Supplementary Tables for Host-specificity and sensitivity of the established and novel sewage-associated marker genes in human and non-human fecal samples

Warish Ahmed[a,*], Pradip Gyawali[b], Shuchen Feng[c], Sandra L. McLellan[c]

[a] CSIRO Land and Water, Ecosciences Precinct, 41 Boggo Road, Dutton Park, QLD 4102, Australia

[b] Institute of Environmental Science and Research Ltd (ESR), Porirua, 5240, New Zealand

[c] School of Freshwater Sciences, University of Wisconsin-Milwaukee, 600 E. Greenfield Ave, Milwaukee, WI 53204, USA.

Supplementary Table S1

Concentration of sewage-associated marker genes in non-human fecal samples

Hosts	Unit of measurements	Total no. of quantifiable samples (Mean ± SD $\log_{10}$)			
		HF183	BacV6-21	Lachno3	crAssphage CPQ_056
Birds	Per gm of feces		5 (4.78 ± 0.51)		
	Per ng of DNA		5 (1.54 ± 0.74)		
Cat	Per gm of feces	1(6.88)		1 (9.24)	5 (6.88 ± 0.99)
	Per ng of DNA	1(2.48)		1 (3.33)	5 (2.46 ± 0.72)
Cow	Per gm of feces				
	Per ng of DNA				
Chicken	Per gm of feces	3 (4.49 ± 0.15)	8 (6.42 ± 1.94)		
	Per ng of DNA	3 (1.56 ± 0.57)	8 (3.22 ± 1.78)		
Deer	Per gm of feces			3 (5.82 ± 0.96)	
	Per ng of DNA			3 (1.85 ±1.10)	
Dog	Per gm of feces		1 (9.99)		
	Per ng of DNA		1 (4.82)		
Emu	Per gm of feces	4 (4.91 ± 0.42)			
	Per ng of DNA	4 (1.89 ± 0.24)			
Horse	Per gm of feces			11 (5.73 ± 0.54)	
	Per ng of DNA			11 (2.48 ± 0.68)	
Koala	Per gm of feces	2 (6.38 ± 0.99)			
	Per ng of DNA	2 (2.68 ± 0.86)			
Sheep	Per gm of feces		1 (4.99)		
	Per ng of DNA		1 (1.99)		

Supplementary Table S2

Concentrations of sewage-associated marker genes/L of sewage or g of human feces

Sewage markers	Types of samples	Concentrations GC/L or g		
		Mean ± SD	Min - Max	95% CI (lower - upper)
HF183	Untreated sewage	9.27 ± 0.13	9.13 - 9.38	8.97 - 9.58
	Human feces	7.83 ± 0.86	6.59 - 9.38	7.47 - 8.19
BacV6-21	Untreated sewage	9.47 ± 0.19	9.24 - 9.60	8.97 - 9.96
Lachno3	Untreated sewage	9.22 ± 0.69	8.45 - 9.77	7.51 - 10.9
	Human feces	7.33 ± 1.13	4.53 - 9.38	6.80 - 7.87
HAdV	Untreated sewage	6.50 ± 0.12	6.37 - 6.62	6.19 - 6.81
HPyV	Untreated sewage	6.72 ± 0.43	6.34 - 7.19	5.64 - 7.79
crAssphage CPQ_056	Untreated sewage	9.37 ± 0.36	9.01 - 9.73	8.47 - 10.3
	Human feces	7.36 ± 0.82	6.25 - 8.92	6.80 - 7.91

Supplementary Table S3
Concentrations of sewage-associated marker genes/ng of DNA extracted from sewage or human feces

Sewage markers	Types of samples	Concentrations/ng of DNA		
		Mean ± SD	Min - Max	95% CI (lower - upper)
HF183	Untreated sewage	2.74 ± 0.31	2.46 - 3.08	1.95 - 3.53
	Human feces	3.23 ± 0.59	1.73 - 4.12	2.96 - 3.50
BacV6-21	Untreated sewage	3.97 ± 0.39	3.60 - 4.39	3.00 - 4.95
Lachno3	Untreated sewage	4.24 ± 0.38	3.82 - 4.59	3.28 - 5.21
	Human feces	3.15 ± 0.90	1.47 - 4.78	2.70 - 3.60
HAdV	Untreated sewage	1.02 ± 0.09	0.94 - 1.12	0.78 - 1.25
HPyV	Untreated sewage	1.22 ± 0.48	0.66 - 1.56	0.01 - 2.43
crAssphage CPQ_056	Untreated sewage	3.54 ± 0.39	3.10 - 3.83	2.57 - 4.52
	Human feces	3.72 ± 0.55	2.43 - 4.39	3.35 - 4.10

Supplementary Table S4

P value obtained using analysis of variance analysis (ANOVA) among the concentrations of sewage-associated marker genes/L sewage or g of feces.

Marker genes	HF183[a]	BacV6-21[a]	Lachno3 [a]	CPQ_056 [a]	HAdV [a]	HPyV [a]	HF183[b]	Lachno3 [b]	CPQ_056 [b]
HF183 [a]	NA	0.22	0.91	0.67	0.00001	0.0006	NA	NA	NA
BacV6-21 [a]		NA	0.59	0.70	0.00002	0.0005	NA	NA	NA
Lachno-3 [a]			NA	0.76	0.002	0.005	NA	NA	NA
CPQ_056 [a]				NA	0.0002	0.001	NA	NA	NA
HAdV [a]					NA	0.45	NA	NA	NA
HPyV [a]						NA	NA	NA	NA
HF183[b]							NA	0.10	0.13
Lachno3 [b]								NA	0.95
CPQ_056 [b]									NA

[a]: Concentration in sewage; [b]: concentration in faeces; NA: not applicable.

Supplementary Table S5

P value obtained using analysis of variance analysis (ANOVA) among the concentrations of sewage-associated marker genes per ng sewage or feces.

Marker genes	HF183[a]	BacV6-21[a]	Lachno3 [a]	CPQ_056 [a]	HAdV [a]	HPyV [a]	HF183[b]	Lachno3 [b]	CPQ_056 [b]
HF183 [a]	NA	0.01	0.006	0.05	0.0008	0.01	NA	NA	NA
BacV6-21 [a]		NA	0.44	0.25	0.0002	0.001	NA	NA	NA
Lachno-3 [a]			NA	0.09	0.0001	0.001	NA	NA	NA
CPQ_056 [a]				NA	0.0004	0.003	NA	NA	NA
HAdV [a]					NA	0.51	NA	NA	NA
HPyV [a]						NA	NA	NA	AN
HF183[b]							NA	0.58	0.03
Lachno3 [b]								NA	0.05
CPQ_056 [b]									NA

[a]: Concentration in sewage; [b]: concentration in faeces; NA: not applicable.

Supplementary Table S6
Primers and probes used in this study

Assays	Primers and probes	Primer and probe concentration	Cycling parameters	References
Bacteroides HF183	F:ATC ATG AGT TCA CAT GTC CG R:CTT CCT CTC AGA ACC CCT ATC C P:FAM-CTA ATG GAA CGC ATC CC-BHQ-1	1000 nM 1000 nM 100 nM	10 min at 95°C, 40 cycles of 15 s at 95°C, 60 s at 60°C	(1)
BacV6-21	F: GC TTG AAT TGC AGA GGA ATA R: GCA GTC TCA CTA GAG TCC TCA G P:FAM-AGT TGA AAG ATT ATG GCC GCA-MGBNFQ	1000 nM 1000 nM 100 nM	2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C, 60 s at 60°C	(4)
Lachno3	F:CAA CGC GAA GAA CCT TAC CAA A R: CCC AGA GTG CCC ACC TTA AAT P:FAM-CTC TGA CCG GTC TTT AAT CGG A-TAMRA	1000 nM 1000 nM 100 nM	2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C, 60 s at 64°C	(3)
HAdV	F:GCC ACG GTG GGG TTT CTA AAC TT R:GCC CCA GTG GTC TTA CAT GCA CAT C P:FAM-TGC ACC AGA CCC GGG CTC AGG TAC TCC GA-TAMRA	200 nM 200 nM 100 nM	10 min at 95°C, 40 cycles of 10 s at 95°C, 10 s for 55°C, 60 s at 65°C	(6)
HPyV	F:AGT CTT TAG GGT CTT CTA CCT TT R:GGT GCC AAC CTA TGG AAC AG P:FAM-TCA TCA CTG GCA AAC AT-MGBNFQ	200 nM 200 nM 100 nM	10 min at 95°C, 40 cycles of 15 s at 95°C, 20 s for 55°C, 60 s at 60°C	(7)
CrAssphage CPQ_056	F:CAG AAG TAC AAA CTC CTA AAA AAC GTA GAG R:GAT GAC CAA TAA ACA AGC CAT TAG C P:Quasar 670-AAT AAC GAT TTA CGT GAT GTA AC-BHQ-3	1000 nM 1000 nM 100 nM	10 min at 95°C, 40 cycles of 15 s at 95°C, 60 s at 60°C	(2)
Sketa22	F:GGT TTC CGC AGC TGG G R:CCG AGC CGT CCT GGT CTA P:FAM-AGT CGC AGG CGG CCA CCG T-TAMRA	500 nM 500 nM 400 nM	10 min at 95°C, 40 cycles of 15 s at 95°C, 45 s at 63°C	(5)

Supplementary Table S7
MST marker gene performance criteria

Performance characteristics	Formula
Host-prevalence	$\frac{\mathrm{TP}}{\mathrm{FN}+\mathrm{TP}}$
Host-specificity	$\frac{\mathrm{TN}}{\mathrm{FP}+\mathrm{TN}}$
Accuracy	$\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}+\mathrm{TN}}$

TP: True positives
FN: False negatives
TN: True negatives
FP: False positives

References

1. Green HC, Haugland RA, Varma M, Millen HT, Borchardt MA, Field KG, Walters WA, Knight R, Sivaganesan M, Kelty CA, Shanks OC. 2014. Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Appl Environ Microbiol. 80:3086-3094.
2. Stachler E, Kelty C, Sivaganesan M, Li X, Bibby K, Shanks OC. 2017. Quantitative CrAssphage PCR assays for human fecal pollution measurement. Environ. Sci. Technol. 51:9146-9154.
3. Feng S, Bootsma M, McLellan SL. 2018. Novel human-associated *Lachnospiraceae* genetic markers improve detection of fecal pollution sources in urban waters. Appl Environ Microbiol (on-line early).
4. Feng S, McLellan SL. 2019. Highly specific sewage-derived Bacteroides qPCR assays target sewage polluted waters. Appl Environ Microbiol (online early).
5. Haugland RA, Siefring SC, Wymer LJ, Brenner KP, Dufour AP. 2005. Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Res. 39:559-568.
6. Heim A, Ebnet C, Harste G, Pring-Akerblom P. 2003. Rapid and quantitative detection of human adenovirus DNA by real-time PCR. J Med Virol. 70, 228-229.
7. McQuaig S, Scott T, Lukasik J, Paul J, Harwood V. 2009. Quantification of human polyomaviruses JC virus and BK virus by TaqMan quantitative PCR and comparison to other water quality indicators in water and fecal samples. Appl Environ Microbiol. 75:3379-3388.