Canadian Journal of Microbiology

Impact of sample collection on prokaryotic and eukaryotic diversity of niche environments of the oil-sand mining impacted Athabasca River.

Journal:	<i>Canadian Journal of Microbiology</i>
Manuscript ID	cjm-2021-0058.R1
Manuscript Type:	Article
Date Submitted by the Author:	25-Apr-2021
Complete List of Authors:	<p>Bergsveinson, Jordyn; Environment and Climate Change Canada, Watershed Hydrology and Ecology Research Division</p> <p>Lawrence, John; Environment and Climate Change Canada, Watershed Hydrology and Ecology Research Division</p> <p>Schebel, Alixandra; Environment and Climate Change Canada, Watershed Hydrology and Ecology Research Division</p> <p>Wasserscheid, Jessica; National Research Council, Energy, Mining and Environment Research Centre</p> <p>Roy, Julie; Environment and Climate Change Canada, Watershed Hydrology and Ecology Research Division</p> <p>Conly, F; Environment and Climate Change Canada, Watershed Hydrology and Ecology Research Division</p> <p>Sanschagrin, Sylvie; National Research Council, Energy, Mining and Environment Research Centre</p> <p>Korber, Darren; University of Saskatchewan, Department of Food and Bioproduct Science</p> <p>Tremblay, Julien; National Research Council, Energy, Mining and Environment Research Centre</p> <p>Greer, Charles W.; National Research Council, Energy, Mining and Environment Research Centre</p> <p>Droppo, Ian; Environment and Climate Change Canada, Canada Centre for Inland Waters</p>
Keyword:	biofilm, planktonic, suspended load, Oil Sands, rivers
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

SCHOLARONE™
Manuscripts

Impact of sample collection on prokaryotic and eukaryotic diversity of niche environments of the oil-sand mining impacted Athabasca River.

Jordyn Bergsveinson^{a,#}, John R. Lawrence^a, Alixandra Schebel^a, Jessica Wasserscheid^b, Julie Roy^a, F. Malcolm Conly^a, Sylvie Sanschagrin^b, Darren R. Korber^c, Julien Tremblay^b, Charles W. Greer^b and Ian G. Droppo^d

^a Environment and Climate Change Canada, National Hydrology Research Centre, Saskatoon, SK Canada

^b Energy, Mining and Environment Research Centre, National Research Council Canada, Montreal, QC, Canada

^c Department of Food and Bioproduct Sciences, University of Saskatchewan, Saskatoon, SK, Canada

^d Environment and Climate Change Canada, Canada Centre for Inland Waters, Burlington, ON, Canada

Running Head: Collection method impacts aquatic microbial community

#Address correspondence to Jordyn Bergsveinson, jordyn.broadbent@canada.ca

Keywords: riverine ecology, biofilm, planktonic, suspended load, hydrocarbon, integrated sampling, oil sands.

Abstract

Microbial communities are an important aspect of overall riverine ecology; however, appreciation of the effects of anthropogenic activities on unique riverine microbial niches, and how the collection of these samples affects the observed diversity and community profile is lacking. We analyzed prokaryotic and eukaryotic communities from surface water, biofilm, suspended load niches along a gradient of oil sands-related contamination in the Athabasca River (Alberta, Canada), with suspended load or particle-associated communities collected either via Kenney Sampler or centrifugation manifold. At the level of phyla, different niche communities were highly similar to one another and across locations. However, there were significant differences in the abundance of specific genera amongst different niches and across sampling locations. A generalized linear model revealed that use of the Kenney Sampler resulted in more diverse bacterial and eukaryotic suspended load community than centrifugal collection, though “suspended load” communities collected by any means remained stably diverse across locations. Though there was influence of water quality parameters on community composition, all sampled sites support diverse bacterial and eukaryotic communities regardless of the degree of contamination, highlighting the need to look beyond ecological diversity as means of assessing ecological perturbations, and consider collecting samples from multiple niche environments.

Keywords: riverine ecology, biofilm, planktonic, suspended load, hydrocarbon, integrated sampling, oil sands.

Introduction

There are a variety of niches of riverine ecosystems that may harbour specific microbial communities, including: planktonic, sediment, flocs, biofilms, as well as those associated with suspended sediments or the suspended load (sediment particles maintained in suspension). The taxonomic profile of riverine bacterioplanktonic communities (Portillo et al. 2012; Staley et al. 2013, 2014) and sediment bacterial communities (Drury et al. 2013; Reid et al. 2019, 2020) have been previously examined. The influence of a variety of physical and chemical factors on these communities (Portillo et al. 2013; Fortunato et al. 2012) including pollution gradients and land use (Staley et al. 2014) has also been assessed. Essential conclusions have been drawn with regard to suspended bacterioplankton in that they appear to be dominated by several phyla, including α -, β -, and γ -*Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, and *Cyanobacteria* (Zwart et al. 2002), which provides some insight into the potential functional capacity of these microbial niches. This is of utility for riverine systems impacted by anthropogenic activities. For instance, a few studies have applied genomic approaches to assess the bed and suspended sediments in Athabasca River tributaries (Reid et al. 2020; Yergeau et al. 2012), demonstrating that there were changes in the microbial processes of bed and suspended sediments associated with increasing polycyclic aromatic compound concentrations resulting from Oil Sands mining activities (Reid et al. 2020). The utility of exploring the diversity of different microbial niches within and between sample sites as a reliable indicator of anthropogenic effects is unclear, given the potential stability and response to environmental parameters of each niche community.

An essential component of ecology involves the modeling of specific taxa distributions throughout an environment, meaning that how these taxa are sampled is of utmost importance (Komori et al. 2020). It is generally assumed that presence-only data for organisms (i.e.,

amplicon sequencing of microbial taxonomic markers) is obtained randomly from the region of interest; however, it is often the case that this collection is already inherently biased toward the areas easily accessed (Komori et al. 2020). By extrapolation, this would also include those organisms that are collected by various means within these preferred areas. Of particular interest is the how the observed community diversity of riverbed sediments and bulk total suspended solids (TSS) may be influenced by the means of collection. These communities are frequently collected via filling of sterile cryotubes (i.e., plastic 5 mL to 50 mL cryotubes) (Reid et al. 2020), or simply by collection with stainless steel shovels or spatulas (Yergeau et al. 2012). In the case of the suspended load, samples may be obtained through passive sampling using distinct devices such as the Phillips Tube (Phillips et al. 2000), Kenney sampler (Kenney 1985), or by active pumping and centrifugation processes (Reid et al. 2020). This prompts speculation that much like anthropogenic contamination, active vs. passive sampling of suspended load communities would have a significant impact on the observed community in both “pristine” and/or contaminated areas.

In the present study, we analyzed the sequences of both the bacterial (16S rRNA gene) and eukaryotic (18S rRNA gene) communities of the planktonic, biofilm, and suspended load niches along the Athabasca River main stem. These sample types allowed for assessment of how these niche communities differ from each other along a gradient of hydrocarbon contamination and in the case of the suspended load, if and how the taxonomic distribution and diversity of the suspended load community is influenced by the specific method of sampling.

91 **Materials and Methods**

92 *Sample collection scheme and methodology*

93 Figure 1 shows the primary sampling sites implemented by Environment and Climate
94 Change Canada (ECCC) along the Athabasca River, including M4 (upstream and downstream of
95 the Muskeg River); M7 (within the area of oil sands (OS) activity, downstream of MacKay and
96 Ells rivers); and M9 (the furthest downstream and prior to river entry into the Peace-Athabasca
97 Delta). Additional sampling sites were established upstream of Fort McMurray (M0), which
98 serves as a “pristine” reference site for areas of OS activity. Samples of each type (surface water,
99 biofilm and suspended load) were collected in triplicate at sites M0, M4, M7 and M9 throughout
100 the 2013 and 2014 year according to the schedule described in File S2. All samples were
101 transported in iced coolers to the laboratory for storage at -80°C until processing for DNA
102 extraction.

103 *Biofilm tile*

104 Unglazed ceramic tiles (2×2 inch) were mounted within a stainless-steel flow-through
105 chamber and suspended in the river at sites (M0, M4, M7 and M9) attached to a floatation buoy
106 and anchored to the river bottom. The incubation system was designed to allow water flow
107 through the device while protecting the developing biofilm tile (BT) from physical disturbance
108 by floating debris (Figure S1).

109 *Centrifugation*

110 Suspended sediments or suspended load samples (CSL) were collected by portable
111 continuous-flow centrifugation (Alfa-Laval model WSB 103B). Water was pumped (4 L min⁻¹)
112 to the centrifuge by a 5C-MD March submersible pump located in the middle of the river at mid-

depth. Recovery of suspended loads was greater than 90% (calculated by the % difference between inflow and outflow). Sediment collected in the centrifuge bowl was immediately removed and placed in sterile 50 mL Falcon Tubes (Fisher Scientific, ON).

Kenney sampler

The manifold used was modified from the suspended load collection system originally described by Kenney (Kenney 1985; Kenney and Mollison 1980) (Figure S2). In short, the system consists of an outer tube with a collection aperture (3/8th inch) orientated to the flow and directing suspended materials into an inner glass collection chamber.

Surface water collection

Surface waters were collected from each location by filling sterile 2 L bottles and filtered through 0.22 mm filters. It should be noted that the biofilm tile, centrifugation, and Kenney sampler systems are not isokinetic (i.e., do not collect at consistent speed) in their approach.

Sample processing, DNA extraction, and sequencing

DNA was extracted from planktonic, biofilm, sediment, and suspended load samples stored at -80 °C, using the MoBio Powersoil Kit (Qiagen Sciences Inc., MD, USA) according to the manufacturer's protocol. DNA was used to prepare 16S rRNA gene sequencing libraries using the primers 515F-Y (5'-GTG YCA GCM GCC GCG GTA A-3') and 926R (5'-CCG YCA ATT YMT TTR AGT TT-3') targeting archaeal and bacterial V4-5 region (Parada et al. 2016), and 18S rRNA gene libraries using 565F (5'-CCA GCA SCY GCG GTA ATT CC-3') and 948R (5'-ACT TTC GTT CTT GAT YRA-3') targeting the V4 region (Stoeck et al. 2010), according to the protocol of Illumina's 16S Metagenomic Sequencing Library Preparation (Part Number

134 15044223 Rev. B). Libraries were sequenced on the Illumina MiSeq System using the paired-end
135 V2 (500 cycle) kit, using a 10% PhiX spike-in.

136 *16S and 18S rRNA gene amplicon sequencing read analysis*

137 Sequencing data was analyzed using AmpliconTagger, the National Research Council
138 Canada's amplicon pipeline (Tremblay and Yergeau 2019). Briefly, raw reads were scanned for
139 sequencing adapters and PhiX spike-in sequences and remaining reads were merged using their
140 common overlapping part with FLASH (Magoc and Salzberg 2011). Merged sequences were
141 filtered based on quality (Phred) score such that reads having an average qscore lower than 30 or
142 having more than one undefined base (N) or having 7 bases lower than phred score 15 were
143 discarded. Remaining sequences were clustered at 100% identity and then clustered/denoised at
144 99% identity (VSEARCH v2.7.1) (Rognes et al. 2016). Clusters having abundances lower than 3
145 were discarded. Remaining clusters were scanned for chimeras with VSEARCH's version of
146 UCHIME denovo and UCHIME reference (Rognes et al. 2016; Edgar et al. 2011) and clustered
147 at 97% (VSEARCH) to form the final clusters/OTUs. Bacterial and eukaryote OTUs were then
148 assigned a taxonomic lineage with the RDP classifier (Wang et al. 2007) using the
149 AmpliconTagger 16S and 18S training sets based on the Silva R128 database (Quast et al. 2013),
150 respectively. The RDP classifier gives a score (0 to 1) to each taxonomic depth of each OTU.
151 Each taxonomic depth having a score ≥ 0.5 was kept, reconstructing the final lineage.
152 Taxonomic lineages were combined with the cluster abundance matrix obtained above to
153 generate a raw OTU table. From that raw OTU table, an OTU table containing only bacterial or
154 eukaryote organisms was generated. Five hundred iterative 1,000 read rarefactions were
155 performed on this latter OTU table and the average number of reads of each OTU of each sample
156 then computed to obtain a consensus rarefied OTU table.

Chemical analysis

All collected samples were analyzed for various chemical parameters (Figure 2; Table S2) including nutrients, metals, major ions, total and alkylated PACs, and BTEX fractions, at several Environment Climate Change Canada (ECCC) National Laboratory for Environmental Testing (NLET) facilities according to their specific analyte specialization. Published analytical protocols were used for all analyses, as described in (Glozier et al. 2013).

Statistical and graphical analyses

For all analysis, rarefied data was used for both bacterial and eukaryotic communities (File S1). The Inverse Simpson diversity metric was calculated for all samples and analysed using a general linearized model (GLM) in order to determine the effect of sampling year, sample location, and sample collection method on microbial community diversity (File S1). Generalized linear hypothesis tests of differences between selected treatment effects were performed using contrast coding and P-values adjusted for multiple comparisons via the multcomp package (v 1.4-13) (30) for R (v 3.5.1) (R Core Team 2018) and diversity values plotted via ggplot2 (Wickham 2016). Tukey's Honest Significant Test HSD) was performed as a post-hoc analysis in R to confirm the results of the GLM (Miller 1981; Yandell 1997) (File S1). Calypso (v 8.84) (Zakrzewski et al. 2016) was used to perform LEfSe analysis (Segata et al. 2011) of site biomarkers and redundancy analysis (RDA) of different samples.

Results

A total of 440 environmental samples were collected across two field seasons, by four distinct collection methods, from three sites (M4, M7, M9) along a gradient of hydrocarbon contamination within the Athabasca River (Glozier et al. 2018; Oil Sands Information Portal 2020). A reference site (M0) located upstream of oil sands mining activity and the bituminous McMurray Formation was also sampled (Figure 1). Site M4 is directly downstream of the city of Fort McMurray WWTP outflow and corresponds with an increase in nutrient levels, which decrease in concentration at downstream sites (Figure 2; File S2; Glozier et al. 2018). Site M7 is directly within the area of oil sands mining activities and is associated with an increase in hydrocarbon (i.e., polycyclic aromatic hydrocarbon, etc.) contamination, though concentrations within water samples are frequently below analyte detection levels (Figure 2; File S2; Glozier et al. 2018). Site M9 is considerably downstream of both WWTP and mining activity exposures and thus considered as a “recovery” site, though specific metals are found at concentrations similar to sites M4 and M7 (Glozier et al. 2018; Oil Sands Information Portal <http://osip.alberta.ca/map/>).

The microbial (16S) and eukaryotic (18S) taxonomic DNA markers from all samples were processed and sequenced, with both types of communities receiving good sequencing coverage as shown by read distribution and rarefaction (File S1). A total of 61 bacterial and 32 eukaryotic phyla were identified across all samples.

Community diversity of different sample types across years

To assess the broad impact of sampling year, method, and site on microbial diversity, the Inverse Simpson’s diversity metric was calculated for both 16S and 18S communities, and the data fit to a generalized linear model (GLM) to assess the effect of study variables on microbial diversity through linear regression. Analyses demonstrated that sampling locations along a

gradient of hydrocarbon contamination had no significant impact on community diversity (i.e., the number and proportions of unique bacterial or eukaryotic organisms) despite the difference in contamination present at each site as derived from the Joint Canada-Alberta Oil Sands Monitoring program (Glozier et al. 2018) (File S2). Effect of sampling year was detected within the eukaryotic data via Tukey Post-Hoc analysis ($p < 0.005$) however the contribution of this effect to the overall GLM was negligible (File S1). The most notable contribution to observed differences in bacterial diversity was the means of suspended load collection (Figure 3). Specifically, suspended load bacterial communities collected via the Kenney sampler (KSL) were the only communities with significantly different diversity scores from biofilm tiles (BT), surface water (SW) and from centrifuged suspended load (CSL) communities collected from the same area (all $p < 0.0005$), as confirmed by both the GLM and Tukey Post-Hoc analysis (File S1). Bacterial BT, SW and CSL communities did not differ from one another in terms of diversity. For eukaryotic communities, the diversity of CSL and SW communities were not statistically different from one another, however all other sample types differed from one another in terms of diversity ($p < 0.05 - 0.005$) (Figure 4; File S1). Though statistically significant, the smallest difference in diversity scores between the eukaryotic communities exists between CSL and KSL samples (File S1).

Redundancy analysis (RDA) further explored how much of the observed variation in OTU abundance across different samples was related to, or explained by, different study parameters, using constrained ordination. When comparing the abundance of OTUs within different niche types (Figure 4a,c) or from different locations (Figure 4b,d), it was evident that sample type had a more pronounced effect on the spatial separation (i.e., variance) of samples than location for both bacterial and eukaryotic communities. For both organism groups, BT and

KSL samples clustered distinctly from SW and CSL samples which overlapped in the ordination space, indicating they were more similar, or less variable, from one another (Figure 4a,c). In addition, the bacterial and eukaryotic communities from the “pristine” reference site M0 had the most pronounced spatial separation from all three other sites (Figure 4b,d), in line with the relative difference in contamination across sites (Figure 2; Glozier et al. 2018). The variance in OTU abundance explained by the RDA ordination space was small (between 1 – 16% for all plots); however, the effect of both sample type and location on relative OTU abundance was significant (RDA $p < 0.001$). Thus, while all microbial niches across all sample locations supported roughly the same number or diversity of different OTUs (Figure 3), it is expected that different types of OTUs (i.e., taxa) would be selected for at different sampling locations given oil sands activity.

Community profile across sample types and locations

At the broad taxonomic level of phyla, the apparent impact of KSL sampling on community diversity indicated by GLM analysis was not immediately intuitive, as the same phyla persisted in similar abundances relative to other sample types (Figure 5). For the eukaryotic communities, the effect of KSL sampling was relatively less profound on taxa abundance, with CSL and SW communities appearing more similar; however, it is acknowledged that direct comparison of averaged abundance values should be made with caution (Figure 5). Proteobacteria, Bacteroidetes, Actinobacteria, Planctomycetes and Firmicutes dominated the bacterial communities collected along the Athabasca River, and Ocrophyta, Cercozoa, Ciliophora, and unclassified Eukaryota were predominant in the protozoan community.

Examining the composition of specific community types from distinct sample locations at a finer taxonomic resolution, the selective effect of location and sample type was more readily

observed (Figure 6). For reference site M0 and M4, which experiences possible contamination from OS and WWTP activity and from naturally eroded bitumen from the main stem and tributaries, all four sample-types were collected (BT, CSL, KSL, SW). Comparing sample types across these locations demonstrated clearly that community niches vary in their taxonomic profile at the level of genus. The relative abundance of bacterial and eukaryotic genera collected by KSL at both M0 and M4 were distinct from CSL and SW samples. Further, BT communities were distinct in terms of top abundance organisms from other sample types and between sites M0 and M4 (Figure 6). Several of the same bacterial and eukaryotic genera were found across all sample types and were present at both pristine and contaminated sites; however, the abundance of these taxa fluctuated across conditions. Importantly, the proportion of “all other” taxa (i.e., not highly abundant taxa), which can contain unique but low-frequency OTUs, fluctuated between samples, varied across sample type and location, and contributed to calculated differences in community (Figure 5 – 6).

Selection of genera by sampling site

Select examination of OTU relative abundance prompted the observation that abundant genera within the Athabasca River were similarly present across the different community niches and locations, thus tempting the conclusion that the communities were largely similar in terms of dominant microorganisms. However, several “biomarker” taxa were determined for each location through linear discriminant analysis effect size (LEfSe) (Figure 7). Results indicated that specific bacterial and eukaryotic genera and/or other taxa were differentially abundant at statistically relevant levels across sampling sites. Specifically, several genera were relatively more abundant at the pristine reference site M0, and at sites M4 and M7, which were directly downstream of WWTP effluents and OS mining activity and experience elevated levels of

hydrocarbons and metals (Figure 2; File S2; Glozier et al. 2018). Thus, while the bacterial and eukaryotic communities of the Athabasca River did not statistically differ in terms of diversity (the number and proportion of unique organisms) at different sampling sites along the gradient of contamination as expected, these differences in contamination exposure did select for the differential abundance of specific taxa.

Influence of Kenney Sampler and centrifugation on suspended load communities

Samples collected by CSL from M0 were further processed so that the community present in the water entering the centrifuge or Kenney sampler (“CSL-IN” or “KSL-IN”) and for the water exiting the centrifuge manifold (“CSL-OUT”) were collected, allowing for detailed assessment of the collection method effect on community composition. Ordination analysis of the communities present in the suspended load obtained by either centrifuge manifold (CSL) or the Kenney Sampler (KSL) clearly demonstrated that both methods influenced the abundance and membership of the “collected” communities (Figure 8). Specifically, we observed that both the bacterial and eukaryotic aquatic communities that “enter” either apparatus (CSL-IN and KSL-IN) were highly similar to one another in terms of OTU abundance, and distinct from the communities that were ultimately collected and deemed the “suspended load” communities (CSL and KSL). Importantly, the “final” suspended load community collected via CSL and KSL were distinct from one another for both community types, with limited (if any) overlap in the ordination space (Figure 8a,b), indicating that there was variance in the CSL and KSL communities explained by these collection methods ($p < 0.001$). Strikingly, the use of centrifugation influenced the abundance of specific bacterial and eukaryotic genera relative to the water that entered and was discharged from the manifold for both bacterial and eukaryotic communities (Figure 8c,d).

Discussion

Unique microbial niches within riverine systems, which include planktonic (surface water), biofilm, sediment, floc (river snow), and suspended load, differ in terms of the physical and chemical parameters of the water they may interact with, and the associations they may establish with other organisms and particulate materials (Paerl 1975). For instance, suspended sediment in rivers are traditionally described as consisting of both organic and inorganic particles in the range of 0.45-2000 μm (Meade et al. 1990; Owens et al. 2005). These values have been noted previously as being somewhat arbitrary cut-offs between “dissolved fractions”, the bed load moving at the base of the river, and suspended materials in the water column (Håkanson 2008; Owens et al. 2008). Nonetheless, it is well documented that there is a cohesive nature to suspended load materials in which the materials are flocculated or aggregated (Droppo 2001). These flocs and aggregates have both a viable and non-viable biological component, described in a conceptual model that includes five sub-components: (a) inorganic particles, (b) biota and bioorganic particles, (c) fibrils, (d) water and (e) pores (Droppo 2001; Leppard and Droppo 2005) and thus may be considered as “microecosystems”. Indeed, analysis of the taxa and diversity of free-living and particle-associated bacterial communities along the Mississippi River reported distinct dominant phyla signatures of these two different community types and found both communities to be susceptible to changes in environmental conditions (Payne et al. 2017).

A number of devices have been implemented to collect these suspended load materials, including point samplers, integrating samplers, pumps, and centrifugation. A commonly used device is that of Phillips et al., originally developed to trap sediment to assess physical properties of the suspended load materials (Phillips et al. 2000; Russel et al. 2000). Another point-

integrating sampler, which samples suspended sediment from one plane within the water column, is the Kenney sampler (Kenney 1985). These passive samplers both use apertures to entrap suspended particles and were developed to study the physical, chemical, geochemical and magnetic properties of transported material as well as flux rates and sources (Perks 2013). Importantly, the Phillips sampler is horizontally installed above the bottom sediments whereas the Kenney sampler is suspended vertically in the water column. Further, the Kenney collection method is a time-integrated means of sampling, whereby sediment is collected over a period of days, which increases the likelihood that organisms may proliferate in number within the apparatus. Centrifugation in contrast, involved more active sampling through use of a pump to draw water through the manifold, collecting sediment within hours. These methods are both in contrast to the collection of surface water (SW), which is performed by capturing a set volume at the water surface, followed by 0.22 μ M filtration, and is typically dominated by the bacterioplanktonic community. Horowitz (1991) indicated that from a physical-chemical sampling perspective, integrating samplers obtain more representative suspended load, and indeed taxonomic assessment of these samples does indicate KSL samples are distinct from both CSL and SW (Figure 3 – 4).

This finding agrees with a number of studies that have demonstrated significant differences in composition between particle-associated and free-living bacterial communities (Rieck et al. 2015; Bižić-Ionescu et al. 2014; Ortega-Retuerta et al. 2013). Indeed, recent metatranscriptional analysis of bed and suspended sediment communities within the Athabasca River basin revealed that suspended communities respond to elevated hydrocarbon contamination in a delayed manner (via expression of energy metabolism and hydrocarbon-degradation genes) relative to bed sediment communities, attributable to the erosion and

transport dynamics of particles within the lotic system (Reid et al. 2020). Taken together, these relevant findings confirming distinct community identity and function lend further credence to the finding that use of the Kenney sampler is an effective means of collecting a distinct suspended load community.

Within the Athabasca River, comparisons between sites M4 and M7 and reference site M0, suggest that the presence of both dissolved and particulate-phase selects for similarly diverse communities that are unique in the abundance of the most-prevalent genera (Figure 5 – 7). This is in agreement with previous analysis of riverine planktonic and particle-associated or suspended solid communities that note diversity remains relatively stable along the distance of the Mississippi river, excepting at distinct areas of urban contamination (Payne et al. 2017). At the level of phyla, the bacterial and eukaryotic communities identified in the present study are in agreement with previous assessments of riverine microbial communities, and of the Athabasca River in particular (Figure 5). Specifically, Proteobacteria and Bacteroidetes were the main phyla found in aquatic systems impacted by hydrocarbon contamination (Jurelevicius et al. 2013; Reid et al. 2020), and in addition to Firmicutes, Actinobacteria, and Chloroflexi, have been found to dominate in the Athabasca system (Foght et al. 2017; Reid et al. 2020). Planctomycetes sequences have been previously identified in hydrocarbon-impacted and anaerobic environments (Elshahed et al. 2007). For eukaryotic communities, organisms belonging to Ochrophyta can exist in a wide range of forms and are commonly found wherever free water is (Harper et al. 2012), and Cercozoa sequences were found to be the dominant group in surface sediments compared with planktonic ecosystems in the South China Sea (Wu and Huang 2019). Similarly, Eukaryota and Ciliophora were abundant in a recent study examining petroleum-contaminated water aquifers (Kachienga et al. 2018).

Of the several discriminatory genera identified to be dominant at particular sites, several of them were related to environmental hydrocarbon contamination (Figure 7). Specifically, *Albidiferax* (Brown et al. 2015), *Methylobacter* (Saidi-Mehrabad et al. 2013), *Nocardiodes* (Brown et al. 2017) at site M4 and *Thiobacillus* (Penner and Foght 2010) at site M7 have all been found to be associated with hydrocarbon contamination; *Thiobacillus* and *Methylobacter* have also been previously shown to be abundant within oil sand tailing ponds and settling basins of the Athabasca River (Saidi-Mehrabad et al. 2013; Penner and Foght 2010). Interestingly, several known hydrocarbon degrading genera were found to be discriminatory at the M0 reference site (Figure 7), specifically *Luteolebacter*, *Potolliolus* and *Methylothermus*, as they were in previous assessments of other “pristine” locations along the Athabasca River (Foght et al. 2017). These microbes were hypothesized to represent descendants of the original microbiota that colonized kerogen-rich sediments during burial and have persisted at the oil–water transition zone (Foght et al. 2017). Eukaryotic genera of *Stentor* and *Vorticella* have been previously found to be rare and common residents, respectively, of activated sludge from a WWTP in Italy (Madoni 2011) and were found to be discriminatorily abundant at site M4, which is immediately downstream of a WWTP. Interestingly, captured within the eukaryotic community data was the presence of *Juncus* DNA (discriminatory at site M7), which is a helophytic plant found to show promising phytoremediative action against organic pollutants, municipal and industrial heavy metal-contaminated wastewater (Gruber et al. 2008; Syranidou et al. 2016).

Despite the apparent agreement of the taxonomic profile and findings of this study with previous assessments of the Athabasca River, there are multiple, important sources of variation to consider for this (and any) microbial community data set. These include data normalization techniques, and the relative strength and age of the databases against which bacteria and

eukaryotic sequences are being identified. However, the present study also indicates the need to be cognizant of the types of communities (i.e., microbial niches) that one seeks to profile in aquatic systems, as there are distinct differences between the abundance and taxonomic profile, and in the instance of suspended load, significant differences in the diversity of bacterial and eukaryotic communities found in the Athabasca River.

The observation that the use of an integrative sampling apparatus had a direct impact on the bacterial suspended load community diversity, and not the influence of contamination across different sites is initially surprising given the literature that supports contamination as directly influencing microbial diversity (Payne et al. 2017; Quero et al. 2015; Zeglin 2015). This speaks, however, to a particular pitfall of microbial analysis studies that solely focus on diversity metrics to assess ecological interactions. As Shade (2016) has argued, within-sample diversity, which is often considered the cause of a community's success or failure, is truly more nuanced and is not the "answer" to community outcomes, but instead should be used as a starting point for further inquiry regarding ecological mechanisms. In order to gain a complete picture of microbial dynamics, the statistical differences in diversity, community similarity, and the relative abundances and profiles of taxa must be examined. In this instance, examination of differential abundances at the level of genus, in conjunction with diversity analysis, reveals patterns of influence of sampling method, microbial niche, and contamination.

While it has been frequently demonstrated and anticipated that anthropogenic contamination will influence environmental microbial communities, the use of aquatic sampling apparatuses, manifolds, or tools also appear to undoubtedly impact the nature of bacterial and eukaryotic suspended load communities. Important questions remain as to whether the functional diversity and capacity of microbial communities were similarly influenced or "selected" for by

different sampling techniques. The potential impact of sampling on ecological observations is an underappreciated reality and critical consideration for environmental sampling/monitoring programs, thus it is recommended that effort be made to collect and profile “reference” communities (i.e., such as CSL-IN and CSL-OUT) are important to help corroborate observations and conclusions.

Conclusions

Different bacterial and eukaryotic communities exist within different niches (biofilm, surface water and suspended load) of the Athabasca River at the level of genus. Importantly, this study did not find that communities of unique niches or locations differed significantly in terms of their diversity – i.e., the number and proportion of unique organisms they support. These observations demonstrate the need to examine the usefulness of biodiversity as a response metric for monitoring contamination or for specific studies. We found the collection of suspended load communities in particular, which have been historically difficult to define, to be directly influenced by the means of collection. The use of active vs passive sample collection through a centrifugation manifold and a Kenney Sampler, respectively, resulted in “suspended load” communities with different genera abundance profiles, and in the case of the Kenney Sampler, actually significantly increased the observed bacterial diversity relative to other methods, which may due to its time-integrative nature. Thus, in future studies that seek to assess important conclusions about the ecological or functional character of a riverine system, especially in a temporal manner, must necessarily consider using a non-time integrative means of sampling and/or consider including an alternative means of collection as a reference. Ultimately, this data provides strong support for the need to consider the nature of the targeted microbial niche under

study and to undertake analysis that does not solely focus on biodiversity when conducting riverine microbial ecology studies.

Acknowledgements

The authors would like to thank Environment and Climate Change Canada, the Strategic Applications of Genomics in the Environment Program (STAGE), and the Genomics Research and Development Initiative (GRDI) for their funding of the project and the Oilsands Monitoring program for logistical and field sampling support. We thank Compute Canada for access to the Graham system (University of Waterloo).

Data availability

The raw sequencing files for this study have been deposited into NCBI's Sequence Read Archive (SRA) under Project Accession Number PRJNA644613.

Appendices

File S1. Biofilm tile collection apparatus; rarefaction curve analysis of bacterial and eukaryotic samples; results of GLM and Tukey statistical tests.

File S2. Associated metadata of sequenced samples.

References

- Bižić-Ionescu, M., Zeder, M., Ionescu, D., Orlic, S., Fuchs, B.M., Grossart, H-P., et al. 2014. Comparison of bacterial communities on limnic versus coastal marine particles reveals profound differences in colonization. *Environ Microbiol* 17:3500–3514. doi: 10.1111/1462-2920.12466
- Brown, B.L., LePrell, R.V., Franklin, R.B., Rivera, M.C., Carbral, F.M., Eaves, H.L., et al. 2015. Metagenomic analysis of planktonic microbial consortia from a non-tidal urban-impacted segment of James River. *Stand Genomic Sci* 10:65. <https://doi.org/10.1186/s40793-015-0062-5>

- 453 Brown, L.M., Gunasekera, T.S., Ruiz, O.N. 2017. Draft genome Sequence of *Nocardioides*
454 *luteus* strain BAFB, an alkane-degrading bacterium isolated from JP-7-polluted oil. *Gen Ann*
455 5:e01529–16. <https://doi.org/10.1128/genomeA.01529-16>
- 456 Droppo, I.G. 2001. Rethinking what constitutes suspended sediment. *Hydrol Proc* 15:1551–
457 1564. <https://doi.org/10.1002/hyp.228>
- 458 Drury, B., Rosi-Marshall, E., Kelly, J.J. 2013. Wastewater treatment effluent reduces the
459 abundance and diversity of benthic bacterial communities in urban and suburban rivers. *Appl*
460 *Environ Microbiol* 79:1897–1905. doi: 10.1128/AEM.03527-12
- 461 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. 2011. UCHIME improves
462 sensitivity and speed of chimera detection. *Bioinformatics* 27:2194–2200. doi:
463 10.1093/bioinformatics/btr381
- 464 Elshahed, M.S., Youssef, N.H., Luo, Q., Najar, F.Z., Roe, B.A., Sisk, T.M., et al. 2007.
465 Phylogenetic and metabolic diversity of Planctomycetes from anaerobic, sulfide- and sulfur-rich
466 Zodletone Spring, Oklahoma. *Appl Environ Microbiol* 73 :4707–4716.
467 <https://doi.org/10.1128/AEM.00591-07>
- 468 Foght, J.M., Gieg, L.M., Siddique, T. 2017. The microbiology of oil sands tailings: past, present,
469 future, *FEMS Microbiol Ecol* 93:fix034, <https://doi.org/10.1093/femsec/fix034>
- 470 Fortunato, C.S., Herfort, L., Zuber, P., Baptista, A.M., Crump, B.C. 2012. Spatial variability
471 overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient.
472 *ISME J.* 6:554–563. doi: 10.1038/ismej.2011.135
- 473 Glozier, N.E., Pippy, K., Levesque, L., Ritcey, A., Armstrong, B., Tobin, O., et al. 2018. Surface
474 water quality of the Athabasca, Peace and Slave Rivers and riverine waterbodies within the
475 Peace-Athabasca Delta p 64. In *Oil Sands Monitoring Program Technical Report Series Part 1.4.*
476 Government of Canada.
- 477 Gruber, H., Wiessner, A., Kusch, P., Kaestner, M., Appenroth, K.J. 2008. Physiological
478 responses of *Juncus effusus* (Rush) to chromium and relevance for wastewater treatment in
479 constructed wetlands. *Int J Phytoremed* 10:77–88. doi:10.1080/15226510801913306
- 480 Harper, M.A., Cassie, V.C., Chang, F., Nelson, W., Broady, P. 2012. Chapter 10: Phylum
481 Ochrophyta: brown and golden-brown algae, diatoms, silicoflagellates, and kin, p 114– 163. In
482 Gordon DP (ed), *New Zealand Inventory of Biodiversity*, Vol. 3, Kingdoms Bacteria, Protozoa,
483 Chromista, Plantae, Fungi, 1st Edition. Canterbury University Press, Christchurch, NZ.
- 484 Horowitz, A. 1991. A Primer on Sediment-Trace Element Chemistry; 2nd ed, p 36. In U.S.
485 Geological Survey; Books and Open-File Reports Section. Open-File Report 91-76.
486 <https://doi.org/10.3133/ofr9176>
- 487 Hothorn, T., Bretz, F., Westfall, P. 2008. Simultaneous inference in general parametric models.
488 *Biometric J* 50:346–363.

- 489 Jurelevicius, D., Alvarez, V.M., Marques, J.M., de Sousa Lima, L.R.F., de Almeida Dias, F.,
490 Seldin, L. 2013. Bacterial community response to petroleum hydrocarbon amendments in
491 freshwater, marine, and hypersaline water-containing microcosms. *Appl Environ Microbiol*
492 79:5927–5935. doi: 10.1128/AEM.02251-13
- 493 Kachienga, L., Jitendra, K., Momba, M. 2018. Metagenomic profiling for assessing microbial
494 diversity and microbial adaptation to degradation of hydrocarbons in two South African
495 petroleum-contaminated water aquifers. *Sci Rep* 8:7564 [https://doi.org/10.1038/s41598-018-](https://doi.org/10.1038/s41598-018-25961-0)
496 25961-0
- 497 Kenney, B.C. 1985. Sediment resuspension and currents in Lake Manitoba. *J Great Lakes Res*
498 11:85–96.
- 499 Kenney, B.C., Mollison, J.C. 1980. Physical limnology instrumentation system for shallow lake
500 towers: Design and calibration. NWRI Report PR-80-3, National Water Research Institute,
501 Winnipeg, Canada.
- 502 Komori, O., Eguchi, S., Saigusa, Y., Kusumoto, B., Kubota, Y. 2020. Sampling bias correction
503 in species distribution models by quasi-linear Poisson point process. *Ecol Inform* 55:101015.
504 <https://doi.org/10.1016/j.ecoinf.2019.101015>
- 505 Leppard, G.G., Droppo, I.G. 2005. Overview of flocculation processes in freshwater ecosystems,
506 p 25–46. In Droppo IG, Leppard GG, Liss SN, Milligan TG. (ed), *Flocculation in natural and*
507 *engineered environmental systems*. CRC Press, Boca Raton, FL.
- 508 Madoni, P. 2011. Protozoa in wastewater treatment processes: A mini review. *Italian J. Zoo*
509 78:3–11, doi: 10.1080/11250000903373797
- 510 Magoc, T., Salzberg, S.L. 2011. FLASH: fast length adjustment of short reads to improve
511 genome assemblies. *Bioinformatics* 27:2957–2963. doi: 10.1093/bioinformatics/btr507
- 512 Meade, R.H., Yuzyk, T.R., Day, T.J. 1990. Chapter 11: Movement and storage of sediment in
513 the rivers of the United States and Canada, p 255–280. In Wolman MG, Riggs HC (ed), *Surface*
514 *Water Hydrology: The Geology of North America*. Geological Society of America, Boulder,
515 Colorado. <https://doi.org/10.1130/DNAG-GNA-O1.255>
- 516 Miller, R.G. 1981. *Simultaneous Statistical Inference*. Springer, New York, NY.
- 517 Oil Sands Information Portal, Government of Alberta. <http://osip.alberta.ca/map/>. Accessed June
518 1, 2020.
- 519 Ortega-Retuerta, E., Joux, F., Jeffrey, W.H., Ghiglione, J.F. 2013. Spatial variability of particle-
520 attached and free-living bacterial diversity in surface waters from the Mackenzie River to the
521 Beaufort Sea (Canadian Arctic). *Biogeosciences* 10:2747–2759. doi: 10.5194/bg-10-2747-2013
- 522 Owens, P.N. 2008. Sediment behaviour, functions and management in river basins. *Sust. Manag*
523 *Sed Res* 4:1–29. [https://doi.org/10.1016/S1872-1990\(08\)80003-7](https://doi.org/10.1016/S1872-1990(08)80003-7)

- Owens, P.N., Batalla, R.J., Collins, A.J., Gomez, B., Hicks, D.M., Horowitz, A.J., et al. 2005. Fine-grained sediment in river systems: environmental significance and management issues. *River Res Appl* 21:693–717. doi: 10.1002/rra.878
- Paerl, H.W. 1975. Microbial attachment to particles in marine and freshwater ecosystems. *Microb Ecol* 2:73–83.
- Parada, A.E., Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* 18 :1403–1414.
- Payne, J.T., Millar, J.J., Jackson, C.R., Ochs, C.A. 2017. Patterns of variation in diversity of the Mississippi River microbiome over 1,300 kilometers. *PloS One*. <https://doi.org/10.1371/journal.pone.0174890>
- Penner, T.J., Foght, J.M. 2010. Mature fine tailings from oil sands processing harbour diverse methanogenic communities. *Can J Microbiol* 56:459–470. <https://doi.org/10.1139/W10-029>
- Perks, M.T. 2013. Chapter 3: Suspended sediment sampling. In Cook SJ, Clarke LE, Nield JM (eds), *Geomorphological Techniques* (online edition). British Society for Geomorphology, London, UK.
- Phillips, J.M., Russell, M.A., Walling, D.E. 2000. Time-integrated sampling of fluvial suspended sediment: a simple methodology for small catchments. *Hydrol Proc* 14:2589–2602. DOI: 10.1002/1099-1085(20001015)14:14<2589::AID-HYP94>3.0.CO;2-D
- Portillo, M.C., Anderson, S.P., Fierer, N. 2012. Temporal variability in the diversity and composition of stream bacterioplankton communities. *Environ Microbiol* 14:2417–2428. doi: 10.1111/j.1462-2920.2012.02785.x
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl Acids Res* 41:D590–596. doi: 10.1093/nar/gks1219
- Quero, G.M., Cassin, D., Botter, M., Perini, L., Luna, G.M. 2015. Patterns of benthic bacterial diversity in coastal areas contaminated by heavy metals, polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs). *Front Microbiol* 6:1053. DOI.10.3389/fmicb.2015.01053.
- R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reid, T., Droppo, I.G., Weisener, C.G. 2020. Tracking functional bacterial biomarkers in response to a gradient of contaminant exposure within a river continuum. *Water Res* 168:115167. <https://doi.org/10.1016/j.watres.2019.115167>
- Reid, T., Chaganti, S.R., Droppo, I.G., Weisener, C.G. 2018. Novel insights into freshwater hydrocarbon-rich sediments using metatranscriptomics: opening the black box. *Water Res* 136:1e11. <https://doi.org/10.1016/j.watres.2018.02.039>.

- Reid, T., Droppo, I.G., Chaganti, S.R., Weisener, C.G. 2019. Microbial metabolic strategies for overcoming low-oxygen in naturalized freshwater reservoirs surrounding the Athabasca Oil Sands: a proxy for end-pit lakes? *Sci Total Environ* 665:113e124. <https://doi.org/10.1016/j.scitotenv.2019.02.032>.
- Rieck, A., Herlemann, D.P.R., Jürgens, K., Grossart, H.P. 2015. Particle-associated differ from free-living bacteria in surface waters of the Baltic Sea. *Front Microbiol* 6 :1297. doi: 10.3389/fmicb.2015.01297
- Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahe, F. 2016. VSEARCH: a versatile open source tool for metagenomics. *Peer J* 4:e2584. doi: 10.7717/peerj.2584
- Russell, M.A., Walling, D.E., Hodgkinson, R.A. 2000. Appraisal of a simple sampling device for collecting time-integrated fluvial suspended sediment samples. *Intern Assoc Hydrol Sci* 236:119–127.
- Saidi-Mehrabad, A., He, Z., Tamas, I., Sharp, C.E., Brady, A.L., Rochman, F.F., et al. 2013. Methanotrophic bacteria in oilsands tailings ponds of northern Alberta. *ISME J* 7:908–921. <https://doi.org/10.1038/ismej.2012.163>
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., et al. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. doi: 10.1186/gb-2011-12-6-r60
- Staley, C., Gould, T.J., Wang, P., Phillips, J., Cotner, J.B., Sadowsky, M.J. 2014. Core functional traits of bacterial communities in the Upper Mississippi River show limited variation in response to land cover. *Front Microbiol* 5:414. doi: 10.3389/fmicb.2014.00414
- Staley, C., Unno, T., Gould, T.J., Jarvis, B., Phillips, J., Cotner, J.B., et al. 2013. Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *J Appl Microbiol* 115:1147–1158. doi: 10.1111/jam.12323
- Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D., Brenner, H.W., et al. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* 19:21–31. doi: 10.1111/j.1365-294X.2009.04480.x
- Syranidou, E., Christofilopoulos, S., Kalogerakis, N. 2016. *Juncus* spp.—The helophyte for all (phyto)remediation purposes? *New Biotechnol* 38. 10.1016/j.nbt.2016.12.005
- Tremblay, J., Yergeau, E. 2019. Systematic processing of ribosomal RNA gene amplicon sequencing data. *GigaScience*, 8:giz146. <https://doi.org/10.1093/gigascience/giz146>
- Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R. 2007. Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–7
- Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York, NY. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>

- 597 Wu, W., Huang, B. 2019. Protist diversity and community assembly in surface sediments of the
598 South China Sea. *Microbiol Open* 8:e891. <https://doi.org/10.1002/mbo3.891>
- 599 Yandell, B.S. 1997. *Practical Data Analysis for Designed Experiments*: 1st Ed. CRC Press, Boca
600 Raton, FL.
- 601 Yergeau, E., Lawrence, J.R., Sanschagrin, S., Waiser, M.J., Korber, D.R., Greer, C.W. 2012.
602 Next-generation sequencing of microbial communities in the Athabasca river and its tributaries
603 in relation to oil sands mining activities. *Appl Environ Microbiol* 78:e7626–7637.
604 <https://doi.org/10.1128/AEM.02036-12>.
- 605 Zakrzewski, M., Proietti, C., Ellis, J., Hasan, S., Brion, M.J., Berger, B., et al. 2016. Calypso: A
606 user-friendly web-server for mining and visualizing microbiome-environment interactions.
607 *Bioinformatics* 33:782–783. doi: 10.1093/bioinformatics/btw725
- 608 Zeglin, L.H. 2015. Stream microbial diversity in response to environmental changes: review and
609 synthesis of existing research. *Front Microbiol* 6:454. DOI.10.3389
- 610 Zwart, G., Crump, B.C., Agterveld, M.P.K.V., Hagen, F., Han., S.K. 2002. Typical freshwater
611 bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers.
612 *Aquat Microb Ecol* 28:141–155. doi: 10.3354/ame028141

613

614 **Figures**

615 **Figure 1.** Schematic of sampling locations along the Athabasca River.

616 **Figure 2.** Select water quality parameters for locations.

617 **Figure 3.** Alpha diversity of sampled communities as measured by Inverse Simpson's metric for
618 a) prokaryotic (16S) and b) eukaryotic (18S) communities.

619 **Figure 4.** RDA plot of the top 500 abundant OTUs across (a, b) prokaryotic (16S) and (c, d)
620 eukaryotic (18S) samples.

621 **Figure 5.** Relative abundance of top 10 most abundant a) prokaryotic and b) eukaryotic phyla
622 across communities collected by different methods.

623 **Figure 6.** Relative abundance of most prevalent a,b) prokaryotic (16S) and c,d) eukaryotic (18S)
624 genera across communities collected from two locations.

625 **Figure 7.** Linear discriminant analysis effect size (LEfSe) biomarker analysis for a) prokaryotic
626 (16S) and b) eukaryotic (18S) genera across sample locations.

627 **Figure 8.** Effect on OTU abundance of various stages of collecting suspended load.

628

Draft

Figure 1. Schematic of sampling locations along the Athabasca River. Samples were collected by various methods from sites M0, M4, M7, and M9. Note that M0 (at end of dashed line) is located farther south, in the mainstem of the Athabasca River at the town of Athabasca (see inset), serves as a reference site outside the bituminous McMurray Formation and the oil sands mining area.

Figure 2. Select water quality parameters for locations. Winter = January, February, March; Summer = June, July, Aug. Mean and standard deviation are plotted; $n=3$ for each seasonal period. Some parameter data for sites M0 and M7 is unavailable for described period, though comprehensive data for Athabasca River is available through Oil Sands Information portal (<http://osip.alberta.ca/map/>).

Figure 3. Alpha diversity of sampled communities as measured by Inverse Simpson’s metric for a) prokaryotic (16S) and b) eukaryotic (18S) communities. BT = biofilm tile; CSL = centrifuged-collected suspended load; KSL = Kenney Sampler suspended load; SW = surface water. Mean and 95% CI plotted for all sample types.

Figure 4. RDA plot of the top 500 abundant OTUs across (a, b) prokaryotic (16S) and (c, d) eukaryotic (18S) samples. BT – Biofilm Tile; CSL – Centrifuged Suspended Load; KSL – Kenney Sampler Suspended Load; SW – Surface Water. Dotted line ellipses indicate a 95% confidence interval. For a,c,d the RDA significance $p < 0.001$; for b $p < 0.005$.

Figure 5. Relative abundance of top 10 most abundant a) prokaryotic and b) eukaryotic phyla across communities collected by different methods. Samples collected from all sites and over both sampling years by each method are averaged. Legend provided describes the top 10 most abundant phyla across all samples (includes all OTUs identified to a particular phyla), and then "all other" phyla detected. BT = biofilm tile (n=109); CSL = centrifuged-collected suspended load (n=74); KSL = Kenney Sampler suspended load (n=97); SW = surface water (n=162).

Figure 6. Relative abundance of most prevalent a,b) prokaryotic (16S) and c,d) eukaryotic (18S) genera across communities collected from two locations. M0 = reference site; M4 = OS activity and related contamination. For eukaryota, the finest resolution of taxa possible is provided (i.e., not all OTUs are able to be identified to the level of genus). Legend provided describes the most abundant genera across all samples, and "all other" genera detected. BT = biofilm tile (M0 n=23; M4 n=28); CSL = centrifuge-collected suspended load (M0 n=21; M4 n=26); KSL = Kenney Sampler suspended load (M0 n=12; M4 n=28); SW = surface water (M0 n=37; M4 n=42).

Figure 7. Linear discriminant analysis effect size (LEfSe) biomarker analysis for a) prokaryotic (16S) and b) eukaryotic (18S) genera across sample locations. A logarithmic LDA (linear discriminant analysis) score greater than 2 is considered significant. The lowest possible taxonomic resolution for discriminant taxa within the eukaryotic community are displayed. M0 n=114; M4 n=112; M7 n=95; M9 n=120.

Figure 8. Effect on OTU abundance of various stages of collecting suspended load. RDA analysis of the top 500 abundant OTUs for a) prokaryotic (16S) and b) eukaryotic (18S) communities that are present in various stages of collecting suspended load. RDA is significant at $p < 0.001$. Relative abundance of top genera for c) prokaryotic and d) eukaryotic communities at site M0. CSL-IN = water entering centrifuge manifold ($n=20$); CSL = centrifuge-collected suspended load ($n=39$); CSL-OUT = water flowing through centrifuge manifold ($n=15$); KSL-IN = water entering Kenney Sampler ($n=42$); KSL = Kenney Sampler-collected suspended load ($n=67$). RDA for a,b significance $p < 0.001$.