

Metagenomic Shotgun Sequencing Provides Prevalence Data for Pathogens, and Source-Tracking Indices Useful in Public Health Risk Assessment of Environmental Waters

Brian Mercer¹, Karim Dawkins¹, Lisa Meday², Nwadiuto Esiobu^{1*}

¹Microbial Biotech Lab Sanson Building, Florida Atlantic University, Boca Raton, USA

²Environmental Resource Management Department Water Quality, Seminole Tribe of Florida, Hollywood, USA

Email: bmercer2012@fau.edu, dawkinsk2014@fau.edu, LisaMeday@seminoletribe.com, *nesiobu@fau.edu

How to cite this paper: Mercer, B., Dawkins, K., Meday, L., & Esiobu, N. (2020). Metagenomic Shotgun Sequencing Provides Prevalence Data for Pathogens, and Source-Tracking Indices Useful in Public Health Risk Assessment of Environmental Waters. *Journal of Geoscience and Environment Protection*, 8, 115-129.

<https://doi.org/10.4236/gep.2020.86009>

Received: May 27, 2020

Accepted: June 27, 2020

Published: June 30, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

State-approved membrane filtration (MF) techniques for water quality assessments were contrasted with metagenomic shotgun sequencing (MSS) protocols to evaluate their efficacy in providing precise health-risk indices for surface waters. Samples from a freshwater receiving pond (ABI-1002) and two upstream storm water ditches (ABI-1003) and (ABI-1004) yielded alarmingly high *Fecal coliform* MF densities of 220, >2000 and >2000 CFU/100ml respectively. The indicator, *Enterococcus* bacteria exceeded allowable limits in all but the equipment control (ABI-1001). Using MSS, the relative numerical abundance of pathogenic bacteria, virulence and antibiotic resistance genes revealed the status and potential pollution sources of each ditch. High levels of *Shigella sp.* (0 (ABI-1001), 4945 (ABI-1002), 55,008 (ABI-1003), and 2221 (ABI-1004) genomic reads/100ml) correlated with virulence genes and antibiotic resistance genes found in fecal samples for ABI1003 and not ABI1004. Traditional culture methods (TCM) showed possible fecal contamination in two of the four samples, and no contamination in the others. MSS clearly distinguished between fecal and environmental bacteria contamination sources, and pinpointed actual risks from pathogens. Our data underscore the potential utility of MSS in precision risk assessment for public and biodiversity health and tracking of environmental microbiomes shifts by field managers and policy makers.

Keywords

Metagenomic Shotgun Sequencing, Water Quality, Risk Assessment, Indicator Bacteria, Environmental Microbiomes

1. Introduction

For over 50 years, microbiological risk assessment of environmental samples has been based on monitoring the prevalence of indicator organisms that are generally not harmful, but indicate the possible presence of pathogenic bacteria, viruses and protozoans (EPA 822-R-10-005, 2010). The widely used indicator bacteria include easily cultured heterotrophic bacteria—*Enterococcus spp.*, *Escherichia coli*, *Fecal coliforms* and more recently *Bacteroidetes spp.*, *Clostridia spp.*, some phages and bio-marker genes (Esiobu et al., 2004; Wade et al., 2006; Noble et al., 2003; Coakley et al., 2015). The current regulations of surface water quality standards for Florida are referenced in document 62 - 302 (Water Quality Standards Variances. Epa.gov). Enteric organisms such as *E. coli*, *Fecal coliforms*, and *Enterococci spp.* are the only well-regulated indicators of microbiological health risks. The safety thresholds vary and are based on the class of water and organism being tested. Issues related to the sensitivity and specificity of the indicator detection systems have been continually revised and improved with advances in chromogenic media and one-step assays (Odonkor & Ampofo, 2013; Ferguson et al., 2013). However, numerous limitations remain. Reviews and discussions of these challenges abound in literature (Rochelle-Newall et al., 2015; Tan et al., 2015; Figueras & Borrego, 2010; Evangelista & Coburn, 2010), and include but not limited to questions about the suitability of enteric indicators for respiratory/skin illnesses, the interpretation of prevalence data in non-point source environments, the difficulty in distinguishing between environmental strains and real indicators (fecal strains), the lack of indicators for biodiversity in preserved areas and finally, the utility of the numbers (Esiobu et al., 2013). The US Environmental Protection Agency (EPA) manual for monitoring (EPA 5.11 Fecal Bacteria, 2012; EPA-820-R-14-010, 2014; EPA-820-R-14-011, 2014) justified the indicator system because “it is difficult, time-consuming, and expensive to test directly for the presence of a large variety of pathogens; water is usually tested for coliforms and fecal *Streptococci* instead”. In some instances, these relatively inexpensive culture techniques provide data considered sufficient to accurately assess (at least in part) the public health risk of water samples. On the other hand, interpretation of results from non-point sources of pollution is not always clear-cut. To address the many challenges that remain unsolved, emerging technologies such as the metagenomic shotgun sequencing, which can detect virulence genes and all microbial life forms—viruses, bacteria, fungi and protozoans in a single assay are being developed. These permit a direct rather than an indirect assessment of public health risks. Techniques such as amplicon metagenomic analysis using 16S rRNA pyrosequencing (Gomez-Alvarez et al., 2012) are being used to rapidly and effectively monitor different disinfection treatments of drinking water samples. Similarly, Cabral et al. (2018) employed sequencing technologies to successfully characterize microbial communities in water ways while other recent studies (Mohiuddin et al., 2017; Roy et al., 2018; Cocolin et al., 2018; Li et al., 2018) have demonstrated the utility of next genera-

tion sequencing techniques and metagenomic shotgun sequencing in gaining insight to the complex problems associated with surface water quality, and direct detection of possible pathogenic organisms in beach sand. Furthermore, whole genome sequencing has been employed to study functional genes in potable water treatment systems and associated biofilms (Douterelo et al., 2018). Wide-scale use of these emerging sequencing technologies for environmental monitoring will require experimentally modelled interpretations of the big-data generated, as well as comparative analysis of the results with validated culture methods and other epidemiological indices. In this study we determined and compared the microbiological water quality of four samples using the traditional EPA standard methods and metagenomic shotgun sequencing of genomic DNA extracted directly from water to evaluate their relative efficacy in predicting actual health risks to the public and biodiversity in the environment.

2. Materials & Methods

2.1. Sampling

Four water samples were collected in September of 2016 aseptically from: (A) deionized water used to ensure sampling equipment was contaminant free (ABI-1001)—(quality control), (B) a receiving freshwater pond used for recreational activities (ABI-1002), a storm water ditch whose effluent enters the freshwater pond (ABI-1003), and a second storm water ditch located the furthest from the pond and whose effluent merges with those of ABI-1003 before emptying into the pond (ABI-1004.) The exact locations of sample points are kept confidential for privacy purposes. All samples were collected at the same event after rainfall to ensure storm water ditches held water. *Fecal coliform* membrane filtration and *Enterococcus* membrane filtration samples were collected in 100 ml sterile vessels containing sodium thiosulfate, sufficient to neutralize any free chlorine present. Metagenomic shotgun sequencing samples were also collected from the same locations, but in 1 L amber bottles (sterilized by rinsing with 10% bleach) with no preservatives. All samples after collection were placed immediately on ice and stored at 4°C in the laboratory until analysis.

2.2. Traditional Culturing Techniques

2.2.1. *Fecal coliform* Enumeration

Fecal coliform bacteria were enumerated following the standard methods SM9222D method (Clesceri et al., 1998). Filtration volumes were based on multiple years of historical data for these sample sites, and results were reported in Colony Forming Units (CFU)/100ml. Two dilutions were prepared for each sample (10 ml and 50 ml) following NELAP protocol of recovering 20 - 60 CFU. Blanks were also prepared to ensure sterility of membrane filter equipment using sterile deionized (DI) water. All samples were processed within the Florida Department of Environmental Protection (DEP) guidelines of 8 hours after collection.

2.2.2. Enterococci Enumeration

Enterococcus spp. were detected using the EPA 1600 method for enumeration of *Enterococci* from water (EPA-821-R-06-009, 2006). Dilutions were carried out as in *Fecal coliform* above, based on multiple years of historical data for these sample sites, and results were reported in CFU/100ml. Two dilutions were prepared for each sample (10 ml and 50 ml) following Standard Methods protocol of recovering 20 - 60 CFU. Blanks were also prepared to ensure sterility of membrane filter equipment using sterile DI water. All samples were processed within the Florida DEP guidelines of 8 hours after collection.

2.3. Metagenomic Shotgun Sequencing

Five hundred (500) ml of each water sample was thoroughly mixed and vacuum filtered through 0.2 μm pore-sized polycarbonate filters, and a replicate sample filtered with the remaining 500 ml. All cells were washed out and concentrated into 5 ml of sterile DI water by placing the polycarbonate filter into a 50 ml centrifuge tube containing 5 ml of sterile DI water and vortexing. The suspension of cells was then harvested by ultra-centrifugation at 15,000 rpm in 1.5 ml aliquots. The pelleted cells were pooled into one tube before lysis and DNA extraction using the Qiagen DNeasy protocol according to manufacturer instructions (Qiagen, Valencia, CA). Metagenomic DNA concentration, purity and quality were then measured with NanoDrop 2000c and gel electrophoresis. After standardized library preparations, sequencing was outsourced to COSMOSID[®] (1600 E Gude Dr. Rockville, MD 20850 United States). DNA libraries were constructed using COSMOSID[®] proprietary Library Prep Kit. The PCR products were purified using 1.0 \times speed beads and eluted in 15 μL of nuclease-free water, then quantified by PicoGreen fluorometric assay. Libraries were then pooled and loaded onto a high sensitivity chip run on a Caliper LabChipGX (Perkin Elmer, Waltham, MA) to estimate the size. Libraries were sequenced using Illumina NextSeq/HiSeq platform. Sequencing primers targeted a full repertoire of bacteria, fungi, viruses, and protozoan parasites; and also included virulence factors and antibiotic resistance markers. All sequences obtained were quality filtered and subsequently aligned and classified using the GenBank[®] database, a highly curated database comprising over 150,000 microorganisms' genomes and gene sequences. This database consists of both private and publicly available genomes from sources such as NCBI-RefSeq/WGS/SRA/nr, PATRIC, M5NR, IMG, ENA, DDBJ, etc. Sequence data were deposited in the NCBI Sequence Read Archive (SRA) under accession number SRP259849.

3. Results

Results obtained from this study are all reported as CFU/100ml unless otherwise stated, and the samples are labeled as ABI1001-ABI1004, with ABI1001 being the equipment control (blank) as previously mentioned.

Figure 1 displays the results obtained for traditional indicators—*Fecal coliforms*

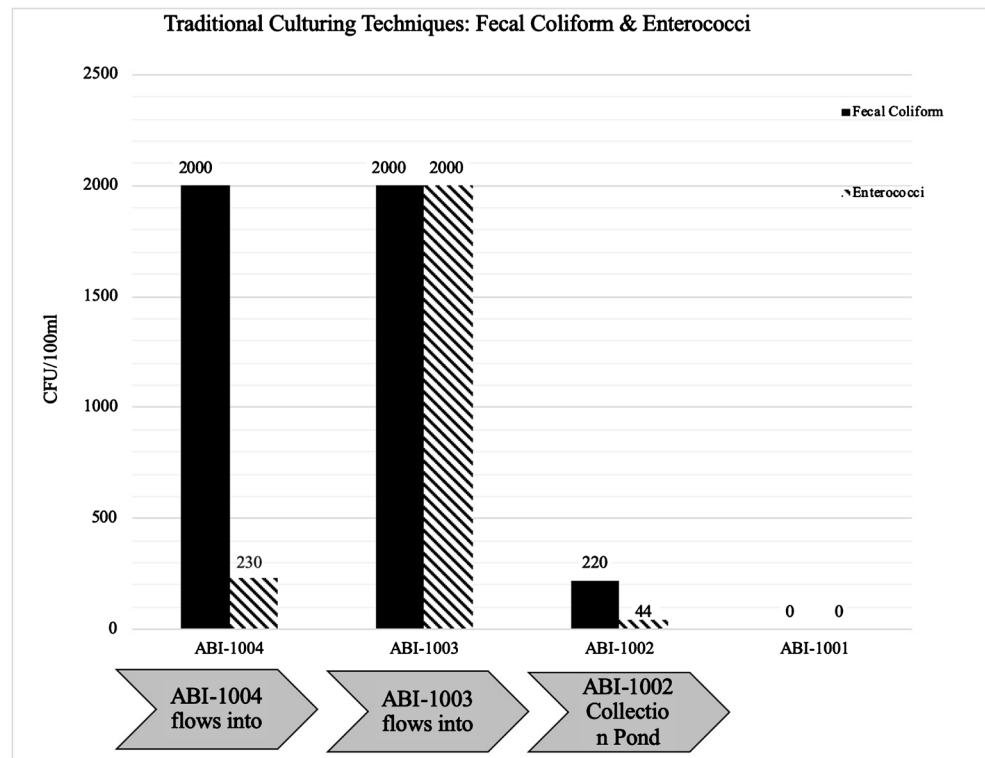


Figure 1. Population of classic indicator bacteria (*Fecal coliforms* and *Enterococci*) in a receiving pond and storm water ditches using the traditional culture techniques CFU/100ml (Colony Forming Units per 100 ml) based in standard Methods guidelines. ABI1001 is an equipment/reagent QC blank.

and *Enterococci* using traditional culture techniques—SM9222D and EPA 1600. The lowest run dilution for these samples was 10 ml, making the highest possible detection limit to be 2000 CFU/100ml. As such all results that are at 2000 CFU/100ml are presumed to be greater than 2000 CFU/100ml, because the plates were too numerous to count. Results reported for these culturing techniques are based on the most accurate number calculated from the volume of sample filtered. It is noteworthy that while *Fecal coliforms* (a nearly ubiquitous enteric group of bacteria in animals) exceeded detection limits in ABI-1003 and 1004; enterococcal levels were contrasting; being significantly less in the latter. The numbers in the receiving pond and control samples were within expected levels.

Figure 2 shows the distribution of the traditional indicator organisms in each sample using metagenomic shotgun sequencing calculations (genomic reads/100ml). Indicator organisms labeled as coliform represents all the genera of the coliform group (*Citrobacter*, *Enterobacter*, *Klebsiella*, and *Escherichia*) found in each sample. The *Enterococcus* value is the sum of all species of *Enterococcus* found in each sample. Species for each category were added up and calculated based on relative abundance, then converted to genomic reads/100ml. MSS numbers sharply contrast the values obtained from culture techniques. Whereas ABI 1003 and 1004 showed a similar trend in **Figure 1**, MSS detected more than 24,000

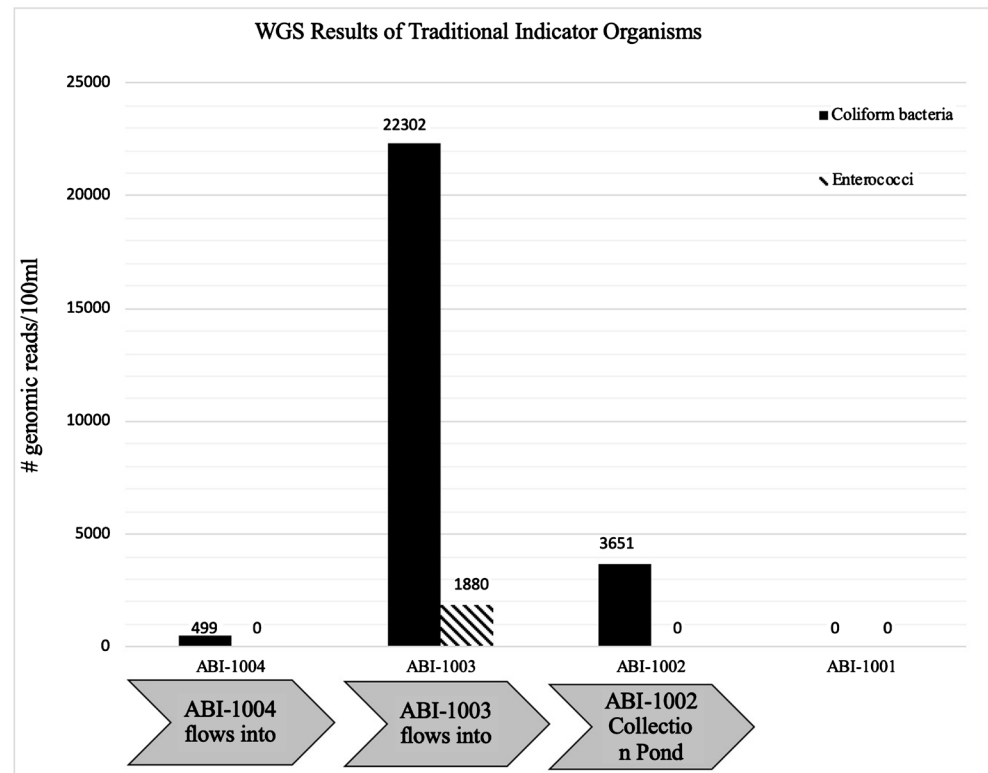


Figure 2. The traditional indicator organisms detected using metagenomic shotgun sequencing MSS. All genera of the Coliform group, and all species of the *Enterococcus* genus were summed up for a final count.

Fecal coliforms in ABI 1003; only 499 for ABI 1004 and more than 3000 for the receiving pond. Although the equivalency of genomic reads and colony forming units is not necessarily linear, the remarkable difference between both techniques underscore the challenge with sensitivity and specificity of risk assessment assays.

In **Figure 3**, the prevalence of some potentially pathogenic bacteria species found in each sample is presented in genomic reads/100ml of sample to allow comparisons with results obtained from traditional methods. *Shigella sp.* were detected at the following densities—(0 (ABI-1001), 4945 (ABI-1002), 55,008 (ABI-1003), and 2221 (ABI-1004) genomic reads/100ml). With the exception of *Pseudomonas aeruginosa* which is ubiquitous in nature, residing in soils, water, and human skin; sample ABI 1003 contained 100 to 1000 times more potential pathogens than ABI 1004 even though both samples were incorrectly rated of similar pollution trend based on traditional indicator culture approach (**Figure 1**). Most of the pathogens detected in high numbers include *Vibrio cholerae* and *Staphylococcus spp.* and other enteric organisms. None of the common pathogens was detected in the receiving pond.

In addition to specific pathogens, indicator genera, a single assay of MSS also detects virulence and antibiotic resistance genes. In **Figure 4(a)**, the distribution and abundance of commonly known bacterial virulence genes found in the samples

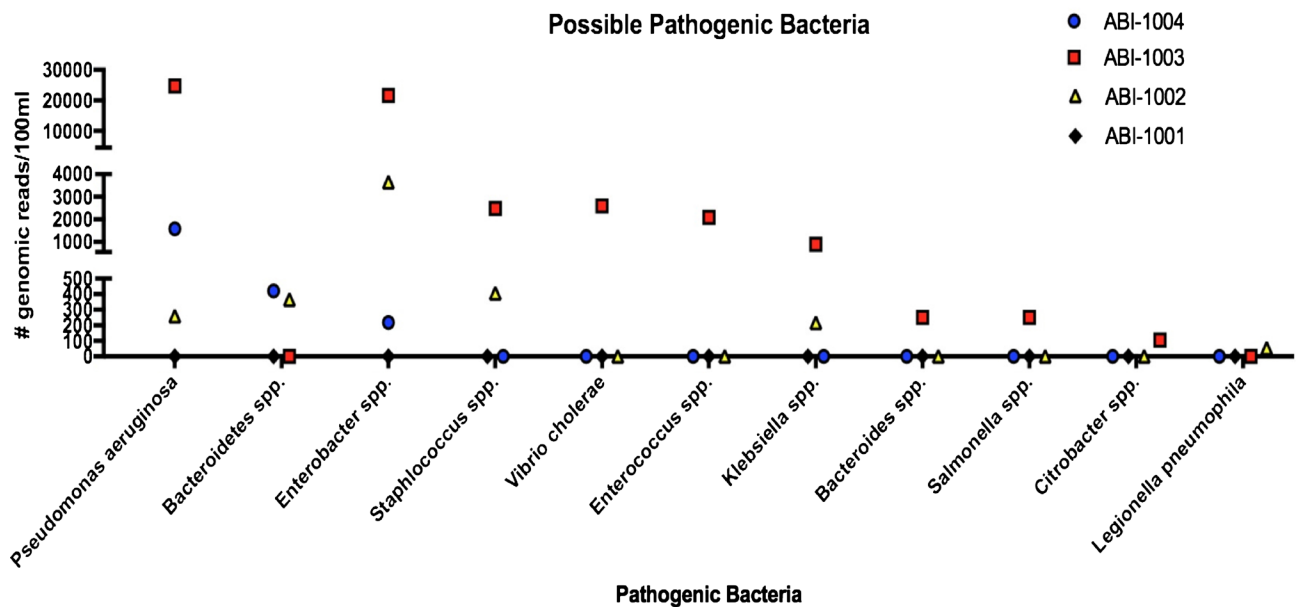


Figure 3. A representation of the potential pathogenic bacteria in each sample expressed in genomic reads/100ml of sample. Those expressed as spp. Had multiple species detected, and all were added together.

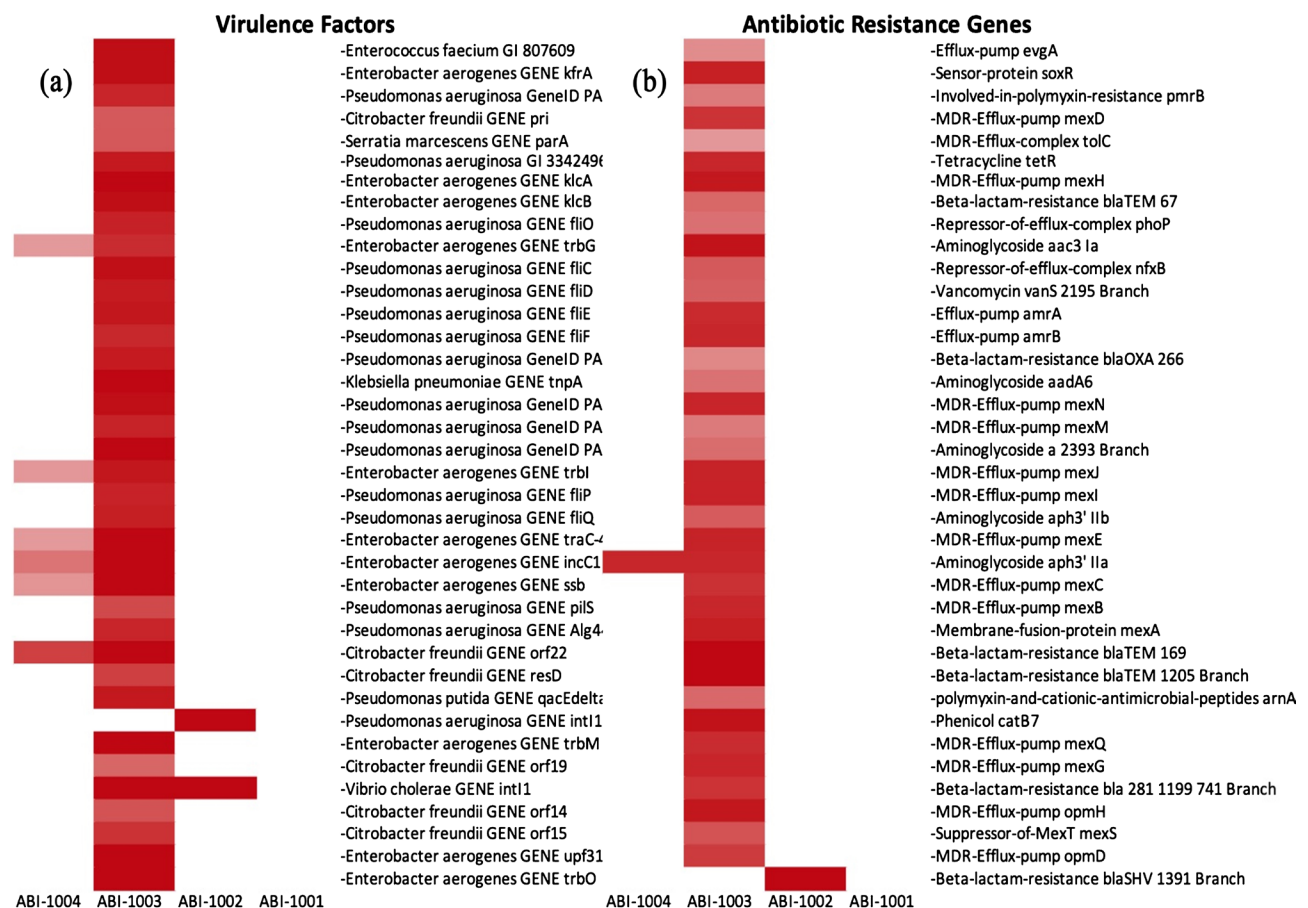


Figure 4. (a) Virulence genes detected (from filtered above threshold data) in each of the 4 samples. ABI-1003 represents the top 10% of virulence genes detected. The more maroon the bar is, the higher incidence that gene was detected. (b) A represents the top 50% of antibiotic resistance genes detected. The more maroon the bar is, the higher incidence that gene was detected.

are displayed Sample ABI-1003 contained additional virulence genes, too numerous to include in one figure. **Figure 4(b)** displays a strikingly important distribution of antibiotic resistance genes associated with the samples. ABI-1003, again, had much more diversity and prevalence of antibiotic resistance genes than could be included in this chart. There is a clear correlation between virulence/antibiotic resistance genes and the density of *Shigella sp.* and other potential pathogens presented in **Figure 3**. The absence or relatively low density of these genes in sample ABI1004 is congruent with results displayed in **Figure 2** and **Figure 3** but sharply contrast traditional culture results which recorded high abundance of *Fecal coliforms* and enterococci (**Figure 1**).

The relative occurrence of bacteria phyla detected in all samples is shown in **Figure 5**. This reveals the relative community structure and diversity of the environmental water sample. Each value is the sum of all the taxonomic ranks for that phylum, present in the sample. The receiving pond-ABI 1002 and sample ABI 1004 have the highest natural bacteria diversity while ABI1003 seemed enriched with Proteobacteria, many of which are enteric organisms.

A summary of the various parameters derivable from the TCM and the MSS are presented in **Table 1**. Whereas TCM yields numerical values of mere indicators of the potential presence of pathogens, the MSS quantifies indicators, pathogenic bacteria and viruses as well as virulence and antibiotic resistance genes together provide a highly reliable measure of the not just the quality but also the source of pollution.

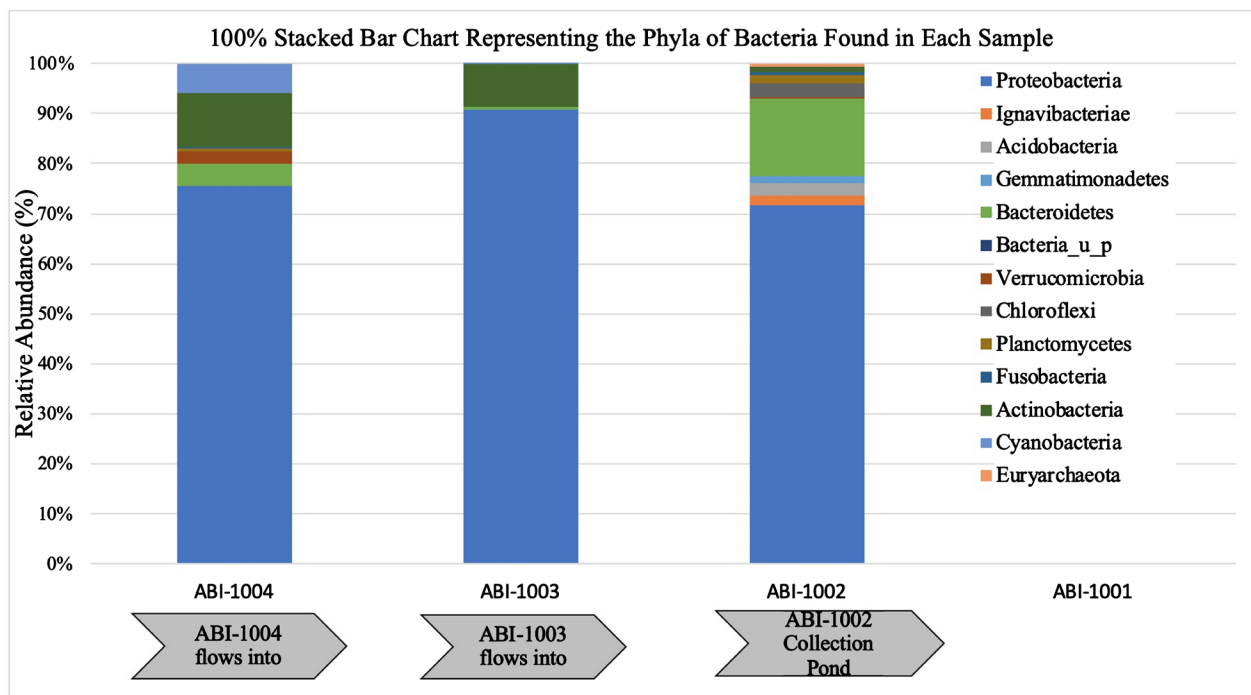


Figure 5. Bacteria community structure of a fresh water receiving pond and two storm water drains and a control, all bacteria phyla detected are expressed in percent relative abundance of total genomic reads. These results correspond to the total 500 ml of each sample.

4. Discussion

4.1. Traditional Culture Methods

Using traditional culturing methods alongside of metagenomic shotgun sequencing allowed us to compare these methods for the first time to understand the microbial community indices versus membrane plate count results. The traditional culturing methods (**Figure 1**) showed that samples ABI-1003 and ABI-1004 exceeded the Florida *Fecal coliform* limit of not to exceed 800 CFU/100ml on any one day) as per SM 9222D for *Fecal coliform* organisms (Clesceri et al., 1998). Furthermore, ABI-1002 showed relatively low concentrations of traditional indicator organisms, and traditional test results were within the acceptable criteria of Florida surface water limits. These methods are beneficial due to their cost effectiveness and quick production of results. Although this information is quick and easy to obtain, traditional indicator organisms tend to adapt and survive for prolonged periods in the humid and warm Florida climate (Bonilla et al., 2007; Hartz et al., 2008). Besides, the phenotypic and biochemical assays are prone to yield false positive results due to the enormous versatility of environmental bacteria metabolism. Also, many environmental factors can influence the results of these methods such as: runoff of soil from rainfall, debris entering sampling point, and natural conditions such as shade from sunlight.

4.2. Metagenomic Shotgun Sequencing

It was possible to obtain a plethora of data from metagenomic shotgun sequencing not discernible from the traditional culture methods used in this study. MSS results enabled the construction of a clear risk assessment by more than one layer of evidence: human pathogens, human pathogen relative abundance, antibiotic resistance genes, virulence genes, viruses, and also all other microorganisms such as environmental bacteria. The relative abundance of human associated pathogens can indicate source and potential risk of an environmental sample (Wade et al., 2006). All of the parameters obtained in a single test (including strain level detection) allowed for potential source-tracking of any fecal pollution present in a given sample. Metagenomic shotgun sequencing can also be applied to define a microbiome for particular environmental water bodies. Understanding the water body microbiome condition could allow for fewer repeats of traditional culture testing and create a more comprehensive fingerprint for long term monitoring of changes in user behaviors and climate.

Metagenomic shotgun sequencing data for traditional indicators are depicted in **Figure 2**. Agreeing with the TCM (**Figure 1**) MSS showed very high levels of indicator organisms in ABI1003. However, it sharply contrasted the results for samples ABI1002 and ABI1004, where elevated levels in the TCM were low in MSS and vice versa. Furthermore, where traditional methods limit total indicator bacteria levels to an estimated maximum value (**Figure 1**), the MSS data in **Figure 2** provided numeric levels, exceeding 20,000 reads/100ml in some cases. **Figure 2** also demonstrates the pattern of indicator bacteria over space, where

higher levels were found in ABI-1003 and present in lesser amounts at the pond (ABI-1002). Indeed, no possible pathogenic bacteria were present at the pond (ABI-1002) at higher levels than were found in the two storm water ditches (**Figure 3**), and in several cases, pathogenic bacteria strains found in the ditches were not present in the pond. This degree of specificity could be very useful for environmental or public health scientists investigating non-point pollution sources.

Pathogenic bacteria in **Figure 3** showed some potentially harmful bacteria in the samples, but ABI-1003 contained the highest density of pathogenic bacteria. Organisms such as *Vibrio cholerae* have been found in known contaminated surface waters in Haiti (Roy et al., 2018) using MSS technology. This organism was found in high quantities in sample ABI-1003 suggesting some source of fecal contamination. In **Figure 4(a)** and **Figure 4(b)**, sample ABI-1003 clearly contained the most virulence and antibiotic resistance genes. This is a striking confirmation of the high prevalence data of actual pathogens for ABI-1003 while most of the other samples contained environmental microorganisms (**Figures 1-3**). The most remarkable evidence of the potential fecal contamination of sample ABI 1003 is the abundance and type of virulence and antibiotic resistance genes associated with human enteric organisms. A correlation plot based on sequence reads shows that virulence genes had a low correlation with viral load except for the fecal polluted sample ABI-1003. On the other hand, a high pathogen density almost always correlated with elevated total bacterial population in exactly the same way as low pathogen counts. While this is not surprising because most of these organisms (pathogen and non-pathogens) are largely heterotrophic, the result underscores the problem of relying on just numbers in risk assessment. Distinguishing fresh pollution by pathogenic strains of bacteria and physiologically similar ecotypes requires rigorous discriminatory tests like the MSS which provides multiple indices of the water quality.

Most of the bacteria found in sample ABI-1003 in **Figure 3** are enteric bacteria of fecal origin. Close-up scrutiny revealed that only one *Bacteroides* spp. was found in ABI-1003, the classic *Bacteroides fragilis* str S6L5. This strain of *Bacteroides fragilis* is also possibly associated with human gastro-intestinal disease, although research is still being conducted. No specific human *Bacteroides* sp. was detected in samples ABI-1002 and ABI-1004. The only bacteroides were members of the Bacteroidetes phylum with no specific annotation, suggesting environmental animal sources. This deduction is fully congruent with the virulence and antibiotic resistance profiles in **Figure 4** which depicts a clear-cut difference between the samples, with only ABI-1003 having more virulence genes associated with enteric bacteria. Taken together, all the results from the MSS hone in on a human fecal bacteria presence at site ABI-1003.

Analysis of taxonomic data from the sequence results of ABI-1002, contrasted the results obtained from the traditional culture methods. It is interesting to note in **Figure 2** that no *Enterococcus* spp. were reported in the major taxa found in samples ABI-1002, and ABI-1004 using metagenomic sequencing, while tradi-

tional methods detected low numbers of Enterococci. This is possibly due to a low sensitivity of the MSS assay where only a few nanograms of DNA is used for sequencing, regardless of amount of sample DNA available. This limitation could result in organisms and entities in very low relative abundance not being detected. Another interpretation of the apparent discrepancy is that the organisms were well below risk levels and within safety thresholds, and so not detecting such microbes would not alter the utility of the MSS. Nevertheless, experimental determination of the actual threshold sensitivity test will enable its large-scale application in food and environmental quality control.

4.3. Benefits and Downfalls of TCM and MSS

Unlike TCM, metagenomic shotgun sequencing allows for coupling multiple layers of risk assessment indices obtained concurrently to make predictions with a very high level of confidence. The profile of virulence genes detected in ABI 1004 is more diverse and abundant than the relatively uncontaminated ABI 1002, and both samples were completely devoid of the genes present in ABI 1003. When coupled with data in **Figure 3**, it becomes apparent that the former contained non-pathogenic naturally occurring organisms. Given the larger dataset of pathogenic bacteria, virulence genes, and antibiotic resistance genes associated with human enteric organisms, the MSS flags ABI-1003 as a suspect.

Table 1 is a summary of the detection capabilities of the traditional methods vs. metagenomic shotgun sequencing. Despite its versatility and specificity, another limitation of metagenomic shotgun sequencing of total DNA from a sample is the possibility of sequencing DNA from dead non-viable cells. This can skew assessments of public health risks and source-tracking. Notwithstanding, coupling

Table 1. Overview of samples and parameters Assessed by detection capabilities of traditional culturing techniques and Metagenomic shotgun sequencing.

Test Protocol	Metagenomic Shotgun Sequencing				EPA Approved Traditional Culturing Techniques	
Expected Results	Levels and Diversity of Traditional Indicator Bacteria, Pathogenic Bacteria, Virulent Genes, Antibiotic Resistant Genes				Levels of Fecal Coliform Bacteria, <i>Enterococci spp.</i> , Possible Contamination	
Sample ID	Traditional Indicators (genomic reads/100ml)	Known Pathogens (genomic reads/100ml)	Total Number of Virulence Genes	Total Number of Antibiotic Resistance Genes	<i>Fecal coliform</i> (CFU/100ml)	<i>Enterococci spp.</i> (CFU/100ml)
ABI1004 (Storm Water Drain)	499	2600	14	1	>2000	230
ABI1003 (Storm Water Drain)	24182	58700	289	70	>2000	>2000
ABI1002 (Freshwater Collection Pond)	3651	4750	6	1	220	44
ABI1001 (Equipment Blank/Control)	0	0	0	0	0	0

information from viruses, virulence factors and antibiotic resistance should allow appropriate conclusions to be reached. Sequencing environmental RNA instead of DNA, in order to establish only viable organisms in the samples would eliminate any uncertainties. Also increasing the depth of sequencing could increase sensitivity of the MSS technique (Rashkin et al., 2017).

There are various benefits and downfalls for both methods tested. Traditional culture techniques can be performed for a couple hundred dollars per sample, providing indicator results for bodies of waters with 24 - 48 hours. Normal laboratory reporting turnaround time for commercial labs for traditional methods ranges from 7 - 21 days. Metagenomic shotgun sequencing of water samples is an niche that is yet unravelling. Low demand for this type of in-depth testing greatly increases the cost of analysis to over a couple thousand dollars for this intense analysis. Turnaround time is about 4 weeks at best currently. Test time and cost will sharply decline as test becomes more popular. The wealth of information obtained from MSS is hard to ignore in a progressive scientific world where precision and confident prediction of risk are critical. An entire microbiome of water bodies can be characterized using this method when sampled over a period of time, giving far greater sensitivity for assessing the public health of a body of water, and also providing insight of potential sources of contamination. Once established, any shift in the microbiome of the sample site will provide detail as to what potential problems may have occurred (fecal pollution, protentional algae blooms, nutrient influx, change in natural conditions, etc.). The utilization of new technologies to assess public health impacts from water bodies is an overlooked benefit.

5. Conclusion

This study confirmed the gross limitations of the traditional culture gold standard for assessing water quality of surface waters to include false positives for indicator bacteria and lack of source-tracking components. The numbers of real positives did not always correlate with actual pathogen presence. However, the TCM is a relatively inexpensive assay with short turnaround time. Metagenomic shotgun sequencing technology on the other hand provided a powerful resolution for a water body where state-mandated thresholds for safe recreation were exceeded, and for which a non-point or point source was not clearly evident. Multiple risk assessment parameters obtained from the relatively expensive MSS analysis include specific identification/quantification of all microbiological entities, virulence genes and antibiotic resistance genes; allowing for a robust health risk evaluation in one step. In addition, the comprehensive data on the microbiomes of the niche provide important base-line reference for early detection and intervention in cases of anthropogenic or climate change perturbation. The need to define and standardize the sensitivity of MSS (assuring the detection of only viable bacteria, even at low relative abundance) is critical to its wide applications in environmental health management. Until the high cost of the MSS is

tampered, environmental managers could utilize this powerful technology for decision-making on preserved and protected surface waters as well as major drainage and hydrological activities that could impact biodiversity other than humans. In all other cases, a periodic application of MSS could be coupled with routine TCM monitoring of sites, thereby establishing a baseline for MSS and creating a database overtime for sample sites, to have a reference of protentional changes to the microbiome.

Authors' Contributions

BM organized sample collection, assisted with wet-lab microbiology, contributed to data analysis and wiring of the manuscript. KD helped with preparation of metagenomic DNA samples. LM assisted with sample collection, contributed to manuscript write up and data presentation. NE conceived and designed study; supervised experimental procedures, directed data analysis and interpretation; and helped with writing of the manuscript.

Acknowledgements

We would like to thank Florida Atlantic University for assistance ships for Brian Mercer and Karim Dawkins. We would also like to thank the Seminole Tribe of Florida for providing partial funding and assistance during this experiment. We would also like to thank Lyle Johnson and Florida Spectrum Environmental Services Inc. for supporting the traditional culturing techniques of this experiment. Finally, we would like to thank COSMOS ID for technical support during this experiment.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

Brian Mercer declares he has no conflict of interest. Karim Dawkins declares he has no conflict of interest. Lisa Meday declares she has no conflict of interest. Nwadiuto Esiobu declares she has no conflict of interest.

References

- Bonilla, T. D., Nowosielski, K., Cuvelier, M., Hartz, A., Green, M., Eiobu, N., McCorquodale, D. S., Fleisher, J. M., & Rogerson, A. (2007). Prevalence and Distribution of Fecal Indicator Organisms in South Florida Beach Sand and Preliminary Assessment of Health Effects Associated with Beach Sand Exposure. *Marine Pollution Bulletin*, 54, 1472-1482. <https://doi.org/10.1016/j.marpolbul.2007.04.016>
- Cabral, B. C. A., Hoffmann, L., Budowle, B., Ürményi, T. P., Moura-Neto, R. S., Azevedo, S. M. F. O., & Silva, R. (2018). Planktonic Microbial Profiling in Water Samples from a Brazilian Amazonian Reservoir. *Microbiology Open*, 7. <https://doi.org/10.1002/mbo3.523>

- Clesceri, L. S., Greenberg, A. E., & Eaton, A. D. (1998). *Standard Methods for the Examination of Water and Wastewater* (20th ed.). Washington DC: APHA American Public Health Association.
- Coakley, T., Brion, G. M., & Fryar, A. E. (2015). Prevalence of and Relationship between Two Human-Associated DNA Biomarkers for Bacteroidales in an Urban Watershed. *Journal of Environment Quality*, 44, 1694. <https://doi.org/10.2134/jeq2014.11.0494>
- Cocolin, L., Mataragas, M., Bourdichon, F., Doulgeraki, A., Pilet, M.-F., Jagadeesan, B., & Phister, T. (2018) Next Generation Microbiological Risk Assessment Meta-Omics: The Next Need for Integration. *International Journal of Food Microbiology*, 287, 10-17. <https://doi.org/10.1016/j.ijfoodmicro.2017.11.008>
- Douterelo, I., Calero-Preciado, C., Soria-Carrasco, V., & Boxall, J. B. (2018). Whole Metagenome Sequencing of Chlorinated Drinking Water Distribution Systems. *Environmental Science: Water Research & Technology*, 4, 2080-2091. <https://doi.org/10.1039/C8EW00395E>
- Environmental Protection Agency (2006). *Method 1600: Enterococci in Water by Membrane Filtration Using Membrane-Enterococcus Indoxyl- β -D-Glucoside Agar (mEI)*. EPA-821-R-06-009
- Environmental Protection Agency (2012). *EPA 5.11 Fecal Bacteria*. <https://archive.epa.gov/water/archive/web/html/vms511.html>
- Environmental Protection Agency (2014). *Overview of Technical Support Materials: A Guide to the Site-Specific Alternative Recreational Criteria TSM Documents*. EPA-820-R-14-010. <https://www.epa.gov/wqc/2012-recreational-water-quality-criteria-documentsEPA-820-R-14-010>
- Environmental Protection Agency (2014). *Site-Specific Alternative Recreational Criteria Technical Support Materials for Alternative Indicators and Methods*. EPA-820-R-14-011. <https://www.epa.gov/wqc/2012-recreational-water-quality-criteria-documentsEPA-820-R-14-011>
- Esiobu, N., Green, M., Echeverry, A., Bonilla, T. D., Stinson, C. M., Hartz, A., & McCorquodale, D. S. (2013). High Numbers of *Staphylococcus aureus* at Three Bathing Beaches in South Florida. *International Journal of Environmental Health Research*, 23, 46-57. <https://doi.org/10.1080/09603123.2012.699027>
- Esiobu, N., Mohammed, R., Echeverry, A., Green, M., Bonilla, T., Hartz, A., McCorquodale, D., & Rogerson, A. (2004). The Application of Peptide Nucleic Acid Probes for Rapid Detection and Enumeration of Eubacteria, *Staphylococcus aureus* and *Pseudomonas aeruginosa* in Recreational Beaches of S. Florida. *Journal of Microbiological Methods*, 57, 157-162. <https://doi.org/10.1016/j.mimet.2003.12.005>
- Evangelista, K. V., & Coburn, J. (2010). *Leptospira* as an Emerging Pathogen: A Review of Its Biology, Pathogenesis and Host Immune Responses. *Future Microbiology*, 5, 1413-1425. <https://doi.org/10.2217/fmb.10.102>
- Ferguson, D. M., Griffith, J. F., McGee, C. D., Weisberg, S. B., & Hagedorn, C. (2013). Comparison of *Enterococcus* Species Diversity in Marine Water and Wastewater Using Enterolert and EPA Method 1600. *Journal of Environmental and Public Health*, 2013, Article ID: 848049. <https://doi.org/10.1155/2013/848049>
- Figueras, M. J., & Borrego, J. J. (2010). New Perspectives in Monitoring Drinking Water Microbial Quality. *International Journal of Environmental Research and Public Health*, 7, 4179-4202. <https://doi.org/10.3390/ijerph7124179>
- Gomez-Alvarez, V., Revetta, R. P., & Santo, D. J. W. (2012). Metagenomic Analysis of Drinking Water Receiving Different Disinfection Treatments. *Applied and Environ-*

- mental Microbiology*, 78, 6095-6102. <https://doi.org/10.1128/AEM.01018-12>
- Hartz, A., Cuvelier, M., Nowosielski, K., Bonilla, T. D., Green, M., Esiobu, N., & Rogerson, A. (2008). Survival Potential of and *Enterococci* in Subtropical Beach Sand: Implications for Water Quality Managers. *Journal of Environment Quality*, 37, 898. <https://doi.org/10.2134/jeq2007.0312>
- Li, B., Saingam, P., Ishii, S., & Yan, T. (2018). Multiplex PCR Coupled with Direct Amplicon Sequencing for Simultaneous Detection of Numerous Waterborne Pathogens. *Applied Microbiology and Biotechnology*, 103, 953-961. <https://doi.org/10.1007/s00253-018-9498-z>
- Mohiuddin, M. M., Salama, Y., Schellhorn, H. E., & Golding, G. B. (2017). Shotgun Metagenomic Sequencing Reveals Freshwater Beach Sands as Reservoir of Bacterial Pathogens. *Water Research*, 115, 360-369. <https://doi.org/10.1016/j.watres.2017.02.057>
- Noble, R., Moore, D., Leecaster, M., Mcgee, C., & Weisberg, S. (2003). Comparison of Total Coliform, *Fecal Coliform*, and *Enterococcus* Bacterial Indicator Response for Ocean Recreational Water Quality Testing. *Water Research*, 37, 1637-1643. [https://doi.org/10.1016/S0043-1354\(02\)00496-7](https://doi.org/10.1016/S0043-1354(02)00496-7)
- Odonkor, S. T., & Ampofo, J. K. (2013). *Escherichia coli* as an Indicator of Bacteriological Quality of Water: An Overview. *Microbiology Research*, 4, 2. <https://doi.org/10.4081/mr.2013.e2>
- Office of Water, Environmental Protection Agency (2010). *Quantitative Microbial Risk Assessment to Estimate Illness in Freshwater Impacted by Agricultural Animal Sources of Fecal Contamination*. EPA 822-R-10-005.
- Rashkin, S., Jun, G., Chen, S., & Abecasis, G. R. (2017). Optimal Sequencing Strategies for Identifying Disease-Associated Singletons. *PLOS Genetics*, 13, e1006811. <https://doi.org/10.1371/journal.pgen.1006811>
- Rochelle-Newall, E. E., Nguyen, T. M., Le, T. P., Sengtaheuanghoung, O., & Riboizi, O. (2015). A Short Review of Fecal Indicator Bacteria in Tropical Aquatic Ecosystems: Knowledge Gaps and Future Directions. *Frontiers in Microbiology*, 6, 308. <https://doi.org/10.3389/fmicb.2015.00308>
- Roy, M., Arnaud, J., Jasmin, P., Hamner, S., Hasan, N., Colwell, R., & Ford, T. (2018). A Metagenomic Approach to Evaluating Surface Water Quality in Haiti. *International Journal of Environmental Research and Public Health*, 15, 2211. <https://doi.org/10.3390/ijerph15102211>
- Tan, B. F., Charmaine, N., Jean, P. N., Lay, L. L., Karina, G. Y.-H., & Janelle, T. R. (2015). Next-Generation Sequencing (NGS) for Assessment of Microbial Water Quality: Current Progress, Challenges, and Future Opportunities. *Frontiers in Microbiology*, 6, 1027. <https://doi.org/10.3389/fmicb.2015.01027>
- Wade, T. J., Calderon, R. L., Sams, E., Beach, M., Brenner, K. P., Williams, A. H., & Dufour, A. P. (2006). Rapidly Measured Indicators of Recreational Water Quality Are Predictive of Swimming-Associated Gastrointestinal Illness. *Environmental Health Perspectives*, 114, 24-28. <https://doi.org/10.1289/ehp.8273>
- Water Quality Standards Variances. *Epa.gov*. <https://www.epa.gov/sites/production/files/2019-10/documents/flwqs-variances>