

Drinking Water Microbiome as a Screening Tool for Nitrification in Chloraminated Drinking Water Distribution Systems

V. Gomez-Alvarez and R. P. Revetta

U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH
45268

Keyword: Biomarkers, machine learning; microbial communities; drinking water distribution system; nitrification

ABSTRACT

Many water utilities in the US using chloramine as disinfectant treatment in their distribution systems have experienced nitrification episodes, which detrimentally impact the water quality. Here, we used 16S rRNA sequencing data to generate high-resolution taxonomic profiles of the bulk water (BW) microbiome from a chloraminated drinking water distribution system (DWDS) simulator. The DWDS was operated through four successive operational schemes, including two stable events (SS) and an episode of nitrification (SF), followed by a ‘chlorine burn’ (SR) by switching disinfectant from chloramine to free chlorine. Specifically, this study focuses on biomarker discovery and their potential use to classify SF episodes. Principal coordinate analysis identified two major clusters (SS and SF; PERMANOVA, $p < 0.0001$) consistent with the effect of disturbance in the relative abundances of the core microbiome. The linear discriminant analysis (LDA) effect size algorithm (LEfSe) revealed both an enrichment and depletion of various bacterial populations associated with episodes of SS and SF. A supervised machine learning approach (naïve Bayes classifier) was used to classify samples from SS and SF. Classification models were trained with different types of biological and chemical datasets, including biomass (ATP), OTUs, genus-level taxonomic groups, and water quality (NH_2Cl and Free- NH_3). Performance of each model was examined using the area under the curve (AUC) from the receiver-operating characteristic (ROC) and precision-recall (PR) curves (AUC of 1.0 indicates an excellent classifier). AUC using biomass data were determined to be 0.596, which is equivalent to a random classification of the samples. The AUC gradually increased to 0.663 when genus-level taxonomic

membership data were used in the classification model and increased significantly using OTU-level membership (0.884). Combining membership with distribution (i.e. community structure) significantly improved the predictive ability of the OTU and genus-level taxonomic model beyond that of membership only (AUC >0.976, $p < 0.01$). Furthermore, models were able to correctly predict 95% (AUC = 0.983, $n = 104$) and 96% (AUC = 0.973, $n = 72$) of samples of the DWDS (community structure of two published studies) and water quality datasets, respectively. The results from this study demonstrate the feasibility of selected BW microbiome signatures as predictive biomarkers of nitrification in DWDS. This new information can be used to optimize current nitrification monitoring plans.