Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jes

JES JOURNAL OF ENVIRONMENTAL SCIENCES www.jesc.ac.cn

Evaluation of bacterial pathogen diversity, abundance and health risks in urban recreational water by amplicon next-generation sequencing and quantitative PCR

*Qijia Cui, Tingting Fang, Yong Huang, Peiyan Dong, Hui Wang**

State Key Joint Laboratory on Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China. E-mail: cuiqj11@gmail.com

ARTICLE INFO

Article history:
Received 4 July 2016
Revised 26 September 2016
Accepted 10 November 2016
Available online 29 November 2016

Keywords:
Human pathogens
Urban water
Recreational activities
16S rRNA gene targeted NGS
qPCR
Quantitative microbial risk assessment

ABSTRACT

The microbial quality of urban recreational water is of great concern to public health. The monitoring of indicator organisms and several pathogens alone is not sufficient to accurately and comprehensively identify microbial risks. To assess the levels of bacterial pathogens and health risks in urban recreational water, we analyzed pathogen diversity and quantified four pathogens in 46 water samples collected from waterbodies in Beijing Olympic Forest Park in one year. The pathogen diversity revealed by 16S rRNA gene targeted next-generation sequencing (NGS) showed that 16 of 40 genera and 13 of 76 reference species were present. The most abundant species were *Acinetobacter johnsonii*, *Mycobacterium avium* and *Aeromonas* spp. Quantitative polymerase chain reaction (qPCR) of *Escherichia coli* (*uidA*), *Aeromonas* (*aerA*), *M. avium* (16S rRNA), *Pseudomonas aeruginosa* (*oaa*) and *Salmonella* (*invA*) showed that the *aerA* genes were the most abundant, occurring in all samples with concentrations of 10^{4-6} genome copies/100 mL, followed by *oaa*, *invA* and *M. avium*. In total, 34.8% of the samples harbored all genes, indicating the prevalence of these pathogens in this recreational waterbody. Based on the qPCR results, a quantitative microbial risk assessment (QMRA) showed that the annual infection risks of *Salmonella*, *M. avium* and *P. aeruginosa* in five activities were mostly greater than the U.S. EPA risk limit for recreational contacts, and children playing with water may be exposed to the greatest infection risk. Our findings provide a comprehensive understanding of bacterial pathogen diversity and pathogen abundance in urban recreational water by applying both NGS and qPCR.

Introduction

Water has been reported as a probable reservoir for pathogens and a route of transmission (Aw and Rose, 2012; Brookes et al., 2004; Falkinham, 2015; Gonzales-Siles and Sjoling, 2016; Leclerc et al., 2002). Water bodies in high-density urban areas, such as lakes, rivers and artificial fountains, play important roles in providing recreational opportunities for citizens. Humans may be frequently exposed to water during recreational activities, and public health might be threatened by microbial pathogens found in the water (Hlavsa et al., 2015). Therefore, the distribution of pathogens in urban water is an appealing global research target (Ahmed et al., 2014; de Man et al., 2014; Ekklesia et al., 2015; Liu et al., 2009; Xiao et al., 2013). However, these studies mainly focused on fecal indicator bacteria (FIB), enteric viruses, or several selected bacterial pathogens, such as pathogenic *Escherichia coli*, *Salmonella* spp., *etc.*

* *Corresponding author*. E-mail: wanghui@tsinghua.edu.cn (Hui Wang).

http://dx.doi.org/10.1016/j.jes.2016.11.008
1001-0742/

(Allmann et al., 2013; He et al., 2012; Hong et al., 2010; Li et al., 2014; Oster et al., 2014; Sales-Ortells and Medema, 2014; Ye et al., 2012).

To reasonably manage the microbial quality of urban water, efficient and comprehensive pathogen monitoring is crucial for the discovery of the dominant pathogens. However, there are hundreds of waterborne or living-in-water bacterial pathogens that can be transmitted through water ingestion, direct contact or aerosol inhalation (Woolhouse and Gowtage-Sequeria, 2005). Thus, it is challenging to conduct comprehensive pathogen monitoring in water by means of frequently used methods, such as culture-dependent methods and quantitative polymerase chain reaction (qPCR), because of their low throughput (Aw and Rose, 2012). Due to its better specificity, sensitivity and the reduced time requirements compared with culture-dependent methods or pathogen agent detections, qPCR has been widely and routinely used to directly detect pathogens in research and clinical diagnosis (Ahmed et al., 2009, 2014; Aw and Rose, 2012; Liu et al., 2009); however, qPCR only identifies targeted pathogens. Since those targets are often somewhat blindly chosen by the surveyors, some key pathogens might be ignored, especially in a complex environment. Recently, a high-throughput and open method, 16S ribosomal ribonucleic acid (rRNA) gene targeted next-generation sequencing (NGS), has been applied to explore bacterial pathogen diversity in wastewater (Kumaraswamy et al., 2014; Lu et al., 2015; Ye and Zhang, 2011), biosolids (Bibby et al., 2010; Luo and Angelidaki, 2014), drinking water (Vierheilig et al., 2015) and a watershed in the middle Santa Ana River in the USA (Ibekwe et al., 2013). For example, in a study on pathogen monitoring of wastewater samples (Kumaraswamy et al., 2014), 21 unique sequences affiliated to 16 pathogen species were screened out in 4225 reads from six samples after alignment to 259 annotated sequences of pathogenic bacteria, and it was found that *Salmonella enterica* and *E. coli* were dominant before treatment while *Bacillus anthracis* became more abundant after disinfection. Even though this method can detect hundreds of pathogens simultaneously, the results only present a semiquantitative profile of potential pathogens. In practical monitoring, the main pathogens need to be accurately quantified to estimate the risk levels. Therefore, utilizing the complementary advantages of both methods to investigate the diversity of pathogens and quantifying targets of concern simultaneously is a promising means to conduct comprehensive and accurate monitoring of pathogens in complex environmental samples.

In this study, 46 recreational water samples from a popular urban park in Beijing with multiple entertainment functions, Olympic Forest Park, were collected over one year. The 16S rRNA gene targeted NGS was applied to analyze the diversity of human pathogenic bacteria. In addition, four pathogen species of concern, including *Salmonella*, *Mycobacterium avium*, *Pseudomonas aeruginosa* and *Aeromonas*, and one FIB, *E. coli*, were quantified by qPCR. Furthermore, the infection risks of the first three targets were evaluated. To the best of our knowledge, this is the first study to apply amplicon NGS and qPCR in tandem to determine pathogen diversity and abundances in urban water.

1. Materials and methods

1.1. Study sites and sampling

Beijing Olympic Forest Park is located in the north of an urban part of Beijing and covers an area of 6.8 km^2 (Fig. 1). The park is free for all citizens and is very popular, with an estimated annual visitor flow of about one hundred million (no exact data could be acquired). We studied the water areas in the south part of the park, including the main lake "Aohai Sea" and a series of wetlands to the northwest of the lake (Fig. 1). The lake covers an area of 15 ha with an average depth of approximately 3 m, and contains approximately 300,000 m^3 of water. It has artificial fountains near the middle of the south boundary and many ornamental fishes. It also provides a service for boating. Every spring and early summer, reclaimed water from the Qinghe and Beixiaohe Reclaimed Water Plants is supplemented through the special distribution system connected to the southwest corner of the lake. Both plants apply membrane filtration, ozonation and chlorine/ultraviolet disinfection as advanced treatment processes. Flow direction of water in the lake is from the west to the east and the exit is at the northeast corner. The wetlands were designed to be a landscape and purify the circulating lake water by combining subsurface and surface flow constructed wetlands successively, with water capacities of 60,000 and 15,000 m^3, respectively. In the wetlands, reeds and calami are planted in abundance. Lake water is transported to the wetlands through pipes from the exit of the lake.

The sampling was conducted monthly on sunny days from spring to fall in 2013 (March to October). A total of 46 water samples were collected, including two samples (RW) at the release point of reclaimed water in April and May where the distribution pipe is submerged in lake water; 28 samples at four locations in the lake, lake west (LW), lake center (LC), lake east (LE) and lake outlet (LO), and 16 samples at two sites in the wetlands, outlets of subsurface wetland (W1) and surface wetland (W2). Details are shown in Fig. 1 and Table 1. Supplemental addition of reclaimed water to the main lake lasted continuously from 12 March to17 June in 2013, with a flow of approximately 4800 m^3/day. The theoretical flow rate of recycled water to the wetlands was 11,000 m^3/day. At each site, 5 L. water was collected at 50 cm below the water surface in sterile containers and transported on ice to the lab within 4 hr. For deoxyribonucleic acid (DNA) analysis, 1.2 to 3.9 L of each water sample was immediately filtered on three pieces of sterile 0.22 μm mixed cellulose membranes (GSWP29325, Merck Millipore, USA), and the filters were stored in sterile 50 mL centrifuge tubes at −20°C until DNA analysis.

1.2. DNA extraction, sequencing and data processing

In total, the following fifteen samples were chosen for sequencing analysis (bold in Table 1): eight at the lake center of each month, five from the remaining sites in lake in May, representing temporal and spacial variations in the park water, respectively, and two at the entrance of the reclaimed water in April and May. Genomic DNA from the fifteen water samples was extracted from the 0.22 μm filters using