Section 1.0 Detailed Description of Statistical Analyses and Associated Results

Upon completion of Phase II, Level I sampling, the *Hillsborough River BMAP Fecal Database* contained 241 unique entries consisting of fecal coliform, *E. coli*, and enterococci bacterial concentrations (i.e., log abundance), MST test results, and a host of abiotic parameters (e.g., pH, temperature, conductivity, salinity, turbidity) recorded during 8 sampling events (May-December 2007) at 55 stations throughout 6 WBIDs in west-central Florida. Stations falling within five miles of a daily rainfall gauge (three operated by COT, one by USGS; Map 1) were appended with time-lagged rainfall data (1-day, 2-day, 4-day, 7-day, and 14-day totals) matching the periods of record. All database entries were vetted for missing, duplicate, or erroneous values prior to further analysis. A detailed description of each analysis and the associated findings are provided below.

Section 1.1 Summary Statistics

Each of the six watersheds sampled had sites with indicator organism (IO) concentrations which regularly exceeded the state standard for fecal coliforms (*http://www.dep.state.fl.us/legal/Rules/shared/62-302/62-302.pdf*; Table 1) designated at 400 CFU/100mL as well as the EPA standard for enterococci, designated as 33 CFU/100mL(*http://www.epa.gov/waterscience/criteria/library/ambientwqc/bacteria1986 .pdf*). Geometric mean (+/- SE) and maximum recorded values were calculated for each of the indicator suite organisms (hereafter, 'IO Suite') under flowing conditions and were graphically displayed by station and by WBID (Figures 1 and 2). In most cases, the IO Suite abundance was higher in the sediment than the water samples, with *E. coli* typically showing depressed values compared to the other two IO Suite organisms (*water samples only*). Generally, data coverage was similar amongst five of the six WBIDs, with slightly more fecal coliform samples from Hillsborough River (fecal coliforms; N=100) than the other water bodies (fecal coliforms; New River, N=22; Spartman Branch, N=32; Flint Creek, N=24; Baker Creek, N=24; Blackwater Creek, N=37); other species remained similar among the WBIDs.

Comparison of the mean IO concentrations for each WBID (calculated as the average of counts from all sites with IOs sampled five or more times) using one-way analysis of variance (ANOVA) revealed no significant differences between fecal coliform and *E. coli* concentrations in water samples. In contrast, the mean enterococci concentrations were significantly different between WBIDs ($p < 0.0157$) and enterococci concentrations, on average, were significantly higher in the water samples collected from Spartman Branch than in than those from the Lower Hillsborough River ($p < 0.05$). Differences between mean fecal coliform and *E. coli* levels in the sediments collected from each WBID were not significant when analyzed by ANOVA ($p > 0.05$). In contrast, mean enterococci concentrations in sediments were significantly different between WBIDs ($p = 0.0366$) and were significantly higher in the Lower Hillsborough River than in Flint Creek ($p < 0.05$). All three IO concentrations were significantly positively correlated between

samples collected from water and sediment (Table 1). All post hoc analyses of ANOVA results to determine differences between groups were done by the Tukey post hoc test.

Table 1. Correlation of IO concentrations in water and sediment samples.

IO	r^2	Significance
Fecal Coliforms	0.4905	$p < 0.0001$
E. coli	0.4736	$p < 0.0001$
Enterococci	0.2749	$p = 0.0006$

Only those subsets of data that had adequate coverage, count summaries of IO Suite abundance and concurrence with abiotic, MST, and/or rainfall data, were calculated on a station-specific basis. Only those stations with six or more ($N \geq 6$) observations were used in further study. All stations were considered to be independent and data were not combined for any set of locations.

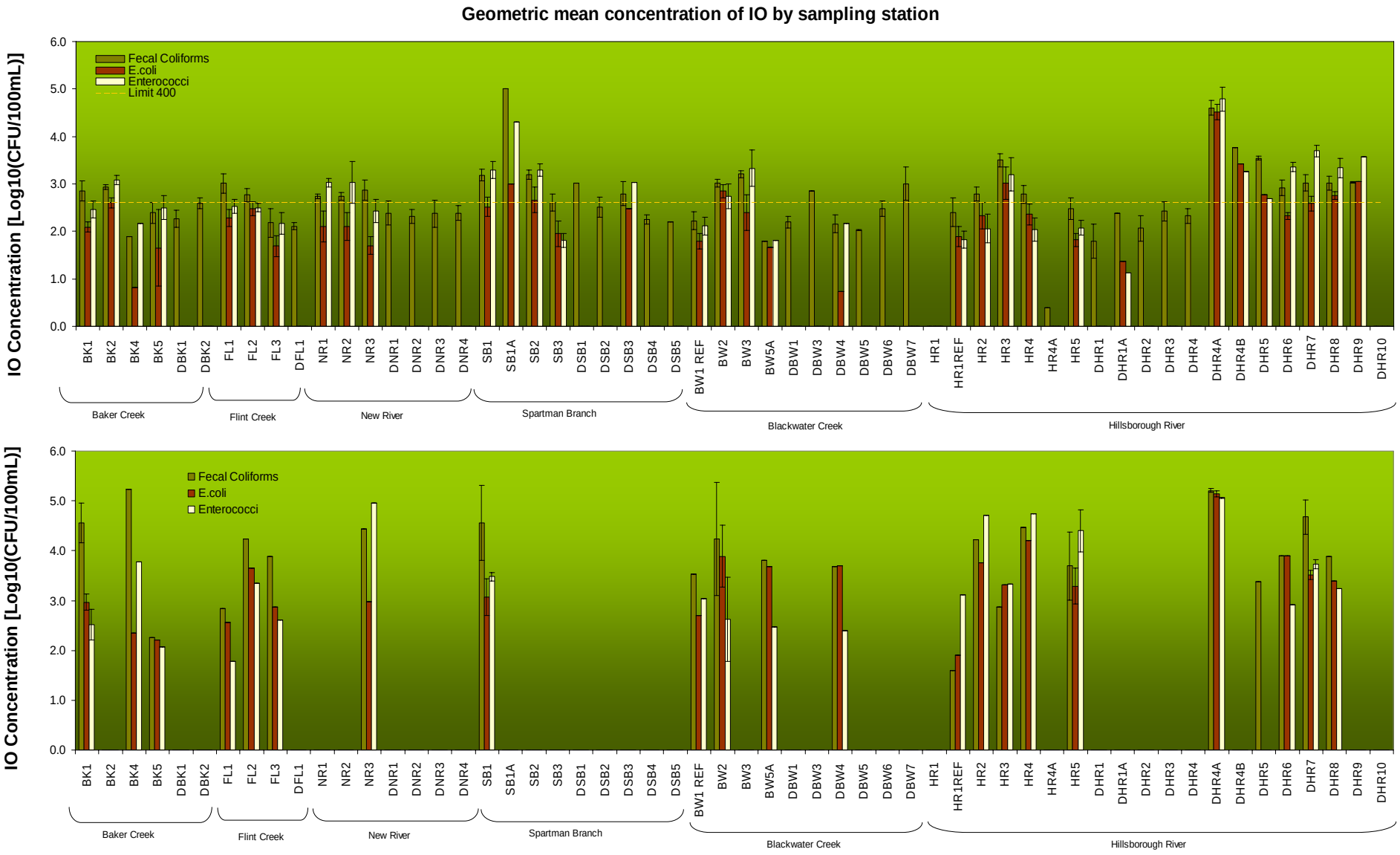


Figure 1. Geometric mean concentration of indicator organisms (IOs) by sampling station in water samples (top graph) and sediment samples (bottom graph). Calculations were performed on all available data collected under flowing conditions.

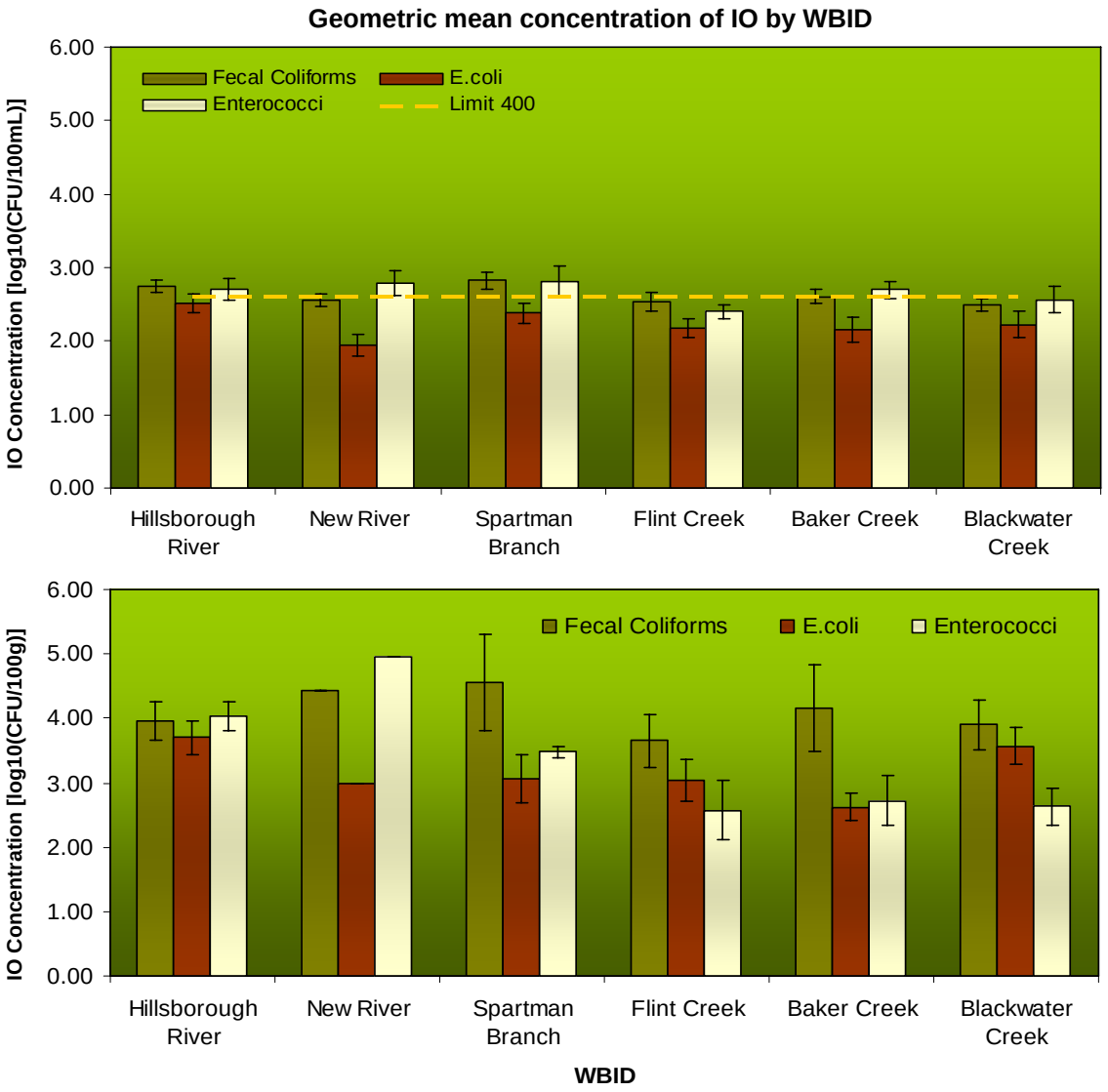


Figure 2. Geometric mean concentration of indicator organisms (IOs) by WBID in water samples (top graph) and sediment samples (bottom graph). Calculations were performed on all available data collected under flowing conditions.

Section 1.2 Indicator Suite Abundance and the Abiotic Environment

IO Suite abundance values from 19 stations representing 5 WBIDs (Blackwater Creek, Flint Creek, Spartman Branch, Baker Creek and Hillsborough River; New River did not meet the $N \geq 6$ criterion for analysis) were then tested for correlations with available water quality parameters; i.e., turbidity, salinity, pH and temperature. The significance of the resulting Pearson Product Moment Correlation Coefficients were independently assessed at an alpha of 0.05; all data were square-root or log-transformed where appropriate. The resulting correlation matrix can be seen in Table 2. Significant linear relationships were found for fecal coliform and *E. coli* abundance with pH at various stations; however, because the correlation itself and the direction of the correlation are not consistent, the detection of significance may merely be the result of chance after numerous tests were performed or the presence of an interacting factor. As expected, fecal coliforms were generally negatively correlated with salinity at several sites within the Lower Hillsborough basin. This is likely due to the high inactivation ("die-off") rate of both fecal coliforms and *E. coli* in saline waters (Anderson et al. 1979, Solic & Krstulovic 1992, Bordalo et al. 2002, Anderson et al. 2005). It should be noted that a small subset of abiotic parameter values were excluded from analysis due to apparent instrument error.

Table 2. Summary of results for the correlative analysis of abiotic parameters with indicator organism concentration for stations within Baker Creek, Flint Creek, Spartman Branch, Blackwater Creek, and the Lower Hillsborough River. Significance was assessed at an alpha of 0.05. Reported values are Pearson correlation coefficients.

Station	Indicator Organism	Temp.	pH	Salinity	Turbidity	Comments
BK1	Fecal Coliforms	-	**-0.808**	-	-	Marginally significant at p=0.052
	E.coli	-	-	-	-	
	Enterococci	X	X	X	X	
BK2	Fecal Coliforms	-	-	-	-	
	E.coli	-	**0.841**	-	-	
	Enterococci	X	X	X	X	
DBK1	Fecal Coliforms	-	-	-	n/a	
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	

Station	Indicator Organism	Temp.	pH	Salinity	Turbidity	Comments
FL2	Fecal Coliforms	-	**-0.892**	-	-	
	E.coli	-	-	-	-	
	Enterococci	-	-	-	-	
FL3	Fecal Coliforms	-	-	X	n/a	
	E.coli	-	n/a	n/a	n/a	
	Enterococci	-	n/a	n/a	n/a	

Station	Indicator Organism	Temp.	pH	Salinity	Turbidity	Comments
SB3	Fecal Coliforms	-	**0.804**	-	n/a	Marginally significant at p=0.054
	E.coli	-	-	-	n/a	
	Enterococci	-	-	-	n/a	
DSB2	Fecal Coliforms	-	-	X	n/a	
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	

Station	Indicator Organism	Temp.	pH	Salinity	Turbidity	Comments
BW2	Fecal Coliforms	-	-	-	-	
	E.coli	-	-	-	n/a	
	Enterococci	-	-	-	n/a	
DBW1	Fecal Coliforms	-	-	-	n/a	
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	
DBW4	Fecal Coliforms	-	**-0.745**	X	n/a	Marginally significant at p=0.054
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	

Station	Indicator Organism	Temp.	pH	Salinity	Turbidity	Comments
HR1REF	Fecal Coliforms	n/a	-	n/a	-	
	E.coli	n/a	-	n/a	-	
	Enterococci	n/a	-	n/a	-	
HR2	Fecal Coliforms	-	-	-	X	
	E.coli	-	-	-	n/a	
	Enterococci	-	-	-	n/a	
HR3	Fecal Coliforms	-	-	**-0.857**	X	
	E.coli	n/a	-	n/a	n/a	
	Enterococci	n/a	-	n/a	n/a	
HR4	Fecal Coliforms	-	-	-	-	
	E.coli	n/a	-	n/a	n/a	
	Enterococci	n/a	-	n/a	n/a	
HR5	Fecal Coliforms	-	X	-	-	
	E.coli	-	X	-	-	
	Enterococci	-	X	-	-	Positive trend with turbidity (r=0.720,p=0.068)
DHR2	Fecal Coliforms	-	-	-	n/a	
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	
DHR3	Fecal Coliforms	-	X	**-0.742**	n/a	
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	
DHR4	Fecal Coliforms	-	-	**-0.908**	n/a	
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	
DHR5	Fecal Coliforms	X	-	**0.853**	n/a	
	E.coli	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	

' - ' = Not significant
X = Not normally distributed
n/a = N<6

Section 1.3 Comparison of Indicator Bacteria

Comparisons among IO concentrations are important as the persistence, as well as the potential for multiplication, *esp*ecially in warm subtropical waters, may differ among species. For example, higher concentrations of enterococccci may be indicative of an older source of pollution (e.g., stormwater runoff) as opposed to a more recent source (e.g., SSO). Although it has not yet been scientifically documented, enterococci appear to demonstrate greater persistence and growth under certain circumstances, such as within enclosures and underground stormwater storage units (personal communication, Dr. Valerie J. Harwood, October 3, 2007).

Correlations amongst IO Suite organisms within each WBID were investigated for all locations with normally-distributed abundance data (log-transformed bacterial counts). In all cases other than New River, bacterial species were positively correlated with each other (Table 3). Pearson Product Moment Correlation Coefficients were particularly high for the Lower Hillsborough River, with greater than 80% of the variation explained by individual correlative pairings.

Table 3. Summary of results for the correlative analysis among IO Suite organism concentrations within each WBID. Significance was assessed at an alpha of 0.05. Reported values are Pearson correlation coefficients.

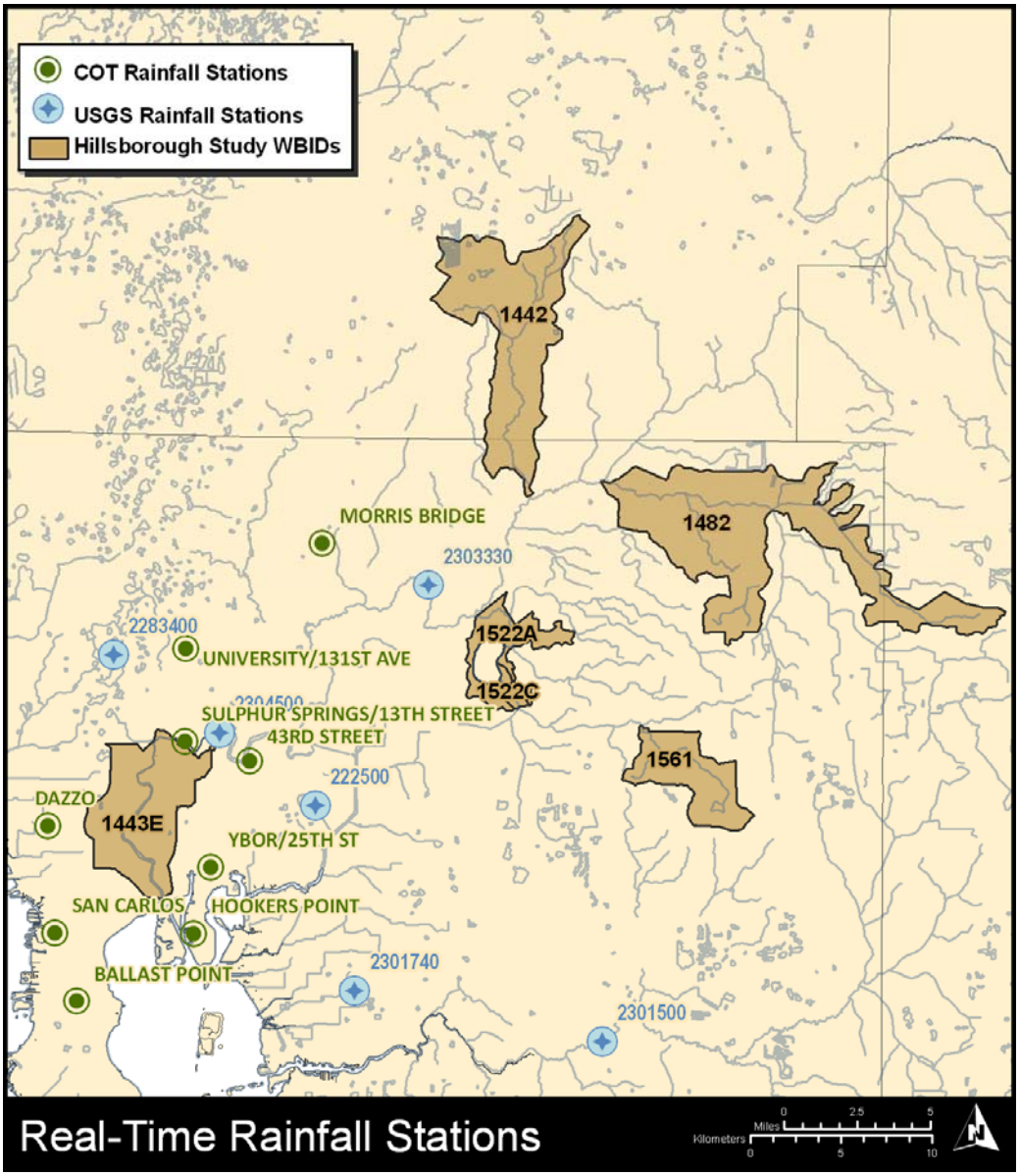
WBID	IO	Fecal Coliforms	E.coli
BK	Fecal Coliforms		
	E.coli	**0.715**	
	Enterococci	-	X (r=0.661,p=0.007)
BW	Fecal Coliforms		
	E.coli	**0.681**	
	Enterococci	**0.745**	**0.622**
FL	Fecal Coliforms		
	E.coli	**0.666**	
	Enterococci	**0.576**	**0.596**
HR	Fecal Coliforms		
	E.coli	**0.903**	
	Enterococci	**0.807**	**0.853**
NR	Fecal Coliforms		
	E.coli	-	
	Enterococci	-	-
SB	Fecal Coliforms		
	E.coli	X	
	Enterococci	X (r=0.747,p=0.001)	**0.743**

' - ' = Not significant
X = Not normally distributed

Section 1.4 Seasonal Evaluation

The occurrence of bacterial numbers with rainfall was investigated for those sampling locations within five miles of an actively-reported rain gauge (Map 1). Nine stations in three WBIDs (Hillsborough River, Flint Creek and Baker Creek) were evaluated using

time-lagged rainfall totals and each of the IO Suite organisms. Data were log-transformed as needed to meet the assumptions of normality; however, low rainfall rates contributed to non-normality in the majority of 1-day to 4-day totals. In many cases, recorded rainfall remained below an inch until week-long blocks were examined. Significant correlations with bacterial abundances were only observed in the 7-day and 14-day totals for stations within the Lower Hillsborough River; these trends were always positive in nature (Table 4). Most commonly, they reflected fecal coliform and enterococci concentrations; however, station HR2 yielded a strong correlation with 14-day rainfall and all three IO Suite organisms (fecal coliform, r=0.838; *E .coli*, r=0.907; enterococci, r=0.900).



Map 1. USGS and COT rainfall stations and Lower Hillsborough River MST project WBIDs.

Table 4. Summary of results for the correlative analysis of rainfall data with indicator organism concentration in Flint Creek, Baker Creek and the Lower Hillsborough River. Significance was assessed at an alpha of 0.05.

Station	Indicator Organism	1d	2d	4d	7d	14d	Comments
HR1REF	Fecal Coliforms	X	X	X	X	-	
	E.coli	X	X	X	X	-	
	Enterococci	X	X	X	**0.712**	-	Despite non-normal 7d rainfall data
HR2	Fecal Coliforms	X	X	-	**0.722**	**0.838**	
	E.coli	X	X	-	-	**0.907**	
	Enterococci	X	X	-	**0.908**	**0.900**	
HR4	Fecal Coliforms	X	X	**-**	**0.698**	-	
	E.coli	X	X	**-**	**-**	-	
	Enterococci	X	X	-	**0.895**	-	
DHR2	Fecal Coliforms	X	X	-	-	-	
	E.coli	n/a	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	n/a	
DHR3	Fecal Coliforms	X	X	-	**0.850**	**0.905**	
	E.coli	n/a	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	n/a	
DHR5	Fecal Coliforms	X	X	-	-	-	
	E.coli	n/a	n/a	n/a	n/a	n/a	
	Enterococci	n/a	n/a	n/a	n/a	n/a	

Note: Low rainfall led to non-normality for 1-day and 2-day totals. Normality was marginal for 4-day totals.

' - ' = Not significant
X = Not normally distributed
n/a = N<6

Station	Indicator Organism	1d	2d	4d	7d	14d	Comments
FL1	Fecal Coliforms	X	X	-	-	-	
	E.coli	X	X	-	-	-	
	Enterococci	X	X	-	-	-	
FL3	Fecal Coliforms	X	X	X	X	-	
	E.coli	X	X	X	X	-	
	Enterococci	X	X	X	X	-	

Note: Low rainfall led to non-normality of data.

' - ' = Not significant
X = Not normally distributed
n/a = N<6

Station	Indicator Organism	1d	2d	4d	7d	14d	Comments
BK1	Fecal Coliforms	X	X	-	-	-	
	E.coli	X	X	-	-	-	
	Enterococci	X	X	X	X	X	

Note: Low rainfall led to non-normality of data.

' - ' = Not significant
X = Not normally distributed
n/a = N<6

Section 1.5 Multivariate Analysis of Indicator Suite Abundance and MST Results

Multivariate analyses of bacterial community data were made using the PRIMER 6 (Windows XP) software for analysis of similarities (ANOSIM) and similarity percentages (SIMPER). The multidimensional scaling (MDS) of sample composition was based on resemblance matrices of Bray-Curtis similarities (bacterial counts were log-transformed). Global R statistics were assessed for significant MST effects (presence or absence by marker type) at $p = 0.05$. All data from both "fixed" and "flexible" stations were utilized after being filtered for missing values. Significant clusters based on MST test results (i.e., 'hits') were further examined for factors contributing to these differences using the SIMPER procedure. These approaches utilize a host of distribution free, primarily permutation-based tests, to quantify differences in multivariate datasets. A lack of restrictive assumptions make them particularly versatile in the assessment of environmental data and provide a unique alternative to more traditional parametric testing, for which many of PRIMER's routines are direct analogs. Therefore, the following analyses ask questions about sample positions in a multivariate space, thereby

incorporating considerably more of the acquired data into the resulting test statistic than would standard univariate approaches.

The effect of MST results on per-sample bacterial communities were found to be significant for human-specific *Bacteroides* (hereafter, 'human *Bacteroides*') at 'all WBIDs' and the Lower Hillsborough River; and for ruminant-specific *Bacteroides* (hereafter, 'ruminant *Bacteroides*') at Blackwater and Baker Creeks (Table 5; Figure 3). These effects were relatively small with Global R values of less than 0.537; however, SIMPER identified *E. coli* concentration as a major driver of group dissimilarity in each of the significant ruminant *Bacteroides* tests (Figure 4). In both cases, higher *E. coli* levels were correlated with the detection of the ruminant marker.

Table 5. Summary of the results showing effect of MST results on per-sample bacterial communities. Global R statistics were assessed for significant MST effects at p = 0.05. All data were filtered for missing values and grouped by WBID prior to analysis. HPyV = human polyomavirus. *esp* = enterococcal surface protein.

WBID	HPyV	Human Bacteroides	*esp*	Ruminant Bacteroides	Comments
All	-	**0.145 (A)**	-	-	
HR	-	**0.115 (B)**	-	-	
BW	-	-	-	**0.537 (C)**	SIMPER dissimilarity driven by E.coli (46.81%)
NR	-	n/a	n/a	-	
FL	-	-	-	n/a	
BK	-	-	-	**0.354 (D)**	SIMPER dissimilarity driven by E.coli (45.54%)
SB	-	n/a	-	-	

Significant global R values (ANOSIM) presented in bold; letter denotes MDS plot
' - ' = Not significant
'n/a' = no positive hits

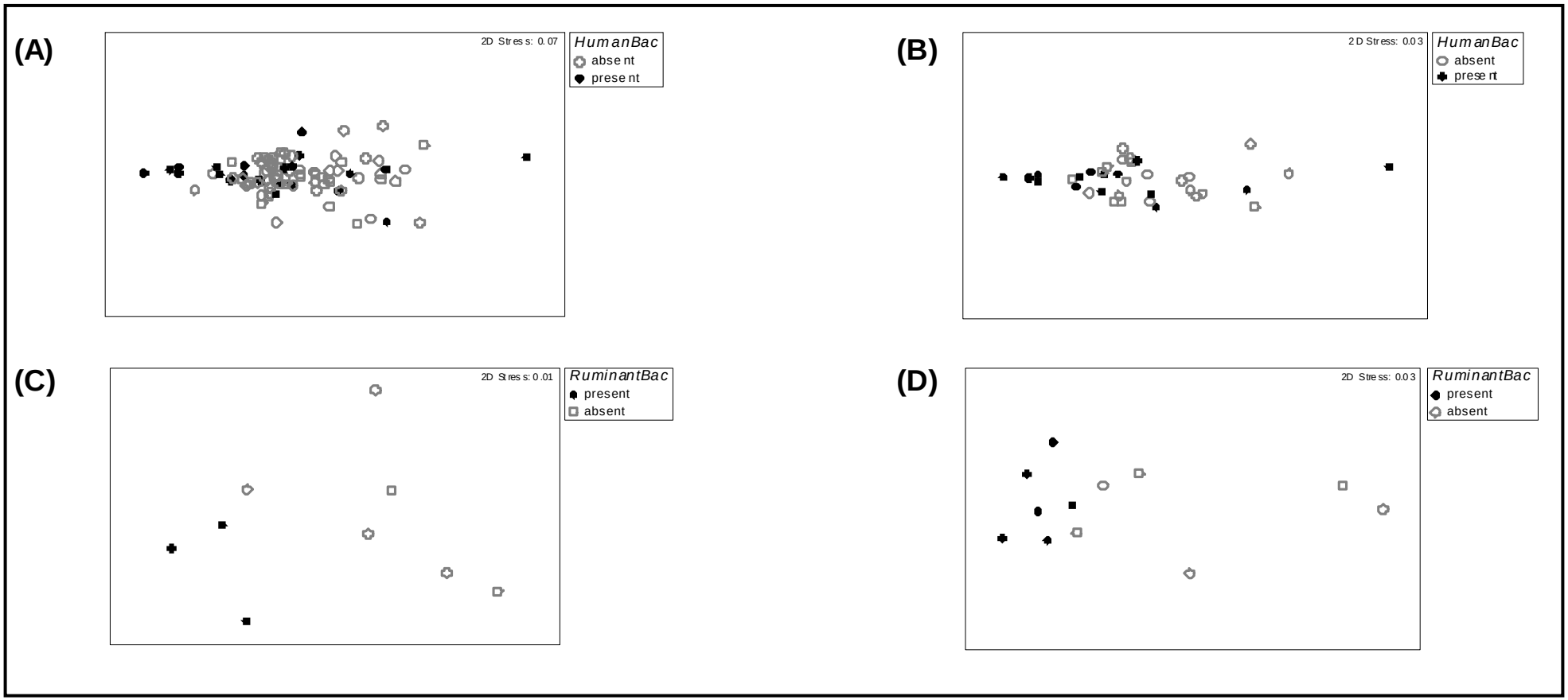
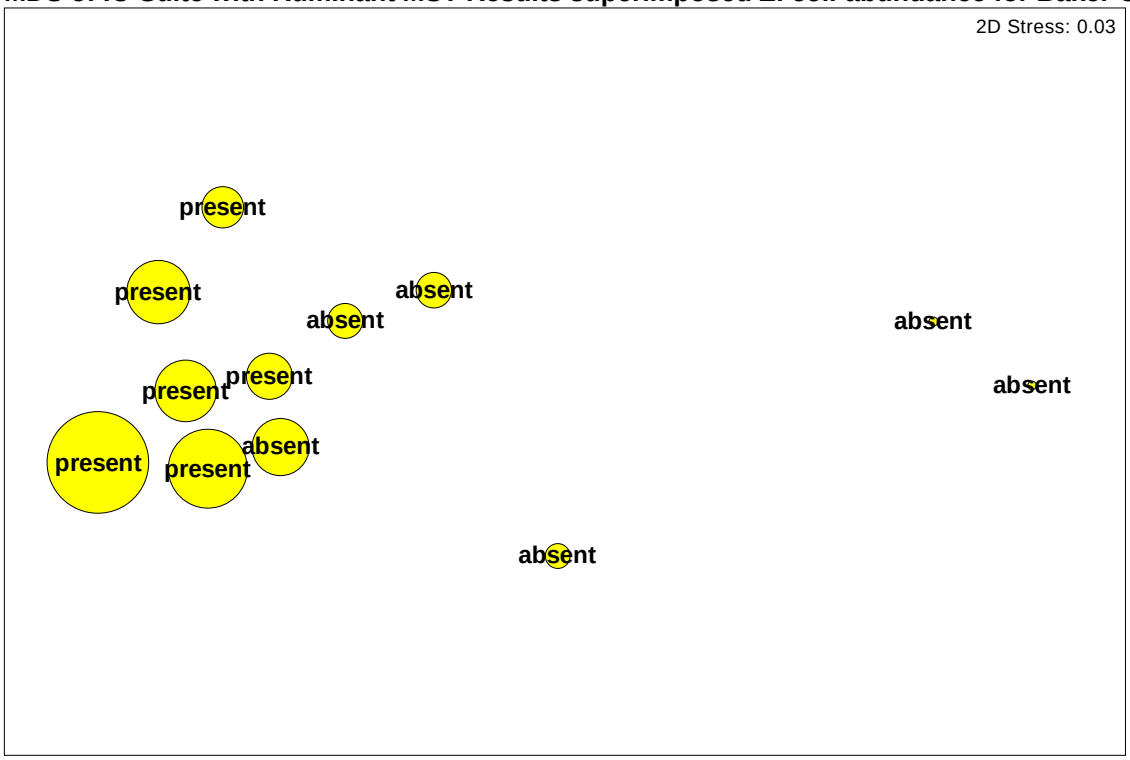


Figure 3. The effect of MST results on per-sample bacterial communities were found to be significant for human *Bacteroides* at 'all WBIDs' (A) and the Lower Hillsborough River (B); and for ruminant *Bacteroides* at Blackwater (C) and Baker (D) Creeks. Global R statistics were assessed for significant MST effects at p = 0.05. All data were filtered for missing values and grouped by WBID prior to analysis.

(A) MDS of IO Suite with Ruminant MST Results superimposed *E. coli* abundance for Baker Creek



(B) MDS of IO Suite with Ruminant MST Results superimposed *E. coli* abundance for Blackwater Creek

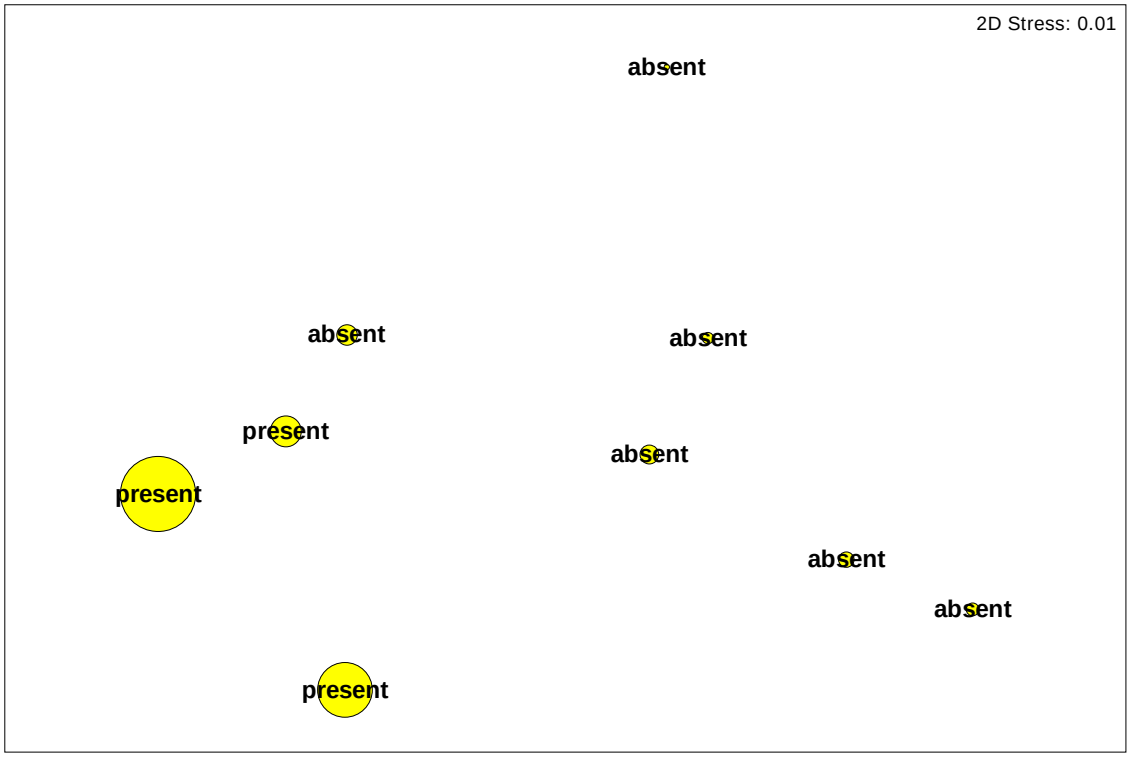


Figure 4. The effect of MST results (presence or absence by marker type) on per-sample bacterial communities was found to be significant for ruminant *Bacteroides* at (A) Baker Creek and (B) Blackwater Creek. SIMPER identified *E. coli* concentration as a major driver of group dissimilarity in both instances where higher concentrations of bacteria correlated with the detection of the ruminant marker.

A chi-squared goodness of fit test was used to further investigate the relationship between fecal coliform concentration and the MST test results (i.e., presence or absence for human polyomavirus (HPyV), *esp*, human *Bacteroides*, ruminant *Bacteroides*, and horse *Bacteroides* markers). Theoretical ratios of positive 'hits' were calculated from the number of MST sampling events with greater than 400 CFU/100mL and were recorded on a per WBID basis. These ratios were then applied to the number of MST samples (n) to produce the expected counts for positive marker occurrence. Since the null hypothesis for each WBID was constructed based on MST concordance with the state-recognized 400 CFU/100mL concentration level, a significant chi-squared test statistic (df=1, $P<0.05$) indicated a departure from this relationship. Thus, chi-squared sums less than the critical value (alpha = 0.05) for a distribution with 1-df signified that MST hits occurred in the same ratios as did samples breaching the 400 CFU/100mL limit.

Evidence for this relationship was found for four of 18 tests; however, 17 test combinations (MST versus WBID) had greater than 20% of expected values under five and were subsequently removed from the analysis. Those MST tests that showed some relation to the 400-limit were HPyV at Flint Creek ($X^2=2.79$, df=1, $P=0.094$), human *Bacteroides* at Blackwater Creek ($X^2=1.33$, df=1, $P=0.249$), and *esp* at Blackwater Creek ($X^2=3.00$, df=1, $P=0.083$) and Flint Creek ($X^2=2.65$, df=1, $P=0.105$). No other individual WBIDs, or all WBIDs combined, showed a significant relationship with the 400-limit.

Section 1.6 Co-occurrence of MST Markers

The use of several MST markers for human contamination is one means of increasing the confidence in results. Because the markers are different in terms of sensitivity, specificity, and fate in the environment (e.g. *esp* does not survive well in septic systems; viruses are smaller and more mobile in surbsurface flow than bacteria), their results in the same sample frequently differ. The confidence with which one can conclude that "microorganisms from human sources are present" (or are likely to be absent) at a given site is greatly increased when multiple markers are observed across more than one sample event.

The horse-specific *Bacteroides* marker produced some false-positive results at the beginning of the study (notably at HR sites) due to the formation of PCR artifacts. Further testing showed that the spurious PCR products were due to the formation of primer dimers. The protocol was altered to lower the primer concentration and remove the "touchdown" component of the cycle. Subsequent testing against target and non-target feces and water samples showed the reaction to be sensitive and specific for horse feces. The new protocol was used from the October sampling event on, and produced no results that appeared to be false-positives (all positives were consistent with land use).

Percentages of MST marker co-occurrence were calculated over all "fixed" sites sampled as well as stations DHR4A-10 in the Lower Hillsborough River (Table 6). Differences in average marker frequency over the sites analyzed were significantly different from one another ($p < 0.0001$). Human *Bacteroides* was detected significantly more frequently than any other marker over the study period ($p < 0.001$), possibly due to higher

concentrations in water than either *esp* or the HPyV. Although a complete understanding of the performance of the individual methods in the variety of complex environmental scenarios encountered in this study is still required, the relatively high level of detection of the human *Bacteroides* marker is also possibly due to a greater level of sensitivity of this assay as compared to the other human-specific markers (the HPyV marker is likely the least sensitive of the assays utilized in this study). In addition, occurrence of this marker was detected in 42.1% of the samples where the ruminant *Bacteroides* marker was identified, the largest percentage of co-occurrence for the project.

It is also important to note that some of the lowest levels of co-occurrence were found between *esp* and the other human-specific markers [Probability (P)(*esp*|HPyV) = 11.8%; P(*esp*|H) = 12.1%]. Although the *esp* gene is commonly found in human sewage resulting from sanitary sewer systems, it less frequently survives through the storage tanks and associated drainfields of septic systems [personal communication, Dr. Valerie J. Harwood, September 24, 2007; (Whitman et al. 2007)]. This disparity has not been observed for human *Bacteroides*. The lack of positive *esp* indicators, together with the presence of human *Bacteroides* and HPyV markers, lends support to the idea that OSTDS are a probable source. As a result, the low level of co-occurrence between *esp* and the other human-specific markers would be expected if some of the sources identified in the project WBIDs include OSTDS.

Table 6. Percentages of MST marker co-occurrence. H = human *Bacteroides*, *esp* = enterococcal surface protein, HPyV = human polyomavirus, R = ruminant *Bacteroides*. P = Probability. The percentage indicates the chance of finding the marker listed first given the detection of that shown second.

Markers Compared	Raw Data	Percentage Observed
P(H|*esp*)	4/20	20.0%
P(HPyV|*esp*)	2/20	10.0%
P(H|HPyV)	3/17	17.6%
P(*esp*|HPyV)	2/17	11.8%
P(*esp*|H)	4/33	12.1%
P(HPyV|H)	3/33	9.1%
P(H|R)	8/19	42.1%
P(*esp*|R)	2/19	10.5%
P(HPyV|R)	3/19	15.8%

Literature Cited

Anderson IC, Rhodes M, Kator H (1979) Sublethal stress in *Escherichia coli*: a function of salinity. Applied and Environmental Microbiology 38:1147-1152

Anderson KL, Whitlock JE, Harwood VJ (2005) Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Applied and Environmental Microbiology 71:3041-3048

Bordalo A, Onrassami R, Dechsakulwatana C (2002) Survival of faecal indicator bacteria in tropical estuarine waters (Bangpakong River, Thailand). Journal of Applied Microbiology 93:864-871

Solic M, Krstulovic N (1992) Separate and combined effects of solar radiation, temperature, salinity and pH on the survival of faecal coliforms in seawater. Marine Pollution Bulletin 24:411-416

Whitman R, Przybyla KK, Shively D (2007) Incidence of the enterococcal surface protein *(esp)* gene in human and animal fecal sources. Environmental science & technology 41:6090-6095