CFID Safe Drinking Water Pilot Project Grant 2015

Summary Report: **Metagenomic analysis of drinking well water for rural and marginalized populations in regional Ontario – implications for testing, including human, public, and environmental health**

Principal Investigator: Dr. Anna Majury PhD DVM (AMMI member)

Co-Investigators: Dr. Gerald Evans MD (AMMI member), Dr. Paul Belanger PhD, Dr. Xudong Liu PhD

Abstract

Given that water is essential to sustaining life, contamination of drinking water sources poses a serious potential threat to human, animal, and environmental health. Private well water supplies, which are the primary source of drinking water for the majority of rural Ontarians, are at risk of becoming microbially contaminated, which is of utmost concern given water remains a significant pathway of infectious disease transmission, both locally and globally. Despite over a century's worth of accumulated knowledge, there remain several gaps in our understanding of the microbiology of well and groundwater, such as the occurrence or abundance of pathogens and their impact on aquifer microbial communities. Scientists estimate that a limited fraction - only about 1% of all microbes, have the capacity to be cultured due to their complex symbiosis with other microorganisms. Traditional indicators of water health and quality, particularly the limited scope of traditional culture and PCR methods, are only successful for a small portion of contaminating microbes and limit study of the diversity of groundwater health indicators. Our prior research, as well that of others, indicates enormous ground and well water microbial diversity in need of further investigation. Genomic studies of drinking well water are limited, although warranted given the public health significance as a consequence of exposure to this water source. Next Generation Sequencing (NGS) technology bypasses the need for isolation or cultivation of microorganisms, providing insight into the environmental metagenome; this technology promises to give rise to a grander understanding of microbial communities and pathogens inhabiting drinking water environments.

Aims & Objectives

The aim of this study is to pilot metagenomic analyses of bacterial communities in drinking well waters of rural and regional southern Ontario.

Prior research has identified three regions in southern Ontario of increased relative risk of *E. coli* contamination in private wells. The regions identified were Kingston/Belleville, Hamilton/Niagara, and London/Bruce, as depicted in Figure 1 (Krolik *et al.*, 2013). By applying metagenomic sequencing and statistical analyses, the objective is to investigate the bacterial diversity and richness in the drinking well water microbiome and investigate the potential change in microbial signature across the different regions of southern Ontario.

A trial of 40 samples was analyzed by metagenomic sequencing and bioinformatics at the Queen's Genomics Laboratory at Ongwanada (QGLO) to explore richness and diversity of the bacterial community in the samples, Ion Personal Genome Machine (PGM™) System (Ion Torrent, 4462921) was used. For the trial samples, a targeted metagenomic approach was adopted, using Ion 16S™ Metagenomics Kit, using 16S Primer Set V3-6, 7-9 targeting 4 hypervariable regions of the 16S rRNA gene.

Findings to Date

The trial consisted of 40 private well water samples, with 16S Primer Set V3-6, 7-9, targeting 4 hypervariable regions of the 16S rRNA gene. Metagenomic sequencing and bioinformatic analyses were conducted at QGLO.

Based on the trial data, members of the phylum *Proteobacteria* evidently are the most abundant bacterial group of southern Ontario drinking well water microbial communities (Figure 3), correlating with previous associated studies. Culture-independent methods suggest that the phylum *Proteobacteria* is the most abundant bacterial group in surface and drinking water (Eichler *et al.*, 2006; Gomez-Alvarez *et al.*, 2012; Hoefel *et al.*, 2005; Kahlisch *et al.*, 2012; Kormas *et al.*, 2010; Vaz-Moreira *et al.*, 2013; Williams *et al.*, 2004), even chlorinated drinking water (Eichler *et al.*, 2006; Gomez-Alvarez *et al.*, 2012; Hoefel *et al.*, 2005; Kormas *et al.*, 2010; Poitelon *et al.*, 2009; Vaz-Moreira *et al.*, 2013; Revetta *et al.*, 2010), and groundwater (Hemme *et al.*, 2010; Smith *et al.*, 2012).

Upon further analysis at the family, genus and species levels, the 16S sequence becomes more conserved between closely related bacteria, resulting in the inability to determine the origin of that particular read. With the 16S Primer Set V3-6, 7-9 targeting 4 hypervariable regions, the *Enterobacteriaceae* family of interest was identified.

Next Steps

Over 700 water samples were collected during the designated sample period from the three predetermined regions in southern Ontario of increased relative risk of *E. coli* contamination. A total of 76 samples, all TC positive with an *E.coli* CFU count of 10 to >81, have been selected for further metagenomic analyses (Figure 5). Both sets of primers will be used to amplify the corresponding hypervariable regions of the 16S rDNA gene in bacteria of the final sequenced samples: Primer set V2-4-8 and Primer set V3-6,7-9, allowing 7 of 9 hypervariable regions (V2-3-4-6-7-8-9) of the 16S rRNA to be amplified.