Metagenomes of Microbial Communities in Arsenic- and Pathogen-Contaminated Well and Surface Water from Bangladesh

Alice C. Layton,[a,b,c] Archana Chauhan,[a,d] Daniel E. Williams,[a] Brian Mailloux,[e] Peter S. K. Knappett,[f] Andrew S. Ferguson,[g] Larry D. McKay,[c] M. Jahangir Alam,[h] Kazi Matin Ahmed,[h] Alexander van Geen,[i] Gary S. Sayler[a,b,d]

Center for Environmental Biotechnology, University of Tennessee, Knoxville, Tennessee, USA[a]; Department of Microbiology, University of Tennessee, Knoxville, Tennessee, USA[b]; Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, Tennessee, USA[c]; Joint Institute for Biological Sciences, University of Tennessee, Knoxville, and Oak Ridge National Laboratory, Knoxville, Tennessee, USA[d]; Department of Environmental Science, Barnard College, New York, New York, USA[e]; Department of Geology and Geophysics, Texas A&M University, College Station, Texas, USA[f]; Civil Engineering & Engineering Mechanics, Columbia University, New York, New York, USA[g]; Department of Geology, University of Dhaka, Dhaka, Bangladesh[h]; Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York, USA[i]

The contamination of drinking water from both arsenic and microbial pathogens occurs in Bangladesh. A general metagenomic survey of well water and surface water provided information on the types of pathogens present and may help elucidate arsenic metabolic pathways and potential assay targets for monitoring surface-to-ground water pathogen transport.

Bangladesh has two microbial-based public health water resource crises. Diarrheal diseases (1), often attributable to poor sanitary conditions and fecal contamination of drinking water, remain a leading cause of mortality for children younger than 5 years (2, 3). In locations with poor sanitation, groundwater is a cleaner source than surface water; however, shallow wells may still contain waterborne pathogens (4). In parts of Bangladesh, the drinking water situation is further complicated by the presence of geogenic arsenic above drinking water standards (>10 μg/liter), which causes serious and debilitating diseases (5, 6). The goal of this study was to generate metagenomic sequences from arsenic- and pathogen-contaminated ground and surface waters in Bangladesh (7).

Surface and ground water samples were collected 4 times between March 2009 and November 2012 in a 1-km^2 area of Bara Haldia, MatLab, Bangladesh (25.467°N 89.517°E) (4) representing both wet and dry seasons (8). In wells, the arsenic concentrations ranged from <1 to 218 parts per billion (ppb), and culturable *Escherichia coli* concentrations ranged from <1 to 100 most probable number (MPN)/100 ml. DNA was extracted from water filters using the Fast DNA spin kit (MP Biomedicals, CA) (2). DNA from 9 to 12 wells or 3 surface water samples was combined from each date to provide a composite sample. Metagenomic libraries were prepared using the Illumina Nextera DNA library preparation kit (Illumina, Inc., CA), and 2 × 100 bp reads were sequenced on an Illumina HiSeq 2000. The raw sequence reads from 8 metagenomic data sets were annotated using MG-RAST version 3.3.7 (9).

In surface water metagenomes, *Bacteria* was the dominant domain (94.2%), followed by *Eukaryota* (3.6%), DNA-based viruses (1.7%), and *Archaea* (0.4%). The most abundant genera in the surface water samples were *Acidovorax* (4.8%), *Mycobacterium* (4.1%), and *Burkholderia* (2.9%). Sequences matching well-known waterborne bacterial pathogens (10) found in surface water included: *Vibrio cholerae* and *Vibrio parahaemolyticus* (0.08%), *Salmonella enterica* (0.07%), *Clostridium difficile* (0.03%), *Cronobacter sakazakii* (0.03%), *Shigella flexneri* and *Shigella dysenteriae* (0.03%), *Staphylococcus aureus* (0.02%), *Campylobacter jejuni* (0.01%), and *Helicobacter pylori* (0.01%),

Consistent with the anaerobic conditions in ground water, archaea sequences comprised 20.4% of the sequences, with the genus *Methanosarcina* representing 4.1% of all well water metagenome sequences. Bacteria comprised 78.1%, eukaryotes 1.2%, and DNA-based viruses <0.1% of the sequences, with *Geobacter* (3.5%) and *Clostridium* (2.0%) being the second and third most abundant genera. Sequences matching the surface water pathogens were found in groundwater but generally at lower concentrations, which was expected due to die-off and filtering. Unexpectedly, four pathogens had 3- to 4-fold higher relative sequence abundances in ground than surface water: *C. difficile* (0.12%), *S. aureus* (0.04%) *C. jejuni* (0.03%), and *H. pylori* (0.03%).

In both surface and well water metagenomes, genes for arsenic metabolism, including arsenate reductase and arsenite oxidation, along with those for arsenic resistance, were present. These genetic pathways may affect arsenic mobility and toxicity (11). In summary, these metagenomic data sets have helped and can continue to further help elucidate the mechanisms of groundwater pathogen and arsenic contamination in Bangladesh.

Nucleotide sequence accession number. Nucleotide sequences obtained were deposited at the NCBI Sequence Read Archive under the accession no. SRP047074.

REFERENCES

1. Walker CL, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, O'Brien KL, Campbell H, Black RE. 2013. Global burden of childhood pneumonia and diarrhoea. Lancet **381**:1405–1416. http://dx.doi.org/10.1016/S0140-6736(13)60222-6.

2. Kosek M, Bern C, Guerrant RL. 2003. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bull. World Health Organ. **81**:197–204. http://www.scielosp.org/scielo.php?script=sci_arttext&pid=S0042-96862003000300010&lng=en&nrm=iso&tlng=en.

3. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, Jha P, Campbell H, Walker CF, Cibulskis R, Eisele T, Liu L, Mathers C. 2008. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet **375**:1969–1987. http://dx.doi.org/10.1016/S0140-6736(10)60549-1.

4. Ferguson AS, Layton AC, Mailloux BJ, Culligan PJ, Williams DE, Smartt AE, Sayler GS, Feighery J, McKay LD, Knappett PS, Alexandrova E, Arbit T, Emch M, Escamilla V, Ahmed KM, Alam MJ, Streatfield PK, Yunus M, van Geen A. 2012. Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater. Sci. Total Environ. **431**:314–322. http://dx.doi.org/10.1016/j.scitotenv.2012.05.060.

5. van Geen A, Ahmed EB, Pitcher L, Mey JL, Ahsan H, Graziano JH, Ahmed KM. 2014. Comparison of two blanket surveys of arsenic in tubewells conducted 12 years apart in a 25-km² area of Bangladesh. Sci. Total Environ. **488–489**:484–492. http://dx.doi.org/10.1016/j.scitotenv.2013.12.049.

6. Mailloux BJ, Trembath-Reichert E, Cheung J, Watson M, Stute M, Freyer GA, Ferguson AS, Ahmed KM, Alam MJ, Buchholz BA, Thomas J, Layton AC, Zheng Y, Bostick BC, van Geen A. 2013. Advection of surface-derived organic carbon fuels microbial reduction in Bangladesh groundwater. Proc. Natl. Acad. Sci. USA **110**:5331–5335. http://dx.doi.org/10.1073/pnas.1213141110.

7. van Geen A, Ahmed KM, Akita Y, Alam MJ, Culligan PJ, Emch M, Escamilla V, Feighery J, Ferguson AS, Knappett P, Layton AC, Mailloux BJ, McKay LD, Mey JL, Serre ML, Streatfield PK, Wu J, Yunus M. 2011. Fecal contamination of shallow tubewells in Bangladesh inversely related to arsenic. Environ. Sci. Technol. **45**:1199–1205. http://dx.doi.org/10.1021/es103192b.

8. Leber J, Rahman MM, Ahmed KM, Mailloux B, van Geen A. 2011. Contrasting influence of geology on E. coli and arsenic in aquifers of Bangladesh. Ground Water **49**:111–123. http://dx.doi.org/10.1111/j.1745-6584.2010.00689.x.

9. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA. 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics **9**:386. http://dx.doi.org/10.1186/1471-2105-9-386.

10. Cabral JPS. 2010. Water microbiology. Bacterial pathogens and water. Int. J. Environ. Res. Public Health **7**:3657–3703. http://dx.doi.org/10.3390/ijerph7103657.

11. Cai L, Liu G, Rensing C, Wang G. 2009. Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. BMC Microbiol. **9**:4. http://dx.doi.org/10.1186/1471-2180-9-4.