# Microbial source tracking of private well water samples across at-risk regions in southern Ontario and analysis of traditional fecal indicator bacteria assays including culture and qPCR

Julia Krolik, Allison Maier, Shawna Thompson and Anna Majury

## ABSTRACT

Many people living in rural areas rely on privately owned wells as their primary source of drinking water. These water sources are at risk for fecal contamination of human, wildlife, and livestock origin. While traditional bacteriological testing involves culture-based methods, microbial source tracking (MST) assays present an opportunity to additionally determine the source of fecal contamination. This study investigated the main host sources of contamination in private well water samples with high levels of *Escherichia coli* (*E. coli*), using MST with human and multi-species specific markers. Fecal contamination of human origin was detected in approximately 50% of samples, indicating that current contamination prevention strategies require reconsideration. The relationship between cattle density and fecal contamination of bovine origin was investigated using a Bovine Bacteroidales specific MST assay. Regional variations of microbial sources were examined, and may inform local primary prevention strategies. Additionally, in order to assess MST and *E. coli* quantitative real time polymerase chain reaction (qPCR) assays as indicators of fecal contamination, these were compared to *E. coli* culture methods. Variation in results was observed across all assay methods investigated, suggesting the most appropriate routine bacteriological testing methodology cannot be determined without comparison to a method that directly detects the presence of fecal contamination.

**Key words** | fecal indicator bacteria, microbial source tracking, PCR, private water supplies, public health

**Julia Krolik**
**Allison Maier**
**Anna Majury** (corresponding author)
Public Health Ontario,
181 Barrie Street, P.O. Box 240,
Kingston,
Ontario K7L 3K2,
Canada
E-mail: *anna.majury@oahpp.ca*

**Shawna Thompson**
**Anna Majury**
Department of Biomedical and Molecular Sciences,
Botterell Hall, Queen's University,
18 Stuart Street,
Kingston,
Ontario K7L 3N6,
Canada

**Anna Majury**
Department of Public Health Sciences,
Carruthers Hall, Queen's University,
Kingston,
Ontario K7L 2N8,
Canada

## INTRODUCTION

Globally, millions of people in developing and affluent nations rely on private groundwater wells as their primary drinking water source. It is well established that aquifers supplying these wells can be become contaminated with human and/or animal feces, which may contain pathogens leading to serious human health risks such as gastrointestinal (GI) illness (Charrois 2010). Microbial source tracking (MST) methods can be used as delineators of fecal contamination through attribution to specific host species, potentially identifying the origin of the contamination and thereby allowing for targeted prevention and remediation strategies. Traditionally, MST methods have been used to evaluate fecal pollution of localized surface waters, including: storm, stream and recreational water sources (Lee *et al.* 2010; Sauer *et al.* 2011; Haack *et al.* 2013; Krentz *et al.* 2013). However, to date, few studies have utilized these techniques in groundwater derived drinking water, particularly from privately owned wells.

At this time, it is not feasible to directly detect pathogens from water samples due to their low levels, sporadic and

doi: 10.2166/wh.2016.024

erratic presence, and the complex methodologies required to isolate them (Cabral 2010). Traditionally, MST methods are not used as routine measures to indicate fecal contamination. Rather, the current standard of practice uses fecal indicator bacteria (FIB) as markers of fecal contamination when testing environmental water supplies for potability, given that normal intestinal inhabitants are typically present alongside certain pathogenic species (Leclerc *et al.* 2001). A recent systematic review by Cabral (2010) defined six (three major and three minor) ideal criteria for a good bacterial fecal indicator. The first three require that the FIB have a high concentration in feces, are generally non-pathogenic to humans, and are amenable to cost-effective and timely detection techniques. In addition, the indicator organism should not reproduce external to the intestinal tract, be present in higher quantities than associated pathogens in water and possess a similar decay rate as the pathogens. Utilizing these criteria, previous investigations have determined that, of all intestinal flora, *Escherichia coli* (*E. coli*) remains the ideal, reliable fecal pollution indicator for environmental waters (Leclerc *et al.* 2001).

While utilizing *E. coli* as FIB is the current best practice, weaknesses have been recently identified. For example, novel evidence documents environmental *E. coli* populations that originated from animal intestines and adapted to proliferate in environmental conditions (Perchec-Merien & Lewis 2012). Moreover, *E. coli* has been documented to enter into a viable but non-culturable state, which may yield false negative results during water testing (Liu *et al.* 2008; Lothigus *et al.* 2009). In general, the use of FIB does not reveal possible sources of fecal contamination and the use of MST methods may provide insight into these potential contributing host species. Additionally, the use of MST assays alongside traditional FIB will allow for the evaluation of MST as a potential indicator of fecal contamination.

In Ontario, three distinct areas with a higher burden of *E. coli* contamination among private wells have been identified (Krolik *et al.* 2013). In one of these regions, the possible origins of fecal contamination were investigated using MST methods (Krolik *et al.* 2014). Half of the samples had a positive signal for human markers, suggesting septic tanks as the potential origin. In this study, two additional areas with higher *E. coli* contamination rates were investigated, as well as a region with high cattle density. Traditional culture methods, MST assays using quantitative real time polymerase chain reaction (qPCR) and *E. coli* qPCR were performed. The specific aims were to: (1a) determine the predominant host-specific origins of contamination among private well water samples from the remaining two regions where burden of *E. coli* contamination is greater, (1b) determine whether a higher cattle density region exhibits higher levels of bovine contamination in private well water samples, (1c) look for general regional variations in predominant primary host contamination; and (2) evaluate, through comparison, MST and *E. coli* qPCR assays alongside *E. coli* culture methods, as indicators of fecal contamination.

## MATERIALS AND METHODS

### Study area and sample sets

Multiple convenience samples of private water supplies (predominantly private wells), collected over multiple years and regions, were utilized in this investigation. The first group was comprised of waters submitted for bacteriological analysis to the Public Health Ontario Laboratory in 2012 from Region A as described in Figure 1, and which tested positive for *E. coli*. The 2013 group consisted of water samples submitted during the summer period from Region A, and which tested negative for *E. coli*. The final sample set, from 2014, included submissions originating from the other two regions with higher *E. coli* contamination (Figure 1); namely Regions B and C (Region A1 was investigated previously). Additionally, samples from an area with high cattle density were included (Region D) (Agriculture Survey Stats Canada 2011).

### Sample processing and nucleic acid extraction

For bacteriological analysis, 100 mL of sample water was filtered using a 0.45 μm pore size mixed cellulose esters filter (Millipore Billerica, MA) using a partial vacuum. *E. coli* colony forming units (CFUs) were enumerated from each filter, after 24 h ± 2 h at 35 °C incubation on differential coliform (DC) media (Oxoid, ON) with 5-bromo-4-chloro-3-indolyl-D-glucuronide (BCIG). Colonies are colored blue

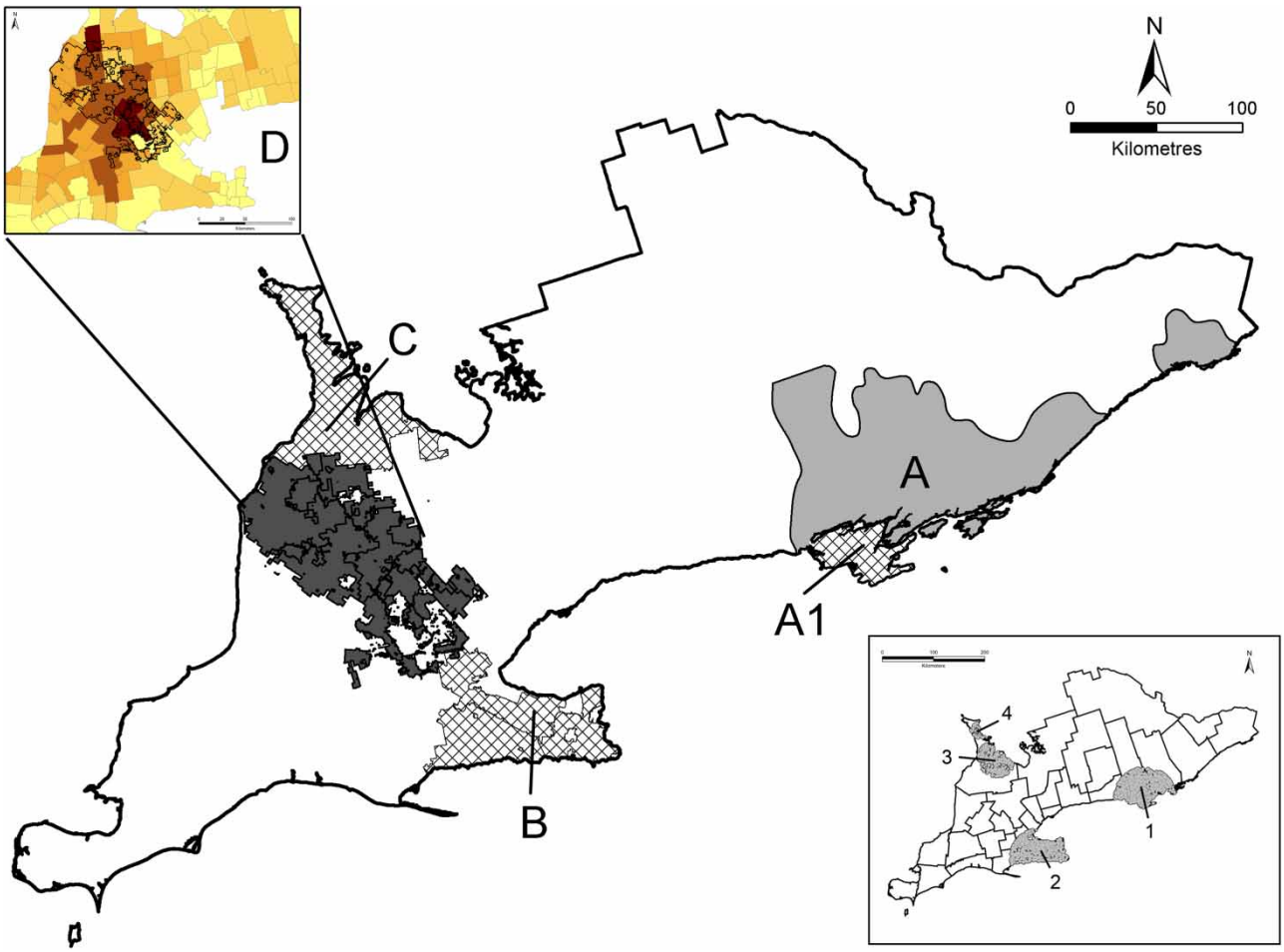


**Figure 1** | Map of sample collection sites: A. Submission catchment for 2012 and 2013 samples; A1. Subset of A that was an area of high *E. coli* contamination. B and C. Areas of high *E. coli* contamination. D. High cattle density region. Map of southern Ontario provided in bottom right demonstrates original regions of elevated risk of *E. coli* contamination in private water submissions (Krolik *et al.* 2013), which were used to define collection sites A1, B and C.

when BCIG is released into the medium via glucuronidase activity, as it is insoluble, accumulating within the cell; these cells are considered positive for *E. coli* (Ogden & Watt 1991). Additionally, total coliforms (TC), which include environmental and gut microbiome bacteria, appear pink on DC media (Oxoid, ON). *E. coli* and TC culture results are counted up to 80 CFU/100 mL, after which they are designated overgrown. Additionally, below or equal to 15 CFU/100 mL the method is considered not quantifiable (ISO 2000).

A second 100 mL of water was filtered through a mixed cellulose esters filter (Millipore Billerica, MA), rolled and placed in 2 mL of Lysis Buffer (bioMérieux, QC) for 1 h at 37 °C on rotation. The DNA was extracted using a magnetic silica bead system (bioMérieux, QC) and eluted to either a 60 μL of a 100 μL volume. The difference in final volumes was accounted for when results were expressed in cells per 100 mL. Several collection sites performed membrane filtration, bacteriological analysis and DNA extraction occurred within 24 hours of collection. DNA samples were frozen at −80 °C and shipped on dry ice to a single site for all qPCR analysis.

### qPCR assays – Bacteroidales and *E. coli*

The MST assay used in this study was previously described by Lee *et al.* (2010) for surface water, and subsequently applied to private well water by Krolik *et al.* (2014). Briefly, the MST

assay is comprised of three singleplex assays: a human specific Bacteroidales assay, a bovine specific Bacteroidales assay, and a general Bacteroidales assay, all of which target the 16S rRNA gene of Bacteroidales and are, henceforth, referred to as BacHuman, BacBovine and BacGeneral, respectively. The BacGeneral assay is designed to detect this gene in a broad array of hosts (human, cow, pig, deer, horse, dog, cat, gull, goose and raccoon). Assay primer and Taqman probes, as well as reaction conditions, have been previously described (Lee *et al.* 2010; Krolik *et al.* 2014). It has been shown that an adjusted limit of detection (LOD) of 10 gene copies per reaction is appropriate (Krolik *et al.* 2014). The EC23S857 *E. coli* assay was previously described by Chern *et al.* (2011) and targets the 23S rRNA gene region. It is a Taqman assay with primer and probe concentrations of 1 μmol/L and 80 nmol/L, respectively, and 5 μL of DNA, for a total volume of 25 μL. An initial incubation at 50 °C for 2 min followed by 95 °C for 10 min was followed by 40 PCR cycles of 95 °C for 15 s and 60 °C for 60 s. All assays were performed using Environmental Mix 2.0 Master Mix on the ViiA7 platform (Thermofischer Scientific, ON). The master mix was chosen to accommodate potential presence of inhibitors and is designed to analyze environmental samples (Thermofischer Scientific, ON).

Standard curves for MST assays were created as previously described (Krolik *et al.* 2014) using plasmids with concentrations ranging from 6.5 million to 6.5 gene copies per reaction. The *E. coli* standards were prepared using genomic DNA from ATCC (American Type Culture Collection) strain 25922 and ranged from 1 million to 1 cell per reaction. Subsequently *E. coli* standards, to a limit of 1 gene copy per reaction, were run independently.

## Data manipulation and statistical analyses

The seven datasets (spanning the three sample collection years) that were used to investigate assay relationships are shown in Table 1.

Based on which species specific markers were detected (for BacGeneral, BacBovine and BacHuman MST assays), each sample was assigned a status – positive for general markers only, positive for human markers only, positive for bovine markers only, positive for both human and bovine markers, or negative for all assays. These statuses were used to tabulate contributing host-distributions for each region. To further illuminate the predominant fecal source, Spearman correlations between host specific markers and general markers were performed for each of the four regions as well as for the A_2012 dataset.

Previous findings (Krolik *et al.* 2014) determined that human sourced contamination was the predominant source for Region A. The MST host abundance results were divided into two categories: human sourced and non-human sourced. In order to investigate geographic differences in fecal origins, a Chi-square test was used to detect significant differences across the four regions. Subsequently, the Benjamin, Hochbery and Yekutieli (BHY) *post-hoc* comparison was used to determine where the differences occurred.

A possible relationship between the regional rates of human sourced fecal contamination and the density of

**Table 1** | Dataset descriptors

| Dataset | CFU/100 mL result | *n* | Region |
|---|---|---|---|
| A_2012 | *E. coli* positive | 716 (MST assay)<br>706 (*E. coli* assay) | Any sample submitted to Region A in 2012 |
| A1_2012 | *E. coli* positive | 102 | Any sample submitted to Region A1 in June–August 2012 (subset of A_2012) |
| A_2013_NEG | Negative | 93 | Any sample submitted to Region A in 2013 |
| A_2013_TC | TC positive | 73 | Any sample submitted to Region A in 2013 |
| B_2014 | *E. coli* positive | 66 | Any sample submitted to Region B in 2014 |
| C_2014 | *E. coli* positive | 129 | Any sample submitted to Region C in 2014 |
| D_2014 | *E. coli* positive | 59 | Any sample submitted to Region D in 2014 |

MST and *E. coli* qPCR analyses were performed on all samples with the exception of A_2012. A_2012 represents submissions from a full year; the rest of the datasets were captured for June–August period.

septic systems, a previously hypothesized source of said human fecal contamination, was explored. Unfortunately, septic density is not available as records are not maintained provincially, and thus, a proxy measure – domestic well density – was utilized. The assumption is that owners of a private well would also require private wastewater disposal, and therefore a 1:1 ratio is reasonable. Using the boundaries developed for sample collection, each of the four regions was assigned a set of Statistics Canada-developed dissemination areas (DAs) (small, relatively stable geographic census units) (Statistics Canada 2011). Well densities were determined using well records from the Well Water Information System (WWIS 2013) and expressed as number of domestic-use wells divided by DA area ($km^2$) using ArcGIS 10.2 (ESRI Inc., Redlands, USA). For each region, a median across all the DAs of wells/$km^2$ was calculated, transformed logarithmically and graphically visualized.

In order to compare the three fecal indicator methods (*E. coli* culture, *E. coli* and MST qPCR), two different statistical analyses were used. The comparison between CFU and PCR tests required categorical analysis given the majority of samples were within the non-quantifiable CFU range. As such, Chi-square tests with odds ratios were utilized, where the break between categorical groups was determined by the quantifiable limits. Quantifiable ranges existed for both qPCR assays; Spearman correlations were used to investigate possible associations as the data did not follow a normal distribution. All statistical analyses were performed in R 3.2.0 (R Core Team, Vienna).

## RESULTS

The MST assays have a natural cut-off of 10 gene copies per reaction as previously published (Krolik *et al.* 2014); this was confirmed with pooled 2014 data (data not shown). The *E. coli* assay has a LOD of 1 cell (7 gene copies) per reaction. Both qPCR assays, for all tested samples, had a PCR efficiency between 90 and 110% and an $R^2$ value above 0.995.

### Host-specific contamination

In this analysis, the previously published 2012 MST dataset was limited to samples from the region of elevated risk for *E. coli* contamination and the summer months, in order to most closely match samples from the other regions (Region A1, $n = 103$ for June–August). Approximately 50% of these samples contained human fecal contamination. The host (human, bovine and general) distributions for this data alongside the other three areas B, C and D are shown in Figure 2.

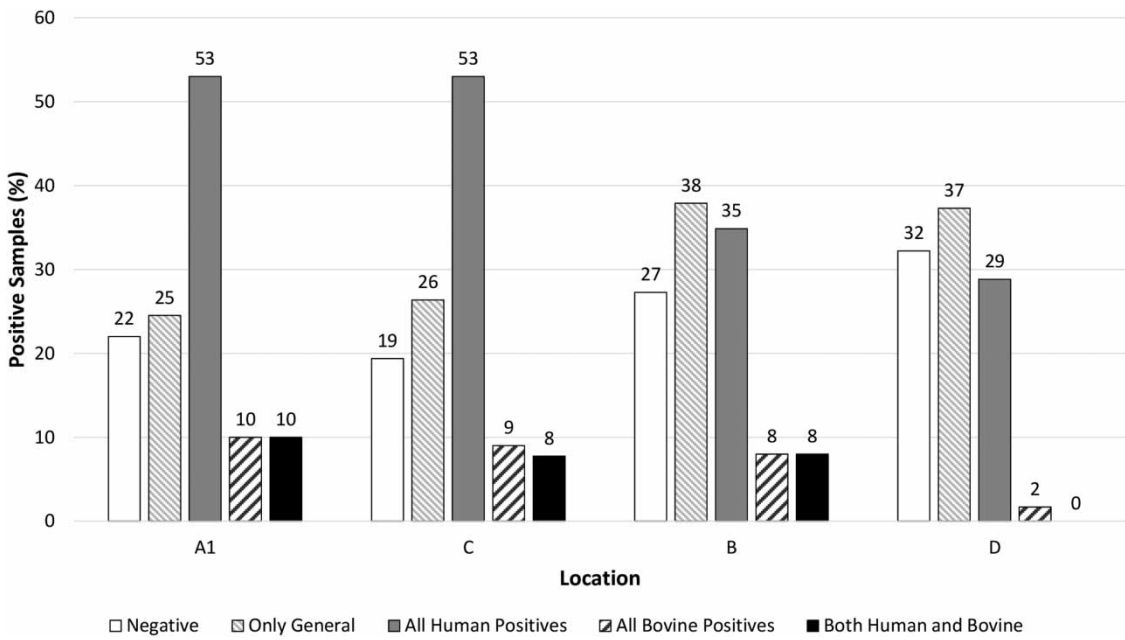


**Figure 2** | Distribution of microbial sources stratified by region. Datasets: A1_2012, C_2014, B_2014, D_2014.

Table 2 shows the number of human and non-human sourced samples (of those that tested positive for BacGeneral) by region. A statistically significant difference between groups was found by the Chi-square test (Chi-square = 12.48, $df = 3$, $p < 0.01$). Given the initial significant result, a *post-hoc* comparison was performed with an adjusted *p*-value alpha of 0.05 (Table 3). Statistically significant differences between Region A1 compared to Region B and D, as well as C compared to B and D were found. As such, Regions A1 and C have statistically significantly higher rates of human sourced fecal contamination.

**Table 2** | Number of samples positive and negative for BacHuman out of all samples positive for BacGeneral for each of the four regions

| MST positives | A1 | C | B | D |
|---|---|---|---|---|
| Human | 57 (70%) | 69 (67%) | 23 (48%) | 17 (44%) |
| Non-human | 25 (30%) | 34 (33%) | 25 (52%) | 22 (56%) |

Datasets: A1_2012, C_2014, B_2014, D_2014.

**Table 3** | Raw and adjusted *p*-values from the BHY *post-hoc* comparisons of regional rates of human sourced fecal contamination

| Comparison | Raw *p*-value | Adjusted *p*-value |
|---|---|---|
| A1 vs C | 0.7526 | 0.8290 |
| A1 vs B | 0.0164 | 0.0328* |
| A1 vs D | 0.0092 | 0.0328* |
| B vs C | 0.0318 | 0.0477* |
| B vs D | 0.8290 | 0.8290 |
| C vs D | 0.0130 | 0.0328* |

*Signifies statistically significant for adjusted *p*-values. Datasets: A1_2012, C_2014, B_2014, D_2014.

The frequency distributions for each region of domestic well density by DA on a logarithmic scale are shown in Figure 3. It demonstrates the data is non-normally distributed and therefore medians per region are appropriate for comparison purposes (Table 4). Both visually, by the frequency distributions, and by comparing medians, it is evident that Regions A1 and C have higher domestic well densities.

Spearman correlations between BacHuman and BacGeneral MST assays, used to further illuminate the predominant source, were found for each of the four areas. The results revealed that there is always a strong relationship between BacHuman and BacGeneral assay results (Table 5). Additionally for the 2012 dataset (where *n* was substantially larger), BacBovine and BacGeneral had a moderate correlation.

## Comparing CFU, *E. coli* and MST qPCR

The MST assay gene copy maxima per 100 mL of water filtered were: 20,367,074, 179,916, and 99,885,456 for BacHuman, BacBovine and BacGeneral assays respectively. The *E. coli* qPCR assay maximum cell number was 18,757 expressed per 100 mL.

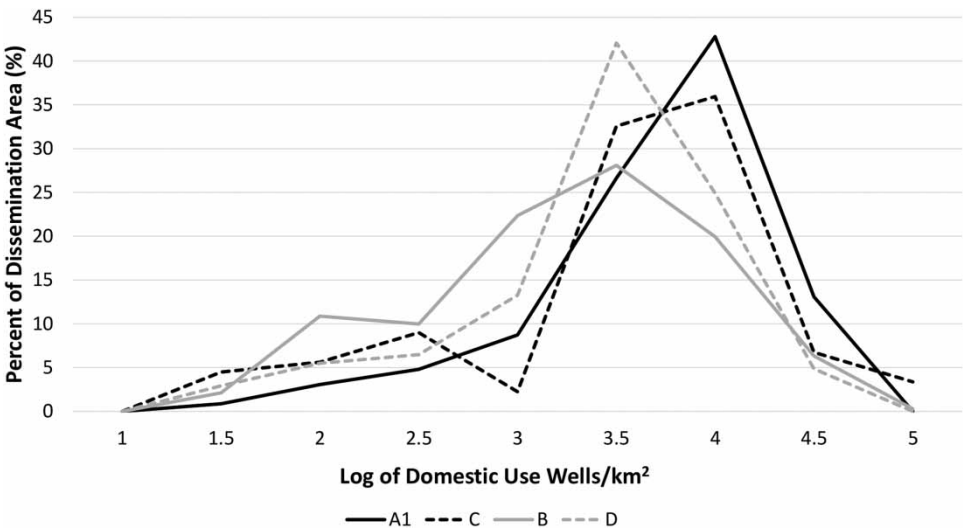


**Figure 3** | Regional frequency distribution of domestic-use well densities by DA expressed as well count per km$^2$ and logarithmically transformed.

**Table 4** | Domestic-use well densities per DA area (km$^2$) expressed as medians for each region

| | Median domestic-use wells/DA area (km$^2$) |
|---|---|
| A1 | 3,905 |
| B | 1,347 |
| C | 2,955 |
| D | 2,100 |

**Table 5** | MST assay correlations

| Dataset | Test | *n* | Spearman's $\rho$ | *p*-value |
|---|---|---|---|---|
| A_2012 | BacHuman vs BacGeneral | 349 | 0.790 | <0.001 |
| A_2012 | BacBovine vs BacGeneral | 88 | 0.509 | <0.001 |
| A1_2012 | BacHuman vs BacGeneral | 57 | 0.823 | <0.001 |
| B_2014 | BacHuman vs BacGeneral | 23 | 0.704 | <0.001 |
| C_2014 | BacHuman vs BacGeneral | 69 | 0.859 | <0.001 |
| D_2014 | BacHuman vs BacGeneral | 17 | 0.914 | <0.001 |

A comparison of *E. coli* qPCR positives between quantifiable and non-quantifiable CFU/100 mL ranges revealed that the probability of obtaining an *E. coli* positive qPCR result are 2.98 (95% CI 1.98–4.58) times greater in the quantifiable range (Table 6) (Chi-square = 26.94, $df = 1$, $p < 0.0001$).

Initial analysis showed that the percentages of positives for the BacGeneral assay were 76.2% and 75.3% respectively for *E. coli* and TC CFU positives (Table 7). In addition, *E. coli* and TC positives can be combined to represent contaminated samples, and compared against culture negatives (the non-contaminated samples) using a Chi-square test. The test revealed that there was a 2.62 (95% CI 1.68–4.08) times greater probability of observing a BacGeneral positive result for fecally contaminated samples compared to non-fecally contaminated samples (Chi-square = 18.40, $df = 1$, $p < 0.0001$).

**Table 6** | *E. coli* PCR results stratified by quantifiable/non-quantifiable cultured counts of *E. coli* per 100 mL of filtered water

| CFU/100 mL<br>*E. coli* PCR result | 1–15<br>*n* (%) | > 15<br>*n* (%) |
|---|---|---|
| Positive | 280 (49.1) | 101 (74.3) |
| Negative | 290 (50.9) | 35 (25.7) |

Dataset: A_2012.

**Table 7** | MST BacGeneral results stratified by membrane filtration contamination results of water samples, where negative represents no growth on the filter

| Culture result<br>MST result | *E. coli* positive<br>*n* (%) | TC positive<br>*n* (%) | Pooled *E. coli* TC data<br>*n* (%) | Negative<br>*n* (%) |
|---|---|---|---|---|
| Positive | 538 (76.2) | 55 (75.3) | 593 (76.1) | 51 (54.8) |
| Negative | 168 (23.8) | 18 (24.7) | 186 (23.9) | 42 (45.2) |

Datasets: A_2012, A_2013_NEG, A_2013_TC.

A comparison between positive and negative results for *E. coli* and MST qPCR assays revealed that the probability of obtaining an *E. coli* positive is 3.71 (95% CI 2.57–5.43) times greater for samples that also tested positive for the MST BacGeneral assay compared to samples that tested negative for the BacGeneral marker (Table 8) (Chi-square = 49.57, $df = 1$, $p < 0.0001$).

Overall, a very weak to weak correlation was found between both BacHuman and BacGeneral positive results and *E. coli* positive results. However, when investigated further, it was observed that correlation strength varied with levels of contamination as determined by traditional *E. coli* culture methods. The relationship strengthened to a moderate relationship for both MST assays in the overgrown group (Table 9).

## DISCUSSION

### Regional MST results

The predominant host contributing to fecal contamination of private wells varies by region; notably, A1 and C had statistically higher rates of human fecal contamination. It

**Table 8** | Number of samples positive for *E. coli* qPCR assay with respect to MST BacGeneral assay results

| MST<br>*E. coli* PCR result | Positive<br>*n* (%) | Negative<br>*n* (%) |
|---|---|---|
| Positive | 331 (61.3) | 50 (29.9) |
| Negative | 208 (38.7) | 117 (70.1) |

Dataset: A_2012.

**Table 9** | Correlations between *E. coli* and MST BacHuman/BacGeneral qPCR assays stratified by bacteriological analysis results

| Data | Culture group (CFU/100 mL) | *n* | Spearman's *ρ* | *p*-value |
|---|---|---|---|---|
| *E. coli* vs BacHuman | 1–15 | 152 | 0.051 | 0.531 |
| | 16–80 | 42 | 0.194 | 0.217 |
| | >80 | 32 | 0.469 | 0.007* |
| | All | 227 | 0.164 | 0.013* |
| *E. coli* vs BacGeneral | 1–15 | 240 | 0.233 | <0.001* |
| | 16–80 | 52 | 0.242 | 0.083 |
| | >80 | 39 | 0.443 | 0.005* |
| | All | 331 | 0.295 | <0.001* |

Dataset: A_2012.

should be noted that, while the Chi-square test revealed a difference in human contamination abundance among the regions, the *post-hoc* test applied (BHY) to identify where the differences were is a less conservative test. The Bonferroni *post-hoc* test was considered, but deemed inappropriate given the test's inherent over-conservative nature and the number of comparisons required. As such, the noted differences in rates of human fecal contamination between A1 and C compared to B and D (the high cattle region) should be interpreted with some caution.

Septic systems are the most likely source of the observed human fecal contamination in private wells given that most rural residents do not have access to municipal wastewater services. The exploratory analysis employed in this study revealed that greater well density may be related to greater abundance of human fecal contamination in private wells at a regional level. As domestic well density was chosen as a proxy for septic tank density, this provides initial evidence that directly implicates septic systems as the regional source of fecal contamination. Previous research in other regions supports this finding. A directly proportional relationship between fecal coliform concentration and septic tank distance was demonstrated by Arnade (1999). Additionally, the likelihood of pumping septic leachate has been found to increase with increasing septic system spatial density (Bremer & Harter 2012). Furthermore, where properties are smaller (and therefore densities are higher), domestic wells are more likely to be at a smaller distance from septic systems on the same and neighboring properties. Therefore, owners of private wells may have their wells contaminated from septic tanks owned by different individuals. As such, these owners cannot directly protect their well, but instead are dependent on local septic regulations. Consumption of human sourced fecal contamination in water has many negative health implications as various potential pathogens have been found in human waste (Cabral 2010; Li *et al.* 2015). As such, a review of septic system regulations should be conducted to ensure that private well owners are not being placed at risk.

The majority of fecal contamination is non-human sourced in both Regions B and D (the high cattle region). This suggests that the fecal animal sources could be either bovine or originating from any of the other eight species (pig, deer, horse, dog, cat, gull, goose and raccoon) targeted by the BacGeneral MST assay (Lee *et al.* 2010). Furthermore, the initial hypothesis stated that a higher rate of bovine contamination would be observed in the high cattle region. However, low rates were observed across all regions and, as such, it was not possible to test this hypothesis statistically. A possible reason for the low bovine rates observed is that private wells, as a groundwater source, require modes of entry for the bovine feces (as well other animal-sourced feces) to enter the aquifer. Typically this occurs due to groundwater under the direct influence (GUDI) of surface water which contains bovine fecal contamination via agricultural runoff (Health Canada 2013). These results suggest that most private wells studied were not from GUDI. In general, this study's observation of regional variation in primary fecal source suggests that primary prevention for groundwater contamination has to be regionally informed. Recent work by Åström *et al.* (2015) suggests that MST marker prevalence in combination with regional experts' judgements may more accurately predict the origin of fecal contamination.

The strong correlations between human and general markers in all regions indicate that where samples are contaminated with human feces, human feces are the predominant source of fecal contamination with little interference from the other nine species targeted by the BacGeneral assay. Conversely, given the weaker relationship between BacBovine and BacGeneral, samples contaminated with bovine feces may also be contaminated by other fecal sources (including human). Lee *et al.* (2014) investigated host-specific fecal contamination in a river watershed and found significant correlations between

human-specific and general markers where the primary contamination source was a wastewater treatment plant. In addition, they found no statistically significant relationship between bovine and general markers in the same samples. Reverse results were found in a study of surface water samples, where the primary contamination source was agricultural run-off (Lee *et al.* 2014). It is noteworthy that the water samples in this study originate from various locations, as previous MST studies are traditionally performed using water originating from a point location (Blanch *et al.* 2011), and yet still exhibit strong marker-to-marker correlation.

## Indicators of fecal contamination

Minor, but significant, correlations were observed between *E. coli* PCR and BacHuman and BacGeneral results (using A_2012 data) with a stronger relationship observed between the latter. When stratified by levels of contamination (defined by culture methods), variation in correlation was observed. For BacHuman, the correlation was only statistically significant for the overgrown group (despite having the lowest statistical power). For BacGeneral, the strength of the correlation varied with level of contamination. The relationships observed for both *E. coli* and BacHuman and BacGeneral assays on overgrown samples were moderate in strength. These findings indicate that a more stable relationship exists between indicators when *E. coli* contamination is greater according to traditional culture methods. Numerous previous studies have attempted to correlate MST and *E. coli* as indicators of fecal contamination with varying results (and also varying power to detect) across a variety of watershed types. For instance, Layton *et al.* (2006) found a significant correlation ($r^2 = 0.85$) between results for an assay detecting human, cattle and horse feces and *E. coli* concentrations for creek water samples. Lee *et al.* (2014) also reported a significant correlation ($r = 0.73$) between BacGeneral gene copies and culturable *E. coli* results when pooled for multiple sampling sites of a river watershed. When assessed individually, correlation results varied depending on sampling site. In addition, Lee *et al.* (2010) found no significant correlation between general Bacteroidales markers and *E. coli* cell numbers for creek and river water samples. Conversely, Sauer *et al.* (2011) found no significant correlation between human Bacteroides qPCR markers and culturable *E. coli* and a weak correlation ($r^2 = 0.34$) between a wide-ranging Bacteroides marker with culturable *E. coli*.

The previously noted studies investigated point-source surface waters and were able to quantitatively correlate CFU/100 mL and MST qPCR results, given a much larger range of values for culturable *E. coli*. Given the lower range of CFU counts per 100 mL in private well waters, the *E. coli* qPCR assay was used in this study to compare across species. Sauer *et al.* (2011) found significant correlations ($r^2 = 0.16$ and $r^2 = 0.46$) between *E. coli* qPCR and human and wide-ranging Bacteroides markers, which were within the same range as the $r^2$ values found by this study.

Previous studies have only investigated samples that were culturable for *E. coli*. This study investigated the presence/absence of MST markers in conjunction with bacteriological results for private well water samples, which included findings of TC only (no evidence of *E. coli*) when cultured, or absolute negative results. While, overall, culturing for *E. coli* (or other coliforms) led to greater probability of testing positive using MST, 54.8% of *E. coli* culture negative samples tested positive for the BacGeneral assay. The same phenomenon occurred with the *E. coli* qPCR assay, where a qPCR *E. coli* positive sample had greater probability of having a BacGeneral positive result, but 38.7% were *E. coli* qPCR negative and BacGeneral positive. These findings raise questions about the mechanisms involved in obtaining a BacGeneral positive and *E. coli* negative result. Overall in this study, 90.2% of samples that were culture negative, but BacGeneral positive samples (46 samples), tested positive using the *E. coli* qPCR assay. Disinfection of water at the well may explain these discordances. Potential water treatment would render any fecal-sourced *E. coli* cells non-culturable, but bacterial DNA would still be present and detectable within all qPCR assays. However, the remaining 9.8% of samples ($n = 5$) tested negative for *E. coli* in both culture and qPCR assay, but positive for BacGeneral; hence, water treatment alone cannot explain the disagreement between indicators. Additionally, 55% of samples positive for TC only, by culture, tested positive for the BacGeneral assay. These specimens are a further indication of the limits of *E. coli* as an indicator, as the viability of TC prevents treatment as an explanation. The relationship between TC and recent

fecal contamination was demonstrated by Savichtcheva *et al.* (2007) who found a strong correlation ($r^2 = 0.8$) between TC/100 mL and Bacteroides gene copies/100 mL.

Furthermore, 23.8% of samples culture-positive for *E. coli* and 29.9% of *E. coli* qPCR samples were negative for the BacGeneral assay. This finding is similar to results from a previous study by this research group (Krolik *et al.* 2014), where a quarter of the samples tested negative for the BacGeneral assay yet positive for *E. coli* by culture. The possibility of false positives by culture cannot be eliminated, although the high prevalence of BacGeneral negatives rules this out as the sole explanation. In instances where the CFU/100 mL count for *E. coli* is low, it is possible that the MST assay was not sensitive enough to detect trace amounts of host-specific fecal contamination. Additionally, the general MST markers include only 10 species whereas the source of fecal contamination may be attributed to other hosts. Finally, several studies have documented the existence of naturalized *E. coli* populations, capable of surviving for extensive periods in environmental matrices and exhibiting genetic differences from animal GI tract populations (Byappanahalli *et al.* 2006; Ishii *et al.* 2006; Ishii & Sadowsky 2008; VanderZaag *et al.* 2010). These *E. coli* subpopulations are not of recent fecal origin and therefore not representative of direct fecal contamination in private well water samples.

The probability of having an *E. coli* PCR positive result were higher if the sample had a higher level of culturable *E. coli.* As per the MST assay, it was not possible to further quantitatively relate the two tests, though the literature suggests a moderate correlation (Ahmed *et al.* 2012). However, there was a high level of disagreement between *E. coli* CFU/100 mL and *E. coli* qPCR assay (46% of samples that tested positive by culture were negative by qPCR). Additionally, when stratified to only include the quantifiable range (>15 CFU/100 mL), 65% of samples were positive by both *E. coli* culture and qPCR methods. This increase in agreement between the two methods once culturable *E. coli* reaches the quantifiable range suggests that the qPCR assay is unable to detect a signal in samples with low *E. coli* CFU counts. The assay was performed using a single cell (7 gene copy) cut-off, which, when tested against samples that were culture positive for *E. coli*, TC positive only (no *E. coli*), and negative for both *E. coli* and TC, yielded the following results: 55%, 25% and 15% positive, respectively. Additionally, a large portion of samples had a signal below the LOD and a supplemental analysis was conducted to further investigate this phenomenon. It was determined that 2.5 gene copies was a stable LOD, and at this level all datasets were at least 70% positive for *E. coli,* potentially indicating limits with the bacteriological testing. However, the 30–45% of *E. coli* culture positives (depending on the qPCR assay cut-off) that did not test positive by the PCR assay remain to be of concern. The high prevalence of signal below the LOD, alongside the correlation results, may also indicate that the *E. coli* qPCR assay used in this study is not sufficiently sensitive for small-volume water samples (i.e., low *E. coli* counts are difficult to distinguish from background signal). One of the reasons for the low sensitivity may be the volume of eluted sample used for qPCR, however these volumes are standard practice for qPCR methodology.

Traditionally, when evaluating laboratory diagnostic methodologies, methodologies are compared to definitive tests that are used to designate true positives and true negatives. This principle could not be applied in this study, because all tests are indicators of fecal contamination, not direct evidence of fecal contamination as it relates to the presence of pathogens and the risk of illness, the true aim of testing for fecal indicators. The analyses employed by this study compared each potential indicator to one another and noted their agreement of positives and negatives (where appropriate quantities allowed), but made no direct conclusions regarding the cases where the indicators disagreed given the uncertainties found in the tests themselves. Given that disagreement was found between all three methods, it is necessary to find and rigorously test techniques that are capable of designating 'true positives and negatives' for fecal contamination. Newer molecular techniques offer a promising new avenue and should be further investigated. Although not expected to be feasible for routine application in laboratories, use of these sophisticated technologies could enable discernment regarding which indicator assay is best.

Moreover, before MST methods can become integrated into routine laboratory testing as indicators of fecal contamination, characteristic ranges must be established. Currently, a general lack of MST method standardization creates a

barrier for cross-comparison studies. Thus, guidelines need to be established regarding not only acceptable, standardized methodology, but also the limits or concentrations of various molecular markers that are required in order to deem water sufficiently safe to drink.

## CONCLUSIONS

This study was among the first to track fecal contamination sources using Bacteroidales in groundwater non-point sources across multiple regions. While, in general, human Bacteroidales markers were the most common individual marker, statistically significant variation across regions was found. This indicates that primary prevention against groundwater contamination must be informed by investigations at a regional level. Additionally, this study utilized the breadth of samples collected for MST to investigate the effectiveness of fecal indicator tests (the traditional *E. coli* culture methods versus *E. coli* and MST qPCR assays). While there were overall trends toward agreement between methodologies, there was still considerable disagreement. Better assessment tools for determining the presence of fecal contamination in water samples are required.

## ACKNOWLEDGEMENTS

The authors wish to thank the employees of Public Health Ontario for their kind support, in particular, staff at the Kingston, London and Hamilton laboratory locations. A special thank you also to Dai Young Lee and Susan Weir who shared their molecular expertise and to Kayla Poirier and Sophie Felletier. The funding for this study was provided in part by the Physicians' Services Incorporated Foundation.

## REFERENCES

Ahmed, W., Masters, N. & Toze, S. 2012 Consistency in the host specificity and host sensitivity of the Bacteroides HF183 marker for sewage pollution tracking. *Letters in Applied Microbiology* **55** (4), 283–289.

Arnade, L. J. 1999 Seasonal correlation of well contamination and septic tank distance. *Groundwater* **37** (6), 920–923.

Åström, J., Pettersson, T. J., Reischer, G. H., Norberg, T. & Hermansson, M. 2015 Incorporating expert judgments in utility evaluation of bacteroidales qPCR assays for microbial source tracking in a drinking water source. *Environmental Science and Technology* **49** (3), 1311–1318.

Blanch, A. R., Hagedorn, C. & Harwood, V. J. 2011 *Microbial Source Tracking: Methods, Applications, and Case Studies.* Springer, New York.

Bremer, J. E. & Harter, T. 2012 Domestic wells have high probability of pumping septic tank leachate. *Hydrology and Earth Systems Science* **16** (8), 2453–2467.

Byappanahalli, M. N., Whitman, R. L., Shively, D. A., Sadowsky, M. J. & Ishii, S. 2006 Population structure, persistence, and seasonality of autochthonous Escherichia coli in temperate, coastal forest soil from a Great Lake watershed. *Environmental Microbiology* **8**, 504–513.

Cabral, J. P. 2010 Water microbiology. Bacterial pathogens and water. *International Journal of Environmental Research and Public Health* **7** (10), 3657–3703.

Census of Agriculture, Farm and Farm Operator Data-95-640-XWE 2011 Ottawa, Statistics Canada, http://www.statcan.gc.ca/eng/ca2011/index (accessed 1 April 2013).

Charrois, J. W. 2010 Private drinking water supplies: challenges for public health. *Canadian Medical Association Journal* **182**, 1061–1064.

Chern, E. C., Siefring, S., Paar, J., Doolittle, M. & Haugland, R. A. 2011 Comparison of quantitative PCR assays for Escherichia coli targeting ribosomal RNA and single copy genes. *Letters in Applied Microbiology* **52** (3), 298–306.

Haack, S. K., Fogarty, L. R., Stelzer, E. A., Fuller, L. M., Brennan, A. K., Isaacs, N. M. & Johnson, H. E. 2013 Geographic setting influences Great Lakes beach microbiological water quality. *Environmental Science and Technology* **47** (21), 12054–12063.

Health Canada 2013 *Guidance on the use of the microbiological drinking water quality guidelines – Catalogue No H144-12/2013E*. Water and Air Quality Bureau, Healthy Environments and Consumer Safety Branch. Ottawa, Ontario.

International Organization for Standardization (ISO) 2000 Water quality – Guidance on validation of microbiological methods – ISO/TR 13843. Switzerland. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=22599 (accessed 22 May 2014).

Ishii, S. & Sadowsky, M. J. 2008 *Escherichia coli* in the environment: implications for water quality and human health. *Microbes and Environments* **23**, 101–108.

Ishii, S., Ksoll, W. B., Hicks, R. E. & Sadowsky, M. J. 2006 Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. *Applied and Environmental Microbiology* **72**, 612–621.

Krentz, C. A., Prystajecky, N. & Isaac-Renton, J. 2013 Identification of fecal contamination sources in water using host-associated markers. *Canadian Journal of Microbiology* **59**, 210–220.

Krolik, J., Maier, A., Evans, G., Belanger, P., Hall, G., Joyce, A. & Majury, A. 2013 A spatial analysis of private well water Escherichia coli contamination in southern Ontario. *Geospatial Health* **8** (1), 65–75.

Krolik, J., Evans, G., Belanger, P., Maier, A., Hall, G., Joyce, A., Guimont, S., Pelot, A. & Majury, A. 2014 Microbial source tracking and spatial analysis of *E. coli* contaminated private well waters in southeastern Ontario. *Journal of Water and Health* **12** (2), 348–357.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R. & Sayler, G. 2006 Development of Bacteroides 16S rRNA Gene TaqMan-Based Real-Time PCR assays for estimation of total, human and bovine fecal pollution in water. *Applied and Environmental Microbiology* **72** (6), 4214–4224.

Leclerc, H., Mossel, D. A., Edberg, S. C. & Struijk, C. B. 2001 Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annual Review of Microbiology* **55**, 201–234.

Lee, D. Y., Weir, S. C., Lee, H. & Trevors, J. T. 2010 Quantitative identification of fecal water pollution sources by TaqMan real-time PCR assays using Bacteroidales 16S rRNA genetic markers. *Applied Microbiology and Biotechnology* **8**, 1373–1383.

Lee, D. Y., Lee, H., Trevors, J. T., Weir, S. C., Thomas, J. L. & Habash, M. 2014 Characterization of sources and loadings of fecal pollutants using microbial source tracking assays in urban and rural areas of the Grand River Watershed, Southwestern Ontario. *Water Research* **53**, 123–131.

Li, X., Harwood, V. J., Nayak, B., Staley, C., Sadowsky, M. J. & Weidhaas, J. 2015 A novel microbial source tracking microarray for pathogen detection and fecal source identification in environmental systems. *Environmental Science and Technology* **49** (12), 7319–7329.

Liu, Y., Gilchrist, A., Zhang, J. & Li, X. F. 2008 Detection of viable but nonculturable *Escherichia coli* o157:H7 bacteria in drinking water and river water. *Applied and Environmental Microbiology* **74** (5), 1502–1507.

Lothigus, A., Sjoling, A., Svennerholm, A. M. & Bolin, I. 2009 Survival and gene expression of enterotoxigenic *Escherichia coli* during long-term incubation in sea water and freshwater. *Journal of Applied Microbiology* **108**, 1441–1449.

Ogden, I. D. & Watt, A. J. 1991 An evaluation of fluorogenic and chromogenic assays for the direct enumeration of *Escherichia coli*. *Letters in Applied Microbiology* **13**, 212–215.

Perchec-Merien, A. M. & Lewis, G. D. 2012 Naturalized *Escherichia coli* from New Zealand wetland and stream environments. *FEMS Microbiology Ecology* **83**, 494–503.

Sauer, E. P., VandeWalle, J. L., Bootsma, M. J. & McLellan, S. L. 2011 Detection of the human specific *Bacteroides* genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. *Water Research* **45**, 4081–4091.

Savichtcheva, O., Okayama, N. & Okabe, S. 2007 Relationships between Bacteroides 16S rRNA genetic markers and presence of bacterial enteric pathogens and conventional fecal indicators. *Water Research* **41** (16), 3615–3628.

Statistics Canada 2011 Dissemination area (DA) – Census Dictionary. Available from: https://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo021-eng.cfm (accessed 1 June 2015).

VanderZaag, A. C., Campbell, K. J., Jamieson, R. C., Sinclair, A. C. & Hynes, L. G. 2010 Survival of *Escherichia coli* in agricultural soil and presence in tile drainage and shallow groundwater. *Canadian Journal of Soil Science* **90**, 495–505.

Well Water Information System (WWIS) 2013 Ontario Ministry of the Environment, Canada. http://www.ontario.ca/environment-and-energy/well-record-data (accessed 1 February 2013).

First received 24 January 2016; accepted in revised form 12 August 2016. Available online 1 October 2016