

SCCWRP #0771

## **Evaluation of molecular community analysis methods for discerning fecal sources and human waste**

Yiping Cao<sup>1</sup>, Laurie C. Van De Werfhorst<sup>2,3</sup>, Eric A. Dubinsky<sup>4</sup>, Brian D. Badgley<sup>5</sup>, Michael J. Sadowsky<sup>5</sup>, Gary L. Andersen<sup>4</sup>, John F. Griffith<sup>1</sup> and Patricia A. Holden<sup>2,3</sup>

<sup>1</sup>*Southern California Coastal Water Research Project, Costa Mesa, CA*

<sup>2</sup>*Bren School of Environmental Science & Management, University of California, Santa Barbara, CA*

<sup>3</sup>*Earth Research Institute, University of California, Santa Barbara, CA*

<sup>4</sup>*Lawrence Berkeley National Laboratory, Earth Sciences Division, Berkeley, CA*

<sup>5</sup>*BioTechnology Institute, University of Minnesota, St. Paul, MN*

### **ABSTRACT**

Molecular microbial community analyses provide information on thousands of microorganisms simultaneously, and integrate biotic and abiotic perturbations caused by fecal contamination entering water bodies. A few studies have explored community methods as emerging approaches for microbial source tracking (MST), however, an evaluation of the current state of this approach is lacking. Here, we utilized three types of community-based methods with 64 blind, single- or dual-source, challenge samples generated from 12 sources, including: humans (feces), sewage, septage, dogs, pigs, deer, horses, cows, chickens, gulls, pigeons, and geese. Each source was a composite from multiple donors from four representative geographical regions in California. Methods evaluated included terminal restriction fragment polymorphism (TRFLP), phylogenetic microarray (PhyloChip), and next generation (Illumina) sequencing. These methods correctly identified dominant (or sole) sources in over 90% of the challenge samples, and exhibited excellent specificity regardless of source, rarely detecting a source that was not present in the challenge sample. Sensitivity, however, varied with source and community analysis method. All three methods distinguished septage from human feces and sewage, and identified deer and horse with 100% sensitivity and 100% specificity. Method performance improved if the composition of blind dual-source reference samples were defined by DNA contribution of each single source within the mixture, instead of by *Enterococcus* colony forming units. Data analysis approach also influenced method performance, indicating the need to standardize data interpretation. Overall, results of this study indicate that community analysis methods hold great promise as they may be used to identify any source, and they are particularly useful for sources that currently do not have, and may never have, a source-specific single marker gene.

**Due to distribution restrictions, the full-text version of this article is available by request only.**

Please contact [pubrequest@sccwrp.org](mailto:pubrequest@sccwrp.org) to request a copy.