Tracking the downstream impacts of inadequate sanitation in central Appalachia

Jacob Cantor, Leigh-Anne Krometis, Emily Sarver, Nicholas Cook and Brian Badgley

ABSTRACT

Poor sanitation in rural infrastructure is often associated with high levels of fecal contamination in adjacent surface waters, which presents a community health risk. Although microbial source tracking techniques have been widely applied to identify primary remediation needs in urban and/or recreational waters, use of human-specific markers has been more limited in rural watersheds. This study quantified the human source tracking marker *Bacteroides*-HF183, along with more general fecal indicators (i.e. culturable *Escherichia coli* and a molecular *Enterococcus* marker), in two Appalachian streams above and below known discharges of untreated household waste. Although *E. coli* and *Enterococcus* were consistently recovered in samples collected from both streams, *Bacteroides*-HF183 was only detected sporadically in one stream. Multiple linear regression analysis demonstrated a positive correlation between the concentration of *E. coli* and the proximity and number of known waste discharge points upstream; this correlation was not significant with respect to *Bacteroides*-HF183, likely due to the low number of quantifiable samples. These findings suggest that, while the application of more advanced source targeting strategies can be useful in confirming the influence of substandard sanitation on surface waters to justify infrastructure improvements, they may be of limited use without concurrent traditional monitoring targets and on-the-ground sanitation surveys.

Key words | fecal indicator bacteria, HF183, rural watersheds, sanitation, source tracking

Jacob Cantor
Leigh-Anne Krometis (corresponding author)
Biological System Engineering,
Virginia Tech,
155 Ag Quad Lane, Seitz Hall,
Blacksburg, VA 24060,
USA
E-mail: *krometis@vt.edu*

Emily Sarver
Mining and Minerals Engineering,
Virginia Tech,
108A Holden Hall,
Blacksburg, VA 24061,
USA

Nicholas Cook
Forest Ecohydrology and Watershed Management,
Department of Forest Engineering,
Resources, and Management,
College of Forestry, Oregon State University,
215 Peavy Hall,
Corvallis, OR 97731,
USA

Brian Badgley
Crop and Soil Environmental Sciences,
Virginia Tech,
RB1880 Suite 1129 Room 1121,
Blacksburg, VA 24061,
USA

INTRODUCTION

Despite major advancements in wastewater treatment in the past century, inadequate sanitation remains an issue for some communities in the United States. Infrastructure for sanitation is particularly lacking in remote communities where there are fewer users to cover infrastructure costs and technical assistance is not readily available locally (Tarricone 1989; Gasteyer & Vaswani 2004). This can result in inadequate wastewater management strategies that direct untreated household waste to local surface waters with minimal pre-treatment (e.g. soakaways, straight pipes), posing potential health risks to adjacent and downstream communities.

Improper treatment of household wastewater has long been recognized as a public health hazard. In the most recently identifiable national survey of medical care providers in rural communities, surface water pollution was identified as a primary environmental health concern (Robson & Schneider 2001). Waste entering common waterways used for recreation or drinking water can be a point of human exposure for infectious disease, particularly in low population areas (Denno *et al.* 2009; Collier *et al.* 2012). Relative to less developed countries, the incidence of waterborne diseases such as hepatitis A, salmonellosis, and typhoid is very low in the United States, but real risks do exist in

doi: 10.2166/wh.2017.005

areas without proper water and sanitation (Gasteyer & Vaswani 2004). Eliminating these sources of pathogen contamination could result in both significant public health and economic benefits, as waterborne illnesses have multiple adverse economic effects including the cost of hospitalizations, losses in productivity at places of employment, and absences from school (Collier *et al.* 2012).

A particularly unique set of conditions exists in rural communities in central Appalachia. In this region, underlying karst soils are often inadequate to support traditional on-site wastewater treatment strategies such as septic systems, i.e. the soils are too thin to provide sufficient biological treatment prior to discharge to the environment. Furthermore, in narrow Appalachian valleys, residences are often located too close to streams to permit the construction of proper drain fields (Cook *et al.* 2015). Other on-site solutions, such as package treatment plants or sand filters have seen limited use, but are often precluded by distinctly local challenges (e.g. the necessity of shared responsibility for installation and maintenance in already resource-stressed communities; Tarricone 1989). When traditional municipal or onsite septic systems are not readily available, untreated household waste may simply be 'straight piped' into nearby surface waters (Banks *et al.* 2005). The number of straight pipes in the entire region is unknown, but the practice is not uncommon, e.g. in Letcher County, KY, one survey estimated 3,000 straight pipes serving about 40% of 30,000 total residents (Glasmeier & Farrigan 2003). Though the discharge of untreated household wastewater is intuitively expected to result in significant deterioration of receiving water quality, little monitoring data describing the magnitude of impact is available.

Standard monitoring practices to identify water quality impairments generally only target fecal indicator bacteria (FIB) such as total coliforms or *Escherichia coli*. While generally non-pathogenic themselves, these are considered sentinels of potential pathogen presence and associated human health risks. However, because these organisms are common to the intestinal tract of all warm-blooded animals, their presence provides limited information regarding the relative contributions of specific sources of fecal contamination (e.g. human vs. animal). Though the transfer of zoonotic pathogens is important, fecal contamination by human sources is generally considered of greatest health concern since many waterborne pathogens (viruses in particular) exhibit host specificity (Harwood *et al.* 2014). The inclusion of library-independent, established source-specific markers in monitoring efforts can greatly aid in watershed assessment: understanding the origin of fecal contamination can improve local understanding of potential health risks and inform decisions about appropriate remediation actions to meet water quality goals (Scott *et al.* 2002). At present, the genetic marker HF183 from the anaerobic *Bacteroides* spp. is well established as strongly indicative of the presence of human fecal contamination in waters with multiple potential contamination (Seurinck *et al.* 2005; Harwood *et al.* 2009, 2014; Boehm *et al.* 2013; Sidhu *et al.* 2013). Because the majority of prior studies have tracked *Bacteroides*-HF183 concentrations downstream from relatively large centralized wastewater treatment facilities or highly urbanized areas, the relationship between FIB, source tracking markers, and health risks in less populated regions is poorly understood (Harwood *et al.* 2009). However, application of these more advanced monitoring techniques in rural regions with known sanitation issues could be useful in tracking the downstream reach of contamination from seemingly isolated communities and justifying and/or evaluating infrastructure improvements. In addition, these relatively simple and sparsely populated systems represent a unique opportunity to evaluate the robustness of common microbial source tracking (MST) markers compared to settings in which they are more commonly used.

This study aimed to directly link microbial measures of fecal contamination with the prevalence of straight pipes in two central Appalachian streams. Specific objectives included: (1) detection of the human source tracking marker *Bacteroides*-HF183 in receiving waters polluted by untreated household waste from rural communities; (2) assessment of the co-occurrence of *Bacteroides*-HF183 with *E. coli* and *Enterococcus*, which are currently recommended by regulatory agencies as sentinels of fecal contamination and associated public health risks based on culturable and molecular targets, respectively (USEPA 2012; Haugland *et al.* 2013); (3) exploration of correlations between microbial indicators and physiochemical water quality parameters; and (4) development of a spatial statistical strategy to link straight pipe locations to

longitudinal downstream measures of *E. coli* in these streams.

METHODS

Site description

The study area consists of two 12-digit hydrological unit code (HUC) watersheds in southwestern Virginia (Figure 1): one drains to Callahan Creek and the other to Roaring Fork. Both of these streams are tributaries to the Powell River basin in Virginia. Per the National Land Cover Database, the land cover mainly consists of hardwood deciduous forest (66%). There is very little agricultural activity (0.5%), but significant surface mining (17%). The remainder of the land cover is composed of grassland (14.5%) and development (2%). Development in both watersheds is mainly concentrated adjacent to the streams in the flatter, more habitable areas known as mountain hollows. Samples were collected directly above and below residential areas or any known discharges of fecal contamination, as well as at the top of the watershed and watershed outlet, for a total of 14 in-stream sampling sites in each watershed.

Sample collection

Water samples were collected monthly from August 2012 until August 2014 and from January 2016 to April 2016. For all months and at all sites, samples were collected for *E. coli* analysis ($n = 28$ for each site), and in-stream temperature, dissolved oxygen, pH and specific conductivity were measured on-site using a YSI Quattro Pro Plus (YSI, Yellow Springs, OH, USA). Between April 2014 and August 2014 and January 2016 to April 2016, an additional sample aliquot was collected for subsequent molecular

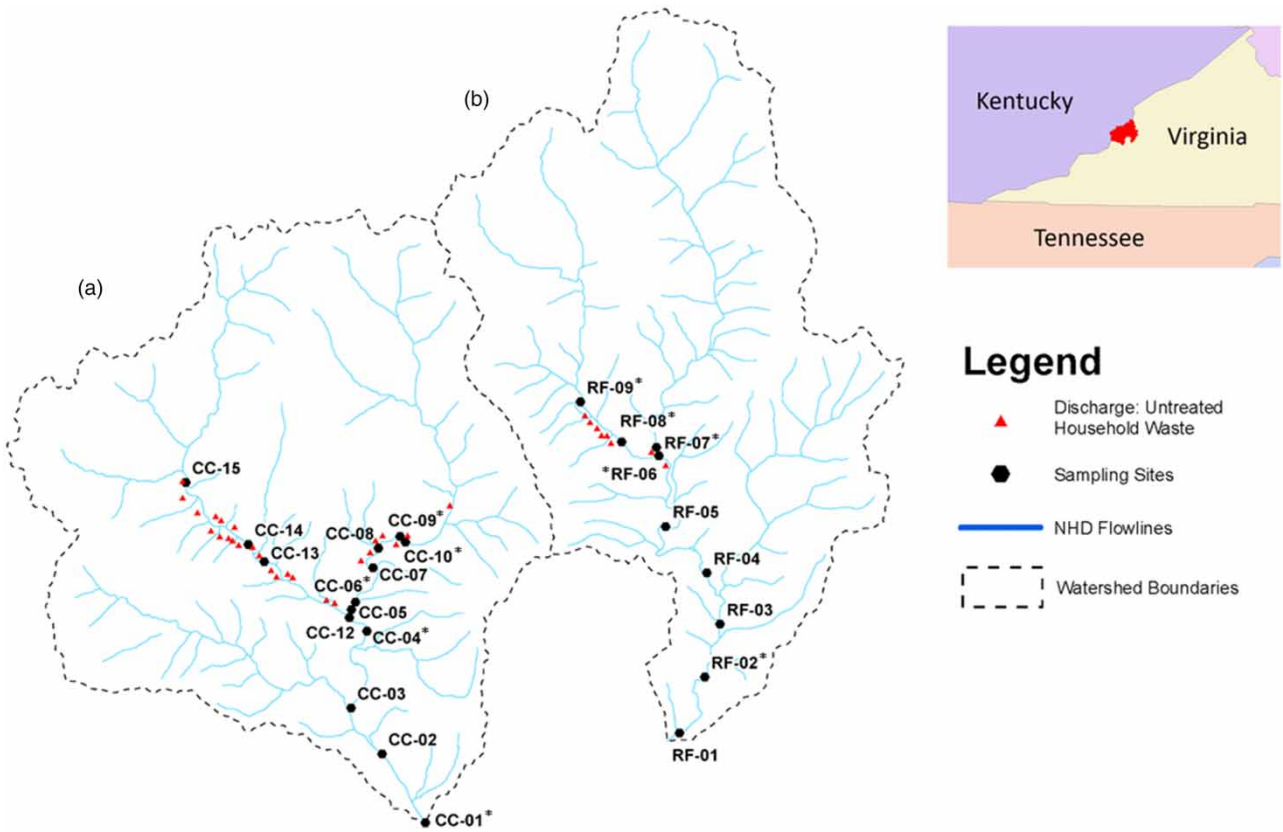

Figure 1 | Sampling sites and known discharges for 12-digit HUC watersheds Callahan Creek (a) and Roaring Fork (b). *Indicates the site was tested for both molecular markers (*Bacteroides*-HF183 and *Enterococcus*) as well as *E. coli*.

analyses ($n = 9$) from five sites in each watershed. All samples for microbiological analysis were collected at the thalweg in 250 mL pre-sterilized bottles, and transported on ice back to the Biological Systems Engineering Seitz Hall laboratory at Virginia Tech for prompt analysis (i.e. within 6 hours for culturable microorganisms; filtration for molecular preservation within 48 hours).

Microbial analyses

Analysis for *E. coli* concentrations occurred immediately upon return to the laboratory via the Colilert defined substrate method using the Quanti-Tray/2000 (IDEXX, Westbrook, ME, USA). Following 24 hours incubation at 37 °C, a most probable number (MPN) of bacteria could be determined based on the number of positive wells (i.e. fluorescent under UV light).

For samples collected for molecular analyses, two additional 100 mL aliquots from each sample were vacuum filtered through a Millipore 0.4 μm filter (Merck KGaA, Darmstadt, Germany). Filter effluent was discarded and filters were stored in cryotubes at −80 °C prior to DNA extraction. Within six months, DNA from the filters was extracted using a PowerWater DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). Copy numbers of *Enterococcus* were quantified via quantitative polymerase chain reaction (qPCR) as described previously by Haugland *et al.* (2005). Copy numbers of the human-specific *Bacteroides*-HF183 MST marker were quantified by the procedures outlined in Griffith *et al.* (2013) and Seurinck *et al.* (2005) (Table 1). Samples were run in triplicate, with each well containing 20 μL of Master Mix and 5 μL of sample. A no-template control was tested along with the samples to indicate any sources of contamination. A base plasmid containing the target marker at known amounts was used to create a standard calibration curve. Results were considered valid if the run-specific standard curve had an R^2 value above 0.99 and efficiency between 0.90 and 1.10. The *Bacteroides*-HF183 master mix also included a plasmid containing an internal amplification control (IAC) sequence to quantify inhibition (Griffith *et al.* 2013). The limit of detection (LOD) for *Bacteroides*-HF183 was determined by scaling the lowest detectable amount of *Bacteroides*-HF183 in tested samples to the total amount of eluted sample from DNA extraction. The same process was repeated with the lowest value on the standard curve (10 copies/mL) to determine the limit of quantification (LOQ).

Statistical methods

All statistical analyses were conducted in JMP 12 (SAS, Cary, NC, USA). Normality/non-normality of each dataset

Table 1 | Master mix composition for *Bacteroides*-HF183 and *Enterococcus* qPCR procedures

Target gene	Primer sequences and sources	Reaction mixture	Volume (μL)	Thermal profile
Bact. HF183		Primer probe mix	3	50 °C – 2 min,
	5′- ATCATGAGTTCACATGTCCG -3′	HF-183	10	95 °C – 10 min,
	5′- CTTCCTCTCAGAACCCCTATCC -3′	BacR287	10	40 Cycles
	[6FAM] - 5′- CTAATGGAACGCATCCC –MGB	BacP234MGB	4	95 °C – 15 sec
	[VIC]-5′- AACACGCCGTTGCTACA –MGB	Bac234IAC	4	60 °C – 1 min
		PCR-grade water	4	
	(Seurinck *et al.* 2005; Griffith *et al.* 2013)	SsoAdvanced Universal Probes Mix	12.5	
		Bovine serum albumin (BSA)	2.5	
		Plasmid with IAC	1	
Ent.		Primer probe mix	3	50 °C – 2 min,
	5′-GAGAAATTCCAAACGAACTTG-3′	ECST748F	10	95 °C – 10 min,
	5′-CAGTGCTCTACCTCCATCATT-3′	ENC854R	10	40 Cycles
	5′-TGGTTCTCTCCGAAATAGCTTTAGGGCTA-3′	GLP813TQ	4	95 °C – 15 sec
		PCR grade water	2	60 °C – 1 min
	(Haugland *et al.* 2005)	BSA	1.5	
		iTaq Universal Probe Mix	12.5	

was determined via observation of the normal quantile plot prior to analysis. Correlations between molecular targets, *E. coli*, and general measures of water quality were assessed via Spearman's rank correlation.

A statistical mixed-effects model was developed to determine if concentrations of microbial markers were significantly related to the number of waste discharges upstream of each site, the distance from each site to known waste discharges upstream, and the season (winter/'high flow' vs not-winter/'low flow'). Previous studies have performed multidimensional regression to create predictive models linking concentrations of bacteria to explanatory factors (e.g. water quality, land use (Hampson *et al.* 2010; Herrig *et al.* 2015)). The model used here is not intended to be predictive, but to identify and define variables that are statistically explanatory:

$$\mu_{ij} = \beta_0 + \beta_1 S + \beta_2 D + \alpha_i + \varepsilon_{ij} \quad (1)$$

Fixed variables included the sum of inverse distances to all upstream discharges of untreated household waste (D), as well as a binary variable indicating whether the season was winter (S; where December, January and February were considered winter months). Although the location of known waste discharges included both septic tanks and straight pipes, the effects from these sources were not differentiated, as local health department sanitary surveys suggest most local septics do not meet codes and/or are failing. Random variables were normally distributed and centered at 0, and accounted for dependence of measurements taken within sites (α_i) as well as possible sources of random error (ε_{ij}). Coefficients (β_0, β_1 and β_2) were determined by the model to describe the behavior of the S and D variables and the general direction of the correlations.

RESULTS AND DISCUSSION

Detection and longitudinal trends of microbial indicators

Bacteroides-HF183 was detected in samples collected in Callahan Creek (Figure 2(a)) but was not found in any samples collected in Roaring Fork; analysis of the IAC did not indicate significant inhibition in any of the samples. The genetic marker, *Bacteroides*-HF183, is known to be present at varying concentrations in different individuals (Field 2002). When tracking sewage contamination from large concentrated communities (e.g. combined sewer overflows in cities) the marker is present in the population at high enough concentrations to be reliable. However, because HF183 is not ubiquitous in the human intestinal tract, its use in tracking waste from more isolated rural communities may be more limited. Though both study watersheds are quite rural, Callahan Creek does have roughly double the population of Roaring Fork with a similar area (180 vs 80

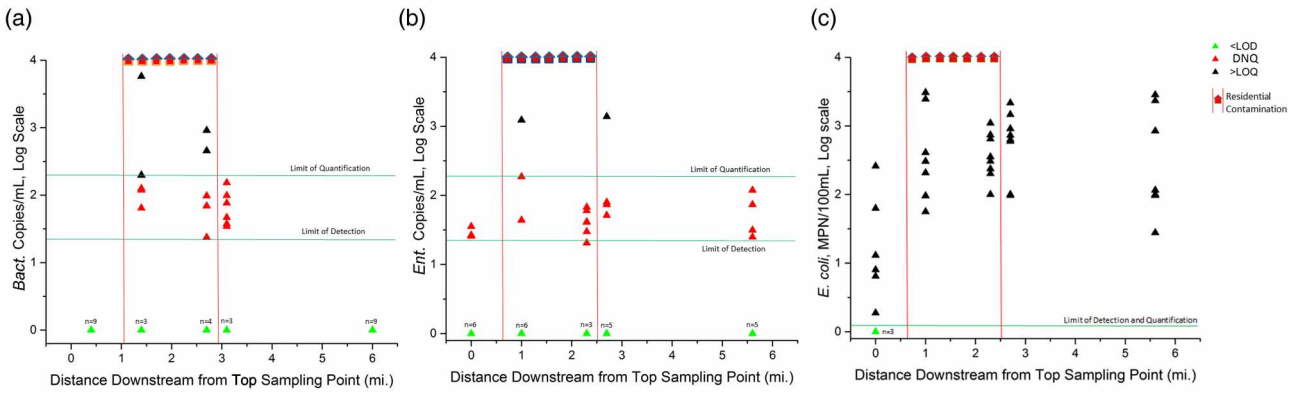

Figure 2 | Longitudinal profile of *Bacteroides*-HF183 (a), *Enterococcus* (b) and *E. coli* (c) in Callahan Creek. Distance is measured along the stream starting from the most upstream sampling point. The data points are extrapolated below the LOQ. The number of samples below the LOD at each site is indicated by an n value ($n = 9$ samples in total for each site). This does not include samples taken directly from household wastewater discharges.

households, counted on Google maps satellite imagery). The population of Roaring Fork may simply be too low to produce measurable levels of HF183 downstream.

As illustrated in Figure 2(a), *Bacteroides*-HF183 concentrations were negligible at the most upstream point in Callahan Creek, which was expected given that this sampling site is above the known discharge points of untreated household waste. The next sampling site is directly below a residential community, and shows a spike in the concentrations of *Bacteroides*-HF183. Samples taken directly from the household waste effluent discharge to Callahan Creek ranged from detectable but not quantifiable (DNQ) to 59,400 copies per mL. Though this value may appear high, it is notably lower than typical values of *Bacteroides*-HF183 in raw sewage which have been observed between 10^6 and 10^7 copies/mL (e.g. Seurinck *et al.* 2005). In Callahan Creek itself, *Bacteroides*-HF183 levels remained between 10 and 10^3 copies/mL (43% positive samples and 13% quantifiable samples) until the furthest downstream sampling site, where it was again below the LOD. Observed levels are similar to those reported by Tambalo *et al.* (2012), in a rural environment with known discharges of human waste in Saskatchewan, Canada. In this study, *Bacteroides* was detected via qPCR of the genetic marker BacH, at concentrations at or below 10^4 copies/mL in different sites at positive detection rates between 25 and 47% of collected samples. In contrast, a recent urban study by Sidhu *et al.* (2013), reported *Bacteroides*-HF183 in over 90% of the samples adversely impacted by untreated or partially treated human waste. Denser population centers with sewage intrusion appear to result in more successful detection of *Bacteroides*-HF183 and higher concentrations in receiving waters.

The *Enterococcus* molecular marker was detected from samples in both Callahan Creek and Roaring Fork (Figures 2(b) and 3(a) respectively). Similar to *Bacteroides*-HF183 in Callahan Creek, *Enterococcus* showed a spike in detection in samples directly downstream from residential communities. However, unlike *Bacteroides*-HF183, *Enterococcus* did not return to background concentrations at the watershed outlet. This trend is similar to that of *E. coli*, which also was elevated directly downstream from residential areas, and remained higher than its background level at the watershed outlet in both Callahan Creek and Roaring Fork (Figures 2(c) and 3(b) respectively).

Both *Enterococcus* and *E. coli* are universally present in the human digestive system, and may also originate from wildlife or domestic animals, however, additional non-human significant discharges of *E. coli* are considered to be negligible in this area. Agricultural land use is low

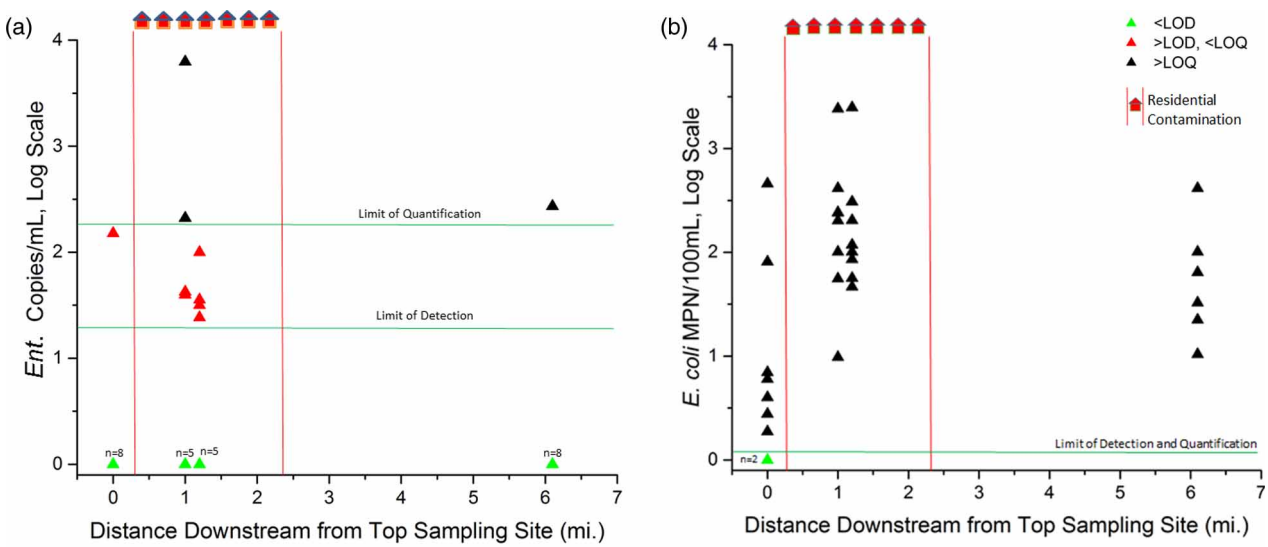

Figure 3 | Longitudinal profile of *Enterococcus* (a) and *E. coli* (b) in Roaring Fork. Distance is measured along the stream starting from the most upstream sampling point. The data points are extrapolated below the LOQ. The number of samples below the LOD at each site is indicated by an *n* value. This does not include samples taken directly from household wastewater discharges.

(0.5%), and samples collected in wholly forested areas upstream from the straight-piped communities generally showed barely detectable *E. coli* (geometric mean = 4 MPN/100 mL), suggesting low levels of background/wildlife. Longer longitudinal persistence in *Enterococcus* and *E. coli* levels therefore may be the result, at least partially, of longitudinal decay. Unlike *Bacteroides*, which is an obligatory anaerobe, *E. coli* and *Enterococcus* are facultative anaerobic bacteria, and therefore can survive in the aerobic surface water environment. Highly oxygenated mountain streams may cause particular stress to *Bacteroides*.

Temporal trends of microbial markers

Seasonal trends are not immediately apparent in the recovery of *Bacteroides*-HF183 and *Enterococcus* from Callahan Creek, though values varied considerably by month (Table 2). *E. coli* was generally lower during winter and early spring months (i.e. January–April) than late spring and summer months (i.e. May–August). There were two notably high values in June 2014 in consecutive sampling sites (i.e. CC09 and CC06). This is consistent with expectations since flow rate should generally increase during the colder/wetter month versus the warmer/drier months while bacterial loads remain constant, thereby resulting in negative correlations with the measured concentration of microbial indicators. However, positive samples were collected even during winter months, when the highest streamflow values would be expected to dilute discharges, and only weak correlations were found between *E. coli* and flow rate in Callahan Creek (Table 3). Therefore, while seasonal water quality parameters may provide some anecdotal explanation of variability in *Bacteroides*-HF183, *Enterococcus* and *E. coli*, additional factors such as sample collection timing are also likely important (i.e. collection directly following straight pipe discharge).

Correlation of microbial indicators with water quality and flow rate

To investigate further, a correlational analysis was conducted between microbial indicators and several general water quality parameters and flow rate (Spearman's rho values, Table 3). *Bacteroides*-HF183 and conductivity were significantly and positively correlated in Callahan Creek (i.e. Spearman's rho = 0.37). This is consistent with expectations since both parameters can originate from the same source: untreated household waste. However, in Roaring Fork, conductivity and *Enterococcus* were inversely correlated (i.e. Spearman's rho = −0.44). This seems contradictory to expectations and perhaps suggested additional unattributed sources of conductivity (i.e. other than the untreated household waste discharges) in this stream (e.g. additional residences, adjacent railroad tracks).

A weak but significant correlation was observed between *E. coli* and pH in this stream (i.e. Spearman's

Table 2 | Monthly *Bacteroides*-HF183, *Enterococcus* and *E. coli* in Callahan Creek

***Bacteroides*-HF183 (copies/mL)**						***Enterococcus* (copies/mL)**					***E. coli* (MPN/100 mL)**				
Date	**CC10**	**CC09**	**CC06**	**CC04**	**CC01**	**CC10**	**CC09**	**CC06**	**CC04**	**CC01**	**CC10**	**CC09**	**CC06**	**CC04**	**CC01**
4/14	−	−	−	−	−	+	−	+	−	−	0	303	99	99	99
5/14	−	200	+	+	−	−	+	+	+	−	62	2,461	1,099	631	2,846
6/14	−	5,770	913	−	−	−	−	+	−	+	7	303	305	99	99
7/14	−	+	+	+	−	−	1,230	−	−	−	12	3,068	201	1,463	98
8/14	−	−	−	−	−	+	+	+	+	+	259	408	737	730	2,333
1/16	−	+	−	+	−	−	−	+	1,380	+	0	206	236	96	115
2/16	−	+	454	+	−	−	−	+	+	+	5	95	647	2,162	27
3/16	−	−	+	+	−	−	−	−	−	−	0	94	645	912	96
4/16	−	200	−	+	−	+	−	−	−	−	0	55	352	596	846

For *Bacteroides*-HF183 and *Enterococcus*, negative samples (i.e. less than 20 copies/mL) are indicated by '−' and samples with 20–200 copies/mL (i.e. detected but not quantifiable) are indicated by '+'. Quantifiable samples show the number of copies.

Table 3 | Results of Spearman's rho correlation analysis between microbial indicators and water quality parameters or flow

	Temp	Cond	DO	pH	Depth	Flow rate[a]
Callahan Creek						
Bact. HF183	0.02	0.37[b]	−0.16	0.06	0.01	0.11
E. coli	0.24[b]	0.20	−0.25[b]	0.21[b]	−0.23[b]	−0.19[b]
Enterococcus	0.14	0.01	−0.25	–0.01	0.24	−0.16
Roaring Fork						
E. coli	0.26[b]	0.03	−0.22[b]	–0.13	0.13	−0.03
Enterococcus	0.302	–0.44[b]	−0.31[c]	–0.05	0.23	0.10

Bacteroides-HF183 is not shown for Roaring Fork since this marker was undetectable in all samples.
[a]Flow rate data from USGS gauge station 03529500 in the Powell River.
[b]Significant correlation: p value <0.05.
[c]p value $= 0.06$.

rho = 0.21). In Callahan Creek, a slight spike in pH values was observed directly below the untreated household waste discharges, and then not far downstream a return to similar values as in the headwaters. The elevated pH near the waste discharges may be related to household chemical use or high pH in the water source for these residences (relative to the native stream pH). No significant correlation was found between pH and microbial indicators in Roaring Fork.

In both streams *E. coli* and temperature were significantly positively correlated (i.e. Spearman's rho of 0.24 in Callahan Creek and 0.26 in Roaring Fork) which may reflect higher detection during low flow warmer months. In keeping with this, flow rate was identified as an important seasonal variable in Callahan Creek. A negative correlation between the flow rate and *E. coli* (i.e. Spearman's rho = −0.19), as well as depth and *E. coli* (i.e. Spearman's rho = −0.23), was observed for that stream. In both streams, a significant and negative correlation was also observed between *E. coli* and dissolved oxygen (i.e. Spearman's rho of −0.25 in Callahan Creek and −0.22 in Roaring Fork), which could be related to biological and/or chemical oxygen demand contributed by untreated household waste.

The fact that multiple significant correlations were found between *E. coli* and other parameters, whereas the genetic markers were only correlated to conductivity, is likely due to the relatively higher number of quantifiable data points available for *E. coli*. The higher number of stronger correlations observed between microbial indicators and other parameters in Callahan Creek as compared to Roaring Fork may simply be related to the relatively larger number of untreated household waste discharges, and thus presumably higher waste volume.

Correlation analysis of microbial indicators

No positive correlations between *E. coli* and the molecular markers were identified when using only those data points above the LOQ; this is not surprising, as few samples were quantifiable via qPCR for either *Bacteroides*-HF183 or *Enterococcus*. When DNQ points (e.g. 20–200 copies/mL) were extrapolated to values below the LOQ, significant ($p < 0.05$) positive correlations were identified in Callahan Creek between *Bacteroides*-HF183 and *E. coli* ($\rho = 0.4$) and *Enterococcus* and *E. coli* ($\rho = 0.37$). The relationship between *Enterococcus* and *E. coli* ($\rho = 0.43$) was also significant in Roaring Fork (Figure 4).

All of the microbial indicators targeted can originate from human fecal contamination, and therefore a positive correlation between them would be expected; however, previous examinations of potential correlations between culturable indicator bacteria and genetic markers used in source tracking have been inconsistent, which may be a result of the inherent differences in analytical targets (Harwood *et al.* 2014). Culture-based methods only indicate the presence of viable organisms while non-culture based methods may detect non-viable strands of DNA; the two types of targets may have different rates of decay and/or transport mechanisms. In a laboratory study in Gulf of Mexico, Harwood *et al.* (2009) did not find any correlation between

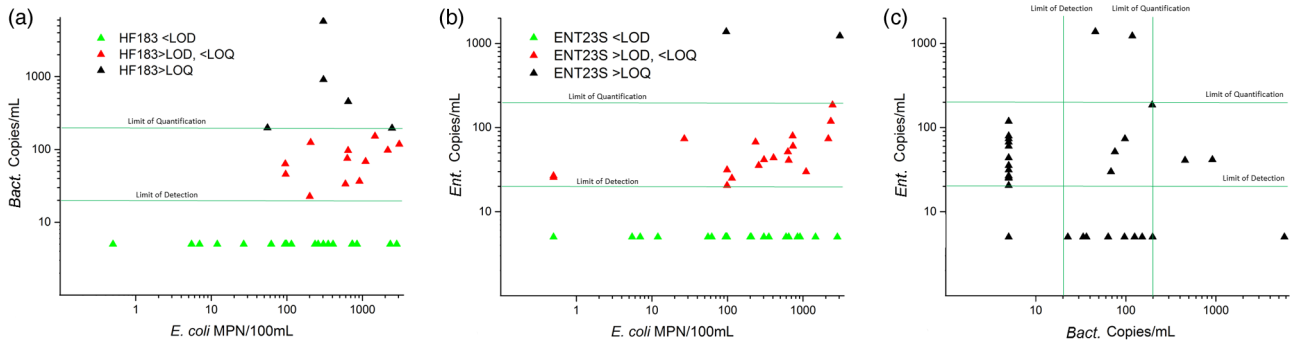

Figure 4 | *Bacteroides*-HF183 vs. *E. coli* (a), *Enterococcus* vs. *E. coli* (b) and *Enterococcus* vs. *Bacteroides*-HF183 (c) in Callahan Creek. Data is extrapolated below the LOQ.

Enterococcus and human associated *Bacteroidales*, nor did Sidhu *et al.* (2013) find a significant correlation between *E. coli* and *Enterococcus* in a field study in commercial and residential catchments. Kapoor *et al.* (2013) found a moderate correlation ($r^2 = 0.17$) between human associated *Bacteroidales* and *E. coli* genetic markers (i.e. uidA gene) downstream from combined sewage overflows in an urban environment. The only attempt at a similar analysis in a rural environment (Tambalo *et al.* 2012) found no correlation between the *Bacteroides* marker BacH and *E. coli*.

Association of microbial levels with untreated household waste via regression model

The sum of the inverse distances to all upstream sources of untreated household waste and the concentration of *E. coli* was significantly and positively correlated ($p < 0.05$) for both Callahan Creek and Roaring Fork (Table 4). There was also a significant, positive relationship between the concentrations of *E. coli* and seasonality for Callahan Creek, though this relationship was not significant for Roaring Fork. Water temperature in the streams and flow in the Powell River (the receiving river for both Callahan Creek and Roaring Fork) were tested in both models but did not show a significant relationship with the concentrations of *E. coli* and therefore were not included. Similar analyses with *Bacteroides*-HF183 or *Enterococcus* as the dependent variables were not significant due to the high level of non-detectable results that skewed the dataset.

Table 4 | Callahan Creek and Roaring Fork regression model coefficients

Variable	Callahan Creek coefficient	Callahan Creek P value	Roaring Fork coefficient	Roaring Fork P value
Intercept	2.69	<0.0001	0.89	<0.0001
Not winter[a]	0.84	<0.0001	N/A	>0.05
Waste discharge value[b]	0.08	<0.0001	0.07	<0.0001

[a]Categorical variable indicates whether the season is not winter, when true the season is not winter, when false the season is winter.
[b]Variable represents the sum of the inverse distances to all known upstream sources of untreated household waste.

The significant waste discharge number variable indicated that with more discharges upstream from a sampling site, the concentration of *E. coli* is likely to be higher, as would be expected. It also indicated that *E. coli* concentrations decrease as distance from discharges increases, which may be attributable to die-off and/or dilution, given the previously mentioned lack of other appreciable inputs of fecal contamination (e.g. wildlife). The model also confirmed that concentrations of *E. coli* are significantly lower during the winter in Callahan Creek, which is consistent with the expectation that these point sources should be diluted by generally higher streamflow values in the winter months and/or lower temperatures may be affecting levels of persistence. However, it is important to recognize that the model does not explain all variability in longitudinal measures of *E. coli* concentrations on Callahan Creek or Roaring Fork ($R^2 = 0.29$ and 0.24, respectively).

CONCLUSIONS

Although the genetic marker *Bacteroides*-HF183 has routinely been detected downstream from highly populated

areas contaminated by human sewage (Seurinck *et al.* 2005; Sidhu *et al.* 2013; Harwood *et al.* 2014), there has been little previous demonstration of its application to rural areas. This study demonstrated that human-specific *Bacteroides*-HF183 can be found in rural watersheds with known upstream direct discharges of untreated household waste (e.g. via straight-pipes). However, given its sporadic recovery, it may be most appropriate to use this marker in conjunction with more traditional culture-based FIB following initial sanitary surveys to identify areas most at risk. Used concurrently with *E. coli* (as has been previously suggested by Harwood *et al.* (2009) and Sidhu *et al.* (2013)), human-specific *Bacteroides* may provide strong evidence of human fecal contamination, and contribute to justification or evaluation of infrastructure improvements. Ideally, epidemiological studies and water quality information at the most common point of exposure (i.e. private drinking water sources) alongside water surveys would allow for more informed decisions on the scale of investment in wastewater infrastructure.

ACKNOWLEDGEMENTS

Funding for this project was provided by the Appalachian Research Initiative for Environmental Science. The views here are those of the authors and do not necessarily represent the views of the sponsoring organization, nor has the sponsoring organization taken any role in the research. Statistical assistance was provided by the Laboratory for Interdisciplinary Statistical Analysis, as well as Karen Kline from the Department of Biological Systems Engineering at Virginia Tech. The Virginia Department of Health provided information on the location of straight pipes and septic systems in the study watersheds.

REFERENCES

Banks, B. A., Jones, A. & Blakeney, A. 2005 Headwaters : a student/faculty participatory research project in an Eastern Kentucky. *J. Appalachian Stud.* **11** (1), 104–132.

Boehm, A. B., Van De Werfhorst, L. C., Griffith, J. F., Holden, P. A., Jay, J. A., Shanks, O. C., Wang, D. & Weisberg, S. B. 2013 Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. *Water Res.* **47** (18), 6812–6828.

Collier, S. A., Stockman, L. J., Hicks, L. A., Garrison, L. E., Zhou, F. J. & Beach, M. J. 2012 Direct healthcare costs of selected diseases primarily or partially transmitted by water. *Epidemiol. Infect.* **140** (11), 2003–2013.

Cook, N., Sarver, E. & Krometis, L. 2015 Putting corporate social responsibility to work in mining communities: exploring community needs for central Appalachian wastewater treatment. *Resources* **4** (2), 185–202.

Denno, D. M., Keene, W. E., Hutter, C. M., Koepsell, J. K., Patnode, M., Flodin-Hursh, D., Stewart, L. K., Duchin, J. S., Rasmussen, L., Jones, R. & Tarr, P. I. 2009 Tri-County comprehensive assessment of risk factors for sporadic reportable bacterial enteric infection in children. *J. Infect. Dis.* **199** (4), 467–476.

Field, K. G. 2002 Fecal source tracking with bacteroides. In: *Microbiological Source Tracking Workshop: Workgroup Findings and Recommendations*, Technical Report 363 (B. B. Bernstein, J. F. Griffith & S. B. Weisberg, eds). Southern California Coastal Water Research Project, Irvine, CA.

Gasteyer, S. & Vaswani, R. 2004 *Still Living Without the Basics in the 21st Century: Analyzing the Availability of Water and Sanitation Services in the United States*. Rural Community Assistance Partnership, Washington, DC. Available from: www.rcap.org/sites/default/files/rcap-files/StillLiving/Still_Living_full.pdf (accessed November 2016).

Glasmeier, A. K. & Farrigan, T. L. 2003 Poverty, sustainability, and the culture of despair: can sustainable development strategies support poverty alleviation in America's most environmentally challenged communities? *Ann. Am. Acad. Pol. Social Sci.* **590** (1), 131–149.

Griffith, J. F., Layton, B. A., Boehm, A. B., Holden, P. A., Jay, J. A., Hagedorn, C., McGee, C. D. & Weisberg, S. B. 2013 *The California Microbial Source Identification Manual: A Tiered Approach to Identifying Fecal Pollution Sources to Beaches*, Technical Report 804, Southern California Coastal Research Project, Costa Mesa, CA.

Hampson, D., Crowther, J., Bateman, I., Kay, D., Posen, P., Stapleton, C., Wyer, M., Fezzi, C., Jones, P. & Tzanopoulos, J. 2010 Predicting microbial pollution concentrations in UK rivers in response to land use change. *Water Res.* **44** (16), 4748–4759.

Harwood, V. J., Brownell, M., Wang, S., Lepo, J., Ellender, R. D., Ajidahun, A., Hellein, K. N., Kennedy, E., Ye, X. & Flood, C. 2009 Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. *Water Res.* **43** (19), 4812–4819.

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K. & Korajkic, A. 2014 Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiol. Rev.* **38** (1), 1–40.

Haugland, R. A., Siefring, S. C., Wymer, L. J., Brenner, K. P. & Dufour, A. P. 2005 Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by

quantitative polymerase chain reaction and membrane filter culture analysis. *Water Res.* **39**, 559–568.

Haugland, R., Oshima, K. & Dufour, A. P. 2013 *Method 1609: Enterococci in Water by Taqman Quantitative Polymerase Chain Reaction (qPCR) with Internal Amplification Control (IAC) Assay*. EPA-820-R-13-005, EPA, Washington, DC.

Herrig, I. M., Böer, S. I., Brennholt, N. & Manz, W. 2015 Development of multiple linear regression models as predictive tools for fecal indicator concentrations in a stretch of the lower Lahn River, Germany. *Water Res.* **85**, 148–157.

Kapoor, V., Smith, S., Santo Domingo, J. W., Lu, T. & Wendell, D. 2013 Correlative assessment of fecal indicators using human mitochondrial DNA as a direct marker. *Environ. Sci. Technol.* **47**, 10485–10493.

Robson, M. & Schneider, D. 2001 Environmental health issues in rural communities. *J. Environ. Health* **63** (10), 16–19.

Scott, T. M., Rose, J. B., Jenkins, T. M., Samuel, R., Lukasik, J. & Farrah, S. R. 2002 Microbial source tracking : current methodology and future directions. *Appl. Environ. Microbiol.* **68** (12), 5796–5803.

Seurinck, S., Defoirdt, T., Verstraete, W. & Siciliano, S. D. 2005 Detection and quantification of the human-specific HF183 Bacteroides 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environ. Microbiol.* **7** (2), 249–259.

Sidhu, J. P. S., Ahmed, W., Gernjak, W., Aryal, R., McCarthy, D., Palmer, A., Kolotelo, P. & Toze, S. 2013 Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *Sci. Total Environ.* **463–464**, 488–496.

Tambalo, D. D., Fremaux, B., Boa, T. & Yost, C. K. 2012 Persistence of host-associated Bacteroidales gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. *Water Res.* **46**, 2891–2904.

Tarricone, P. 1989 Big trouble in little America. *Civil Eng.* **59** (8), 57–59.

USEPA 2012 Recreational Water Quality Criteria. Available from: www.epa.gov/wqc/2012-recreational-water-quality-criteria. (accessed November 2016).

First received 30 November 2016; accepted in revised form 26 April 2017. Available online 26 May 2017