A Customized DNA Microarray for Microbial Source Tracking in Environmental Systems

Xiang Li[1, 2], Valerie J. Harwood[3], Bina Nayak[3], Christopher Staley[4], Michael J. Sadowsky[5] and Jennifer Weidhaas[2]

• [1]U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, 26 West Martin Luther King Drive, Cincinnati, OH 45268, USA • [2]Department of Civil and Environmental Engineering, West Virginia University, P.O. Box 6103, Morgantown, West Virginia 26506, USA

• [3]Department of Integrative Biology, University of South Florida, Tampa, Florida 33620, United States • [4]BioTechnology Institute, University of Minnesota, St. Paul, Minnesota 55108, United States • [5]Department of Soil, Water, and Climate, BioTechnology Institute, University of Minnesota, St. Paul, Minnesota 55108, United States

Contact information

Xiang Li, Ph.D
ORISE Post-Doctoral Fellow
ORD/NRMRL/WSWRD
Tel: 513-569-7567
Email: Li.Xiang@epa.gov

Introduction

The environmental waters in the United States are impaired due to the presence of waterborne pathogens. [1] These pathogens typically originate from human and other animal fecal pollution sources. These impaired waters are usually evaluated using analyses of fecal indicator bacteria (FIB), [2] but the presence of FIB does not necessarily indicate that pathogens are present in the environmental system. Therefore, various microbial source tracking (MST) methods have been proposed for identification of different fecal pollution sources in environmental systems. [3] However, these MST methods suffer from low throughput as only a small number of targets can be tested at one time. Microarrays can overcome this issue by simultaneously testing hundreds to thousands of targets. As a result, MST microarrays are of increasing interest in the research and regulatory realms. [4]

Objectives

- Design and validate an MST microarray for detecting fecal contamination and pathogens in environmental samples, while simultaneously indicating the source(s) of fecal pollution.
- Evaluate the correlation between microarray measurements and other more established methods, namely qPCR and culture based methods.
- Next generation sequencing to characterize the relative coverage of the fecal microbiome represented by the novel microarray approach.

Materials and Methods

Customized MST microarray design

Probe selection principle:

- Previously published and validated microarray probes.
- Previously published qPCR probes and/or PCR primers that could be lengthened or shortened to 60 mers without cross hybridizing with known environmental microbes.
- Probes designed CommOligo 2.0 [5], genes from select fecal microorganisms.

Probes include:

- Common waterborne pathogens including bacteria, eukaryotes and viruses
- Previously published MST marker genes and organisms
- Antibiotic resistance genes,
- Fecal indicator bacteria
- mtDNA
- Universal bacterial

Sample handling, experimental process and data analysis (Figure 1).

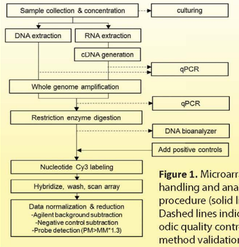

Figure 1. Microarray sample handling and analysis procedure (solid lines). Dashed lines indicate periodic quality control and method validation testing.

Results

- Eight of nine genes or organisms in the spiked positive control sample were detected in 22 individual probes, for an 89% detection rate.
- There were a total of 70 and 132 probes detected in the raw sewage on microarray 1 and 2, respectively. Similarly, 73 and 111 probes were detected in the cattle feces, 68 and 110 probes detected in the swine feces, 104 and 130 probes detected in the poultry litter on microarray 1 and 2, respectively. Ninety-one probes were detected from the mixed avian fecal sample, and 100 probes were detected in the spinach sample dipped in raw sewage on microarray 1. Eight different viruses and 25 different pathogens were detected in the sewage, feces, and poultry litter samples.
- Microarray detections decrease with serial dilutions of raw sewage samples (**Figure 2**).
- Semi-quantitative ability of microarray (**Figure 3**).
- Fecal source determination via microarray detections and nonmetric multidimensional scaling (NMDS) analysis (**Figure 4**).
- Next generation sequencing (NGS) collaborates with microarray (**Figure 5**).

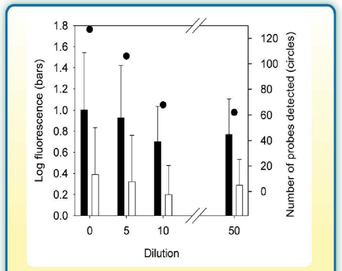

Figure 2. Mean ± SD of relative fluorescence of perfect match (■) and mismatch (□) probes in a dilution series of raw wastewater (0, 5, 10, and 50 to 1 dilution) and number of probes detected (●) on each microarray.

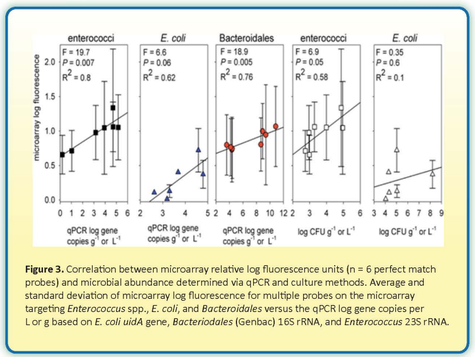

Figure 3. Correlation between microarray relative log fluorescence units (n = 6 perfect match probes) and microbial abundance determined via qPCR and culture methods. Average and standard deviation of microarray log fluorescence for multiple probes on the microarray targeting *Enterococcus* spp., *E. coli*, and *Bacteroidales* versus the qPCR log gene copies per L or g based on *E. coli uidA* gene, *Bacteriodales* (Genbac) 16S rRNA, and *Enterococcus* 23S rRNA.

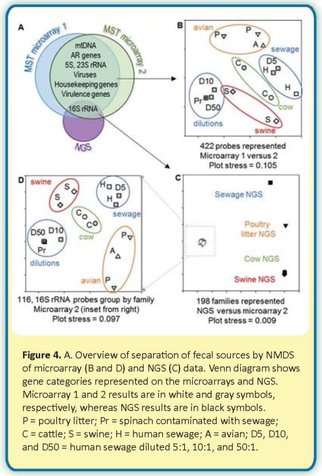

Figure 4. A. Overview of separation of fecal sources by NMDS of microarray (B and D) and NGS (C) data. Venn diagram shows gene categories represented on the microarrays and NGS. Microarray 1 and 2 results are in white and gray symbols, respectively, whereas NGS results are in black symbols. P = poultry litter; Pr = spinach contaminated with sewage; C = cattle; S = swine; H = human sewage; A = avian; D5, D10, and D50 = human sewage diluted 5:1, 10:1, and 50:1.

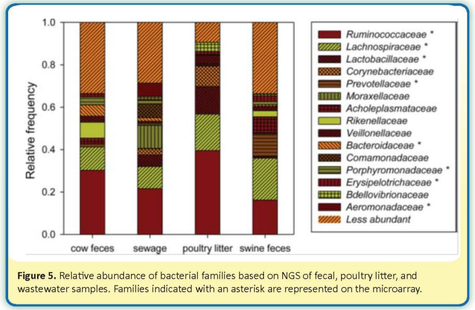

Figure 5. Relative abundance of bacterial families based on NGS of fecal, poultry litter, and wastewater samples. Families indicated with an asterisk are represented on the microarray.

Conclusions

- Common waterborne pathogens (bacterial, eukaryotic and viral) can be detected via microarrays in fecal samples, raw sewage, poultry litter and contaminated produce.
- Microarray data may be correlated with qPCR and culture based enumeration of FIB.
- Microarray data reproducibly separated fecal material from different sources into different clusters.
- The current MST microarray contains probes accounting for four of the five most dominant phyla and eight of the 15 most dominant families found in fecal samples as determined via NGS.

References

1. USEPA Water Quality Reporting Database (ATTAINS). In 2015.
2. Hlavsa, M. C.; Roberts, V. A.; Kahler, A. M.; Hilborn, E. D.; Wade, T. J.; Backer, L. C.; Yoder, J. S., Recreational water-associated disease outbreaks--United States, 2009-2010. *MMWR. Morbidity and mortality weekly report* **2014,** *63*, (1), 6-10.
3. Harwood, V. J.; Staley, C.; Badgley, B. D.; Borges, K.; Korajkic, A., Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS microbiology reviews* **2014,** *38*, (1), 1-40.
4. Berthet, N.; Dickinson, P.; Filliol, I.; Reinhardt, A. K.; Batejat, C.; Vallaeys, T.; Kong, K. A.; Davies, C.; Lee, W.; Zhang, S.; Turpaz, Y.; Heym, B.; Coralie, G.; Dacheux, L.; Burguiere, A. M.; Bourhy, H.; Old, I. G.; Manuguerra, J. C.; Cole, S. T.; Kennedy, G. C., Massively parallel pathogen identification using high-density microarrays. *Microbial biotechnology* **2008,** *1*, (1), 79-86.
5. He, Z.; Gentry, T. J.; Schadt, C. W.; Wu, L.; Liebich, J.; Chong, S. C.; Huang, Z.; Wu, W.; Gu, B.; Jardine, P.; Criddle, C.; Zhou, J., GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *The ISME journal* **2007,** *1*, (1), 67-77.