Electronic Supplementary Information

Surveillance and seasonal correlation of Rotavirus A with coliphages and coliforms in two lakes in highly urbanized regions of western India

by

Lakshmi Pisharody[1], Sumathi Suresh[1], and Suparna Mukherji[1],*

[1]Environmental Science and Engineering Department (ESED),
IIT Bombay, Powai, Mumbai - 400 076, India

Submitted to
Environmental Science: Water Research & Technology
November, 2021

Text S1. Isolation and concentration of coliphage SUSP2 from Powai Lake

SUSP2 was isolated from Powai lake. One litre of Powai lake water was collected between 7.30 am to 8.30 am. The isolation protocol was started within an hour of collection. Multiple cycles of purification and concentration were required for isolation. Powai lake water (1L) was centrifuged at 10,000 g for 10 min to remove large particles. Thereafter, the sample was filtered using a 0.22 µm membrane filter to remove bacteria and fungi. The sample was subsequently enriched using 271 B media containing *E.coli* C3000 (ATCC® 15597™) in an exponential growth phase for 4 h at 37°C. Samples were filtered (0.22 µm) and subsequently plaque assay was performed to obtain discrete plaques for isolation of coliphages. After 24 h of incubation, a single discrete plaque was picked employing an autoclaved Pasteur pipette and the plaque was dispensed in 100 µL of SM buffer, and a few drops of $CHCl_3$ was added to prevent bacterial contamination [1]. Various dilutions of the virus sample were then incubated in a heating block (37°C, 10 min) along with 300 µL of the host, and subsequently plaque assay was performed. Multiple rounds of isolation were conducted to obtain a pure culture of the virus. Finally, the virus sample was concentrated using a high speed ultracentrifuge (85,940 g 4°C, 2 h). After discarding the supernatant, the pellet was hydrated using 500 µL of SM buffer and was stored

at 4°C. The DNA of the virus was thereafter isolated by the phenol-chloroform method [2] and was sent for identification to SciGenom, Cochin, India. SciGenome performed next-generation sequencing for the phage and on blast analysis, the coliphage isolated from Powai lake showed maximum similarity (85%) with the phage SUSP2 which was identified as a 'superspreader' by Keen et al., (2017). DNA concentration was quantified by first measuring the absorbance at 260 nm using 8453 UV-Visible spectrophotometer (Agilent, USA) in a 50 µL quartz cuvette. The concentration of the DNA was thereafter quantified using equation S1 [2,3]

$$Concentration\ of\ DNA\ (\mu g\ mL) = O.D \times dilution\ factor \times 50 \qquad Eqn.\ S1$$

Table S1 Composition of SM buffer

Reagent	Final concentration (mM)
NaCl	100
$MgSO_4.7H_20$	8
Tris-Cl (pH 7.5)	50

Table S2 Primer sequence for target viruses

Primers	Sequence	Target and species	Reference
Forward (5'-3')	ACGCAATCACGACCTTCA	RVA NSP5 gene	Ye et al. 2012[4]
Reverse (5'-3')	CTCGGGTAGTGTTTCCTACTTT	RVA NSP5 gene	Ye et al. 2012[4]
Forward (5'-3')	GTCGCGGTAATTGGCGC	MS2 Assembly protein	O'Connell et al. 2006[5]
Reverse (5'-3')	GGCCACGTGTTTTGATCGA	MS2 Assembly protein	O'Connell et al. 2006[5]
Forward (5'-3')	ATGCAGTAGACCGTACCAGC	SUSP2 Putative tail protein	This study
Reverse (5'-3')	AGGTCAGCGTACAGAACACC	SUSP2 Putative tail protein	This study

Table S3 Standard curve equations for the viruses

Species	Equation	R^2	% Efficiency
SUSP2	y = -3.1745x + 42.458	0.999	106
MS2	y = -3.0455x + 37.071	0.998	112
RVA	y = -3.0552x + 38.404	0.994	112

Where: y is Ct values, x is copies/reaction for RVA and MS2; x is copies/mL for SUSP2
Each reaction corresponds to 20 µL which can be calculated for 1 mL and can be expressed as copies/mL

Table S4 Concentration of viruses across the lakes for a period of two years (Jan 2018-January 2020)

Months	RVA (Ln N)		MS2 (Ln N)		SUSP2 (Ln N)	
Masunda Lake						
	Year 1	**Year 2**	**Year 1**	**Year 2**	**Year 1**	**Year 2**
January	6.93±0.15	BDL	6.03±0.06	3.01±0.96	6.64±0.16	2.96±0.14
February	8.87±0.23	13.12±0.10	6.31±0.23	5.36±0.10	7.99±0.11	8.11±0.04
March	7.70±0.14	BDL	1.79±0.50	5.74±0.06	5.44±0.02	7.97±0.11
April	3.78±0.52	BDL	1.70±0.82	BDL	3.46±0.39	1.39±0
May	5.21±0.37	BDL	1.50±0.56	BDL	1.39±0	4.54±0.29
June	4.80±0.06	4.78±0	1.79±0.50	2.08±0	BDL	BDL
July	5.26±0.03	0.94±1.00	4.11±.10	1.080.29	7.59±0.12	BDL
August	2.71±1	4.38±0.35	1.61±0.40	BDL	4.72±0.04	BDL
September	3.79±0.13	4.76±0	1.20±0.61	2.08	3.79±0.13	4.62±0
October	BDL	BDL	2.40±0.45	BDL	4.54±0.06	BDL
November	4.79±0.01	BDL	2.08±0	BDL	4.63±0.01	BDL
December	4.79±0	5.82±0.20	2.08±0	5.19±0.46	4.63±0	BDL
Powai Lake						
	Year 1	**Year 2**	**Year 1**	**Year 2**	**Year 1**	**Year 2**
January	BDL	5.65±0.70	BDL	2.04±0.92	3.19±0.11	2.29±0.18
February	8.03±0.16	8.55±0.15	4.88±0.57	0.70±0.53	5.89±0.48	5.87±0.25
March	9.87±0.09	9.24±0.02	9.57±0.82	0.07±0.07	8.30±0.12	9.24±0.02
April	5.34±0.75	5.04±.36	6.04±0.33	3.29±0.99	2.44±0.04	9.96±0.04
May	BDL	BDL	3.70±0.31	BDL	3.30±0	7.34±0.38
June	5.61±0.18	BDL	2.62±0.02	0.92±1.00	BDL	3.94±0.17
July	3.93±1.00	4.04±0.19	1.50±0.78	1.94±1.00	4.30±0.12	BDL
August	BDL	1.00±0.48	1.61±0.20	2.12±1.00	3.28±0.09	BDL
September	BDL	BDL	1.45±0.11	1.24±0.01	BDL	BDL
October	6.20±0.14	BDL	2.21±0.94	BDL	BDL	BDL
November	2.48±1.00	BDL	0.64±1.00	4.01±0.48	1.71±0.06	BDL
December	5.13±0.04	0.16±1.00	6.77±0.14	4.70±1.00	3.27±0.03	BDL

N: Number of copies of virus/ 100 mL, *BDL: Below Detection Limit, SE represents variation in virus concentration, Ln (N), from the same composite samples

Text S2. Principal component analysis (PCA) of the fecal indicators

PCA was performed to further understand the major parameters that affect fecal contamination in lake water samples. The matrix for the PCA contained variables, such as, RVA, MS2, SUSP2, total coliforms, fecal coliforms, and COD. The concentration of coliforms and coliphages (N) obtained from both the lakes were transformed to ln N before performing PCA [6]. The first two principal components (PCs), i.e., principal component 1 and principal component 2 (PC1 and PC2) in both the lakes, Powai lake, and Masunda lake, showed eigenvalues greater than one. The eigenvalues for Powai lake were 2.21 and 1.71 for PC1 and PC2, respectively. For Masunda lake the eigenvalues were 2.07 and 1.92, respectively, for PC1 and PC2. Since the eigenvalues for PC1 and PC2 were greater than 1, these PCs captured the maximum variance in the data [7]. More than 65% of the variance in the data was explained by PC1 and PC2 for Powai lake. For Masunda lake 66% of the variance in the data was explained by the first two PCs (PC1 and PC2). PC1 which explained the maximum variance in the data for Powai lake was influenced mainly by MS2, RVA, and SUSP2 populations with loading in the range of 0.48-0.60. For Masunda lake the maximum variance was influenced mainly by SUSP2, RVA, COD, and MS2 with loadings (eigenvectors) ranging from 0.36-0.55 (Table S5, in ESI). The extracted eigenvector (Table S5) provides the coefficients for Eqs S1-S4 representing PC1 and PC2, where Eq. 1 and 2 are for Powai lake and Eq. 3 and 4 are for Masunda lake.

$$PC1 = 0.59 \times RVA + 0.60 \times MS2 + 0.48 \times SUSP2 + 0.13 \times TC + 0.06 \times FC + 0.20 \times COD \quad (Eqn\ S1)$$

$$PC2 = -0.02 \times RVA + 0.05 \times MS2 - 0.37 \times SUSP2 + 0.67 \times TC + 0.61 \times FC + 0.18 \times COD \quad (Eqn\ S2)$$

$$PC1 = 0.51 \times RVA + 0.36 \times MS2 + 0.55 \times SUSP2 - 0.01 \times TC - 0.19 \times FC + 0.52 \times COD \quad (Eqn\ S3)$$

$$PC2 = 0.20 \times RVA + 0.45 \times MS2 - 0.13 \times SUSP2 + 0.61 \times TC + 0.59 \times FC - 0.14 \times COD \quad (Eqn\ S4)$$

Fig. S1 the biplot for lakes is essentially a 2D visualization, which includes one point for each of the observations, with the coordinates indicating the scores for each observation for the two principal components represented in the plot. In the biplot, all the six variables (RVA, MS2, SUSP2, TC, FC, and COD) are plotted as vectors, wherein the length and the direction of each vector represent the contribution of the corresponding variable to each of the principal components. PC1 which is depicted on the horizontal axis comprises positive coefficients for all variables for Powai lake. Thus, the vectors corresponding to these variables lie on the right side of the graph, the largest positive value corresponds to MS2 and RVA (0.60). The PC2 on the vertical axis has positive coefficients for 4 (MS2, TC, FC, and COD) variables. Similarly, for Masunda lake, principal component 1 comprises positive coefficients for all 4 variables, RVA, SUSP2, MS2, and COD. The two largest values correspond to RVA (0.55) and SUSP2 (0.44).

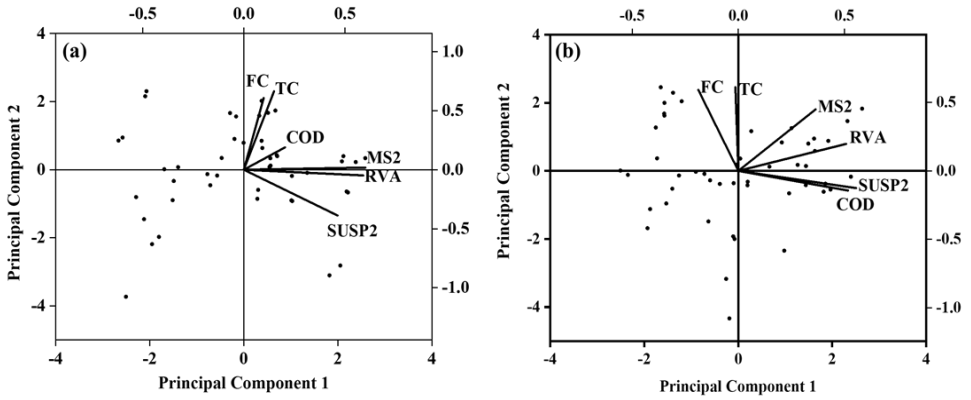

Fig. S1 PCA correlation biplot for (a) Powai lake and, (b) Masunda lake

The first two PCs account for a total variation of 65.18% (PC1: 36.77%, PC2: 28.41%) in Powai Lake and 66.59% (PC1:34.54%, PC2:32.05%) in Masunda Lake. The secondary X-axis and Y-axis represent the PC1 loadings and PC2 loadings, respectively for the variables

Table S5 Extracted eigenvectors for principal component 1 (PC1) and principal component 2 (PC2)

	Coefficient of PC1	Coefficient of PC2
Powai lake		
RVA (Copies/100 mL)	0.59	-0.02
MS2 (Copies/100 mL)	0.60	0.05
SUSP2 (Copies/100 mL)	0.47	-0.37
TC (CFU/100 mL)	0.13	0.67
FC (CFU/100 mL)	0.06	0.61
COD (mg/L)	0.20	0.18
Masunda lake		
RVA (Copies/100 mL)	0.51	0.20
MS2 (Copies/100 mL)	0.36	0.45
SUSP2 (Copies/100 mL)	0.55	-0.13
TC (CFU/100 mL)	-0.01	0.61
FC (CFU/100 mL)	-0.19	059
COD (mg/L)	0.52	-0.14

References

1. K. E. Wommack, K. E.Williamson, R. R. Helton, S. R., Bench, and D. M. Winget, in *Methods for the isolation of viruses from environmental samples*, Bacteriophages methods and protocols, isolation, characterization, and interactions. , ed. M. R. J.,Clokie, and A. M. Kropinski, Humana Press, New York, USA, 2009, pp.3–15.,

2. J. Sambrook, and D. W. Russell, Molecular cloning laboratory manual, ed. M. R. Green, and J. Sambrook, J. Cold spring harbour laboratory press, Cold Spring Harbour, New York, 2012, pp. 1-15.

3. Biosystems, A. Creating Standard Curves with Genomic DNA or Plasmid DNA Templates for Use in Quantitative PCR. (2003).

4. X. Y. Ye, X. Ming, Y. L. Zhang, Xiao, W. Q. Huang, X. N. Cao, Y. G. and K. D. Gu, *Curr. Microbiol.*, 2012, **65**, 244–253.

5. K. P. O'Connell, J. R. Bucher, P. E. Anderson, C. J. Cao, A. S. Khan, M. V. Gostomski, and J. J. Valdes, *Appl. Environ. Microbiol.*, 2006, **72**, 478–83.

6. A. Ahmad, A. C. Dada, G. Usup, and L. Y. Heng, *Springerplus* **2**, 1–18 (2013).

7. Jolliffe, I. T. Principal Component Analysis, *Encycl. Stat. Behav. Sci.*, 2002.