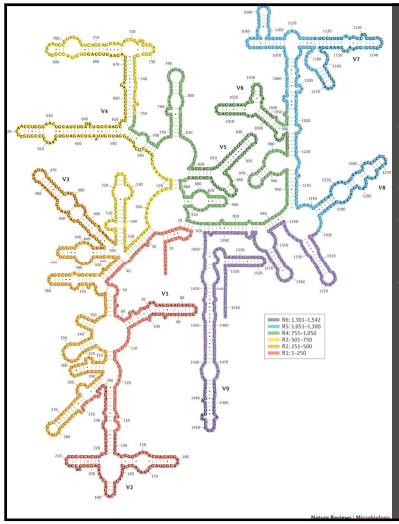
Use of Emerging Technologies for Monitoring of Water Quality and Ecosystem Health

Dr Megan J. Huggett
Dr Troy Gaston
Associate Professor Bill Leggat

Background

- Funded by a Lack Macquarie City Council grant
  - Emerging tools for monitoring water quality
  - Provide information on sources of faecal contamination (both human and non-human)
  - Detect putative pathogens and whole microbial population shifts
- This project tested the applicability of two recently developed approaches to detect human and marine pathogens and identify the source of faecal contamination (dog vs human) in Lake Macquarie.



16s ribosomal region

- Molecular profiling of bacterial populations
- 16s rDNA found in all bacteria
- Consists of a variety of variable and conserved regions
- Variable regions (V1-V9) can be used as a marker to bacterial taxa (sequence variants)

- The two methods that were tested, to evaluate as possible additions to the water quality monitoring program of Lake Macquarie, are:
  - 16S gene sequencing, to detect organisms such as *Vibrio*, *Salmonella* and other pathogenic organisms (not only *E. coli*) and to enable microbial community diversity to be determined.
  - Indicators of human faecal contamination via *Bacteriodetes* Polymerase Chain Reaction (PCR) and sequencing.

Pro and cons of the approches

16s gene sequencing		PCR	
Gives relative abundance of all bacteria	Expensive (~$60/sample)	Cheap ($1-5 /sample)	Specific for 1 (or small group of bacterial types)
Examines known and unknown bacteria	Slow (1-4 weeks for results)	Can be semi-quantitative	Can have false positive
Can detect early changes to bacterial populations	Performed by a specialized lab	Rapid 1 day to results	
		Can be done in any general molecular lab	

Approach

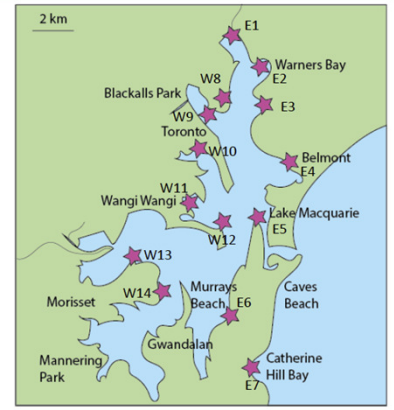
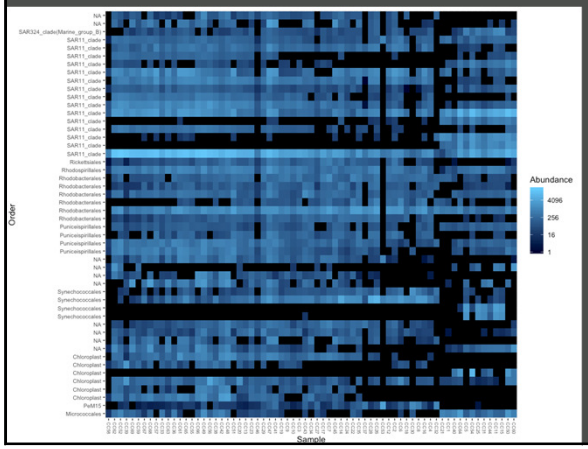


Figure 1: Map of Lake Macquarie showing sample sites

- Water samples taken from 14 sites that are monitored by the Lake Macquarie City Council water quality program
- Samples taken at
  - 4th December 2018
  - 14th January 2019
  - 5th February 2019
  - 5th March 2019
  - 2nd April 2019
- Normal water quality sampling season November – April
- 1 L water samples collected, placed on ice, filtered through a 0.2 µm filter, stored at -80ºC and then DNA extracted
- Water quality measures
- Sequencing – utilised the V3-V4 region of the 16s rDNA gene and sequenced using an Illumina MiSeq
- Bioinformatically analysed
- Putative faecal and pathogenic bacteria further characterised

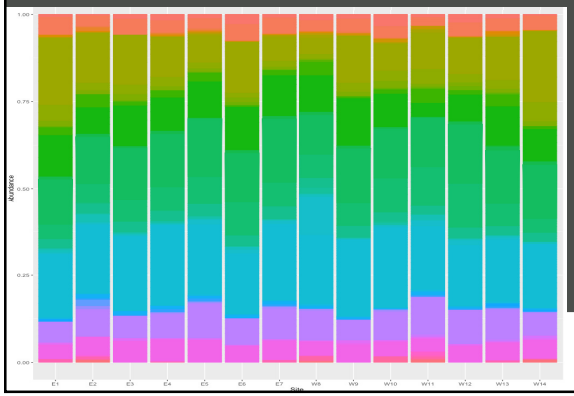
Heatmap of amplicon sequence variants (ASV) present in at least one samples above 5%



- Microbial population of Lake Macquarie is dominated by ”typical” marine and coastal aquatic communities
  - SAR11 Clade
  - SAR86 Clade
  - *Synechococcus*
  - Oceanspiralles

Site differences

Family level mean relative abundance



- At the ASV level there are significant difference between sites ($p<0.05$)
  - 30/92 pairwise comparisons
  - However few obvious geographic pattern

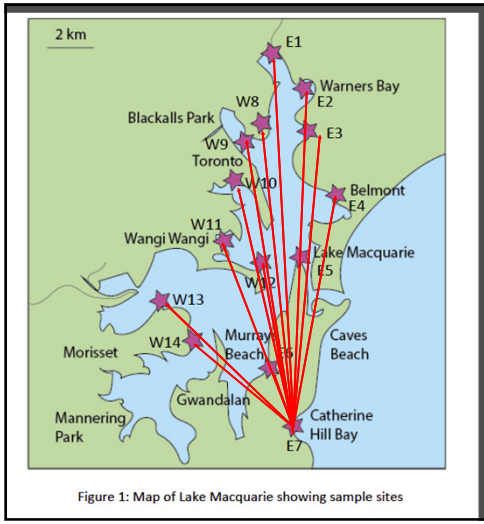


Figure 1: Map of Lake Macquarie showing sample sites

- Open ocean effect
- Catherine Hill Bay significantly different to 12 sites
- Not significantly different to Swansea (E5)

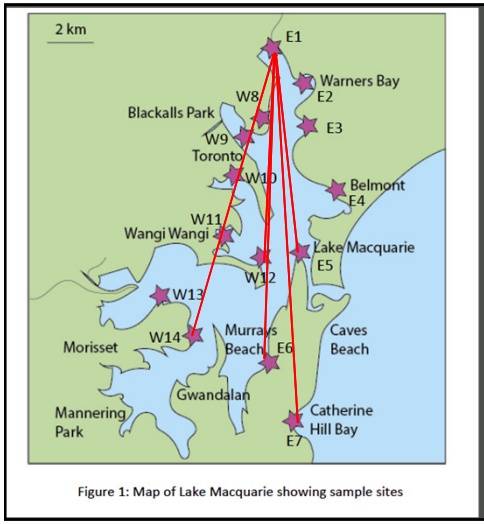
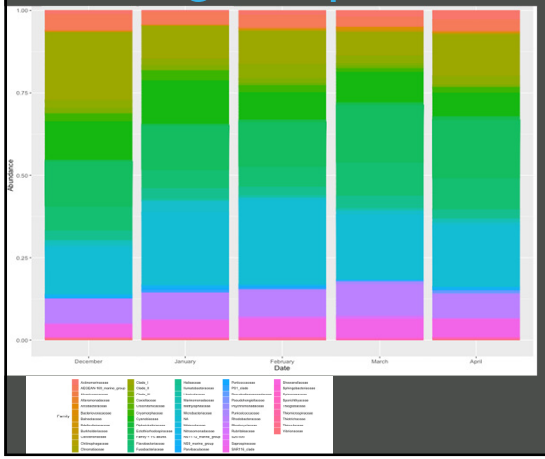


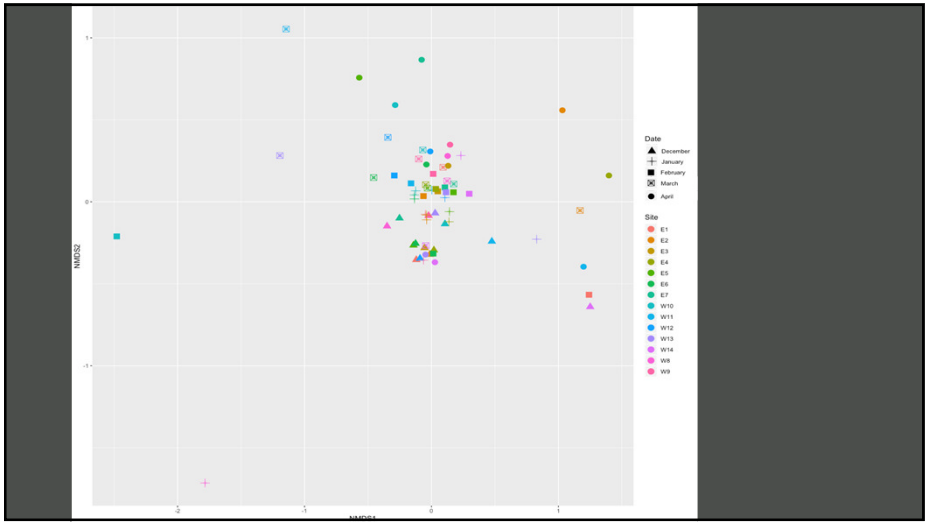
Figure 1: Map of Lake Macquarie showing sample sites

- Some north to South effect
- Spears Point Park significantly different to 5/13 other sites

Strong temporal effects



- Every sampling time point was significantly different at the ASV level ($p<0.001$)



Environmental drivers

- Measured Temperature, Conductivity, Salinity, pH, Turbidity, DO, TDS
- These accounted for 31.4% of the variation in the microbial dataset
- Two main drivers
  - Temperature (19.72 ºC Belmont (E4) April, 30 ºC Balcolyn (W13) January) and
  - Salinity – substantial variability - (16.57 Belmont (W13), 38.83 ppt Eleebana (E2) March)
    - Western sites high in February, eastern sites high in March and January
  - Rainfall over study period
    - December, January and March – dry periods (4 days of no rain)
    - Feb and April followed/during light rains (10-20 mm over previous 4 days)
    - March (38 mm)

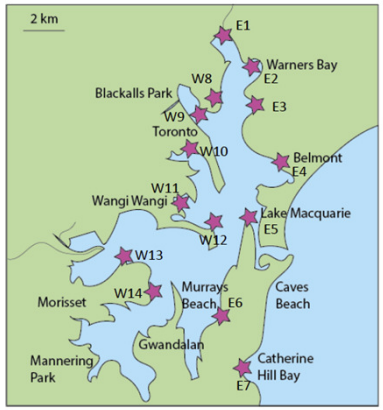


Figure 1: Map of Lake Macquarie showing sample sites

Did sequencing results match *Enterococcus* counts?

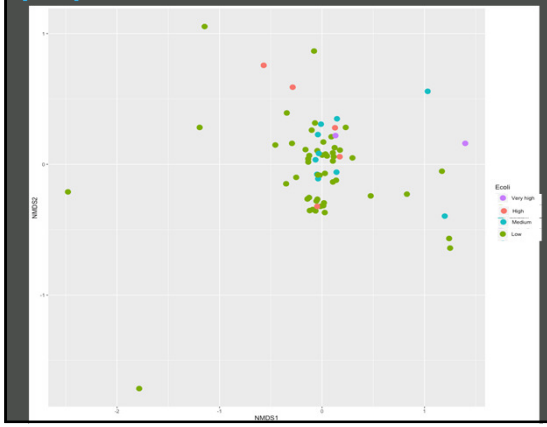
Site	E1	E2	E3	E4	E5	E6	E7	W8	W9	W10	W11	W12	W13	W14
Site	Speers Pt Park	Eleebana	Croudace Bay	Belmont	Swansea	Cams Wharf	Catherine Hill Bay	Bolton Pt	Toronto	Kilaben Bay	Arcadia Vale	Wangi	Balcolyn	Sunshine
December	0	4	4	4	2	2	0	0	0	4	2	2	8	0
January	14	48	0	0	44	10	4	0	6	2	2	0	6	0
February	2	12	2	0	180	2	2	2	2	8	0	8	6	2
March	0	0	110	4	18	0	6	0	0	2	0	0	6	0
April	170	32	600	1600	200	16	0	420	30	110	34	26	260	0

ASV no.	SILVA assigned taxonomy	Accession no. of nearest match in NCBI	Species designation of nearest match in NCBI database	% identity to nearest match in NCBI	Potential pathogen	Possible source
8217	*Bacillus aquimaris*	MK934550 MK262997	*Bacillus aquimaris* *Bacillus vietnamensis*	100% 100%	No	Environmental (aquatic) Environmental (sediment)
6038	*Bacillus* sp.	MN216227 MN213373	*Bacillus cereus* *Bacillus cereus*	100% 100%	Yes	Environmental (soil) Environmental (worm GI* tract)
6316	*Bacillus* sp.	KX898175 NR_024690	Uncultured sequence *Bacillus carboniphilus*	100% 99.53%	No	Environmental (limpet) Unidentified
3401	*Exiguobacterium* sp.	MG190720 MG190711	*Exiguobacterium profundum* *Exiguobacterium aestuarii*	100% 100%	No	Environmental (sediment) Environmental (sediment)
8930	*Exiguobacterium* sp.	MN192433 MN100070	*Exiguobacterium mexicanum* *Exiguobacterium aurantiacum*	100% 100%	No	Unidentified Environmental (plant)
5291	*Listeria* sp.	CP041213 CP041211	*Listeria monocytogenes* *Listeria monocytogenes*	100% 100%	Yes	Unidentified Unidentified
2547	*Planococcus* sp.	MK104526 KU726519	*Planococcus maritimus* *Planococcus rifietoensis*	100% 100%	No	Unidentified Environmental (sheepskin)
3366	*Planococcus* sp.	MN187272 MK879392	*Planococcus maitriensis* *Planococcus maritimus*	100% 100%	No	Environmental (plant) Environmental (biofilm)
10882	Unidentified Bacilli	HM437410 JQ579821	Uncultured sequence Uncultured sequence	100% 100%	No	Environmental (marine) Environmental (sediment)

* GI= gastrointestinal tract

- Sequence analysis did not identify any ASV within the *Enterococcus*
- Identified 9 in the class Bacilli, including two potential pathogens, *Listeria* and *Bacillus*

Was there a relationship between population structure and counts?



- No significant groupings based upon *Enterococcus* load
  - Low <10
  - Medium 10-100
  - High 100-500
  - Very high >500

Counts vs sequencing reads for Bacilli

Site	E1	E2	E3	E4	E5	E6	E7	W8	W9	W10	W11	W12	W13	W14
Site	Speers Pt Park	Eleebana	Croudace Bay	Belmont	Swansea	Cams Wharf	Catherine Hill Bay	Bolton Pt	Toronto	Kilaben Bay	Arcadia Vale	Wangi	Balcolyn	Sunshine
December	0	4	4	4	2	2	0	0	0	4	2	2	8	0
January	14	48	0	0	44	10	4	0	6	2	2	0	6	0
February	2	12	2	0	180	2	2	2	2	8	0	8	6	2
March	0	0	110	4	18	0	6	0	0	2	0	0	6	0
April	170	32	600	1600	200	16	0	420	30	110	34	26	260	0

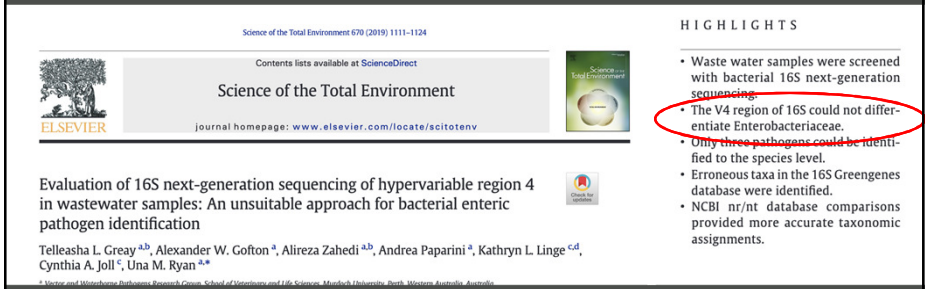
Table 4: Percent abundance and SILVA assigned taxonomy of Bacilli amplicon sequence variants (ASVs) in samples. Missing samples had no Bacilli ASVs.

Site	E4	W11	E5	E3	E5	W14	W8	W13	E5	E6	W12	SILVA assigned taxonomy		
Month	December		January		February		March		April			Family	Genus	Species
ASV8217	0	0	0	0	0	0	0	0	0	0	0	Bacillaceae	*Bacillus*	*aquimaris*
ASV6038	0	0	0	0	0	0.013967	0	0	0	0	0	Bacillaceae	*Bacillus*	NA
ASV6316	0	0	0	0	0.011679	0	0	0	0	0	0	Bacillaceae	*Bacillus*	NA
ASV3401	0	0	0	0	0	0	0	0.021658	0.025147	0.00883	0	Family_XII	*Exiguobacterium*	NA
ASV8930	0	0	0	0	0	0	0.004839	0	0	0	0	Family_XII	*Exiguobacterium*	NA
ASV5291	0.016957	0.005567	0	0	0	0	0	0	0	0	0	Listeriaceae	*Listeria*	NA
ASV2547	0	0	0.007208	0	0	0	0	0	0.066296	0.022074	0	Planococcaceae	*Planococcus*	NA
ASV3366	0	0	0	0	0	0	0	0	0	0	0.04579	Planococcaceae	*Planococcus*	NA
ASV10882	0	0	0	0.002785	0	0	0	0	0	0	0	NA	NA	NA

Conclusions

- Sequencing of the V3-V4 region of the 16S rDNA gene of samples within the Lake revealed a typical marine/coastal aquatic assemblage, which was driven primarily by changes in temperature and salinity.
- Some geographic patterns of microbial populations
- Sequencing also revealed a diverse number of both faecal indicator bacterial taxa, as well as putative pathogens. These were generally present in very low mean relative abundance within samples.
- The indicator taxa and putative pathogens were most abundant and most diverse during the same sample time points that *Enterococcus* counts were highest within the system, that is, during April 2019.

Future directions



Science of the Total Environment 670 (2019) 1111–1124

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Evaluation of 16S next-generation sequencing of hypervariable region 4 in wastewater samples: An unsuitable approach for bacterial enteric pathogen identification

Telleasha L. Greay [a,b], Alexander W. Gofton [a], Alireza Zahedi [a,b], Andrea Paparini [a], Kathryn L. Linge [c,d], Cynthia A. Joll [c], Una M. Ryan [a,*]

HIGHLIGHTS

- Waste water samples were screened with bacterial 16S next-generation sequencing.
- The V4 region of 16S could not differentiate Enterobacteriaceae.
- Only three pathogens could be identified to the species level.
- Erroneous taxa in the 16S Greengenes database were identified.
- NCBI nr/nt database comparisons provided more accurate taxonomic assignments.

- Currently resequencing subset of samples using a different variable region
- Expanding to other sites in conjunction with the Central Coast Council to examine effectiveness

Acknowledgments

UoN

Tess Moriarty

Alessandra Suzzi

Volunteers

LMCC

Andrew Irelend

LMCC officers

PCR approach