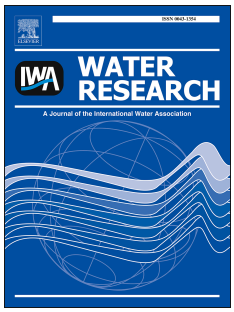
Accepted Manuscript



Spatial and Hydrologic Variation of *Bacteroidales,* Adenovirus and Enterovirus in a Semi-arid, Wastewater Effluent-Impacted Watershed

Dustin G. Bambic, Beverly J. Kildare-Hann, Veronica B. Rajal, Belinda S.M. Sturm, Chris B. Minton, Alexander Schriewer, Stefan Wuertz

PII: S0043-1354(15)00096-2

DOI: 10.1016/j.watres.2015.02.023

Reference: WR 11159

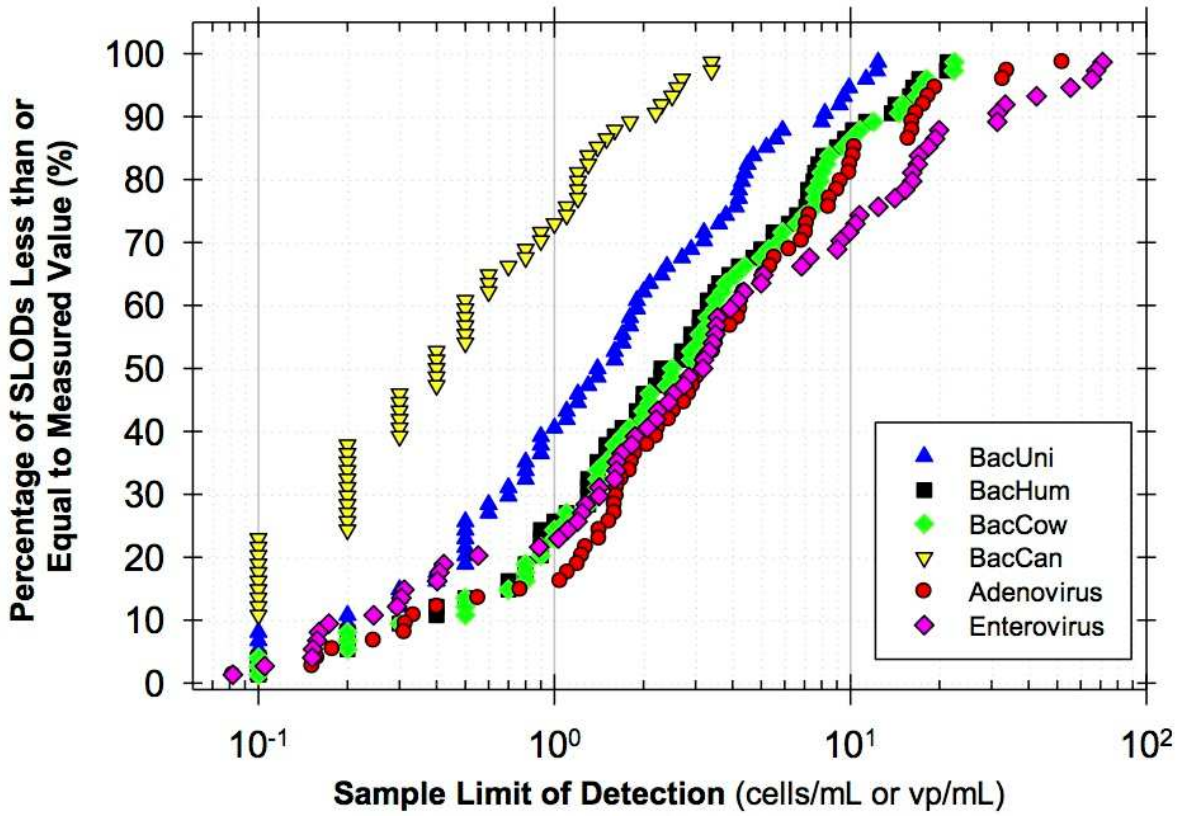
To appear in: *Water Research*

Received Date: 6 July 2014

Revised Date: 8 February 2015

Accepted Date: 10 February 2015

Please cite this article as: Bambic, D.G., Kildare-Hann, B.J., Rajal, V.B., Sturm, B.S.M., Minton, C.B., Schriewer, A., Wuertz, S., Spatial and Hydrologic Variation of *Bacteroidales,* Adenovirus and Enterovirus in a Semi-arid, Wastewater Effluent-Impacted Watershed, *Water Research* (2015), doi: 10.1016/j.watres.2015.02.023.



Spatial and Hydrologic Variation of *Bacteroidales,* Adenovirus and Enterovirus in a Semi-arid, Wastewater Effluent-Impacted Watershed

Dustin G. Bambic[1,2,Φ], *Beverly J. Kildare-Hann*[1,Π], *Veronica B. Rajal*[1,#], *Belinda S. M. Sturm*[1,§], *Chris B. Minton*[2], *Alexander Schriewer*[1], *and Stefan Wuertz*[1,3*]

[1] Department of Civil and Environmental Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616, U.S.A.

[2] Larry Walker Associates, 707 Fourth Street Suite 200, Davis, CA, 95616, U.SA.

[3] Singapore Centre on Environmental Life Sciences Engineering (SCELSE) and School of Civil and Environmental Engineering, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551

* Corresponding author, E-mail: swuertz@ucdavis.edu

[Φ] Current affiliation: Paradigm Environmental, 377 Saint Andrews Drive, Franklin, TN 37069, U.S.A.

[Π] Current affiliation: Carollo Engineers, 2880 Gateway Oaks Drive Suite 300, Sacramento, CA 95833, U.S.A.

[#] Current affiliation: INIQUI (CONICET, UNsa) – Facultad de Ingeniería, Universidad Nacional de Salta. Av. Bolivia 5150, Salta 4400, Argentina.

[§] Current affiliation: Department of Civil, Environmental, and Architectural Engineering, The University of Kansas, 2150 Learned Hall, 1530 W 15th Street Lawrence, Kansas 66045-7609 U.S.A.

ABSTRACT

Bacteroidales and viruses were contemporaneously measured during dry and wet weather conditions at a watershed-scale in a semi-arid watershed impacted by a mixture of agricultural runoff, municipal wastewater effluent and municipal runoff. The results highlight the presence of municipal wastewater effluent as a confounding factor for microbial source tracking (MST) studies, and thus data were segregated into groups based on whether they were impacted by wastewater effluent. In semi-arid environments such as the Calleguas Creek watershed, located in southern California, the relative contribution of municipal wastewater effluent is dependent on hydrology as storm events lead to conditions where agricultural and municipal stormwater dominate receiving waters (rather than municipal wastewater, which is the case during dry weather). As such, the approach to data segregation was dependent on hydrology / storm conditions. Storm events led to significant increases in ruminant- and dog-associated *Bacteroidales* concentrations, indicating that overland transport connects strong non-human fecal sources with surface waters. Because the dataset had a large number of non-detect samples, data handling included the Kaplan-Meir estimator and data were presented graphically in a manner that reflects the potential effect of detection limits. In surface water samples with virus detections, *E. coli* concentrations were often below (in compliance with) the recreational water quality criteria. In fact, sites downstream of direct inputs of municipal wastewater effluent exhibited the lowest concentrations of *Escherichia. coli*, but the highest concentrations of human-associated *Bacteroidales* and highest detection rates of human viruses. The toolkit, comprised of the four *Bacteroidales* assays and human virus assays used, can be successfully applied to inform watershed managers seeking to comply with recreational water quality criteria.

However, care should be taken when analyzing data to account for the effect of non-detect samples, sources with differing microbial viability, and diverging hydrologic conditions.

Keywords: microbial source tracking; *Bacteroidales*; enterovirus; adenovirus; quantitative PCR; total maximum daily load (TMDL)

1. Introduction

Over 12,000 waterbodies in the United States are categorized as impaired by fecal indicator bacteria (FIB) discharges, and have been subject to total maximum daily loads (TMDLs), which describe the water quality improvement strategy to address FIB sources in the watershed (USEPA, 2009). Compliance with recreational water quality (REC) criteria in developed watersheds, both in the U.S. and elsewhere, represents a significant challenge to responsible agencies, as a myriad of non-point bacteria sources contribute to impairment. Some watersheds that are only subject to natural bacteria sources (e.g., birds) have been found to exceed REC criteria (Tiefenthaler et al., 2008), and some waterbodies have been subject to extensive remediation efforts yet exceedances of criteria persist (POLA, 2006). During storm events in urbanized watersheds, which may represent >99% of the annual bacteria discharge (Reeves et al., 2004), loading rates can be extraordinarily high – several times greater than the equivalent daily fecal loading from the entire human population within the watershed (Surbeck et al., 2006). The United States Environmental Protection Agency (USEPA) recently conducted extensive research including epidemiological studies and adopted revised federal REC criteria (Wade et al., 2006,USEPA, 2012). The revised criteria underscore the importance of the type of fecal source when evaluating potential REC health risks. Health risks associated with recreating

in waters impacted by non-human sources can be orders of magnitude less than those with human sources (Colford et al., 2007, Soller et al., 2010).

Given the immense challenges involved with complying with REC criteria, and the importance of fecal source type to the level of risks, watershed managers often desire data regarding the fecal sources that are driving levels of FIB. Collectively referred to as microbial source tracking (MST), a plethora of methods have been developed to characterize the contribution of fecal discharges from different host populations to surface waters and are applied throughout the world (Field and Samadpour, 2007, Santo Domingo et al., 2007, Boehm et al., 2013). The most widely-applied and tested of these approaches targets host-associated 16S rRNA genes of the *Bacteroidales*, and assays based on quantitative PCR (qPCR) can be used to estimate genomic concentrations (Kildare et al., 2007, Shanks et al., 2008, Shanks et al., 2009). Multiple comparison studies have tested and confirmed that, while not 100% sensitive or specific, many *Bacteroidales* markers are sufficiently sensitive and specific for detecting host-associated contamination (Boehm et al., 2013, Layton et al., 2013, Raith et al., 2013, Schriewer et al., 2013), are repeatable/reproducible (Ebentier et al., 2013), and the stable populations required for marker-based MST are present around the globe (Reischer et al., 2013).

Statistical and modeling approaches have been evaluated for using ratios of host-associated to universal *Bacteroidales* markers to quantify the contribution of human versus non-human sources on levels of FIB in watersheds (Harwood, 2007, Wang et al., 2010, Wang et al., 2013, Stoeckel and Russell et al., 2013). Applications of these ratios, which should account for differences in fate and transport characteristics along with the fact that MST assays are imperfect, are emerging as a tool for quantitative MST. Ratios and concentrations are interpreted differently; all host-associated concentrations represent the potential impact of that host

91 population on downstream waters, while host-associated:universal ratios highlight the effect of
92 that host population on the total *Bacteroidales* loading at the monitored site. Suppose that a
93 runoff site has very high levels of the human marker BacHum (when compared to other sites) but
94 a very low ratio of BacHum:BacUni. In this case, the site might pose an elevated risk to
95 recreational users who come into contact with a waterbody impacted by human fecal sources, but
96 on the other hand an agency that is responsible for remediating that site should also target
97 potential non-human sources.

98 To support REC risk assessment, MST assays can be coupled with pathogen assays,
99 particularly those for human viruses (McBride et al., 2013, Harwood et al., 2014). Virus assays
100 with qPCR have been shown to be highly specific for mixed human fecal sources (Harwood et
101 al., 2013), though they are often absent in individual fecal samples (Noble et al., 2003).
102 Enterovirus, a single-stranded RNA virus, has been readily detected with qPCR during several
103 studies of the coastal ocean and coastal watersheds in the western U.S. (Fuhrman et al., 2005;
104 Noble et al., 2006, Viau et al., 2011). Adenovirus, a double-stranded DNA virus, is often
105 detected in these same environments (Choi and Jiang, 2005, Sassoubre et al., 2012), and has been
106 reported to have prolonged survival time and increased resistance to UV treatments (Nwachuku
107 et al., 2005). Prior to this study, no known studies have contemporaneously measured
108 *Bacteroidales* and viruses over the long-term at watershed-scale in waterbodies impacted by a
109 mixture of agricultural runoff, municipal wastewater and municipal stormwater.

110 The objectives of this study were to (i) evaluate the abundance of four validated fecal
111 *Bacteroidales* genetic markers (universal [BacUni], human- [BacHum], dog- [BacCan], and
112 ruminant-associated [BacCow]) in treated and untreated municipal wastewater, (ii) compare
113 quantitative data on host-associated fecal source identifiers based on *Bacteroidales* and human

enteroviruses and adenoviruses with FIB measurements in surface waters, and (iii) utilize the spatial and hydrologic variations of these quantitative MST markers to elucidate the predominant FIB in Calleguas Creek Watershed (CCW), a multi-use coastal watershed in southern California. We hypothesized that concentrations of *Bacteroidales* and viruses would relate to certain types of discharges in the watershed (e.g., agricultural runoff, urban runoff, and municipal wastewater), and expected our results to assist stakeholders with development and implementation of a bacteria TMDL.

To test this hypothesis, the four *Bacteroidales* specific assays (BacUni, BacHum, BacCow, and BacCan), *Escherichia. coli* and human-associated viruses (enteroviruses and adenoviruses) were monitored at multiple CCW sites for one year. To our knowledge, this was the first long-term, watershed-scale study to quantitatively measure *Bacteroidales* and human viruses in water samples. Our approach consisted of combining MST and pathogen methodologies. First, we filtered large volume (100-liter) samples and spiked water samples with surrogates in order to increase the accuracy of quantitation by accounting for DNA losses that occur during filtration and extraction (Rajal et al., 2007a). Then we used qPCR to quantify genomic concentrations of human viruses – adenovirus and enterovirus (Rajal et al., 2007b) – and universal and host-associated *Bacteroidales* markers and their ratios to the universal marker (Kildare et al., 2007). Our approach to data synthesis incorporates tools not often used by MST studies including application of a Montel Carlo model to account for imperfect MST assays and using statistical approaches that robustly account for datasets that are dominated by non-detect results.

2. Methods and Materials

2.1 Watershed Description

Calleguas Creek watershed is subject to a mixture of land uses including agricultural (25%), urban land use (25%) and open space (50%) (Ventura County, 2014). Three subwatersheds of the CCW were monitored: Arroyo Simi, Conejo Creek, and Revolon Slough (**Figure 1**). Arroyo Simi and Conejo Creek were investigated with transects, each having three sampling sites, while Revolon Slough was investigated with a single site. Each of the sampled sites is listed as "impaired" by the State of California due to impacts from *E. coli* sourcess, meaning that a TMDL will be developed for these sites under federal requirements. For both investigated transects, the predominant land uses in the immediate vicinity of the three sampled sites, from upstream to downstream, ranged from open space (limited development) to urban (residential, commercial and industrial land uses) to agricultural (row crops and orchards). Tertiary-treated, chlorine-disinfected effluent ("effluent") from municipal wastewater treatment plants (WWTPs) is discharged at three locations within CCW, upstream of the intermediate site of the Arroyo Simi transect, upstream of the most upstream site of the Conejo Creek transect, and upstream of the intermediate site of the Conejo Creek transect. During dry weather, a majority of the flow rate at locations downstream of the WWTP outfalls is effluent. The land use of Revolon Slough is predominantly irrigated agriculture, though discharges of urban runoff are also present. While there is potential for seepage from wastewater collection systems to flow through storm drains into receiving waters (Haile et al., 1999, Sercu et al., 2009), the wastewater and stormwater systems in CCW are separate and there were no reported sewage spills during sampling events.

2.2 Sample Collection

2.2.1 Collection of Samples from Surface waters and Weather Definition

Samples consisting of 100 liters of surface water were collected in five autoclaved, rinsed, 20-liter polypropylene carboys for pathogen analysis and microbial source tracking. A total of 73 grab samples were generally collected monthly from the seven surface water sites between June 2004 and May 2005 (Kundu et al., 2013). Samples were transported on ice and processed for ultrafiltration as stated below.

In southern California, wet weather is traditionally defined as days with greater than 0.1 inches plus the three following days. For this study, all wet weather samples were collected during active storm events when it was raining and flows were elevated. Dry weather samples were collected after at least one week of non-rain days.

2.2.2. Collection of Primary Influent and Disinfected Effluent Samples from Municipal Wastewater Treatment Plants

Primary influent (minimally-treated sewage at the headworks) samples were collected in sterile 250-mL bottles, and transported on ice to the laboratory on the same day. Samples of disinfected effluent were collected in 2-liter bottles. A total of 14 samples were collected each of primary influent and disinfected effluent.

2.3 Traditional Indicator and Chemical Methods

E. coli concentrations [most probable number per 100mL, MPN per 100mL] were measured according to Standard Method 9223, which is based on chromogenic substrate (IDEXX Colilert). An additional water sample was collected at each site and analyzed for total suspended solids according to Standard Method 2540D [milligram per liter, mg/L].

The following parameters were measured at the time of water collection: water temperature, turbidity, conductivity, dissolved oxygen, and pH (Hach Quanta, Loveland, CO). When measurable but not hazardous (e.g. storm) flow conditions were present, flow rate measurements were performed with an electromagnetic flow meter.

2.4 Processing of Samples for Bacteroidales and Virus Analysis

Details regarding sample processing methods can be found in the Supplemental Information. Viruses and bacteria in 100 liter water samples were concentrated by ultrafiltration using two sequential hollow fiber modules as described previously (Rajal et al., 2007a). Real-time QPCR for surrogate PP7, adenovirus and enterovirus was performed as described in Rajal et al. (2007b). Real-time QPCR for the fecal *Bacteroidales* assays (universal) BacUni, (human-associated) BacHum, (ruminant-associated) BacCow, and (dog-associated) BacCan was performed as described in Kildare et *al.* (2007).

Detection of target nucleic acids by real-time QPCR (which was based on TaqMan assays) was found to be strongly affected by the presence of inhibitors, and the multiple dilution approach was used to address inhibition in all wastewater and surface water samples, as described previously (Rajal *et al*., 2007a). For each sample, a unique sample limit of detection (SLOD) was calculated that accounts for varying inhibition, concentration factors, and filtration recovery (**Figure 2**).

2.5 Statistical Analysis

Statistical analyses were performed using R software version 2.12.0 and the NADA library. Tests were selected based on the fact that our MST and pathogen datasets were highly censored (large number of non-detect samples). In general, non-parametric tests were used that

can handle varying detection limits (without substitution). For summary statistics, estimates were generated using Kaplan-Meir statistics (Kaplan and Meier, 1958), which are commonly used in survival analysis and readily-adaptable to environmental statistics to handle datasets with a large numbers of non-detects, as described by Helsel, 2005 and Helsel, 2012. To highlight the effect of test selection, summary statistics were also generated using regression-on-order statistics using a jackknife procedure based on SLOD for non-detect samples (Shumway et al., 2002) and maximum likelihood estimation (MLE). For comparison of water quality among sites, the Mann-Whitney-Wilcoxon test was used to determine if any of the sites exhibited distributional differences. For all detected significant differences among groups at $p<0.05$, the Kruskal-Wallis test was also reported $p<0.05$. The pairwise p-values were corrected with a correction factor to determine the individual error rate. Both the Bonferroni (highly conservative) and Benjamin & Hocklenberg (B&H; less conservative) correction factors were applied (Helsel, 2012). Correlation analyses were based on tests of Kendall's tau.

A Monte Carlo model developed by Wang et al. (2010) was used to calculate "true" ratios of BacHum:BacUni, BacCow:BacUni, and BacDog:BacUni. These true ratios are referred to as Hum_{ratio}, Cow_{ratio}, and Dog_{ratio}, respectively. The Monte Carlo model accounts for the fact that the markers are not 100% specific and sensitive. The model also accounts for the fact that the raw ratios are not equal to unity for feces and sewage (e.g., BacHum:BacUni is less than one in sewage because there are *Bacteroidales*-specific markers in human feces other than BacHum). Note the fecal samples used for model validation in Wang et al. (2010) were collected from the Calleguas Creek watershed and also used to validate the *Bacteroidales* assays by Kildare et al. (2007) that are applied herein. As such, application of the Monte Carlo model for this study is well-vetted.

3. Results

3.1 Municipal Wastewater

As described in the Supplemental Information, analysis of untreated and treated wastewater samples provided data regarding baseline levels and ratios of *Bacteroidales* in illicit discharges (untreated wastewater) and just downstream of WWTP outfalls (treated wastewater). The BacHum:BacUni ratio was found to be geographically-dependent but relatively stable within a region, and levels of *Bacteroidales* in tertiary-treated, disinfected effluent were found to be relatively high compared to ambient surface water.

3.2 Surface waters

3.2.1 Prevailing Rates in Surface Waters

The prevailing rate, or positive detection frequency, allows for a simple assessment of the predominance of investigated sources. The microbial indicators that were assayed with qPCR during this study varied widely in their prevailing rates (**Table 1)**. The universal *Bacteroidales* marker (BacUni) was detected in all 74 surface water samples (detection frequency of 100%), while enterovirus and adenovirus were only detected in one and eight samples (1% and 11%), respectively. Of the host-associated *Bacteroidales* markers, the human-associated marker (BacHum) was detected most frequently (detection frequency of 90%) and the cow-associated marker (BacCow) least frequently (55%). BacCow was only detected in two of eight samples (20%) from Revolon Slough (4-B), which is dominated by agriculture (row crops, not livestock).

247 The climate of Calleguas Creek watershed is arid, with storms generally being limited to
248 the winter and spring seasons. Annual rainfall is approximately 15 inches. The prevailing rates
249 of all host-associated *Bacteroidales* markers were higher during wet weather, and one of eight
250 (12.5% detection frequency) adenovirus detections occurred during wet weather. The mean
251 estimated percent recovery of each *Bacteroidales* marker from wet weather CCW samples was
252 not significantly different [p = 0.50] from dry weather samples (data not shown). The wide
253 range of detection frequencies for qPCR targets suggests that our corresponding estimates of
254 sample-specific limits of detection (SLOD; **Figure 2** shows the SLODs for each marker) were
255 important to ensure data analysis and interpretation reflects varying SLODs. The ubiquity (i.e.,
256 high frequency of detection) of BacUni, BacHum, and BacCan suggests that qualitative
257 (presence/absence) PCR would not provide much insight with regards to the impact of these
258 bacteria sources on CCW.

259 The fact that a large portion of the collected samples were non-detect suggests that data
260 handling of non-detects can effect report summary statistics. The potential effect of data handling
261 is illustrated in the reported summary statistics for BacCow during dry weather, which was not
262 detected in 39% of samples, for three different approaches: Kaplan-Meier, ROS, and MLE
263 (**Table 2)**. The estimated mean and median by the different tests differ by up to a factor of 4.1
264 (median of Kaplan-Meier versus ROS). The effect of non-detects should also be considered
265 when graphically presenting datasets; the cumulative distribution plots in Figure 3 present the
266 potential range of non-detect samples.

267

268 *3.2.2 Spatial and hydrologic variations in abundance*

269

Variations over space (site-by-site) and hydrology (wet versus dry weather) were assessed to elucidate the characteristics of FIB sources that are impacting the CCW. Concentrations of BacUni, BacCow, BacCan, and *E. coli* were significantly higher during wet weather ($p<0.005$). For surface water data, rather than using the "raw" host-associated:universal marker ratios, a statistical model described by Wang et al. (2010) was used to generate "true" ratios, referred to as Hum_{ratio}, Cow_{ratio}, and Dog_{ratio}, and $Other_{ratio}$. These true ratios reflect conditional probabilities that incorporate the rate of false positives and negatives inherent in MST assays, and provide a more quantitative MST framework compared to the raw ratios. During wet weather (**Table 3**), Hum_{ratio} was significantly lower [$p=0.02$] compared to dry weather while Cow_{ratio} and Dog_{ratio} were significantly higher [$p<0.018$] (**Figure 3** shows the distributions of measured concentrations during dry and wet weather while accounting for non-detects in the dataset).

Due to the significant differences in detection frequencies and abundance during wet versus dry weather, and lower number of samples available for the wet weather condition, spatial variations were only assessed for the dry weather condition (**Figure 4**). The only statistically significant spatial difference in MST marker abundance among sites was for BacHum and Hum_{ratio} at the intermediate Conejo Creek site (9A-B), with at least one being significantly higher [$p<0.05$ with B&H correction] than all other sites except the upstream Conejo Creek site (10-B). None of the other marker-site or ratio-site combinations exhibited significant differences. Virus detections were too rare to reliably assess spatial differences; adenovirus was only detected more than once at the upstream and intermediate sites along the Conejo Creek transect (10-B and 9A-B). Concentrations of adenovirus were significantly higher during wet weather [$p<0.001$]. The only enterovirus detection was at the Revolon Slough site (4-B).

293 *3.2.3 Relationships among measurements*

294 Correlations among the measured parameters were evaluated using the 59 samples
295 collected during dry weather. The 15 wet weather samples were not included because the
296 concentrations of most of the *Bacteroidales* targets, *E. coli* and TSS were significantly higher
297 during wet weather, possibly leading to dry and wet weather "clusters" that could induce less-
298 meaningful correlations. Correlations were based on tests of Kendall's τ, which incorporates
299 SLODs.

300 During dry weather, the *Bacteroidales* measurements were weakly correlated to one
301 another [$\tau > 0.275$, $p < 0.001$], but not to *E. coli* or TSS. The strongest correlation among the
302 *Bacteroidales* markers was for BacUni and BacHum [$\tau = 0.596$, $p < 0.001$]. The fact that
303 *Bacteroidales* markers correlated with one another, but not with *E. coli*, is likely a reflection of
304 the differences in both organism ecology (e.g., facultatively anaerobic versus anaerobic) and
305 quantification methodology (e.g., viability- versus genome-based methods). In addition, this
306 suggests that sites along our transects were subject to discharges from multiple source types
307 simultaneously (e.g., inputs from both cow and human sources occurred). Adenovirus
308 concentrations were not correlated to any other variable.

309

310 **4. Discussion**

311

312 This is the first study known to contemporaneously analyze on a watershed-scale
313 *Bacteroidales* and human virus concentrations in flowing freshwater impacted by municipal
314 wastewater. Overall, our study design was based on evaluating relative differences in universal
315 and host-associated *Bacteroidales* and human virus concentrations over space and time (or

316 weather condition), as elevated host-associated *Bacteroidales* concentrations were assumed to be
317 due to fecal discharges from that host population (e.g., BacCow is due to impacts by cows in the
318 watershed). Such an assumption is warranted based on the efforts taken to develop and validate
319 the applied MST markers and ultrafiltration method (Kildare et al., 2007, Rajal et al., 2007a,
320 Rajal et al., 2007b), but as discussed below, there are a number of confounding factors, such as
321 decay rates and viability, that should be considered when designing, conducting, and analyzing
322 the results of MST studies.

323 Significantly elevated concentrations and ratios of BacCow and BacCan during wet
324 weather, along with significantly lower concentrations and true ratios of BacHum during wet
325 weather (**Figure 3**, **Table 1** and **Table 3**), indicate that non-human sources may be responsible
326 for the significantly elevated BacUni, and perhaps *E. coli*, concentrations that occur during storm
327 events in the CCW. The non-human sources responsible for elevated *Bacteroidales* loading
328 during storm events in CCW are likely contributing to the corresponding exceedances of *E. coli*
329 criteria and should be an important consideration for local stakeholders during TMDL
330 implementation.

331 The fact that sites 9A-B and 10-B along Conejo Creek receive direct inputs of treated
332 WWTP effluent increases the likelihood that non-viable (disinfected) cells may be responsible
333 for the elevated concentrations of BacHum and the detections of adenovirus at these sites.
334 Evidence of the influence of non-viable cells is provided by that fact that sites 9A-B and 10-B
335 exhibited relatively high concentrations of BacHum (genome-based measurement) but relatively
336 low concentrations of *E. coli* (viability-based measurement)(**Figure 4**).

337 Detection frequency of MST and pathogen marker can be quite low, as reflected by
338 adenovirus, enterovirus and BacCow in this study. As shown with the simple comparison of

summary statistics produced by Kaplan-Meier, ROS and MLE methods, the handling of these non-detects can affect findings and conclusions regarding sources. In the field of MST, the handling of non-detects has generally been rudimentary, for example, often replacing non-detect values with one-half the LOD. This study demonstrates that statistical methods and graphical procedures from the field of survival analysis can be readily employed to handle the high rate of non-detects and sample-specific LODs (Helsel, 2005).

4.1 Applicability of host-associated-to-universal Bacteroidales ratios

Bacteroidales concentrations were analyzed both individually and with respect to the true ratios of the host-associated-to-universal marker (**Table 3**). However, additional research is needed before host-associated:universal *Bacteroidales* ratios can be used in a truly quantitative manner (e.g., cows versus dogs) to assess the dominant source(s) to collected water samples (Wang et al., 2013).

To use host-associated:universal *Bacteroidales* ratios for fecal load allocations the following three relationships should be evaluated, discussed further below: (i) the environmental persistence of the host-associated marker when compared to the universal marker and other host-associated markers, (ii) the value of the host-associated to universal ratio and its variability (i.e., stability) in fecal sources, and (iii) if the ratios are to be used for source apportionment of FIB and/or pathogens, then the relative abundance and environmental persistence of *Bacteroidales* versus FIB and/or pathogens. In addition, the specificity and sensitivity of the applied MST assays should be incorporated, which was the purpose of generating true ratios Hum_{ratio}, Cow_{ratio}, and Dog_{ratio} with the Monte Carlo model.

With regards to (i), the persistence of the universal and host-associated *Bacteroidales* markers used for this study is known to be comparable among the four markers studied here in both freshwater and seawater environments. The markers were previously evaluated using flow-through, open-air microcosms in seawater and freshwater under dark and sunlit (diurnal cycle) conditions (Bae and Wuertz, 2009b; Bae and Wuertz, 2012). It was concluded that decay rates among universal (BacUni) and host-associated *Bacteroidales* markers (BacHum, BacCow, and BacCan) were not significantly different, suggesting that differential persistence is not a limiting factor for quantifying relative source contribution.

Relationship (ii) was partially addressed in the present study for untreated sewage discharges; the BacHum:BacUni ratio appeared to be relatively stable in regional sewage (**Table S1**), suggesting that it can be used as a "signature" of human fecal impacts, but the ratio might vary geographically. For fecal discharges from individual humans, however, the BacHum:BacUni ratio was highly variable, possibly limiting its utility for areas subject to individual as opposed to mixed human fecal sources (e.g., areas with homeless persons).

Finally, relationship (iii) is especially critical for studies related to TMDLs – the linkage between *Bacteroidales* and FIB hinges on the relative abundance of *Bacteroidales* in fecal sources and the relative persistence. *Bacteroidales* may be relatively abundant in the fecal samples from a given host, while *E. coli* are relatively low. Based on the fecal samples analyzed during the watershed-specific validation of the qPCR markers applied herein, this is likely the case for seagulls (Kildare et al., 2007), which are more amendable to MST with *Catellicoccus* (Sinigalliano et al., 2013). With regards to relative environmental persistence, the most critical relationship for human risk assessment is relative decay rates of pathogens versus *Bacteroidales*. Walters et al. (2009) found that BacHum exhibited similar survival characteristics to infectious

384 enteroviruses in a sunlight-exposed, sewage-derived microcosm, with both being detected
385 through 8 days of experiment. Bae and Wuertz (2012) found that *Bacteroidales* and
386 *Campylobacter* cells exposed to sunlight exhibited similar survival rates, and host-associated
387 *Bacteroidales* DNA and waterborne pathogen DNA were degraded at comparable rates.

388 Because of these remaining data gaps (and others), the *Bacteroidales* ratios calculated
389 herein were only used in a "within-host" framework among sites and weather conditions, just as
390 with the corresponding concentrations, instead of attempting to quantify the relative contribution
391 of fecal discharges from the different host populations (e.g., BacCow:BacUni is not compared to
392 BacHum:BacUni). Furthermore, it is acknowledged that the *Bacteroidales* ratios do not
393 necessarily reflect the relative abundance of sources of FIB.

394 ***4.2 Influence of treated WWTP effluent on qPCR-based MST***

395 Our results demonstrate that the relatively high concentrations of *Bacteroidales* and
396 human virus cells in WWTP effluent confound qPCR-based MST efforts. MST with qPCR does
397 not distinguish between treated and untreated sources of human feces, which is disconcerting for
398 stakeholders seeking to identify sources of bacteria in an attempt to reduce human health risks in
399 recreational waters. Source trackers should either (a) segregate sites that do and do not receive
400 treated WWTP effluent during statistical analyses of the relative values of BacUni, BacHum and
401 BacHum:BacUni or (b) apply laboratory or field techniques that remove/attenuate non-viable
402 cells from water samples prior to performing qPCR assays.

403 With regards to approach (a), MST study designs and data analysis should evaluate
404 samples collected downstream of WWTP effluent discharges separately from samples collected
405 either upstream of the WWTP discharge or from untreated discharges to the waterbody (e.g.,
406 urban runoff). For instance, considering the sites within the CCW that do not receive treated

WWTP effluent, it should be disconcerting to watershed managers that site 4-B had higher levels of BacHum and $Human_{ratio}$ (and a higher virus detection rate) when compared to site 8-B. However, for most MST applications the concentrations measured at site 8-B should not be directly compared to site 10-B, which receives treated WWTP effluent.

For approach (b) above, the use of propidium monoazide (PMA) with qPCR (PMA-qPCR) has been found to show promise for distinguishing between viable and non-viable *Bacteroidales* cells in sewage and treated WWTP effluent (Bae and Wuertz, 2009a). That PMA-qPCR approach was optimized using the four assays applied during this study, and concentrations of BacUni, BacHum, BacCow, and BacCan measured by PMA-qPCR decayed much more rapidly in freshwater and seawater when compared to concentrations reported by qPCR (Bae and Wuertz, 2009b; Bae and Wuertz, 2012). Future MST and pathogen studies of watersheds influenced by WWTP effluent should consider the application of PMA-qPCR. The CCW study was performed prior to optimization of the PMA-qPCR approach, and thus future applications of this dataset for source assessment should rely on approach (a) describe above (data segregation).

4.3 Occurrence of human viruses in surface waters

In CCW, prevailing rates of adenovirus and enterovirus are much lower (11% and 1%, respectively) when compared to those for *Bacteroidales* (**Table 1**). The much higher detection rate of *Bacteroidales* when compared to human virus may be expected, as *Bacteroidales* are abundant in the feces of a majority of hosts (Menaja et al., 1996), while viruses are only shed by hosts that are infected. The presumed low abundance of human virus was the motivation for collecting 100-liter samples during this study; *Bacteroidales* could be readily detected using

much smaller samples volumes (Dick and Field, 2004). As in this study of the CCW, other viral studies in southern California have detected adenovirus more frequently than enterovirus (Jiang and Chu, 2004; Choi and Jiang, 2005). However, these studies detected human virus more frequently during the winter months, while six of nine (67%) detections in CCW were during the summer months. Like previous viral studies that demonstrated the lack of relationship among virus occurrence and compliance with microbial water quality criteria (Gerba et al., 1979; Jiang et al., 2001; Noble and Fuhrman, 2001; Griffin et al., 2003), eight of nine (89%) human virus detections in CCW occurred when *E. coli* concentrations were below the single sample criteria of 235 MPN/100mL.

As with human *Bacteroidales*, the presence of treated WWTP effluent may confound attempts to identify high-risk human virus sources. Other studies have shown human virus genomes to be readily detected in treated WWTP effluent, while corresponding viable virus titers were typically quite low (Boehm et al., 2005). In the present study, seven of nine (78%) human virus detections occurred in waters dominated by treated WWTP effluent discharges but the viability/infectivity of these viruses are uknown. A recent QMRA study based on the adenovirus concentrations in the CCW reported here, which assumed various proportions of detected viruses were infectious, estimated that human health risks associated with primary and secondary water contact were lower than acceptable thresholds by USEPA (Kundu et al., 2013).

5. Conclusions

This study combined (i) large-volume hollow fiber ultrafiltration of surface water samples using a multiple replicate dilution approach and incorporating estimates of SLODs based on

453 spiked surrogates, (ii) quantification of multiple serotypes of adenovirus and enterovirus, (iii)
454 application of four validated probe-based *Bacteroidales* assays, and (iv) data analysis with a
455 Monte Carlo model and statistical routines that account for non-detects and sample-specific
456 LODs.

457 The results demonstrate that MST based on *Bacteroidales* assays can inform watershed
458 managers seeking to develop strategies to comply with REC criteria, but it is critical to handle
459 non-detects with appropriate statistical methods and to acknowledge the underlying assumptions
460 of qPCR-based MST. While MST shows promise for providing quantitative source
461 apportionment, there are still data gaps including relative decay rates of FIB, *Bacteroidales* and
462 pathogens in effluent-impacted surface waters and lack of qPCR assays for viruses that reflect
463 viable/infective concentrations (e.g., using PMA). Eventually, MST markers may support not
464 only source apportionment but also risk assessment, given additional epidemiological data and/or
465 empirical descriptions of pathogen-*Bacteroidales* relationships.

466

467 **Acknowledgements**

468 The study was funded by Calleguas Creek Watershed (CCW) stakeholders. We thank Dean
469 Messer who helped initiate the project and provided insights into site selection and other
470 logistical aspects. Assistance by Timothy Bartrand with development of the cumulative
471 distribution plots was much appreciated.

472

473 **References**

474 APHA-AWWA-WEF (1998) Standard Methods for the Examination of Water and Wastewater,
475 Washington, DC.

476 Bae, S. and Wuertz, S. (2012) Survival of host-associated bacteroidales cells and their
477 relationship with *Enterococcus* spp., *Campylobacter jejuni*, *Salmonella enterica* serovar
478 Typhimurium, and adenovirus in freshwater microcosms as measured by propidium monoazide-
479 quantitative PCR. Applied and Environmental Microbiology 78(4), 922-932.

480 Bae, S.W. and Wuertz, S. (2009) Discrimination of viable and dead fecal *Bacteroidales* bacteria
481 by quantitative PCR with propidium monoazide. Applied and Environmental Microbiology
482 75(9), 2940-2944.

483 Bernhard, A.E. and Field, K.G. (2000) A PCR assay to discriminate human and ruminant feces
484 on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Applied
485 and Environmental Microbiology 66(10), 4571-4574.

486 Boehm, A.B., Keymer, D.P. and Shellenbarger, G.G. (2005) An analytical model of enterococci
487 inactivation, grazing, and transport in the surf zone of a marine beach. Water Research 39(15),
488 3565-3578.

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

492 Boom, R., Sol, C.J.A., Salimans, M.M.M., Jansen, C.L., Wertheimvandillen, P.M.E. and
493 Vandernoordaa, J. (1990) Rapid and simple method for purification of nucleic acids. Journal of
494 Clinical Microbiology 28(3), 495-503.

495 Choi, S. and Jiang, S.C. (2005) Real-time PCR quantification of human adenoviruses in urban
496 rivers indicates genome prevalence but low infectivity. Applied and Environmental
497 Microbiology 71(11), 7426-7433.

498 Colford, J.M., Wade, T.J., Schiff, K.C., Wright, C.C., Griffith, J.F., Sandhu, S.K., Burns, S.,
499 Sobsey, M., Lovelace, G. and Weisberg, S.B. (2007) Water quality indicators and the risk of
500 illness at beaches with nonpoint sources of fecal contamination. Epidemiology 18(1), 27-35.

501 Dick, L.K. and Field, K.G. (2004) Rapid estimation of numbers of fecal *Bacteroidetes* by use of
502 a quantitative PCR assay for 16S rRNA genes. Applied and Environmental Microbiology 70(9),
503 5695-5697.

504 Dick, L.K., Stelzer, E.A., Bertke, E.E., Fong, D.L. and Stoeckel, D.M. (2010) Relative decay of
505 *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater
506 microcosms. Applied and Environmental Microbiology 76(10), 3255-3262.

507 Domingo, J.W.S., Bambic, D.G., Edge, T.A. and Wuertz, S. (2007) Quo vadis source tracking?
508 Towards a strategic framework for environmental monitoring of fecal pollution. Water Research
509 41(16), 3539-3552.

510 Ebentier, D.L., Hanley, K.T., Cao, Y., Badgley, B.D., Boehm, A.B., Ervin, J.S., Goodwin, K.D.,
511 Gourmelon, M., Griffith, J.F., Holden, P.A., Kelty, C.A., Lozach, S., McGee, C., Peed, L.A.,
512 Raith, M., Ryu, H., Sadowsky, M.J., Scott, E.A., Santo Domingo, J., Schriewer, A., Sinigalliano,

C.D., Shanks, O.C., Van De Werfhorst, L.C., Wang, D., Wuertz, S. and Jay, J.A. (2013) Evaluation of the repeatability and reproducibility of a suite of qPCR-based microbial source tracking methods. Water Research 47(18), 6839-6848.

Field, K. and Samadpour, M. (2007) Fecal source tracking, the indicator paradigm, and managing water quality. Water Research 41(16), 3517-3538.

Fiksdal, L., Make, J.S., LaCroix, S.J. and Staley, J.T. (1985) Survival and detection of *Bacteroides* spp., prospective indicator bacteria. Applied and Environmental Microbiology 49, 148-150.

Fuhrman, J.A., Liang, X.L. and Noble, R.T. (2005) Rapid detection of enteroviruses in small volumes of natural waters by real-time quantitative reverse transcriptase PCR. Applied and Environmental Microbiology 71(8), 4523-4530.

Gerba, C., Goyal, S., LaBelle, R.L., Cech, I. and Bodgan, G.F. (1979) Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters. American Journal of Public Health 69(11), 1116-1119.

Griffin, D.W., Donaldson, K.A., Paul, J.H. and Rose, J.B. (2003) Pathogenic human viruses in coastal waters. Clinical Microbiology Reviews 16(1), 129-143.

Haile, R. W., J. S. Witte, M. Gold, R. Cressey, C. McGee, R. C. Millikan, A. Glasser, N. Harawa, C. Ervin, P. Harmon, J. Harper, J. Dermand, J. Alamillo, K. Barrett, M. Nides, and G.-Y. Wang. (1999) The health effects of swimming in ocean water contaminated by storm drain runoff. Epidemiology (10), 355–363.

Harwood, V.J., Boehm, A.B., Sassoubre, L.M., Vijayavel, K., Stewart, J.R., Fong, T.T., Caprais, M.P., Converse, R.R., Diston, D., Ebdon, J., Fuhrman, J.A., Gourmelon, M., Gentry-Shields, J., Griffith, J.F., Kashian, D.R., Noble, R.T., Taylor, H. and Wicki, M. (2013) Performance of viruses and bacteriophages for fecal source determination in a multi-laboratory, comparative study. Water Research 47(18), 6929-6943.

Harwood, V.J., Staley, C., Badgley, B.D., Borges, K. and 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.

Helsel, D. (2005) More than Obvious: Better Methods for Interpreting Nondetect Data. Environmental Science and Technology. 39 (20), 419A-423A.

Helsel, D. (2012) Statistics for Censored Environmental Data Using Minitab and R. 2nd edition, Wiley.

Jiang, S., Noble, R. and Chui, W.P. (2001) Human adenoviruses and coliphages in urban runoff-impacted coastal waters of Southern California. Applied and Environmental Microbiology 67(1), 179-184.

548 Jiang, S.C. and Chu, W. (2004) PCR detection of pathogenic viruses in southern California urban
549 rivers. Journal of Applied Microbiology 97(1), 17-28.

550 Kaplan, E. L.; Meier, P. (1958). Nonparametric estimation from incomplete observations". J.
551 Amer. Statist. Assn. 53 (282): 457–481.

552 Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B. and Wuertz, S.
553 (2007) 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-
554 specific fecal *Bacteroidales*: A Bayesian approach. Water Research 41(16), 3701-3715.

555 Kreader, C.A. (1998) Persistence of PCR-detectable *Bacteroides distasonis* from human feces in
556 river water. Applied and Environmental Microbiology 64(10), 4103-4105.

557 Kundu, A., McBride, G. and Wuertz, S. (2013) Adenovirus-associated health risks for
558 recreational activities in a multi-use coastal watershed based on site-specific quantitative
559 microbial risk assessment. Water Research 47(16), 6309-6325.

560 Layton, B.A., Cao, Y., Ebentier, D.L., Hanley, K., Balleste, E., Brandao, J., Byappanahalli, M.,
561 Converse, R., Farnleitner, A.H., Gentry-Shields, J., Gidley, M.L., Gourmelon, M., Lee, C.S.,
562 Lee, J., Lozach, S., Madi, T., Meijer, W.G., Noble, R., Peed, L., Reischer, G.H., Rodrigues, R.,
563 Rose, J.B., Schriewer, A., Sinigalliano, C., Srinivasan, S., Stewart, J., Van De Werfhorst, L.C.,
564 Wang, D., Whitman, R., Wuertz, S., Jay, J., Holden, P.A., Boehm, A.B., Shanks, O. and Griffith,
565 J.F. (2013) Performance of human fecal anaerobe-associated PCR-based assays in a multi-
566 laboratory method evaluation study. Water Research 47(18), 6897-6908.

567 McBride, G.B., Stott, R., Miller, W., Bambic, D. and Wuertz, S. (2013) Discharge-based QMRA
568 for estimation of public health risks from exposure to stormwater-borne pathogens in recreational
569 waters in the United States. Water Research 47(14), 5282-5297.

570 Menaja, J.A.G.F., Simoes, F., Sousa, A.T., Moura, P. and Amaraj Collaco, M.T. (1996)
571 *Bacteroides* spp. as alternative indicator organisms: monitoring through PCR 16S- rRNA
572 amplification. OECD Workshop Molecular Methods for Safe Drinking Water, 1996.

573 Noble, R.T., Allen, S.M., Blackwood, A.D., Chu, W.-P., Jiang, S.C., Lovelace, G.L., Sobsey,
574 M.D., Stewart, J.R. and Wait, D.A. (2003) Use of viral pathogens and indicators to differentiate
575 between human and non-human fecal contamination in a microbial source tracking comparison
576 study. Journal Water Health 1, 195-207.

577 Noble, R.T. and Fuhrman, J.A. (2001) Enteroviruses detected by reverse transcriptase
578 polymerase chain reaction from the coastal waters of Santa Monica Bay, California: low
579 correlation to bacterial indicator levels. Hydrobiologia 460, 175-184.

580 Noble, R.T., Griffith, J.F., Blackwood, A.D., Fuhrman, J.A., Gregory, J.B., Hernandez, X.,
581 Liang, X.L., Bera, A.A. and Schiff, K. (2006) Multitiered approach using quantitative PCR to
582 track sources of fecal pollution affecting Santa Monica Bay, California. Applied and
583 Environmental Microbiology 72(2), 1604-1612.

584 Nwachuku, N., Gerba, C.P., Oswald, A. and Mashadi, F.D. (2005) Comparative inactivation of
585 adenovirus serotypes by UV light disinfection. Applied and Environmental Microbiology 71(9),
586 5633-5636.

587 Okabe, S. and Shimazu, Y. (2007) Persistence of host-specific *Bacteroides-Prevotella* 16S rRNA
588 genetic markers in environmental waters: effects of temperature and salinity. Applied
589 Microbiology and Biotechnology 76(4), 935-944.

590 POLA (2006) Inner Cabrillo Beach Bacterial Water Quality Improvement Project
591 Implementation Plan, Port of Los Angeles.

592 Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M.,
593 Reischer, G.H., Farnleitner, A.H., Ervin, J.S., Holden, P.A., Ebentier, D.L., Jay, J.A., Wang, D.,
594 Boehm, A.B., Aw, T.G., Rose, J.B., Balleste, E., Meijer, W.G., Sivaganesan, M. and Shanks,
595 O.C. (2013) Comparison of PCR and quantitative real-time PCR methods for the characterization
596 of ruminant and cattle fecal pollution sources. Water Research 47(18), 6921-6928.

597 Rajal, V.B., McSwain, B.S., Thompson, D.E., Leutenegger, C.M., Kildare, B.J. and Wuertz, S.
598 (2007a) Validation of hollow fiber ultrafiltration and real-time PCR using bacteriophage PP7 as
599 surrogate for the quantification of viruses from water samples. Water Research 41(7), 1411-
600 1422.

601 Rajal, V.B., McSwain, B.S., Thompson, D.E., Leutenegger, C.M. and Wuertz, S. (2007b)
602 Molecular quantitative analysis of human viruses in California stormwater. Water Research
603 41(19), 4287-4298.

604 Reeves, R.L., Grant, S.B., Mrse, R.D., Oancea, C.M.C., Sanders, B.F. and Boehm, A.B. (2004)
605 Scaling and management of fecal indicator bacteria in runoff from a coastal urban watershed in
606 southern California. Environmental Science and Technology 38(9), 2637-2648.

607 Reischer, G.H., Ebdon, J.E., Bauer, J.M., Schuster, N., Ahmed, W., Astrom, J., Blanch, A.R.,
608 Bloschl, G., Byamukama, D., Coakley, T., Ferguson, C., Goshu, G., Ko, G., de Roda Husman,
609 A.M., Mushi, D., Poma, R., Pradhan, B., Rajal, V., Schade, M.A., Sommer, R., Taylor, H., Toth,
610 E.M., Vrajmasu, V., Wuertz, S., Mach, R.L. and Farnleitner, A.H. (2013) Performance
611 characteristics of qPCR assays targeting human- and ruminant-associated bacteroidetes for
612 microbial source tracking across sixteen countries on six continents. Environmental Science and
613 Technology 47(15), 8548-8556.

614 Reischer, G.H., Kasper, D.C., Steinborn, R., Mach, R.L. and Farnleitner, A.H. (2006)
615 Quantitative PCR method for sensitive detection of ruminant fecal pollution in freshwater and
616 evaluation of this method in alpine karstic regions. Applied and Environmental Microbiology
617 72(8), 5610-5614.

618 Russell, T.L., Sassoubre, L.M., Wang, D., Masuda, S., Chen, H., Soetjipto, C., Hassaballah, A.
619 and Boehm, A.B. (2013) A coupled modeling and molecular biology approach to microbial
620 source tracking at Cowell Beach, Santa Cruz, CA, United States. Environmental Science and
621 Technology 47(18), 10231-10239.

622 Sassoubre, L.M., Nelson, K.L. and Boehm, A.B. (2012) Photoinactivation mechanisms of
623 *Enterococcus faecalis* in seawater. Applied and Environmental Microbiology.

624 Schriewer, A., Goodwin, K.D., Sinigalliano, C.D., Cox, A.M., Wanless, D., Bartkowiak, J.,
625 Ebentier, D.L., Hanley, K.T., Ervin, J., Deering, L.A., Shanks, O.C., Peed, L.A., Meijer, W.G.,
626 Griffith, J.F., SantoDomingo, J., Jay, J.A., Holden, P.A. and Wuertz, S. (2013) Performance
627 evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study.
628 Water Research 47(18), 6909-6920.

629 Schriewer, A., Miller, W.A., Byrne, B.A., Miller, M.A., Oates, S., Conrad, P.A., Hardin, D.,
630 Yang, H.H., Chouicha, N., Melli, A., Jessup, D., Dominik, C. and Wuertz, S. (2010) Presence of
631 *Bacteroidales* as a predictor of pathogens in surface waters of the central California coast.
632 Applied and Environmental Microbiology 76(17), 5802-5814.

633 Sercu, B., Van De Werfhorst, L. C., Murray, J., & Holden, P. A. 2009. Storm drains are
634 sources of human fecal pollution during dry weather in three urban southern california
635 watersheds. Environmental Science and Technology, 43(2), 293-298.

636 Seurinck, S., Defoirdt, T., Verstraete, W. and Siciliano, S. (2005) Detection and quantification of
637 the human-specific HF 183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for
638 assessment of human faecal pollution in freshwater. Environmental Microbiology 7(2), 249-259.

639 Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J.R., Noble, R.T., Domingo, J.S., Seifring, S.,
640 Sivaganesan, M. and Haugland, R.A. (2008) Quantitative PCR for detection and enumeration of
641 genetic markers of bovine fecal pollution. Applied and Environmental Microbiology 74(3), 745-
642 752.

643 Shanks, O.C., Kelty, C.A., Sivaganesan, M., Varma, M. and Haugland, R.A. (2009) Quantitative
644 PCR for genetic markers of human fecal pollution. Applied and Environmental Microbiology
645 75(17), 5507-5513.

646 Shumway, R.H., Azari, R.S. and Kayhanian, M. (2002) Statistical approaches to estimating mean
647 water quality concentrations with detection limits. Environmental Science and Technology
648 36(15), 3345-3353.

649 Silkie, S.S. and Nelson, K.L. (2009) Concentrations of host-specific and generic fecal markers
650 measured by quantitative PCR in raw sewage and fresh animal feces. Water Research 43(19),
651 4860-4871.

652 Sinigalliano, C.D., Ervin, J.S., Van De Werfhorst, L.C., Badgley, B.D., Balleste, E., Bartkowiak,
653 J., Boehm, A.B., Byappanahalli, M., Goodwin, K.D., Gourmelon, M., Griffith, J., Holden, P.A.,
654 Jay, J., Layton, B., Lee, C., Lee, J., Meijer, W.G., Noble, R., Raith, M., Ryu, H., Sadowsky,
655 M.J., Schriewer, A., Wang, D., Wanless, D., Whitman, R., Wuertz, S. and Santo Domingo, J.W.
656 (2013) Multi-laboratory evaluations of the performance of *Catellicoccus marimammalium* PCR
657 assays developed to target gull fecal sources. Water Research 47(18), 6883-6896.

658 Soller, J.A., Schoen, M.E., Bartrand, T.A., Ravenscroft, J.E. and Ashbolt, N.J. (2010) Estimated
659 human health risks from exposure to recreational waters impacted by human and non-human
660 sources of faecal contamination. Water Research 44(16), 4674-4691.

661 Stapleton, C.M., Kay, D., Wyer, M.D., Davies, C., Watkins, J., Kay, C., McDonald, A.T., Porter,
662 J. and Gawler, A. (2009) Evaluating the operational utility of a *Bacteroidales* quantitative PCR-
663 based MST approach in determining the source of faecal indicator organisms at a UK bathing
664 water. Water Research 43(19), 4888-4899.

665 Stoeckel, D.M. and Harwood, V.J. (2007) Performance, design, and analysis in microbial source
666 tracking studies. Applied and Environmental Microbiology 73(8), 2405-2415.

667 Surbeck, C.Q., Jiang, S.C., Ahn, J.H. and Grant, S.B. (2006) Flow fingerprinting fecal pollution
668 and suspended solids in stormwater runoff from an urban coastal watershed. Environmental
669 Science and Technology 40(14), 4435-4441.

670 Tiefenthaler, L.L., E.D., S. and Lyon, G.S. (2008) Fecal Indicator Bacteria (FIB) Levels During
671 Dry Weather from Southern California Reference Streams. Southern California Coastal Water
672 Research Project. Technical Report 542.

673 USEPA (2009) Fact Sheet: Total Maximum Daily Loads
674 http://water.epa.gov/lawsregs/lawsguidance/cwa/tmdl/upload/2009_09_09_tmdl_results_22tmdl
675 _progress_factsht.pdf.

676 USEPA (2012) Recreational Water Quality Criteria. Office of Water 820-F-12-058.

677 Vaneechoutte, M., Young, D.M., Ornston, L.N., De Baere, T., Nemec, A., Van Der Reijden, T.,
678 Carr, E., Tjernberg, I. and Dijkshoorn, L. (2006) Naturally transformable *Acinetobacter* sp strain
679 ADP1 belongs to the newly described species *Acinetobacter baylyi*. Applied and Environmental
680 Microbiology 72(1), 932-936.

681 Ventura County Watershed Coalition (2014). www.ventura.org/wcvc/calleguas.htm (accessed
682 November 2014).

683 Viau, E.J., Lee, D. and Boehm, A.B. (2011) Swimmer risk of gastrointestinal illness from
684 exposure to tropical coastal waters impacted by terrestrial dry-weather runoff. Environ Science
685 and Technology 45(17), 7158-7165.

686 Wade, T.J., Calderon, R.L., Sams, E., Beach, M., Brenner, K.P., Williams, A.H. and Dufour,
687 A.P. (2006) Rapidly measured indicators of recreational water quality are predictive of
688 swimming-associated gastrointestional illness. Environmental Health Perspectives 114(1), 24-28.

689 Walters, S.P. and Field, K.G. (2006) Persistence and growth of fecal *Bacteroidales* assessed by
690 bromodeoxyuridine immunocapture. Applied and Environmental Microbiology 72(7), 4532-
691 4539.

692 Wang, D., Farnleitner, A.H., Field, K.G., Green, H.C., Shanks, O.C. and Boehm, A.B. (2013)
693 *Enterococcus* and *Escherichia coli* fecal source apportionment with microbial source tracking
694 genetic markers--is it feasible? Water Research 47(18), 6849-6861.

695 Wang, D., Silkie, S.S., Nelson, K.L. and Wuertz, S. (2010) Estimating true human and animal
696 host source contribution in quantitative microbial source tracking using the Monte Carlo method.
697 Water Research 44, 4760-4775.

698

699

700
701 **Figure legends**
702

703 Fig. 1 – Map of the Calleguas Creek Watershed and monitoring locations (yellow circles).
704 Waterbodies are shown with yellow lines, and the three subwatersheds analyzed along transects
705 during this study are highlighted (pink border, Conejo Creek; blue border, Arroyo Simi; green
706 border, Revolon Slough). The State of California designates 10 distinct reaches in the
707 watershed, which are shown with black bars. Revolon Slough and Conejo Creek do not mix prior
708 to discharge to the estuary. Treated WTP effluent discharges occur upstream of sites 10-B, 9A-B
709 and 7-B. The blue lines in southwestern portion of watershed show major agricultural drainage
710 ditches.

711

712 Fig. 2 – Distribution of sample limits of detection for samples collected from CCW, specific to
713 qPCR markers (blue triangle up, BacUni; black square, BacHum; green diamond, BacCow,
714 yellow triangle down, BacCan; red circle, adenovirus; pink diamond, enterovirus). All samples
715 analyzed by qPCR were 100 liters. The 50% value of the y-axis axis represents the study median.

716

717 Fig. 3 – Cumulative distribution plots of *Bacteroidales*, adenovirus, and *E. coli* concentrations
718 measured during dry weather (blue triangles) and wet weather (red squares). Non-detect (BDL)
719 samples are plotted at the SLOD without fill and the dotted lines show the potential range of
720 percentiles for the sample (samples with low SLODs have a smaller range than samples with
721 high SLODs). The dotted lines in these figures transparently show the potential effect of non-
722 detect samples on estimated percentiles (and summary statistics). The 50^{th} percentile value of
723 the x-axis represents the study median.

724 Fig. 4 – Geometric mean *Bacteroidales* and *E. coli* concentrations (bottom plot) and host-
725 associated to universal *Bacteroidales* ratios (top plot) measured during dry weather along three
726 transects in CCW (Arroyo Simi, left; Conejo Creek, center; Revolon Slough, right). Ratios were
727 calculated for each sample as the host-associated *Bacteroidales* marker concentration (open
728 circle, BacHum; filled triangle, BacCow; open triangle, BacCan) divided by the BacUni
729 concentration (filled circle, BacUni). *E. coli* concentrations are shown in the bottom plot (filled
730 square). Error bars are not shown to allow for plotting within a single figure.

731
732 Table 1 – Summary of Kaplan-Meier statistics for *Bacteroidales* and adenovirus concentrations measured in the CCW, grouped by
733 hydrologic condition (dry versus wet weather).
734

Statistic	BacUni (cell eq/mL)				BacHum (cell eq/mL)				BacCow (cell eq/mL)				BacCan (cell eq/mL)				Adenovirus [1] (genomes/mL)			
	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry
N[2]	15	59			14	59			15	59			15	59			15	58		
N detected	15	59			13	53			13	36			15	47			1	8		
% detected	100	100			92.9	89.8			86.7	61.0			100	79.7			6.7%	13.8%		
Median			2747	1111			95	43			2.5	0.3			36	1.0			NA[3]	NA
Mean			12019	3673			300	500			57	1.6			47	1.7			NA	NA
Std. Dev.			19092	11353			680	2211			151	3.5			55	2.7			NA	NA
10th %ile			384	67			46	0.0001			0.7	0.02			5	0.030			NA	NA
25th %ile			1649	320			66	7.0			1.5	0.03			17	0.120			NA	NA
50th %ile			2747	1111			95	43			2.5	0.33			36	1.0			NA	NA
75th %ile			14284	3089			147	162			44	1.6			59	3.0			NA	NA

90th %ile	40096	6439	341	957	64	4.6	67	4.0	1432	613		

735
736 [1] Enterovirus was detected in 1 of 58 (1.7%) dry weather samples and zero wet weather samples.
737 [2] Number of samples
738 [3] Not applicable because of insufficient number of detects to reliably estimate summary statistic
739
740
741
742

743 Table 2 – Comparison of Kaplan-Meier, ROS, and MLE summary statistics for concentrations of
744 BacCow measured in the CCW during dry weather.

745

Statistic	Statistical method					
	Kaplan-Meier		Regression on Order Statistics		Maximum Likelihood Estimator	
		(cell eq/mL)		(cell eq/mL)		(cell eq/mL)
N[1]	59		59		59	
N detected	36		36		36	
% detected	61.0		61.0		61.0	
Median		0.33		0.08		0.23
Mean		1.6		1.5		3.5
Std. Dev.		3.5		3.5		53
10th %tile		0.02		0.02		0.01
25th %tile		0.03		0.04		0.05
50th %tile		0.33		0.08		0.2
75th %tile		1.6		1.6		1.1
90th %tile		4.6		3.9		4.5

746
747 [1] Number of samples
748

749

750 Table 3 – Summary of Kaplan-Meier statistics for host-associated to universal *Bacteroidales*
751 ratios measured in the CCW, grouped by hydrologic condition (dry versus wet weather).

752

Statistic	Hum_{ratio}		Cow_{ratio}		Can_{ratio}	
	Wet	Dry	Wet	Dry	Wet	Dry
N[1]	14	59	15	59	15	59
N detected	13	53	13	36	15	47
% detected	92.9	89.8	86.7	61.0	100	79.7
Median	0.02	0.05	0.01	0.002	0.07	0.01
Mean	0.03	0.08	0.04	0.02	0.19	0.03
Std. Dev.	0.03	0.07	0.07	0.04	0.25	0.08
10^{th} %tile	0.0001	0.002	0.0001	0.0001	0.017	0.001
25^{th} %tile	0.001	0.01	0.005	0.0001	0.04	0.003
50^{th} %tile	0.018	0.05	0.015	0.002	0.07	0.01
75^{th} %tile	0.04	0.13	0.04	0.01	0.30	0.02
90^{th} %tile	0.08	0.20	0.15	0.07	0.61	0.05

753
754 [1] Number of samples

755

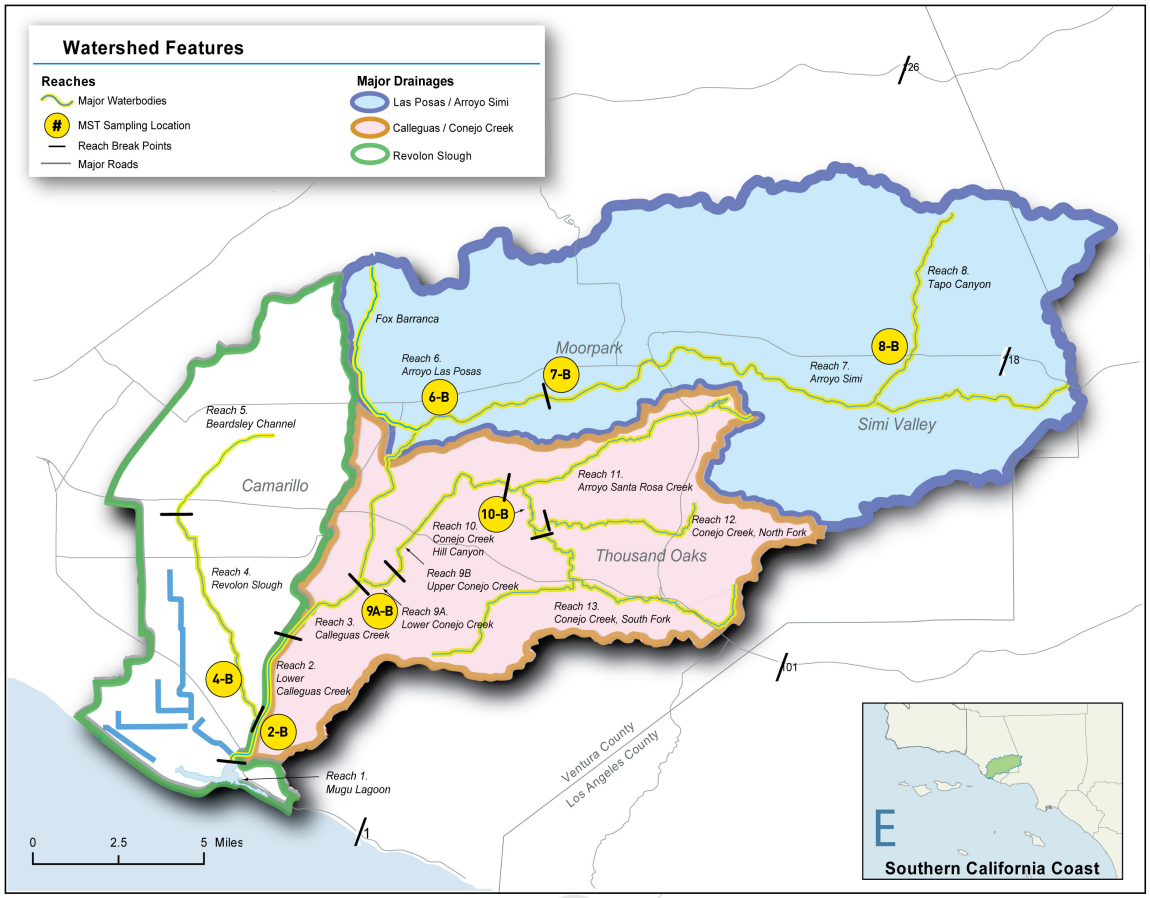


Fig. 1 – Map of the Calleguas Creek Watershed and monitoring locations (yellow circles). Waterbodies are shown with yellow lines, and the three subwatersheds analyzed along transects during this study are highlighted (pink border, Conejo Creek; blue border, Arroyo Simi; green border, Revolon Slough). The State of California designates 10 distinct reaches in the watershed, which are shown with black bars. Revolon Slough and Conejo Creek do not mix prior to discharge to the estuary. Treated WTP effluent discharges occur upstream of sites 10-B, 9A-B and 7-B. The blue lines in southwestern portion of watershed show major agricultural drainage ditches.

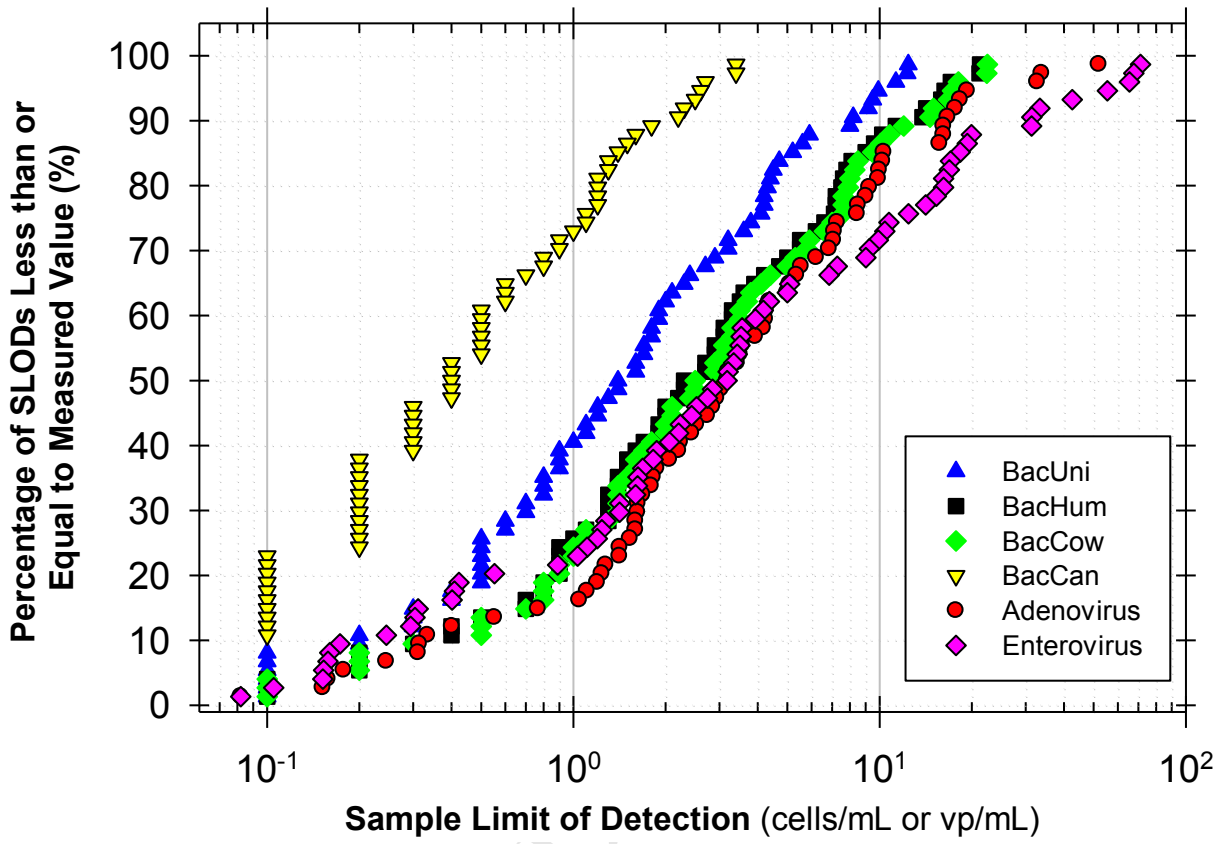
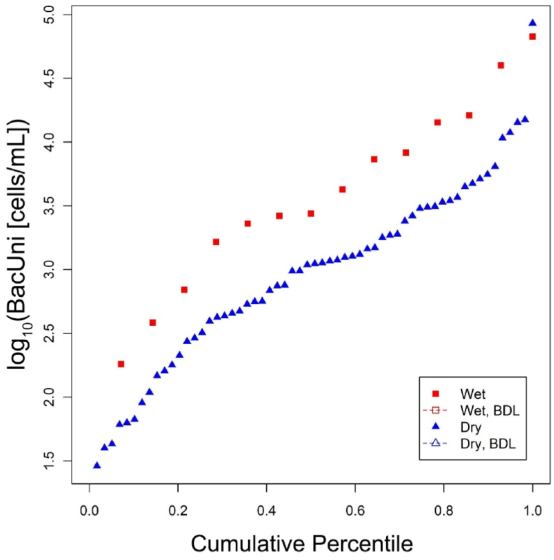
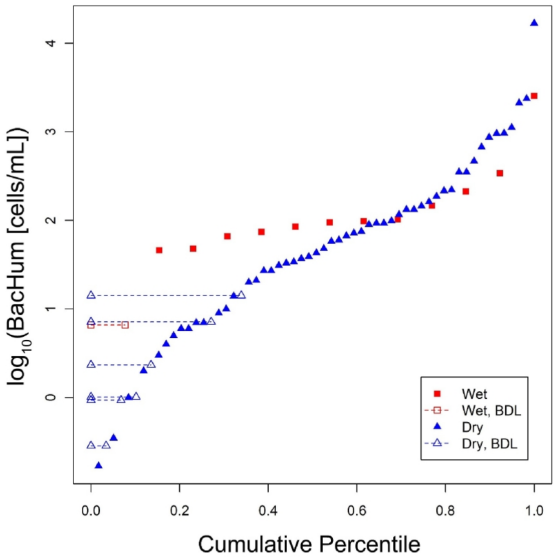
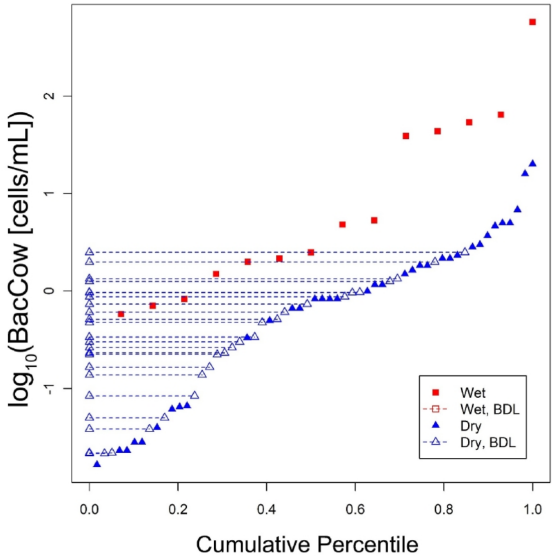
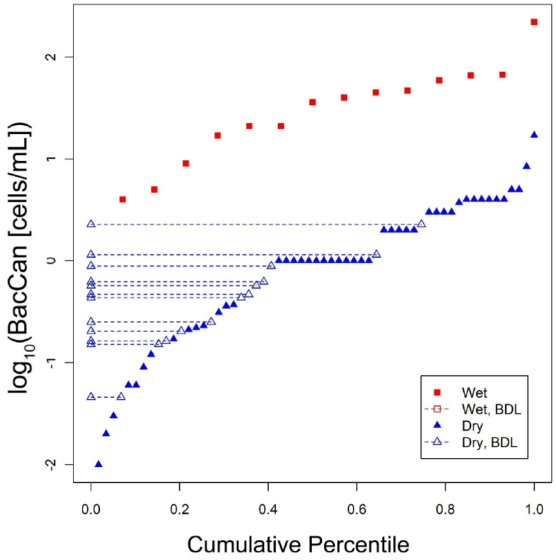


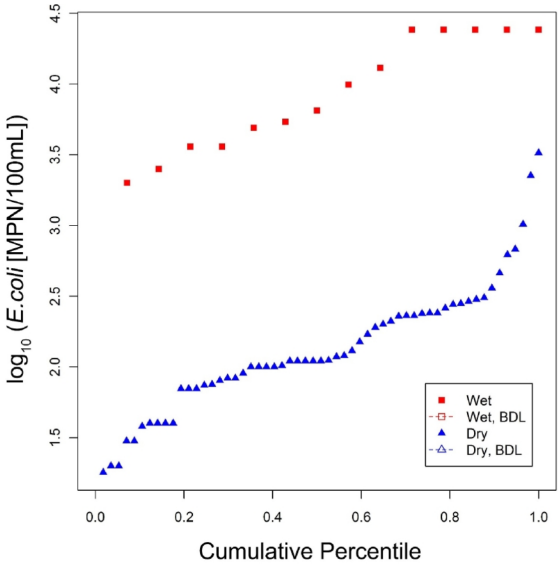
Fig. 2 – Distribution of sample limits of detection for samples collected from CCW, specific to qPCR markers (blue triangle up, BacUni; black square, BacHum; green diamond, BacCow, yellow triangle down, BacCan; red circle, adenovirus; pink diamond, enterovirus). All samples analyzed by qPCR were 100 liters. The 50% value of the y-axis axis represents the study median.


log10(BacUni [cells/mL])
Cumulative Percentile
Wet
Wet, BDL
Dry
Dry, BDL


log10(BacHum [cells/mL])
Cumulative Percentile
Wet
Wet, BDL
Dry
Dry, BDL


log10(BacCow [cells/mL])
Cumulative Percentile
Wet
Wet, BDL
Dry
Dry, BDL


log10(BacCan [cells/mL])
Cumulative Percentile
Wet
Wet, BDL
Dry
Dry, BDL


log10 (E.coli [MPN/100mL])
Cumulative Percentile
Wet
Wet, BDL
Dry
Dry, BDL

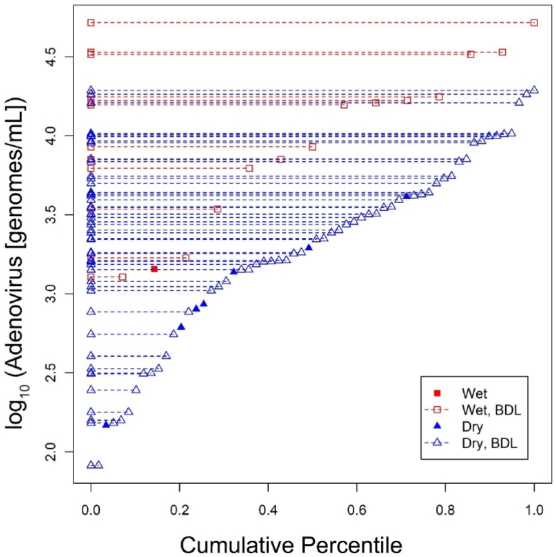

log10 (Adenovirus [genomes/mL])
Cumulative Percentile
Wet
Wet, BDL
Dry
Dry, BDL

Fig. 3 – Cumulative distribution plots of *Bacteroidales,* adenovirus, and *E. coli* concentrations measured during dry weather (blue triangles) and wet weather (red squares). Non-detect (BDL) samples are plotted at the SLOD without fill and the dotted lines show the potential range of percentiles for the sample (samples with low SLODs have a smaller range than samples with high SLODs). The dotted lines in these figures transparently show the potential effect of non-detect samples on estimated percentiles (and summary statistics). The 50^{th} percentile value of the x-axis represents the study median.

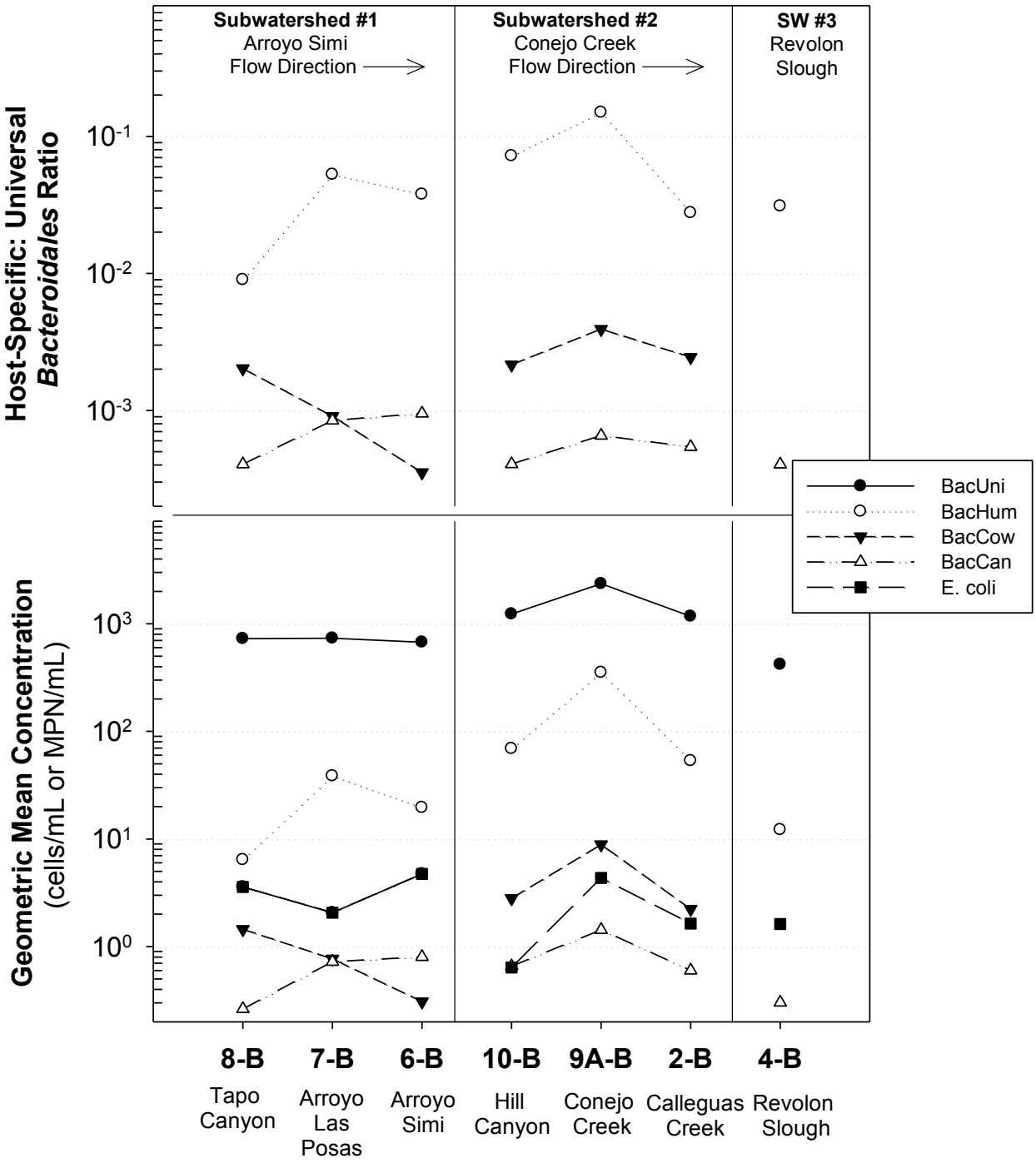

Subwatershed #1
Arroyo Simi
Flow Direction
Subwatershed #2
Conejo Creek
Flow Direction
SW #3
Revolon
Slough
Host-Specific: Universal
Bacteroidales Ratio
10-1
10-2
10-3
Geometric Mean Concentration
(cells/mL or MPN/mL)
10³
10²
10¹
10⁰
BacUni
BacHum
BacCow
BacCan
E. coli
8-B
Tapo
Canyon
7-B
Arroyo
Las
Posas
6-B
Arroyo
Simi
10-B
Hill
Canyon
9A-B
Conejo
Creek
2-B
Calleguas
Creek
4-B
Revolon
Slough

Fig. 4 – Geometric mean *Bacteroidales* and *E. coli* concentrations (bottom plot) and host-associated to universal *Bacteroidales* ratios (top plot) measured during dry weather along three transects in CCW (Arroyo Simi, left; Conejo Creek, center; Revolon Slough, right). Ratios were calculated for each sample as the host-associated *Bacteroidales* marker concentration (open circle, BacHum; filled triangle, BacCow; open triangle, BacCan) divided by the BacUni concentration (filled circle, BacUni). *E. coli* concentrations are shown in the bottom plot (filled square). Error bars are not shown to allow for plotting within a single figure.

Highlights

- Municipal wastewater effluent was confounding factor for microbial source tracking.
- Showed effect of treatment of non-detects on data analysis in monitoring studies.
- Used Monte Carlo simulations to correct *Bacteroidales* concentrations.

SUPPLEMENTAL INFORMATION

Spatial and Hydrologic Variation of *Bacteroidales*, Adenovirus and Enterovirus in a Semi-arid, Wastewater Effluent-Impacted Watershed

Dustin G. Bambic[1,2,Φ], *Beverly J. Kildare-Hann*[1,Π], *Veronica B. Rajal*[1,#], *Belinda S. M. Sturm*[1,§], *Chris B. Minton*[2], *Alexander Schriewer*[1], *and Stefan Wuertz*[1,3*]

[1] Department of Civil and Environmental Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616, U.S.A. Phone: (530) 754 6407, Fax: (530) 752 7872

[2] Larry Walker Associates, 707 Fourth Street Suite 200, Davis, CA, 95616, U.SA. Phone (530) 753-6400

[3] Singapore Centre on Environmental Life Sciences Engineering (SCELSE), School of Biological Sciences, and School of Civil and Environmental Engineering, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551

* Corresponding author, E-mail: swuertz@ucdavis.edu

[Φ] Current affiliation: Paradigm Environmental, 377 Saint Andrews Drive, Franklin, TN 37069, U.S.A. Phone (615) 490-3602

17 Π Current affiliation: Carollo Engineers, 2880 Gateway Oaks Drive Suite 300, Sacramento, CA
18 95833, U.S.A. Phone (916) 565-4888
19 # Current affiliation: INIQUI (CONICET, UNsa) – Facultad de Ingeniería, Universidad Nacional
20 de Salta. Av. Bolivia 5150, Salta 4400, Argentina. Phone and Fax: (54-387) 425 1006
21 § Current affiliation: Department of Civil, Environmental, and Architectural Engineering, The
22 University of Kansas, 2150 Learned Hall, 1530 W 15th Street Lawrence, Kansas 66045-7609
23 U.S.A. Phone: (785) 864-1739, Fax: (785) 864-5379
24
25

OVERVIEW

The following supplemental information is presented below:

- Section 1: Detailed methods for sample processing
- Section 2: Results and brief discussion regarding analysis of wastewater samples for *Bacteroidales*

1. Methods for Processing and Analyzing Samples for Bacteroidales and Virus Analysis

1.1 Concentration of Surface Water and Wastewater Samples

Viruses and bacteria in surface water samples were concentrated by ultrafiltration using two sequential hollow fiber modules as described previously (Rajal et al., 2007a). Briefly, 100 liters of each water sample was sieved and spiked with a known amount of the surrogates PP7 (ATCC 15692-B2), a bacteriophage of *Pseudomonas aeruginosa* (Bolback and Helsenbeck, 2001), and *Acinetobacter baylyi ADP1* (Vaneechoutte et al., 2006). The water (feed, FLS) was pumped through the first ultrafiltration unit with a 50,000 MW membrane cut-off (Microza AHP 2013, Pall Life Sciences, East Hills, NY), until the volume was reduced to 1.5 L. Two elution steps with 0.05M for glycine/NaOH and 0.1% Tween 80 were performed to increase the recovery of microorganisms. The supernatant obtained after centrifuging the retentate from the large filtration module was used as the feed for a second smaller filtration unit (Microza AHP 1013, also 50,000MW cut-off). The final concentrated water sample (RF), 50–100 mL, consisted of the mixture of the eluate from the small unit plus the final retentate. The recovery efficiency of viruses and bacteria in the filtration system was determined based on real-time qPCR of spiked surrogates as described below.

For samples of primary influent and disinfected effluent, once in the laboratory, samples were centrifuged at 4,000 × g for 10 min at 4°C to pelletize the bacterial matter in the sample. The pellet was removed from the bottle with a sterile utensil, and bacterial DNA was extracted immediately.

1.2 Nucleic Acid Extraction and PCR Assays

1.2.1 Nucleic Acid Extraction from Water and Effluent Samples

In order to analyze a large representative fraction of the original sample, 10 mL of the feed or final retentate of the second filtration step were each added to a 200 mL conical plastic centrifuge bottle containing 40 mL of lysis buffer (Boom et al., 1990), and the solution was pulse vortexed 15 times. After a 10-minute incubation period at room temperature, the samples were either stored at -20°C, or extracted immediately. For extraction, 40 mL of absolute ethanol was added, and again pulse vortexed 15 times. The resultant lysate was centrifuged for 10 min at 5,000 × g to pellet solids. The entire supernatant was added to a QIAamp Maxi Spin column (Qiagen,Valencia, CA) and processed according to the manufacturer's instructions. Nucleic acid was eluted twice with 600 µL of DEPC treated water at 4,000 × g for 5 min. The volume of the final eluent was noted for later calculations.

1.2.2 Nucleic Acid Extraction from Primary Influent Samples

Fecal DNAs, and DNA from the resultant pellet of the centrifugation of the primary influent samples were extracted using the QIAamp DNA Stool kit (Qiagen, Valencia, CA) according to the manufacturer's directions. Final eluted volumes were approximately 200 µl.

1.2.3 Real-Time PCR for Viruses and Bacteroidales

Real-time QPCR for surrogate PP7, adenovirus and enterovirus was performed as described in Rajal et al. (2007b). Real-time QPCR for the fecal *Bacteroidales* assays (universal) BacUni, (human-associated) BacHum, (ruminant-associated) BacCow, and (dog-associated) BacCan was performed as described in Kildare et *al.* (2007). For all genomic DNA (gDNA) involving TaqMan probe-based assays (*Bacteroidales* assays and adenovirus), standard amplification conditions were used: 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 min at 60°C. For all RNA-based QPCRreactions, the amplification conditions were: 30 min at 48°C and 10 min at 95°C, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min.

1.2.4 Surrogate Assay for Bacteria using Acinetobacter qPCR assays

Each 25-µL PCR reaction contained 12.5 µL of commercially available QPCR mastermix (Eurogentec) with 400 nM each of forward and reverse primers and 80 nM probe for the respective QPCR system (Schriewer et al., 2010). For all QPCRreactions, 10 µL of the diluted gDNA sample was assayed in a final reaction volume of 25 µL. Four serial dilutions were performed to assess inhibition factors (see below). Cycling conditions were 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C and 60 s at 60°C, using an ABI Prism 7000 (Applied Biosystems).

1.3 Calculation of Sample Limits of Detection

Detection of target nucleic acids by real-time QPCR (based on TaqMan assays) was found to be strongly affected by the presence of inhibitors, and the multiple dilution approach

was used to address inhibition in all wastewater and surface water samples, as described previously (Rajal *et al*., 2007a). Concentrations and sample limits of detection (SLOD) were analyzed according to Rajal et al. (2007a). Each sample has multiple limits of detection, one for each tested marker. Each water sample has a unique set of SLODs due to varying inhibition, concentration factors, and filtration recovery (**Figure 2**). The SLOD (gene copies/milliliter [gc /mL]) values are calculated as follows:

$$SLOD = \left(\frac{1000 \cdot ALOD \cdot I}{R \cdot C_{extr} \cdot C_{filtr} \cdot V_T}\right), \qquad \text{Eq. 1}$$

where ALOD (gc/µL) is the assay limit of detection for the applied assay and specific conditions, I is the dilution factor required to relieve QPCR inhibition [unitless], V_T is the volume of nucleic acid template added to QPCR reaction [µL], and C [unitless] indicates concentration factors for filtration (C_{filtr}) or nucleic acid extraction (C_{extr}). The overall recovery proportion for bacteria and viruses, R, is assessed by measurement of known spike doses of either a bacterial surrogate, *Acinetobacter baylyi* strain ADP1 (Vaneechoutte et al., 2006), previously referenced as *Acinetobacter* sp. strain ADP1 (Juni and Janik, 1969) or the bacteriophage PP7.

The assay limits of detection (ALOD) for adenovirus and enterovirus are presented in Rajal et al. (2007b) and the ALOD for each *Bacteroidales* assay is presented in Kildare et al. (2007).

1.4 Calculation of Virus and Bacterial Concentrations in Water Samples

When the real-time qPCR assays produced a positive reading for the target being assayed, the concentration of target organisms (*Concentration* [gc/mL]) in the original water sample was calculated according to the following equation:

$$Concentration = \left(\frac{T \cdot I}{R \cdot C_{extr} \cdot C_{filtr} \cdot V_T}\right), \qquad \text{Eq. 2}$$

where T is the viral particles or bacterial cells measured in the real-time QPCR reaction [gene copies per reaction for virus assays, or corresponding cells per reaction for bacterial assays] and other variables are as defined for the previous equation.

Since the concentration provided by the standard curve is in units of gene copy numbers measured per volume of reaction, an assumption was made in order to convert the copy numbers found (based on real-time QPCR analysis of a sample) into an estimated concentration of *Bacteroidales* cells for the sample. The assumption, which has been used previously by others (Bernhard and Field, 2000; Seurinck et al., 2005), is that there are an average of five 16S rRNA operons per *Bacteroidales* cell (rRNA Operon Copy Number Collection http://rrndb.cme.msu.edu/rrndb).

2. Results and Brief Discussion regarding analysis of Bacteroidales in Municipal Wastewater

3.1 Untreated Wastewater

Samples of untreated influent to ten municipal wastewater treatment plants (WWTPs) in northern and southern California were tested for universal (BacUni) and human (BacHum) *Bacteroidales* (**Table 1**). Three of these WWTPs were sampled twice, while the others were sampled once. All samples were taken during the months of September and October. Both BacUni and BacHum were detected in 100% of the untreated wastewater samples, with concentrations ranging from 8.1 to 102 x 10^4 cells/ml and from 0.7 to 17 x 10^4 cells/ml, respectively. While variable, these concentrations serve as "expected values" of *Bacteroidales* in illicit discharges. The calculated ratios of BacHum:BacUni may serve as the basis for a quantitative framework to assess host-associated impacts on surface waters (e.g., the contribution of human versus cow fecal inputs; Wang et al., 2010). Ratios of BacHum:BacUni were less variable than the corresponding BacHum concentrations (coefficient of variation [CV]of 0.72 for BacHum:BacUni compared to a CV of 0.88 for BacHum). When compared to analyses of 18 individual human fecal samples (data not shown, CV >2.0), it appears that the BacHum:BacUni ratio in untreated wastewater (a "mixed" human fecal source) is much less variable. It is also noted that analyses of eight cow and eight dog fecal samples from individual hosts yielded highly variable ratios of BacCow:BacUni and Can:BacUni, respectively (data not shown).

The BacHum:BacUni ratio in untreated wastewater samples was significantly different ($p < 0.01$) and also less variable when grouped by watershed – the lower Sacramento River watershed (n = 8) and CCW (n = 4) (**Table S1**). The mean and CV of BacHum:BacUni ratio in the lower Sacramento River Watershed were 0.07 and 0.30, respectively, compared to 0.25 and

0.14 for CCW. Note that the BacUni and BacHum concentrations (as opposed to ratios) in these watersheds were not significantly different. Although a limited number of samples were analyzed, these results suggest that while it may be possible to use the BacHum:BacUni ratio as a "signature" of human-waste impacted waters, the ratio may be geographically-dependent.

3.2 Tertiary-treated, disinfected wastewater

Following the testing of surface waters (discussed in the next section), two samples each of tertiary-treated, disinfected effluent ("treated effluent") from three WWTPs in the CCW were tested for BacUni and BacHum (**Table S2**). The main goal of treated effluent testing was to establish an expected "baseline" of *Bacteroidales* in CCW surface water just downstream of WWTP outfalls. Both BacUni and BacHum were detected in 100% of the effluent samples, with measured concentrations, in units of 10^4 cells per mL, ranging from 0.0022 to 2.5 and 0.0016 to 1.7 cells/ml, respectively. Note that some samples of treated effluent exhibited concentrations similar to those in untreated sewage. One of the three WWTPs – WWTP #2 – exhibited concentrations of BacUni and BacHum that were over two orders of magnitude lower, but the corresponding BacHum:BacUni ratios were the highest. There were no obvious differences in the treatment processes at WWTP #2 that might have lead to significantly lower *Bacteroidales* concentrations.

172 Table S1 – Universal and human-specific *Bacteroidales* concentrations and ratios in municipal
173 WTP influent (sewage) measured across California, grouped by geography. Note that
174 concentration units are 10^4 cell equivalents/ml (cell eq/ml).

175

Watershed Name	**WWTP Location**	**BacUni** (10^4 cell eq/ml)	**BacHum** (10^4 cell eq/ml)	**BacHum: BacUni**
Lower Sacramento (Northern California)	Woodland, CA	8.1	0.7	0.08
	Lincoln, CA	70.4	3.5	0.05
	Woodland, CA	62.6	3.8	0.06
	Davis, CA	15.2	1.4	0.09
	Vacaville, CA	36.3	1.8	0.05
	Lincoln, CA	58.1	2.3	0.04
	Davis, CA	53.8	2.7	0.05
Lower American (Northern California)	Roseville, CA	102.4	17.4	0.17
Calleguas Creek (Southern California)	Moorpark, CA	43.5	7.9	0.23
	Simi Valley, CA	29.3	7.6	0.26
	Hill Canyon, CA	24.4	7.1	0.29
	Camarillo, CA	19.9	4.2	0.21
Oxnard (Southern California)	Oxnard, CA	46.6	2.3	0.05

176
177
178
179

Table S2 – Universal and human-associated *Bacteroidales* concentrations and ratios in tertiary-treated, disinfected effluent collected from three WWTPs in CCW. Note that concentration units are 10^4 cell equivalents/ml (cell eq/ml).

WWTP	**Sample Timing**	**BacUni** (10^4 cell eq/ml)	**BacHum** (10^4 cell eq/ml)	**BacHum: BacUni**
#1	Morning	1.4	0.42	0.29
#1	Afternoon	2.3	0.38	0.17
#2	Morning	0.0018	0.0016	0.89
#2	Afternoon	0.0022	0.0016	0.73
#3	Morning	2.5	0.59	0.23
#3	Afternoon	2.4	1.67	0.69