magbanua et al. 2015, Kotalik and Clements 2019), or regression-based analyses with normalized timing data (Rogers et al. 2016, Schmidt et al. 2018). For example, the time to 50% of emergence (TE_{50}) can be calculated by fitting a logistic equation to time (days) versus the number of individuals emerging each day (Basic OriginPro, Origin Lab, Northampton, MA). The equation to estimate time to reach 50% emergence is:

$$y = a_2 + (a_1 + a_2)/(1 + \left(\frac{x_i}{x_0}\right)^p)$$

where y is the number of emerging adults at time x_i, a_1 is the model-predicted minimum number of emergers, a_2 is the model-predicted maximum number of emerging adults, $x_0$is the time at 50% emergence, and p is the power of the equation.

Emergence dynamics (i.e., the cumulative number of emerging insects over time) can be calculated using cumulative deviance of emergence through time. The equation to estimate difference in cumulative adult emergence between treatments is:

$$y_i = \sum T_i - \sum C_i$$

where y_i is the difference in cumulative emergence between treatment (T_i) and control (C_i), and T_i is calculated as the daily sum of emerging adults since the start of the experiment.

Alterations in community structure resulting from differences in population-level sensitivity (e.g., direct mortality) or organism emigration (e.g., emergence) can be the most sensitive indicators of toxicant stress (Cairns and Niederlehner 1995). Thus, a major strength of mesocosm studies is that communities of aquatic organisms are exposed to aquatic contaminants and community-level responses are evaluated. Statistical procedures have been specifically developed for large-scale mesocosm studies to analyze community treatment effects (Van den Brink and Ter Braak 1999), and the multivariate power given community assemblage structure (Timmerman and Ter Braak 2006). The Principal Response Curve (PRC) method analyzes time-dependent multivariate responses of communities to stress, and is frequently applied to emergence data. PRC is based on the Redundancy Analysis (RDA) technique and is a constrained form of principal component analysis (Van den Brink and Ter Braak 1999). PRC works by separating three sources of variation in community composition: 1) changes in time, 2) differences between replicates, and 3) effects of treatments. Importantly, PRC can include hypothesis testing using Monte Carlo permutation with RDA that is capable of determining community-level “Lowest Observed Effect Concentration” and “No Observed Effect Concentration.” Functional endpoints such as rates of emergence can exhibit high intra-treatment variability, and analyses at the community level with PRC are particularly valuable and have been shown to be more sensitive in detecting changes over time compared to univariate analyses (Van den Brink and Ter Braak 1998). Permutation testing for “significant” treatment differences should employ a low numbers of categorical variables (in contrast to regression-based designs), and experimental duration must be adequate to gather numerous observations. **Figure 1.2** is an example of PRC from Joachim et al. (2017), demonstrating the effects of copper through time on the abundance of emerging aquatic insects.

Conceptual Model and Conclusion

Regulatory frameworks such as aquatic life criteria could benefit from the integration of mesocosm studies generally, and aquatic insect emergence is a particularly relevant endpoint for protecting aquatic dependent life through its A-T linkage. Our hope is that aquatic insect emergence endpoints, along with other endpoints that link A-T contaminant fate and effects, are incorporated into mesocosm testing procedures and are used in a “primary role” in criteria development. Of course, mesocosm studies represent just one line of evidence in evaluating contaminant effects, and researchers and regulators could benefit by working together to employ a comprehensive approach utilizing laboratory, mesocosm, and field data to develop aquatic life criteria that are inclusive of linked terrestrial effects. As mesocosms become more useful as predictive tools, their utility will no doubt be tested, but we should address rather than shelve these challenges if we wish to make meaningful progress in modernizing aquatic toxicity testing approaches that include the fate and effects of contaminants in linked A-T ecosystems.

Integrating Aquatic Insect Emergence into Aquatic Criteria Development

Here we provide a very basic conceptual workflow for integrating aquatic insect emergence into the aquatic life criteria development process (**Figure 1.3**; modified from Buchwalter et al. 2017). First, single-species toxicity tests using a diverse representation of holometabolous and hemimetabolous aquatic insects can be used to evaluate contaminant effects across multiple developmental life stages, including metamorphosis to adults. Second, life stage sensitivity can be precisely determined using contaminant exposure gradients, and contaminant transfer from larvae and adults can be quantified. Third, species sensitivity distribution (SSD) models (or other normalizing models) can generate, for a specified protection level (e.g., 95%), concentrations for both larvae and adult life stages. Mesocosm testing can then evaluate the laboratory-derived SSD models by comparing single-species SSD results to the responses of natural aquatic communities. Moreover, normalizing models to generate protective concentrations can be applied to mesocosm data, and importantly, they can include the larval and adult life stage responses of indigenous taxa obtained from mesocosm testing, as well as community-level treatment effect derivations.

Lastly, aquatic (e.g., benthic and emergence) and terrestrial (e.g., riparian consumer density and contaminant concentrations) field sampling methods can be used to calibrate the relationship of altered insect emergence to linked terrestrial effects and (or) contaminant transfer and to verify criteria protectiveness. Reductions in emerging biomass may disproportionally affect riparian consumers (e.g., growth rates, presence or absence) (Sabo and Power 2002a, Murakami and Nakano 2002, Paetzold et al. 2011, Kraus et al. 2014a), or effects may be less than predicted due to the adaptability of terrestrial consumers to subsidy perturbations (Sabo and Power 2002b, Kato et al. 2003, Spiller et al. 2010). The regulatory community can utilize these different approaches together to ultimately decide what rates of aquatic insect adult contaminant transfer and subsidy alterations are acceptable.

Aquatic insect emergence endpoints vary in their ability to be quantified in the laboratory, with mesocosms, or in the field, because of their respective spatiotemporal scales and abilities to control and replicate. Nevertheless, these differences should not be seen as a limitation, rather researchers can leverage all applicable ecotoxicological methods, identify their respective strengths and weaknesses, and use these multiple lines of evidence to develop aquatic life criteria that are inclusive of contaminant fate and effects in linked A-T ecosystems. Differences between laboratory and mesocosm results may also be seen as a weakness in applying mesocosm results to aquatic life standards. However, these differences can be used to generate hypotheses of the mechanisms influencing these differences (e.g., indirect effects, seasonality, or other factors), and additional research can refine criteria derivations.

Concluding Remarks

To standardize the effects that A-T linkages have to aquatic ecosystems, we need experimentally robust methods that integrate the appropriate biological complexity. We focused on adult insect emergence because of its linkage to aquatic impairment and terrestrial effects. However, we recognize that subsidy alterations and contaminant delivery to terrestrial ecosystems is affected by other processes such as behavioral drift, aquatic primary productivity, and secondary production. Without field validations, the application of A-T data gathered from mesocosm studies will be limited in deriving aquatic life criteria. More field biomonitoring is needed to link aquatic impairment to terrestrial effects among differing ecosystems if we wish to meaningfully test these relationships in the future. To conclude, mesocosm experiments can successfully quantify A-T linkages, and data obtained from these experiments have great potential to contribute to the aquatic life criteria development process.

Figures and Tables from Chapter 1

Figures 1.1 to 1.3

Table 1.1

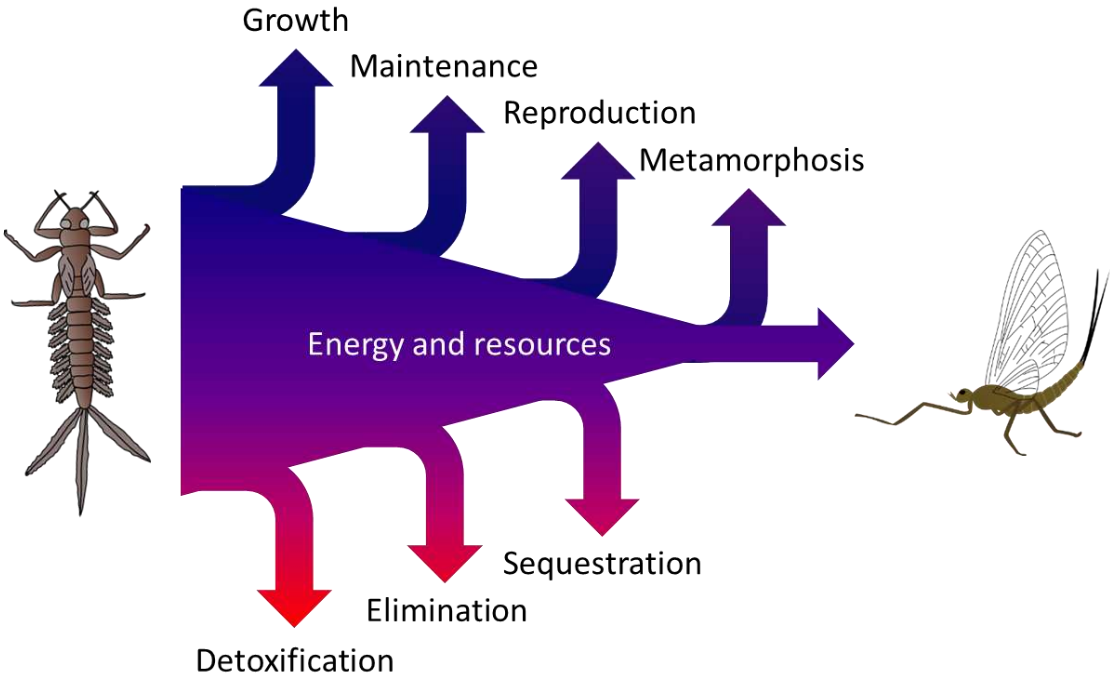

Figure 1.1 Bioenergetics of aquatic insects developing from larvae to their adult life stages. Upward arrows represent energy and resources used by individuals for growth, maintenance, reproduction, and metamorphosis. In the presence of contaminants, energy and resource “costs” are associated with detoxification, elimination, and sequestration.

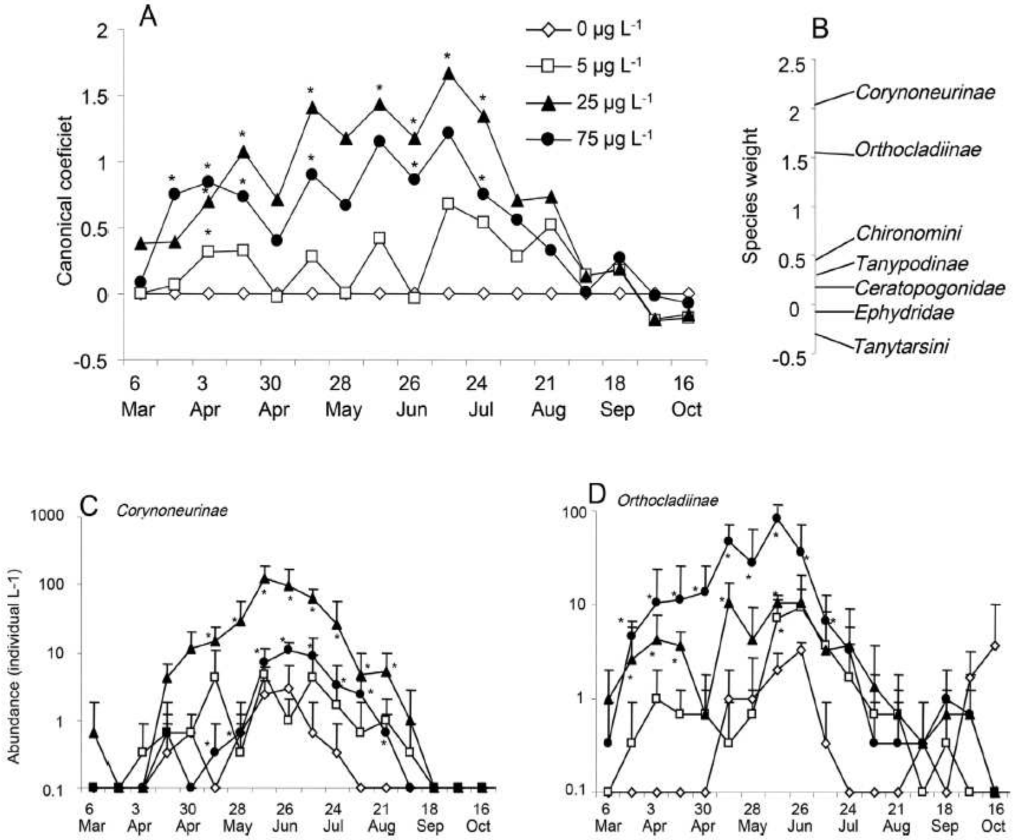

Figure 1.2 An example of the use of Principal Response Curves to compare community shifts in treatments against controls over time, including the influence of specific species. (A) Shows the effect of copper through time based on the canonical coefficient of the emerging insect community. Curves deviating from the reference value of 0 indicate treatment effects. The species weight can be interpreted as the affinity of the taxon with the curves (B). Notice that species weight between 0.5 and –0.5 are not presented, as they were likely to show a weak or unrelated response. Dynamics of the two most important species are presented in a logarithmic scale: Corynoneurinae spp. (C), Orthocladiinae spp. (D). Asterisks indicate significant difference relative to controls (Williams test, $P < 0.05$). From Joachim et al. (2017), reprinted with permission.

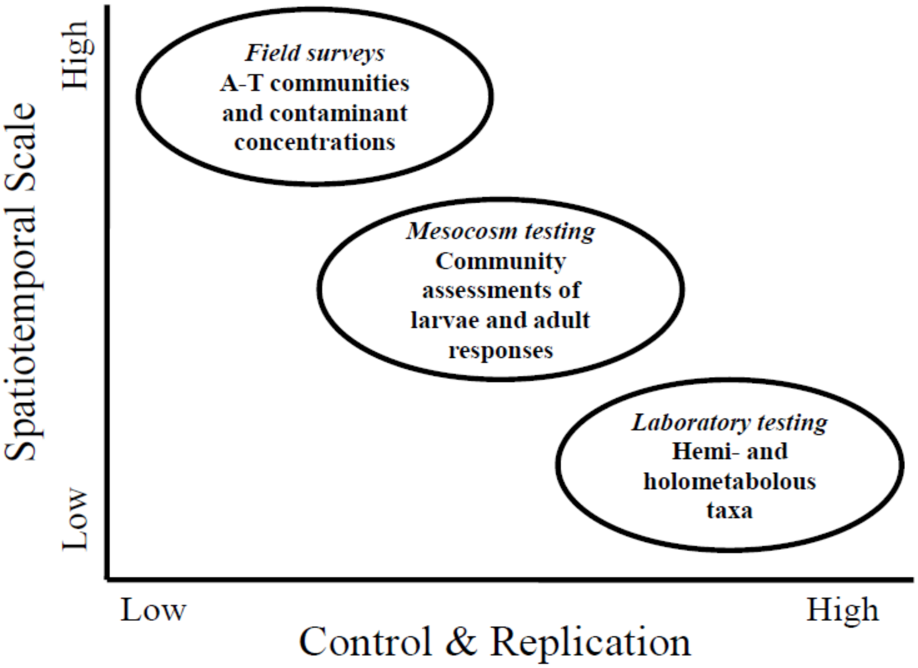

Figure 1.3 Modified from Buchwalter et al. (2017). Relationship of spatiotemporal scale, control, and replication of field surveys, mesocosm testing, and laboratory toxicity tests. This conceptual workflow can be applied for integrating aquatic insect emergence into develop aquatic life criteria.

Table 1.1 Summary of emergence assessment endpoints using aquatic mesocosm contaminant exposures to characterize aquatic-terrestrial outcomes

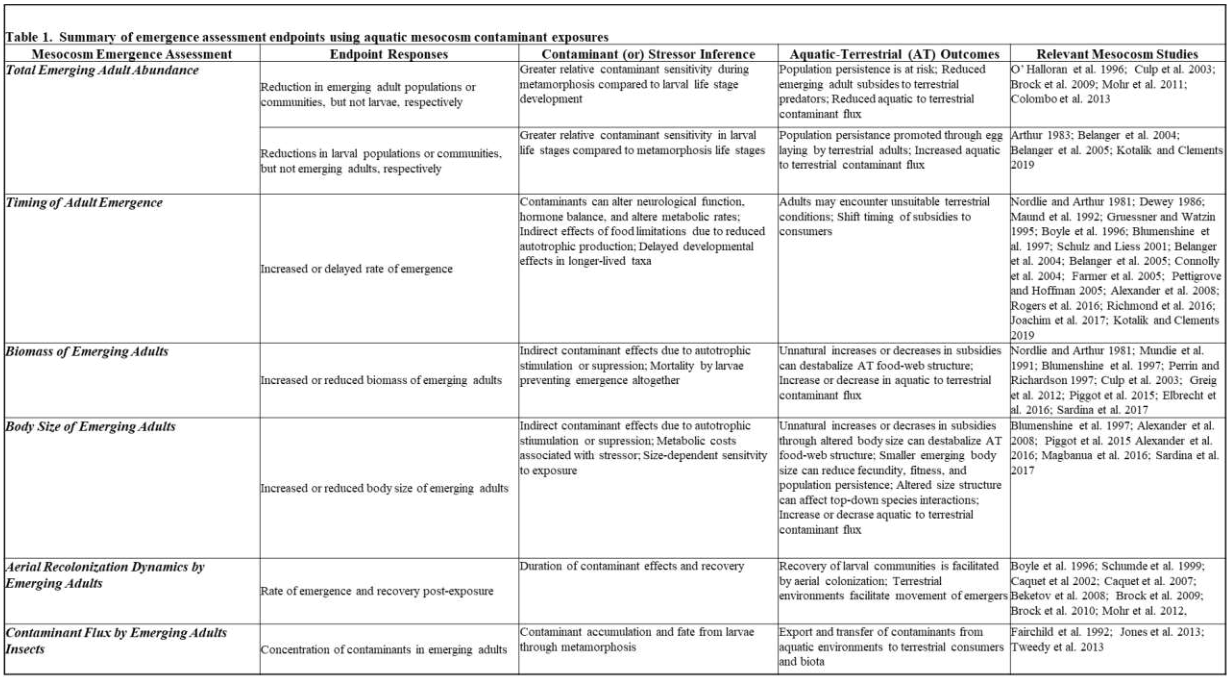

Table 1. Summary of emergence assessment endpoints using aquatic mesocosm contaminant exposures

Mesocosm Emergence Assessment	Endpoint Responses	Contaminant (or) Stressor Inference	Aquatic-Terrestrial (AT) Outcomes	Relevant Mesocosm Studies
Total Emerging Adult Abundance	Reduction in emerging adult populations or communities, but not larvae, respectively	Greater relative contaminant sensitivity during metamorphosis compared to larval life stage development	Population persistence is at risk; Reduced emerging adult subsides to terrestrial predators; Reduced aquatic to terrestrial contaminant flux	O' Halloran et al. 1996; Culp et al. 2003; Brock et al. 2009; Mohr et al. 2011; Colombo et al. 2013
	Reductions in larval populations or communities, but not emerging adults, respectively	Greater relative contaminant sensitivity in larval life stages compared to metamorphosis life stages	Population persistance promoted through egg laying by terrestrial adults; Increased aquatic to terrestrial contaminant flux	Arthur 1983; Belanger et al. 2004; Belanger et al. 2005; Kotalik and Clements 2019
Timing of Adult Emergence	Increased or delayed rate of emergence	Contaminants can alter neurological function, hormone balance, and altere metabolic rates; Indirect effects of food limitations due to reduced autotrophic production; Delayed developmental effects in longer-lived taxa	Adults may encounter unsuitable terrestrial conditions; Shift timing of subsidies to consumers	Nordlie and Arthur 1981; Dewey 1986; Maund et al. 1992; Gruessner and Watzin 1995; Boyle et al. 1996; Blumenshine et al. 1997; Schulz and Liess 2001; Belanger et al. 2004; Belanger et al. 2005; Connolly et al. 2004; Farmer et al. 2005; Pettigrove and Hoffman 2005; Alexander et al. 2008; Rogers et al. 2016; Richmond et al. 2016; Joachim et al. 2017; Kotalik and Clements 2019
Biomass of Emerging Adults	Increased or reduced biomass of emerging adults	Indirect contaminant effects due to autotrophic stimulation or supression; Mortality by larvae preventing emergence altogether	Unnatural increases or decreases in subsidies can destabalize AT food-web structure; Increase or decrease in aquatic to terrestrial contaminant flux	Nordlie and Arthur 1981; Mundie et al. 1991; Blumenshine et al. 1997; Perrin and Richardson 1997; Culp et al. 2003; Greig et al. 2012; Piggot et al. 2015; Elbrecht et al. 2016; Sardina et al. 2017
Body Size of Emerging Adults	Increased or reduced body size of emerging adults	Indirect contaminant effects due to autotrophic stiumulation or supression; Metabolic costs associated with stressor; Size-dependent sensitvity to exposure	Unnatural increases or decrases in subsidies through altered body size can destabalize AT food-web structure; Smaller emerging body size can reduce fecundity, fitness, and population persistence; Altered size structure can affect top-down species interactions; Increase or decrase aquatic to terrestrial contaminant flux	Blumenshine et al. 1997; Alexander et al. 2008; Piggot et al. 2015 Alexander et al. 2016; Magbanua et al. 2016; Sardina et al. 2017
Aerial Recolonization Dynamics by Emerging Adults	Rate of emergence and recovery post-exposure	Duration of contaminant effects and recovery	Recovery of larval communities is facilitated by aerial colonization; Terrestrial environments facilitate movement of emergers	Boyle et al. 1996; Schumde et al. 1999; Caquet et al 2002; Caquet et al. 2007; Beketov et al. 2008; Brock et al. 2009; Brock et al. 2010; Mohr et al. 2012,
Contaminant Flux by Emerging Adults Insects	Concentration of contaminants in emerging adults	Contaminant accumulation and fate from larvae through metamorphosis	Export and transfer of contaminants from aquatic environments to terrestrial consumers and biota	Fairchild et al. 1992; Jones et al. 2013; Tweedy et al. 2013

References

Arthur, J. W., J. A. Zischke, K. N. Allen, and R. O. Hermanutz. 1983. Effects of diazinon on macroinvertebrates and insect emergence in outdoor experimental channels. Aquatic Toxicology 4:283-302.

Baxter, C. V., K. D. Fausch, and W. Carl Saunders. 2005. Tangled webs: reciprocal flows of invertebrate prey link streams and riparian zones. Freshwater biology 50:201-220.

Beketov, M. A., R. B. Schäfer, A. Marwitz, A. Paschke, and M. Liess. 2008. Long-term stream invertebrate community alterations induced by the insecticide thiacloprid: effect concentrations and recovery dynamics. Science of the Total Environment 405:96-108.

Belanger, S. E., D. M. Lee, J. W. Bowling, and E. M. LeBlanc. 2004. Responses of periphyton and invertebrates to a tetradecyl-pentadecyl sulfate mixture in stream mesocosms. Environmental Toxicology and Chemistry 23:2202-2213.

Belanger, S. E., J. W. Bowling, E. M. LeBlanc, B. B. Price, R. R. Herzog, and E. Bozso. 2005. Comprehensive assessment of aquatic community responses to a new anionic surfactant, high-solubility alkyl sulfate. Ecotoxicology and Environmental Safety 62:75-92.

Belanger, S., J. Bowling, D. Lee, E. LeBlanc, K. Kerr, D. McAvoy, S. Christman, and D. Davidson. 2002. Integration of aquatic fate and ecological responses to linear alkyl benzene sulfonate (LAS) in model stream ecosystems. Ecotoxicology and Environmental Safety 52:150-171.

Bloesch, J., P. Bossard, H. Bührer, H. Bürgi, and U. Uehlinger. 1988. Can results from limnocorral experiments be transferred to in situ conditions? Hydrobiologia 159:297-308.

Bloomquist, J. R. 1996. Ion channels as targets for insecticides. Annual review of entomology 41:163-190.

Blumenshine, S., Y. Vadeboncoeur, D. Lodge, K. Cottingham, and S. Knight. 1997. Benthic-pelagic links: responses of benthos to water-column nutrient enrichment. Journal of the North American Benthological Society 16:466-479.

Boyle, T. P., J. F. Fairchild, P. S. Haverland, J. A. Lebo, and E. F. Robinson-Wilson. 1996. Ecological restructuring in experimental aquatic mesocosms due to the application of diflubenzuron. Environmental Toxicology and Chemistry 15:1806-1814.

Brock, T., I. Roessink, J. D. M. Belgers, F. Bransen, and S. J. Maund. 2009. Impact of a benzoyl urea insecticide on aquatic macroinvertebrates in ditch mesocosms with and without non-sprayed sections. Environmental Toxicology and Chemistry 28:2191-2205.

Buchwalter, D. B., W. H. Clements, and S. N. Luoma. 2017. Modernizing water quality criteria in the United States: A need to expand the definition of acceptable data. Environmental Toxicology and Chemistry 36:285-291.

Cadmus, P., H. Guasch, A. T. Herdrich, B. Bonet, G. Urrea, and W. H. Clements. 2018. Structural and functional responses of periphyton and macroinvertebrate communities to ferric Fe, Cu, and Zn in stream mesocosms. Environmental Toxicology and Chemistry 37:1320-1329.

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

Cairns Jr, J. 1994. Ecological toxicity testing: scale, complexity, and relevance. CRC Press.

Campbell, P., D. Arnold, T. Brock, N. Grandy, W. Heger, F. Heimbash, S. Maund, and M. Streloke. 1999. Guideline document. Higher-tier Aquatic Risk Assessment for Pesticides (HARAP), SETAC-Europe Publication, Brussels.

Campero, M., M. De Block, F. Ollevier, and R. Stoks. 2008. Metamorphosis offsets the link between larval stress, adult asymmetry and individual quality. Functional Ecology 22:271-277.

Caquet, T., L. Lagadic, G. Monod, J.-C. Lacaze, and A. Couté. 2001. Variability of physicochemical and biological parameters between replicated outdoor freshwater lentic mesocosms. Ecotoxicology 10:51-66.

Caquet, T., M. L. Hanson, M. Roucaute, D. W. Graham, and L. Lagadic. 2007. Influence of isolation on the recovery of pond mesocosms from the application of an insecticide. II. Benthic macroinvertebrate responses. Environmental Toxicology and Chemistry 26:1280.

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

CCME, A. 2007. Protocol for the Derivation of Water Quality Guidelines for the Protection of Aquatic Life. Winnipeg, Manitoba: Canadian Council of Ministers of the Environment.

Clements, W. H. 2000. Integrating effects of contaminants across levels of biological organization: an overview. Journal of Aquatic Ecosystem Stress and Recovery 7:113-116.

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

Clements, W. H., and P. M. Kiffney. 1994. Assessing contaminant effects at higher levels of biological organization. Environmental Toxicology and Chemistry 13:357-359.

Clements, W. H., and P. M. Kiffney. 1994. Integrated laboratory and field approach for assessing impacts of heavy metals at the Arkansas River, Colorado. Environmental Toxicology and Chemistry: An International Journal 13:397-404.

Clements, W. H., and P. M. Kiffney. 1995. The influence of elevation on benthic community responses to heavy metals in Rocky Mountain streams. Canadian Journal of Fisheries and Aquatic Sciences 52:1966-1977.

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

Clements, W. H., C. W. Hickey, and K. A. Kidd. 2012. How do aquatic communities respond to contaminants? It depends on the ecological context. Environmental Toxicology andChemistry 31:1932-1940.

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

Clements, W. H., J. H. Van Hassel, D. S. Cherry, and J. Cairns. 1989. Colonization, variability, and the use of substratum-filled trays for biomonitoring benthic communities. Hydrobiologia 173:45-53.

Colombo, V., S. Mohr, R. Berghahn, and V. J. Pettigrove. 2013. Structural changes in a macrozoobenthos assemblage after imidacloprid pulses in aquatic field-based microcosms. Archives of Environmental Contamination and Toxicology 65:683-692.

Connolly, N., M. Crossland, and R. Pearson. 2004. Effect of low dissolved oxygen on survival, emergence, and drift of tropical stream macroinvertebrates. Journal of the North American Benthological Society 23:251-270.

Culp, J. M., K. J. Cash, N. E. Glozier, and R. B. Brua. 2003. Effects of pulp mill effluent on benthic assemblages in mesocosms along the Saint John River, Canada. Environmental Toxicology and Chemistry 22:2916-2925.

Debecker, S., K. V. Dinh, and R. Stoks. 2017. Strong delayed interactive effects of metal exposure and warming: latitude-dependent synergisms persist across metamorphosis. Environmental Science & Technology 51:2409-2417.

Dewey, S. L. 1986. Effects of the herbicide atrazine on aquatic insect community structure and emergence. Ecology 67:148-162.

Eberhardt, L., and J. Thomas. 1991. Designing environmental field studies. Ecological monographs 61:53-73.

Echols, K. R., D. E. Tillitt, J. W. Nichols, A. L. Secord, and J. P. McCarty. 2004. Accumulation of PCB congeners in nestling tree swallows (Tachycineta bicolor) on the Hudson River, New York. Environmental Science & Technology 38:6240-6246.

Elbrecht, V., A. J. Beermann, G. Goessler, J. Neumann, R. Tollrian, R. Wagner, A. Wlecklik, J. J. Piggott, C. D. Matthaei, and F. Leese. 2016. Multiple-stressor effects on stream invertebrates: a mesocosm experiment manipulating nutrients, fine sediment and flow velocity. Freshwater biology 61:362-375.

European Union. 1997. Council Directive 97/57/EC of September 21, 1997; Establishing annex VI to Directive 91/414/EEC concerning the placing of plant protection products on the market. Official Journal of the European Communities L265:87–109.

European Union. 2002. Guidance document on aquatic toxicology in the context of the Directive 91/414EEC. Working Document of the European Commission Health & Consumer Protection Directorate-General. Brussels, Belgium

Farmer, D., I. R. Hill, and S. J. Maund. 1995. A comparison of the fate and effects of two pyrethroid insecticides (lambda-cyhalothrin and cypermethrin) in pond mesocosms. Ecotoxicology 4:219-244.

Giberson, D. J., and D. M. Rosenberg. 1994. Life histories of burrowing mayflies (Hexagenia limbata and H. rigida, Ephemeroptera: Ephemeridae) in a northern Canadian reservoir. Freshwater biology 32:501-518.

Grace, J. B. 2006. Structural equation modeling and natural systems. Cambridge University Press.

Graney, R. L. 1993. Aquatic mesocosm studies in ecological risk assessment. CRC Press.

Gratton, C., and M. J. Vander Zanden. 2009. Flux of aquatic insect productivity to land: comparison of lentic and lotic ecosystems. Ecology 90:2689-2699.

Greig, H. S., P. Kratina, P. L. Thompson, W. J. Palen, J. S. Richardson, and J. B. Shurin. 2012. Warming, eutrophication, and predator loss amplify subsidies between aquatic and terrestrial ecosystems. Global Change Biology 18:504-514.

Gruessner, B., and M. C. Watzin. 1996. Response of aquatic communities from a Vermont stream to environmentally realistic atrazine exposure in laboratory microcosms. Environmental Toxicology and Chemistry 15:410-419.

Hastings, A. 2012. Temporally varying resources amplify the importance of resource input in ecological populations. Biology letters:rsbl20120669.

Hickey, C. W., and L. A. Golding. 2002. Response of macroinvertebrates to copper and zinc in a stream mesocosm. Environmental Toxicology and Chemistry 21:1854-1863.

Hocking, M. D., and J. D. Reynolds. 2011. Impacts of salmon on riparian plant diversity. Science 331:1609-1612.

Huxel, G. R., and K. McCann. 1998. Food web stability: the influence of trophic flows across habitats. The American Naturalist 152:460-469.

Hynes, H. 1975. The stream and its valley: Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen 19:1-15.

Iwasaki, Y., T. S. Schmidt, and W. H. Clements. 2018. Quantifying differences in responses of aquatic insects to trace metal exposure in field studies and short-term stream mesocosm experiments. Environmental Science & Technology 52:4378-4384.

Joachim, S., H. Roussel, J. M. Bonzom, E. Thybaud, C. A. Mebane, P. Van den Brink, and L. Gauthier. 2017. A long-term copper exposure in a freshwater ecosystem using lotic mesocosms: Invertebrate community responses. Environmental Toxicology and Chemistry 36:2698-2714.

Jones, T. A., M. M. Chumchal, R. W. Drenner, G. N. Timmins, and W. H. Nowlin. 2013. Bottom-up nutrient and top-down fish impacts on insect-mediated mercury flux from aquatic ecosystems. Environmental Toxicology and Chemistry 32:612-618.

Kato, C., T. Iwata, S. Nakano, and D. Kishi. 2003. Dynamics of aquatic insect flux affects distribution of riparian web-building spiders. Oikos 103:113-120.

Kefford, B. J., D. Nugegoda, L. Zalizniak, E. J. Fields, and K. L. Hassell. 2007. The salinity tolerance of freshwater macroinvertebrate eggs and hatchlings in comparison to their older life-stages: a diversity of responses. Aquatic Ecology 41:335-348.

Kiffney, P. M., and W. H. Clements. 1994. Effects of heavy metals on a macroinvertebrate assemblage from a Rocky Mountain stream in experimental microcosms. Journal of the North American Benthological Society: 511-523.

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

Kim, K., D. Funk, and D. Buchwalter. 2012. Dietary (periphyton) and aqueous Zn bioaccumulation dynamics in the mayfly Centroptilum triangulifer. Ecotoxicology 21:2288-2296.

Kotalik, C. J., W. H. Clements, and P. Cadmus. 2017. Effects of magnesium chloride road deicer on montane stream benthic communities. Hydrobiologia 799:193-202.

Kotalik, C. J., P. Cadmus, and W. H. Clements. 2019. Indirect effects of iron oxide on stream benthic communities: capturing ecological complexity with controlled mesocosm experiments. Environmental Science & Technology.

Kotalik, C. J., and W. H. Clements. 2019. Stream mesocosm experiments show significant differences in sensitivity of larval and emerging adults to metals. Environmental Science & Technology.

Kraus, J. M., T. S. Schmidt, D. M. Walters, R. B. Wanty, R. E. Zuellig, and R. E. Wolf. 2014a. Cross-ecosystem impacts of stream pollution reduce resource and contaminant flux to riparian food webs. Ecological Applications 24:235-243.

Kraus, J. M., D. M. Walters, J. S. Wesner, C. A. Stricker, T. S. Schmidt, and R. E. Zuellig. 2014b. Metamorphosis alters contaminants and chemical tracers in insects: implications for food webs. Environmental Science & Technology 48:10957-10965.

Kraus, J. M., J. F. Pomeranz, A. S. Todd, D. M. Walters, T. S. Schmidt, and R. B. Wanty. 2016. Aquatic pollution increases use of terrestrial prey subsidies by stream fish. Journal of Applied Ecology 53:44-53.

Lee, S. S., A. M. Paspalof, D. D. Snow, E. K. Richmond, E. J. Rosi-Marshall, and J. J. Kelly. 2016. Occurrence and potential biological effects of amphetamine on stream communities. Environmental Science & Technology 50:9727-9735.

Liber, K., K. L. Schmude, and T. D. Corry. 1996. Effects of the insect growth regulator diflubenzuron on insect emergence within littoral enclosures. Environmental entomology 25:17-24.

Liess, M., and P. C. Von Der Ohe. 2005. Analyzing effects of pesticides on invertebrate communities in streams. Environmental Toxicology and Chemistry 24:954-965.

Macneale, K. H., P. M. Kiffney, and N. L. Scholz. 2010. Pesticides, aquatic food webs, and the conservation of Pacific salmon. Frontiers in Ecology and the Environment 8:475-482.

Magbanua, F. S., C. R. Townsend, K. J. Hageman, J. J. Piggott, and C. D. Matthaei. 2016. Individual and combined effects of fine sediment and glyphosate herbicide on invertebrate drift and insect emergence: a stream mesocosm experiment. Freshwater Science 35:139-151.

Matthews, R. A., W. G. Landis, and G. B. Matthews. 1996. The community conditioning hypothesis and its application to environmental toxicology. Environmental Toxicology and Chemistry 15:597-603.

Maund, S., A. Peither, E. Taylor, I. Jüttner, R. Beyerle-Pfnür, J. Lay, and D. Pascoe. 1992. Toxicity of lindane to freshwater insect larvae in compartments of an experimental pond. Ecotoxicology and Environmental Safety 23:76-88.

Mebane, C. A., T. S. Schmidt, and L. S. Balistrieri. 2017. Larval aquatic insect responses to cadmium and zinc in experimental streams. Environmental Toxicology and Chemistry 36:749-762.

Menon, J., and R. Rozman. 2007. Oxidative stress, tissue remodeling and regression during amphibian metamorphosis. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 145:625-631.

Merritt, R. W., and K. W. Cummins. 1996. An introduction to the aquatic insects of North America. Kendall Hunt.

Mohr, S., R. Berghahn, R. Schmiediche, V. Hübner, S. Loth, M. Feibicke, W. Mailahn, and J. Wogram. 2012. Macroinvertebrate community response to repeated short-term pulses of the insecticide imidacloprid. Aquatic Toxicology 110:25-36.

Mundie, J., K. Simpson, and C. Perrin. 1991. Responses of stream periphyton and benthic insects to increases in dissolved inorganic phosphorus in a mesocosm. Canadian Journal of Fisheries and Aquatic Sciences 48:2061-2072.

Murakami, M., and S. Nakano. 2002. Indirect effect of aquatic insect emergence on a terrestrial insect population through by birds predation. Ecology Letters 5:333-337.

Nakano, S., and M. Murakami. 2001. Reciprocal subsidies: dynamic interdependence between terrestrial and aquatic food webs. Proceedings of the National Academy of Sciences 98:166-170.

Nation Sr, J. L. 2015. Insect physiology and biochemistry. CRC press.

Nebeker, A. V. 1971. Effect of high winter water temperatures on adult emergence of aquatic insects. Water Research 5:777-783.

Nebeker, A. V. 1972. Effect of low oxygen concentration on survival and emergence of aquatic insects. Transactions of the American Fisheries Society 101:675-679.

Nordlie, K. J., and J. W. Arthur. 1981. Effect of elevated water temperature on insect emergence in outdoor experimental channels. Environmental Pollution Series A, Ecological and Biological 25:53-65.

O'Halloran, S., K. Liber, K. Schmude, and T. Corry. 1996. Effects of diflubenzuron on benthic macroinvertebrates in littoral enclosures. Archives of Environmental Contamination and Toxicology 30:444-451.

Paetzold, A., M. Smith, P. H. Warren, and L. Maltby. 2011. Environmental impact propagated by cross-system subsidy: Chronic stream pollution controls riparian spider populations. Ecology 92:1711-1716.

Palmquist, K., P. Jepson, and J. Jenkins. 2008. Impact of aquatic insect life stage and emergence strategy on sensitivity to esfenvalerate exposure. Environmental Toxicology and Chemistry 27:1728.

Peckarsky, B. L., B. W. Taylor, A. R. McIntosh, M. A. McPeek, and D. A. Lytle. 2001. Variation in mayfly size at metamorphosis as a developmental response to risk of predation. Ecology 82:740-757.

Perceval, O., T. Caquet, L. Lagadic, A. Bassères, D. Azam, G. Lacroix, and V. Poulsen 2009. Mesocosms: Their value as tools for managing the quality of aquatic environments. Pages 14-16. Recap of the Ecotoxicology Symposium, October.

Perrin, C. J., and J. S. Richardson. 1997. N and P limitation of benthos abundance in the Nechako River, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 54:2574-2583.

Pettigrove, V., and A. Hoffmann. 2005. Effects of long-chain hydrocarbon-polluted sediment on freshwater macroinvertebrates. Environmental Toxicology and Chemistry 24:2500-2508.

Piggott, J. J., C. R. Townsend, and C. D. Matthaei. 2015. Climate warming and agricultural stressors interact to determine stream macroinvertebrate community dynamics. Global Change Biology 21:1887-1906.

Pond, G. J., M. E. Passmore, F. A. Borsuk, L. Reynolds, and C. J. Rose. 2008. Downstream effects of mountaintop coal mining: comparing biological conditions using family-and genus-level macroinvertebrate bioassessment tools. Journal of the North American Benthological Society 27:717-737.

Pontasch, K. W., and J. Cairns. 1989. Establishing and maintaining laboratory-based microcosms of riffle insect communities: their potential for multispecies toxicity tests. Hydrobiologia 175:49-60.

Pontasch, K. W., and J. Cairns. 1991. Multispecies toxicity tests using indigenous organisms: predicting the effects of complex effluents in streams. Archives of Environmental Contamination and Toxicology 20:103-112.

Postma, J., M. Buckert-de Jong, N. Staats, and C. Davids. 1994. Chronic toxicity of cadmium to Chironomus riparius (Diptera: Chironomidae) at different food levels. Archives of Environmental Contamination and Toxicology 26:143-148.

Poteat, M. D., and D. B. Buchwalter. 2014. Phylogeny and size differentially influence dissolved Cd and Zn bioaccumulation parameters among closely related aquatic insects. Environmental Science & Technology 48:5274-5281.

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

Rasmussen, N. L., and V. H. Rudolf. 2015. Phenological synchronization drives demographic rates of populations. Ecology 96:1754-1760.

Richardson, J. S., and C. J. Perrin. 1994. Effects of the bacterial insecticide Bacillus thuringiensis var. kurstaki (Btk) on a stream benthic community. Canadian Journal of Fisheries and Aquatic Sciences 51:1037-1045.

Richmond, E. K., E. J. Rosi-Marshall, S. S. Lee, R. M. Thompson, and M. R. Grace. 2016. Antidepressants in stream ecosystems: influence of selective serotonin reuptake inhibitors (SSRIs) on algal production and insect emergence. Freshwater Science 35:845-855.

Rogers, H. A., T. S. Schmidt, B. L. Dabney, M. L. Hladik, B. J. Mahler, and P. C. Van Metre. 2016. Bifenthrin Causes Trophic Cascade and Altered Insect Emergence in Mesocosms: Implications for Small Streams. Environmental Science & Technology 50:11974-11983.

Rosseel, Y. 2012. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). Journal of statistical software 48:1-36.

Sabo, J., and M. Power. 2002a. River–watershed exchange: effects of riverine subsidies on riparian lizards and their terrestrial prey. Ecology 83:1860-1869.

Sabo, J. L., and M. E. Power. 2002b. Numerical response of lizards to aquatic insects and short-term consequences for terrestrial prey. Ecology 83:3023-3036.

Sardiña, P., J. Beardall, J. Beringer, M. Grace, and R. M. Thompson. 2017. Consequences of altered temperature regimes for emerging freshwater invertebrates. Aquatic Sciences 79:265-276.

Schmidt, T. S., J. M. Kraus, D. M. Walters, and R. B. Wanty. 2013. Emergence flux declines disproportionately to larval density along a stream metals gradient. Environmental Science & Technology 47:8784-8792.

Schmidt, T. S., H. A. Rogers, J. L. Miller, C. A. Mebane, and L. S. Balistrieri. 2018. Understanding the captivity effect on invertebrate communities transplanted into an experimental stream laboratory. Environmental Toxicology and Chemistry.

Schmude, K. L., K. Liber, T. D. Corry, and F. S. Stay. 1999. Effects of 4-nonylphenol on benthic macroinvertebrates and insect emergence in littoral enclosures. Environmental Toxicology and Chemistry 18:386-393.

Schriever, C. A., and M. Liess. 2007. Mapping ecological risk of agricultural pesticide runoff. Science of the Total Environment 384:264-279.

Sheldon, A. L. 1999. Emergence patterns of large stoneflies (Plecoptera: Pteronarcys, Calineuria, Hesperoperla) in a Montana river. The Great Basin Naturalist:169-174.

Sibley, P. K., G. T. Ankley, A. M. Cotter, and E. N. Leonard. 1996. Predicting chronic toxicity of sediments spiked with zinc: An evaluation of the acid-volatile sulfide model using a life-cycle test with the midge Chironomus tentans. Environmental Toxicology and Chemistry 15:2102-2112.

Spiller, D. A., J. Piovia-Scott, A. N. Wright, L. H. Yang, G. Takimoto, T. W. Schoener, and T. Iwata. 2010. Marine subsidies have multiple effects on coastal food webs. Ecology 91:1424-1434.

Statzner, B., and V. H. Resh. 1993. Multiple-site and-year analyses of stream insect emergence: a test of ecological theory. Oecologia 96:65-79.

Stephan, C.E. 1985. Are the 'Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Life and its Uses’ based on sound judgments? Pages 515-526 in R. C. Bahner, and D. J. Hansen, editors. Aquatic Toxicology and Hazard

Assessment: Seventh Symposium: ASTM STP 891. American Society for Testing and Materials (ASTM), Philadelphia, PA.

Stephan, C. E., D. I. Mount, D. J. Hansen, J. Gentile, G. A. Chapman, and W. A. Brungs. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. US Environmental Protection Agency Washington, DC.

Takimoto, G., T. Iwata, and M. Murakami. 2002. Seasonal subsidy stabilizes food web dynamics: balance in a heterogeneous landscape. Ecological Research 17:433-439.

Touart, L. W. 1988. Hazard Evaluation Division Technical Guidance Document--aquatic mesocosm test to support pesticide registrations.

Tweedy, B. N., R. W. Drenner, M. M. Chumchal, and J. H. Kennedy. 2013. Effects of fish on emergent insect-mediated flux of methyl mercury across a gradient of contamination. Environmental Science & Technology 47:1614-1619.

U.S. EPA. 1993. Questions and answers: Improvements to EPA's program to prevent adverse environmental effects of pesticides. H7506-C. Washington, DC.

U.S. EPA. 2016. Aquatic life ambient water quality criterion for selenium in freshwater 2016. 822-R-16-006, U.S. Environmental Protection Agency.

Van den Brink, P.J. 2006. Response to recent criticism on aquatic semifield experiments: Opportunities for new developments in ecological risk assessment of pesticides. Integrated Environmental Assessment and Management. 2(3): 202-203.

Van den Brink, P. J. 2013. Assessing aquatic population and community-level risks of pesticides. Environmental Toxicology and Chemistry 32:972-973.

Van den Brink, P. J., and C. J. Ter Braak. 1998. Multivariate analysis of stress in experimental ecosystems by principal response curves and similarity analysis. Aquatic Ecology 32:163-178.

Van den Brink, P. J., and C. J. Ter Braak. 1999. Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress. Environmental Toxicology and Chemistry 18:138-148.

Van den Brink, P. J., R. P. Van Wijngaarden, W. G. Lucassen, T. C. Brock, and P. Leeuwangh. 1996. Effects of the insecticide Dursban® 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: II. Invertebrate community responses and recovery. Environmental Toxicology and Chemistry 15:1143-1153.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130-137.

Versteeg, D.J., S.E. Belanger, and G.J. Carr. 1999. Understanding single-species and model ecosystem sensitivity: Data-based comparison. Environmental Toxicology and Chemistry. 18(6): 1329-1346

Walters, D. M., K. M. Fritz, and R. R. Otter. 2008. The dark side of subsidies: adult stream insects export organic contaminants to riparian predators. Ecological Applications 18:1835-1841.

Walters, D. M., J. S. Wesner, R. E. Zuellig, D. A. Kowalski, and M. C. Kondratieff. 2018. Holy flux: spatial and temporal variation in massive pulses of emerging insect biomass from western US rivers. Ecology 99:238-240.

Wesner, J. S., J. M. Kraus, T. S. Schmidt, D. M. Walters, and W. H. Clements. 2014. Metamorphosis enhances the effects of metal exposure on the mayfly, Centroptilum triangulifer. Environmental Science & Technology 48:10415-10422.

Wieczorek, M. V., D. Kötter, R. Gergs, and R. Schulz. 2015. Using stable isotope analysis in stream mesocosms to study potential effects of environmental chemicals on aquatic-terrestrial subsidies. Environmental Science and Pollution Research 22:12892-12901.

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

Yang, L. H., J. L. Bastow, K. O. Spence, and A. N. Wright. 2008. What can we learn from resource pulses. Ecology 89:621-634.

Yang, L. H., K. F. Edwards, J. E. Byrnes, J. L. Bastow, A. N. Wright, and K. O. Spence. 2010. A meta-analysis of resource pulse–consumer interactions. Ecological monographs 80:125-151.

Zhao, M., C. Wang, K. K. Liu, L. Sun, L. Li, and W. Liu. 2009. Enantioselectivity in chronic toxicology and accumulation of the synthetic pyrethroid insecticide bifenthrin in Daphnia magna. Environmental Toxicology and Chemistry 28:1475-1479.

Zwick, P. 1990. Emergence, maturation and upstream oviposition flights of Plecoptera from the Breitenbach, with notes on the adult phase as a possible control of stream insect populations. Hydrobiologia 194:207-223

CHAPTER 2

STREAM MESOCOSM EXPERIMENTS SHOW SIGNIFICANT DIFFERENCES IN SENSITIVITY OF LARVAL AND EMERGING ADULTS TO METALS

Introduction

Evaluations of aquatic insect responses to contaminants predominately use immature life stages (hereafter referred to as "larvae") to characterize taxa sensitivity, but the morphological transition from larvae to terrestrial adults has received considerably less attention. Most aquatic insects metamorphose to winged-terrestrial adults to reproduce and complete their life cycle (Merritt et al. 2004). Emerging adults provide crucial prey subsidies to linked riparian consumers such as birds, bats, and spiders (Sabo and Power 2002, Baxter et al. 2005), and they supply detrital material and nutrients to terrestrial ecosystems (Nakano and Murakami 2001, Gratton and Vander Zanden 2009). Adults are also a vector for contaminant transfer from aquatic to terrestrial ecosystems (Echols et al. 2004, Walters et al. 2008, Paetzold et al. 2011, Kraus et al. 2014a). Similar to the morphological and physiological development during early life stages of fish and other aquatic vertebrates, invertebrate metamorphosis is a biologically stressful process because of the greater energy requirements and cell differentiation that occur during tissue reorganization (Campero et al. 2008). Therefore, assessment of insect larval responses alone may not adequately characterize the effects of contaminants on adult emergence (Pascoe et al. 1989, Schmidt et al. 2013, Wesner et al. 2014, Soucek and Dickinson 2015).

Field and laboratory studies of aquatic insect emergence have shown greater sensitivity of emerging adults to trace metals compared to larvae (Schmidt et al. 2013, Wesner et al. 2014). Energy used by larvae to detoxify metals at sublethal concentrations may decrease the metabolic resources available to successfully complete metamorphosis (Hare 1992, Liber et al. 1996, Sibley et al. 1997, Rainbow 2007). Metal exposure can reduce insect emergence and affect linked terrestrial ecosystems that are dependent on aquatic subsidies (Kraus et al. 2014a). However, metal flux to riparian systems by emerging adults is influenced by the strength of the resource linkage (e.g., number of emerging larvae), and the decrease of metal body burden from larvae to adults after metamorphosis (Timmermans and Walker 1989, Kim et al. 2012; Kraus et al. 2014b, Wesner et al. 2017).

Quantifying the relationship of metal exposure and emergence propensity in the field is complicated by the seasonality of emergence, colonization, and the inability to control for extraneous variables (Buchwalter et al. 2017). Traditional testing methodology employed in single-species toxicity tests is also limited in evaluating adult insect emergence. Because organisms commonly used in laboratory toxicity tests (e.g., *Daphnia* spp., *Ceriodaphnia dubia*) do not metamorphose into a terrestrial adult life stage, there is no equivalent endpoint in these studies. Full-life cycle tests with the parthenogenetic mayfly *Neocloeon triangulifer* have shown differences in sensitivity to metals among life stages through metamorphosis, demonstrating the importance of evaluating adult emergence (Wesner et al. 2014, Soucek and Dickinson 2015). Community-level mesocosm experiments provide an opportunity to quantify the relationship between metal exposure and emergence because a diversity of indigenous aquatic insect taxa are exposed through multiple instar sizes and life stages (Clements 2004).

To test the effects of metal exposure on larvae and emerging adults, we conducted two stream mesocosm experiments. We compared the life stage responses of a developmentally mature, late-instar mayfly (Ephemeroptera: Heptageniidae) to the responses of three dominant taxa obtained from a naturally colonized stream benthic community. Both experiments exposed organisms to environmentally relevant concentration gradients of dissolved Cu and Zn. The primary goals of this study were to test the following hypotheses: 1) adult insect emergence is correlated with larval responses across a metal-exposure gradient; 2) responses to metals differbetween larval and adult life stages; and 3) exposure to sublethal concentrations of metals alters timing of emergence and sex ratios in adult aquatic insects.

Materials and Methods

Mesocosm Experiments

Two mesocosm experiments were conducted at the Stream Research Laboratory (SRL) located at Colorado State University Foothills Campus, Fort Collins, Colorado, USA. One mesocosm experiment used late-instar, developmentally mature *Rhithrogena robusta* (Heptageniidae: Ephemeroptera) collected from Joe Wright Creek (~2,500 m ASL), Colorado, an undisturbed 2nd-order stream in the Roosevelt National Forest. Individuals were collected from the underside of cobble substrate using a flat-angled, fine-bristled paint brush to minimize organism handling stress. We selected *R. robusta* that were developmentally mature based on wing-pad development, as well as other morphological features that indicate maturation (e.g., size, color, pigmentation). Individuals were placed in aerated insulated coolers filled with site water and transported to the SRL. Four un-colonized rock trays were placed in each mesocosm to provide habitat, and 32 or 34 individuals were assigned to each replicate mesocosms. To reduce potential density dependent effects, we used relatively low densities of *R. robusta* compared to the densities of benthic macroinvertebrates typically exposed in the mesocosms.

The second mesocosm experiment used naturally colonized stream benthic communities collected from the Arkansas River (AR), Colorado, a high-elevation (~3000 m ASL) 4th-order mountain stream. AR was historically contaminated by acid mine drainage (including Cu and Zn), but remediation in the last 25 years has significantly improved water quality and metal concentrations rarely exceed water quality criteria (Clements et al. 2010). Communities were collected from the AR using a technique described by Clements et al. (1988). Trays (10 x 10 x 6 cm) filled with pebble and small cobble were deployed for 35 d and colonized with a diverse benthic assemblage that is similar in abundance and species richness observed on naturally occurring substrate (Clements 2004). The trays were colonized by a range of life stages of aquatic insects that included early to mid-instar taxa, as well as mature taxa nearing completion of their life cycle as emerged adults. Five trays were randomly collected from the colonization racks, placed in insulated coolers, transported to the SRL, and assigned treatments.

Dilution water in the SRL is supplied from a deep mesotrophic reservoir. Water quality in the SRL is representative of cold water Rocky Mountain streams with cool temperatures (10-15 ºC), low hardness (30-35 mg/L $CaCO_3$), alkalinity (25-30 mg/L $CaCO_3$), and dissolved organic carbon (2.5-3.0 mg/L), circumneutral pH (7.0-7.8), and low conductivity (60-90 µS/cm) (Clements 2004, Clements et al. 2013). Current in the 18 mesocosms (20 L) was provided by submersible water pumps (EcoPlus®) at a rate of 2233 L per h (i.e., 1.9 hydraulic circulations per min). Each mesocosm was covered with insect netting (220 micron mesh size) to contain emerged adults. To ensure that most emerging insects were removed, daily adult collection was conducted between 17:00 and 19:00 for both experiments. The flow-through mesocosms (76 x 46 x 12 cm) received water at 1.0 L/min, and metal salt solutions were pumped to mesocosms using peristaltic pumps at a rate of 10 mL/min. To prevent emigration by drifting invertebrates, standpipes were covered with fine mesh.

Rhithrogena robusta were exposed to a concentration gradient of Cu and Zn at 1:10, with 3 replicate mesocosms (6 total treatments including controls) per treatment ranging from 10 to 100 µg/L Cu, and 100 to 1000 µg/L Zn. The AR community was exposed to a gradient of Cu and

Zn at approximately 1:15, with 3 replicate mesocosms per treatment ranging from 10 to 100 µg/L Cu, and 15 to 1500 µg/L Zn. These mass ratios were chosen to represent the Cu and Zn mixtures observed in Colorado mountain streams affected by AMD (Schmidt et al. 2010). To confirm dissolved Cu and Zn concentrations, water samples (15 mL) were taken from each mesocosm during the exposure duration (3x for *R. robusta*; 4x for AR community). Water samples were filtered through a 0.45-µm filter, acidified to a pH < 2.0 using analytical grade nitric acid, and analyzed using flame or furnace atomic absorption spectrophotometry. Cumulative criterion units (CCU), defined as the ratio of the measured dissolved metal concentrations to the United States Environmental Protection Agency (U.S. EPA) hardness-adjusted criterion value and summed for each metal, were used to quantify the metal mixture of Cu and Zn as one value. We chose to use hardness-adjusted criteria because previous experiments have found that the water chemistry of the dilution water in the SRL is highly consistent. In addition, the U.S. EPA biotic ligand model (BLM) and hardness-based criteria generate very similar values (Clements et al. 2013. Because the BLM requires additional chemical analyses (e.g., DOC) that were unavailable for these experiments, we chose to use the hardness-based criteria to benchmark metal mixture concentrations.

Rhithrogena robusta were exposed for 10 d, and surviving larvae that did not emerge were preserved at the end of the experiment. The AR community exposure was 14 d. At the end of the experiment, colonization trays were rinsed through a 350-µm sieve, with organisms and detrital material preserved in 80% ethanol. Benthic community samples were sorted in the laboratory. All larvae were enumerated and identified to the lowest taxonomic resolution possible, typically to genus, except for chironomids that were identified to subfamily or tribe. Emerged adult insects were enumerated and identified to family or subfamily, and sexed. For biomass estimates, larval and adult taxa were dried at 60 ºC for approximately 72 h and then weighed to the nearest 0.00001 g. To test for differences in emerging-adult biomass over time, daily samples were combined for days 1-4, 5-8, 9-13. Daily adult biomass estimates were not possible due to very low biomass in the high-concentration treatments near the end of the experiment. Larval biomass samples were sorted from the benthic community samples that were exposed for 14 d.

Statistical Analyses

All statistical analyses were conducted using R statistical computing (R Core Team v3.5.1). We used a linear model ("LM" function; package 'car') (Fox et al. 2012) to test the responses of *R. robusta* larvae and emerging adults to metals. A linear model was also fit to *Rhithrogena* spp. larvae abundance in the AR community to compare to *R. robusta* larval abundance results. For the AR community experiment, linear models were used to test concentration-response relationships between CCU and life stage abundances, biomass, and adult sex ratios of the three dominant taxa that were present at sufficient densities (i.e., Baetidae, Chironomidae, and Simuliidae). To determine if the concentration-response relationships of larvae and adults were statistically different, an ANCOVA interaction model was used ("LM" function), with CCU as the continuous predictor and life stage as the categorical predictor of either abundance or biomass. Statistical significance of the interaction term indicates that the slopes of larvae and adults are different from one another (i.e., the effect of CCU is dependent on life stage).

To test for differences in the timing of adult emergence abundance and biomass, CCU treatments were categorized (6 treatments, 3 replicates each) to support single-factor repeated measures ANOVA analyses. For emergence data, daily observations were successively combined to estimate cumulative emergence, whereby daily emergence abundance was additively summed over the duration of the experiment. We summed the daily results because we wanted to test if the mean difference in the number of emerging individuals over the entire exposure duration differed among treatments. Adult biomass data were analyzed in the three time groups as previously described. A single-factor repeated measures ANOVA (function "lmer," package 'lme4"; Bates et al. 2014) was fit to the cumulative emergence data to test if the effects of metals on cumulative emergence abundance and biomass varied over the exposure duration. We chose a covariance structure of compound symmetry, which assumes constant correlation between observations on the same sample replicate (i.e., cumulative observations are dependent on the previous observation). *Post hoc* comparisons with "estimated marginal means of linear trends" function from package 'emmeans' (Lenth 2018) were used to compare treatment means of emergence abundance and biomass to control values with Dunnett's Test. We chose to not conduct date-by-date treatment contrasts because we wanted to test for differences in average cumulative emergence abundance over the 14 d exposure. All data were log transformed to meet the assumptions of parametric statistics and to improve fit for all statistical models (i.e., linear models, ANCOVA, single-factor repeated measures ANOVA). Statistical significance for all models was $p < 0.10$. The *a priori* selection of a slightly higher p-value was motivated by the high within-treatment variation observed with naturally colonized communities, relatively low number of replicates (n = 3), and the goal of reducing the risk of Type II errors (i.e., incorrectly determine that there was no effect of metals). Standardized effect sizes were estimated by partial eta-squared (Cohen 1973) values with 90% confidence intervals for all linear and ANCOVA models.

Results

Measured concentrations of Cu and Zn were variable (**Table S2.1**), but generally approximated target values (**Table 2.1**). Measured treatment concentrations of Cu and Zn for the *R. robusta* experiment ranged from 14-119 µg/L Cu and 48-1068 µg/L Zn (**Table 2.1**). In the AR community experiment, concentrations ranged from 10-86 µg/L Cu and 86-1566 µg/L Zn. Water hardness and alkalinity were approximately 30 mg/L $CaCO_3$ (s.d. = 0.52) and 29 mg/L $CaCO_3$ (s.d. = 0.89), respectively. The U.S. EPA acute hardness-adjusted criterion values for Cu (U.S. EPA, 1985) was 4.3 µg/L and for Zn (U.S. EPA, 2002) 42.2 µg/L. CCU calculations among all mesocosm observations for the *R. robusta* experiment ranged from 3.6-52.7 CCU, and for the AR community ranged from 3.4-54.5 CCU (**Table 2.1**). Temperature, pH, and conductivity were consistent among treatments and between experiments (**Table S2.1**).

Emergence (**Figure 2.1**) and larval survival (**Figure S2.1**) of *R. robusta* did not significantly (p-value < 0.10) decrease in any treatment (**Table S2.2**). For two suspected outliers with relatively low adult emergence and high larval mortality compared to other treatments, we ran the linear models with and without these points and identified no significant relationships; therefore, we did not remove these data from the analysis. The total number of individuals emerging was greater than 80% for most treatments, and approximately 75% of individuals that did not emerge over the 10 d exposure were alive at the end of the experiment. In contrast, *Rhithrogena* spp. from the AR community experiment were not developmentally mature to the point of emergence, but *Rhithrogena* spp. larvae were significantly reduced (**Figure S2.1**).

Although metal exposure had no effect on total emergence of *R. robusta*, the timing of emergence over the 10 d experiment varied among metal treatments (**Figure 2.1; Table S2.3**). Average cumulative emergence greater than controls was interpreted as emergence stimulation, while average cumulative emergence less than controls was considered a delay if a significant decrease in treatment abundance was not detected among treatments. Cumulative emergence observed below control values but with significant decreases in overall abundance is the result of larval mortality preventing adult emergence. Because we observed no treatment effect on total emergence of *R. robusta*, timing of emergence was delayed among metal treatments. Statistically significant *post hoc* relationships (compared to controls) were observed for the 4- and 13-CCU treatments.

The AR benthic community was highly diverse, with control mesocosms averaging (±s.e.) 28.3 (±0.8) taxa and 864 (±135) individuals. Number of taxa and abundance of emerging adults in controls averaged 7.3 (±0.3) and 328 (±29), respectively. Three dominant taxa groups (Baetidae, Chironomidae and Simuliidae) were present at sufficiently high larval and adult densities in controls for statistically valid comparisons. Abundance of Baetidae larvae and emerging adults were negatively ($p < 0.0001$) correlated with exposure concentration (**Figure 2.2; Table S2.2**). Although abundance of larval Chironomidae was also negatively correlated with metal exposure ($p < 0.0001$), the number of emerging adults was not. For both Baetidae and Chironomidae, the ANCOVA interaction term was highly significant (p-value < 0.0001), indicating that the response to metals was dependent on life stage (**Table S2.4**). For both groups, effects of metals were greater on larval survival than emergence. In contrast to baetids and chironomids, there were no effects of metals on larval survival or emergence of Simuliidae.

In addition to measuring effects of metals on abundance of larval and adult aquatic insects, we assessed effects on biomass, which is a more direct measure of subsidy export to riparian ecosystems. Effects of metals on the biomass of mayflies, chironomids and blackflies were generally similar to those observed for abundance (**Figure 2.3; Table S2.2**). Across all metal treatments, biomass of adult mayflies, chironomids and simuliids was greater than larval biomass. Ephemeroptera, Chironomidae, and community biomass all displayed significant ANCOVA interaction terms, with greater effects of metals on larval biomass compared to biomass of emerging adults (**Table S2.4)**. The timing of adult biomass varied among treatments and taxonomic groups (**Figure S2.2; Table S2.3**). Biomass of adult baetid mayflies was significantly reduced at the 2 highest treatments compared to control observations. In contrast, biomass of adult chironomids increased over time, with no differences among treatments over the three time periods. Maximum biomass of Simuliidae was generally achieved by day 8 for all treatments, but pairwise comparisons showed no treatment differences compared to controls. Biomass of the AR community was significantly different overall, with the 24-CCU treatment significantly different from controls.

Cumulative emergence of the 3 dominant groups was altered during the 14 d exposure (**Figure 2.4; Table S2.3**). Relative to controls, abundance of emerging baetids was significantly stimulated at 7 CCU, but decreased in the highest treatments. Abundance of emerging chironomids was stimulated in all treatments from days 6-9, and the lowest metal treatments were statistically different from the controls. Although the timing of chironomid emergence was clearly altered in the 2 highest treatments, because we tested for differences relative to controls over the entire exposure period, these treatments were not statistically different from controls. In contrast to the other groups, we saw no evidence of emergence stimulation by metals for Simuliidae. Emergence of simuliids was consistently less than controls, with emergence in the 7- to 24-CCU treatments significantly decreased. Lastly, the relationship between sex ratio (expressed as female to male adults) and metal concentration was highly significant for Baetidae, with greater abundance of female adults relative to males as metal concentrations increased

(**Figure 2.5**). Exposure to metals did not alter the sex ratios of Chironomidae and Simuliidae (**Figure S2.3**).

Discussion

Exposure of aquatic insect communities to metals significantly decreased abundance and biomass of larval and adult aquatic insects. However, effects of metals differed among dominant taxa and life stages. Chironomidae adult abundances were unaffected by the metal treatments, but larval densities were strongly reduced. Differences in metal sensitivity among body sizes and life stages of aquatic insects have been observed with field (Schmidt et al. 2013), mesocosm (Kiffney and Clements 1996, Clark and Clements 2006), and laboratory approaches (Wesner et al. 2014, Soucek and Dickinson 2015). One explanation regarding these observed differences in sensitivity is aquatic insect phenology, which influences seasonal rates of development and the life stages that are represented during contaminant exposure. For example, previous field and mesocosm experiments showed that large, developmentally mature individuals of the mayfly *Rhithrogena hageni* collected in spring were more tolerant to metal exposure compared to summer populations that were dominated by smaller individuals (Clark and Clements 2006). Populations of chironomids in our control mesocosms were dominated by early to mid-instars (**Figure S2.4**), but these individuals were highly reduced from the high metal treatments. In contrast, developmentally mature larvae were highly tolerant and able to complete their life cycle as emerged adults. Similarly, developmentally mature *R. robusta* from the single-species experiment were tolerant to metals and showed high emergence success, but the abundance of mid-instar *Rhithrogena* spp. in the community experiment significantly decreased as metal concentration increased (**Figure S2.1**). These differences in sensitivity among life stages may help to explain the discrepancies often reported between laboratory and field responses of aquatic insects to metals (Buchwalter et al. 2007, Brix et al. 2011, Clements et al. 2013). Laboratory experiments conducted with aquatic insects routinely employ larger, developmentally mature individuals.

Baetidae larvae and adults were both decreased by metal exposure, but greater metal effects were observed in larvae. The sensitivity of Baetidae to metals has been documented in laboratory and mesocosm experiments (Clements 1994, Fialkowski et al. 2003, Wesner et al. 2014, Mebane et al. 2017). The reduction in abundance of emerging adults was likely due to larval mortality that prevented emergence altogether; however, other authors have reported reduced emergence success during metamorphosis when larvae are exposed to sublethal contaminant concentrations (Liber et al. 1996, Brock et al. 2009, Wesner et al. 2014). Wesner et al. (2014) reported decreased emergence success of the mayfly *N. triangulifer* when larvae were exposed to low concentrations of Zn. Because metamorphosis is an energetically demanding process, sublethal exposure to metals may reduce energy available for emergence, resulting in mortality during this life stage transition. We hypothesize that decreased emergence of baetids in our experiment may have been in part due to mortality occurring during metamorphosis. Given that aquatic insects are the dominant organisms in many freshwater ecosystems, and that metamorphosis is required for most aquatic insects to complete their life cycles, more research is needed to understand effects of sublethal exposure to this sensitive life stage.

Physiological traits such as rates of uptake, elimination, and detoxification can influence variation in metal bioaccumulation among aquatic insect taxa (Buchwalter et al. 2008). Diptera may have greater ability to withstand metal exposures compared to mayflies, in part, because of their higher production of metallothionein-like proteins, greater efflux rates, and production of metal granules (Postma and Davids 1995, Buchwalter and Luoma 2005, Brix et al. 2011). Phylogenetic differences in physiology may explain the greater tolerance of adult chironomids compared to baetids, but it does not account for the greater relative sensitivity of larvae of both groups. One hypothesis for this discrepancy may be body size. Buchwalter et al. (2008) controlled for body size when assessing the influence of phylogeny on sensitivity of aquatic insects to metals. In contrast, our community experiment incorporated a relatively broad size range of baetids and chironomids, enabling us to make comparisons across developmental sizes. The observed difference in sensitivity among life stages demonstrates the importance of accounting for insect development and strongly suggests that phylogeny and phenology influence metal sensitivity.

Rhithrogena robusta in the single-species toxicity test were highly tolerant to metal exposure, with no effects on total emergence or larval survival. In contrast, larvae and adult emergence of the dominant mayfly (Baetidae) in the community experiments were both significantly decreased. One hypothesis to explain this difference in sensitivity is the duration of life-cycle exposure between taxa. *R. robusta* were exposed for a shorter portion of their total larval life cycle compared to Baetidae because *R. robusta* is univoltine (i.e., one generation per year) and Baetidae from our community were multivoltine (i.e., multiple generations per year). Therefore, more-sensitive, less-developed instars of Baetidae were exposed, whereas *R. robusta* were mainly exposed as fully developed late instars. Additionally, *R. robusta* develop to a much larger overall body size compared to Baetidae, potentially increasing its relative tolerance to metals (Kiffney and Clements 1996). Metal exposure in the field is spatiotemporally variable (Clements 1994, Mebane et al. 2015, Herbst et al. 2018) and because aquatic insects exhibit complex life histories (Brittain 1990), our ability to characterize full-life cycle sensitivity with mesocosms is limited by the exposure duration and the phenology of the community. Moreover, our mesocosms are closed systems that do not allow for natural colonization. Open testing systems that allow for population immigration and emigration over longer testing durations (e.g., exposure of multiple generations across multiple seasons) may better predict exposure outcomes in the field (Magbanua et al. 2016).

Effects of metals on larval and adult biomass were generally similar to those observed for abundance. Biomass of adult Chironomidae and Simuliidae was much greater than larvae, and more than 25% of the total community biomass in control mesocosms was comprised of adults. The magnitude of emerging adult biomass was unexpected, and represents a significant export of materials and energy to linked aquatic-terrestrial ecosystems. Individual emerging adults have much greater biomass compared to less developed early instars, and they contribute much greater biomass to consumers. Larval mortality in metal contaminated streams reduces secondary production, and it is likely that our mesocosm results underestimate the effects of metals to aquatic biomass in the field (Iwasaki et al. 2018). Toxicity assessments and field surveys that incorporate emerging adult biomass will more accurately characterize the relationship between contaminant exposure and aquatic subsidy dynamics.

Alterations to the timing of adult insect emergence can reflect larval stress due to direct (i.e., aqueous or dietary exposure) and (or) indirect toxicity (i.e., decreased food resources). Aquatic insects may accelerate or delay their rate of development to emergence to decrease their exposure to contaminants or other environmental stressors (Nordlie and Arthur 1981, Connolly et al. 2005, Pettigrove and Hoffmann 2005, Greig et al. 2012, Piggott et al. 2015, Lee et al. 2016, Richmond et al. 2016, Rogers et al. 2016, Joachim et al. 2017). We observed increased rates of emergence in Chironomidae for all treatments from days 6 through 9, and significantly greater cumulative emergence for metal treatments ranging from 4-13 CCU. The pronounced and immediate stimulation of emergence observed in our study suggests that chironomids can increase their rate of development and transition to their adult life stage. Increased rates of development may decrease aquatic exposure duration and increase the probability of successfully emerging and reproducing (Sibly and Calow 1989). However, irregular timing of emergence may result in adults encountering unsuitable terrestrial conditions (Nebeker 1971), shift the timing of aquatic-terrestrial subsidies that terrestrial consumers are temporally cued to encounter (Takimoto et al. 2002, Yang et al. 2010, Hastings 2012), and affect adult reproduction (Sardina et al. 2017).

Sex ratios were significantly altered in Baetidae, with relatively greater abundance of females emerging as metal concentration increased. Female baetids are generally larger than males (Caudill and Peckarsky 2003), which may result in relatively greater tolerance to exposure (Kiffney and Clements 1996). Differences in gonadal tissues between sexes may also influence rates of metal accumulation, resulting in discrepancies in toxicity (Hare 1992, Kim et al. 2012). Wesner et al. (2017) observed different rates of metal loss between adult male and female mayfly imagos (i.e., sexually mature adults) through metamorphosis, with males retaining greater body burdens of metals. The synchrony of emergence between sexes is necessary for successful reproduction to occur. Adults of some taxa such as Baetidae are short-lived (24-48 h) (Brittain 1990), and can experience high mortality due to predation (Jackson and Fisher 1986, Baxter et al. 2005). Thus, the mismatches in sex ratios observed in our study, coupled with the alterations in timing of emergence, may result in reduced fecundity and population persistence in aquatic ecosystems exposed to sublethal concentrations of metals.

Aquatic insects have complex life histories, yet with few exceptions, most invertebrate test species used in aquatic toxicity testing are short-lived and do not metamorphose. Alternative species that can be cultured in the laboratory such as the parthenogenetic mayfly *N. triangulifer* more accurately reflect the phenology of naturally occurring aquatic insects and should be used more frequently in single-species toxicity testing. Community mesocosm studies are valuable because they evaluate the responses of multiple aquatic insect taxa across a range of instar sizes and life stages under natural conditions, improving the ability to characterize full-life cycle effects. Aquatic insect emergence is a particularly relevant endpoint because it characterizes effects on a sensitive life stage and links aquatic-derived effects to terrestrial ecosystems. Failure to quantify emergence can mischaracterize contaminant effects on aquatic insect population dynamics and aquatic subsidies to terrestrial ecosystems. Regulatory frameworks would benefit by including test results that account for effects of contaminants on metamorphosis and adult insect emergence for the development of aquatic life standards.

Figures and Tables from Chapter 2

Figures 2.1 to 2.5

Table 2.1

See Appendix A for supplementary materials including Figures S2.1 to S2.4, and Tables S2.1 to S2.4

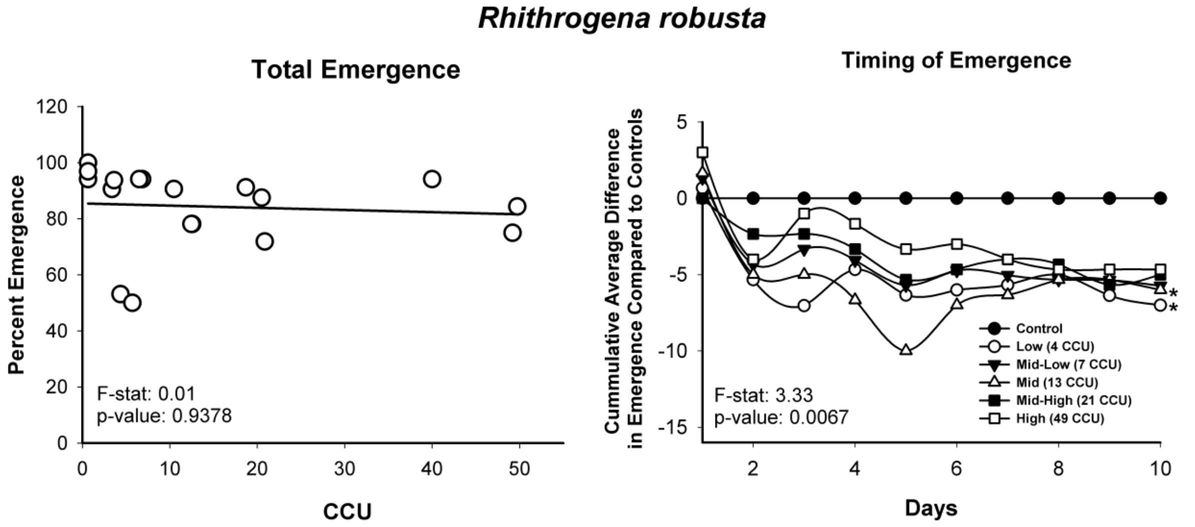

Figure 2.1 Results from the single-species toxicity test conducted with the mayfly *Rhithrogena robusta* (Ephemeroptera: Heptageniidae). The left panel shows the percentage of total emerging adults after 10 d exposure to Cu and Zn. The right panel shows the cumulative average daily difference in emergence compared to the controls over the exposure duration. CCU is the ratio of measured dissolved metal concentrations to the U.S. EPA hardness-adjusted criterion value and summed for each metal. Asterisks (*) indicate significant differences from controls.

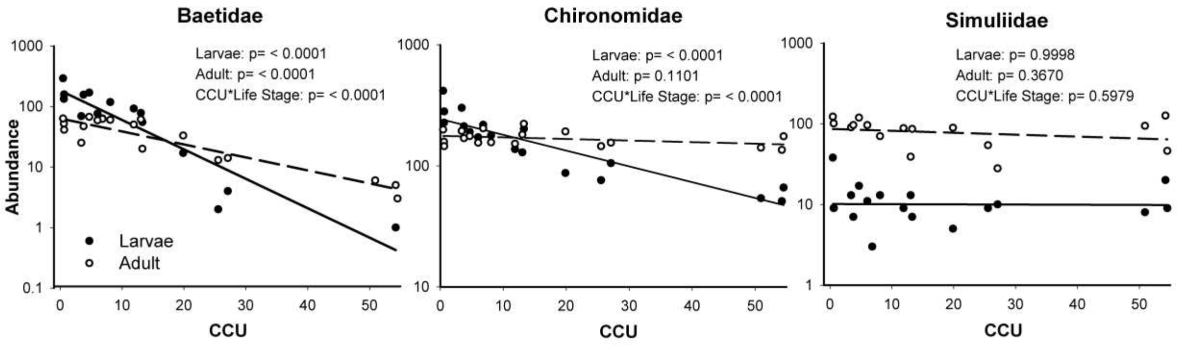

Figure 2.2 Effects of Cu and Zn on total larval (solid circles) and emerging adult (open circles) abundances of Baetidae (Ephemeroptera), Chironomidae (Diptera), and Simuliidae (Diptera) in the 14 d community exposure. Linear regression for larvae (solid line) and adults (dashed line) was used to determine significant concentration-response relationships for each life stage. ANCOVA model was used to test the hypothesis that the slopes for larvae and adults differed, as indicated by the p-value for the interaction term (CCU*Life Stage).

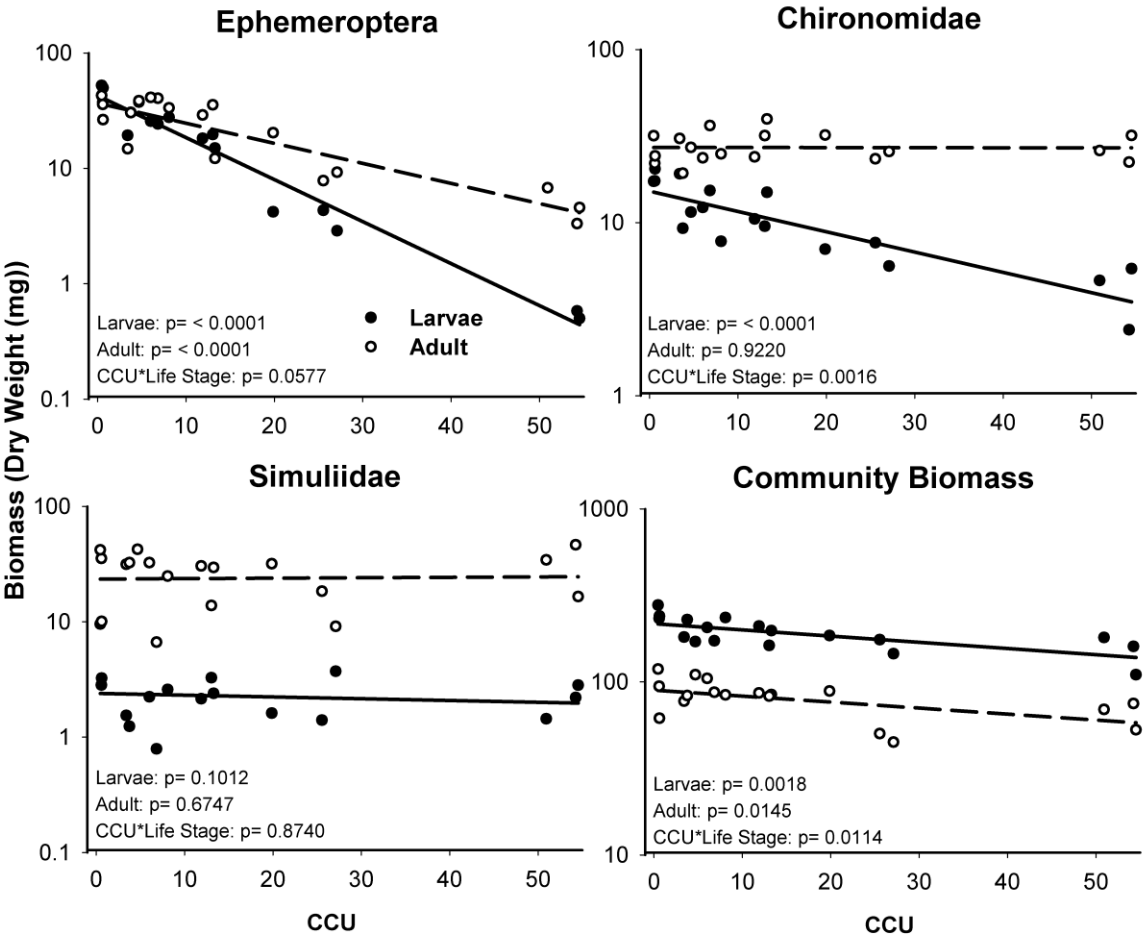

Figure 2.3 Effects of Cu and Zn on total larval (solid circles) and emerging adult (open circles) biomass (dry weight (mg)) of Ephemeroptera, Chironomidae, Simuliidae, and total community biomass in the 14 d community exposure. Linear regression for larvae (solid line) and adults (dashed line) was used to determine significant concentration-response relationships for each life stage. ANCOVA model was used to test the hypothesis that the slopes for larvae and adults differed, as indicated by the p-value for the interaction term (CCU*Life Stage). Note that the log-normalized y-axis may affect the visual interpretation of the life stage interaction.

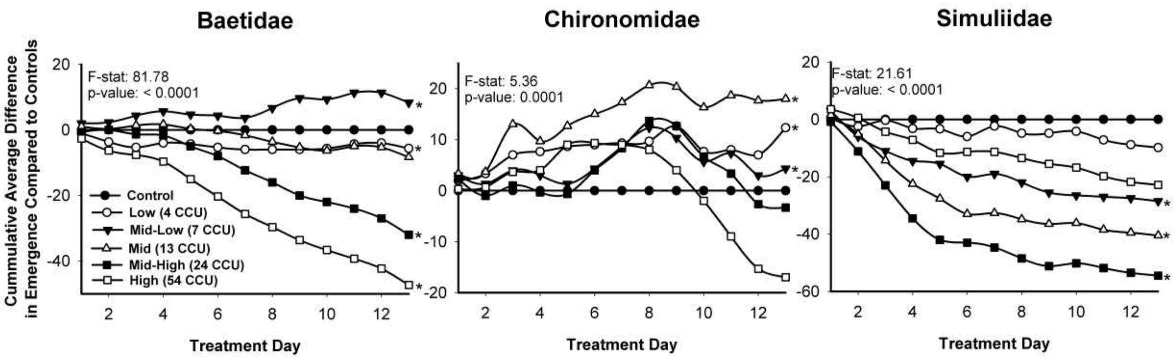

Figure 2.4 Effects of Cu and Zn on the cumulative average daily difference in emergence abundance compared to the controls over the 14 d exposure of Baetidae (Ephemeroptera), Chironomidae (Diptera), and Simuliidae (Diptera) in the community exposure. Single-factor repeated measures ANOVA was used to test for significant differences in cumulative emergence among treatments over the entire exposure duration. Post-hoc comparisons with Dunnett's test were used to test for significant relationships between treatment and control responses. Asterisks (*) indicate significant differences from controls.

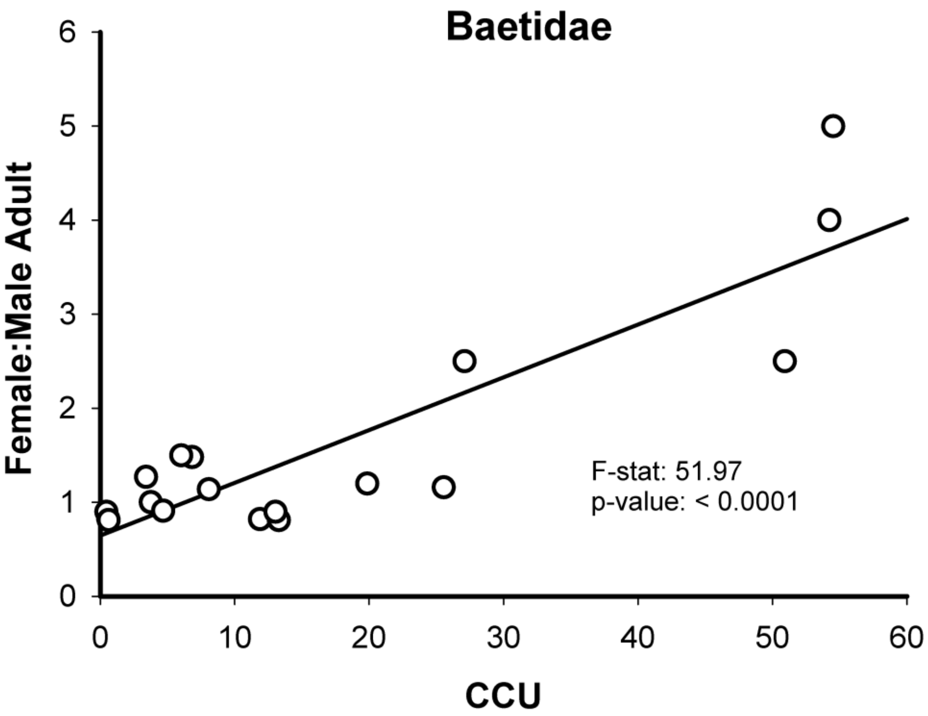

Figure 2.5 Effects of Cu and Zn on emerging adult sex ratios (female:male) of Baetidae in the stream benthic community experiment. Linear regression was used to determine significant concentration-response relationships for each life stage.

References

Bates, D.; Mächler, M.; Bolker, B.; Walker, S., Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823 2014.

Baxter, C. V.; Fausch, K. D.; Carl Saunders, W., Tangled webs: reciprocal flows of invertebrate prey link streams and riparian zones. *Freshwater Biology* 2005, 50, (2), 201-220.

Brittain, J. E., Life history strategies in Ephemeroptera and Plecoptera. In *Mayflies and stoneflies: Life histories and biology*, Springer, 1990; pp 1-12.

Brix, K. V.; DeForest, D. K.; Adams, W. J., The sensitivity of aquatic insects to divalent metals: A comparative analysis of laboratory and field data. *Science of the Total Environment* 2011, 409, (20), 4187-4197.

Brock, T.; Roessink, I.; Belgers, J. D. M.; Bransen, F.; Maund, S. J., Impact of a benzoyl urea insecticide on aquatic macroinvertebrates in ditch mesocosms with and without non-sprayed sections. *Environmental Toxicology and Chemistry* 2009, 28, (10), 2191-2205.

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

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. *Environmental Toxicology and Chemistry* 2017, 36, (2), 285-291.

Buchwalter, D. B.; Luoma, S. N., Differences in dissolved cadmium and zinc uptake among stream insects: mechanistic explanations. *Environmental Science & Technology* 2005, 39, (2), 498-504.

Campero, M.; De Block, M.; Ollevier, F.; Stoks, R., Metamorphosis offsets the link between larval stress, adult asymmetry and individual quality. *Functional Ecology* 2008, 22, (2), 271-277.

Caudill, C. C.; Peckarsky, B. L., Lack of appropriate behavioral or developmental responses by mayfly larvae to trout predators. *Ecology* 2003, 84, (8), 2133-2144.

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., Benthic invertebrate community responses to heavy metals in the Upper Arkansas River Basin, Colorado. *Journal of the North American Benthological Society* 1994, 30-44.

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.; Cadmus, P.; Brinkman, S. F., 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.; Van Hassel, J. H.; Cherry, D. S.; Cairns, J., Colonization, variability, and the use of substratum-filled trays for biomonitoring benthic communities. *Hydrobiologia* 1989, 173, (1), 45-53.

Clements, W. H.; Vieira, N. K.; Church, S. E., 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.

Cohen, J., Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement* 1973, 33, (1), 107-112.

Connolly, N.; Crossland, M.; Pearson, R., Effect of low dissolved oxygen on survival, emergence, and drift of tropical stream macroinvertebrates. *Journal of the North American Benthological Society* 2004, 23, (2), 251-270.

Echols, K. R.; Tillitt, D. E.; Nichols, J. W.; Secord, A. L.; McCarty, J. P., Accumulation of PCB congeners in nestling tree swallows (*Tachycineta bicolor*) on the Hudson River, New York. *Environmental Science & Technology* 2004, 38, (23), 6240-6246.

Fialkowski, W.; Klonowska-Olejnik, M.; Smith, B. D.; Rainbow, P. S., Mayfly larvae (*Baetis rhodani* and *B. vernus*) as biomonitors of trace metal pollution in streams of a catchment draining a zinc and lead mining area of Upper Silesia, Poland. *Environmental Pollution* 2003, 121, (2), 253-267.

Fox, J.; Weisberg, S.; Adler, D.; Bates, D.; Baud-Bovy, G.; Ellison, S.; Firth, D.; Friendly, M.; Gorjanc, G.; Graves, S., Package ‘car’. Vienna: R Foundation for Statistical Computing 2012.

Gratton, C.; Vander Zanden, M. J., Flux of aquatic insect productivity to land: comparison of lentic and lotic ecosystems. *Ecology* 2009, 90, (10), 2689-2699.

Greig, H. S.; Kratina, P.; Thompson, P. L.; Palen, W. J.; Richardson, J. S.; Shurin, J. B., Warming, eutrophication, and predator loss amplify subsidies between aquatic and terrestrial ecosystems. *Global Change Biology* 2012, 18, (2), 504-514.

Hare, L., Aquatic insects and trace metals: bioavailability, bioaccumulation, and toxicity. *Critical Reviews in Toxicology* 1992, 22, (5-6), 327-369.

Hastings, A., Temporally varying resources amplify the importance of resource input in ecological populations. *Biology Letters* 2012, rsbl20120669.

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.

Iwasaki, Y.; Schmidt, T. S.; Clements, W. H., Quantifying differences in responses of aquatic insects to trace metal exposure in field studies and short-term stream mesocosm experiments. *Environmental Science & Technology* 2018**,** 52, (7), 4378-4384.

Jackson, J. K.; Fisher, S. G., Secondary production, emergence, and export of aquatic insects of a Sonoran Desert stream. *Ecology* 1986**,** 67, (3), 629-638.

Joachim, S.; Roussel, H.; Bonzom, J.-M.; Thybaud, E.; Mebane, C. A.; Van den Brink, P.; Gauthier, L., A long-term copper exposure in a freshwater ecosystem using lotic mesocosms: Invertebrate community responses. *Environmental Toxicology and Chemistry* 2017, 36, (10), 2698-2714.

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.

Kim, K.; Funk, D.; Buchwalter, D., Dietary (periphyton) and aqueous Zn bioaccumulation dynamics in the mayfly *Centroptilum triangulifer*. *Ecotoxicology* 2012, 21, (8), 2288-2296.

Kraus, J. M.; Schmidt, T. S.; Walters, D. M.; Wanty, R. B.; Zuellig, R. E.; Wolf, R. E., Cross-ecosystem impacts of stream pollution reduce resource and contaminant flux to riparian food webs. *Ecological Applications* 2014a, 24, (2), 235-243.

Kraus, J. M.; Walters, D. M.; Wesner, J. S.; Stricker, C. A.; Schmidt, T. S.; Zuellig, R. E., Metamorphosis alters contaminants and chemical tracers in insects: implications for food webs. *Environmental Science & Technology* 2014b, 48, (18), 10957-10965.

Lee, S. S.; Paspalof, A. M.; Snow, D. D.; Richmond, E. K.; Rosi-Marshall, E. J.; Kelly, J. J., Occurrence and potential biological effects of amphetamine on stream communities. *Environmental Science & Technology* 2016, 50, (17), 9727-9735.

Lenth, R., Emmeans: Estimated marginal means, aka least-squares means. R Package Version 2018, 1, (2).

Liber, K.; Schmude, K. L.; Corry, T. D., Effects of the insect growth regulator diflubenzuron on insect emergence within littoral enclosures. *Environmental Entomology* 1996, 25, (1), 17-24.

Magbanua, F. S.; Townsend, C. R.; Hageman, K. J.; Piggott, J. J.; Matthaei, C. D., Individual and combined effects of fine sediment and glyphosate herbicide on invertebrate drift and insect emergence: a stream mesocosm experiment. *Freshwater Science* 2016, 35, (1), 139-151.

Mebane, C. A.; Eakins, R. J.; Fraser, B. G.; Adams, W. J., Recovery of a mining-damaged stream ecosystem. *Elemental Science of the Anthropocene* 2015, 3: 000042

Mebane, C. A.; Schmidt, T. S.; Balistrieri, L. S., Larval aquatic insect responses to cadmium and zinc in experimental streams. *Environmental Toxicology and Chemistry* 2017**,** 36, (3), 749-762.

Merritt, R. W.; Cummins, K. W.; Berg M.B., An introduction to the aquatic insects of North America. Kendall Hunt: 2004.

Nakano, S.; Murakami, M., Reciprocal subsidies: dynamic interdependence between terrestrial and aquatic food webs. *Proceedings of the National Academy of Sciences* 2001**,** 98, (1), 166-170.

Nebeker, A. V., Effect of high winter water temperatures on adult emergence of aquatic insects. *Water Research 1971*, 5, (9), 777-783.

Nordlie, K. J.; Arthur, J. W., Effect of elevated water temperature on insect emergence in outdoor experimental channels. *Environmental Pollution Series A, Ecological and Biological* 1981, 25, (1), 53-65.

Paetzold, A.; Smith, M.; Warren, P. H.; Maltby, L., Environmental impact propagated by cross-system subsidy: Chronic stream pollution controls riparian spider populations. *Ecology* 2011, 92, (9), 1711-1716.

Pascoe, D.; Williams, K. A.; Green, D. W., Chronic toxicity of cadmium to *Chironomus riparius* Meigen-effects upon larval development and adult emergence. *Hydrobiologia* 1989, 175, (2), 109-115.

Pettigrove, V.; Hoffmann, A., Effects of long-chain hydrocarbon-polluted sediment on freshwater macroinvertebrates. *Environmental Toxicology and Chemistry* 2005, 24, (10), 2500-2508.

Piggott, J. J.; Salis, R. K.; Lear, G.; Townsend, C. R.; Matthaei, C. D., Climate warming and agricultural stressors interact to determine stream periphyton community composition. *Global Change Biology* 2015, 21, (1), 206-222.

Postma, J. F.; Davids, C., Tolerance induction and life cycle changes in cadmium-exposed *Chironomus riparius* (Diptera) during consecutive generations. *Ecotoxicology & Environmental Safety* 1995, 30, (2), 195-202.

Rainbow, P. S., Trace metal bioaccumulation: models, metabolic availability and toxicity. *Environment International* 2007, 33, (4), 576-582.

Richmond, E. K.; Rosi-Marshall, E. J.; Lee, S. S.; Thompson, R. M.; Grace, M. R., Antidepressants in stream ecosystems: influence of selective serotonin reuptake inhibitors (SSRIs) on algal production and insect emergence. *Freshwater Science* 2016, 35, (3), 845-855.

Rogers, H. A.; Schmidt, T. S.; Dabney, B. L.; Hladik, M. L.; Mahler, B. J.; Van Metre, P. C., Bifenthrin causes trophic cascade and altered insect emergence in mesocosms: Implications for small streams. *Environmental Science & Technology* 2016, 50, (21), 11974-11983.

Sabo, J.; Power, M., River–watershed exchange: effects of riverine subsidies on riparian lizards and their terrestrial prey. *Ecology* 2002, 83, (7), 1860-1869.

Sardiña, P.; Beardall, J.; Beringer, J.; Grace, M.; Thompson, R. M., Consequences of altered temperature regimes for emerging freshwater invertebrates. *Aquatic Sciences* 2017, 79, (2), 265-276.

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.

Schmidt, T. S.; Kraus, J. M.; Walters, D. M.; Wanty, R. B., Emergence flux declines disproportionately to larval density along a stream metals gradient. *Environmental Science & Technology* 2013, 47, (15), 8784-8792.

Sibley, P. K.; Benoit, D. A.; Ankley, G. T., The significance of growth in *Chironomus tentans* sediment toxicity tests: relationship to reproduction and demographic endpoints. *Environmental Toxicology and Chemistry* 1997, 16, (2), 336-345.

Sibly, R.; Calow, P., A life-cycle theory of responses to stress. *Biological Journal of the Linnean Society* 1989, 37, (1-2), 101-116.

Soucek, D. J.; Dickinson, A., 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.

Takimoto, G.; Iwata, T.; Murakami, M., Seasonal subsidy stabilizes food web dynamics: balance in a heterogeneous landscape. *Ecological Research* 2002, 17, (4), 433-439.

Team, R. C., R: A language and environment for statistical computing. 2013.

Timmermans, K. R.; Walker, P. A., The fate of trace metals during the metamorphosis of chironomids (Diptera, Chironomidae). *Environmental Pollution* 1989, 62, (1), 73-85.

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: 1985

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

Walters, D. M.; Fritz, K. M.; Otter, R. R., The dark side of subsidies: adult stream insects export organic contaminants to riparian predators. *Ecological Applications* 2008, 18, (8), 1835-1841

Wesner, J. S.; Kraus, J. M.; Schmidt, T. S.; Walters, D. M.; Clements, W. H., Metamorphosis enhances the effects of metal exposure on the mayfly, *Centroptilum triangulifer*. *Environmental Science & Technology* 2014, 48, (17), 10415-10422.

Wesner, J. S.; Walters, D. M.; Schmidt, T. S.; Kraus, J. M.; Stricker, C. A.; Clements, W. H.; Wolf, R. E., Metamorphosis affects metal concentrations and isotopic signatures in a mayfly (*Baetis tricaudatus*): Implications for the aquatic-terrestrial transfer of metals. *Environmental Science & Technology* 2017, 51, (4), 2438-2446.

Yang, L. H.; Edwards, K. F.; Byrnes, J. E.; Bastow, J. L.; Wright, A. N.; Spence, K. O., A meta-analysis of resource pulse–consumer interactions. *Ecological Monographs* 2010, 80, (1), 125-151.

CHAPTER 3

INDIRECT EFFECTS OF IRON OXIDE ON STREAM BENTHIC COMMUNITIES: CAPTURING ECOLOGICAL COMPLEXITY WITH CONTROLLED MESOCOSM EXPERIMENTS

Introduction

Iron (III) oxyhydroxide precipitates occur in streams receiving acid mine drainage (AMD) under neutral pH and highly oxygenated conditions. Although the dissolved ferrous (Fe^{2+}) form of iron (Fe) is more soluble and is directly toxic to aquatic organisms, ferric iron (Fe^{3+}) is generally considered indirectly toxic due to physical habitat alteration (Cadmus et al. 2018a,b). In lotic systems, ferric Fe (hereafter referred to as Fe) can fill interstitial space (i.e., habitat for benthic organisms), impair primary production (e.g., smother periphyton, decrease light penetration), and reduce benthic invertebrate abundance and richness (McKnight and Feder 1984, Vuori 1995, Wellnitz and Sheldon 1995, DeNicola and Stapleton 2002, Guasch et al. 2012). The US Environmental Protection Agency (USEPA) chronic criterion for Fe (1.0 mg/L total recoverable) was primarily developed with field surveys of trout in a small stream in Colorado (USEPA 1976). Outside of the US, lower freshwater Fe criterion (0.3 mg/L) have been adopted by Canada (MacDonald et al. 2002), Australia and New Zealand (ANZECC, 2000). Although adverse effects typically do not occur at the USEPA criterion concentration in the laboratory, and Fe is generally considered relatively non-toxic (Biesinger and Christensen 1972, Randall et al. 1999), field surveys have demonstrated effects to benthic communities below the USEPA criterion (Linton et al. 2007).

Discrepancies between laboratory-predicted Fe toxicity and field observations are likely due to differences in the effects of Fe at dissimilar spatiotemporal scales. For example, most laboratory approaches are short-duration and do not allow sufficient time for Fe deposition and accumulation into interstitial spaces (Cadmus et al. 2018a). Also, many commonly employed test species are pelagic (e.g., *Daphnia magna*, *Pimephales promelas*) and can avoid Fe deposition, or they are adapted to fine sediment environments similar to deposited Fe precipitates (e.g., *Chironomus riparius*). These test species do not adequately represent the diversity of life histories that naturally occur in stream benthic communities (Buchwalter et al. 2007, Poteat and Buchwalter 2014). In contrast, field approaches are somewhat limited because they usually do not control confounding and extraneous variables (Buchwalter et al. 2017). For instance, Fe generally co-occurs with other metals at AMD sites, temporal differences in water velocity influence rates of Fe deposition, and the presence of other anthropogenic stressors limits the ability to isolate the effects of Fe (Millward and Moore 1982, Gerhardt 1994, Gerhardt and Westermann 1995).

Stream mesocosms provide a unique combination of ecological realism and experimental control, addressing many of the shortcomings of laboratory and field evaluations of Fe oxides (Cadmus et al. 2018a). Mesocosm experiments can test the responses of native benthic communities that contain numerous taxa with complex life histories (Clements 2004, Clark and Clements 2006, Kotalik and Clements 2019). Exposure to Fe under conditions that include natural benthic habitat and continuous flow better simulate the dynamics of Fe deposition in streams compared to most laboratory approaches. These experiments can test the direct and indirect effects of contaminants on primary production and can investigate potential trophic cascades that may occur between benthic invertebrates (i.e., consumers) and autochthonous primary producers (i.e., food resources) (Rogers et al. 2016), while still controlling most physical and chemical characteristics of the exposure system. Importantly, mesocosms have reconciled field and laboratory results for other metals, often with benthic invertebrates responding at lower concentrations compared to single-species assessments (Clements et al. 2013, Mebane et al. 2017, Iwasaki et al. 2018).

Evaluations of benthic communities typically examine immature life stages (e.g., larvae) of aquatic insects to characterize taxa sensitivity, yet most aquatic insects metamorphose to a winged-terrestrial adult life stage to reproduce and complete their life cycle. Emerged adults are crucial for population persistence, and they provide prey subsides to linked riparian consumers (Sabo and Power 2002, Baxter et al. 2005). However, the transition from larvae to the adult life stage is biologically stressful because of the high rates of cell differentiation and energy expenditure during tissue reorganization (Campero et al. 2008). Field and laboratory studies of aquatic insect emergence have shown greater sensitivity of emerging adults to dissolved trace metals compared to larvae (Schmidt et al. 2013, Wesner et al. 2014). Therefore, the evaluation of larval life stages alone risks underestimating the effects of metals on aquatic insects and the export of subsidies to linked-terrestrial ecosystems (Pascoe et al. 1989, Schmidt et al. 2013, Wesner et al. 2014, Soucek and Dickinson 2015). Many commonly used laboratory test species do not metamorphose to an adult life stage (e.g., *Daphnia* spp., *Ceriodaphnia dubia*); thus, these tests lack an equivalent endpoint to evaluate the effects of Fe oxyhydroxides on metamorphosis and emergence. Furthermore, most studies conducted with aquatic insects fail to enumerate emerged adults as they emigrate from testing systems during exposure.

To test the effects of Fe oxide on stream communities, we exposed naturally-colonized benthic communities to a gradient of ferric Fe in stream mesocosms. The Fe exposure simulated natural conditions that maintained constant Fe concentrations in the water column while interstitial space of the benthic habitat filled with Fe precipitates. Specific goals of this study were: 1) to quantify effects of Fe deposition on primary production by measuring algal colonization; 2) to compare lethal and sublethal effects of Fe among aquatic insect taxa and between developmental life stages (i.e., larvae versus emerged adult); and 3) to determine concentrations of Fe that significantly alter benthic community structure and to identify taxa that most contributed to community differences among treatments.

Methods

A mesocosm experiment was conducted at the Stream Research Laboratory (SRL) located at Colorado State University Foothills Campus, Fort Collins, Colorado, USA. Naturally colonized stream benthic communities were collected from the Arkansas River (AR; Leadville, Colorado, USA), a high-elevation (~ 3000 m ASL) 4th-order mountain stream (39.1283 N, 106.3111 W). Communities were collected from the AR using a colonization technique described by Clements et al. (1988). Plastic trays (10 x 10 x 6 cm) filled with pebble and small cobble were deployed in the AR for 30 d and colonized with a diverse benthic assemblage (collected August 26, 2015). The trays were colonized by a range of life stages of aquatic insects that included developmentally immature taxa, as well as taxa nearing the completion of their life cycle as emerged adults. Five trays were randomly placed into each of 18 insulated coolers, transported to the SRL, transferred to 20-L stream mesocosms, and assigned Fe treatments.

Water quality in the SRL is representative of cold water Rocky Mountain streams with cool temperatures (10-15 ºC); low hardness (30-35 mg/L $CaCO_3$), alkalinity (25-30 mg/L

$CaCO_3$), conductivity (60-90 µS/cm) and concentrations of dissolved organic carbon (2.5-3.0 mg/L); and circumneutral pH (7.0-7.8). Current in the 18 mesocosms was driven by submersible water pumps (EcoPlus®) at a rate of 2233 L/h (i.e., 1.9 hydraulic circulations/min). On day 0, four 25-cm^2 unglazed ceramic tiles were placed in each mesocosm directly downstream of the colonization trays to estimate algal colonization. All tiles were pre-soaked in bleach for 24 h and rinsed multiple times with reverse osmosis water before deployment. Each mesocosm was covered with insect mesh netting to capture emerged adults (**Photo S1**). Emerged adults were aspirated from the netting every evening between 17:00 and 19:00 and preserved in 80% ethanol. To prevent emigration by drifting invertebrates, standpipes in the stream mesocosms were covered with fine mesh.

Each mesocosm received dilution water at 1.0 L/min while peristaltic pumps delivered 10 mL/min Fe stock solution. To convert all Fe(II) to Fe(III), ferric chloride stock solutions in 20-L carboys for each stream were vigorously aerated while NaOH was added to attain a pH greater than 6.5. After 60 min of stirring, pH was reexamined before dosing was initiated. Target Fe treatment concentrations were 0.9, 1.8, 3.7, 7.5, and 15.0 mg/L. To maintain a homogenous suspension of Fe oxyhydroxide precipitates in the carboys, stock solutions were vigorously aerated throughout the exposure. Total (unfiltered) and dissolved (0.45-µm filtered) water samples were taken from each mesocosm on 4 dates, acidified to a $pH < 2.0$ using analytical grade nitric acid to dissolve all Fe, and analyzed by flame atomic absorption spectrophotometry. In addition, ferrous Fe concentration was measured by colorimetric analysis using a ferrous Fe reagent (1,10 -phenanthroline indicator; Hach method 8146. Loveland, CO, USA). Temperature, pH, conductivity, and turbidity were measured every 3 d (YSI Pro1030). Alkalinity and hardness were measured by titration (USEPA methods 200.7; 310.7). To estimate deposition of Fe, 50 mL beakers containing 20 mL of glass beads (6-mm diameter) were placed in each stream mesocosm. The Fe was allowed to settle in the interstitial spaces of the beads for 96 h, then 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 (GC50, Sterlitech). The filter and Fe precipitates were then dried and weighed.

After 10 d, we estimated algal colonization (on ceramic tiles) with a BenthoTorch (bbe Moldaenke GmbH, Germany), an *in situ* pulse-amplitude modulated fluorometer that measures benthic algae biomass as chlorophyll *a* and estimates the relative contribution of diatoms, green algae, and cyanobacteria pigment complexes to total algal biomass (Beutler et al. 2002). On day 11, we estimated community metabolism (defined as the difference between photosynthesis and whole community respiration) using diurnal changes in dissolved O_2 concentrations in the water column (YSI ProODO. Yellow Springs, OH, USA). After 14 d, colonization trays were rinsed through a 350-µm sieve, and organisms and detrital material were preserved in 80% ethanol. All larvae were enumerated to the lowest taxonomic resolution possible, typically to genus, except for chironomids that were identified to subfamily or tribe. Emerged adult insects were enumerated and identified to family or subfamily. For biomass estimates, larval and adult taxa were dried at 60 ºC for 72 h and weighed to the nearest 0.00001 g.

Statistical Analyses

All univariate statistical analyses were conducted using R statistical computing (R Core Team v3.5.1). We used a linear model ("LM" function; package 'car'; Fox et al. 2012) to test significant concentration-response relationships between Fe and algae community composition and metabolism. Non-linear concentration response curves were fit for the algal data using

Sigma Plot Ver. 11. A linear model was used to test concentration-response relationships between life-stage abundances and biomass of the three dominant taxa that had sufficient densities of larvae *and* adults (n > 15) (Baetidae, Chironomidae, and Simuliidae) and total macroinvertebrate abundance. To determine if the concentration-response relationships of larvae and adults were statistically different, an ANCOVA interaction model was used (“LM” function), with Fe as the continuous predictor and life stage as the categorical predictor of either abundance or biomass. Statistical significance of the interaction term indicates that the slopes developed for larvae and adults differed (i.e., the effect of Fe was dependent on life stage). Lastly, for all significant linear models we estimated EC20 values (effect concentration that reduced the endpoint by 20%, exclusive of control data) with two-parameter log-logistic concentration-response curves. The only exception was for cyanobacteria colonization that required a two-parameter log-normal concentration-response model because biomass increased with Fe concentration (R package ‘drc’; Ritz et al. 2015). EC50 values were also calculated to compare to EC20 estimates (**Figure S3.1**); however, we chose to report EC20s based on the USEPA risk-management decision to derive Final Chronic Values using EC20 values (USEPA 1999). Concentration-response models were selected based on Akaike Information Criterion (“mselect” function) and visual model fit for all endpoints.

To test for differences in the timing of adult emergence, Fe treatments were categorized (5 treatments, 3 replicates each) to support repeated-measures ANOVA analyses. Daily observations were combined to cumulative emergence, and abundance of emerging adults was summed over the duration of the experiment. We combined daily observations because we wanted to test if the difference in the number of emerging individuals over the entire experiment differed among treatments. A linear mixed model (function “lmer,” package ‘lme4”; Bates et al.

2014) was fit to the cumulative emergence data to test if the effects of Fe on cumulative emergence varied during the exposure. We chose a covariance structure of compound symmetry, which assumes constant correlation between observations on the same sample replicate (i.e., cumulative observations are dependent on the previous observation). *Post hoc* comparisons with "estimated marginal means of linear trends" function from package 'emmeans' (Lenth 2018) were used to compare treatment means of emergence abundance to control values with Dunnett's Test. All data were log transformed to meet the assumptions of parametric statistics and to improve model fit.

Multivariate analyses (PRIMER-e v7; + PERMANOVA; Quest Research Limited, Cambridge, United Kingdom) were used to test for changes in aquatic insect community composition among Fe treatments (Anderson et al. 2008). Larval and adult abundance were log (x+1) transformed and a Euclidean distance matrix was calculated. Rare taxa were removed based on the *a priori* decision to define rare species as less than 10 individuals in any treatment. Treatments were categorized, and a one-way permutational multivariate analysis of variance (PERMANOVA) was run with 999 permutations to test for significant alterations in community composition. Pairwise-comparisons were used to test for differences between treatments and controls. Because of the relatively low number (10) of 'Unique Permutations' generated (i.e., the number of unique values of the test statistic obtained under permutation). Monte Carlo adjusted *p*-values were calculated by random sampling from the asymptotic permutation distribution. A Similarity Percentages (SIMPER) analysis was performed on the log-transformed data to determine which taxa accounted for greater than 50% of the dissimilarity among the Fe treatments compared to the controls. Lastly, a nonmetric-multidimensional scaling (NMDS) plot was used to visualize differences among treatments in ordination space.

Results

Measured total Fe concentrations ranged from 0.85 to 14.85 mg/L Fe and were within 10% of the target values (**Table 3.1**). Among all treatments, filtered Fe and ferrous Fe concentrations were below the detection limit. Water hardness and alkalinity were approximately 31 mg/L $CaCO_3$ (s.d. = 0.55, n = 12) and 28 mg/L $CaCO_3$ (s.d. = 1.01, n = 12), respectively. Temperature and pH showed little variation and were consistent among treatments. Specific conductance (77.6 to 195 µS/cm) and turbidity (1.6 to 9.7 NTU) increased with metal concentration. Rates of deposition in all treated mesocosms were significantly correlated with total Fe measured in the water column (R^2= 0.86; **Figure S3.2**).

The benthic community was diverse, with control mesocosms averaging (± s.e., n = 3) 28 (± 1) taxa and 1093 (± 86) individuals per stream (**Table S3.1**). Number of taxa and abundance of emerging adults in controls averaged 8 (± 0.3) and 520 (± 25), respectively. Three dominant taxonomic groups (Baetidae, Chironomidae and Simuliidae) were present at sufficient larval and adult densities in controls for statistically valid comparisons. Baetidae, Chironomidae, and total larval and emerging adult abundance significantly ($p < 0.10$) decreased as Fe concentration increased (**Figure 3.1; Table S3.2**). The abundance of larval Simuliidae was unaffected by Fe treatments, but emergence was significantly reduced. For all dominant groups and total community abundance, the ANCOVA interaction term was significant, indicating that the response to Fe was dependent on life stage (**Table S3.3**). The effects of Fe were greater on larval survival than emergence for Baetidae, Chironomidae and total community abundance, but Simuliidae emerged adults were more sensitive to Fe than larval responses. Among significant linear models, EC20 values ranged from 0.01 (Baetidae adult abundance) to 1.01 mg/L Fe (Chironomidae adult abundance), respectively (**Table S3.4**). Lastly, the cumulative emergence of the 3 dominant groups was significantly altered during the 14-d exposure (**Figure 3.2**) with decreased emergence relative to controls (**Table S3.5**).

Across the Fe treatment gradient, biomass of adult mayflies, chironomids and simuliids was greater than larval biomass after the 14-d exposure. Effects of Fe on biomass of baetids and simuliids (**Figure 3.3**) were generally similar to those observed for abundance (**Figure 3.1**). However, patterns in biomass of chironomids (larvae and adults) and total macroinvertebrate biomass differed from those observed for abundance, with similar effects of Fe on these group's biomass reduction regardless of life stage (i.e., slopes for larvae and emerged adults were not significantly different) (**Table S3.3**). EC20 values ranged from 0.001 mg/L Fe for Baetidae adult biomass to 0.80 mg/L Fe for Chironomidae adult biomass (**Table S3.4**).

Fe (III) significantly altered community composition based on PERMANOVA results (**Table S3.6**). Visual inspection of the NMDS plot showed clear separation among Fe treatments compared to controls, with increased separation as Fe concentration increased (**Figure 3.4**). The three control mesocosms were very close in NMDS space (i.e., similar community structure), whereas separation within Fe treatments was much greater, suggesting that exposure to Fe increased variability in community composition. Pairwise comparisons between each of the treatments and the controls showed substantial alterations in community composition (**Table 3.2**), including significant alterations at the lowest Fe concentration. The taxa that most contributed to the differences among treatments based on the SIMPER analysis were primarily larval life stages of Baetidae, two groups of chironomids (Chironomini and Tanytarsini), and Glossosomatide (Trichoptera).

Algal colonization was significantly altered among all dominant algal groups (**Figure 3.5**). Biomass of green algae, diatoms, and total chlorophyll *a* decreased significantly across the

Fe exposure gradient, with dramatic reductions at concentrations near or lower than the USEPA criterion of 1.0 mg/L (**Table S3.2**). In contrast, biomass of cyanobacteria significantly increased as the Fe concentration was increased. These shifts in biomass of the major primary producers were associated with significant reductions in community metabolism, likely reflecting lower biomass and primary productivity in Fe treated streams (**Figure S3.3**). Among the significant algae linear models, EC20 values ranged from 0.0004 (chlorophyll *a*) to 1.935 mg/L Fe (cyanobacteria), respectively (**Table S3.4**).

Discussion

Ferric Fe significantly altered community composition and community metabolism after 14 d of exposure in our stream mesocosms, with effects on some endpoints occurring near or below the USEPA chronic criterion for Fe (1.0 mg/L). Abundance of both larval and adult life stages decreased for most aquatic insect taxa, and effects varied between life stages and among taxa. Larval life stages were generally more sensitive to Fe compared to emerging adults. An important exception was for Simuliidae, in which adult emergence significantly decreased, but larval abundance was unaffected. Differences in metal sensitivity among life stages of aquatic insects have been observed with field (Schmidt et al. 2013), mesocosm (Kiffney and Clements 1996, Kotalik and Clements 2019), and laboratory approaches (Wesner et al. 2014, Soucek and Dickinson 2015). Although previous studies have reported decreased emergence resulting from exposure to fine sediment (Elbrecht et al. 2016, Magbanua et al. 2016), to our knowledge, this is the first study to quantify the physical effects of Fe to benthic community emergence. Importantly, our results demonstrate that ferric Fe toxicity is life stage-dependent, and suggests that testing for Fe toxicity should incorporate multiple life stages to characterize taxa sensitivity.

Because biomass is directly related to ecosystem energetics, the effects of Fe on insect biomass is a more direct measure of subsidy export to riparian ecosystems compared to insect abundance. Given that the effects of Fe are generally considered indirect, measuring biomass is important because reduced food availability, food quality, and habitat can influence growth rates that may affect larval development and size of emerging adults (Blumenshine et al. 1997, Grieg et al. 2012, Piggot et al. 2015, Alexander et al. 2016). The greater sensitivity of emerging simuliid adults compared to larvae represents a comparatively greater biomass effect because adults are larger than larvae. In contrast, our observation of greater larval mortality compared to emerging adults for other groups (i.e., baetids) has implications for secondary production because larvae are unable to reach maturation. Because the duration of our study was relatively short (14 d), it is likely that we underestimated potential effects of chronic Fe exposure to invertebrate biomass in the field (Iwasaki et al. 2018). Linked aquatic and terrestrial consumers are dependent on aquatic insect subsidies, and biomass estimates more accurately estimate subsidy dynamics compared to abundances. More research is needed to address the effects of Fe deposition on secondary production and the delivery of aquatic subsidies to terrestrial ecosystems.

We expected to see the greatest effects of Fe on sedentary taxa that were unable to avoid smothering of habitat and on grazing insects that were affected by diminished periphyton quality and biomass. Indeed, sedentary chironomids (i.e., Chironomini, Tanytarsini) and grazing caddisflies (i.e., Glossosomatidae) strongly contributed to the dissimilarity among Fe treatments. While simuliid adults were not the strongest contributors to treatment dissimilarly, we believe that high adult sensitivity was due to their sedentary pupal life stage and branched pupal gills (Merritt et al. 2004) that are susceptible to smothering. Physiological traits of aquatic insects are often used to characterize metal bioaccumulation and to predict metal sensitivity in the field (Buchwalter et al. 2008). However, biological traits such as habit (e.g., burrower, sprawler, swimmer), habitat (e.g., lotic erosional, depositional), and drift propensity may be more appropriate predictors for indirect stressors such as Fe (Rader 1997, Poff et al. 2006, Cadmus et al. 2016, Dabney et al. 2018, Reich et al. 2019).

The dramatic effects of Fe on baetid mayflies were unexpected, because these organisms are strong swimmers and highly mobile (Brittain 1990), and therefore should be capable of avoiding Fe precipitates deposited in the benthos. One possible explanation for the high sensitivity of baetids to Fe exposure was reduced food quantity and quality. Baetidae gut contents in low Fe treatments clearly show ferric precipitates accumulated in the gut column (**Figure 3.6**). Cain et al. (2016) reported that ferric Fe did not inhibit feeding of diatoms by the benthic grazer *Lymnaea stagnalis*, which suggests that periphyton with deposited Fe precipitates are not avoided. It is unclear if the dietary accumulation of indigestible ferric Fe affects aquatic insect growth and survival, or if changes in pH associated with the digestive tract affect exposure outcomes. More research is needed to evaluate the effects of dietary accumulation of Fe.

Overall algal colonization after 10 d was significantly reduced in stream mesocosms. Chlorophyll *a*, diatoms, and green algae biomass were among the most sensitive endpoints we examined, with EC20 values well below the USEPA chronic criterion for Fe (**Table S3.4**). In contrast, Cadmus et al. (2018a) did not observe significant reductions to algal biomass on tiles that were colonized (~ 30 d) in the field before Fe (III) exposure. Because the current study examined effects without prior incubation, the high sensitivity of these groups is likely due to the functional inability of algae to initially colonize and grow. Metal oxides can prevent the attachment of algae to substrate and limit photosynthesis due to direct smothering and reduced light penetration associated with increased turbidity (McKnight and Feder 1984, Wellnitz and Sheldon 1995, Niyogi et al. 2002). However, cyanobacteria biomass showed the opposite trend in our study, with increased biomass associated with greater Fe exposure. Cyanobacteria are generally considered unpalatable and even toxic to some aquatic insect consumers, and its presence in streams is generally considered an indicator of stress (McCormick and Cairns 1994, Jüttner and Wessel 2003, Douterelo et al. 2004, Camargo and Alonso 2006). Cyanobacteria are also the dominant primary producers at AMD sites with neutral or higher pH (Bray et al. 2009). The evaluation of primary producers in streams degraded by Fe (III) is important because Fe precipitates appear to affect some groups of primary producers more than others, and changes in biomass and composition affect aquatic insects that depend on these autochthonous resources.

In conclusion, we characterized the effects of ferric Fe to benthic communities by exposing naturally colonized communities under environmentally realistic habitat and exposure conditions. Importantly, we incorporated indigenous aquatic insects with complex life histories, and demonstrated that adult emergence is a particularly useful response because it characterizes the effects of a sensitive life stage and linked aquatic-derived effects to terrestrial ecosystems. Differing spatiotemporal approaches are needed to predict effects of contaminants to benthic communities using structural and functional endpoints in combination. Because traditional toxicity testing methodologies were not developed to evaluate the indirect effects of contaminants, other approaches such as mesocosm experiments are needed to characterize and predict these types of responses in real-world ecosystems. For these reasons, the development of water quality standards would generally benefit by including mesocosm testing results.

Figures and Tables from Chapter 3

Figures 3.1 to 3.6

Table 3.1

See Appendix B for supplementary materials including Figures S3.1 to S3.4, and Tables S3.1 to S3.6

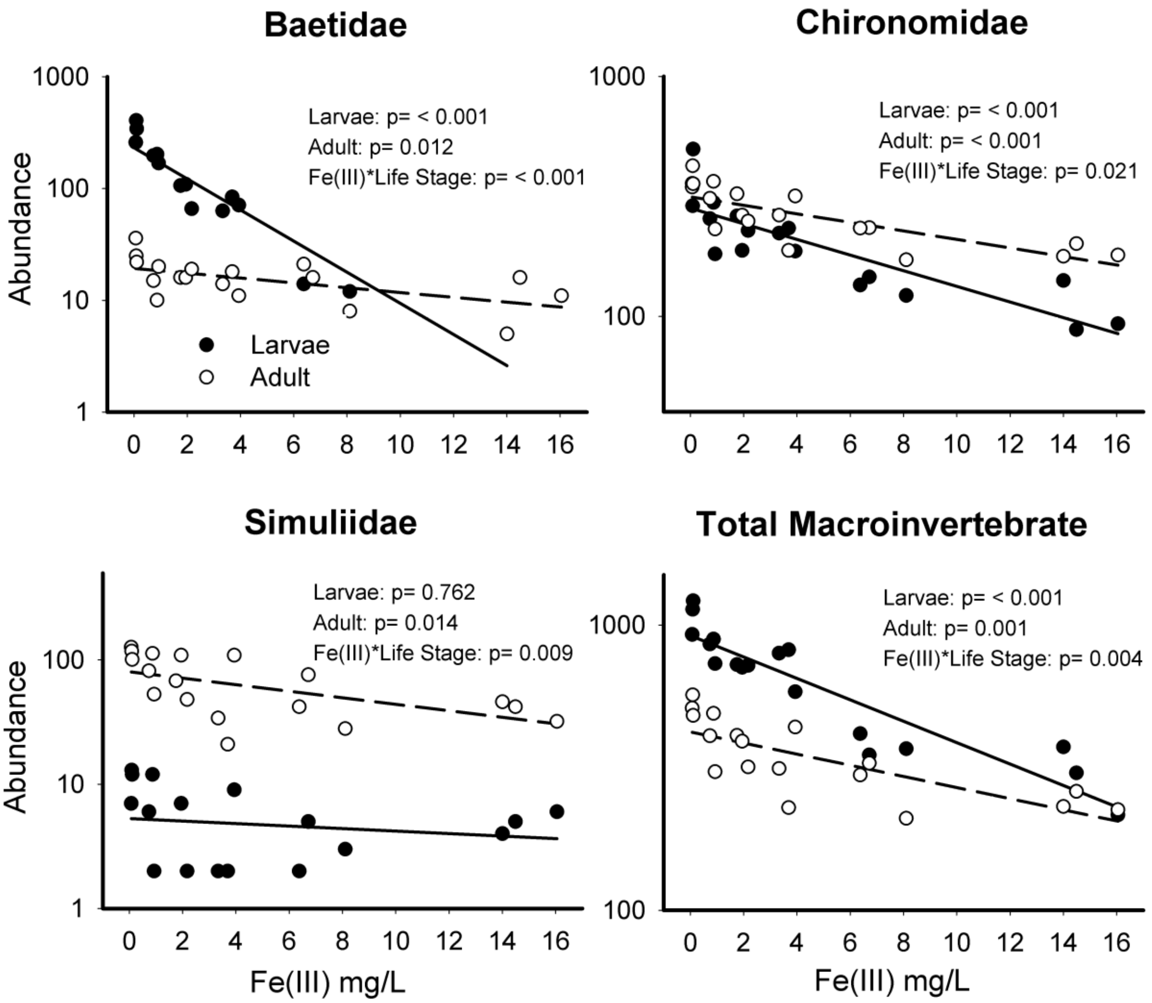

Figure 3.1 Effects of Fe (III) on total larval (solid circles) and emerging adult (open circles) abundances of Baetidae (Ephemeroptera), Chironomidae (Diptera), Simuliidae (Diptera), and the total macroinvertebrate community after the 14-d exposure. Linear models for larvae (solid line) and adults (dashed line) were used to determine significant concentration-response relationships for each life stage. ANCOVA model was used to test the hypothesis that the slopes for larvae and adults differed, as indicated by the p-value for the interaction term (Fe (III)*Life stage).

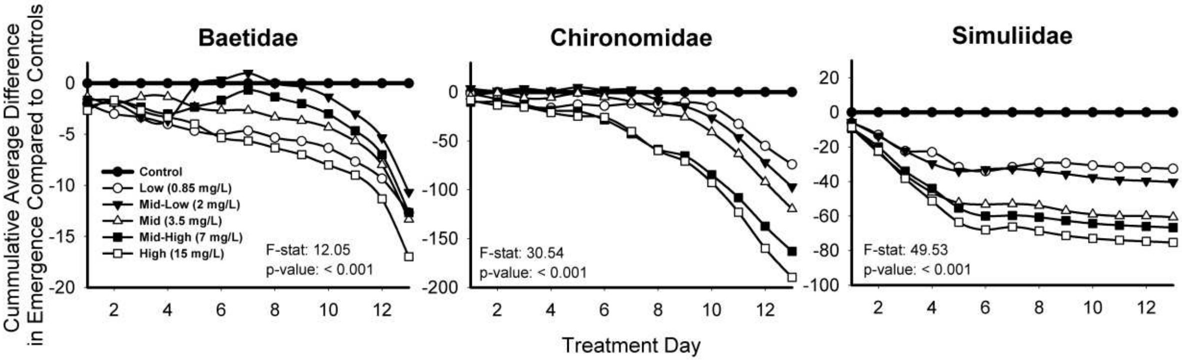

Figure 3.2 Effects of Fe (III) on the cumulative emergence of Baetidae (Ephemeroptera), Chironomidae (Diptera), and Simuliidae (Diptera) during the 14-d exposure. Single-factor repeated measures ANOVA was used to test for significant differences in cumulative emergence among treatments during the exposure.

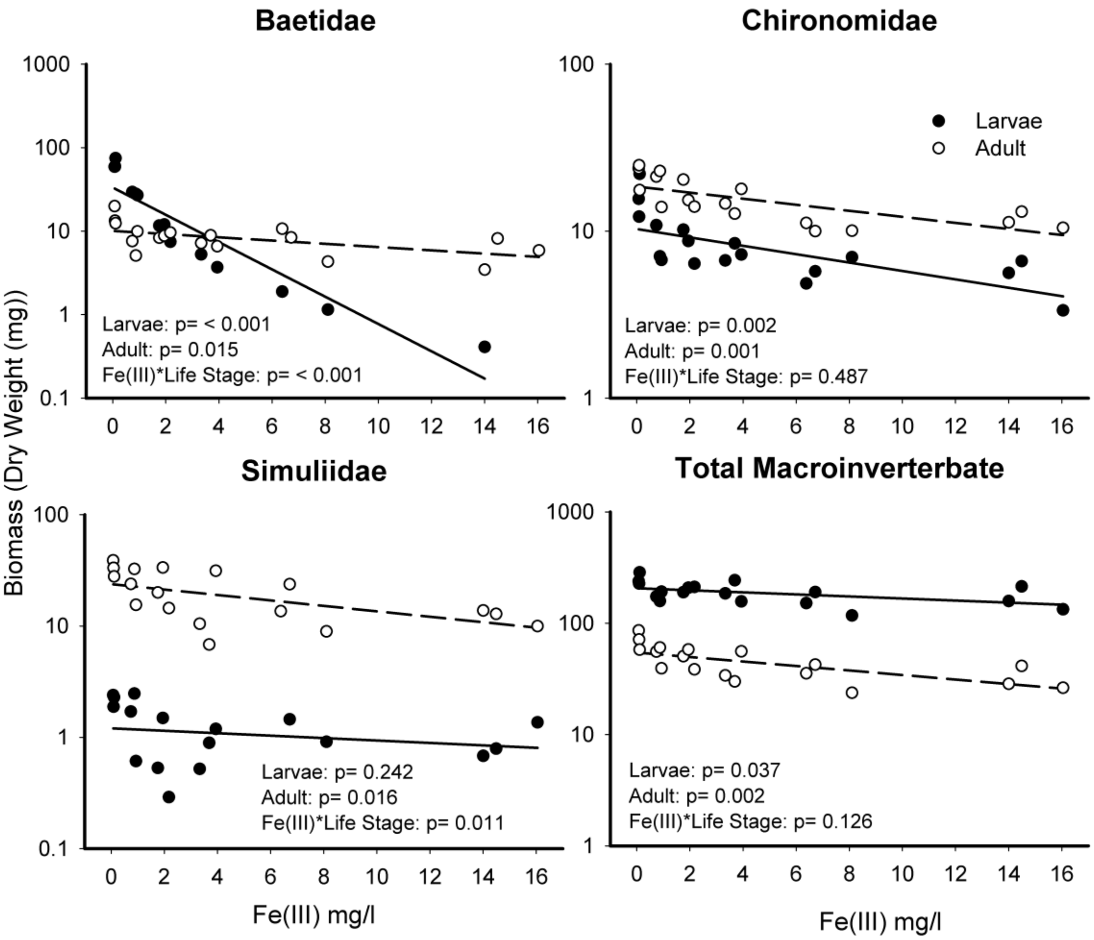

Figure 3.3 Effects of Fe (III) on total larval (solid circles) and emerging adult (open circles) biomass of Baetidae (Ephemeroptera), Chironomidae (Diptera), Simuliidae (Diptera), and the total macroinvertebrate community in the 14-d community exposure. Linear regression for larvae (solid line) and adults (dashed line) was used to determine significant concentration-response relationships for each life stage. ANCOVA model was used to test the hypothesis that the slopes for larvae and adults differed, as indicated by the p-value for the interaction term (Fe (III)*Life stage).

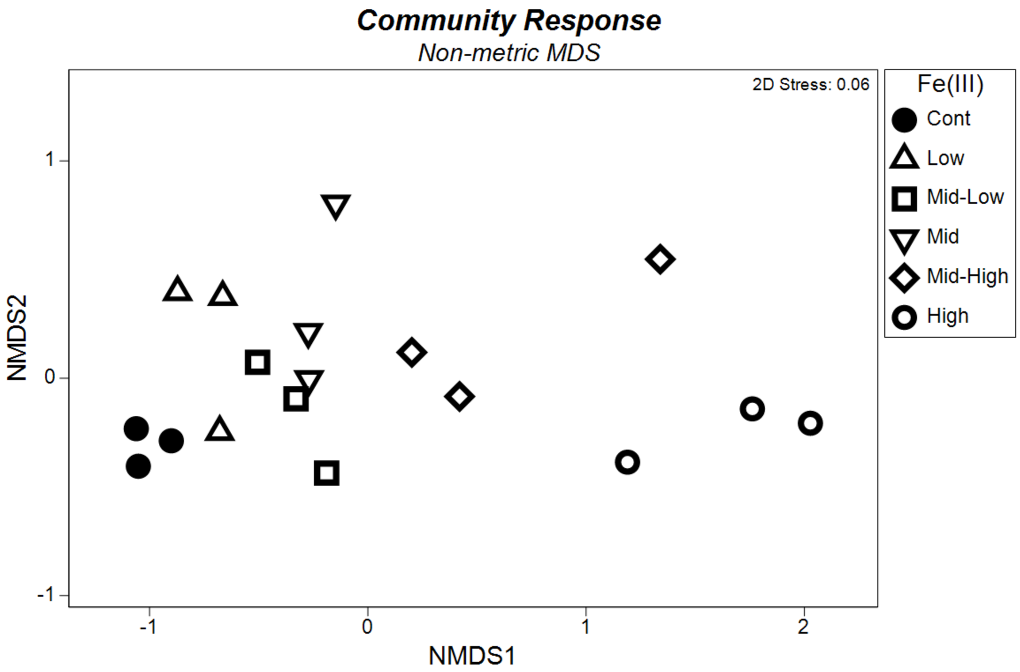

Figure 3.4 Non-metric multi-dimensional scaling (NMDS) plot of aquatic invertebrate community abundance in the five Fe (III) treatments and the control. Permutational multivariate analysis of variance (PERMANOVA) was used to test for significant shifts in community composition among treatments.

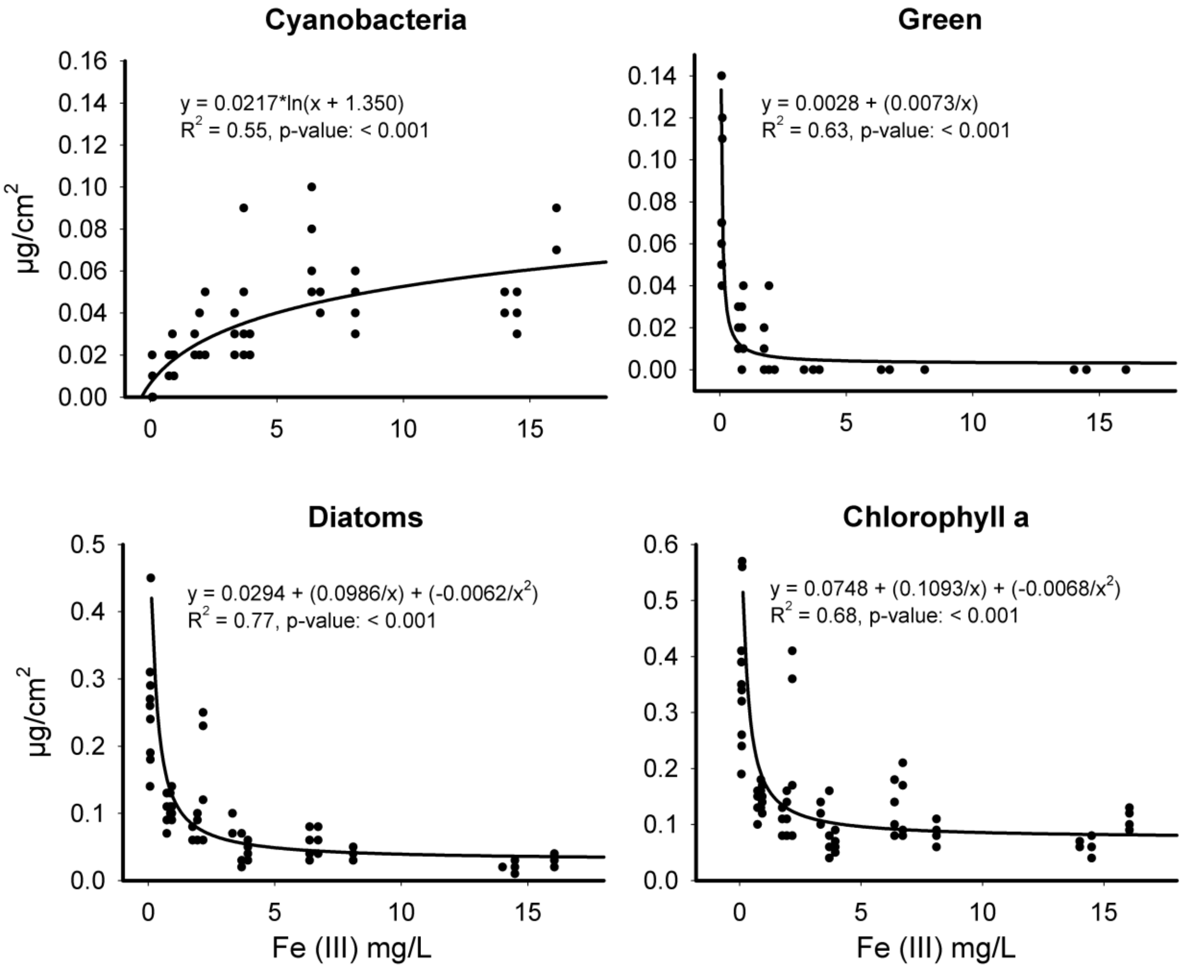

Figure 3.5 Effects of Fe (III) on 10-d algal colonization biomass (µg/cm^2) on ceramic tiles (n=4) in stream mesocosms for major groups of the algal community.

Figure 3.6 Photo of *Baetis* sp. showing ferric Fe accumulation in gut contents. The two *Baetis* sp. on the left are from control treatments, the two middle *Baetis* sp. were from low Fe treatments, and the two *Baetis* sp. from the right were collected from mid-low Fe treatments. At the bottom of the photo, a horizontal 0.5-mm-diameter piece of pencil lead provides a size reference.

Table 3.1 Water quality characteristics in stream mesocosms. Values are reported as mean (± s.d., n = 12). Fe detection limit using flame atomic absorption spectrophotometry was 0.06 mg/L.

Fe (III) Treatment	Target (mg/L)	Dissolved Fe (mg/L)	Total Fe (mg/L)	Temp (°C)	pH	Conductivity (µS/cm)	Turbidity (NTU)
Control	0	bdl	0.08 (0.02)	14.9 (0.5)	7.43 (0.18)	77.6 (1.9)	1.6 (0.1)
Low	0.94	bdl	0.85 (0.10)	14.7 (0.3)	7.42 (0.14)	86.9 (3.6)	2.3 (0.2)
Mid-Low	1.88	bdl	1.95 (0.21)	14.9 (0.6)	7.42 (0.12)	93.6 (1.6)	3.1 (0.2)
Mid	3.75	bdl	3.65 (0.30)	14.7 (0.3)	7.4 (0.07)	107.9 (2.9)	3.9 (0.1)
Mid-High	7.5	bdl	7.06 (0.91)	14.4 (0.4)	7.37 (0.07)	133.4 (7.8)	5.4 (0.8)
High	15	bdl	14.85 (1.06)	14.3 (0.4)	7.39 (0.04)	195.9 (14.4)	9.7 (0.4)

Table 3.2 Single-factor permutational MANOVA pair-wise comparison of each treatment to the control for the aquatic macroinvertebrate community exposed to ferric Fe for 14 d. P-values were calculated using Monte Carlo permutations. SIMPER outputs show the average dissimilarity (Distance) among the exposure levels compared to controls and the species that contributed to the upper 50% of the differences among treatments. Taxa ("L" denotes larvae; "A" denotes emerged adults) are listed from the highest to lowest contribution.

Fe (III) Treatment	p-value	Average Distance to Controls	Taxa
Low	0.072	10.60	Chironomini (L), Tanytarsini (L), Glossosomatidae (L)
Mid-Low	0.011	11.21	Simuliidae (L), Tanytarsini (L), Baetidae (L)
Mid	0.012	17.36	Chironomini (L), Tanytarsini (L), Baetidae (L), Simuliidae (A)
Mid-High	0.01	32.53	Baetidae (L), Tanytarsini (L)
High	0.002	63.09	Baetidae (L), Glossosomatidae (L)

References

Alexander, A. C.; Culp, J. M.; Baird, D. J.; Cessna, A. J., Nutrient–insecticide interactions decouple density-dependent predation pressure in aquatic insects. *Freshwater Biology* 2016, 61, (12), 2090-2101.

Anderson, M.; Gorley, R. N.; Clarke, R. K., Permanova+ for primer: Guide to software and statistical methods. Primer-E Limited: 2008.

Anzecc, A., Australian and New Zealand guidelines for fresh and marine water quality. *Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Canberra* 2000, 1-103.

Bates, D.; Mächler, M.; Bolker, B.; Walker, S., Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823 2014.

Baxter, C. V.; Fausch, K. D.; Carl Saunders, W., Tangled webs: reciprocal flows of invertebrate prey link streams and riparian zones. *Freshwater Biology* 2005, *50*, (2), 201-220.

Beutler, M.; Wiltshire, K. H.; Meyer, B.; Moldaenke, C.; Lüring, C.; Meyerhöfer, M.; Hansen, U.-P.; Dau, H., A fluorometric method for the differentiation of algal populations in vivo and in situ. *Photosynthesis Research* 2002, *72*, (1), 39-53.

Biesinger, K. E.; Christensen, G. M., Effects of various metals on survival, growth, reproduction, and metabolism of *Daphnia magna*. *Journal of the Fisheries Board of Canada* 1972, *29*, (12), 1691-1700.

Blumenshine, S.; Vadeboncoeur, Y.; Lodge, D.; Cottingham, K.; Knight, S., Benthic-pelagic links: responses of benthos to water-column nutrient enrichment. *Journal of the North American Benthological Society* 1997, *16*, (3), 466-479.

Bray, J. P.; Broady, P. A.; Niyogi, D. K.; Harding, J. S., Periphyton communities in New Zealand streams impacted by acid mine drainage. *Marine and Freshwater Research* 2009, 59, (12), 1084-1091.

Brittain, J. E., Life history strategies in Ephemeroptera and Plecoptera. In mayflies and stoneflies: Life histories and biology, *Springer*: 1990; pp 1-12.

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

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. *Environmental Toxicology and Chemistry* 2017, *36*, (2), 285-291.

Cadmus, P.; Brinkman, S. F.; May, M. K., Chronic Toxicity of Ferric Iron for North American Aquatic Organisms: Derivation of a Chronic Water Quality Criterion Using Single Species and Mesocosm Data. *Archives of Environmental Contamination and Toxicology* 2018a, 1-11.

Cadmus, P.; Clements, W. H.; Williamson, J. L.; Ranville, J. F.; Meyer, J. S.; Gutiérrez Ginés, M. J., 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.

Cadmus, P.; Guasch, H.; Herdrich, A. T.; Bonet, B.; Urrea, G.; Clements, W. H., Structural and functional responses of periphyton and macroinvertebrate communities to ferric Fe, Cu, and Zn in stream mesocosms. *Environmental Toxicology and Chemistry* 2018b, *37*, (5), 1320-1329.

Cain, D. J.; Croteau, M.N., Fuller, C. C.; Ringwood, A. H., Dietary uptake of Cu sorbed to hydrous iron oxide is linked to cellular toxicity and feeding inhibition in a benthic grazer. *Environmental Science & Technology* 2016, *50*, (3), 1552-1560.

Camargo, J. A.; Alonso, Á., Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: a global assessment. *Environment International* 2006, *32*, (6), 831-849.

Campero, M.; De Block, M.; Ollevier, F.; Stoks, R., Metamorphosis offsets the link between larval stress, adult asymmetry and individual quality. *Functional Ecology* 2008, *22*, (2), 271-277.

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.; Cadmus, P.; Brinkman, S. F., 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.; Cherry, D. S.; Cairns, J., Structural alterations in aquatic insect communities exposed to copper in laboratory streams. *Environmental Toxicology and Chemistry* 1988, *7*, (9), 715-722.

Dabney, B. L.; Clements, W. H.; Williamson, J. L.; Ranville, J. F., Influence of Metal Contamination and Sediment Deposition on Benthic Invertebrate Colonization at the North Fork Clear Creek Superfund Site, Colorado, USA. *Environmental Science & Technology* 2018, *52*, (12), 7072-7080.

DeNicola, D. M.; Stapleton, M. G., 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.

Douterelo, I.; Perona, E.; Mateo, P., Use of cyanobacteria to assess water quality in running waters. *Environmental Pollution* 2004, *127*, (3), 377-384.

Elbrecht, V.; Beermann, A. J.; Goessler, G.; Neumann, J.; Tollrian, R.; Wagner, R.; Wlecklik, A.; Piggott, J. J.; Matthaei, C. D.; Leese, F., Multiple-stressor effects on stream invertebrates: a mesocosm experiment manipulating nutrients, fine sediment and flow velocity. *Freshwater Biology* 2016, *61*, (4), 362-375.

Fox, J.; Weisberg, S.; Adler, D.; Bates, D.; Baud-Bovy, G.; Ellison, S.; Firth, D.; Friendly, M.; Gorjanc, G.; Graves, S., Package ‘car’. Vienna: R Foundation for Statistical Computing 2012.

Gerhardt, A., Short term toxicity of iron (Fe) and lead (Pb) to the mayfly *Leptophlebia marginata* (L.)(Insecta) in relation to freshwater acidification. *Hydrobiologia* 1994, *284*, (2), 157-168.

Gerhardt, A.; Westermann, F., Effects of precipitations of iron hydroxides on *Leptophlebia marginata* (L.)(Insecta: Ephemeroptera) in the field. *Archiv fur Hydrobiologie*. Stuttgart 1995, *133*, (1), 81-93.

Greig, H. S.; Kratina, P.; Thompson, P. L.; Palen, W. J.; Richardson, J. S.; Shurin, J. B., Warming, eutrophication, and predator loss amplify subsidies between aquatic and terrestrial ecosystems. *Global Change Biology* 2012, *18*, (2), 504-514.

Guasch, H.; Acosta, X. G.; Urrea, G.; Bañeras, L., Changes in the microbial communities along the environmental gradient created by a small Fe spring. *Freshwater Science* 2012, *31*, (2), 599-609.

Iwasaki, Y.; Schmidt, T. S.; Clements, W. H., Quantifying differences in responses of aquatic insects to trace metal exposure in field studies and short-term stream mesocosm experiments. *Environmental Science & Technology* 2018, *52*, (7), 4378-4384.

Jüttner, F.; Wessel, H. P., Isolation of di(hydroxymethyl) dihydroxypyrrolidine from cyanobacteria genus *Cylindrospermum* that effectively inhibits digestive glucosides of aquatic insects and crustacean grazers *Journal of Phycology* 2003, *39*, (1), 26-32.

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.

Kotalik, C. J.; Clements, W.H., Stream mesocosm experiments show significant differences in sensitivity of larval and emerging adults to metals. *Environmental Science & Technology* 2019, doi.org/10.1021 /acs.est.9b00883

Lenth, R., Emmeans: Estimated marginal means, aka least-squares means. R Package Version 2018, *1*, (2).

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

MacDonald, D.; Smorong, D.; Levy, D.; Swain, L.; Caux, P.; Kemper, J., Guidance on the site-specific application of water quality guidelines in Canada: Procedures for deriving numerical water quality objectives. *Report prepared for the Canadian Council of Ministers of the Environment and Environment Canada* 2002.

Magbanua, F. S.; Townsend, C. R.; Hageman, K. J.; Piggott, J. J.; Matthaei, C. D., Individual and combined effects of fine sediment and glyphosate herbicide on invertebrate drift and insect emergence: a stream mesocosm experiment. *Freshwater Science* 2016, *35*, (1), 139-151.

McCormick, P. V.; Cairns, J., Algae as indicators of environmental change. *Journal of Applied Phycology* 1994, (*5-6*), 509-526.

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

Mebane, C. A.; Schmidt, T. S.; Balistrieri, L. S., Larval aquatic insect responses to cadmium and zinc in experimental streams. *Environmental Toxicology and Chemistry* 2017, *36*, (3), 749-762.

Merritt, R. W.; Cummins, K. W.; Berg M.B., An introduction to the aquatic insects of North America. Kendall Hunt: 2004.

Millward, G.; Moore, R., The adsorption of Cu, Mn and Zn by iron oxyhydroxide in model estuarine solutions. *Water Research* 1982, *16*, (6), 981-985.

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

Pascoe, D.; Williams, K. A.; Green, D. W., Chronic toxicity of cadmium to *Chironomus riparius* (meigen)— effects upon larval development and adult emergence. *Hydrobiologia* 1989, *175*, (2), 109-115.

Piggott, J. J.; Salis, R. K.; Lear, G.; Townsend, C. R.; Matthaei, C. D., Climate warming and agricultural stressors interact to determine stream periphyton community composition. *Global Change Biology* 2015, *21*, (1), 206-222.

Poff, N. L.; Olden, J. D.; Vieira, N. K.; Finn, D. S.; Simmons, M. P.; Kondratieff, B. C., Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships. *Journal of the North American Benthological Society* 2006, *25*, (4), 730-755.

Poteat, M. D.; Buchwalter, D. B., Phylogeny and size differentially influence dissolved Cd and Zn bioaccumulation parameters among closely related aquatic insects. *Environmental Science & Technology* 2014, *48*, (9), 5274-5281.

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.

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

Reich, J. K.; Nichols, S. J.; Maher, W. A.; Kefford, B. J., Is metal flocculation from mining activities a previously overlooked mechanism for impairing freshwater ecosystems? *Science of the Total Environment* 2019, *671*, 1108-1115.

Ritz, C.; Baty, F.; Streibig, J. C.; Gerhard, D., Dose-response analysis using R. *PloS one* 2015, *10*, (12), e0146021.

Rogers, H. A.; Schmidt, T. S.; Dabney, B. L.; Hladik, M. L.; Mahler, B. J.; Van Metre, P. C., Bifenthrin causes trophic cascade and altered insect emergence in mesocosms: Implications for small streams. *Environmental Science & Technology* 2016, *50*, (21), 11974-11983.

Sabo, J.; Power, M., River–watershed exchange: effects of riverine subsidies on riparian lizards and their terrestrial prey. *Ecology* 2002, *83*, (7), 1860-1869.

Schmidt, T. S.; Kraus, J. M.; Walters, D. M.; Wanty, R. B., Emergence flux declines disproportionately to larval density along a stream metals gradient. *Environmental Science & Technology* 2013, *47*, (15), 8784-8792.

Soucek, D.; Dickinson, A., 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.

Team, R. C., R: A language and environment for statistical computing. 2013.

US Environmental Protection Agency. 1976. Quality criteria for water. Washington, DC.

US Environmental Protection Agency. 1999. Update of aquatic life ambient water quality criteria for ammonia. EPA 822-R-99-014. Washington, DC.

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

Wellnitz, T.; Sheldon, S., The effects of iron and manganese on diatom colonization in a Vermont stream. *Freshwater Biology* 1995, *34*, (3), 465-470.

Wesner, J. S.; Kraus, J. M.; Schmidt, T. S.; Walters, D. M.; Clements, W. H., Metamorphosis enhances the effects of metal exposure on the mayfly, *Centroptilum triangulifer*. *Environmental Science & Technology* 2014, *48*, (17), 10415-10422.

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). 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* 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 (**Table S4.1**) (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 (**Table S4.3**).

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 (**Figure S4.1**). 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 (**Figures S4.2-S4.5**). 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; **Table S4.6**). Mid-instar 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 (**Table S4.5**). 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**, **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**; **Table S4.6**). 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 (**Tables S4.3**, **S4.6**) 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 our 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 our 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 our 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, our 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 our 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).

Our 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 our experiments. Our 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 C for Figures S4.1 to S4.5 and Tables S4.1 to S4.6.

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.

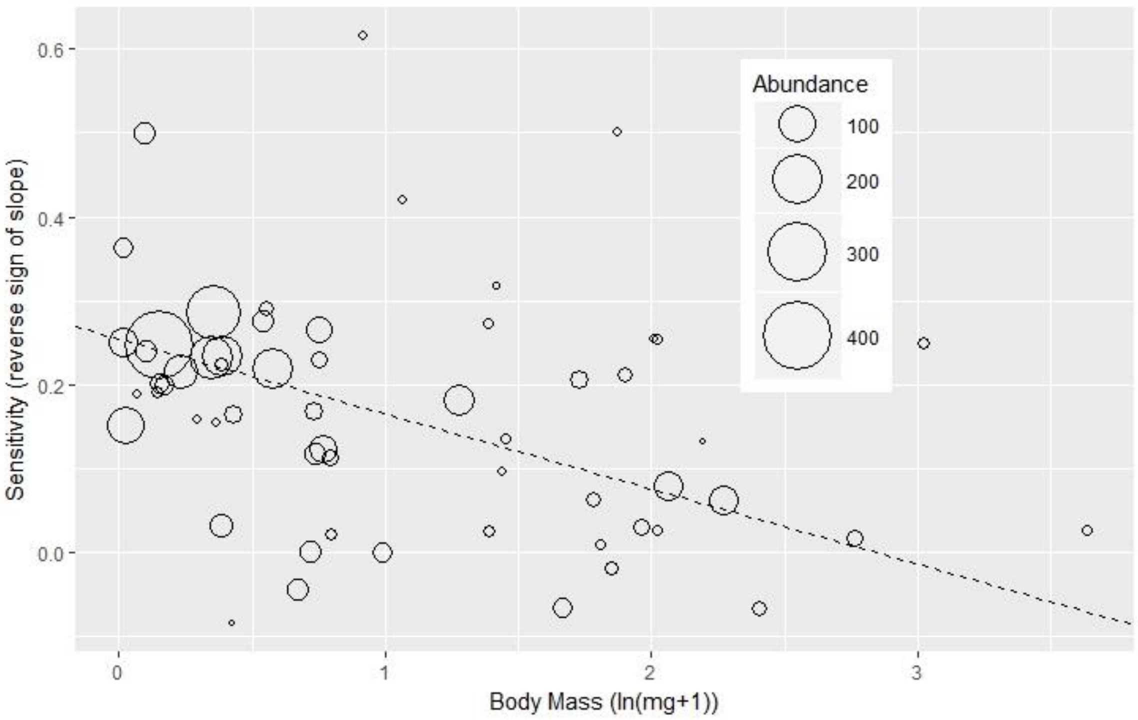

Figure 4.2 Relationship of sensitivity index to body mass of EPT (Ephemeroptera Plecoptera and Trichoptera) larvae in the 2007 (Cu and Cu+Zn) and 2010 (Cu+Zn+ Cd) mesocosm experiments combined. Sensitivity index equals the reverse sign of the slope of ln([Abundance]/[Average Abundance In Controls]+1) 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) $p < 0.0001$.)

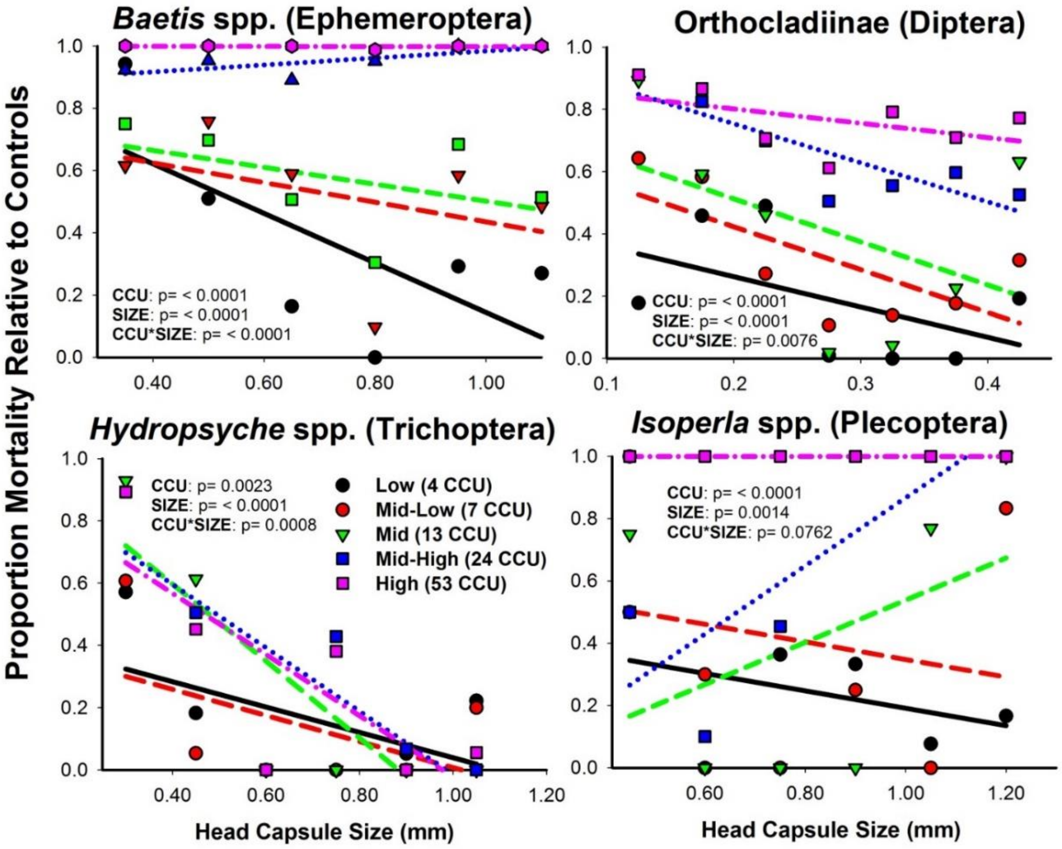

Figure 4.3 Relationships between mortality and head-capsule size in mesocosm experiments (Cu + Zn) from 30 August to 12 September of 2015. Each symbol represents the average proportional mortality of three replicates in each treatment level.

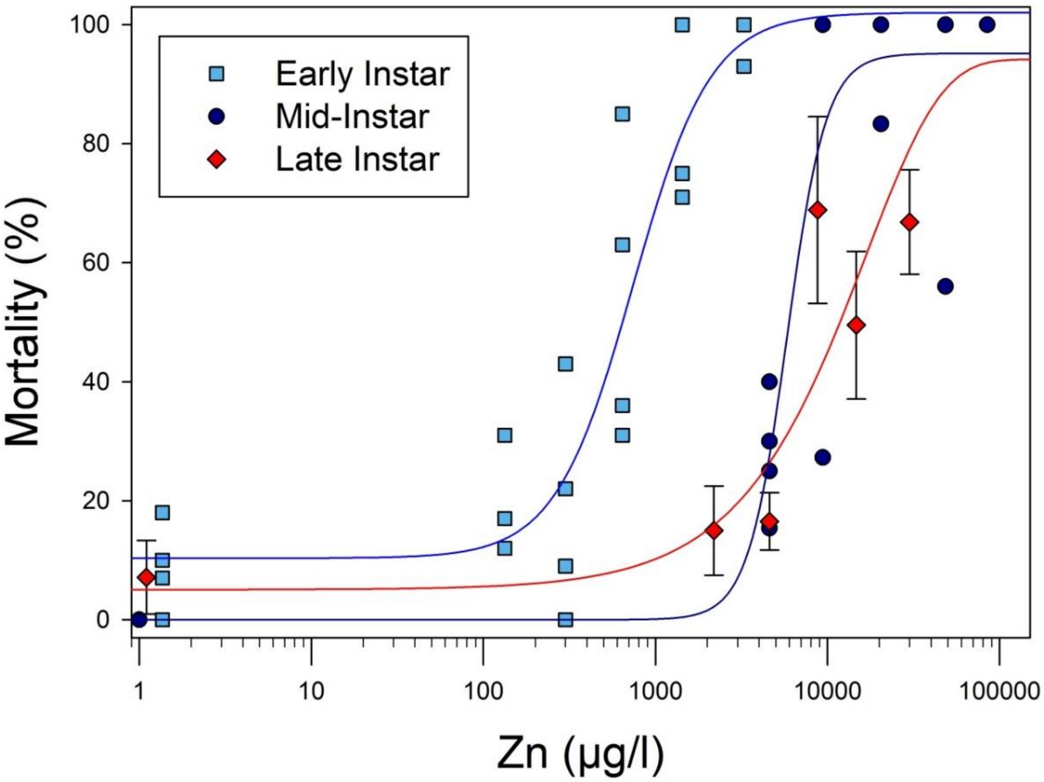

Figure 4.4 Mortality of early instars (light blue squares) and mid-instars (dark blue circles) of *Baetis tricaudatus* after 96 h exposure to Zn. Results from Brinkman and Johnston (2012; red diamonds) are included for a comparison to late instars. (± s.e.; n=4).

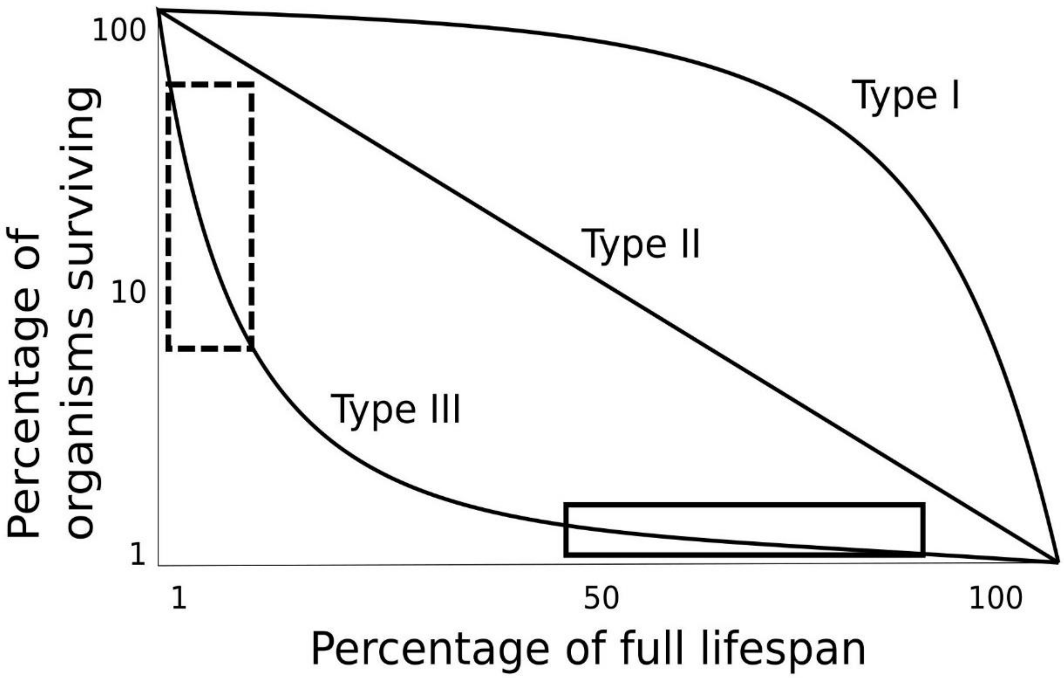

Figure 4.5 Conceptual survivorship curves are commonly used by ecologists to characterize life history traits. Fish and insects generally occupy a Type III curve, whereas longer lived species such as large mammals typically occupy a Type I curve. The dashed box includes early, more sensitive life stages; 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

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

CHARACTERIZING BIOLOGICAL RECOVERY TO ACID MINE DRAINAGE BY RELATING BEFORE-AFTER CONTROL-IMPACT FIELD SURVEYS TO NOVEL EXPERIMENTAL APPROACHES

Introduction

Stream impairment due to acid mine drainage (AMD) results from historical mineral extraction that releases metals and other physical stressors into receiving water bodies (Moore et al. 1991, DeNicola and Stapleton 2002, Niyogi et al. 2002, Clements et al. 2000, Herbst et al. 2018). Depending on the chemical composition of the mining discharge and receiving stream water chemistry, stream communities are exposed to differing mixtures and concentrations of dissolved metals and metal oxides, acidified waters, and fine sediments (DeNicola and Stapleton 2002, Battaglia et al. 2005, Cadmus et al. 2016). Without remediation of stressor inputs, chronic effects on stream life can persist indefinitely depending on watershed geochemistry, receiving stream discharge, and habitat quality (Clements et al. 2010, Schmidt et al. 2010). However, remediation of metal loading through AMD treatment, tailing pile capping or removal, and instream habitat improvements can significantly improve biological integrity (Clements et al. 2010, Mebane et al. 2015, Herbst et al. 2018).

Dissolved metals are directly toxic and accumulated by aquatic organisms primarily through the gill membrane and diet (Hare 1992, Rainbow 2002, Kim et al. 2012). In contrast, at circumneutral pH, metals such as Fe and Mn oxides are indirectly toxic, reducing interstitial space (i.e., benthic habitat) and limiting light penetration required for autochthonous production (McKnight and Feder 1984, Vuori 1995, Wellnitz and Sheldon 1995, Guasch et al. 2012).

Additionally, increased fine sediment deposition resulting from legacy placer mining and overburden from mineral extraction, as well as general stream habitat degradation, result in physical habitat impairment at AMD-impaired sites (Cadmus et al. 2016, Dabney et al. 2018).

Evaluating biological recovery after AMD remediation in the field is essential to assess treatment effectiveness, but various factors affect our ability to associate AMD remediation with biological response. For example, pretreatment biological and chemical data are often lacking, and natural variation in background water chemistry and flow regime influence rates of biological recovery. In circumstances when timing and location of stressor mitigation are known, and pretreatment data are available, a Before-After Control-Impact (BACI) field study design can be implemented (Hulbert 1984, Stewart-Oaten et al. 1986, Underwood 1992). The BACI study design is based on the assumption that an anthropogenic disturbance (i.e. impact) will cause a different rate of change (i.e., recovery) from before, to after, remediation compared to the natural variation observed at control sites. By accounting for natural variation at control sites while obtaining samples before and after stressor mitigation, we improve our ability to causally associate treatment (e.g., reduced in stream metal concentrations) to biological response (Underwood 1992).

Field data are often affected by extraneous variables that have the potential to influence spatiotemporal trends in recovery (Buchwalter et al. 2017); however, other ecotoxicological approaches can supplement BACI models to isolate specific stressors. For example, stream mesocosm and field experiments can integrate naturally occurring aquatic communities that are more representative of the assemblages of organisms (e.g., NFCC) responding to impacts in the field, and with improved control and replication. These experimental approaches at varying spatiotemporal extents coupled with field surveys provide the most comprehensive approach to predicting and assessing AMD remediation effectiveness (Clements 2004, Cadmus et al. 2016, Buchwalter et al. 2017).

North Fork of Clear Creek (NFCC) is an AMD-impacted montane stream located in Gilpin County in Central Colorado, USA, with an approximate 1036 km^2 watershed. The NFCC is a high-gradient montane stream with headwaters originating just east of the Continental Divide and continuing through the town of Black Hawk before reaching the main stem of Clear Creek 10.6 km downstream (**Figure 5.1**). The flow regime is primarily snow-melt driven, but isolated convective storms during the summer monsoon season can rapidly increase flow resulting in a "flashy" hydrograph (**Figure S5.1**). In addition, discharges from abandoned mines and drainage tunnels contribute flow to NFCC.

Clear Creek is a major drinking water supply for over 500,000 people in the Denver Metropolitan area and used for numerous recreational activities (e.g., fishing, boating, swimming), resulting in both human and environmental health concerns. The middle reach of NFCC was placed on the U.S. Environmental Protection Agency (USEPA) National Priority List (Superfund) site in 1983 due to significantly elevated concentrations of metals related to mining operations for gold and silver deposits in the Central City-Black Hawk area from 1860 through the 1970s. In particular, two major legacy mine drainages, Gregory Incline and National Tunnel, entered the NFCC in Black Hawk, resulting in decreased pH and alkalinity, and increased water hardness, dissolved metals (i.e., copper, zinc, cadmium) and metal oxides (iron, manganese). These point-source AMD inputs extensively impaired the NFCC, with biological surveys showing sections of NFCC completely devoid of aquatic life, and others with inhibited primary production, low densities of benthic macroinvertebrates, and no sustained fish populations. In summer 2017, the USEPA and the Colorado Department of Public Health and Environment completed the North Clear Creek Water Treatment Plant (NCCWTP) to treat the two major AMD point-sources using high-density lime-sludge treatment to remove metals by precipitation and flocculation before the treated effluent is returned to NFCC. Remediation on NFCC provided a unique opportunity to assess biological recovery because the majority of the metal loading was directly treated, resulting in rapid improvements in water quality.

To characterize the biological recovery of NFCC to improvements in water quality related to the NCCWTP, we conducted biological monitoring, stream mesocosm and field experiments. Biological monitoring (biomonitoring) data of benthic algae and macroinvertebrate communities, and emerging aquatic insects were gathered over an 8 year time period bracketing before and after the treatment plan installation. During this time period, biannual biomonitoring was conducted at sites located above and below the AMD inputs, with increased sampling intensity in summer 2017 and 2018 following NCCWTP treatment. During summer sampling events, emergence nets were deployed to capture emerged aquatic insects to estimate aquatic subsidy export.

Stream mesocosms were used to expose naturally colonized benthic communities obtained from the reference location above the AMD inputs in NFCC to an environmentally relevant gradient of Cu and Zn (the predominant metals of concern). This experiment was designed to determine differences in metal sensitivity among benthic macroinvertebrates indigenous to NFCC. Next, rates of algal colonization among differing substrata tested if algal attachment and growth were affected by the presence of iron scales “plated” on rock surfaces. Lastly, field experiments estimated benthic macroinvertebrate colonization rates in response to AMD treatment. Because habitat quality is impaired in NFCC due to fine sediments and stream habitat degradation, deployment of trays filled with clean substrate reduces the potential cofounding effect of habitat quality, and improves our ability to separate the effects of physical and dissolved metal stressors. The overall goal of this research was to apply varying descriptive and experimental approaches to predict and characterize biological recovery in NFCC following AMD treatment.

Materials and Methods

Study Site

The North Fork Clear Creek (NFCC) water chemistry below the two main acid mine drainages (AMD) changed dramatically following the diversion and treatment of these inputs to the North Clear Creek Water Treatment Plant (NCCWTP) beginning in June 2017. Prior to treatment, water chemistry resulting from these inputs was characterized as: circumneutral pH (6.5-7.5), low alkalinity (< 15 mg/L as $CaCO_3$), moderate conductivity (300-600 µS/cm), with low concentrations of dissolved organic carbon (1-3 mg C/L), and high concentrations of total Fe (2000-20,000 µg/L), Zn (500-1500 µg/L), Cu (10-50 µg/L), and $CaCO_3$ (200-300 mg/L). After the NCCWTP began treatment, water quality improved, though differences in water chemistry were observed upstream of the NCCWTP compared to downstream of treated effluent discharge (**Table 5.1**). Concentrations of most metals decreased (particularly iron); however, metal concentrations fluctuated throughout the year due mainly to increased flow associated with runoff and residual loading of metals not captured by the NCCWTP through the town of Black Hawk (Meyer et al. 2018).

Field Biomonitoring

Benthic Macroinvertebrates and Emerged Adults

Benthic macroinvertebrate communities were sampled above and below the AMD inputs in NFCC from 2011 through 2018. Biannual sampling (typically early-May and mid-October) was conducted at seven sites from 2011 through 2016 prior to AMD treatment. After NCCTP began treating the AMD effluent, additional monthly sampling was conducted in the summers of 2017 and 2018 to characterize the initial benthic macroinvertebrate response to AMD treatment. Four sites (i.e., Pumphouse, Riviera, Rail-less, and Gauge Station) were sampled in the summer of 2017, and these same sites, in addition to Above Russel Gulch, were sampled in summer 2018. The site Above Russell Gulch was added to summer sampling because biannual sampling data showed that this stream segment was recovering differently from other upstream and downstream sites. Given the Before-After Control-Impact (BACI) design of our study, and our desire to focus specifically on NFCC, we chose to exclude the biomonitoring results from the two sites located on the main-stem of Clear Creek (i.e., Confluence and CC main). The BACI site designation is as follows: “Reference” (Upper Reference, Pumphouse), “Impact” (Riviera, Rail-less, Above Russel Gulch), and “Recovery” (Gauge Station) (**Table 5.2**).

Benthic macroinvertebrate communities were quantitatively sampled using a modified Hess sampler (0.1 m^2). Replicate samples (n= 5) were taken at each site in targeted riffle sections (< 0.5 m depth). Samples were filtered through a 350-µm sieve in the field and preserved in 80% ethanol. In the laboratory, benthic samples were subsampled to a 300-count from detrital material, and identified to the lowest taxonomic resolution possible, typically to genus (except for chironomids that were identified to subfamily). In addition to benthic surveys, emerged aquatic insect communities were sampled in the summer of 2017 and 2018. Pyramid-shaped floating emergence traps (n = 4) were deployed on top of pools and slack water to capture emerging aquatic insects. The emergence traps were 1 m^2 plastic structures that float on the water's surface (Cadmus et al. 2016b) and are covered with 500-*u*m mesh with an external collection bottle filled with 80% ethanol. External sample bottles were collected after 3 d to allow sufficient time to collect emerging aquatic insect adults. In the laboratory, emerged adults were sorted and identified to family, except for chironomids that were identified to subfamily (Merrit et al. 2008).

Benthic Algae

Benthic algal biomass and community composition were measured from 2015 through summer 2019. Biannual samples were gathered from 2015-2019, and monthly sampling was conducted in summer 2017 and 2018. Measurements were taken using a BenthoTorch (bbe Moldaenke GmbH, Germany), an *in situ* pulse-amplitude modulated fluorometer that measures benthic algae biomass as chlorophyll *a* and estimates the relative contribution of diatoms, green algae, and cyanobacteria pigment complexes to total algal biomass (Beutler et al. 2002). Algal measurements (*n* = 15-25) were taken on cobble size substrate located in riffle sections at each site. We chose to increase the number of measurements taken with the BenthoTorch after 2016 because we observed high within-site natural variation in benthic algal biomass.

Field Experiment

To characterize the rates of benthic macroinvertebrate colonization among sites following AMD treatment, we conducted a substrate colonization experiment in NFCC in the summer of 2017 and 2018. Plastic trays (10 x 10 x 6 cm) were filled with clean pebble and small cobble substrate and placed in riffle sections at four sites (Reference, Riviera, Rail-less, Gauge Station) in the summer of 2017 and a 5th site (Above Russel Gulch) in 2018. Colonization trays contained clean substrate with optimal interstitial space that provided ideal habitat for benthic colonization, allowing for comparisons to benthic samples collected on natural substrate (Clements et al. 1989). Trays (n = 24 per site) were deployed and sampled on days 15, 30, 45, and 60 to estimate rates of macroinvertebrate colonization. For each sampling event, trays (n= 6) were randomly selected, with paired trays (n = 2) treated as a single experimental unit. Pairs of colonization trays were rinsed through a 350-µm sieve, and organisms and detrital material were preserved in 80% ethanol. All larvae were enumerated to the lowest taxonomic resolution possible, typically to genus, except for chironomids that were identified to subfamily or tribe.

Stream Mesocosm Experiments

North Fork Clear Creek Benthic Community Exposure

A mesocosm experiment was conducted at the Stream Research Laboratory (SRL) located at Colorado State University Foothills Campus, Fort Collins, Colorado, USA to test the responses of native communities obtained from NFCC to dissolved metals exposure (Cu and Zn). Naturally colonized stream benthic communities were collected from NFCC (Pumphouse) using a colonization technique described by Clements et al. (1988). Plastic trays (10 x 10 x 6 cm) filled with pebble and small cobble were deployed for 30 d and colonized by a diverse benthic assemblage. Five trays were randomly placed into each of 18 insulated coolers, transported to the SRL, transferred to 20-L stream mesocosms, and assigned metal treatments.

Water quality in the SRL is representative of cold water Rocky Mountain streams with cool temperatures (10-15 ºC); low hardness (30-35 mg/L $CaCO_3$), alkalinity (25-30 mg/L $CaCO_3$), conductivity (60-90 µS/cm), and concentrations of dissolved organic carbon (2.5-3.0

mg/L); and circumneutral pH (7.0-7.8). Current in the 18 mesocosms was driven by submersible water pumps (EcoPlus®) at a rate of 2233 L/h (i.e., 1.9 hydraulic circulations/min). Each mesocosm was covered with insect mesh netting (200-µm) to capture emerged adults. Emerged adults were aspirated from the netting every evening between 17:00 and 19:00 and preserved in 80% ethanol. To prevent emigration by drifting invertebrates, standpipes in the stream mesocosms were covered with fine mesh.

Each mesocosm (76 x 46 x 12 cm^2) received dilution water at 1.0 L/min while peristaltic pumps delivered 10 mL/min metal salt solutions. The NFCC community was exposed to a gradient of Cu and Zn at approximately 1:15, with three replicate mesocosms per treatment ranging from 10 to 100 µg/L Cu, and 15 to 1500 µg/L Zn. These mass ratios were chosen to represent the Cu and Zn mixtures observed in NFCC (Meyers et al. 2018). To confirm dissolved Cu and Zn concentrations, water samples (15 mL) were taken from each mesocosm during the experiment. Water samples were filtered through a 0.45-µm filter, acidified to a pH < 2.0 using analytical grade nitric acid, and analyzed using flame or furnace atomic absorption spectrophotometry. Cumulative criterion units (CCU), defined as the ratio of the measured dissolved metal concentrations to the United States Environmental Protection Agency (U.S. EPA) hardness-adjusted criterion value and summed for each metal, were used to quantify the metal mixture of Cu and Zn as one value. We chose to use hardness-adjusted criteria because previous experiments have found that the water chemistry of the dilution water in the SRL is highly consistent. In addition, the U.S. EPA biotic ligand model (BLM) and hardness-based criteria have been shown to generate very similar values based on water quality in our stream mesocosms (Clements et al. 2013). Because the BLM requires additional chemical analyses (e.g., DOC) that were unavailable for these experiments, we chose to use the hardness-based criteria to benchmark metal mixture concentrations.

Temperature, pH, conductivity, were measured every 3 d (YSI Pro1030). Alkalinity and hardness were measured by titration (USEPA methods 200.7; 310.7). After 14 d, colonization trays were rinsed through a 350-µm sieve, and organisms and detrital material were preserved in 80% ethanol. All larvae were enumerated to the lowest taxonomic resolution possible, typically to genus, except for chironomids that were identified to subfamily or tribe. Emerged adult insects were enumerated and identified to family or subfamily.

Algal Colonization to Metal-Plated Substrata

To test the ability of algae to colonize metal-plated substrate, we deployed eighteen 25-cm^2 unglazed ceramic tiles glued to cinder bricks (using silicon) at the Rail-less site in a riffle section before the NCCTP became operational. The tiles were allowed to plate *in situ* with metal oxides (primarily Fe and Mn) for approximately one year (July 2016-August 2017). Metal-plated tiles were transferred to the SRL to estimate rates of algal colonization. Because substrata color has been shown to influence reflectance and photosynthetic performance (Carpentier et al. 2013), we also wanted to test the influence of color on algal colonization. The substrate color in NFCC prior to AMD treatment was a mixture of blood-orange, mustard, and rust-red, whereas substrate from the reference sites was generally dark to light grey. To match these colors using standardized substrata, we deployed white, cinnamon, and pumpkin colored ceramic tiles with the metal-plated tiles.

Each stream mesocosm (*n* = 6) received 3 tiles of each respective tile type, except for the white tiles that received 6, and were randomly assigned a position in a 4 x 3 grid. Each mesocosm received dilution water at 1.0 L/min, and current velocity was provided by paddle wheels. Algal composition and biomass were estimated using a BenthoTorch. To prevent sloughing of algae on tiles while measurements were taken, a tile adapter was developed to fit around the edges of each tile to reduce disturbance (**Figure S5.2**). Measurements were taken at 0, 6, 11, 16, 21, 26, and 31 d.

Statistical Analyses

All univariate statistical analyses were conducted using R statistical computing (R Core Team v3.5.1). To quantify the responses of benthic communities to AMD treatment, we evaluated changes before and after the NCCTP began operation. We chose to examine benthic macroinvertebrate metrics of total abundance, taxa richness, Ephemeroptera-Plecoptera-Trichoptera (EPT) abundance, and EPT taxa richness. These metrics are known to respond to anthropogenic stressors generally (Lenat and Barbour 1994). For algal communities, we examined cyanobacteria, diatoms, green algae, and total chlorophyll-a. Two-factor analysis of variance (ANOVA) tested mean differences before and after treatment ("LM" function; package 'car'; Fox et al. 2012). The factors in these models were site (i.e., Reference, Impact, Recovery) and time (i.e., before and after AMD treatment). A significant interaction term indicates that a metric changed among sites after AMD treatment. Multiple comparisons (Tukey's method) were used to compare sites before and after treatment based on time-interaction coefficient contrasts obtained from "estimated marginal means of linear trends" function from package 'emmeans' (Lenth 2018). For analyzing aquatic insect emergence, linear mixed-effects models (package 'lme'; West et al. 2014) tested significant changes in abundance and biomass from 2017 to 2018, with a random effect of date (month of sampling). We chose a nested design to account for the seasonality of emergence because rates of emergence decrease through the summer into fall, which strongly affects mean comparisons among sites over time.

To analyze benthic macroinvertebrate colonization from 2017 and 2018, total abundance and taxa richness, as well as biomass were calculated. In addition, dominant taxa (*n* >10) abundance represented between both experiments were calculated for Baetidae, *Hydropsyche* spp., Chironomidae, and *Simulium* sp.. We applied two-factor repeated measures ANOVA analyses (function “lmer,” package ‘lme4”; Bates et al. 2014) to test for site differences in colonization metrics between year 1 and year 2 post-treatment operation. The factors in these models were site location and year. A significant interaction term in this model indicates that a colonization metric changed among sites between years. We chose a covariance structure of compound symmetry, which assumes constant correlation between observations on the same sample replicate because rates of colonization are additive over the 60 day experiment and sample observations are correlated with the previous collection event. *Post hoc* colonization metric comparisons (Tukey’s method) among sites were compared between years.

To analyze responses of dominant (*n* >10) benthic macroinvertebrate taxa in the benthic community stream mesocosm experiment, linear models tested for significant concentration-response relationships between metals (Cu and Zn) and macroinvertebrate abundance. For significant linear models, we estimated EC20 values (effect concentration that reduced abundance by 20%, compared to controls) with two-parameter log-logistic concentration-response curves (R package ‘drc’; Ritz et al. 2015). EC50 values were also calculated to compare to EC20 estimates (**Figure S5.1**); however, we chose to report EC20s based on the USEPA risk-management decision to derive Final Chronic Values (USEPA 1999). Lastly, for the algal colonization mesocosm experiment, comparisons among tile treatments were analyzed using two-factor repeated measures ANOVA. For each stream mesocosm, tile treatments were averaged and pairwise comparison of the time-interaction coefficients (obtained from 'emmeans') tested for significant differences in algal colonization biomass among treatments at each time point. All data were log- transformed to meet the assumptions of parametric statistics and to improve model fit.

All multivariate analyses were conducted using PRIMER-e v7 + PERMANOVA (Quest Research Limited, Cambridge, United Kingdom; Anderson et al. 2016). To characterize changes in benthic community structure before and after AMD treatment, all benthic macroinvertebrate monitoring data were compiled from 2011 through 2018. To bracket before and after AMD treatment, data were divided from fall 2011 through spring 2017 (before treatment) and from summer 2017 though fall 2018 (after treatment). For the benthic macroinvertebrate community colonization experiments, alterations in community composition were examined in years 1 and 2 (2017 and 2018) following treatment. For all multivariate analyses, rare taxa were removed (less than 10 individuals among all sample observations). Due to lower taxonomic resolution for Chironomidae for samples obtained from 2011 through 2012, we chose to group chironomid subfamilies for all benthic samples obtained before the NCCWTP became operational (fall 2011-spring 2017). However, because of consistency in the taxonomic resolution of Chironomidae subfamilies in benthic samples obtained after AMD treatment, subfamilies were included in these multivariate analyses (summer 2017- fall 2018). All abundances were log (x+1) transformed and Euclidean distance matrices were calculated. Nonmetric-multidimensional scaling (NMDS) plots visualized differences among sites before and after treatment.

To test for significant alterations in community structure before and after NCCTP operations (2011-2018), we ran a two-factor (site and time) permutational multivariate analysis of variance (PERMANOVA) with 999 permutations. Due to differences in the taxonomic resolution of chironomids among years, subfamily data were grouped to Chironomidae for all samples. Abundances were log (x+1) transformed and Euclidean distance matrices were calculated. Factors were assigned for site (Reference, Above AMD Treatment, Below AMD Treatment, Above Russel Gulch, Upstream of CC Confluence) and time (Before and After NCCTP operations). A significant interaction term indicates that community structure among sites significantly changed after AMD treatment. For the macroinvertebrate colonization experiments, the same biomonitoring sites were used, with time expressed as “year 1” and “year 2.” For both the benthic macroinvertebrate monitoring data and colonization experiment, similarity percentages (SIMPER) analyses were performed on the log-transformed datasets to determine which taxa accounted for greater than 50% of benthic community dissimilarity using paired site comparisons. In addition, the SIMPER analyses determined the average squared distance for each pairwise comparison, with larger values indicating greater site dissimilarity.

Results

The North Clear Creek Water Treatment Plant (NCCWTP) decreased the metal loading into North Fork Clear Creek (NFCC) (**Table 5.1**) by treating the two major sources of Acid Mine Drainage (AMD) that formerly entered NFCC through Black Hawk. However, untreated residual influxes of metals (mainly through the town of Black Hawk continue to contribute to seasonal variations in dissolved metal concentrations that frequently exceed water quality criteria (Meyers et al. 2018). Prior to completion of the NCCWTP, benthic macroinvertebrates were extirpated at the Impact sites downstream through Above Russell Gulch, with some improvement historically observed above the Clear Creek Confluence due to the treated wastewater effluent that discharges above this site (reduced metal bioavailability).

Field Biomonitoring

Total abundance of benthic macroinvertebrates and taxa richness significantly ($p < 0.10$) increased following the NCCWTP operation, with highly significant interaction terms indicating that this increase in abundance significantly differed after AMD treatment (**Figure 5.2**). Average total abundance at the Impact sites prior to AMD treatment was 5 (0.95 ± s.e.) individuals per 0.1 m^2, compared to 76 (23.56 ± s.e.) individuals from July 2017 through October 2018. The abundance of EPT taxa followed similar trends, with large increases in EPT abundance following treatment. Increases in larval abundances were generally associated with greater abundance of emerged aquatic insect adults (**Figure 5.3**); however, our ability to detect a significant interaction was limited due to the high variability among sites. In contrast, emerged adult biomass significantly increased from 2017 to 2018, but again high variability among sites limited our ability to separate site differences between years (**Figure 5.3**). Taxa richness and EPT richness significantly increased after AMD treatment; however, benthic macroinvertebrate richness at the Reference stations was still approximately twice the average number of taxa compared to the Impact and Recovery sites following treatment. Notably, taxa richness significantly increased at both the Impact and Recovery sites, but for EPT richness the Recovery stations did not significantly increase in response to AMD treatment whereas the Impact site significantly increased.

Prior to NCCWTP operations, results from the benthic algal surveys generally showed greater biomass among all algal metrics (except for green algae) at reference sites compared to downstream sites (**Figure 5.4**). Following treatment, significant increases in total chlorophyll-a cyanobacteria, and diatoms were observed at Impact sites, but significant changes in these metrics were not observed at the Recovery site. Green algae biomass showed no change after AMD treatment, and was generally highly variable.

Before AMD treatment (Fall 2011-Spring 2017) the non-metric multidimensional scaling (NMDS) analyses based on macroinvertebrate abundances showed clear site separation from the Reference observations, while all other downstream sites were clustered together (**Figure 5.5**). After AMD treatment (Summer 2017– Fall 2018), separation of downstream sites from reference sites remained distinct; however, comparatively greater separation among downstream sites were observed. This greater separation suggests that the community composition of downstream sites is changing, but that patterns of site-specific clustering indicates that these communities are recovering differently, particularly at the site upstream of the Clear Creek confluence. Community composition of Riviera (i.e. Above AMD treatment) generally appears to be closest in multivariate space to the Reference observations.

Permutational MANOVA results for benthic sampling from 2011 through 2018 showed a highly significant interaction term, indicating that benthic community structure significantly changed after NCCWTP operation (**Table 5.3**). Pairwise comparisons of the similarity percentages (SIMPER) of taxa most contributing to compositional differences in benthic community structure between the Reference and downstream sites before AMD treatment were *Baetis* spp., Chironomidae, *Heterlimnius* sp., *Prostoia*/*Podomosta* spp., and *Rhyacophila* spp. (**Table S5.1**). The taxa most contributing to differences among downstream sites were

Chironomidae, *Baetis* spp., *Rhyacophila* spp., and *Arctopsyche* sp. (**Table S5.1**). After AMD treatment, contributing taxa to differences between Reference and downstream sites were generally similar to those observed before treatment, but with the additions of *Epeorus* sp., and *Zapada* sp. and less contribution from Chironomidae (**Table S5.2**). Lastly, pairwise comparisons among downstream sites showed *Baetis* spp., Orthocladinae, *Hydropsyche* sp., and *Simulium* sp. most contributing to compositional differences in community structure (**Table S5.3**). In general, the distance among downstream sites compared to the Reference sites decreased after AMD treatment, with the exception of Gauge station (i.e., Upstream of Clear Creek Confluence) that actually became more dissimilar after treatment.

Field Experiments

The patterns of benthic macroinvertebrate colonization were markedly different between summer 2017 and summer 2018. The NMDS plot in 2017 (first summer after treatment) showed large separation between the Reference and three downstream sites (**Figure 5.6**), with sites located above and below the NCCWTP clustered together, but with separation from Gauge (the site upstream of the Clear Creek confluence). In contrast, the NMDS plot in 2018 (second summer after treatment) showed more similarity among downstream sites compared to Reference sites, but with distinct separation among the four downstream sites (**Figure 5.6)**. Permutational MANOVA results were highly significant, with a significant interaction term indicating that colonization differed from 2017 to 2018 (**Table 5.4**). In the first summer, the taxa strongly contributing to community dissimilarity between the Reference and downstream sites were *Baetis* spp., *Taenionema* sp., *Zapada* spp., and *Prostoia/Podmosta* spp. (**Table S5.4**). Among downstream sites, taxa strongly contributing to community differences among sites were

Orthocladinae, Chloroperlidae, and *Baetis* spp. (**Table S5.4**). In the second summer, *Zapada* sp., *Epeorus* spp., *Diphetor* sp., *Baetis* spp., *Hydropsyche* spp., and *Simulium* were most frequently contributing to community differences between the Reference site and downstream locations (**Table S5.5**). Taxa contributing to differences among downstream sites in 2018 were similar to 2017, with the addition of mainly *Hydropsyche* spp. and *Simulium* sp. (**Table S5.6**).

Total abundance, taxa richness, and biomass metrics used to characterize benthic macroinvertebrate colonization among sites significantly differed between year 1 and year 2 (**Figure 5.7**). Multiple comparisons showed no change in these respective metrics at the Reference site between years, but statistically significant increases were observed at the Impact sites from Year 1 to 2. Abundance at the Recovery site significantly increased from 2017 to 2018, with significantly greater abundance than the Reference site, but taxa richness and biomass did not significantly increase. Among the dominant taxa (Baetidae, *Hydropscyhe* spp., Chironomidae, and *Simulium* spp.), all repeated-measures interaction terms were significant, indicating that the abundance of these taxa significantly changed from 2017 to 2018 (**Figure 5.8**). Densities of dominant taxa at the Reference site between years were generally similar with the exception of Chironomidae, which had much higher initial colonization density in Year 2. Significant increases in the abundance of dominant taxa at the Impact site were observed for all taxa except for *Simulium* sp.. Lastly, the Recovery site showed significant increases in abundance of *Hydropsyche* spp. and *Simulium* sp., but Baetidae and Chironomidae remained unchanged.

Stream Mesocosm Experiments

North Fork Clear Creek Benthic Community Exposure

Measured concentrations of Cu and Zn generally approximately target values (**Table S5.7**), with measured concentrations of Cu and Zn ranging from 9.3-95.6 µg/L and 79.7-1450.5 µg/L, respectively. Water hardness and alkalinity were approximately 30 mg/L $CaCO_3$ (0.74 ± s.d.) and 29 mg/L $CaCO_3$ (0.80 ± s.d.), respectively. The U.S. EPA acute hardness-adjusted criterion values for Cu (U.S. EPA, 1985) was 4.3 µg/L and for Zn (U.S. EPA, 2002) 42.2 µg/L. Temperature, pH, and conductivity were consistent among treatments and between experiments (**Table S5.8**).

The North Fork Clear Creek benthic community was diverse, with an average of 536 individuals (41 ± s.e.) and 23.6 (0.88 ± s.e.) taxa in control treatments. The dominant larvae ($n > 10$) represented in the benthic exposure were *Baetis* spp., *Ephemerella* sp., *Rhyacophila*, Tanytarsini, Orthocladiinae, and *Epeorus* spp. Among adult aquatic insects, the dominant taxa emerging during our 14 d exposure were *Baetis* spp. and Chironomidae. Estimated EC20 values ranged from 1.08 CCU for *Zapada* sp. to 14.89 CCU for total adult aquatic insect emergence (**Table 5.5**).

Algal Colonization to Metal-Plated Substrata

Water quality in the stream mesocosms over the 31 d experiment was highly consistent, with pH 7.35 (0.09 ± s.d.), conductivity 75.6 (0.98 ± s.d.) µS/cm, and temperature 12.03 (0.54 ± s.d.) °C. Significant differences in algal colonization among tiles was observed, though the repeated-measures interaction term was non-significant, indicating that the tile treatments did not significantly differ in algal biomass (chlorophyll-*a*) over the 31d colonization assessment

(**Figure 5.9**). Pairwise comparisons on the time-interaction coefficients showed significant differences among tile types at days 11 and 16, with statistically higher algal biomass observed on the metal-plated tiles, but non-significant differences among treatments for the remainder of the experiment.

Discussion

Benthic communities in the North Fork Clear Creek responded rapidly following treatment of the two major acid mine drainage (AMD) inputs, Gregory Incline and National Tunnel drainages, by the North Clear Creek Water Treatment Plant (NCCWTP). While residual metal loading continued in the watershed primarily through non-point sources (i.e., tailing piles, ground water seepages), treatment of Gregory Incline and National Tunnel resulted in significant biological recovery one year following continuous metals treatment. Benthic algal biomass increased downstream, and both colonizing and *in situ* benthic invertebrates showed increased abundance and taxa richness downstream of the NCCWTP. Importantly, increased benthic production at downstream sites was associated with increased aquatic biomass export (i.e., terrestrial subsidy) by aquatic insect adult emergence.

Evaluation of benthic macroinvertebrate colonization experiments showed dramatic increases in density and diversity approximately one year after continuous AMD treatment; however, we observed compositional differences between the colonizing and *in situ* benthic assemblages. For example, NMDS plots clearly show site separation in community composition on the colonization trays in summer 2018, but benthic surveys show considerably more site clustering and less separation among downstream sites (**Figure S5.3**). Evaluation of the SIMPER results generally showed similar taxa contributing to site dissimilarity between biomonitoring and colonization studies; therefore, we believe differences in community composition were caused by habitat quality. The colonization trays represented ideal habitat with optimal amounts of interstitial space, whereas our field samples were collected on degraded benthic substrate. This concurrent evaluation of field and experimental approaches provided comparisons of the relative effects of substrate and water quality. Previous studies have confirmed the impacts of impaired substrate quality in mining-impacted streams (Cadmus et al. 2016, Dabney et al. 2018, Kotalik et al. 2019). The principal focus of most AMD remediation projects is to improve water quality, but habitat quality should also receive equal attention.

Characterizing and predicting the relative impacts of physical and chemical stressors on AMD streams is particularly challenging, in part because their co-occurrence varies through space and time. Cadmus et al. (2016) estimated recovery potential of dominant taxa in the NFCC based on drift propensity and mesocosm and field experiments that evaluated sensitivity to physical and chemical AMD stressors. The authors predicted that Rhyacophilidae, Hydropsychidae, and Chloroperlidae would be most likely to recolonize NCCC after remediation. Given results from our benthic field surveys, Baetidae, Hydropsychidae, and Chironomidae appear to be the dominant taxa at downstream sites. Rhyacophilidae and Chloroperlidae continued to be extirpated downstream, and we hypothesize that because both Rhyacophilidae and Chloroperlidae are predators, their absence downstream may be due to insufficient prey availability (Pomeranz et al. 2019). Significant downstream colonization by Chironomidae may be due to life stage sensitivity, habitat preferences by some taxa for fine sediment, and aerial dispersal ability as winged terrestrial adults (Oliver 1971). We have previously shown that aquatic insect development influences sensitivity to metals, and developmentally mature chironomid adults are particularly tolerant to exposure (Kotalik and

Clements 2019). These results highlight the difficulty predicting biological recovery following AMD remediation.

The greatest differences in benthic macroinvertebrate community structure among Impact and Recovery sites was observed directly below the NCCWTP. This result was surprising considering the close proximity to the upstream reference locations that represent a robust colonization source and because the site immediately upstream (Riviera above the AMD treatment plant) displayed the greatest benthic community similarity to the Reference sites. We hypothesize that this difference is the result of water chemistry changes associated with the treated effluent discharge from NCCWTP, which has much higher hardness and conductivity compared to reference water quality (**Table 5.1**). Mesocosm and field studies have shown that some aquatic insects, particularly those that have evolved in low conductivity waters, are sensitive to elevated conductivity. For example, significant alterations in community composition of aquatic insects in the field has been observed at conductivity between 300-500 µS/cm (Pond et al. 2008, Cormier et al. 2013), and EC20 estimates for mayfly abundance obtained from community mesocosm exposures ranged from 280-1300 µS/cm (Clements and Kotalik 2016). This project provided a unique opportunity to study the direct impact of the treated effluent because prior to treatment, the 2 impacted sites (Riviera and Rail-less) experienced the same metals exposure regime, but following treatment Riviera received only upstream "reference" water while Rail-less received a mixture of reference water and treated effluent. Our results suggest that diversion and effluent by NCCWTP significantly improved overall water quality and aquatic life in NFCC, but aquatic life directly downstream of the treated effluent remains degraded because of elevated conductivity.

Downstream community composition and taxa richness were very different from the Reference sites through 2018, and many upstream aquatic insect taxa continued to be extirpated downstream. Stream mesocosm results showed that *Zapada* sp., *Baetis* spp., and chironomids were most sensitive to Cu and Zn exposure. In support of these findings, *Zapada* sp. continued to be extirpated at all downstream sites; however, baetids and chironomids were abundant and even the dominant taxa colonizing downstream. We believe that this dissimilarity between mesocosm-predicted and field-observed results is due to life history attributes and life stage sensitivity. For example, *Zapada* sp., are longer lived and univoltine (one generation per year), resulting in longer larval exposure in the field and lower propensity to colonize. In contrast, both baetids and chironomids can have multiple generations (i.e., multivoltine) in a year, and they have high drift and emergence propensity, resulting in decreased larval life stage exposure and greater potential for downstream colonization.

The stream community mesocosm exposure used in this study was relatively short duration, the mesocosms were closed systems (i.e., limiting outside colonization), and the exposure regime (i.e., stable vs fluctuating metal concentrations) differed from field conditions. These experimental artifacts may, in part, explain discrepancies between field and mesocosm results. Iwasaki et el. (2018) characterized aquatic insect taxa sensitivity between field and mesocosm approaches and determined that taxa were generally more sensitive in the field. The authors suggested that factors such as dietary exposure, exposure duration, recolonization in the field, and effects of metals on sensitive life stages influenced mesocosm and field results. Given the importance of developing predictive models for AMD remediation projects, integrated experimental and field approaches such as those applied in this research, have the potential to improve our ability to predict biological recovery in the future.

Primary production is a critical stream ecosystem process, providing the trophic base for many stream food webs; therefore, algal response to remediation is a key biomonitoring assessment metric (McCormick and Cairns 1994). Benthic algae responded immediately following treatment, with increases in biomass of diatoms and cyanobacteria at Impact sites. This result supports our mesocosm findings that metal-plated or “armored” substrate does not appear to inhibit algal colonization and that even accelerated growth relative to non-plated substrate may occur once acid mine drainage inputs are removed. Considering that algal biomass was undetected at many Impact sites prior to remediation, even a marginal increase in biomass was significant. Increased biomass following AMD treatment suggests that rates of primary production increased; however, some downstream sites shifted to algal communities dominated by cyanobacteria. Grazing aquatic insects generally prefer diatoms (Anderson and Cummins 1979), whereas cyanobacteria are generally considered non-palatable and even toxic to some aquatic organisms. More research is needed to determine the underlying mechanisms driving the dominance of cyanobacteria at AMD impaired sites and the relative influence of this material on macroinvertebrate grazers.

Associating relative changes in algal composition to AMD remediation is challenging because algae respond to numerous abiotic (water and habitat quality) and biotic factors (consumer-resource interactions), which can have both positive and negative effects (Hart 1985, Hill and Knight 1987, Medley and Clements 1998, Kotalik et al. 2019). For example, algal biomass at the Recovery station remained relatively low and did not significantly change after AMD treatment, whereas biomass at Impact sites increased and often exceeded Recovery site observations. We hypothesize that because higher densities of grazers recolonized downstream at the Gauge site, benthic algal standing stock was relatively more suppressed (i.e., top-down grazer control) compared to Impact sites that had much lower densities of recolonizing collector-gathers. In fact, we generally observed decreasing benthic macroinvertebrate abundance and increasing algal biomass in both summer 2017 and 2018, supporting this relationship. Lastly, we believe the Reference sites continued to have the greatest algal biomass given optimal water and habitat quality, and because these communities possessed far greater functional-feeding group diversity and were not dominated by algal consumers (i.e., less top-down control of algae).

Interannual and seasonal fluctuations in metal concentrations and stream community structure can strongly influence the responses of benthic communities to remediation (Clements 2000, Mebane et al. 2015, Herbst et al. 2018). Our ability to characterize biological recovery following NCCWTP operation was greatly improved by the Before-After Control-Impact design used for this study. The six years of NFCC benthic field data collected prior to AMD treatment characterized the extent of downstream impairment and reference site community composition. Multiple years of consistent and continuous sampling before AMD-treatment improved our ability to determine the extent of biological recovery by accounting for annual variation in the NFCC ecosystem. Additionally, because the primary sources of metal loading were contained by two point-source inputs that were diverted, treated, and discharged, we distinctly bracketed our Control and Impact site locations for our BACI study design.

Water quality in the NFCC significantly improved following NCCWTP operations, and given the current trajectory of recovery, we expect continued improvements in biological metrics among downstream sites in the future. However, habitat impairment and residual metals inputs in the watershed will continue to impair aquatic life, and prioritizing remediation of these stressors will be critical moving forward. Our BACI survey coupled with the novel experiments used in this applied research improved our ability to predict and characterize physical and chemical remediation effectiveness in the NFCC. Applying both experimental and field approaches to predict and evaluate AMD-remediation projects in the future will improve our ability to address the physical, chemical, and biological mechanisms influencing stream recovery.

Figures and Tables from Chapter 5

Contains:

Figures 5.1 to 5.9

Tables 5.1 and 5.5

See Appendix D for Figures S5.1 to S5.3 and Tables S5.1 to S5.8

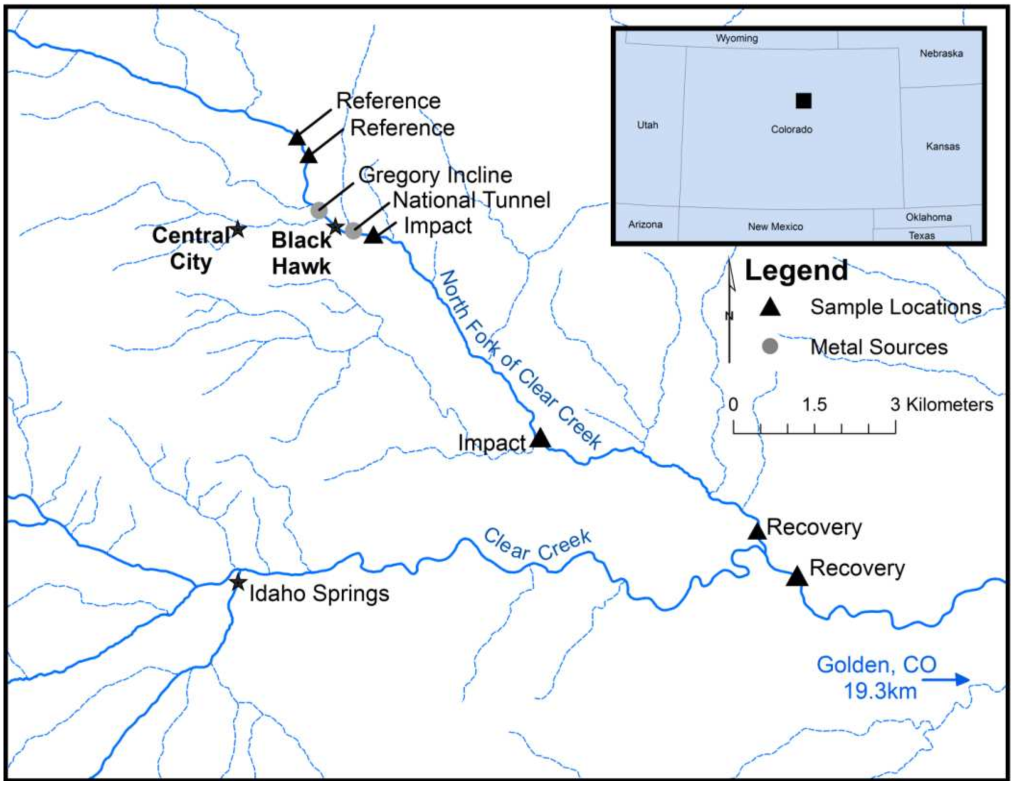

Figure 5.1 Map of the North Fork Clear Creek Watershed located approximately 20 km west of Golden, Colorado USA. Prior to the installation of the North Clear Creek Treatment Plant (NCCTP), point source acid mine drainages (AMD) discharged in the town of Black Hawk at Gregory Incline and National Tunnel. Two reference sites were located above the AMD discharge, impacted sites were located above/below the NCCTP and above Russel Gulch, and the recovery site was located immediately upstream of the Clear Creek Confluence.

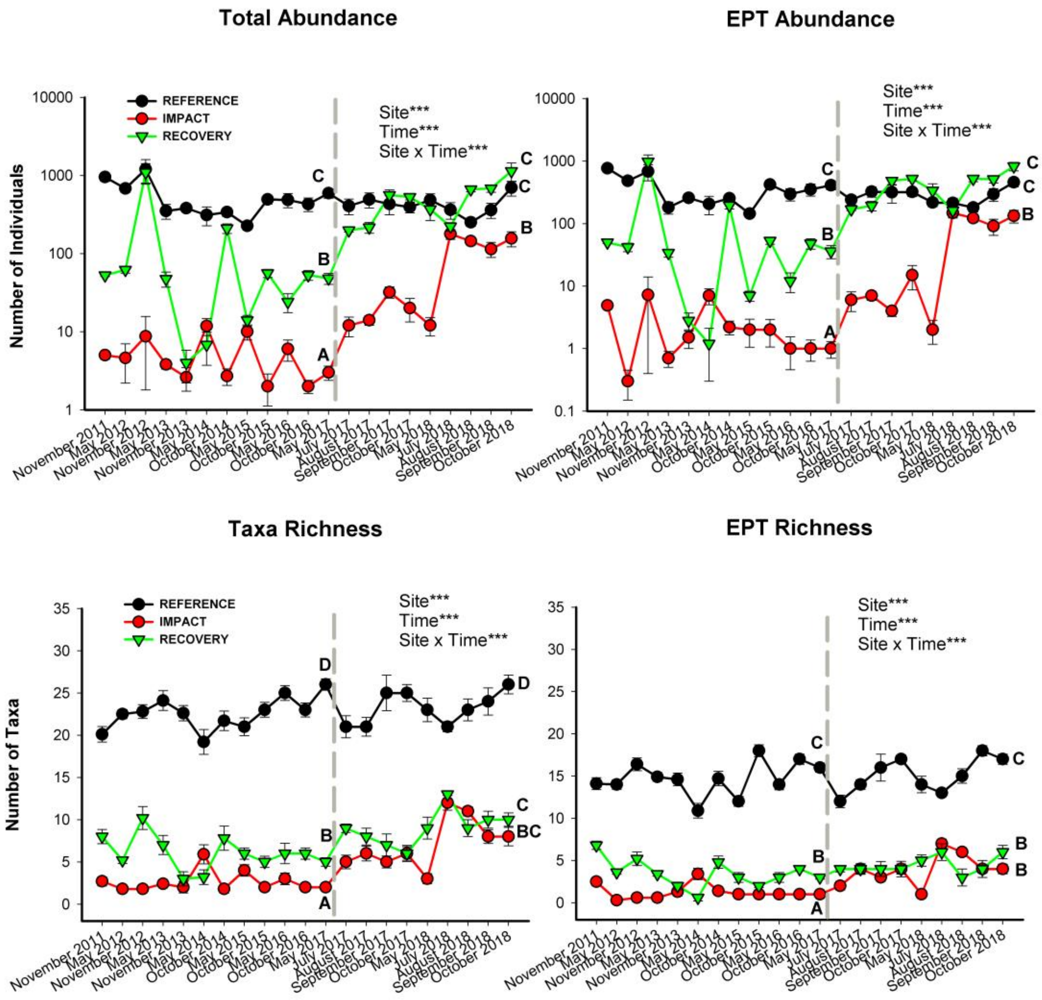

Figure 5.2 Results from benthic macroinvertebrate community surveys in the North Fork Clear Creek, Colorado, USA from 2011-2018. The North Clear Creek Water Treatment Plant became operational the summer of 2017 indicated by the grey dashed line). Results of the statistical analyses are shown for each of the four metrics analyzed, including the interaction term. Multiple comparisons among sites before and after treatment are expressed as letters (A-D), with different letters indicating significant differences.

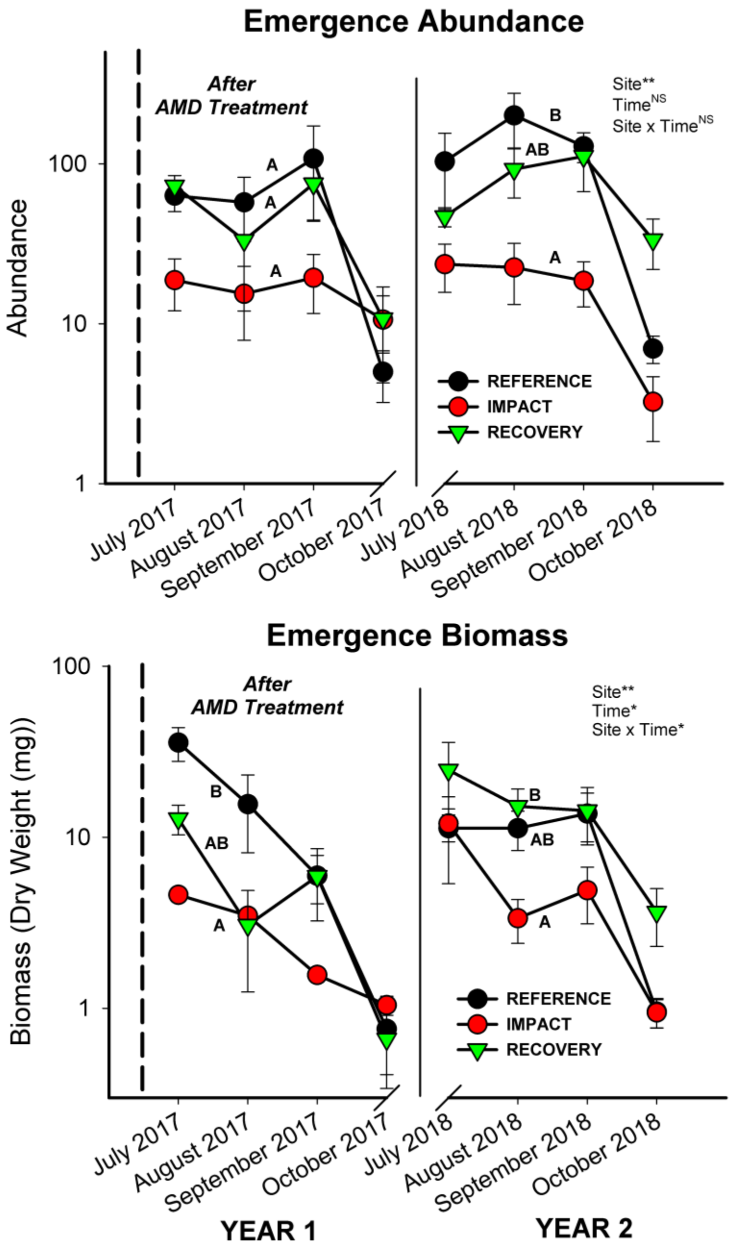

Figure 5.3 Results from aquatic insect emergence sampling in the summers of 2017 and 2018 following operation of the North Clear Creek Water Treatment Plant. Models terms tested the hypothesis that emergence abundance and biomass changed from Year 1 and Year 2, but multiple comparisons of average site metrics were performed for each year separately.

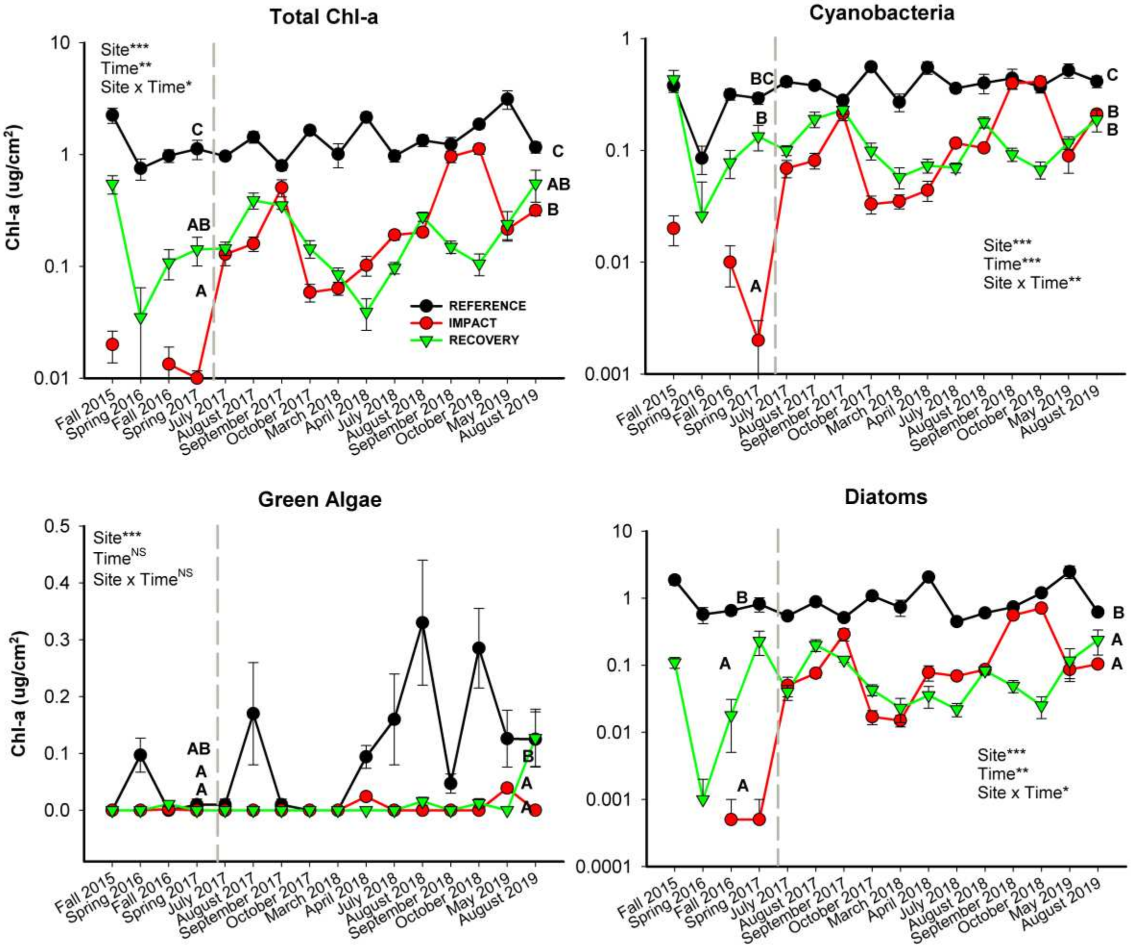

Figure 5.4 Results from benthic algal surveys in the North Fork Clear Creek, Colorado, USA from 2015-2019. The North Clear Creek Water Treatment Plant became operational the summer of 2017 (indicated by the grey dashed line). Results of the statistical analyses are shown for each of the four metrics analyzed, including the interaction term. Multiple comparisons among sites before and after treatment are expressed as letters (A-C), with different letters indicating significant differences.

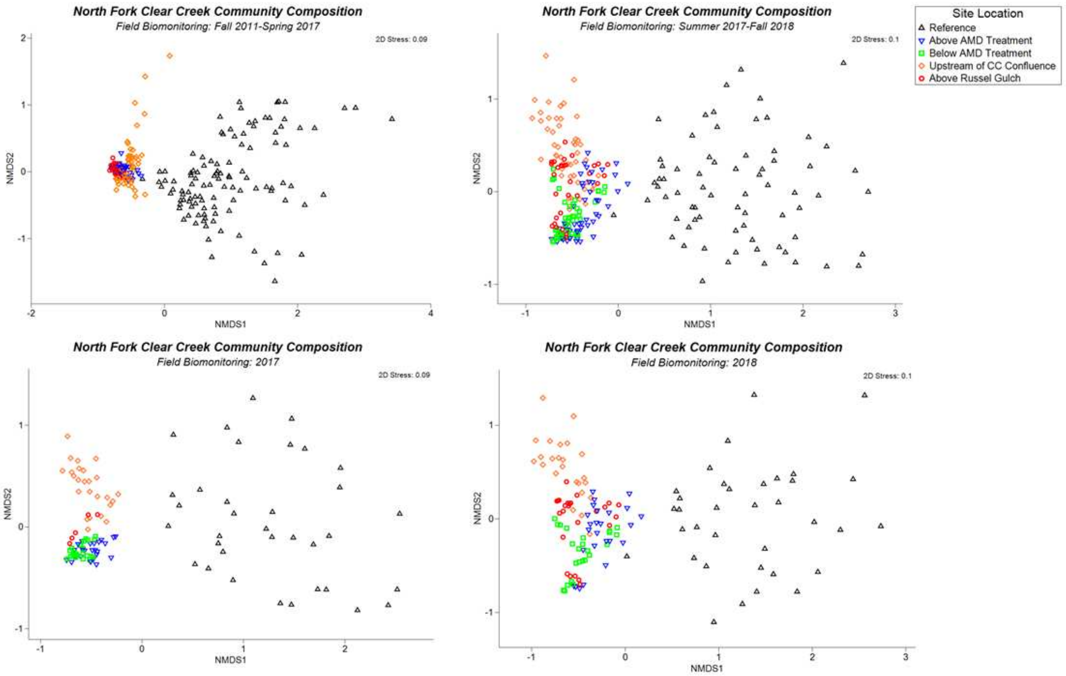

Figure 5.5 Nonmetric Multi-Dimensional Scaling (nMDS) analyses of benthic macroinvertebrate communities in North Fork Clear Creek from fall 2011 through spring 2017, summer 2017 through fall 2018, and annual plots for 2017 and 2018, respectively.

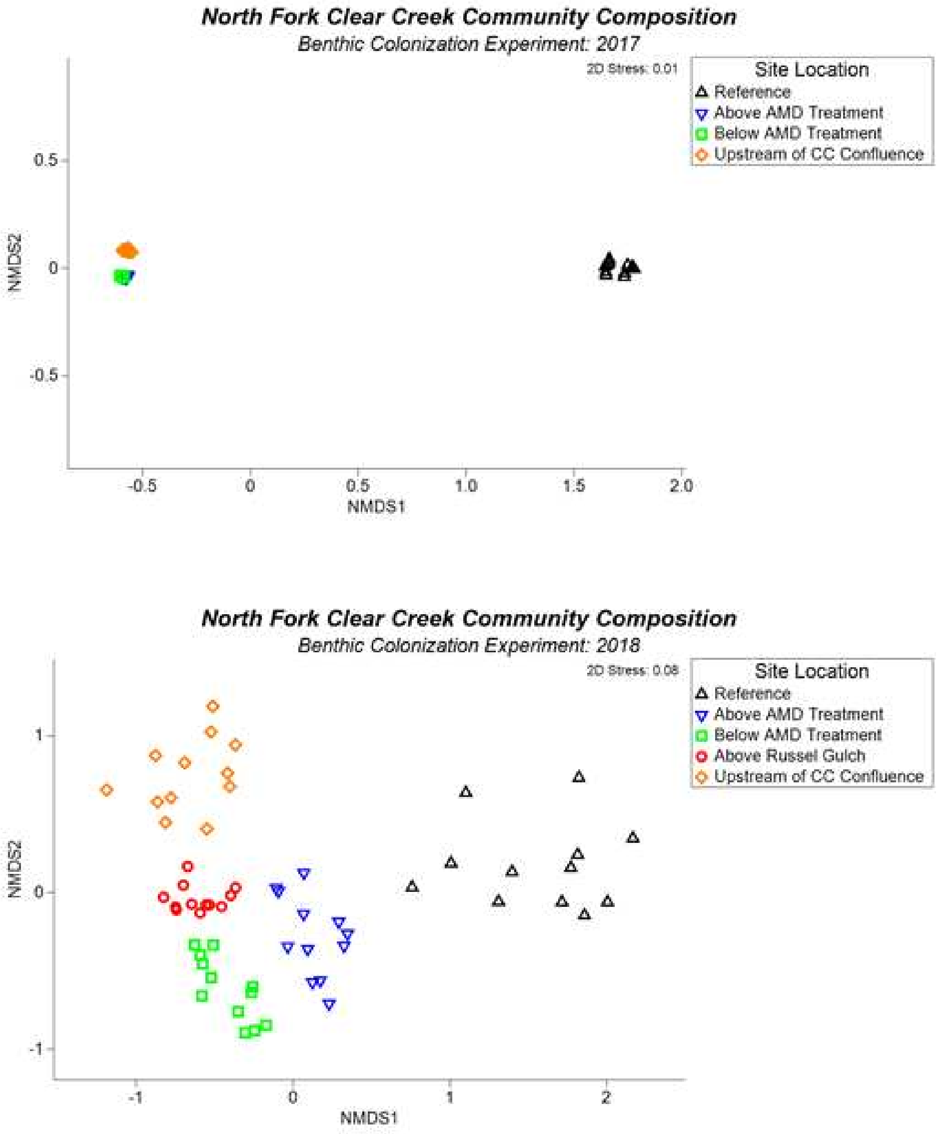

Figure 5.6 Nonmetric Multi-Dimensional Scaling (nMDS) analyses of benthic macroinvertebrate communities observed throughout the 60 d colonization experiments in the summers of 2017 and 2018 following acid mine drainage treatment.

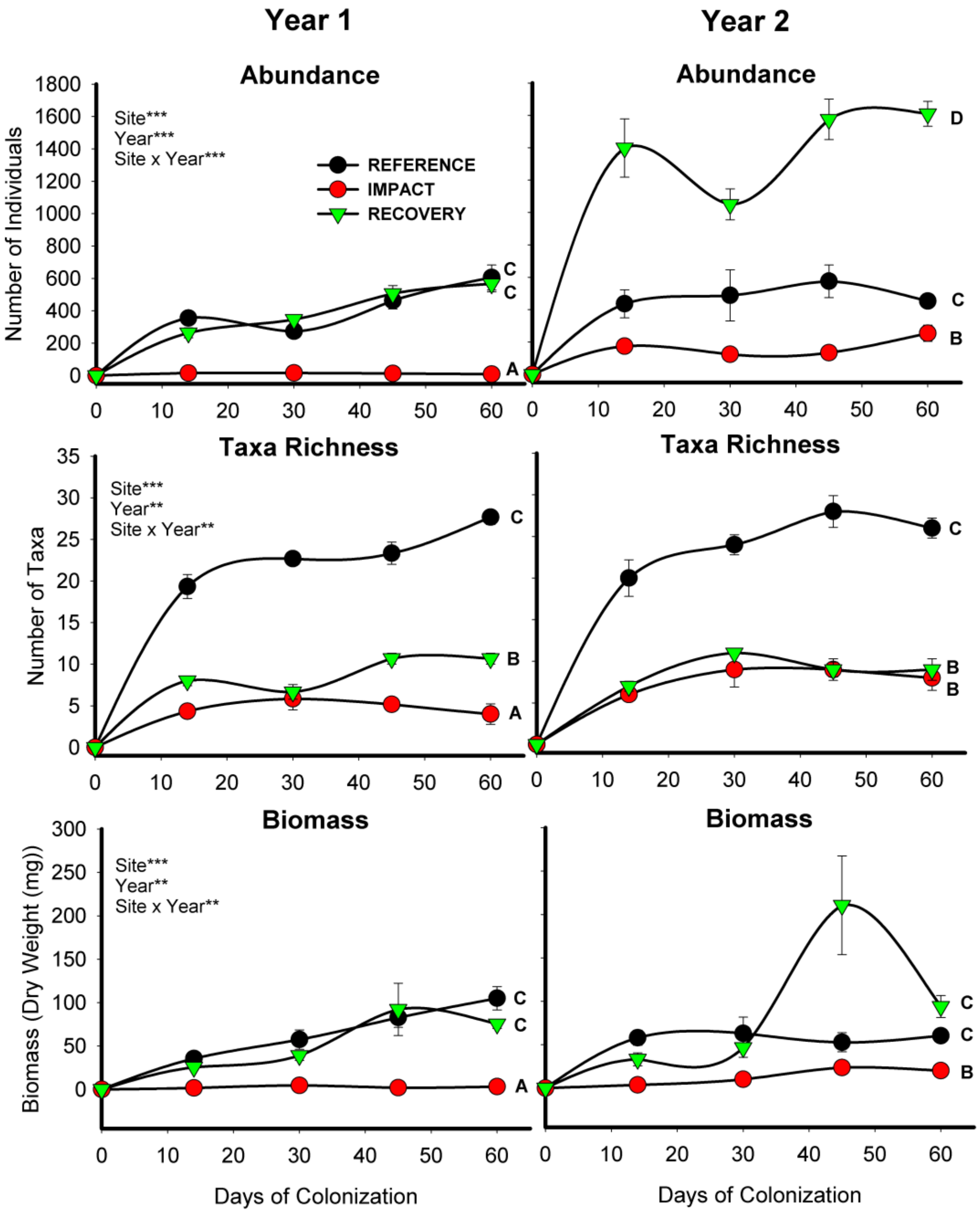

Figure 5.7 Benthic macroinvertebrate abundance, taxa richness, and biomass metrics evaluated in the field colonization experiments in 2017 and 2018 following NCCWTP operation. Substrate filled trays were deployed for 60 d and subsampled every 15 d (3 replicates of 2 trays per sampling event) to estimate rates of macroinvertebrate colonization.

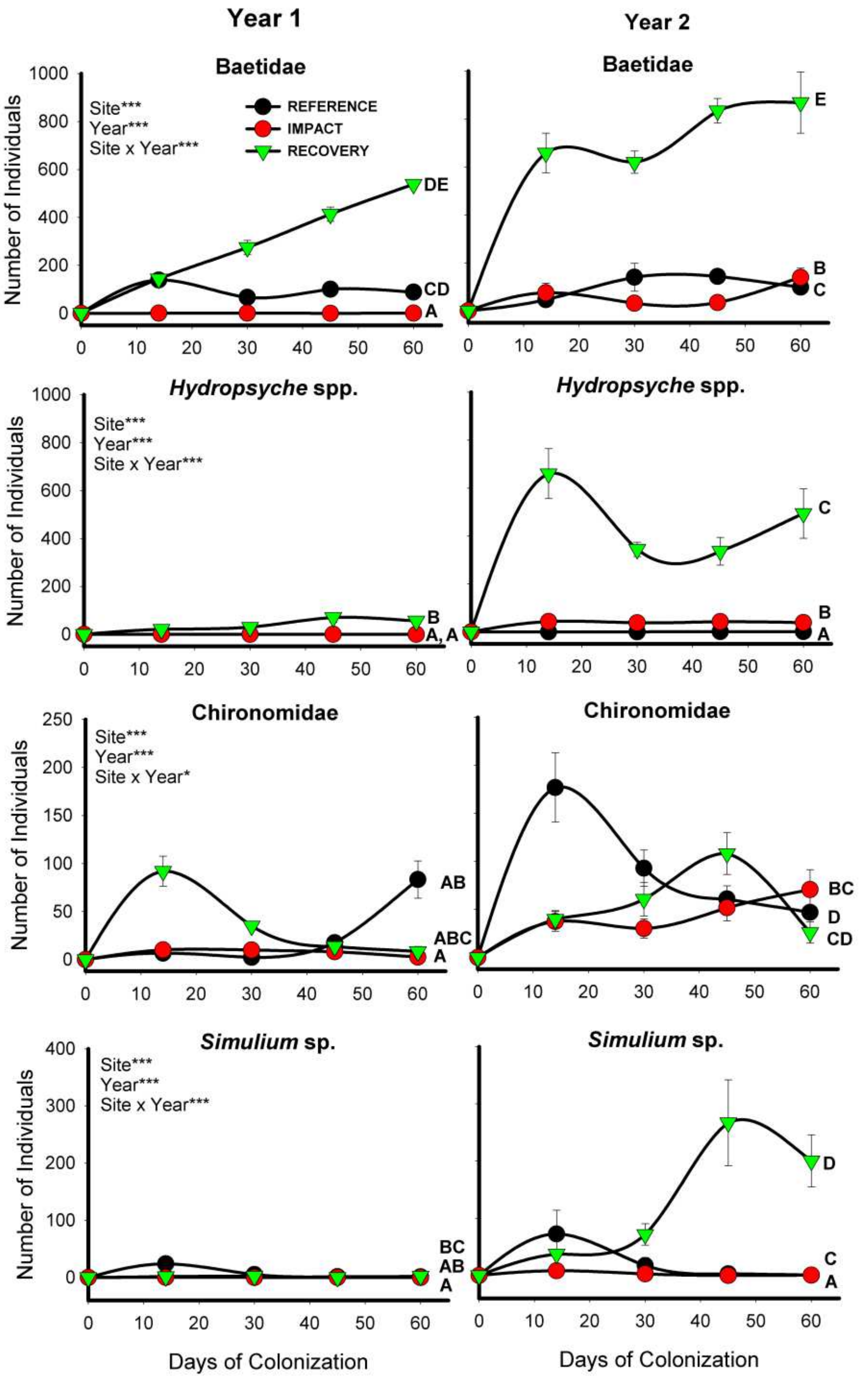

Figure 5.8 Abundance of dominant benthic macroinvertebrate taxa in the field colonization experiments in 2017 and 2018 following NCCWTP operation. Substrate filled trays were deployed for 60 d and subsampled every 15 d (3 replicates of 2 trays per sampling event) to estimate rates of macroinvertebrate colonization.

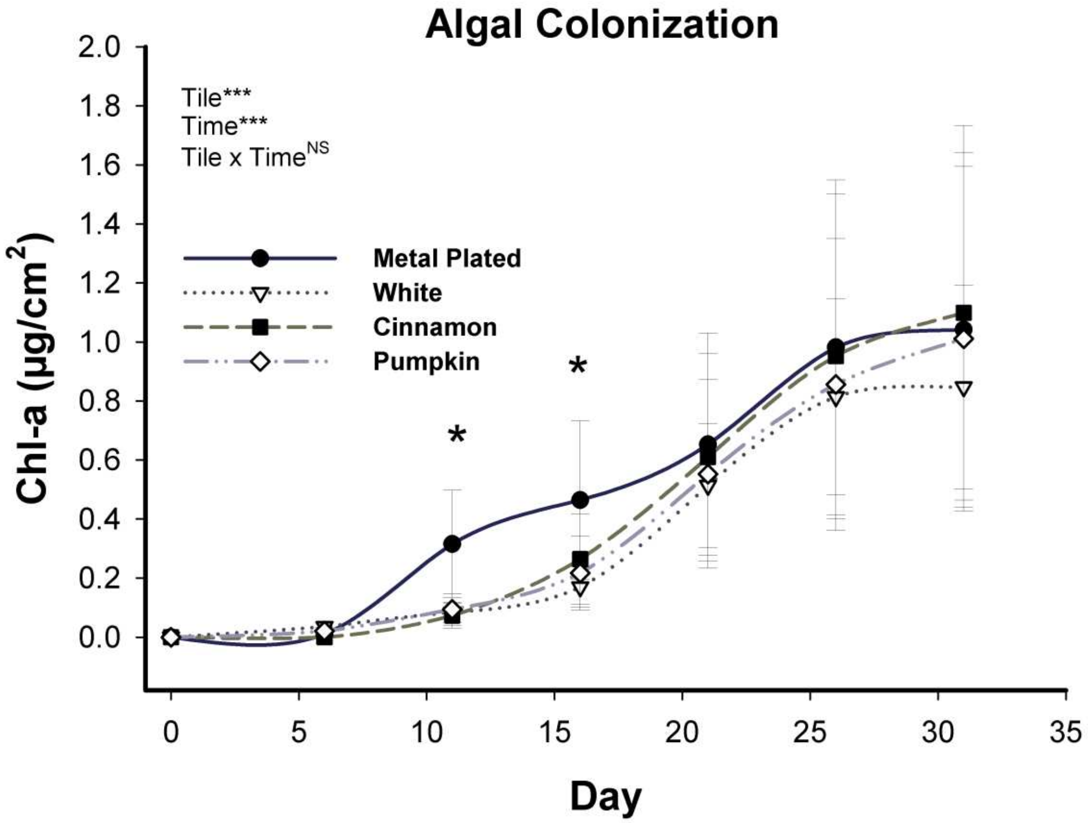

Figure 5.9 Algal colonization of ceramic tiles measured over a 31 d time period in stream mesocosms located at the Stream Research Laboratory, Colorado State University.

Table 5.1 Generalized water chemistry of North Fork Clear Creek after treatment of the two acid mine drainage sources (i.e., Gregory Incline and National Tunnel) at the North Clear Creek Water Treatment Plant. See Meyer et al. 2019 for higher resolution water chemistry.

Parameter	Reference	Above AMD Treatment	Below AMD Treatment	Above Main Stem Confluence
Cu (µg/L)	0 – 1	5 – 20	5 – 20	0 – 10
Zn (µg/L)	0 – 5	200 – 500	200 – 500	100 – 300
Fe (Total) (µg/L)	0 – 250	150 – 500	150 – 500	150 – 500
Alkalinity (mg/L $CaCO_3$)	< 15	< 15	< 15	< 25
Hardness (mg/L $CaCO_3$)	< 50	50 – 100	300 – 500	200 – 250
Conductivity (µS/cm)	70 – 120	150 – 250	500 – 800	500 – 800
pH	6.7 – 7.5	6.7 – 7.5	7.0 –7.7	7.0 – 7.7

Table 5.2 Site descriptions and locations of biological monitoring sites in the North Fork of Clear Creek, Colorado, USA.

Site Description	Label	Site Location	BACI
Upper Reference (0.20 miles upstream of Black Hawk Pump house)	Upper Reference	39°48'55.87"N 105°30'1.45"W	Reference
Pumphouse (at water pump station 0.3 miles upstream from fire station)	Pumphouse	39°48'45.49"N 105°29'53.22"W	Reference
Right Below National Tunnel (just downstream of National Tunnel)	Riviera	39°47'55.40"N 105°28'59.15"W	Impact
Below North Clear Creek Treatment Plant (NCCTP) (0.55 km downstream of the NCCTP)	Rail-less	39°47'32.24"N 105°28'20.66"W	Impact
Above Russel Gulch (Right upstream of the Russel Gulch confluence to NFCC)	Above Russel	39°45'52.58"N 105°26'48.06"W	Impact
Above Clear Creek Confluence (Just upstream of the old Hwy-6-119 bridge and tunnel)	USGS Gauge	39°44'57.51"N 105°23'58.67"W	Recovery
At the Clear Creek Canyon Park (0.5 miles downstream of Hwy-6 and Hwy-119 intersection)	Confluence	39°44'31.45"N 105°23'26.04"W	n/a
Main Stem Clear Creek, (right before Tunnel #6 heading on Hwy-6 to Idaho Springs)	Clear Creek Main	39°44'25.43"N 105°24'41.46"W	n/a

Table 5.3 Permutational multivariate ANOVA results for benthic macroinvertebrate biomonitoring data in North Fork Clear Creek before and after North Clear Creek Water Treatment Plant Operations.

Variable	df	SS	MS	Pseudo-F	P(perm)	Unique perms	P(MC)
Site	4	8145.2	2036.3	107.41	0.001	999	0.001
Time	1	596.13	596.13	31.446	0.001	999	0.001
Site x Time	4	628.81	157.2	8.2924	0.001	998	0.001
Res	515	9763.1	18.957				
Total	524	20876					

Table 5.4 Permutational multivariate ANOVA results for the benthic macroinvertebrate field colonization experiment in North Fork Clear Creek in the summers of 2017 and 2018. North Clear Creek Water Treatment Plant operations.

Variable	df	SS	MS	Pseudo-F	P(perm)	Unique perms	P(MC)
Site	4	273.87	273.87	35.639	0.001	999	0.001
Time	1	2475.7	618.92	80.541	0.001	996	0.001
Site x Time	3	240.43	80.145	10.429	0.001	997	0.001
Res	98	753.08	7.6845				
Total	106	3760					

Table 5.5 Estimates of EC50 and EC20 values (expressed as Cumulative Criterion Units based on summed hardness-adjusted criterion for Cu and Zn) obtained from a 14 d stream mesocosm exposure of field collected benthic communities obtained from the North Fork of Clear Creek Reference location. Dominant taxa as well as total larvae and emerged adults were selected based on our desired minimum number of individuals (n > 10) and statistically significant regression models.

Taxa Metrics	**EC50**	**EC50 CI**	**EC20**	**EC20 CI**
Baetis spp.	5.288	(2.49, 11.21)	2.93	(0.82, 10.41)
Ephemeroptera	6.37	(3.24, 12.52)	3.57	(1.22, 10.43)
Ephemerella spp.	6.388	(3.94, 10.35)	4.4	(2.24, 8.81)
Tanytarsini	7.37	(2.99, 18.14)	3.09	(0.65, 14.66)
Chironomidae	8.67	(3.12, 24.0)	2.97	(0.45, 19.44)
Epeorus spp.	8.813	(4.90, 15.83)	5.31	(2.38, 11.84)
Zapada sp.	9.83	(1.48, 64.93)	1.08	(0.005, 218.77)
Orthocladinae	11.88	(4.33, 32.60)	3.75	(0.6, 23.34)
Ephemeroptera Adult	13.25	(4.78, 36.69)	4.06	(0.65, 25.09)
Chloroperlidae	31.742	(0.59, 1688)	0.53	(0, 90345)
Community Larvae	8.714	(2.56, 29.65)	2.32	(0.19, 28.45)
Community Adult	72.1335	(5.53, 940.6)	14.89	(2.87, 77.15)

References

Anderson, M.; Gorley, R. N.; Clarke, R. K., Permanova+ for primer: Guide to software and statistical methods. Primer-E Limited: 2008.

Anderson, N.; Cummins, K. W., Influences of diet on the life histories of aquatic insects. Journal of the Fisheries Board of Canada 1979, 36, (3), 335-342.

Bates, D.; Mächler, M.; Bolker, B.; Walker, S., Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823 2014.

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

Beutler, M.; Wiltshire, K. H.; Meyer, B.; Moldaenke, C.; Lüring, C.; Meyerhöfer, M.; Hansen, U.-P.; Dau, H., A fluorometric method for the differentiation of algal populations in vivo and in situ. Photosynthesis research 2002, 72, (1), 39-53.

Bray, J. P.; Broady, P. A.; Niyogi, D. K.; Harding, J. S., Periphyton communities in New Zealand streams impacted by acid mine drainage. Marine and Freshwater Research 2009, 59, (12), 1084-1091.

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. Environmental Toxicology and Chemistry 2017, 36, (2), 285-291.

Cadmus, P.; Clements, W. H.; Williamson, J. L.; Ranville, J. F.; Meyer, J. S.; Gutiérrez Ginés, M. a. J. s., 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.

Cadmus, P.; Pomeranz, J. P.; Kraus, J. M., Low-cost floating emergence net and bottle trap: comparison of two designs. Journal of Freshwater Ecology 2016, 31, (4), 653-658.

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

Carpentier, C.; Dahlhaus, A.; van de Giesen, N.; Maršálek, B., The influence of hard substratum reflection and calibration profiles on in situ fluorescence measurements of benthic microalgal biomass. Environmental Science: Processes & Impacts 2013, 15, (4), 783-793.

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.; Cadmus, P.; Brinkman, S. F., 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.; Carlisle, D. M.; Lazorchak, J. M.; Johnson, P. C., Heavy metals structure benthic communities in Colorado mountain streams. Ecological Applications 2000, 10, (2), 626-638.

Clements, W. H.; Kotalik, C., 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.; Van Hassel, J. H.; Cherry, D. S.; Cairns, J., Colonization, variability, and the use of substratum-filled trays for biomonitoring benthic communities. Hydrobiologia 1989, 173, (1), 45-53.

Clements, W. H.; Vieira, N. K.; Church, S. E., 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.; Suter, G. W.; Zheng, L., Derivation of a benchmark for freshwater ionic strength. Environmental Toxicology and Chemistry 2013, 32, (2), 263-271.

Dabney, B. L.; Clements, W. H.; Williamson, J. L.; Ranville, J. F., Influence of Metal Contamination and Sediment Deposition on Benthic Invertebrate Colonization at the North Fork Clear Creek Superfund Site, Colorado, USA. Environmental Science & Technology 2018, 52, (12), 7072-7080.

DeNicola, D. M.; Stapleton, M. G., 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.

Fox, J.; Weisberg, S.; Adler, D.; Bates, D.; Baud-Bovy, G.; Ellison, S.; Firth, D.; Friendly, M.; Gorjanc, G.; Graves, S., Package ‘car’. Vienna: R Foundation for Statistical Computing 2012.

Guasch, H.; Acosta, X. G.; Urrea, G.; Bañeras, L., Changes in the microbial communities along the environmental gradient created by a small Fe spring. Freshwater Science 2012, 31, (2), 599-609.

Hare, L., Aquatic insects and trace metals: bioavailability, bioaccumulation, and toxicity. Critical reviews in toxicology 1992, 22, (5-6), 327-369.

Hart, D. D., Grazing insects mediate algal interactions in a stream benthic community. Oikos 1985, 40-46.

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.

Hill, W. R.; Knight, A. W., Experimental analysis of the grazing interaction between a mayfly and stream algae. Ecology 1987, 68, (6), 1955-1965.

Hurlbert, S. H., Pseudoreplication and the design of ecological field experiments. Ecological monographs 1984, 54, (2), 187-211.

Kim, K.; Funk, D.; Buchwalter, D., Dietary (periphyton) and aqueous Zn bioaccumulation dynamics in the mayfly Centroptilum triangulifer. Ecotoxicology 2012, 21, (8), 2288-2296.

Kotalik, C. J.; Cadmus, P.; Clements, W. H., Indirect effects of iron oxide on stream benthic communities: capturing ecological complexity with controlled mesocosm experiments. Environmental Science & Technology 2019.

Kotalik, C. J.; Clements, W. H., Stream mesocosm experiments show significant differences in sensitivity of larval and emerging adults to metals. Environmental Science & Technology 2019, 53, (14), 8362-8370.

Lenat, D. R.; Barbour, M. T., Using benthic macroinvertebrate community structure for rapid, cost-effective, water quality monitoring: rapid bioassessment. Biological monitoring of aquatic systems. Lewis Publishers, Boca Raton, Florida 1994, 187-215.

Lenth, R., Emmeans: Estimated marginal means, aka least-squares means. R Package Version 2018, 1, (2).

McCormick, P. V.; Cairns, J., Algae as indicators of environmental change. Journal of Applied Phycology 1994, 6, (5-6), 509-526.

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

Mebane, C. A.; Eakins, R. J.; Fraser, B. G.; Adams, W. J., Recovery of a mining-damaged stream ecosystem. Elem Sci Anth 2015, 3.

Medley, C. N.; Clements, W. H., Responses of diatom communities to heavy metals in streams: the influence of longitudinal variation. Ecological Applications 1998, 8, (3), 631-644.

Merritt, R. W.; Cummins, K. W.; Berg, M., An introduction to the aquatic insects of North America. Kendall. Hunt. 2004.

Meyers, J. S., Lloyd, E. H., Ranville, J. F., Water chemistry changes in North Fork Clear Creek, Gilpin County, Colorado, USA, following remediation of acid mine drainage inputs; and associated effects of bioavailability of copper and zinc. Prepared for International Copper Association, Ltd. 2018.

Moore, J. N.; Luoma, S. N.; Peters, D., Downstream effects of mine effluent on an intermontane riparian system. Canadian Journal of Fisheries and Aquatic Sciences 1991, 48, (2), 222-232.

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

Oliver, D., Life history of the Chironomidae. Annual review of entomology 1971, 16, (1), 211-230.

Pomeranz, J. P.; Warburton, H. J.; Harding, J. S., Anthropogenic mining alters macroinvertebrate size spectra in streams. Freshwater biology.

Pond, G. J.; Passmore, M. E.; Borsuk, F. A.; Reynolds, L.; Rose, C. J., Downstream effects of mountaintop coal mining: comparing biological conditions using family-and genus-level macroinvertebrate bioassessment tools. *Journal of the North American Benthological Society* 2008, *27*, (3), 717-737.

Rainbow, P. S., Trace metal concentrations in aquatic invertebrates: why and so what? Environmental Pollution 2002, 120, (3), 497-507.

Ritz, C.; Baty, F.; Streibig, J. C.; Gerhard, D., Dose-response analysis using R. PloS one 2015, 10, (12), e0146021.

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.

Stephan, C. E.; Mount, D. I.; Hansen, D. J.; Gentile, J.; Chapman, G. A.; Brungs, W. A., Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. US Environmental Protection Agency Washington, DC: 1985.

Stewart-Oaten, A.; Murdoch, W. W.; Parker, K. R., Environmental impact assessment:" Pseudoreplication" in time? Ecology 1986, 67, (4), 929-940.

Team, R. C., R: A language and environment for statistical computing. 2013.

Underwood, A., Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world. Journal of experimental marine biology and ecology 1992, 161, (2), 145-178.

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, 1985.

U.S. EPA. 1999. Update of aquatic life ambient water quality criteria for ammonia. EPA 822-R-99-014. Washington, DC.

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

Vuori, K.-M. In Direct and indirect effects of iron on river ecosystems, Annales Zoologici Fennici, 1995; JSTOR: 1995; pp 317-329.

Wellnitz, T.; Sheldon, S., The effects of iron and manganese on diatom colonization in a Vermont stream. Freshwater biology 1995, 34, (3), 465-470.

West, B. T.; Welch, K. B.; Galecki, A. T., Linear mixed models: a practical guide using statistical software. Chapman and Hall/CRC: 2014.

APPENDIX A

SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 2

Contains:

Figures S2.1 to S2.4

Tables S2.1 to S2.4

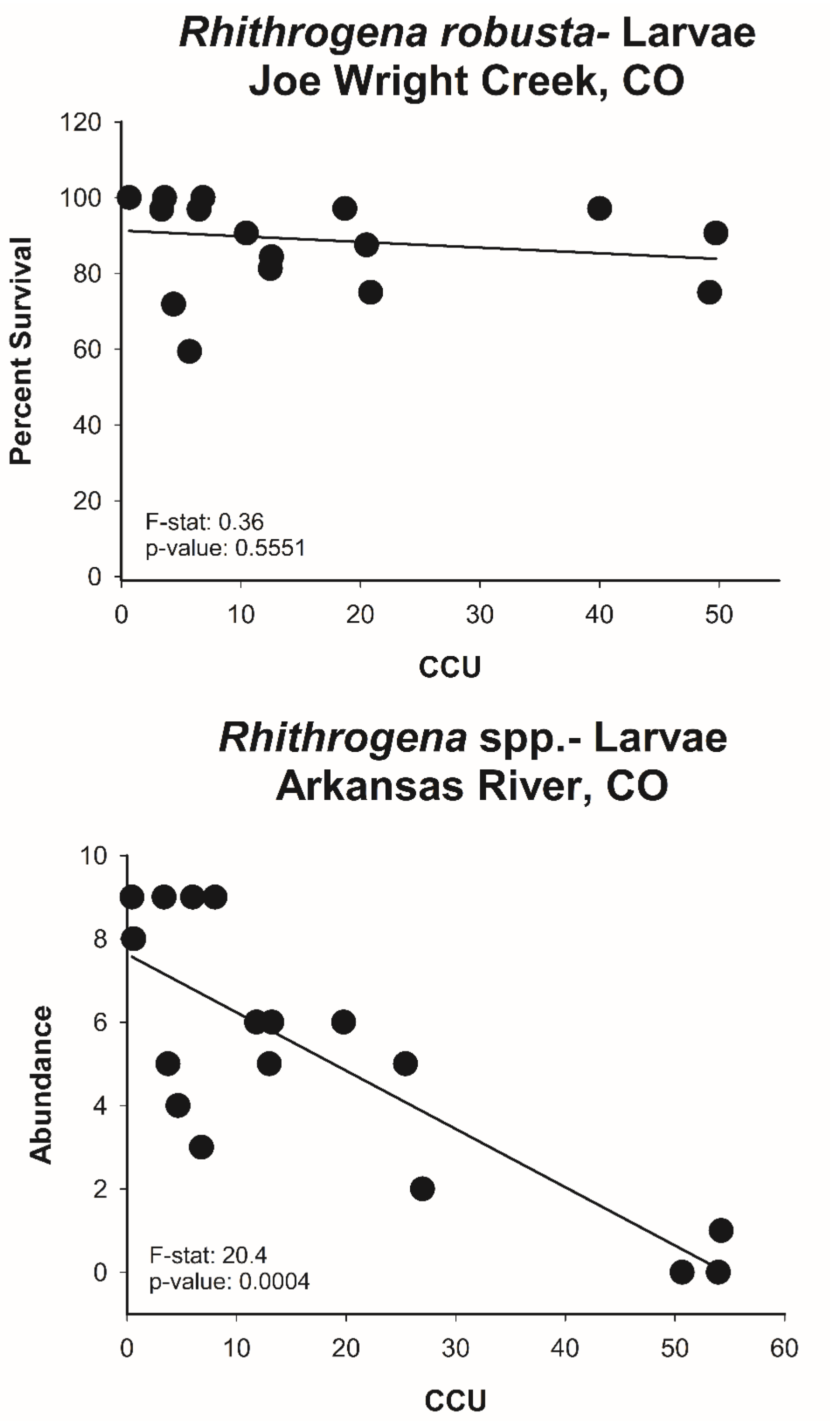

Figure S2.1 Comparison of metal responses of larval *Rhithrogena robusta* from the single-species experiment and *Rhithrogena* spp. from the community mesocosm experiment.

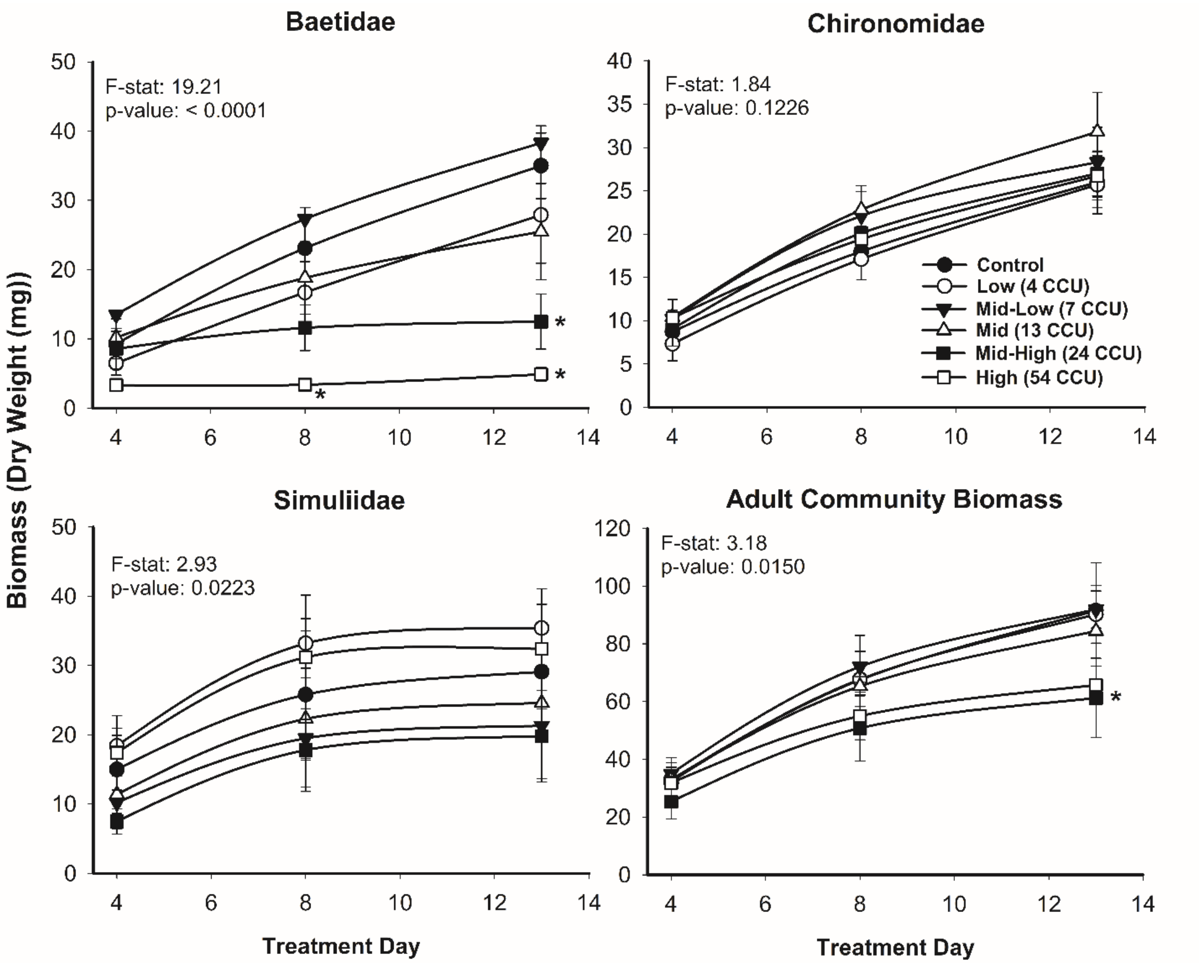

Figure S2.2 Effects of Cu and Zn on cumulative average (± s.e) adult insect emergence biomass of Baetidae (Ephemeroptera), Chironomidae (Diptera), Simuliidae (Diptera), and adult community biomass in the 14-d community exposure. Biomass estimates were conducted for days 1-4, 5-8, and 9-13. Single-factor repeated measures ANOVA was used to test for significant differences in mean cumulative emergence biomass among treatments over the exposure duration. *Post-hoc* pairwise comparisons with Dunnett's test were used to test for significant relationships between treatment and control responses. Asterisks (*) indicate significant differences from controls.

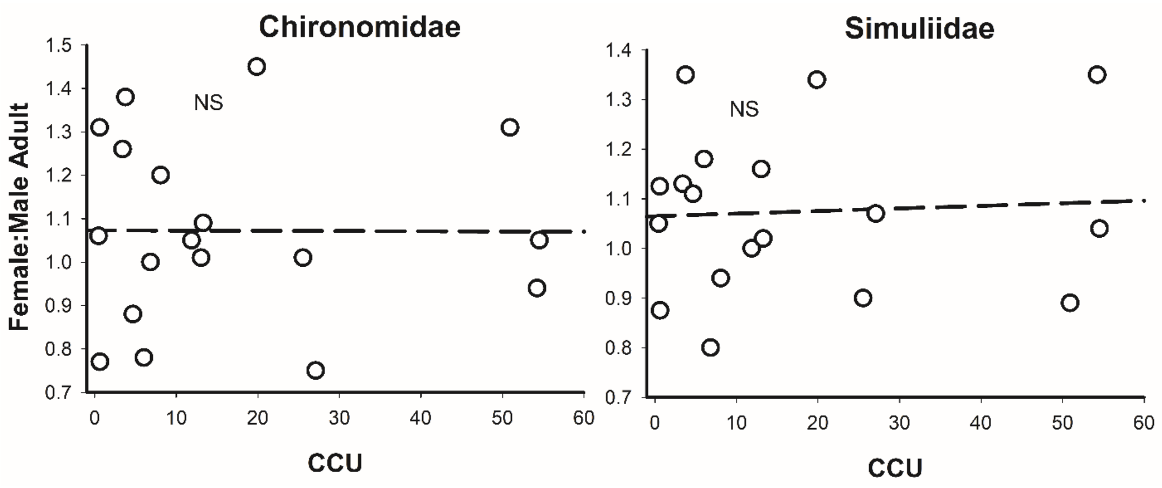

Figure S2.3 Effects of Cu and Zn on emerging adult sex ratios (female:male) of Chironomidae (Diptera) and Simuliidae (Diptera) collected from the stream benthic community experiment. Linear regression was used to develop concentration-response relationships for each group.

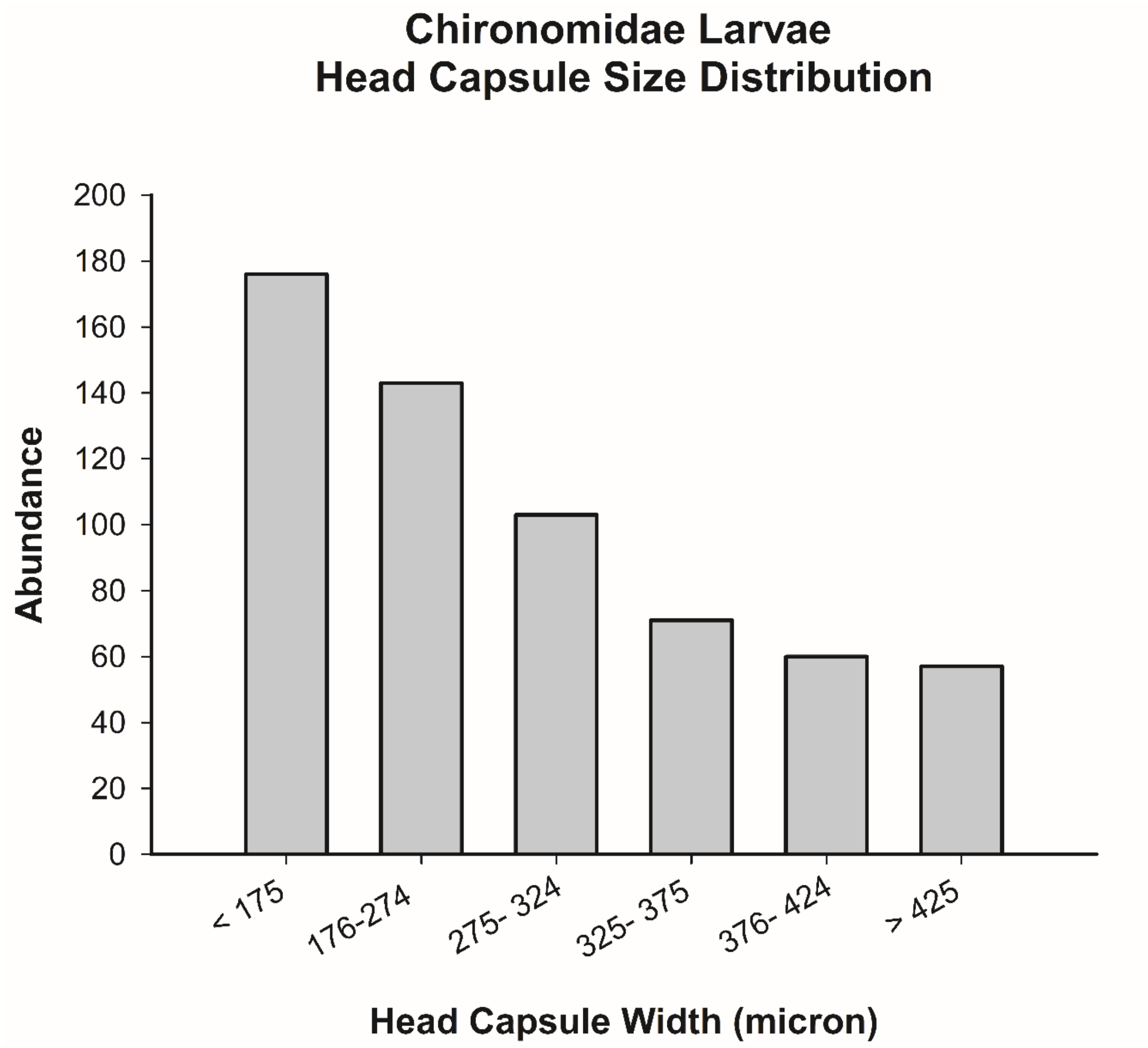

Figure S2.4 Head capsule size distribution of Chironomidae in control treatments from Arkansas River community mesocosm exposure. Approximately 52% of Chironomidae abundance in the control treatments were early to mid-instar sizes, defined as 274 microns and smaller.

Table S2.1 Water quality characteristics in stream mesocosms from the *R. robusta* single-species and Arkansas River whole-community exposures. Values are reported as means (+/- s.d.).

Experiment	Treatment	% of Target Cu	% of Target Zn	Temperature (°C)	pH	Conductivity (µS/cm)
R. robusta	Control	n/a	n/a	12.64 (0.52)	7.31 (0.08)	77.84 (0.54)
	Low	195	79	12.69 (0.41)	7.26 (0.07)	78.24 (0.34)
	Mid-Low	149	82	12.77 (0.98)	7.22 (0.06)	78.81 (0.55)
	Mid	144	71	12.67 (0.41)	7.20 (0.07)	79.37 (0.75)
	Mid-High	106	76	12.42 (0.48)	7.16 (0.10)	79.9 (0.83)
	High	109	100	12.60 (0.47)	7.09 (0.06)	82.62 (0.69)
AR Community	Control	n/a	n/a	12.46 (0.62)	7.35 (0.05)	79.29 (0.34)
	Low	116	102	12.55 (0.58)	7.29 (0.03)	79.86 (0.51)
	Mid-Low	104	89	12.38 (0.77)	7.31 (0.08)	79.87 (0.38)
	Mid	95	81	12.44 (0.65)	7.26 (0.05)	80.5 (0.21)
	Mid-High	75	87	12.32 (0.53)	7.21 (0.05)	81.63 (0.37)
	High	78	98	12.42 (0.57)	7.13 (0.04)	83.95 (1.42)

Table S2.2 Summary of statistical results of linear models analyses used to test for significance from single species and whole community experiments. The predictor variable for all response variables was Cumulative Criterion Unit (CCU). Degrees of freedom were the same for all linear models (d.f. = 1, 16). Standardized effect sizes for each model were estimated using partial-eta squared (η^2).

Experiment	**Response Variable**	**R^2**	**Slope**	**Intercept**	**F-stat**	**p-value**	**η^2**
R. robusta	Percent Emergence	0.06	-0.0004	4.427	0.01	0.9378	0.001
	Percent Larval Mortality	0.02	0.117	9.722	0.36	0.5551	0.022
AR Community	Baetidae Larval Abundance	0.89	-0.098	5.137	150.4	< 0.0001	0.904
	Baetidae Adult Abundance	0.84	-0.0457	4.144	93.33	< 0.0001	0.853
	Chironomidae Larval Abundance	0.81	-0.0297	5.499	75.75	< 0.0001	0.826
	Chironomidae Adults Abundance	0.09	-0.0029	5.180	2.86	0.1101	0.152
	Simuliidae Larval Abundance	0.06	-0.0006	2.424	0.01	0.9998	0.001
	Simuliidae Adult Abundance	0.01	-0.0054	4.470	0.86	0.3670	0.051
	Rhithrogena spp. Larval Abundance	0.54	-0.512	2.857	20.4	0.0004	0.560
	Ephemeroptera Larval Biomass	0.64	-0.0007	0.032	32.04	< 0.0001	0.667
	Ephemeroptera Adult Biomass	0.61	-0.0006	0.034	27.81	< 0.0001	0.635
	Chironomidae Larval Biomass	0.60	-0.0002	0.015	26.61	< 0.0001	0.625
	Chironomidae Adult Biomass	0.05	-0.000006	0.027	0.01	0.9220	0.001
	Simuliidae Larval Biomass	0.11	-0.0006	0.004	3.04	0.1012	0.160
	Simuliidae Adult Biomass	0.05	-0.001	0.029	0.18	0.6747	0.011
	Community Larval Biomass	0.43	-0.0014	0.217	13.95	0.0018	0.466
	Community Adult Biomass	0.27	-0.0006	0.091	7.51	0.0145	0.319
	Baetidae Adult Sex Ratio	0.75	0.056	0.645	51.97	< 0.0001	0.764
	Chironomidae Adult Sex Ratio	0.04	-0.00005	1.071	0.0003	0.9854	0.000
	Simuliidae Adult Sex Ratio	0.05	0.0005	1.065	0.05	0.8174	0.003

Table S2.3 Summary of statistical results of single-factor repeated measures ANOVA used to test for overall differences in the timing of emerging adult abundance and biomass. The table also shows pairwise contrasts among CCU treatments and control observations over the entire exposure duration.

Experiment	Response Variable	Contrast to Controls	t-ratio	p-value
R. robusta	Emergence Abundance	Low	-2.331	0.0860
	(d.f. = 5, 165)	Mid-Low	0.942	0.7689
	(F-stat: 3.33; p-value: 0.0067)	Mid	-2.602	0.0437
		Mid-High	-0.986	0.7439
		High	-0.178	0.9969
AR Community	Baetidae Emergence Abundance	Low	-3.130	0.0092
	(d.f. = 5, 216)	Mid-Low	2.526	0.0523
	(F-stat: 81.78; p-value: < 0.0001)	Mid	-1.088	0.6824
		Mid-High	-4.838	< 0.0001
		High	-15.682	< 0.0001
	Chironomidae Emergence Abundance	Low	3.180	0.0078
	(d.f. = 5, 216)	Mid-Low	2.307	0.0902
	(F-stat: 5.36; p-value: 0.0001)	Mid	4.611	< 0.0001
		Mid-High	1.384	0.4932
		High	1.116	0.6647
	Simuliidae Emergence Abundance	Low	-0.574	0.9314
	(d.f. = 5, 216)	Mid-Low	-2.342	0.0834
	(F-stat: 21.61; p-value: < 0.0001)	Mid	-4.894	< 0.0001
		Mid-High	-8.208	< 0.0001
		High	-2.062	0.1560
	Baetidae Emergence Biomass	Low	-1.587	0.3802
	(d.f. = 5, 46)	Mid-Low	1.266	0.5744
	(F-stat: 19.21; p-value: < 0.0001)	Mid	-0.971	0.7543
		Mid-High	-3.172	0.0121
		High	-7.546	< 0.0001
	Chironomidae Emergence Biomass	Low	-0.464	0.9604
	(d.f. = 5, 46)	Mid-Low	1.423	0.4759
	(F-stat: 1.84; p-value: 0.1226)	Mid	2.201	0.1275
		Mid-High	0.626	0.9144
		High	0.677	0.8957
	Simuliidae Emergence Biomass	Low	1.461	0.4531
	(d.f. = 5, 46)	Mid-Low	-1.361	0.5143
	(F-stat: 2.93; p-value: 0.0223)	Mid	-0.404	0.9724
		Mid-High	-1.617	0.3634
		High	0.862	0.8131
	Community Emergence Biomass	Low	0.230	0.9938
	(d.f. = 5, 46)	Mid-Low	0.764	0.8597
	(F-stat: 3.18; p-value: 0.0150)	Mid	0.131	0.9987
		Mid-High	-2.620	0.0499
		High	-1.329	0.5347

Table S2.4 Summary of statistical results of ANCOVA models used to test for interaction effects between metals (as CCU) and life stage (larvae vs adult) for response variables in the AR community exposure. Degrees of freedom (d.f. = 1, 32) for all ANCOVA models were the same. Standardized effect sizes for each model were estimated using partial-eta squared (η^2).

Response Variable	Predictor Variable	F-stat	p-value	η^2
Baetidae Abundance	CCU	48.43	< 0.0001	0.602
	Life stage	18.74	0.0001	0.369
	CCU x Life stage	31.72	< 0.0001	0.498
Chironomidae Abundance	CCU	1.21	0.2798	0.036
	Life stage	11.23	0.0020	0.260
	CCU x Life stage	48.36	< 0.0001	0.602
Simuliidae Abundance	CCU	0.71	0.4052	0.022
	Life stage	81.19	< 0.0001	0.717
	CCU x Life stage	0.28	0.5979	0.009
Ephemeroptera Biomass	CCU	29.17	< 0.0001	0.477
	Life stage	1.40	0.2446	0.042
	CCU x Life stage	3.87	0.0577	0.110
Chironomidae Biomass	CCU	0.20	0.6549	0.006
	Life stage	3.72	0.0625	0.104
	CCU x Life stage	11.74	0.0016	0.268
Simuliidae Biomass	CCU	0.35	0.5601	0.011
	Life stage	13.76	0.0008	0.300
	CCU x Life stage	0.02	0.8740	0.001
Community Biomass	CCU	4.45	0.0427	0.122
	Life stage	84.75	< 0.0001	0.726
	CCU x Life stage	7.20	0.0114	0.184

APPENDIX B

SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 3Contains:

Figures S3.1 to S3.4

Tables S3.1 to S3.6

Figure S3.1. Photo of stream mesocosms covered with insect screening material to capture emerging adult aquatic insects. Location: Stream Research Laboratory, Colorado State University, Foothill Campus, Fort Collins, CO USA.

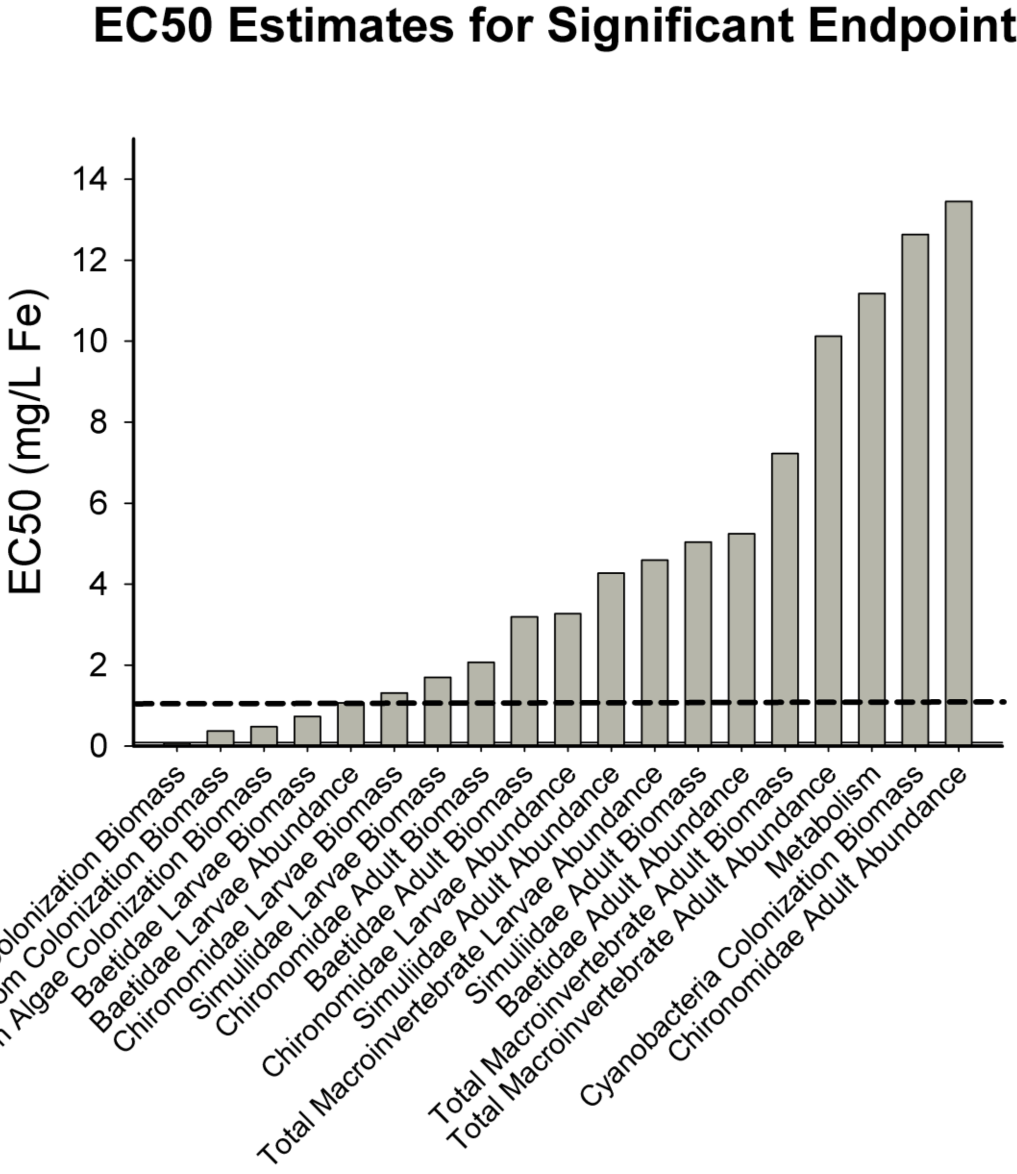

Figure S3.2 EC50 rankings for all significant structural and functional endpoints measured in the study. We estimated EC50 values (effect concentration that reduced endpoint by 50%, exclusive of control data) with two-parameter log-logistic dose response curves, expect for cyanobacteria colonization biomass that required a two-parameter log-normal dose-response model because biomass increased with Fe concentration.

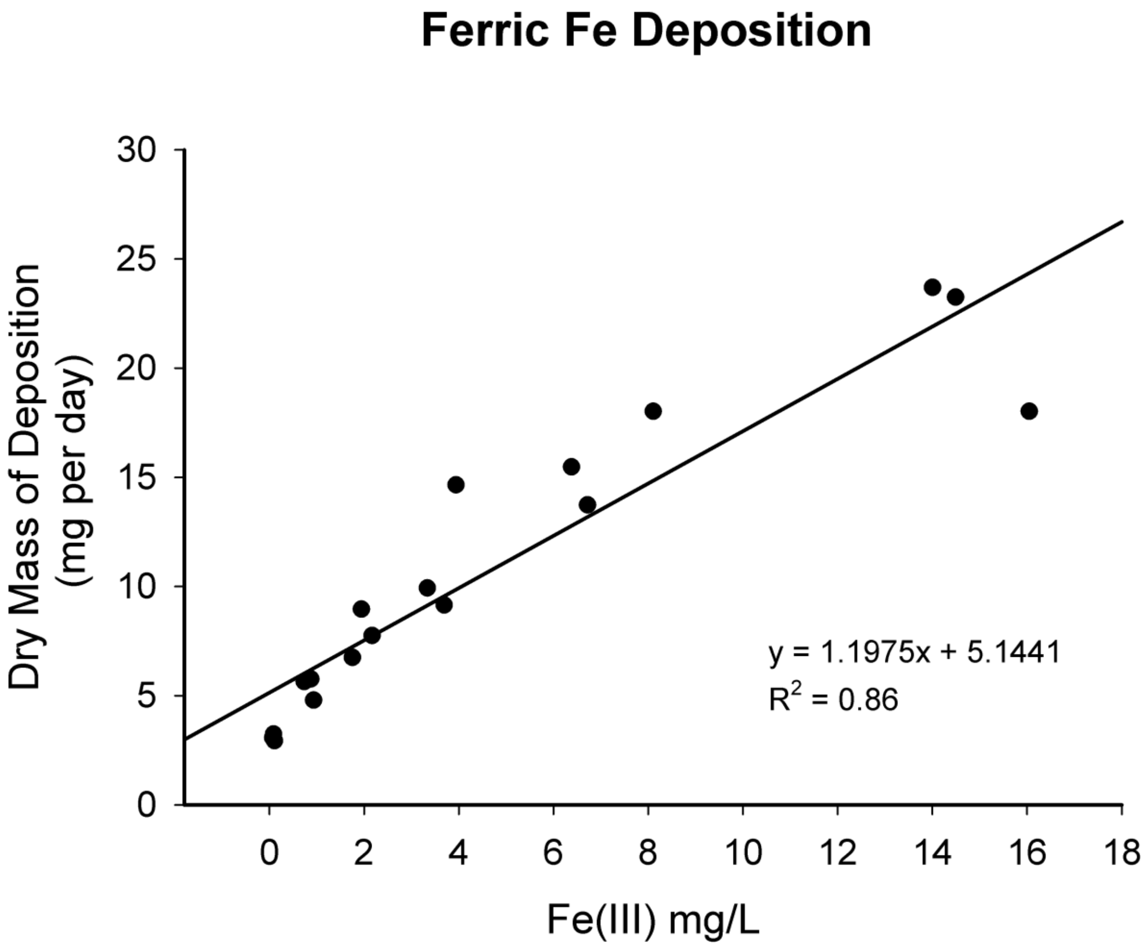

Figure S3.3 Deposition of Fe precipitates in bead (6 mm diameter) samplers. Dry mass of Fe precipitates (mg/day) was highly correlated with total Fe (R^2=0.86) based on the linear model. Deposition was assessed between days 10-13.

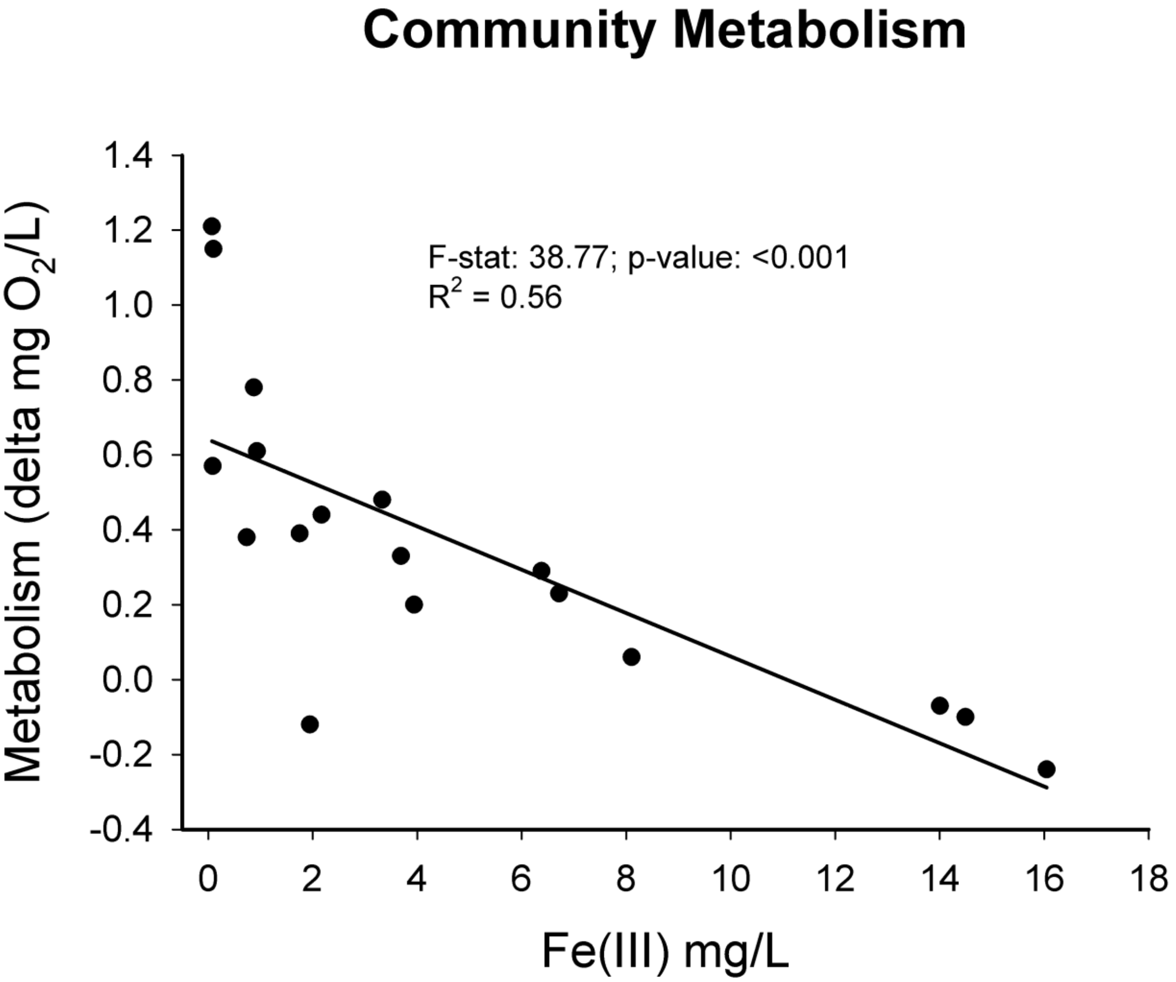

Figure S3.4 Community metabolism after exposure to Fe for 10 d. Metabolism was defined as the difference in dissolved oxygen concentrations between daylight and night observations. A linear model was used to test for a significant relationship of Fe and metabolism. DO = dissolved oxygen concentration.

Table S3.1 List of benthic macroinvertebrate taxa among all control replicates (n=3) after 14 d Fe exposure.

Ephemeroptera	**Plecoptera**	**Trichoptera**	**Chironomidae**	**Diptera**	**Miscellaneous**
Ameletus sp.	*Amphinemura* sp.	*Arctopsyche* sp.	Chironomini	*Antocha* sp.	*Heterlimnius* sp.
Baetis spp.	*Capnia* sp.	*Brachycentrus* sp.	Diamesinae	*Atherix* sp.	*Hydracarina* sp.
Diphetor sp.	*Claassenia* sp.	*Glossosoma* sp.	Orthocladiinae	*Dicranota* sp.	Oligochaeta
Drunella spp.	*Hesperoperla* sp.	*Hydropsyche* spp.	Tanypodinae	*Neoplasta* sp.	*Polycelis* sp.
Epeorus sp.	*Isoperla* sp.	*Hydroptila* spp.	Tanytarsini	*Pericoma* sp.	
Ephemerella sp.	*Pteronarcella* sp.	*Lepidostoma* sp.		*Simulium* sp.	
Paraleptophlebia sp.	*Skwala* sp.	*Oligophlebodes* sp.		*Tipula* sp.	
Rhithrogena sp.	*Sweltsa* sp.	*Rhyacophila* spp.			

Table S3.2 Summary of statistical results of linear models analyses used to test for significance for all endpoints. The predictor variable for all response variables was measured total iron. Degrees of freedom for all macroinvertebrate-associated endpoints and metabolism were d.f. = (1, 16); degrees of freedom for all algal endpoints were d.f = (1, 70).

Response Variable	R^2	Slope	Intercept	F-stat	p-value
Baetidae Adult Abundance	0.29	-0.0458	3.0090	7.94	0.012
Baetidae Adult Biomass	0.27	-0.0393	2.4065	7.41	0.015
Baetidae Larvae Abundance	0.94	-0.3410	5.4995	287.60	<0.001
Baetidae Larvae Biomass	0.76	-0.2514	3.2877	55.59	<0.001
Chironomidae Adult Abundance	0.56	-0.0408	5.7549	23.31	<0.001
Chironomidae Adult Biomass	0.47	-0.0391	2.9703	16.45	<0.001
Chironomidae Larvae Abundance	0.72	-0.0746	5.6468	45.21	<0.001
Chironomidae Larvae Biomass	0.43	-0.0508	2.4276	13.92	0.002
Chl-a Colonization Biomass	0.44	-0.1096	0.5339	56.92	<0.001
Cyanobacteria Colonization Biomass	0.46	0.1777	0.0053	63.14	<0.001
Diatom Colonization Biomass	0.65	-0.1328	0.4815	130.10	<0.001
Green Algae Colonization Biomass	0.59	-0.1008	0.2140	102.70	<0.001
Metabolism	0.69	-0.1561	1.0561	38.77	<0.001
Simuliidae Adult Abundance	0.27	-0.0590	4.4017	7.45	0.014
Simuliidae Larvae Abundance	0.01	-0.0100	1.7250	0.09	0.762
Simuliidae Adult Biomass	0.27	-0.0534	3.2140	7.28	0.015
Simuliidae Larvae Biomass	0.02	-0.0194	0.8207	1.47	0.242
Total Macroinvertebrate Adult Abundance	0.53	-0.0447	6.0469	20.77	<0.001
Total Macroinvertebrate Adult Biomass	0.42	-0.0454	4.0158	13.48	0.002
Total Macroinvertebrate Larvae Abundance	0.83	-0.0862	6.8234	89.14	<0.001
Total Macroinvertebrate Larvae Biomass	0.19	-0.0211	5.3328	5.17	0.037

Table S3.3 Summary of statistical results of ANCOVA models used to test interaction effects between Fe and life stage (larvae vs adult) for response variables in the AR community exposure. Degrees of freedom (d.f. = 1, 32) for all ANCOVA models were the same.

Response Variable	Predictor Variable	F-stat	p-value
Baetidae Abundance	Fe(III)	2.16	0.151
	Life Stage	60.38	<0.001
	Fe(III) x Life Stage	49.23	<0.001
Baetidae Biomass	Fe(III)	2.29	0.139
	Life Stage	11.94	0.001
	Fe(III) x Life Stage	33.42	<0.001
Chironomidae Abundance	Fe(III)	17.12	<0.001
	Life Stage	1.24	0.273
	Fe(III) x Life Stage	5.87	0.021
Chironomidae Biomass	Fe(III)	10.98	0.002
	Life Stage	21.88	<0.001
	Fe(III) x Life Stage	0.49	0.487
Simuliidae Abundance	Fe(III)	16.77	<0.001
	Life Stage	72.79	<0.001
	Fe(III) x Life Stage	7.61	0.009
Simuliidae Biomass	Fe(III)	15.79	<0.001
	Life Stage	78.43	<0.001
	Fe(III) x Life Stage	7.21	0.011
Total Macroinvertebrate Abundance	Fe(III)	22.28	<0.001
	Life Stage	69.3	<0.001
	Fe(III) x Life Stage	9.55	0.004
Total Macroinvertebrate Biomass	Fe(III)	17.23	<0.001
	Life Stage	149.88	<0.001
	Fe(III) x Life Stage	2.46	0.126

Table S3.4 Estimates of EC50 and EC20 values for all significant structural and functional endpoints examined.

Endpoint	EC50	EC50 95% CI	EC20	EC20 95% CI
Chl-*a* Colonization Biomass	0.072	(0.001, 31.186)	0.0004	(0.000, 1.01x10^5)
Diatom Colonization Biomass	0.374	(0.038, 3.670)	0.040	(0.000, 3.705)
Green Algae Colonization Biomass	0.479	(0.172, 1.331)	0.261	(0.050, 1.337)
Baetidae Larvae Biomass	0.729	(0.173, 3.074)	0.312	(0.026, 3.634)
Baetidae Larvae Abundance	1.068	(0.329, 3.461)	0.422	(0.053, 3.322)
Chironomidae Larvae Biomass	1.312	(0.012, 135.92)	0.015	(0.000, 1.53x10^6)
Simuliidae Larvae Biomass	1.700	(0.023, 123.67)	0.016	(0.000, 3.29x10^6)
Chironomidae Adult Biomass	2.067	(0.084, 50.888)	0.795	(0.061, 10. 276)
Baetidae Adult Biomass	3.189	(0.008, 120.73)	0.001	(0.000, 1.49x10^18)
Chironomidae Larvae Abundance	3.269	(0.527, 20.291	0.298	(0.000, 45.426)
Simuliidae Adult Abundance	4.270	(0.684, 26.681)	0.393	(0.004, 39.502)
Total Macroinvertebrate Larvae Abundance	4.591	(1.162, 18.135)	0.824	(0.059, 11.345)
Simuliidae Adult Biomass	5.040	(0.697, 36.48)	0.414	(0.003, 48.438)
Baetidae Adult Abundance	5.242	(0.424, 648.74)	0.011	(0.000, 5.31x10^9)
Total Macroinvertebrate Adult Biomass	7.224	(0.498, 104.71)	0.368	(0.000, 136.780)
Total Macroinvertebrate Adult Abundance	10.124	(0.586, 174.71)	0.694	(0.008, 54.312)
Metabolism	11.180	(0.403, 347.17)	0.626	(0.004, 89.438)
Cyanobacteria Colonization Biomass	12.637	(0.444, 359.08)	1.935	(1.237, 3.035)
Chironomidae Adult Abundance	13.450	(0.551, 326.11)	1.009	(0.024, 41.208)

Table S3.5 Summary of statistical results of single-factor repeated measures ANOVA used to test for overall differences in the timing of emerging adult abundance. The table also shows pairwise contrasts among Fe treatments and control observations over the entire exposure duration.

Response Variable	Contrast to Controls	t-ratio	p-value
Baetidae Emergence Abundance	Low	-5.816	<0.001
(d.f. = 5, 216)	Mid-Low	-2.571	0.046
(F-stat: 12.04; p-value: < 0.0001)	Mid	-3.384	0.004
	Mid-High	-3.909	<0.001
	High	-6.95	<0.001
Chironomidae Emergence Abundance	Low	-3.196	0.007
(d.f. = 5, 216)	Mid-Low	-1.175	0.627
(F-stat: 30.54; p-value:< 0.001)	Mid	-3.716	0.001
	Mid-High	-8.917	<0.001
	High	-9.328	<0.001
Simuliidae Emergence Abundance	Low	-5.035	<0.001
(d.f. = 5, 216)	Mid-Low	-5.574	<0.001
(F-stat: 49.53; p-value: < 0.0001)	Mid	-12.001	<0.001
	Mid-High	-11.011	<0.001
	High	-12.543	<0.001

Table S3.6 Permutational multivariate analysis of variance (PERMANOVA) results for overall Fe treatment effect.

	df	SS	MS	Pseudo-F	P(MC)
Fe (III)	5	193.22	38.64	7.44	0.001*
Residual	12	62.28	5.19		

APPENDIX C

SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 4

Contains:

Figures S4.1 to S4.5

Tables S4.1 to S4.6

Figure S4.1 *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 S4.2 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.

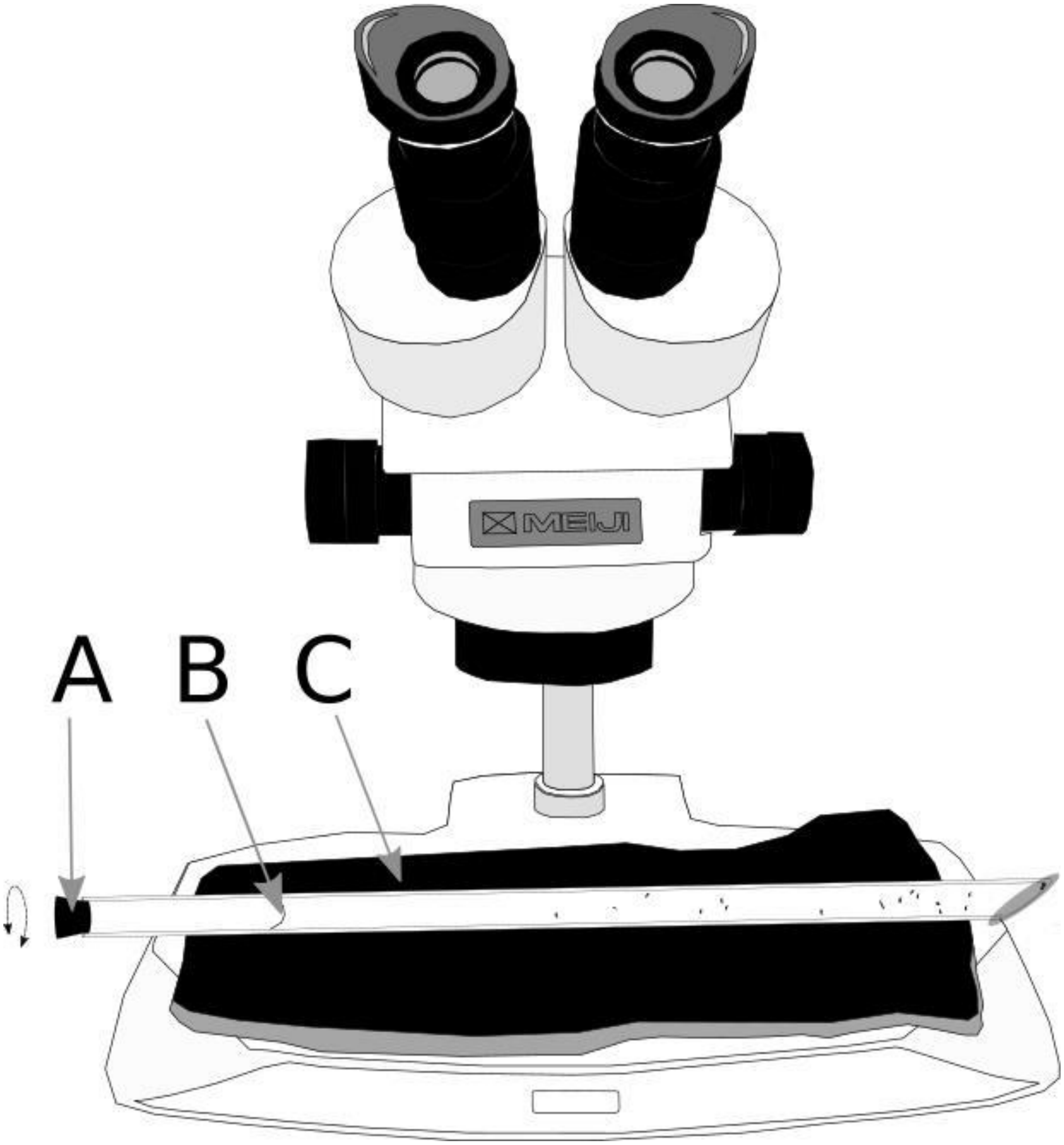

Figure S4.3 Exposure tubes were placed upright in a 500-ml beaker filled with chilled, highly-oxygenated water. Using a dissecting microscope (Meji EMZ-TR with 20x eyepieces), 10-20 organisms were transferred to each of the 24 exposure tubes using a 100-µl pipette. Rubber stoppers (A) were placed in 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

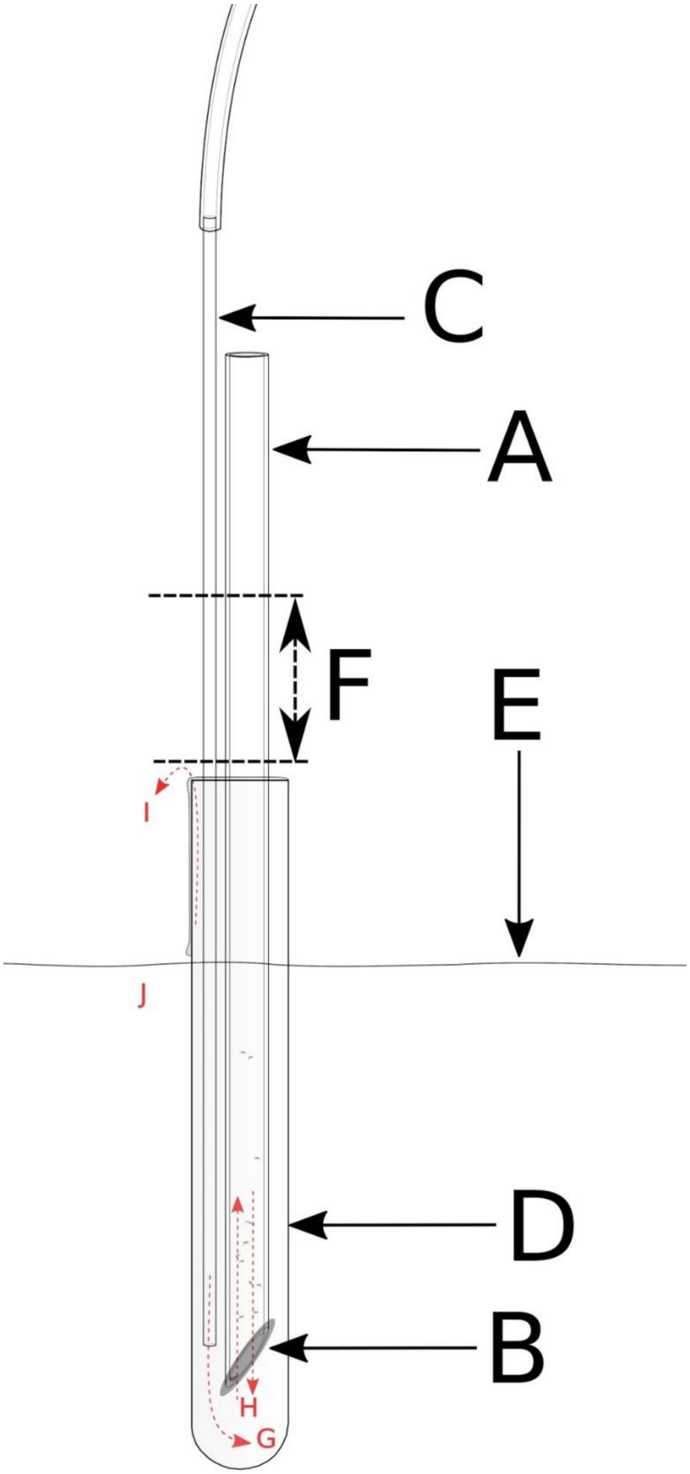

Figure S4.4 Organisms were contained in 24 exposure tubes (A). Fine stainless steel mesh (B) retained organisms during the experiment. Vinyl food-grade tubing and fine glass tubes (C; 3 mm ID, 5 mm OD) were used to transport toxicants and aerated water from a serial diluter to 24 test tubes (D; 23 mm 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 S4.5**) 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.

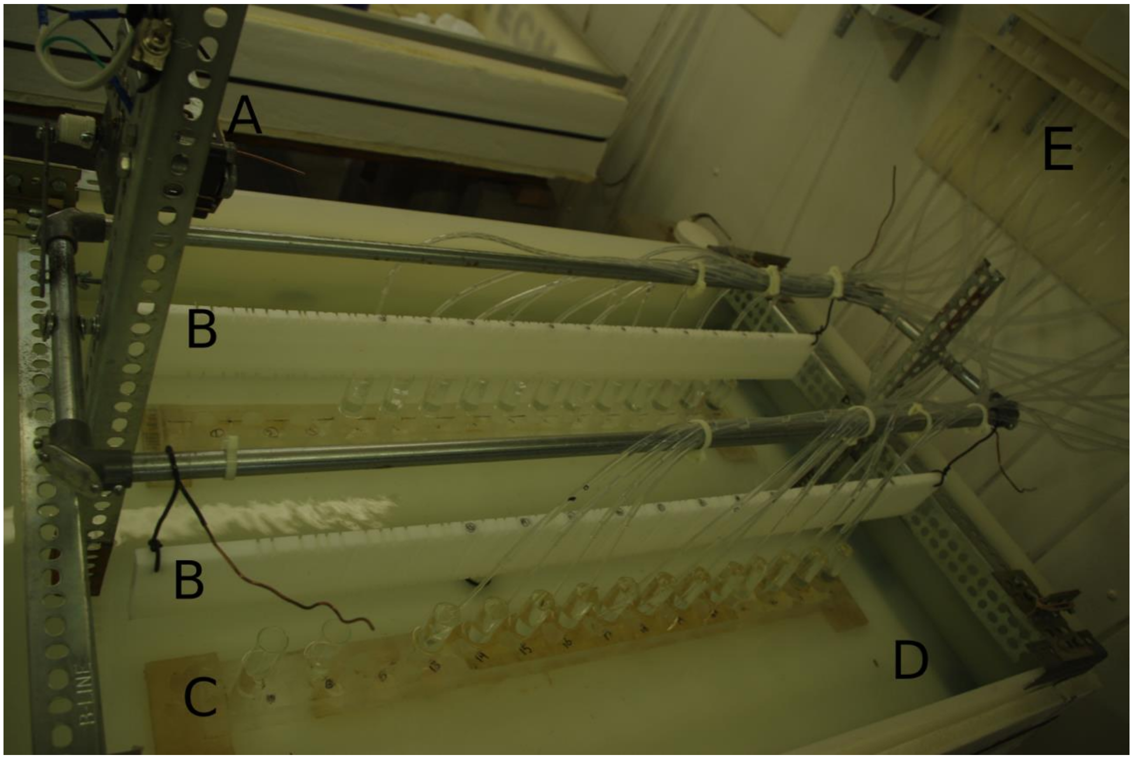

Figure S4.5 Exposure system. A motorized “teeter-totter” powered by a gear motor (A) was constructed of perforated angle iron and electrical conduit. White, high-density polyethylene boards (B) were suspended by wire on each arm of the teeter-totter. Dado cuts (grooves) accommodated rubber bands (not 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 S4.1 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 S4.2 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 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 S4.3 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 S4.4 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 S4.5 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	CCU	11.53	0.0023	<0.7 mm	-3.928	(-0.38, -7.48)	A
(0 - 5.1 CCU)	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	CCU	7.06	0.0137	<0.7 mm	-0.184	(-2.94, 2.57)	A
(0 - 7.0 CCU)	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 S4.6 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.

APPENDIX D

SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 5

Contains:

Figures S5.1 to S5.3

Tables S5.1 to S5.8

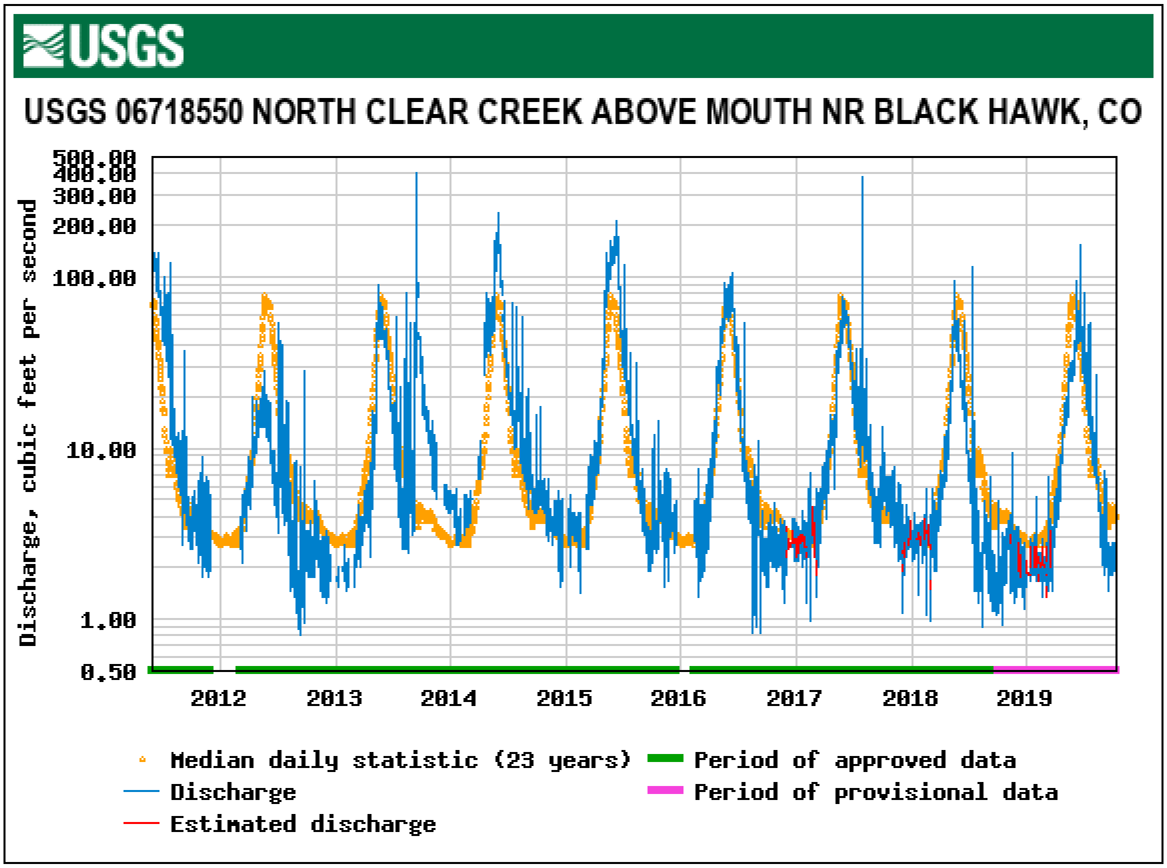

Figure S5.1 Stream discharge hydrograph of the North Fork Clear Creek, CO, USA from June 2011 through September 2019. Data were obtained from the United States Geological Survey Water Resources Division (https://waterdata.usgs.gov/nwis/uv?site_no=06718550). Note the rapid increase or "spikes" in stream discharge often observed during summer monsoon events.

Figure S5.2 Photo of Benthotorch tile adapter used to estimate algal biomass and community structure on different ceramic tile treatments in stream mesocosms. The main purpose of the tile adapter was to reduce sloughing of algae during measurements and consistently sample the same location on each respective tile.

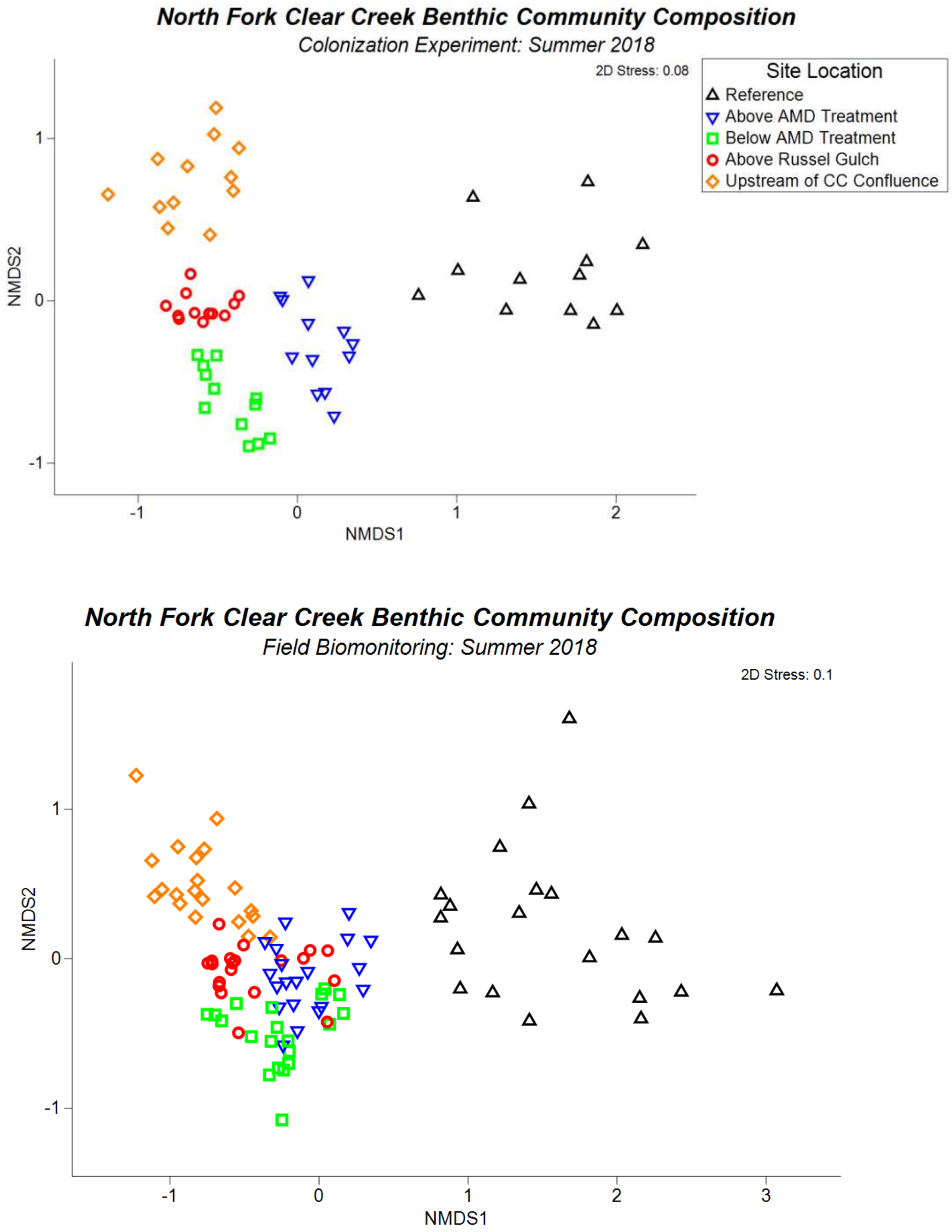

Figure S5.3 Nonmetric Multi-Dimensional Scaling (nMDS) plots comparing benthic macroinvertebrate communities sampled on naturally occurring substrate (lower panel) compared to benthic macroinvertebrate collected from colonization trays (upper panel) in summer 2018.

Table S5.1 Pairwise comparisons by similarity percentages (SIMPER) of the benthic macroinvertebrate taxa contributing to the upper 50% of difference among sites in the North Fork of Clear Creek, CO, prior (Fall 2011-Spring 2017) to the North Clear Creek Water Treatment Plant operation.

Before AMD Treatment						
Reference & Above AMD Treatment						
Average squared distance = 127.72						
	Reference	**Above AMD Treatment**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	4.52	0.438	17.7	2.24	13.85	13.85
Chironomidae	4.13	0.953	11.7	1.4	9.14	22.99
Heterlimnius sp.	3.48	0.27	11.4	1.73	8.95	31.94
Prostoia/Podmosta spp.	3.16	0.161	11.2	1.19	8.74	40.68
Rhyacophila spp.	3.36	0.412	9.88	1.52	7.74	48.42
Epeorus sp.	2.36	0.112	7.59	1.05	5.95	54.36
Reference & Above Russel Gulch						
Average squared distance = 136.71						
	Reference	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	4.52	0.11	20.2	2.74	14.77	14.77
Chironomidae	4.13	0.63	13.6	1.62	9.91	24.69
Heterlimnius sp.	3.48	0.0452	12.8	1.96	9.33	34.02
Rhyacophila spp.	3.36	0.0778	11.7	1.83	8.53	42.55
Prostoia/Podmosta spp.	3.16	0.0704	11.6	1.23	8.51	51.06
Reference & Upstream of CC Confluence						
Average squared distance = 122.53						
	Reference	**Upstream of CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Heterlimnius sp.	3.48	0.15	12.2	1.84	9.95	9.95
Prostoia/Podmosta spp.	3.16	0.0904	11.6	1.22	9.46	19.41
Rhyacophila spp.	3.36	0.257	10.7	1.64	8.76	28.18
Chironomidae	4.13	1.5	9.11	1.08	7.43	35.61
Baetis spp.	4.52	2.71	7.96	0.91	6.5	42.11
Epeorus sp.	2.36	0.063	7.76	1.07	6.33	48.44
Taenionema sp.	2.25	0	6.88	0.99	5.61	54.05
Above AMD Treatment & Above Russel Gulch						
Average squared distance = 4.21						
	Above AMD Treatment	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Chironomidae	0.953	0.63	1.41	0.77	33.62	33.62
Baetis spp.	0.438	0.11	0.551	0.62	13.1	46.72
Rhyacophila spp.	0.412	0.0778	0.55	0.55	13.08	59.8
Above AMD Treatment & Upstream of CC Confluence						
Average squared distance = 24.01						
	Above AMD Treatment	**Upstream of CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	0.438	2.71	9.55	0.81	39.77	39.77
Arctopsyche sp.	0.0252	0.724	2.79	0.4	11.63	51.4
Above Russel Gulch & Upstream of CC Confluence						
Average squared distance = 24.77						
	Above Russel Gulch	**Upstream of CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	0.11	2.71	10.9	0.86	43.91	43.91
Arctopsyche sp.	0.0984	0.724	2.78	0.41	11.23	55.14

Table S5.2 Pairwise comparisons by similarity percentages (SIMPER) of the benthic macroinvertebrate taxa contributing to the upper 50% of difference between Reference and downstream site observations in the North Fork of Clear Creek, CO, following North Clear Creek Water Treatment Plant operation (Summer 2017-Fall 2018).

After AMD Treatment

Reference & Above AMD Treatment

Average squared distance = 115.68

	Reference	Above AMD Treatment				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Epeorus sp.	3	0.466	8.77	1.18	7.58	7.58
Heterlimnius sp.	3.32	0.619	8.34	1.54	7.21	14.79
Rhyacophila spp.	3.27	0.598	8.22	1.53	7.1	21.89
Baetis spp.	4.22	2.1	8	1.04	6.92	28.81
Zapada sp.	2.53	0.097	7.51	1.22	6.5	35.31
Prostoia/Podmosta spp.	2.05	0.115	7.35	0.9	6.35	41.66
Tanytarsini	2.22	0.207	6.26	0.86	5.41	47.07
Taenionema sp.	2.06	0.181	5.48	1.01	4.73	51.8

Reference & Below AMD Treatment

Average squared distance = 128.02

	Reference	Below AMD Treatment				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	4.22	1.37	10.7	1.32	8.38	8.38
Heterlimnius sp.	3.32	0.246	10.3	1.85	8.07	16.45
Rhyacophila spp.	3.27	0.306	9.69	1.81	7.57	24.03
Epeorus sp.	3	0.325	9.4	1.25	7.34	31.37
Zapada sp.	2.53	0.0139	7.88	1.26	6.15	37.52
Prostoia/Podmosta spp.	2.05	0.0139	7.58	0.91	5.92	43.44
Chloroperlidae	2.45	0.105	6.77	1.31	5.29	48.73
Tanytarsini	2.22	0.105	6.59	0.89	5.15	53.88

Reference & Above Russel Gulch

Average squared distance = 125.02

	Reference	Above Russel Gulch				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Heterlimnius sp.	3.32	0.262	10.2	1.84	8.18	8.18
Rhyacophila spp.	3.27	0.427	9.05	1.7	7.24	15.42
Epeorus sp.	3	0.665	8.33	1.12	6.66	22.08
Hydropsyche spp.	0.0974	2.53	7.95	1.18	6.36	28.44
Zapada sp.	2.53	0	7.94	1.26	6.35	34.79
Prostoia/Podmosta spp.	2.05	0	7.63	0.91	6.1	40.9
Tanytarsini	2.22	0.0462	6.8	0.9	5.44	46.33
Taenionema sp.	2.06	0	6.04	1.07	4.83	51.17

Reference & Upstream of CC Confluence

Average squared distance = 131.74

	Reference	Upstream of CC Confluence				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Hydropsyche spp.	0.0974	3.4	13.6	1.55	10.35	10.35
Heterlimnius sp.	3.32	0.287	10.3	1.74	7.82	18.17
Rhyacophila spp.	3.27	0.245	10.1	1.86	7.66	25.83
Epeorus sp.	3	0.561	8.11	1.14	6.16	31.99
Zapada sp.	2.53	0.101	7.51	1.21	5.7	37.69
Prostoia/Podmosta spp.	2.05	0.0539	7.5	0.91	5.69	43.38
Taenionema sp.	2.06	0	6.04	1.07	4.59	47.97
Simulium sp.	0.989	1.97	5.82	0.68	4.42	52.39

Table S5.3 Pairwise comparisons by similarity percentages (SIMPER) of the benthic macroinvertebrate taxa contributing to the upper 50% of difference among downstream site observations in the North Fork of Clear Creek, CO, following North Clear Creek Water Treatment Plant operation (Summer 2017-Fall 2018).

After AMD Treatment						
Above AMD Treatment & Below AMD Treatment						
Average squared distance = 20.09						
	Above AMD Treatment	**Below AMD Treatment**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	2.1	1.37	5.81	0.83	28.91	28.91
Orthocladinae	2.21	1.66	2.71	0.82	13.51	42.42
Hydropsyche spp.	0.413	0.634	2.25	0.45	11.19	53.6
Above AMD Treatment & Upstream of CC Confluence						
Average squared distance = 47.76						
	Above AMD Treatment	**Upstream of CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	2.1	5.37	14.8	1.18	30.89	30.89
Hydropsyche spp.	0.413	3.4	12.1	1.37	25.37	56.26
Above AMD Treatment & Above Russel Gulch						
Average squared distance = 31.23						
	Above AMD Treatment	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	Av.Sq.Dist	Sq.Dist/SD	Contrib%	Cum.%
Baetis spp.	2.1	3.71	10.2	1.04	32.64	32.64
Hydropsyche spp.	0.413	2.53	6.99	1.07	22.37	55.01
Below AMD Treatment & Upstream of CC Confluence						
Average squared distance = 50.23						
	Below AMD Treatment	**Upstream of CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	1.37	5.37	19.2	1.5	38.19	38.19
Hydropsyche spp.	0.634	3.4	11.7	1.35	23.29	61.47
Below AMD Treatment & Above Russel Gulch						
Average squared distance = 29.35						
	Below AMD Treatment	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	1.37	3.71	12.2	1.15	41.55	41.55
Hydropsyche spp.	0.634	2.53	6.95	1.09	23.7	65.24
Upstream of CC Confluence & Above Russel Gulch						
Average squared distance = 32.73						
	Upstream of CC Confluence	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	5.37	3.71	8.22	0.67	25.1	25.1
Simulium sp.	1.97	0.673	6.46	0.68	19.73	44.83

Table S5.4 Pairwise comparisons by similarity percentages (SIMPER) of the benthic macroinvertebrate taxa contributing to the upper 50% of difference among sites in the North Fork of Clear Creek, CO following the 60 d colonization experiment the first summer (Year 1; 2017) following North Clear Creek Water Treatment Plant operations.

Year 1						
Reference & Above AMD Treatment						
Average squared distance = 111.99						
	Reference	**Above AMD Treatment**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	4.55	0.347	18	3.65	16.11	16.11
Taenionema sp.	4.09	0	16.8	7.72	15	31.11
Zapada sp.	3.74	0.0578	13.9	3.39	12.41	43.53
Prostoia/Podmosta spp.	2.78	0	10.9	1.17	9.78	53.3
Reference & Below AMD Treatment						
Average squared distance = 120.51						
	Reference	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	4.55	0.0578	20.3	6.22	16.85	16.85
Taenionema sp.	4.09	0	16.8	7.72	13.94	30.8
Zapada sp.	3.74	0	14.3	3.63	11.86	42.65
Prostoia/Podmosta spp.	2.78	0	10.9	1.17	9.09	51.74
Reference & Upstream of CC Confluence						
Average squared distance = 109.12						
	Reference	**Upstream of CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Taenionema sp.	4.09	0.126	15.9	5.44	14.53	14.53
Hydropsyche spp.	0	3.61	13.4	3.12	12.31	26.84
Zapada sp.	3.74	0.252	12.6	2.91	11.53	38.37
Prostoia/Podmosta spp.	2.78	0.063	10.6	1.15	9.75	48.12
Isoperla sp.	3.16	0.126	9.4	3.59	8.61	56.74
Above AMD Treatment & Below AMD Treatment						
Average squared distance = 3.42						
	Reference	**Above AMD Treatment**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Orthocladinae	2.06	1.88	0.993	0.79	29.02	29.02
Chloroperlidae	0.831	0.231	0.753	0.79	22.01	51.03
Above AMD Treatment & Above CC Confluence						
Average squared distance = 48.90						
	Reference	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	0.347	5.67	28.9	3.84	59.04	59.04
Below AMD Treatment & Above CC Confluence						
Average squared distance = 51.26						
	Reference	**Upstream of CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	0.0578	5.67	31.8	5.36	62.01	62.01

Table S5.5 Pairwise comparisons by similarity percentages (SIMPER) of the benthic macroinvertebrate taxa contributing to the upper 50% of difference between Reference and downstream site observations in the North Fork of Clear Creek, CO, following the 60 d colonization experiment the second summer (Year 2; 2017) following North Clear Creek Water Treatment Plant operations.

Year 2						
Reference & Above AMD Treatment						
Average squared distance = 89.33						
	Reference	**Above AMD Treatment**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Zapada sp.	3.98	0.149	15.1	3.02	16.92	16.92
Epeorus sp.	2.92	0.0578	11.4	1.46	12.74	29.67
Diphetor sp.	2.43	0.0578	8.5	1.24	9.52	39.19
Capnia sp.	2.84	0.173	8.45	1.79	9.46	48.64
Reference & Below AMD Treatment						
Average squared distance = 126.56						
	Reference	**Below AMD Treatment**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Zapada sp.	3.98	0.0578	15.7	3.35	12.44	12.44
Baetis spp.	0	0.731	12.8	1.91	10.08	22.52
Epeorus sp.	2.92	0	11.7	1.48	9.23	31.75
Isoperla sp.	3.38	0.116	10.9	3.79	8.59	40.34
Hydropsyche spp.	0.241	2.74	9.07	1.05	7.16	47.5
Diphetor sp.	2.43	0	8.74	1.25	6.91	54.41
Reference & Above Russel Gulch						
Average squared distance = 125.45						
	Reference	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Zapada sp.	3.98	0	16.2	3.55	12.89	12.89
Hydropsyche spp.	0.241	4.16	16	2.66	12.73	25.62
Epeorus sp.	2.92	0	11.7	1.48	9.31	34.93
Rhyacophila spp.	2.92	0	8.97	2.31	7.15	42.08
Capnia sp.	2.84	0.0578	8.94	1.94	7.13	49.21
Arctopsyche sp.	2.94	0.0578	8.84	2.16	7.04	56.26
Reference & Upstream of CC Confluence						
Average squared distance = 140.35						
	Reference	**Above CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Hydropsyche spp.	0.241	6.06	34.2	5.34	24.38	24.38
Zapada sp.	3.98	0	16.2	3.55	11.52	35.9
Simulium sp.	1.92	4.65	10.7	1.12	7.65	43.54
Epeorus sp.	2.92	0.652	9.05	1.25	6.45	49.99
Diphetor sp.	2.43	0	8.74	1.25	6.23	56.22

Table S5.6 Pairwise comparisons by similarity percentages (SIMPER) of the benthic macroinvertebrate taxa contributing to the upper 50% of difference among downstream sites observations in the North Fork of Clear Creek, CO, following the 60 d colonization experiment the second summer (Year 2; 2017) following North Clear Creek Water Treatment Plant operations.

Year 2						
Above AMD Treatment & Below AMD Treatment						
Average squared distance = 34.79						
	Above AMD Treatment	**Below AMD Treatment**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	3.3	0.731	9.02	1.05	25.91	25.91
Hydropsyche spp.	0.87	2.74	7.08	0.91	20.34	46.25
Arctopsyche sp.	2.03	0.427	5.29	0.96	15.19	61.45
Above AMD Treatment & Above Russel Gulch						
Average squared distance = 42.88						
	Above AMD Treatment	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Hydropsyche spp.	0.87	4.16	12.2	1.65	28.45	28.45
Orthocladinae	4.28	1.51	8.35	1.69	19.48	47.93
Arctopsyche sp.	2.03	0.0578	6.08	1.04	14.18	62.1
Above AMD Treatment & Above CC Confluence						
Average squared distance = 70.85						
	Above AMD Treatment	**Above CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Hydropsyche spp.	0.87	6.06	28	2.71	39.58	39.58
Simulium sp.	1.39	4.65	12.9	1.31	18.25	57.83
Below AMD Treatment & Above Russel Gulch						
Average squared distance = 30.23						
	Below AMD Treatment	**Above Russel Gulch**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	0.731	4.79	17.4	2.26	57.49	57.49
Below AMD Treatment & Above CC Confluence						
Average squared distance = 79.33						
	Below AMD Treatment	**Above CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Baetis spp.	0.731	6.33	32.2	3.18	40.54	40.54
Simulium sp.	0.116	4.65	21.3	2.65	26.86	67.39
Above Russel Gulch & Above CC Confluence						
Average squared distance = 43.73						
	Above Russel Gulch	**Above CC Confluence**				
Variable	**Av.Value**	**Av.Value**	**Av.Sq.Dist**	**Sq.Dist/SD**	**Contrib%**	**Cum.%**
Simulium sp.	0.173	4.65	20.9	2.51	47.77	47.77
Orthocladinae	1.51	3.32	5.33	1.12	12.19	59.95

Table S5.7. Target exposure concentrations and average measured concentrations (µg/L) (±s.d.) of Cu and Zn in the North Fork Clear Creek community exposures.

	Target Cu	Measured Cu	Target Zn	Measured Zn	CCU
Low	6	9.9 (0.32)	94	85 (5.8)	4.3
Mid-Low	12	18.2 (0.72)	188	162 (3.5)	8.1
Mid	25	38.2 (2.21)	375	313 (29.2)	16.3
Mid-High	50	46.1 (1.61)	750	673 (22.2)	26.7
High	100	93.9 (0.96)	1500	1411 (37.9)	55.2

CCU = the sum of the measured concentration for each metal divided by the U.S. EPA hardness-adjusted criterion values for Cu (4.3 µg/L) and Zn (42.3 µg/L).

Table S5.8. Water quality characteristics in stream mesocosms from the North Fork Clear Creek Benthic Community Exposure. Values are reported as mean (+/- s.d.).

Treatment	Temperature (°C)	pH	Conductivity (µS/cm)
Control	10.21 (0.12)	7.33 (0.03)	75.41 (0.79)
Low	10.26 (0.07)	7.27 (0.01)	75.55 (1.29)
Mid-Low	10.2 (0.12)	7.28 (0.03)	74.91 (1.49)
Mid	10.35 (0.04)	7.24 (0.02)	75.16 (1.95)
Mid-High	10.16 (0.09)	7.24 (0.01)	76.73 (0.31)
High	10.26 (0.04)	7.19 (0.04)	77.86 (1.55)