Supplementary Information for

Occurrence of Novel Human Tomato Brown Rugose Fruit Virus and Conventional Microbial Source Tracking Genetic Markers in a Hawaiian Coupled Stream-Beach System

Sarah A. Lowry[1], Adam Diedrich[2], Ella Lum[3], Orin C. Shanks[2], and Alexandria B. Boehm[1*]

[1] Department of Civil and Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, CA 94305, USA
[2] U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH, United States of America
[3] Punahou School, 1601 Punahou St, Honolulu, HI 96822

* Corresponding author: aboehm@stanford.edu

qPCR and dPCR information. Thermocycling conditions for the qPCR assays (HF183/BacR287, CPQ_056, GFD, DG3, and Sketa22) were as follows: 10 min at 95°C followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Thermocycling conditions for the dPCR assays (ToBRFV, BCoV) were as follows: 50°C for 40 min, 95°C for 2 min, followed by 40 cycles of 94°C for 5 s and 55°C for 30 s.

The equivalent sample volume analyzed for qPCR assays was 4 mL sample (2 µL template from 100 µL eluent). The equivalent sample volume analyzed for dPCR assays was 60 mL sample (15µL template in two pooled wells from 100 µL eluent). The dPCR mean partition volume was 0.7 nl. The mean dPCR partitions measured were 50869.7 partitions (variance 6192 partitions). The mean dPCR target copies per partition was 0.00011 target copies per partition (variance 1.64×10^{-7} target copies per partition).

Exclusion of GFD in logistic regression analysis. For GFD, quasi-perfect separation was observed with salinity and enterococci concentration, as the GFD marker was detected in all samples at Sites A and B (i.e., for all salinity values lower than 33.8 ppt and all enterococci concentrations >10 MPN/100mL, GFD was detected). Because these variables almost perfectly predict the detection of GFD, logistic regression is not a suitable method for the analysis of the data.

BCoV as an RNA processing control. BCoV was used as an RNA control in this study with acceptance criteria of falling between 0.1x to 10x of the mean method blank BCoV concentration for each batch. A downward concentration trend over time and difference between batch preparations were observed, suggesting that stock degradation may be occurring during storage at 4°C between uses (Figure S7). Conversely, no decay or difference across batches in the study was observed for the DNA control Sketa22 (Figure S8), suggesting that it is a more stable control. Future work should focus on identifying an RNA control with similar performance and stability as Sketa22.

Tables

Table S1. Primer and probe sequence information for selected assays.

Assay	Sequence (5' to 3')	Reference
HF183/BacR287	F: ATCATGAGTTCACATGTCCG R: CTTCCTCTCAGAACCCCTATCC P: FAM-CTAATGGAACGCATCCC-MGB IAC: VIC-AACACGCCGTTGCTACA-MGB	1
CPQ_056	F: CAGAAGTACAAACTCCTAAAAAACGTAGAG R: GATGACCAATAAACAAGCCATTAGC P: FAM-AATAACGATTTACGTGATGTAAC-MGB	2
ToBRFV	F: TCAGTGTCTGTTTGGTCGATAA R: GGAACGACTTTGAACTGAAACC P: FAM-AGAGCGGACGAGGCAACTCCTG-ZEN/IBHQ	3
GFD	F: TCGGCTGAGCACTCTAGGG R: GCGTCTCTTTGTACATCCCA P: FAM-ACGTCAAGTCATCATGGCCCTTACGC-ZEN/IBHQ	4,5
DG3	F: TTTTCAGCCCCGTTGTTTCG R: TGAGCGGGCATGGTCATATT P: FAM-AGTCTACGCGGGCGTACT-MGB	6
Sketa22	F: GGTTTCCGCAGCTGGG R: CCGAGCCGTCCTGGTC P: FAM-AGTCGCAGGCGGCCACCGT-TAMRA	1
BCoV	F: CTGGAAGTTGGTGGAGTT R: ATTATCGGCCTAACATACATC P: FAM-CCTTCATATCTATACACATCAAGTTGTT-BHQ1	7

Table S2. Temporal sampling campaign univariate analysis results used to screen correlations for inclusion into logistic regression analysis. Variables with correlations with p-values less than 0.25 are bolded and were included in the logistic regression analysis.

MST Marker	Variable Tested	P-Value for Correlation
ToBRFV	Enterococci	0.49
ToBRFV	Tidal Stage	0.34
ToBRFV	Temperature	0.99
ToBRFV	Salinity	0.82
ToBRFV	**Location**	**0.098**
ToBRFV	**Time**	**0.11**
ToBRFV	**Rainfall**	**0.012**
DG3	**Enterococci**	**0.23**
DG3	**Tidal Stage**	**0.15**
DG3	**Temperature**	**0.19**
DG3	**Salinity**	**0.03**
DG3	Location	0.99
DG3	Time	0.46
DG3	Rainfall	1.00

Table S3. Results of multivariable logistic regression model for ToBRFV and DG3 detection. P-values adjusted in each family (ToBRFV, DG3) of tests using the Benjamini-Hochberg (BH) correction. * indicates a significant result.

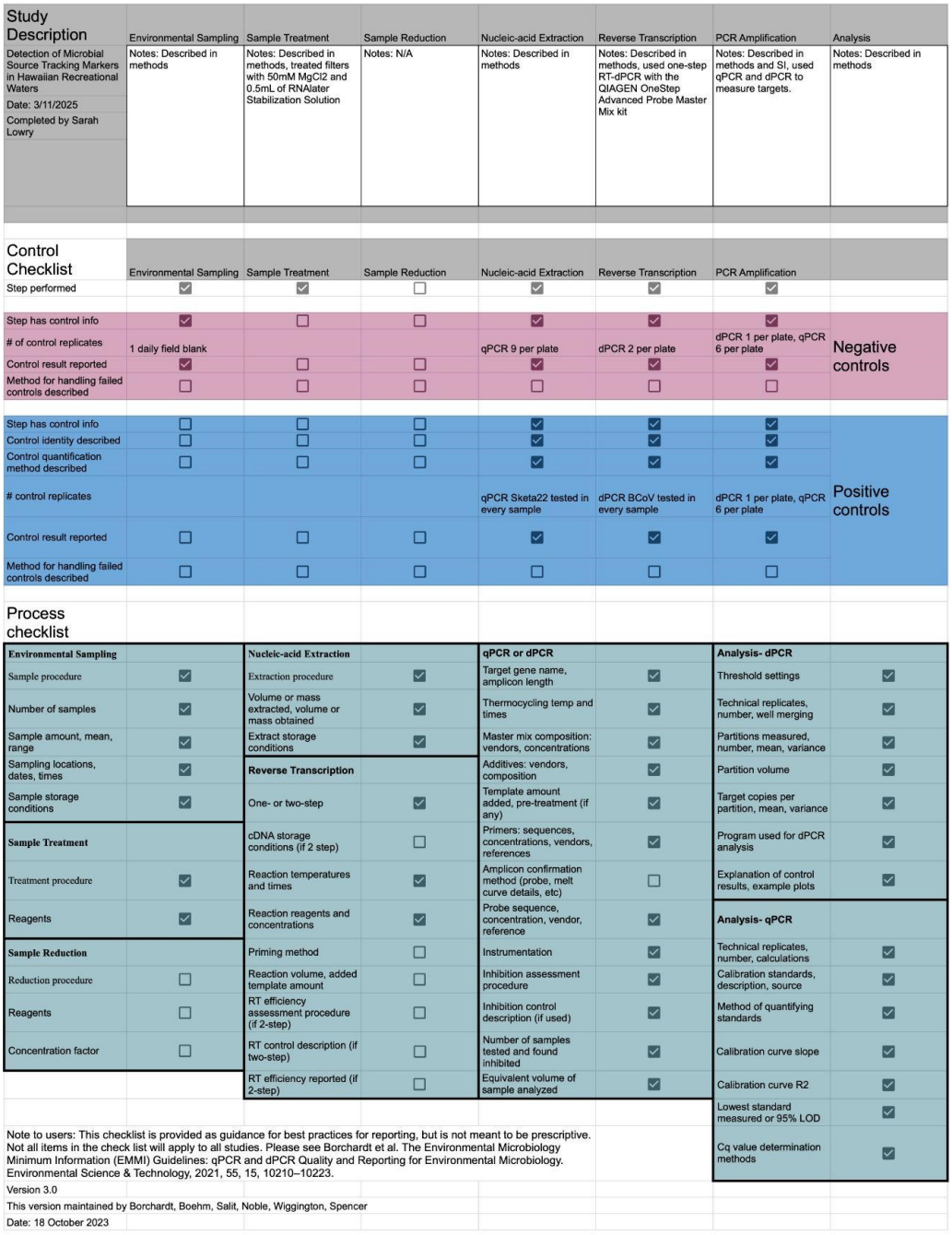
ToBRFV				
Term	**β Estimate**	**Standard Error**	**P-value**	**Adjusted P-value**
Intercept	-0.97	0.57	0.090	N/A
Site: B	1.11	0.63	0.076	0.10
Site: C	-0.42	0.65	0.52	0.52
Time: Morning	-1.04	0.53	0.051	0.10
Rainfall in previous day	**1.49**	**0.54**	**0.0062**	**0.025***
DG3				
Term	**β Estimate**	**Standard Error**	**P-value**	**Adjusted P-Value**
Intercept	-0.86	6.28	0.65	N/A
Enterococci concentration	<-0.001	0.002	0.41	0.41
Water temperature	-0.0026	0.27	0.41	0.41
Salinity	-0.096	0.053	0.27	0.41
Flood Tidal Stage	2.66	1.75	0.13	0.41
Transitional Tidal Stage	1.92	1.5	0.20	0.41

Table S4. Results from random grab samples taken throughout the sampling campaign at sites of interest. Positive MST marker detections are bolded and highlighted in red.

Sample	Enterococci	HF183/Ba	CPQ_056	ToBRFV	DG3	GFD

	concentration (MPN/100mL)	cR287	crAssphage-like			
Stream 1	41	Nondetect	Nondetect	**Detect**	Nondetect	**Detect**
Stream 2	41	Nondetect	Nondetect	Nondetect	Nondetect	**Detect**
Stream 3	216	**Detect**	**Detect**	**Detect**	Nondetect	**Detect**
Outfall	<10	Nondetect	Nondetect	Nondetect	Nondetect	Nondetect
Seaweed	<10	Nondetect	Nondetect	Nondetect	Nondetect	Nondetect

Figures



Study Description	Environmental Sampling	Sample Treatment	Sample Reduction	Nucleic-acid Extraction	Reverse Transcription	PCR Amplification	Analysis
Detection of Microbial Source Tracking Markers in Hawaiian Recreational Waters Date: 3/11/2025 Completed by Sarah Lowry	Notes: Described in methods	Notes: Described in methods, treated filters with 50mM MgCl2 and 0.5mL of RNAlater Stabilization Solution	Notes: N/A	Notes: Described in methods	Notes: Described in methods, used one-step RT-dPCR with the QIAGEN OneStep Advanced Probe Master Mix kit	Notes: Described in methods and SI, used qPCR and dPCR to measure targets.	Notes: Described in methods

Control Checklist	Environmental Sampling	Sample Treatment	Sample Reduction	Nucleic-acid Extraction	Reverse Transcription	PCR Amplification	
Step performed	☑	☑	☐	☑	☑	☑	
Step has control info	☑	☐	☐	☑	☑	☑	Negative controls
# of control replicates	1 daily field blank			qPCR 9 per plate	dPCR 2 per plate	dPCR 1 per plate, qPCR 6 per plate	
Control result reported	☑	☐	☐	☑	☑	☑	
Method for handling failed controls described	☐	☐	☐	☐	☐	☐	
Step has control info	☐	☐	☐	☑	☑	☑	Positive controls
Control identity described	☐	☐	☐	☑	☑	☑	
Control quantification method described	☐	☐	☐	☑	☑	☑	
# control replicates				qPCR Sketa22 tested in every sample	dPCR BCoV tested in every sample	dPCR 1 per plate, qPCR 6 per plate	
Control result reported	☐	☐	☐	☑	☑	☑	
Method for handling failed controls described	☐	☐	☐	☐	☐	☐	

Process checklist

Environmental Sampling		**Nucleic-acid Extraction**		**qPCR or dPCR**		**Analysis- dPCR**	
Sample procedure	☑	Extraction procedure	☑	Target gene name, amplicon length	☑	Threshold settings	☑
Number of samples	☑	Volume or mass extracted, volume or mass obtained	☑	Thermocycling temp and times	☑	Technical replicates, number, well merging	☑
Sample amount, mean, range	☑	Extract storage conditions	☑	Master mix composition: vendors, concentrations	☑	Partitions measured, number, mean, variance	☑
Sampling locations, dates, times	☑	**Reverse Transcription**		Additives: vendors, composition	☑	Partition volume	☑
Sample storage conditions	☑	One- or two-step	☑	Template amount added, pre-treatment (if any)	☑	Target copies per partition, mean, variance	☑
Sample Treatment		cDNA storage conditions (if 2 step)	☐	Primers: sequences, concentrations, vendors, references	☑	Program used for dPCR analysis	☑
Treatment procedure	☑	Reaction temperatures and times	☑	Amplicon confirmation method (probe, melt curve details, etc)	☐	Explanation of control results, example plots	☑
Reagents	☑	Reaction reagents and concentrations	☑	Probe sequence, concentration, vendor, reference	☑	**Analysis- qPCR**	
Sample Reduction		Priming method	☐	Instrumentation	☑	Technical replicates, number, calculations	☑
Reduction procedure	☐	Reaction volume, added template amount	☐	Inhibition assessment procedure	☑	Calibration standards, description, source	☑
Reagents	☐	RT efficiency assessment procedure (if 2-step)	☐	Inhibition control description (if used)	☑	Method of quantifying standards	☑
Concentration factor	☐	RT control description (if two-step)	☐	Number of samples tested and found inhibited	☑	Calibration curve slope	☑
		RT efficiency reported (if 2-step)	☐	Equivalent volume of sample analyzed	☑	Calibration curve R2	☑
						Lowest standard measured or 95% LOD	☑
						Cq value determination methods	☑

Note to users: This checklist is provided as guidance for best practices for reporting, but is not meant to be prescriptive. Not all items in the check list will apply to all studies. Please see Borchardt et al. The Environmental Microbiology Minimum Information (EMMI) Guidelines: qPCR and dPCR Quality and Reporting for Environmental Microbiology. Environmental Science & Technology, 2021, 55, 15, 10210–10223.

Version 3.0

This version maintained by Borchardt, Boehm, Salit, Noble, Wiggington, Spencer

Date: 18 October 2023

Figure S1. Completed Environmental Microbiology Minimum Information (EMMI) guidelines from Borchardt et al. 2021.[8]

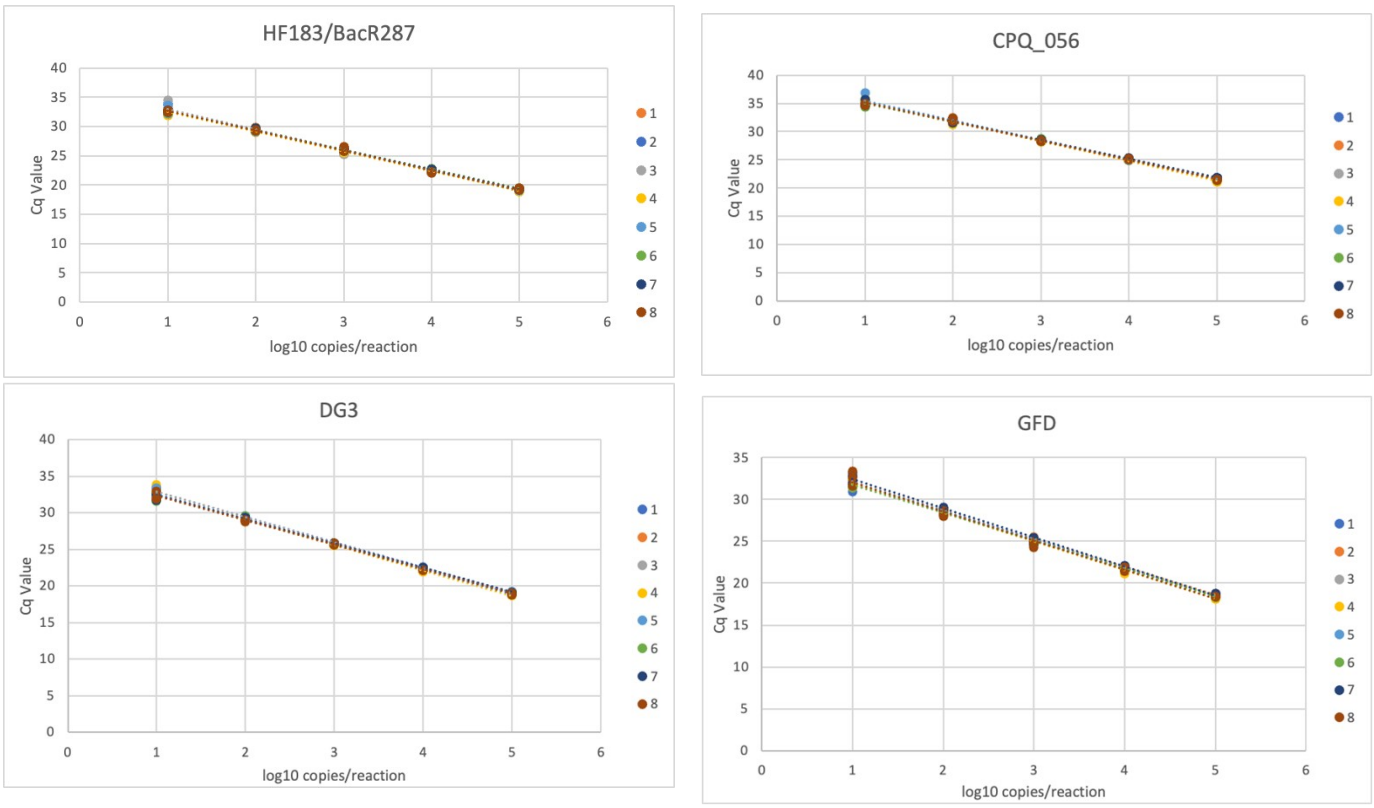


Figure S2. Standard curves for HF183/BacR287, CPQ_056 (crAssphage), DG3, and GFD assays. Eight standard curves produced for each assay. Slope generated from all standard curves, y-intercept generated for each batch using the batch-specific y-intercept control.

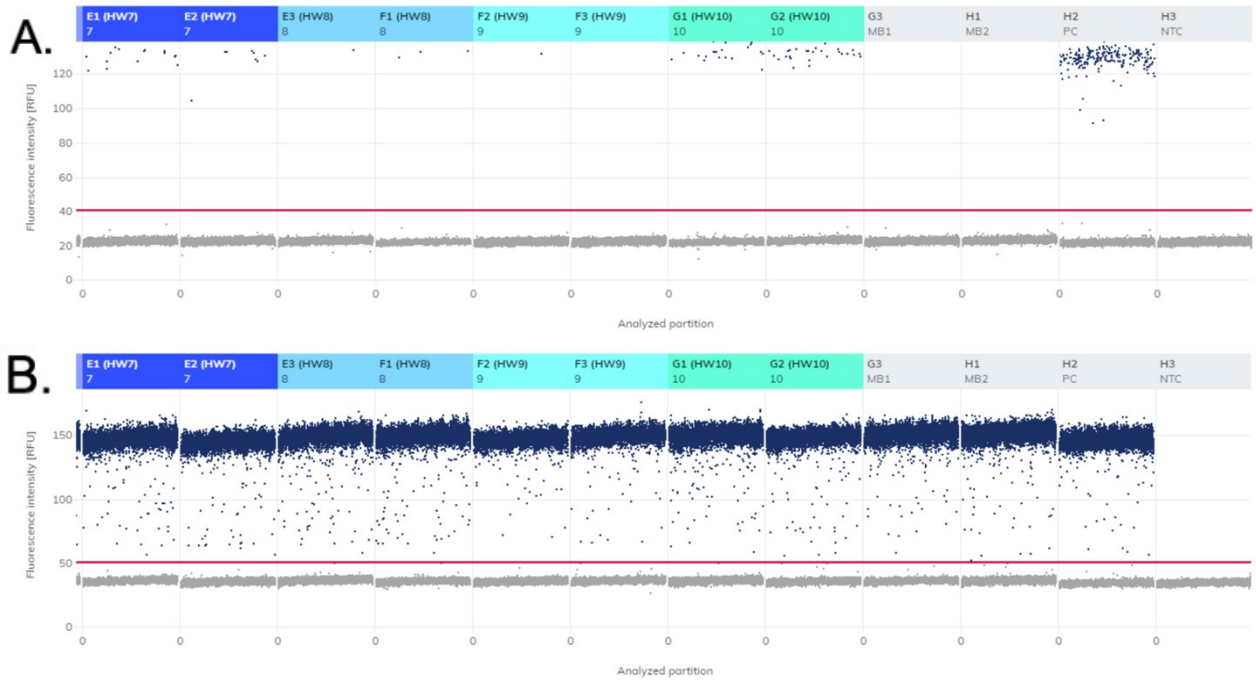


Figure S3. (A) Example thresholding for ToBRFV partition data. Threshold was set at 40 RFU across all instrument runs from precision profile testing. (B) Example thresholding for BCoV partition data. Threshold was set at 50 RFU across all instrument runs from precision profile testing.

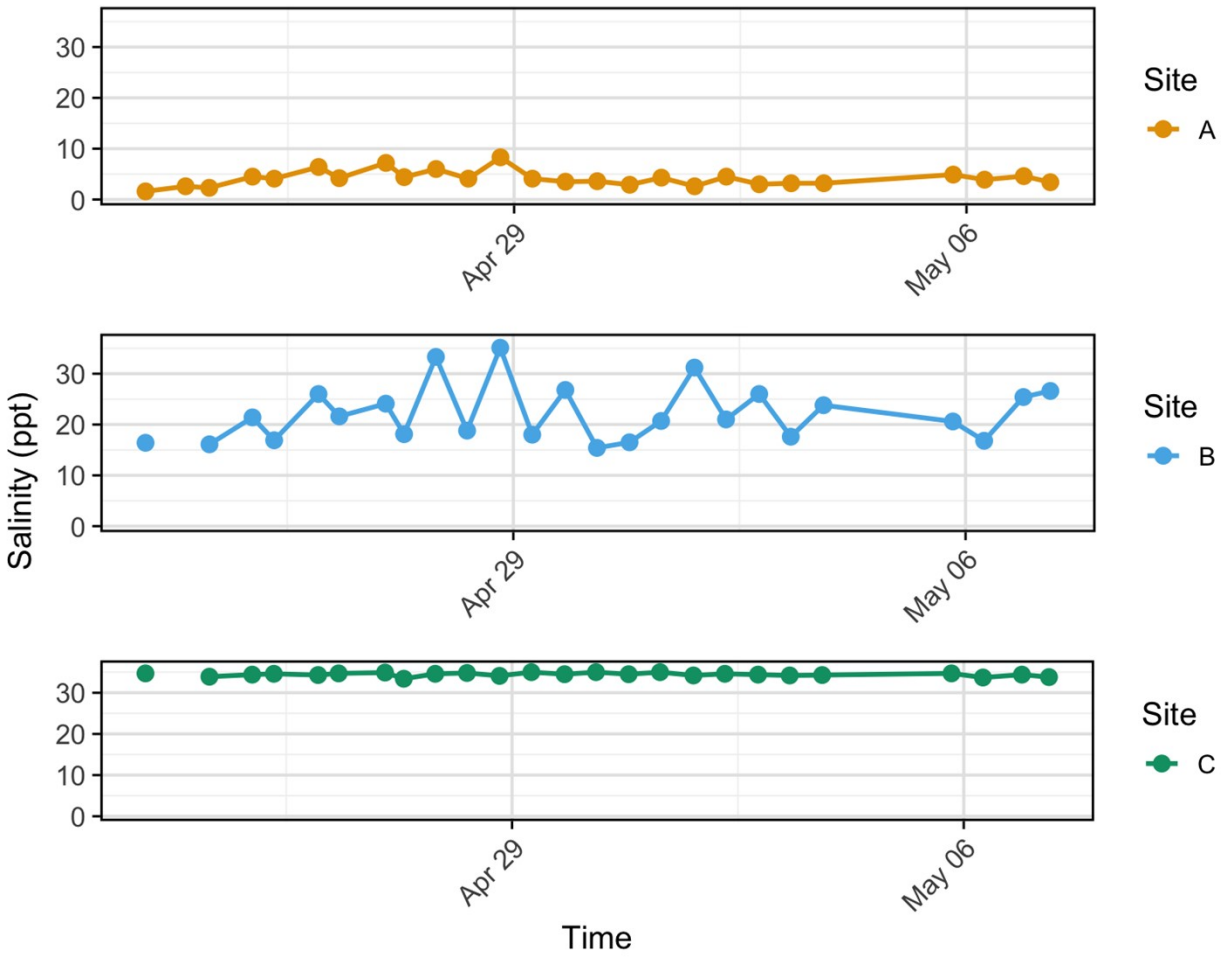


Figure S4. Salinity measurements in ppt over time for each sample in the temporal sampling campaign stratified by site. Two samples (B4-23E and C4-23E) are missing salinity data because of instrument malfunction, shown by a gap in the line.

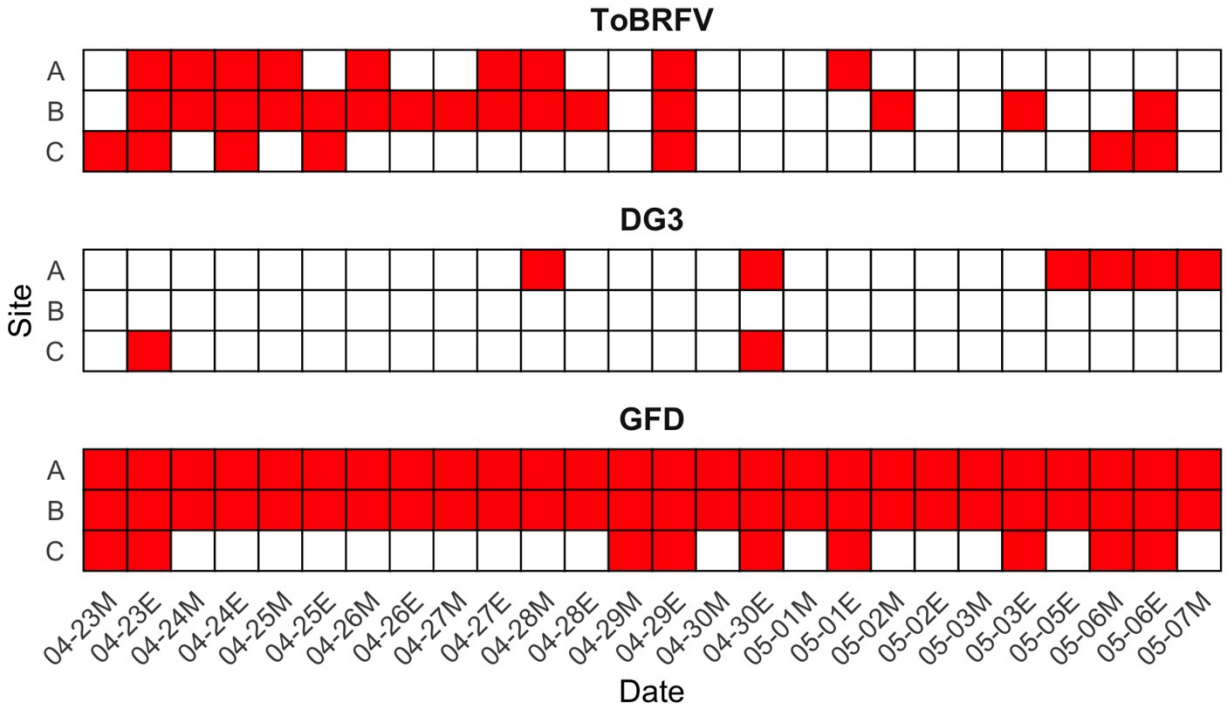


Figure S5. Heatmap of three MST markers across sites and dates of the sampling campaign. Detects are shown in red and non-detects are shown in white. The letter on the end of the date refers to the time of day the sample was collected, with M representing “morning” and E representing “evening”. Site is shown on the x-axis, with A representing the most upstream site (marsh), B representing the stream site, and C representing the beach site.

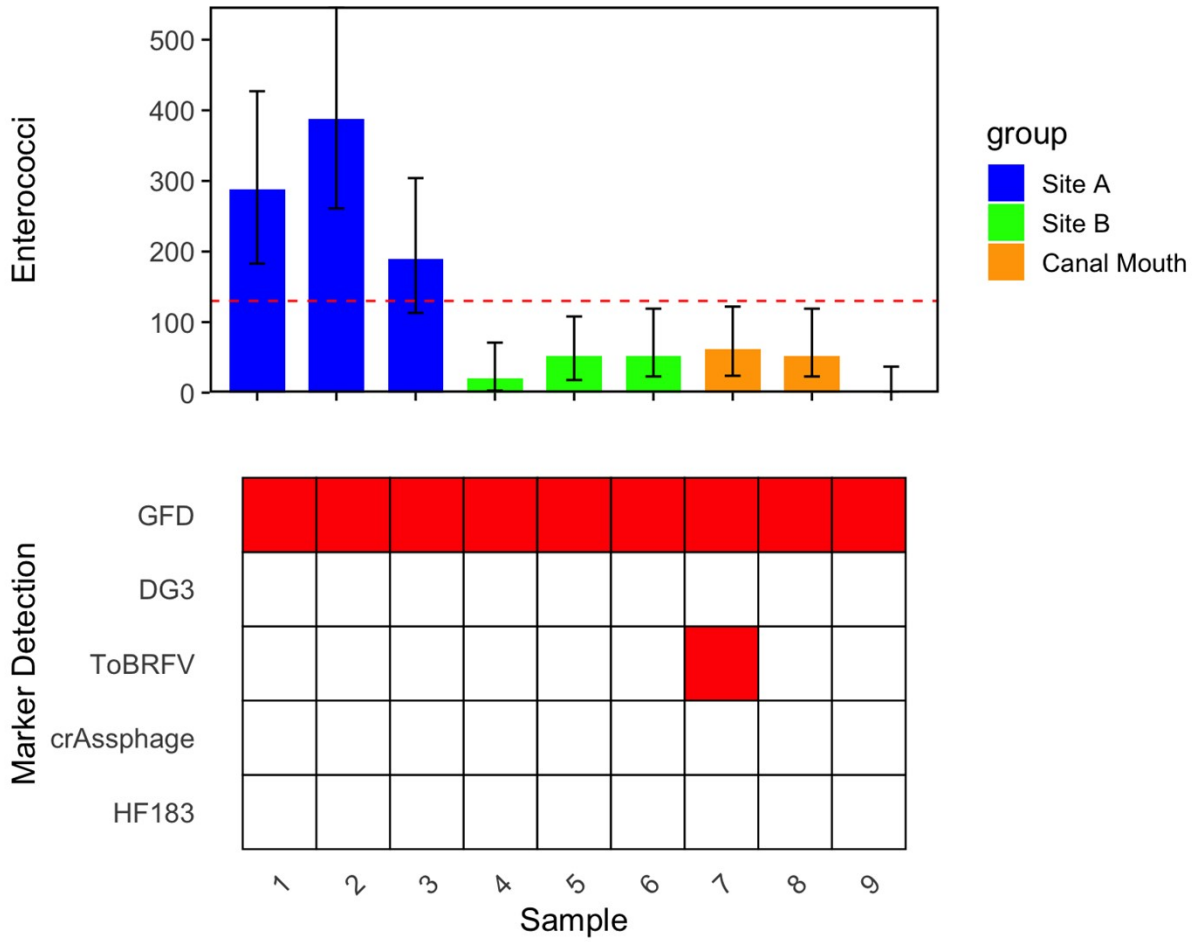


Figure S6. Results from lateral spatial sampling campaign, where samples 1, 2, and 3 were taken across the canal at Site A, samples 4, 5, and 6 were taken across the canal at Site B, and samples 7, 8, and 9 were taken across the canal at the mouth of the canal. The top graph shows enterococci concentration (MPN/100 mL) for each group of lateral samples with error bars representing lower and upper 95% confidence limits, and the bottom heat maps shows the detection of MST markers for each sample.

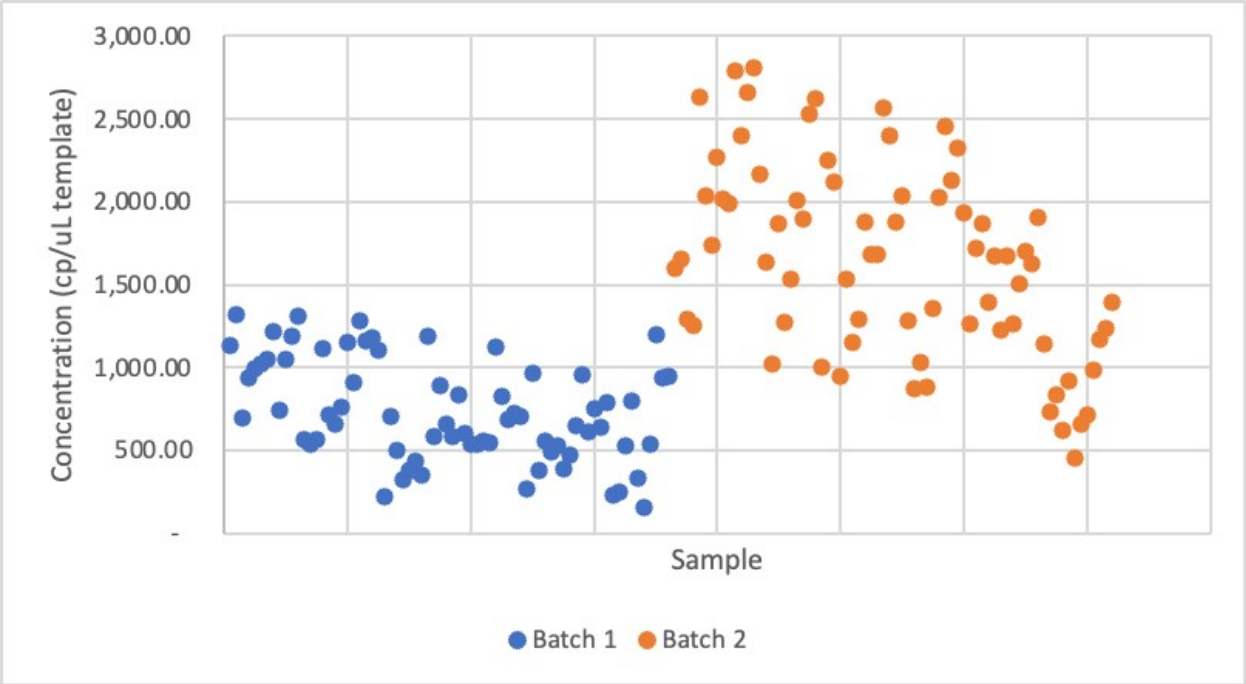


Figure S7. BCoV concentrations (cp/μL template) measured in each sample over the course of the study. Downward trend over time and difference in batches used was observed, suggesting that decay may be occurring while BCoV control was stored in the fridge.

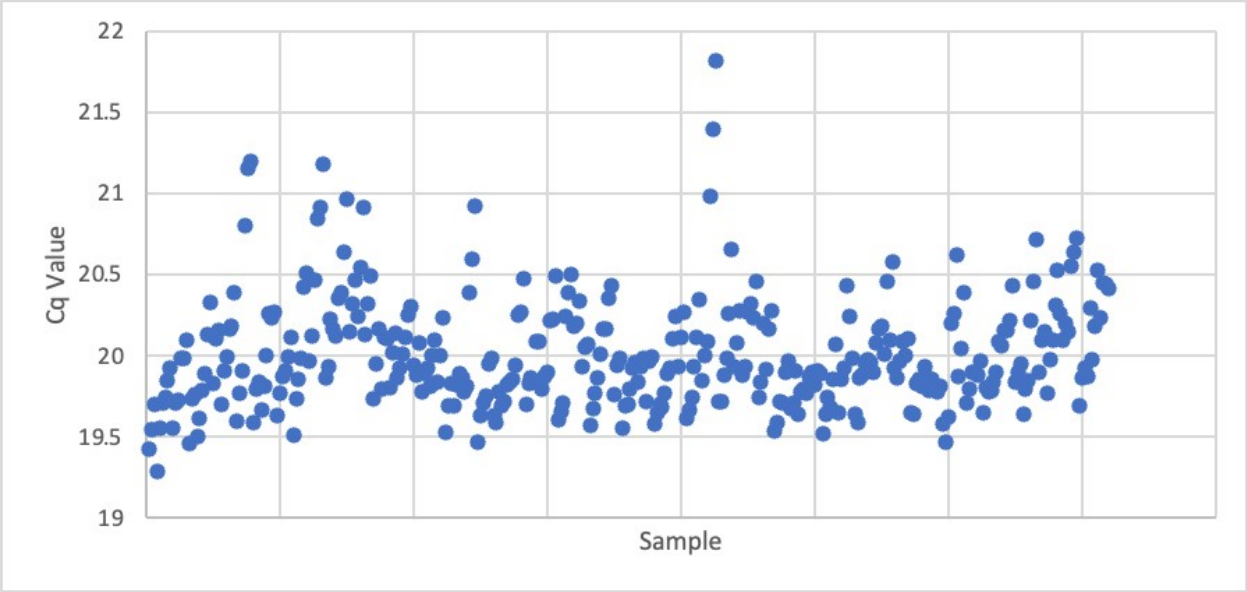


Figure S8. Sketa22 Cq values measured in each sample over the course of the study. No downward trend or difference in batches was observed, and sample Cq values are tightly clustered.

References

1. USEPA. Method 1696.1: Characterization of human fecal pollution in water by HF183/BacR287 TaqMan quantitative polymerase chain reaction (qPCR) assay. United States Environmental Protection Agency; 2021.

2. Stachler E, Kelty C, Sivaganesan M, Li X, Bibby K, Shanks OC. Quantitative CrAssphage PCR Assays for Human Fecal Pollution Measurement. Environ Sci Technol. 2017 Aug 15;51(16):9146–54.

3. Natarajan A, Fremin BJ, Schmidtke DT, Wolfe MK, Zlitni S, Graham KE, et al. The Tomato Brown Rugose Fruit Virus Movement Protein Gene Is a Novel Microbial Source Tracking Marker. Applied and Environmental Microbiology. 2023 July 5;89(7):e00583-23.

4. Green HC, Dick LK, Gilpin B, Samadpour M, Field KG. Genetic Markers for Rapid PCR-Based Identification of Gull, Canada Goose, Duck, and Chicken Fecal Contamination in Water. Applied and Environmental Microbiology. 2012 Jan 15;78(2):503–10.

5. Weller D, Belias A, Green H, Roof S, Wiedmann M. Landscape, Water Quality, and Weather Factors Associated With an Increased Likelihood of Foodborne Pathogen Contamination of New York Streams Used to Source Water for Produce Production. Front Sustain Food Syst [Internet]. 2020 Feb 6 [cited 2025 Apr 16];3. Available from: https://www.frontiersin.orghttps://www.frontiersin.org/journals/sustainable-food-systems/articles/10.3389/fsufs.2019.00124/full

6. Green HC, White KM, Kelty CA, Shanks OC. Development of Rapid Canine Fecal Source Identification PCR-Based Assays. Environ Sci Technol. 2014 Oct 7;48(19):11453–61.

7. Decaro N, Elia G, Campolo M, Desario C, Mari V, Radogna A, et al. Detection of bovine coronavirus using a TaqMan-based real-time RT-PCR assay. Journal of Virological Methods. 2008 Aug 1;151(2):167–71.

8. Borchardt MA, Boehm AB, Salit M, Spencer SK, Wigginton KR, Noble RT. The Environmental Microbiology Minimum Information (EMMI) Guidelines: qPCR and dPCR Quality and Reporting for Environmental Microbiology. Environ Sci Technol. 2021 Aug 3;55(15):10210–23.