Incorporating microbes into environmental monitoring and mine closure programs: river diversions as test beds

Melanie L. Blanchette[1], Mark Lund[1], Michael Moore[2], Digby Short[3]

[1]*Mine Water and Environment Research Centre (MiWER), School of Science, Edith Cowan University, Perth, Western Australia, Australia, m.blanchette@ecu.edu.au*

[2]*Yancoal Australia Ltd., Sydney, New South Wales, Australia 3Ashton Coal (Yancoal Australia), Singleton, New South Wales, Australia*

Abstract

Microbes are rarely part of river assessment and mine closure programs because their ecology is complex and methods are not standardised. The aims of this paper are to 1) develop standardised collection and analysis protocols for riverine microbes, 2) determine if microbial communities correlate to environmental condition in mine-affected rivers, and 3) apply the 'system variability' closure approach to microbes. In two Australian rivers, the user-friendly collection protocol was tested on benthic and pelagic assemblages and pebble biofilms. Benthic and pelagic microbial assemblages varied spatially and temporally which correlated with environmental variables. However, more research is needed before microbes are incorporated into monitoring and closure.

Keywords: freshwater microbe protocol, ecological indicator, system variability, mining rehabilitation, Hunter Valley

Introduction

There has yet to be a systematic effort to incorporate microbes into river assessment programs because their broad-scale spatial and temporal variability is not well-researched, and the identification technology and collection methodology is still not routine. Microbes represent the majority of Earth's biodiversity, play key roles in aquatic biogeochemical processes, and are extremely sensitive to even small inputs of contaminants (Sims et al. 2013). Advances in the field of environmental genomics facilitated rapid identification of microbes by sequencing DNA directly from field samples, free from the selective effects of culturing that hampered previous attempts to understand microbial communities (Whiteley et al. 2012).

Aquatic macroinvertebrates (for example) are well-known ecological indicators for mine-affected rivers, but are not ideal, because identifying macroinvertebrates by their morphological characteristics is time-consuming. Although advances in DNA sequencing have shown great potential for identifying macroinvertebrates more rapidly (Carew et al. 2013), macroinvertebrates may not be ideal indicators. Unlike microbial processes, macroinvertebrates (and other multi-cellular taxa) may exhibit a substantial lag between exposure and effect detection due to lifecycles and/or sampling framework. More critically, macroinvertebrates cannot directly *identify* the effect – unlike the microbes (Druschel et al. 2004). For example, mining-related discharges have created new metal-rich niches that select for metal-tolerant and sulfate-reducing microbes and genes (Kang et al. 2013). Different successional stages of microbial communities may signal when a past effect occurred – knowledge of that would be useful when examining known affected sites. Essentially, we now have potential ecological indicators that can inform faster, cheaper, and more sensitive monitoring protocols but there are no standard methods.

Previously, we proposed a model (the 'system variability' approach) for setting rehabilitation goals and evaluating the condition of highly modified environments without reference sites and pre-impact data (Blanchette et al. 2016). This model used multivariate statistics to compare rehabilitated sites to overall local ecosystem variability, rather than reference site

characteristics. We tested the model using biophysical data from river diversions in the Hunter Valley, New South Wales, Australia, and found that data from diverted and non-diverted sections of river were substantially different, potential management actions could be identified, and site trajectories were mapped – in some instances towards closure (Blanchette and Lund 2017). This model gave us a more holistic view of the ecology of the river system, and lends itself to studying microbial assemblages and biogeochemical processes.

Therefore, the aims of this research were to 1) develop standardised collection and analysis protocols for riverine microbes, 2) determine if microbial communities correlate to environmental condition in mine-affected rivers, and 3) apply the 'system variability' closure approach to microbes (Archaea and Bacteria) in two Australian rivers. A user-friendly standardised collection protocol (Aim 1) is important because it increases data reliability across space and time. Determining if microbial communities statistically correlate to environmental condition (Aim 2) is critical in assessing whether they can be reliable ecological indicators (*sensu* Kurtz et al. 2001)

Methods

Overview – Hydrology and catchment (Hunter Valley, NSW, Australia, Bowman's Creek; BC, and the Goulburn River; GR) were as per Blanchette and Lund (2017). Briefly, these naturally intermittent rivers were diverted to facilitate coal mining. Complex hydrological conditions in both diversion and downstream sites resulted in treefalls and excessive growth of aquatic reeds, benthic siltation and further alteration of in-stream habitats (Blanchette and Lund 2017).

System Variability approach to developing closure criteria – APPENDIX I.

Biophysical data – River sites (Bowman's Creek; 12 sites, Goulburn River; 20 sites) were sampled during an annual hydroperiod (2016-2017, GR; four times, BC; twice). GR had a trapezoidal channel, whereas the BC diversion had been constructed with more 'natural' attributes. Within each river, sites were clustered into a priori groups ('zones') based on gross biophysical similarity (see Blanchette and Lund (2017)). *In situ* water quality (turbidity, pH, conductivity, temperature, ORP) and soil measurements (pH, ORP), flow, water collected for metal and nutrient analysis, percent cover of in-stream habitat and riparian condition was collected as described in Blanchette and Lund (2017). Sediment size was determined according to a phi scale (McMullen et al. 2011).

For each site, benthic sediments (7 to -3 phi) were analysed for bioavailable metals and nutrients (Na, Mg, K, Fe, Al, S, Ca, P, Mn, Cr, Co, Ni, Cu, Zn, Cd, Pb) at the Edith Cowan University Analytical Chemical Laboratory and CSBP Perth. At ECU, sediments were dried (40°C), ground by hand in an agate mortar and pestle, passed through a 500 µm sieve and digested according to US EPA protocol 3050B for elemental analysis (Thermo Fisher Scientific iCAP QICPMS, 7600 ICPOES). Total N (Rayment and Lyons Method 7A5) and total organic C (acidification method 6B1) (Rayment and Lyons 2011) were analysed at CSBP Perth.

Microbe collection protocols – Benthic sediments (7 to -3 phi), pebbles (-5 phi) and water were collected from sites and times as above for DNA sequencing and OTU identification (Goulburn River; n_{water} = 77, $n_{sediment}$ = 78, $n_{pebbles}$ = 16, Bowman's Creek n_{water} = 24, $n_{sediment}$ = 21, $n_{pebbles}$ = 18).

Benthic sediments (for microbes and nutrient analyses) were collected at 25 equidistant points along the 50 m transect (site) using a modified sterile syringe to ensure only the top 1 cm was retained for analysis. In-stream differences (e.g., shading, hydrological habitats including dry sections, aquatic plants) were proportionally represented among the 25 samples, which were manually homogenised in a plastic bag such that each bag represented conditions at one site. Samples were placed on ice, then at -20°C for transportation and -80°C for storage.

Where benthic sediment surface area was dominated (or co-represented) by pebbles, 25 individual stones were collected roughly equidistantly along the 50 m transect. Pebbles from the same site were homogenised, transported and stored as per sediments.

When present, a water sample was collected from each site at approximately 5

cm beneath the surface with a sterile bottle without disturbing the benthos. Water was filtered on site (Durapore membrane filters, 0.22 µm, Merck Millipore) under gentle vacuum pressure until filter clogged. On site, the filter was transferred to a sterile petri dish, finely sliced using an autoclaved blade, and placed into a sterile Eppendorf tube for storage and transport as per sediments and pebbles.

Laboratory processing and OTU identification – DNA extraction, PCR, sequencing and OTU identification was performed by the Australian Genome Research Facility (AGRF). Unlike benthic and pelagic samples, biofilm on pebbles required liberation prior to DNA extraction. All pebbles were scrubbed with an autoclaved toothbrush, rinsed with autoclaved distilled water and elutriate filtered as per pelagic samples. It was impossible to treat samples from different habitats exactly the same prior to DNA extraction.

DNA extraction was performed by the AGRF using a DNeasy PowerLyzer PowerSoil Kit (Qiagen). Briefly, PCR amplicons were generated using specific primers (Target; 341F-806R, Forward(341F): CCTAYGGGRBGCASCAG, Reverse(806R): GGACTACNNGGGTATCTAAT) and conditions (Target region; 16S: V3-V4, Cycle; 29, Initial; 95°C for 7 min., Dissociate; 94°C for 30 s, Anneal; 50°C for 60 s, Extension; 72°C for 60 s, Finish; 72°C for 7 min.), with AmpliTaq Gold 360 mastermix (Life Technologies, Australia) for the primary PCR. A secondary PCR to index the amplicons was performed with TaKaRa Taq DNA Polymerase (Clontech). The resulting amplicons were measured by fluorometry (Invitrogen Picogreen) and normalised. The eqimolar pool was then measured by qPCR (KAPA) followed by sequencing on the Illumina MiSeq (San Diego, CA, USA) with 2 x 300 base pairs paired-end chemistry.

Paired-ends reads were assembled by aligning the forward and reverse reads using PEAR (version 0.9.5) (Zhang et al. 2014). Primers were identified and trimmed. Trimmed sequences were processed using Quantitative Insights into Microbial Ecology (QIIME 1.8) (Caporaso et al. 2010), USEARCH (version 7.1.1090) (Edgar et al. 2011) and UPARSE (Edgar 2013). Using USEARCH, sequences were quality-filtered, full-length duplicate sequences were removed and sorted by abundance. Singletons or unique reads in the data set were discarded. Sequences were clustered followed by chimera filtering using the "rdp_gold" database as the reference (Cole et al. 2014). To obtain the number of reads in each OTU, reads were mapped back to OTUs with a minimum identity of 97%. Using QIIME, taxonomy was assigned to OTUs using Greengenes database (version 13_8, Aug 2013) (DeSantis et al. 2006). Samples were rarefied at 17,000 reads for calculating dissimilarity between samples using weighed & unweighed Unifrac metrics.

Note that the benthic and pelagic samples were completed as part of one independent sequencing run, and the pebble biofilm as a second sequencing run. Therefore, in this paper, data were analysed separately. Future tests will re-run the similarity matrices of all three habitats simultaneously to ascertain significant differences (if any).

Data analysis – Non-metric multidimensional ordination (NMDS) in PRIMER visually displayed the microbial assemblage data (see Blanchette and Lund (2017) for biophysical data).

In this research, hypothesis tests on OTU weighted (abundance and presence/absence) Unifrac genetic distance data were conducted with a PERMANOVA (9999 permutations with a Euclidean distance; zone, time and habitat (pelagic, benthic, pebble) were fixed factors with site as a random factor nested in zone and time) in PRIMER (Anderson et al. 2008)). Where factorial interactions were significant ($p < 0.05$), results were interpreted with caution and within the context of subsequent pairwise analyses (see Anderson et al. 2008). Pseudo-F is reported hereafter as *F*.

Biophysical data was normalised and a resemblance matrix among zones created using Euclidean distance. Microbial assemblage data (relative abundance at the OTU level) similarity among zones was created using Bray-Curtis dissimilarity. Correlations between microbial assemblage data and biophysical parameters were analysed using RELATE (Spearman's rank; rs, $p < 0.05$; 9999 permutations) and BEST (BVSTEP) with a

Results and Discussion

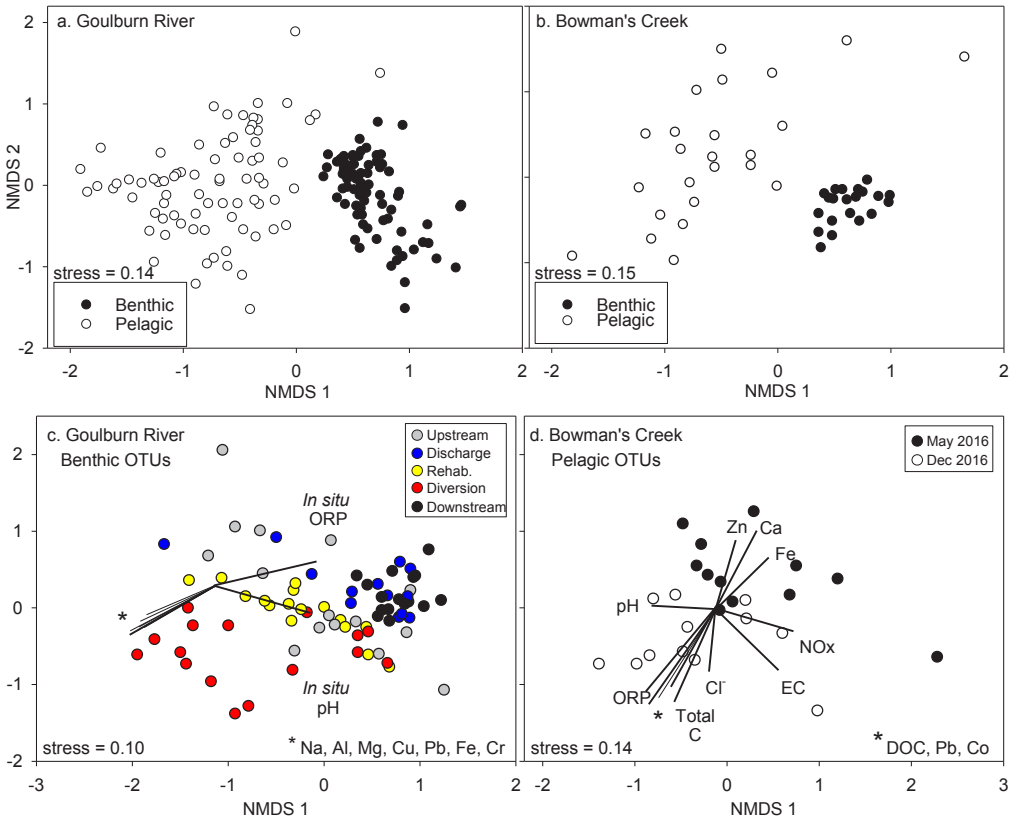

Figure 1 *NMDS ordinations of weighted Unifrac data (riverine OTUs) from the Hunter Valley, New South Wales, Australia. Figs. 1a and 1c - Gouburn River, and Figs. 1b and 1d - Bowman's Creek. Vectors (Figs. 1c and 1d) generated using Pearson's correlation with biophysical variables as per (Blanchette and Lund 2017). * indicates spatially close vectors.*

Pearson's correlation in PRIMER. Biophysical data was log-transformed and a Draftsman's plot allowed detection and subsequent removal of parameters for highly auto-correlated (> 0.95) variables. Characteristics of dominant microbes (relative abundance) were described at the Family level.

Samples from different riverine habitats (benthic, pelagic, pebble biofilm) were collected (Aim 1), genetic material was well-amplified, and OTUs were identified using established databases (Aim 1). Soil and pelagic assemblages were significantly correlated with environmental variables (Aim 2). Examples of how the system variability approach could be applied to the data were explored (Aim 3), which could inform future environmental monitoring and closure programs.

In both the Goulburn River (GR) and Bowman's Creek (BC), there were significant effects of zone (GR; $F_4 = 7.10$, $p < 0.01$, BC; $F_3 = 2.69$, $p < 0.01$), time (GR; $F_3 = 9.78$, $p < 0.01$, BC; $F_1 = 4.62$, $p < 0.01$), and habitat (GR; $F_1 = 102.1$, $p < 0.01$, BC; $F_1 = 25.2$, $p < 0.01$) on microbial assemblages (weighted Unifrac distance OTU) at sites. NMDS showed a strong separation between pelagic and benthic OTUs in both rivers (Figures 1a and 1b). Essentially, microbial assemblages were different enough across space and time to be 'identified' with different zones over the hydroperiod.

A pairwise analysis for zone in the Goulburn River indicated that across all times, most comparisons of benthic microbial assemblages were significantly different ($p < 0.03$) with the exception of those found

upstream and in the 'rehab,' and upstream and in the discharge sites. The relationship between Goulburn River sediment OTUs and sediment chemistry variables was significant (r_s = 0.372, p < 0.01), with separation along NMDS axis 1 correlated to *in situ* ORP (r = 0.71) and *in situ* pH (r = 0.60) to the discharge and downstream sites, and Pb, Cr, Al, Na, Fe, Mg, and Cu (r ≅ - 0.5) to the diversion sites. Across all sites and times, the taxon in Goulburn River sediment samples with the highest relative abundance at the Family level (total n Families was 827) were the Hyphomicrobiaceae (Alphaproteobacteria), of which many Genera are primarily involved in C and N cycling (Wang et al. 2016). Future work would explore the potential presence of indicator taxa in zones, and how they correlate with changing environmental conditions.

In Bowman's Creek, pelagic microbe assemblage composition was different between times (May and December 2016; F_1 = 5.56, p < 0.01) and across space (zone; F_3 = 2.21, p < 0.01). Two diversions were sampled in Bowman's Creek. Pelagic microbial assemblages in the North Diversion were significantly different to those upstream, whereas assemblages in the West Diversion were not significantly different to those upstream or downstream. Pelagic microbial assemblages were significantly correlated to water quality variables (r_s = 0.435, p < 0.01) (Figure 1d); higher concentrations of Ca were associated with May 2016 (r = 0.69) and higher total C (r = -0.68), Pb (r = - 0.67) and DOC (r = - 0.64) were associated with December 2016.

Across all sites and times, the OTUs in Bowman's Creek pelagic samples with the highest relative abundance (*n* Families = 666) were the Synechococcaceae (Phylum Cyanobacteria), which can cause nuisance freshwater blooms with nutrient input and salinity variations (Muir and Perissinotto 2011). The domination of cyanobacteria and correlation of assemblages with Ca, carbon and other variables associated with evapoconcentration, decomposition and catchment run-off suggest that seasonality should play a key role in monitoring and closure planning.

In both the Goulburn River and Bowman's Creek, there was no effect of zone (GR; F_1 = 0.25, p = 0.82, BC; F_3 = 0.99, p = 0.41) or time (GR; F_3 = 0.65, p = 0.65, BC; F_1 = 0.91, p = 0.38) on microbe assemblages from pebble biofilms. Biofilm assemblages were dominated by *Exiguobacterium* sp. (Order Bacillales) of which there are currently two distinct genetic groups: the psychrophiles, (from cold environments such as permafrost) and the thermophiles (from warm environments such as hot springs) (Vishnivetskaya et al. 2009). In contrast with benthic and pelagic OTUs, pebble biofilm assemblages were stable over space and time despite changing environmental variables. Therefore, biofilms may not be ideal ecological indicators in these systems.

Conclusions

Ecological indicators are only useful if there are clear and defensible links between the indicator and ecological function (Kurtz et al. 2001). Microbes are challenging to use as ecological indicators because (for example) they have complex lifecycles, 'species' is a challenging definition, modes of reproduction can be flexible, and community interactions are complex. Here we did not test for gene presence or expression and microbial ecological function could only be speculated. Future work could incorporate metabolomics and proteomics, which would involve laboratory cultures, potentially undercutting the ecological 'gains' made with the culture-free techniques of this study. Cell counts may negate the effects of within-sample relative abundances, although staining techniques (e.g.) can also introduce bias, especially when working with multiple unidentified taxa.

There are management concerns with freshwater benthic ecosystems in mined catchments (*sensu* Kurtz et al. 2001). We explored the use of microbes as ecological indicators, and how they may be used in river monitoring and assessment. We collected and identified OTUs from different habitats in dry and wet river beds (Aim 1); important for seasonal river systems which form a large proportion of global river networks (Datry et al. 2018) yet are often overlooked in monitoring programs (Steward et al. 2018). In the case of soil and pelagic microbes, assemblages correlated to environmental variables (Aim 2). We also provided examples

of how the system variability approach could be applied to the data (Aim 3). By virtue of their fast lifecycles, sensitivity to changing conditions, and ease of collection and identification, benthic and pelagic microbe assemblages hold promise as ecological indicators. However, more research is needed before microbes are incorporated into monitoring and closure programs.

APPENDIX I – System variability approach to developing closure criteria.

The system variability approach to developing closure criteria employs multivariate ordination - a data reduction and clustering technique traditionally used in ecology - as an applied management tool. A successfully rehabilitated site would lie within the river's 'normal' variability, which is sustained and tracked over time (c.f., 'reference sites' in Blanchette et al. 2016). Determining if a rehabilitated site is within the variability of the system is a testable hypothesis; a null hypothesis of no significant difference between sites/assemblages is $p > 0.05$ using permutational MANOVA (Blanchette et al. 2016). Previously, we applied our model to biophysical data in the Hunter Valley rivers (Blanchette and Lund 2017).

Acknowledgements

Govt. of Western Australia Dept. of Ag. and Food (s165 Biosecurity and Agriculture Mgmt Act 2007, ECU QF 0754 QF-LA). Australian Coal Association Research Program, Ulan Coal Mines (Glencore) and Ashton Coal (Yancoal Australia). John Watson, Robyn Stoney, staff at both mines, Australian Genomic Research Facility.

References

Anderson M, Gorley RN, Clarke RK (2008) Permanova+ for Primer: Guide to Software and Statistical Methods.

Blanchette ML, Lund MA (2017) Biophysical closure criteria without reference sites: Evaluating river diversions around mines. – In: Wolkersdorfer, C.; Sartz, L.; Sillanpää, M. & Häkkinen, A.: Mine Water & Circular Economy (Vol I). – p. 437 – 444; Lappeenranta, Finland (Lappeenranta University of Technology).

Blanchette ML, Lund MA, Stoney R, Short D, Harkin C Bio-physical closure criteria without reference sites: realistic targets in modified rivers. – In: Drebenstedt, C. & Paul, M.: IMWA 2016 – Mining Meets Water – Conflicts and Solutions. – p. 586 – 592; Freiberg/Germany (TU Bergakademie Freiberg).

Caporaso JG et al. (2010) QIIME allows analysis of high-throughput community sequencing data Nature methods 7:335-336

Carew ME, Pettigrove VJ, Metzeling L, Hoffmann AA (2013) Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator species Frontiers in zoology 10:1-15

Cole JR et al. (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis Nucleic acids research 42:D633-D642

Datry T et al. (2018) Author Correction: A global analysis of terrestrial plant litter dynamics in non-perennial waterways Nature Geoscience 11:542

DeSantis TZ et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB Applied and environmental microbiology 72:5069-5072

Druschel GK, Baker BJ, Gihring TM, Banfield JF (2004) Acid mine drainage biogeochemistry at Iron Mountain, California Geochemical Transactions 5:13-32

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial ... Nature methods 10:996-998

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection Bioinformatics 27:2194-2200

Kang S, Van Nostrand JD, Gough HL, He Z, Hazen TC, Stahl DA, Zhou J (2013) Functional gene array–based analysis of microbial communities in heavy metals-contaminated lake sediments FEMS Microbiology Ecology 86:200-214 doi:10.1111/1574-6941.12152

Kurtz JC, Jackson LE, Fisher WS (2001) Strategies for evaluating indicators based on guidelines from the EPA's Office of Research and Development Ecological indicators 1:49-60

McMullen K, Poppe L, Ackerman S, Blackwood D, Schaer J, Nadeau M, Wood D (2011) Surficial geology of the sea floor in Central Rhode Island Sound Southeast of Point Judith, Rhode Island. USGS

Muir DG, Perissinotto R (2011) Persistent phytoplankton bloom in Lake St. Lucia (iSimangaliso Wetland Park, South Africa) caused by a cyanobacterium closely associated with the genus Cyanothece (Synechococcaceae, Chroococcales) Applied and environmental microbiology 77:5888-5896

Rayment GE, Lyons DJ (2011) Soil chemical methods: Australasia vol 3. CSIRO publishing,

Sims A, Zhang Y, Gajaraj S, Brown PB, Hu Z (2013) Toward the development of microbial indicators for wetland assessment Water Research 47:1711-1725

Steward AL, Negus P, Marshall JC, Clifford SE, Dent C (2018) Assessing the ecological health of rivers when they are dry Ecological Indicators 85:537-547

Vishnivetskaya TA, Kathariou S, Tiedje JM (2009) The Exiguobacterium genus: biodiversity and biogeography Extremophiles 13:541-555

Wang W et al. (2016) Consistent responses of the microbial community structure to organic farming along the middle and lower reaches of the Yangtze River Scientific reports 6:35046

Whiteley AS et al. (2012) Microbial 16S rRNA Ion Tag and community metagenome sequencing using the Ion Torrent (PGM) Platform J Microbiol Methods 91:80-88 doi:10.1016/j.mimet.2012.07.008

Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: a fast and accurate Illumina Paired-End reAd mergeR Bioinformatics 30:614-620