

Supplementary Information
Potential of an Environmental DNA Metabarcoding Assay as a Tool for Source Tracking of Fecal Pollution in River Waters.

Akihiko Hata^{1,*}, Yuno Meuchi², Keisuke Kuroda¹, Kenshi Sankoda¹, Seiya Hanamoto³

¹ *Faculty of Engineering, Toyama Prefectural University, Imizu, Toyama, Japan*

² *Graduate School of Engineering, Toyama Prefectural University, Imizu, Toyama, Japan*

³ *Environment Preservation Center, Kanazawa University, Kanazawa, Ishikawa, Japan*

***Corresponding author:** Akihiko Hata

Mailing address: *Department of Environmental and Civil Engineering, Faculty of Engineering, Toyama Prefectural University, 5180 Kurokawa, Imizu, Toyama 939-0398, Japan*

Tel: +81-766-56-7500

E-mail: a-hata@pu-toyama.ac.jp

Supplementary Figures

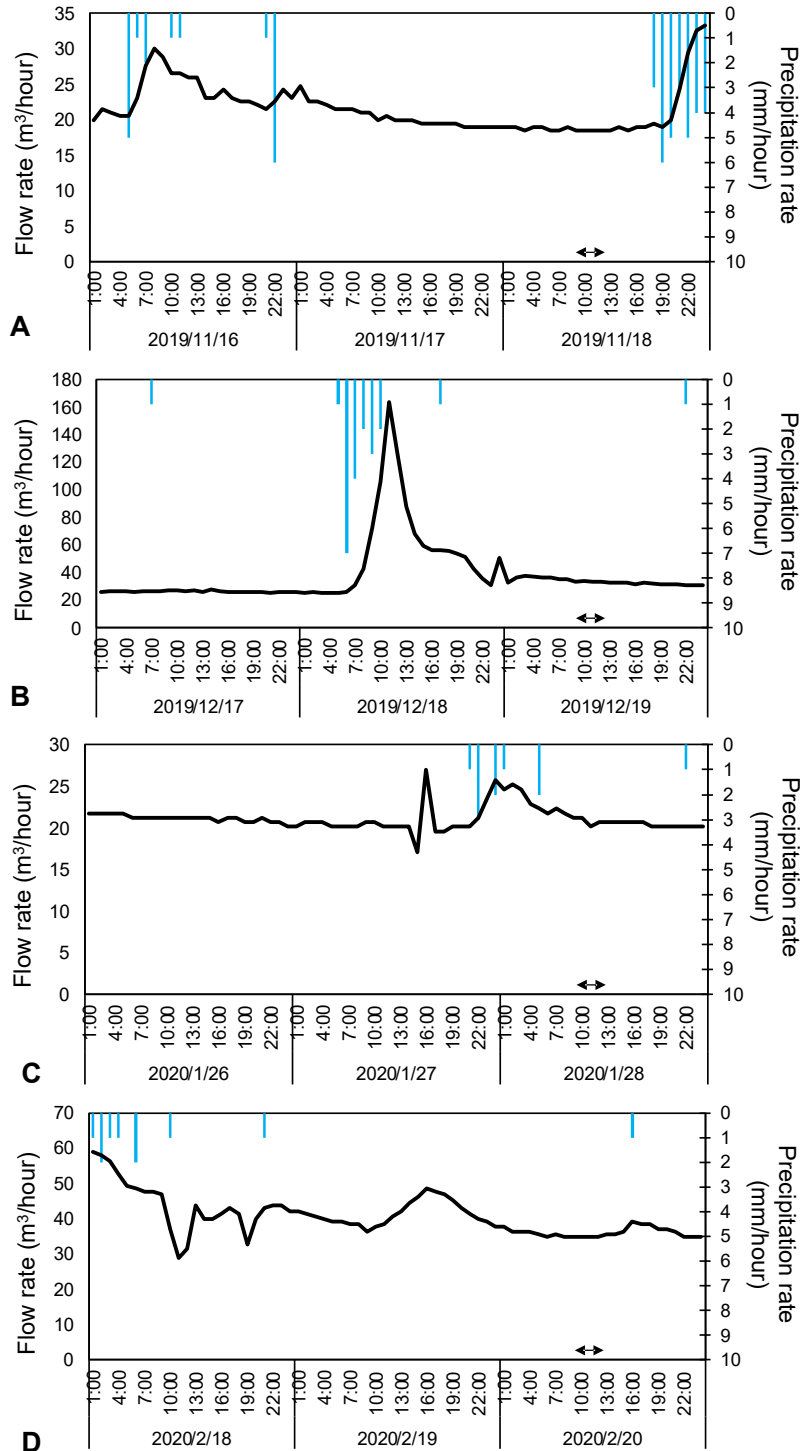


Figure S1-1. Changes in flow rate and precipitation around site O1 from two days before to the day of sample collection in Nov. 2019 (A), Dec. 2019 (B), Jan. 2020 (C), and Feb. 2020 (D). Arrows indicate the time of sampling (10:30-12:30).

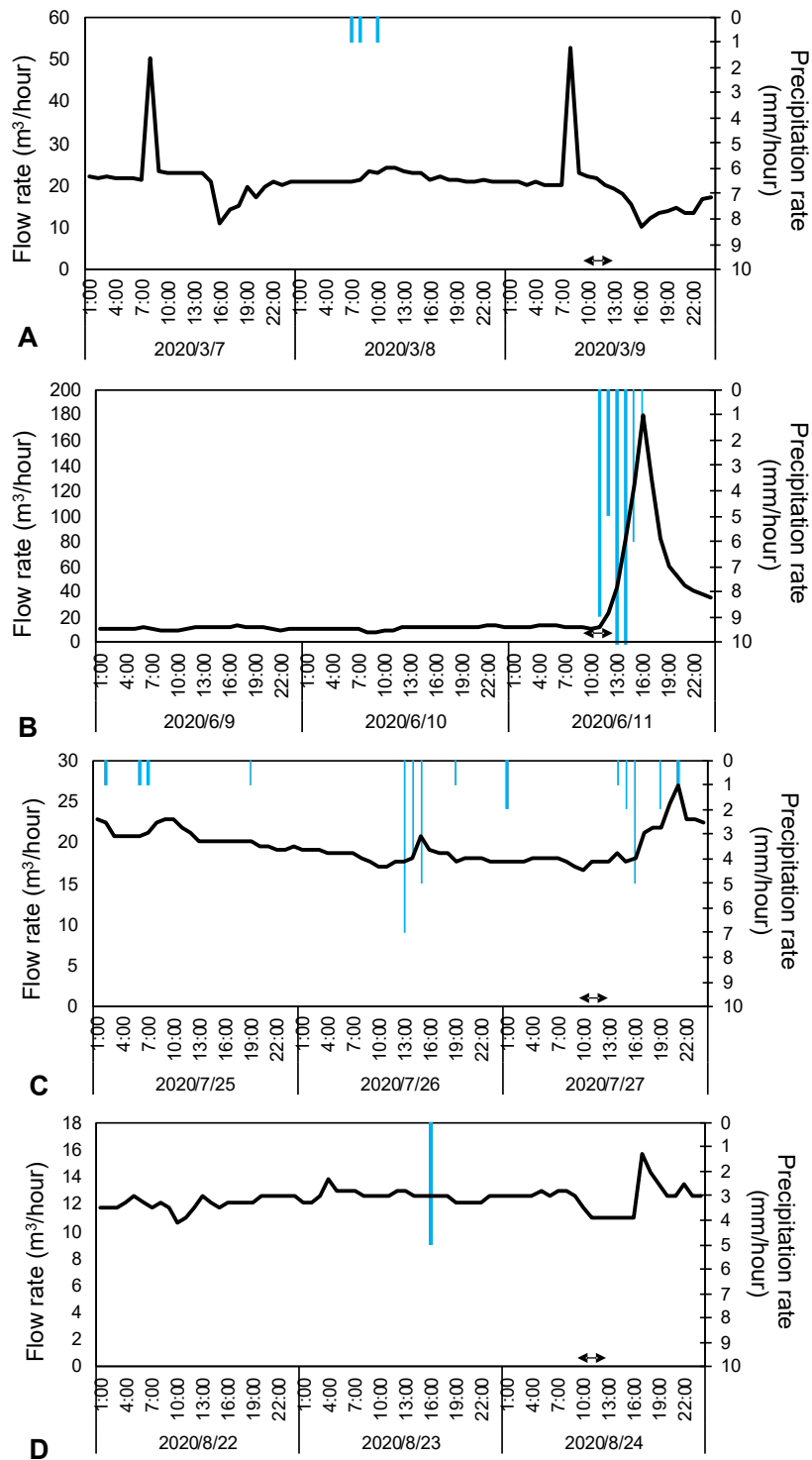


Figure S1-2. Changes in flow rate and precipitation around site O1 from two days before to the day of sample collection in Mar. 2020 (A), June 2020 (B), July 2020 (C), and Aug. 2020 (D). Arrows indicate the time of sampling (10:30-12:30).

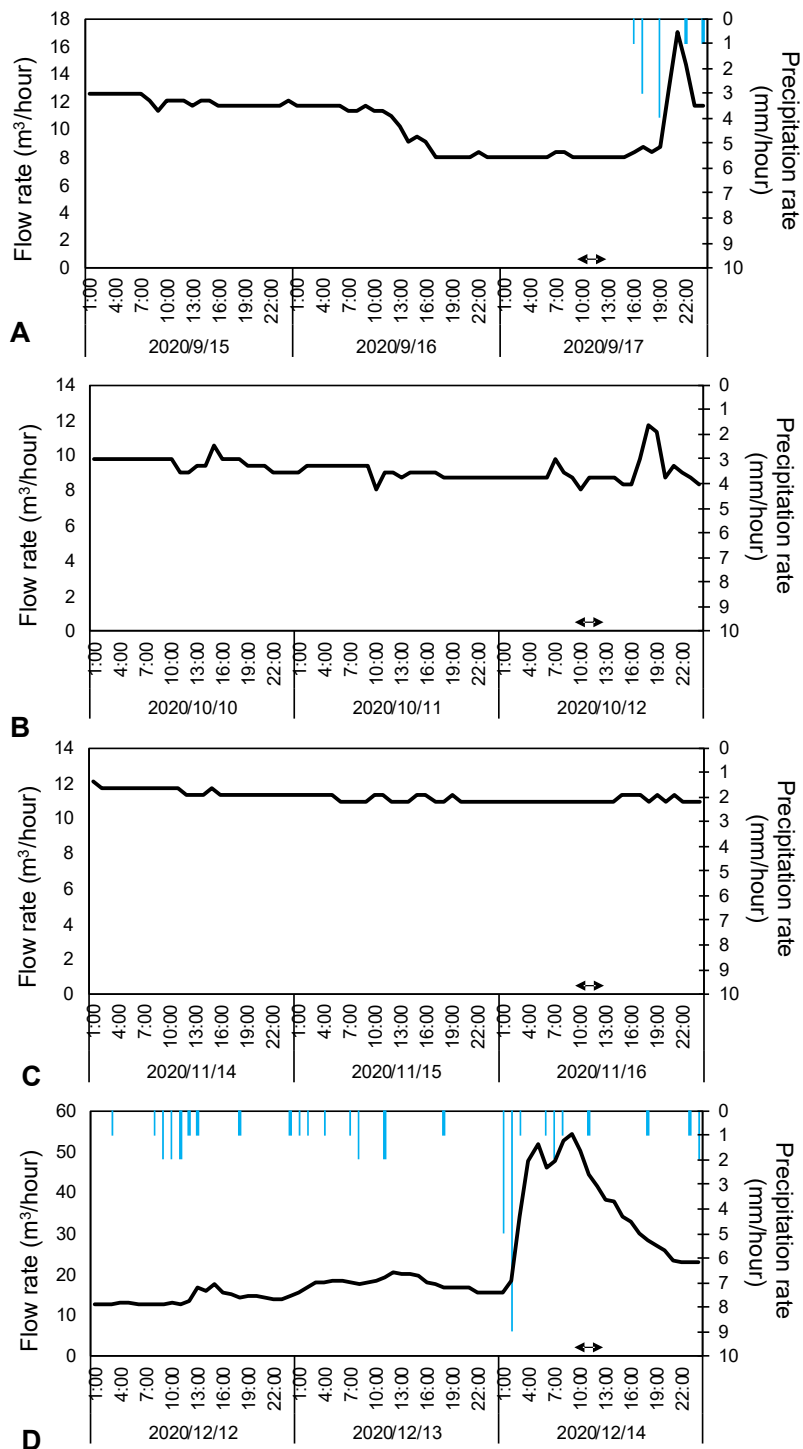


Figure S1-3. Changes in flow rate and precipitation around site O1 from two days before to the day of sample collection in Sept. 2020 (A), Oct. 2020 (B), Nov. 2020 (C), and Dec. 2020 (D). Arrows indicate the time of sampling (10:30-12:30).

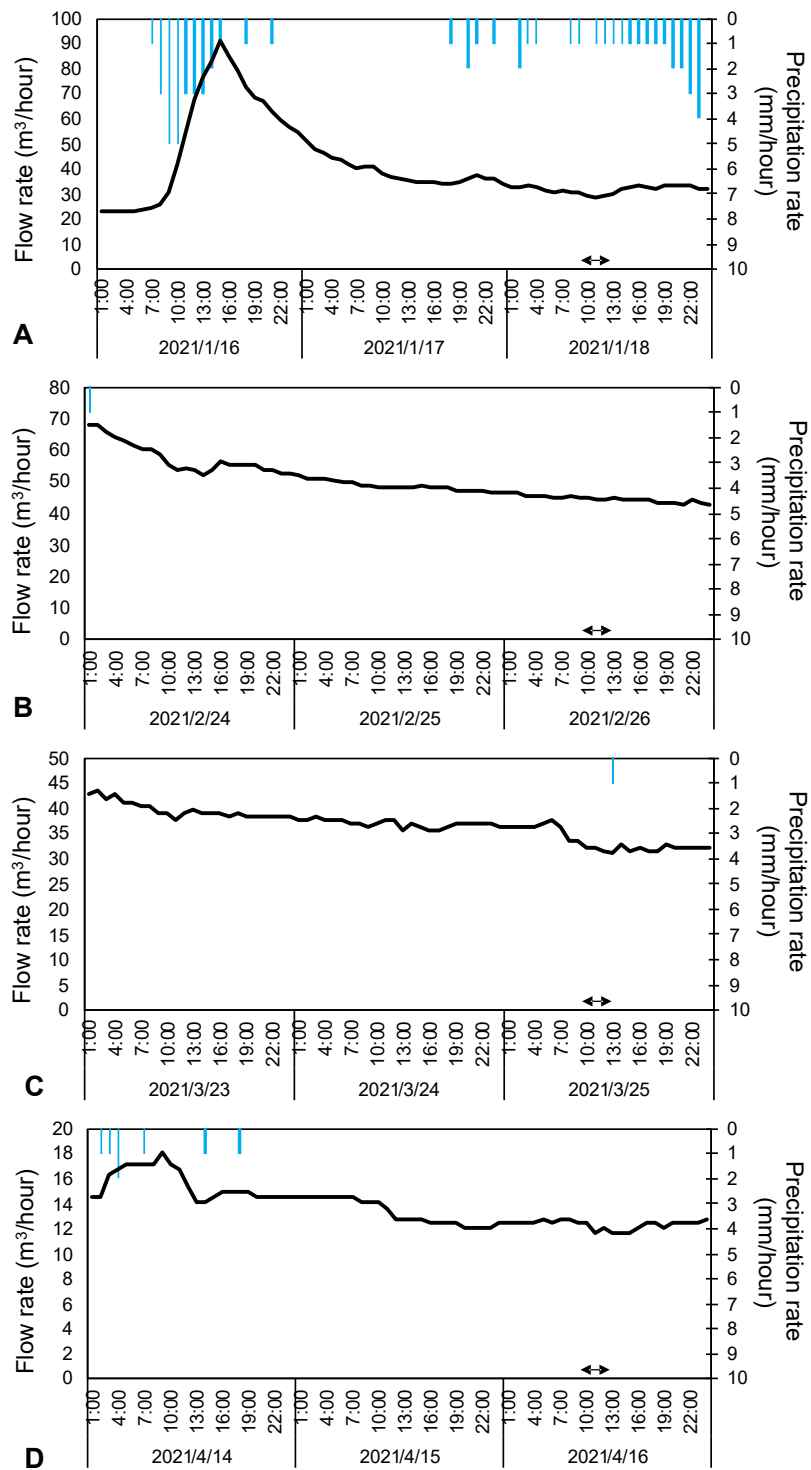


Figure S1-4. Changes in flow rate and precipitation around site O1 from two days before to the day of sample collection in Jan. 2021 (A), Feb. 2021 (B), Mar. 2021 (C), and Apr. 2021 (D). Arrows indicate the time of sampling (10:30-12:30).

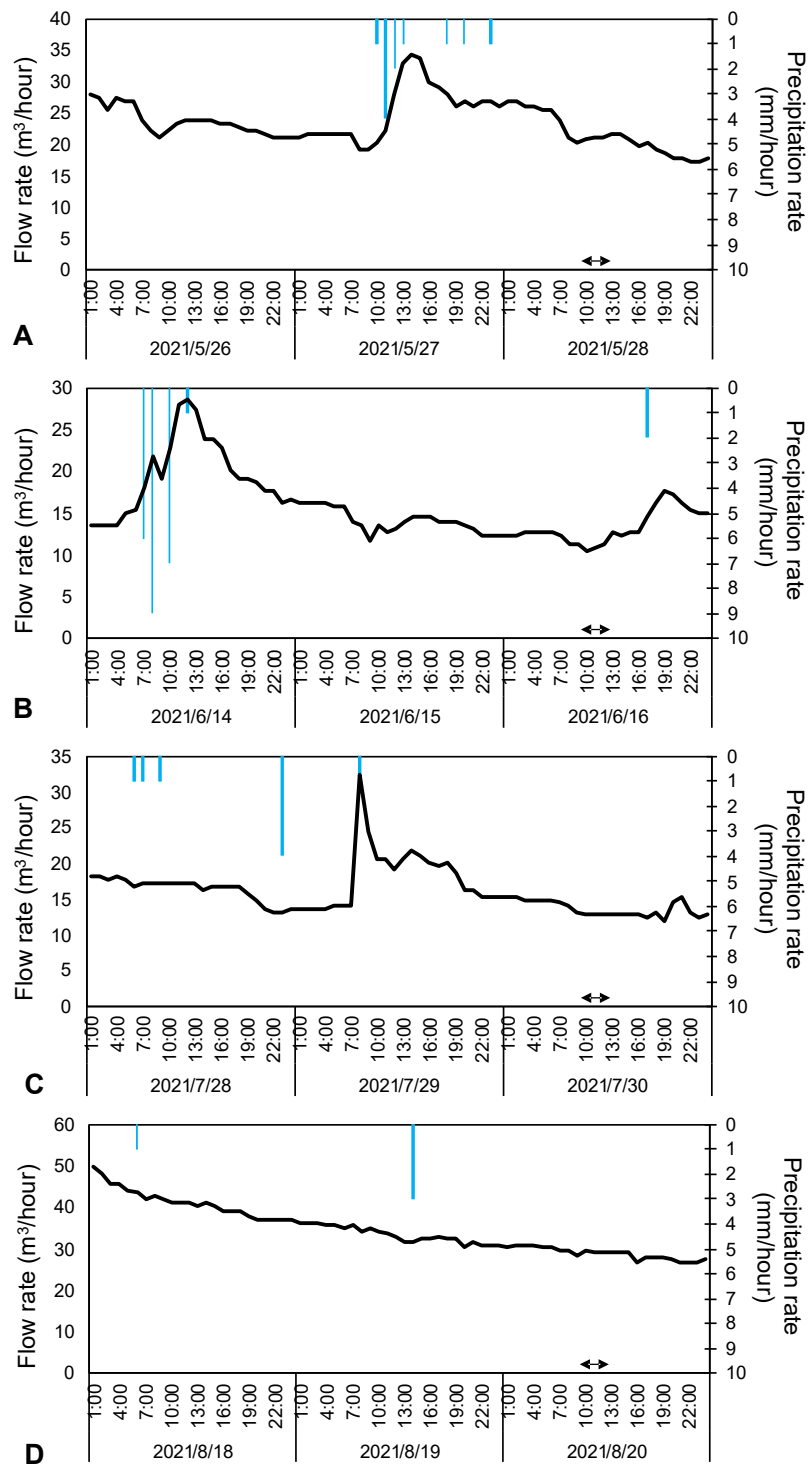


Figure S1-5. Changes in flow rate and precipitation around site O1 from two days before to the day of sample collection in May 2021 (A), June 2021 (B), July 2021 (C), and Aug. 2021 (D). Arrows indicate the time of sampling (10:30-12:30).

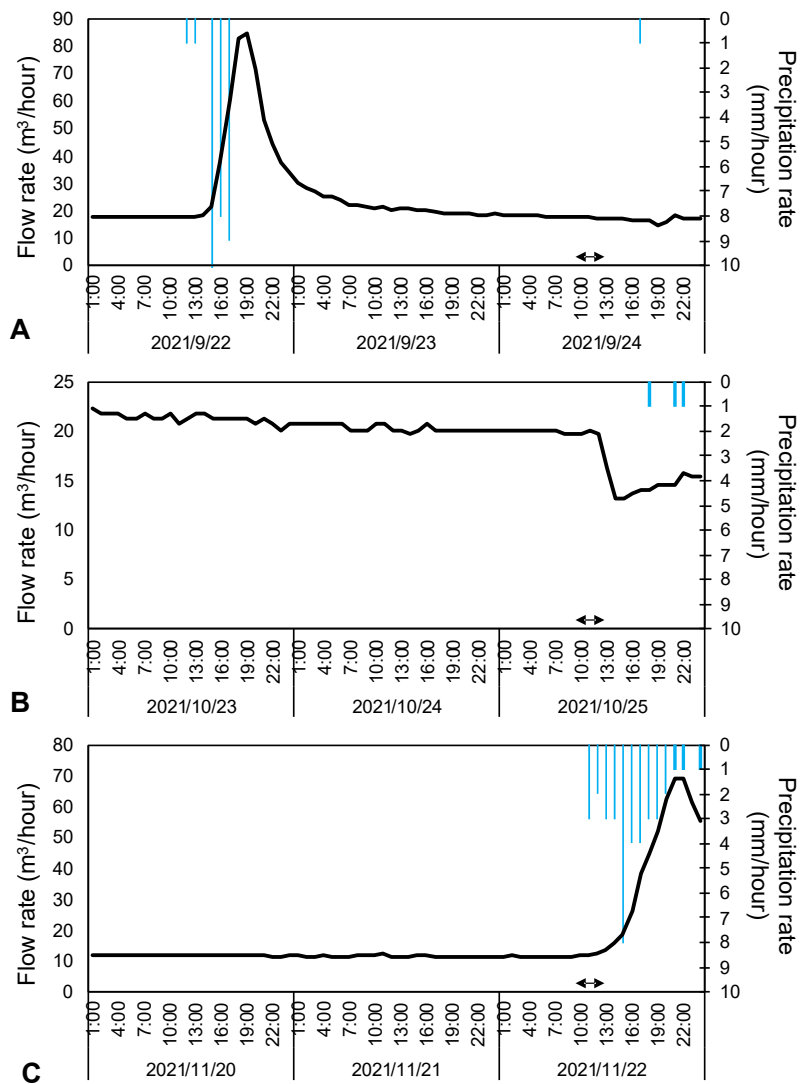


Figure S1-6. Changes in flow rate and precipitation around site O1 from two days before to the day of sample collection in Sept. 2021 (A), Oct. 2021 (B), and Nov. 2021. Arrows indicate the time of sampling (10:30-12:30).

Figure S2. Averaged abundance values of each animal species-like sequence reads obtained from the samples at each site. Only positive samples were used to calculate the averaged abundance values, and if a particular species-like read was not obtained from the sample of interest, the value within that sample was set to 0.

Supplementary Tables

Table S1. Primer sequences for the metabarcoding assay in this study.

Assay	Name	Function ^a	Sequence (5' → 3') ^b	Temperature for annealing and extension		Reference
Metabarcoding	MiMammal-U-F	F	<u>TTTCTGTTGGTGCTGATATTGCGGGTTGGT</u> AAATTTTCGTGCCAGC	65 °C and 72 °C		Ushio et al., 2017
	MiBird-U-F	F	<u>TTTCTGTTGGTGCTGATATTGCGGGTTGGT</u> AAATCTTGTGCCAGC			Ushio et al., 2018
	MiMammal-U-R/ MiBird-U-R	R	<u>ACTTGCCTGTGCTCTATCTTCCATAGTGG</u> GGTATCTAATCCCAGTTTG			Miya et al., 2015
qPCR – phiX 174		F	ACAAAGTTTGGATTGCTACTGACC	60 °C	LOD ^c : < 10 copies PCR eff ^d .: 104% R ^{2e} : 1.00	Verreault et al. 2010
		R	CGGCAGCAATAAACTCAACAGG			
		TP*1	HEX-CTCTCGTGCTCGTCGCTGCGTTGA- TAMRA			
qPCR – HF183	HF183	F	ATCATGAGTTCACATGTCCG	56 °C	LOD: < 10 copies PCR eff.: 105% R ² : 0.99	Green et al., 2014
	BacR287	R	CTTCCTCTCAGAACCCCTATCC			
	BacP234MGB	TP*2	CTAATGGAACGCATCCC			
qPCR – Pig-2-Bac	Pig-2-Bac41F	F	GCATGAATTTAGCTTGCTAAATTTGAT	60 °C	LOD: < 10 copies PCR eff.: 78% R ² : 0.99	Mieszkin et al., 2009
	Pig-2-Bac163Rm	R	ACCTCATAACGGTATTAATCCGC			
	Pig-2-Bac113MGB	TP*2	TCCACGGGATAGCC			
qPCR – BacR	BacR_f	F	GCGTATCCAACCTTCCCG	60 °C	LOD: < 10 copies PCR eff.: 81% R ² : 0.99	Reischer et al., 2006
	BacR_r	R	CATCCCCATCCGTTACCG			
	BacR_p	TP*2	CTTCCGAAAGGGAGATT			

^a: F, R, and TP denote forward primer, reverse primer, and TaqMan probe, respectively. TaqMan probes were labeled with 5'-HEX (hexachlorofluorescein) and 3'-TAMRA (6-carboxytetramethylrhodamine) (*¹) or 5'-FAM (6-carboxyfluorescein) and 3'-MGB-NFQ (minor groove binder-non fluorescent quencher nonfluorescent quencher) (*²).

^b: Underlined sequences were substituted from the original sequences to allow annealing of barcoding primers used in the second round PCR.

^c: LOD stands for the limit of detection, which is defined as the copy number that can be detected with a 95% probability.

^d: PCR eff. stands for PCR efficiency, which was calculated as follows; PCR eff. (%) = $100 \times (10^{(-1/\text{slope})} - 1)$, where “slope” refers to the slope value for the standard curve.
^e: The coefficient of determination for the standard curve.

Table S2. Averaged number of species observed from the samples at each site.

	O1 (<i>N</i> = 23)	O2 (<i>N</i> = 23)	O3 (<i>N</i> = 23)	J (<i>N</i> = 20)	S (<i>N</i> = 19)
No. of positive samples	13	17	15	15	14
No. of positive species (Average ± SD)	6.2 ± 1.8	4.6 ± 1.9	5.1 ± 0.4	6.4 ± 2.4	4.9 ± 1.9

Table S3. Positive rates and abundance values of each human-, livestock (pig, cattle, and chicken)-, and Bacterial 16S rRNA gene-like sequence read obtained by the metabarcoding assay from surface water samples ($N = 108$).

		All sites ($N = 108$)	O1 ($N = 23$)	O2 ($N = 23$)	O3 ($N = 23$)	J ($N = 20$)	S ($N = 19$)
Any ^a	No. of	74 (69%)	13 (57%)	17 (74%)	15 (65%)	15 (75%)	14 (74%)
	No. of reads ^b	5,975 (606-	3,975 (787-	6,994 (2,147-	4,519 (606-9,491)	9,400 (2,189-	4,483 (651-
Human	No. of	73 (68%)	13 (57%)	16 (70%)	15 (65%)	15 (75%)	14 (74%)
	Abundance ^c	44.0% (1.1-	54.4% (5.1-	24.5% (2.6-76.5%)	55.0% (11.8-	45.6% (17.0-	43.3% (1.1-
Pig/ wild boar	No. of	66 (61%)	12 (52%)	17 (74%)	13 (57%)	12 (60%)	12 (63%)
	Abundance	19.3% (0.5-	11.2% (2.5-	54.1% (4.1-93.7%)	7.2% (0.5-37.8%)	4.6% (0.5-18.5%)	5.7% (1.4-19.4%)
Cattle	No. of	23 (21%)	7 (30%)	3 (13%)	5 (22%)	7 (35%)	1 (5%)
	Abundance	1.9% (0.5-10.5%)	1.8% (0.5-7.4%)	0.7% (0.6-0.8%)	0.7% (0.5-1.0%)	3.5% (0.6-10.5%)	0.9%
Chicken	No. of	47 (44%)	11 (48%)	12 (52%)	13 (57%)	10 (50%)	6 (32%)
	Abundance	4.1% (0.5-85.7%)	1.8% (0.5-6.1%)	9.0% (0.6-85.7%)	2.8% (0.6-9.4%)	3.3% (0.9-11.4%)	2.9% (0.8-8.7%)
Bacterial rRNA gene	No. of	15 (14%)	3 (13%)	3 (13%)	3 (13%)	1 (5%)	5 (26%)
	Abundance	4.0% (0.6-19.9%)	2.8% (1.6-5.1%)	0.9% (0.6-1.1%)	5.4% (1.5-12.9%)	5.2%	5.5% (0.9-19.9%)

^a: Including target animal other than human or livestock, i.e., terrestrial mammalian and avian species.

^{b, c}: Averaged values (minimum value-maximum value) among positive samples are shown.

^c: Abundance was calculated by subtracting number of the animal-like reads from number of total reads.

Table S4. Positive rates and abundance values of terrestrial mammal-like sequence read obtained by the metabarcoding assay from surface water samples ($N = 108$).

		All sites ($N = 108$)	O1 ($N = 23$)	O2 ($N = 23$)	O3 ($N = 23$)	J ($N = 20$)	S ($N = 19$)
Any ^a	No. of	74 (69%)	13 (57%)	17 (74%)	15 (65%)	15 (75%)	14 (74%)
	No. of reads ^b	5,975 (606-	3,975 (787-	6,994 (2,147-	4,519 (606-9,491)	9,400 (2,189-	4,483 (651-
Sika deer	No. of	26 (24%)	4 (17%)	5 (22%)	2 (9%)	7 (35%)	8 (42%)
	Abundance ^c	4.0% (0.5-20%)	1.2% (0.5-2.5%)	3.3% (0.5-12.8%)	0.5 (0.5-0.6%)	8.0% (0.7-20%)	3.2% (0.7-10.6%)
Civet	No. of	20 (19%)	4 (17%)	3 (13%)	5 (22%)	6 (30%)	2 (11%)
	Abundance	4.7% (0.5-44.7%)	1.0% (0.8-1.3%)	4.0% (0.6-10.0%)	11.2% (0.8-	3.4% (0.6-13.5%)	0.7% (0.5-0.8%)
Raccoon dog	No. of	8 (7%)	1 (4%)	0 (0%)	3 (13%)	3 (15%)	1 (5%)
	Abundance	20.8% (0.5-	67.6%	-	4.9% (0.7-9.3%)	0.9% (0.5-1.1%)	81.9%
Serow	No. of	13 (12%)	4 (17%)	1 (4%)	1 (4%)	3 (15%)	4 (21%)
	Abundance	2.4% (0.6-11.6%)	3.6% (0.6-11.6%)	7.3%	1.5%	1.5% (0.6-3.3%)	0.8% (0.6-1.1%)
Vole	No. of	11 (10%)	1 (4%)	2 (9%)	1 (4%)	5 (25%)	2 (11%)
	Abundance	2.8% (0.6-9.6%)	0.8%	2.7% (1.8-3.6%)	1.9%	1.5% (1.5-9.6%)	2.9% (0.6-5.2%)
Field	No. of	1 (1%)	0 (0%)	1 (4%)	0 (0%)	0 (0%)	0 (0%)
	Abundance	1.9%	-	1.9%	-	-	-
Rat	No. of	5 (5%)	0 (0%)	3 (13%)	0 (0%)	2 (10%)	0 (0%)
	Abundance	1.6% (0.5-3.7%)	-	1.2% (0.5-2.4%)	-	2.3% (0.9-3.7%)	-
Dog	No. of	8 (7%)	5 (22%)	0 (0%)	1 (4%)	2 (10%)	0 (0%)
	Abundance	2.9% (0.5-13%)	3.5% (0.5-10.9%)	-	0.8%	2.2% (0.7-3.7%)	-

^a: Including target animal other than terrestrial mammalian, i.e., human, livestock, and avian species.

^{b, c}: Averaged values (minimum value-maximum value) among positive samples are shown.

^c: Abundance was calculated by subtracting number of the animal-like reads from number of total reads.

Table S5. Positive rates and abundance values of wild bird-like sequence read obtained by the metabarcoding assay from surface water samples ($N = 108$).

		All sites ($N = 108$)	O1 ($N = 23$)	O2 ($N = 23$)	O3 ($N = 23$)	J ($N = 20$)	S ($N = 19$)
Any ^a	No. of positive	74 (69%)	13 (57%)	17 (74%)	15 (65%)	15 (75%)	14 (74%)
	No. of reads ^b	5,975 (606-78,603)	3,975 (787-10,991)	6,994 (2,147-40,449)	4,519 (606-9,491)	9,400 (2,189-78,603)	4,483 (651-11,138)
Duck	No. of positive	70 (65%)	13 (57%)	14 (61%)	14 (61%)	15 (75%)	14 (74%)
	Abundance ^c	10.5% (0.8-84.1%)	6.9% (1.0-38.4%)	7.4% (0.9-28.2%)	12.5% (1.3-	13.1% (0.8-44.2%)	12.0% (0.8-84.1%)
Coot	No. of positive	6 (6%)	1 (4%)	1 (4%)	0 (0%)	3 (15%)	1 (5%)
	Abundance	2.6% (0.6-5.5%)	1.7%	0.6%	-	3.2% (1.9-5.5%)	3.9%
Mergin	No. of positive	3 (3%)	1 (4%)	0 (0%)	0 (0%)	2 (10%)	0 (0%)
	Abundance	1.4% (0.8-1.9%)	1.5%	-	-	1.4% (0.8-1.9%)	-
Ardea	No. of positive	12 (11%)	3 (13%)	1 (4%)	2 (9%)	3 (15%)	3 (16%)
	Abundance	1.9% (0.5-6.0%)	0.7% (0.5-1.1%)	0.7%	0.6% (0.6-0.7%)	4.3% (1.5-6.0%)	2.0% (0.5-4.3%)
Pigeon	No. of positive	8 (7%)	2 (9%)	2 (9%)	1 (4%)	4 (20%)	0 (0%)
	Abundance	3.3% (0.6-23.3%)	0.6% (0.6-0.6%)	0.6% (0.6-0.6%)	0.6%	6.7% (0.6-23.3%)	-
Crow	No. of positive	1 (1%)	0 (0%)	0 (0%)	1 (4%)	0 (0%)	0 (0%)
	Abundance	1.0%	-	-	1.0%	-	-

^a: Including target animal other than wild bird, i.e., human, livestock, and terrestrial mammalian species.

^{b, c}: Averaged values (minimum value-maximum value) among positive samples are shown.

^c: Abundance was calculated by subtracting number of the animal-like reads from number of total reads.

Table S6. *P* values obtained with Tukey’s multiple-comparison test to determine whether the normalized numbers of MST markers and corresponding animal-like sequence reads were significantly different over the sampling sites.

Combination of the sites	HF183	Human	Pig-2-Bac	Pig	BacR	Cattle
O2-O1	0.000 ^a	0.001	0.594	0.204	0.059	0.100
O3-O1	0.873	0.994	0.823	0.682	0.070	0.164
J-O1	0.108	0.986	0.157	0.057	0.355	1.000
S-O1	0.847	0.988	0.000	0.955	0.836	0.012
O3-O2	0.000	0.000	0.996	0.004	1.000	1.000
S-O2	0.000	0.000	0.000	0.033	0.450	0.860
J-O2	0.000	0.000	0.002	0.000	0.000	0.092
J-O3	0.006	1.000	0.007	0.576	0.000	0.156
S-O3	0.263	1.000	0.000	0.972	0.480	0.794
S-J	0.583	1.000	0.003	0.244	0.033	0.010

^a: The colored columns indicate that the *p* value is < 0.05.

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(10):3086–3094. <https://doi.org/10.1128/AEM.04137-13>
2. Mieszkin S, Furet JP, Corthier G, Gourmelon M (2009) Estimation of pig fecal contamination in a river catchment by real-time PCR using two pig-specific *Bacteroidales* 16S rRNA genetic markers. *Appl Environ Microbiol* 75(10):3045–3054. <https://doi.org/10.1128/AEM.02343-08>
3. Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *R Soc Open Sc* 22(7):150088. <https://doi.org/10.1098/rsos.150088>
4. Reischer GH, Kasper DC, Steinborn R, Mach RL, Farnleitner AH (2006) Quantitative PCR method for sensitive detection of ruminant fecal pollution in freshwater and evaluation of this method in alpine karstic regions. *Appl Environ Microbiol* 72(8):5610–5614. <https://doi.org/10.1128/AEM.00364-06>
5. Ushio M, Fukuda H, Inoue T, Makoto K, Kishida O, Sato K, Murata K, Nikaido M, Sado T, Sato Y, Takeshita M, Iwasaki W, Yamanaka H, Kondoh M, Miya M (2017) Environmental DNA enables detection of terrestrial mammals from forest pond water. *Mol Ecol Resour* 17:e63–e75. <https://doi.org/10.1111/1755-0998.12690>

6. Ushio M, Murata K, Sado T, Nishiumi I, Takeshita M, Iwasaki W, Miya M (2018) Demonstration of the potential of environmental DNA as a tool for the detection of avian species. *Sci Rep* 8(1):4493. <https://doi.org/10.1038/s41598-018-22817-5>
7. Verreault D, Rousseau GM, Gendron L, Massé D, Moineau S, Duchaine C (2010) Comparison of polycarbonate and polytetrafluoroethylene filters for sampling of airborne bacteriophages. *Aerosol Sci Technol* 44(3):197-201. <https://doi.org/10.1080/02786820903518899>