Molecular Analysis for Screening Human Bacterial Pathogens in Municipal Wastewater Treatment and Reuse

Rajkumari Kumaразwamy,† Yamrot M. Amha,† Muhammad Z. Anwar,‡ Andreas Henschel,‡ Jorge Rodríguez,† and Farrukh Ahmad*†

†Institute Center for Water and Environment (iWATER), Masdar Institute of Science and Technology, P.O. Box 54224, Abu Dhabi, UAE
‡Institute Center for Smart and Sustainable Systems (iSmart), Masdar Institute of Science and Technology, P.O. Box 54224, Abu Dhabi, UAE

ABSTRACT: Effective and sensitive monitoring of human pathogenic bacteria in municipal wastewater treatment is important not only for managing public health risk related to treated wastewater reuse, but also for ensuring proper functioning of the treatment plant. In this study, three different 16S rRNA gene molecular analysis methodologies were employed to screen bacterial pathogens in samples collected at three different stages of an activated sludge plant. Overall bacterial diversity was analyzed using next generation sequencing (NGS) on the Illumina MiSeq platform, as well as PCR-DGGE followed by band sequencing. In addition, a microdiversity analysis was conducted using PCR-DGGE, targeting Escherichia coli. Bioinformatics analysis was performed using QIIME protocol by clustering sequences against the Human Pathogenic Bacteria Database. NGS data were also clustered against the Greengenes database for a genera-level diversity analysis. NGS proved to be the most effective approach screening the sequences of 21 potential human bacterial pathogens, while the E. coli microdiversity analysis yielded one (O157:H7 str. EDL933) out of the two E. coli strains picked up by NGS. Overall diversity using PCR-DGGE did not yield any pathogenic sequence matches even though a number of sequences matched the NGS results. Overall, sequences of Gram-negative pathogens decreased in relative abundance along the treatment train while those of Gram-positive pathogens increased.

INTRODUCTION

The demand for water around the globe for agriculture and other forms of irrigation is exponentially growing as a result of population growth and urbanization.1,2 The use of treated wastewater to augment freshwater supply is often projected as a solution to overcome water scarcity.3,4 Consequently, various countries around the world are adopting treated wastewater reuse practices for irrigation, with the highest volumes reported in Mexico and Egypt, with approximately 4 million and 2 million m³/day, respectively.5 However, treated municipal wastewater carries the risk of contamination from pathogenic microorganisms owing to incomplete disinfection of recalcitrant organisms. Therefore, treated wastewater requires vigilant monitoring for the presence of pathogens prior to reuse.

Developing reliable procedures to screen pathogens is crucial for setting guidelines on restricted and unrestricted treated wastewater reuse.6,7 Traditional culture-based microbiological methods, such as Most Probable Number (MPN) and plating have been widely used for wastewater quality assessment.8,9 However, monitoring pathogens using conventional methods based on indicator organisms has been demonstrated to be inadequate at detecting the presence and levels of pathogens, especially in warmer climates.10,11 Nucleic acid–based molecular methods have overcome the shortcomings of culture-based methods, by enabling higher sensitivity of pathogen detection, minimizing long-term costs, and by reducing analysis time.12–15 Molecular methods, such as microarrays (a hybridization assay method) and quantitative polymerase chain reaction (qPCR) have been used as targeted approaches in the detection of specific pathogens in wastewater samples enabling a semi-quantitative and a direct-quantitative analysis, respectively.16–18 However, these molecular approaches are pathogen-specific and therefore fail to detect the presence of nontargeted pathogens, and overlook overall diversity shifts in microbial community composition.19

The use of 16S rRNA gene analysis to study bacterial diversity and taxonomy has been by far the most common approach in environmental microbiology, and is becoming popular in clinical and food microbiology owing to the availability of standardized techniques and comprehensive databases.20,21 Additionally, it is useful as an exploratory tool for studying poorly characterized mixed-culture samples. In this study, two different 16S rRNA gene analysis methodologies,
namely, (i) PCR-denaturing gradient gel electrophoresis (PCR-DGGE) and (ii) Illumina NGS, were adopted and compared to study bacterial diversity and to screen human pathogens in samples obtained along the process train of the central wastewater treatment plant in Abu Dhabi, UAE. PCR-DGGE with universal 16S rRNA gene primers has been widely used to study the structural bacterial diversity in wastewater treatment. In addition, this method has previously been adopted for rapid detection of pathogens and tick-infecting bacteria. Owing to the low-cost and relatively less intensive data analysis, PCR-DGGE was recommended as a prescreening method for a more efficient use of advanced molecular biology tools such as next generation sequencing (NGS). Illumina sequencing, a high-throughput NGS platform utilizing surface-attached automated dye sequencing, enables the identification and characterization of microbial community members in an ecosystem at a depth of up to millions of sequences per sample. Therefore, Illumina sequencing with universal 16S rRNA gene primers can possibly be used to screen potential pathogens, which would otherwise go undetected using targeted methods. Although frequently applied in medical research, only one study to date has adopted Illumina sequencing for pathogen analysis in wastewater samples. While few researchers have compared the outcomes of PCR-DGGE and NGS, to the best of our knowledge, this type of analysis has not been conducted in treated wastewater for quality assessment purposes.

As a third and targeted approach, a nested PCR based DGGE methodology was developed to analyze the microdiversity of pathogenic and nonpathogenic Escherichia coli in the wastewater treatment train samples. In routine treated wastewater quality monitoring, screening for indicator organisms such as Escherichia coli has been considered an alternative to screening for individual pathogens, in order to enhance sensitivity by circumventing low pathogenic abundance. It is important to note that apart from being considered as an indicator microorganism, some of its strains such as E. coli O157:H7, are pathogens causing serious health effects in human beings. Routinely used culture-based E. coli detection methods are inadequate in differentiating between pathogenic and non-pathogenic strains of E. coli. Thus, some researchers have used real-time nested PCR to quantify pathogenic E. coli O157:H7. Nevertheless, the development of 16S rRNA DGGE as a fingerprinting tool in the current study will be helpful in analyzing microdiversity directly from genomic DNA extracted from water samples, bypassing pure culture isolation. Furthermore, subsequent sequencing of DGGE bands (band length of ~550 bp) can reveal the identity of the strains. This study demonstrates three different approaches (Figure 1) to study the presence and distribution of pathogens in the wastewater treatment plant (WWTP) at three different stages of treatment. The study brings a new perspective to wastewater quality monitoring by demonstrating the efficacy of various molecular approaches for the initial screening and diversity analysis of pathogens at different stages of wastewater treatment.

**MATERIALS AND METHODS**

**Sample Collection.** The wastewater treatment plant under study is the central activated sludge municipal wastewater treatment facility of the city of Abu Dhabi, United Arab Emirates, with a design capacity of 260,000 m$^3$/day. The plant employs chlorination for disinfection. A total of 18 one-liter grab samples were collected from the treatment train over two different time points, December 2012 and April 2013 (mean monthly ambient temperatures of 21 and 27 °C, respectively). In each sampling event, a total of three grab samples were collected from each of the three different stages (i.e., pretreated [PT], prechlorinated [PT], and postdisinfected [PD] samples, (Figure 2)) of municipal wastewater treatment.
Figure 2. Sampling locations at three stages of the activated sludge wastewater treatment plant. In the first sampling grab samples were collected in December 2012 (mean ambient temperature 21 °C). The second set of samples collected in April 2013 (mean ambient temperature 27 °C) were composited from three grab samples collected over a 12 h period at each location; the sample collection periods at different points along the treatment train were time-staggered by the hydraulic retention time between them.
chimeras and trimmed using BioEdit software version 7.0.9 (Ibis Biosciences (http://www.mbio.ncsu.edu/BioEdit/bioedit.html), Carlsbad, CA). The sequences (400–500 bp) determined in this study have been deposited in GenBank (KM403162-KM403198). The DGGE sequences were blasted against the Illumina sequencing results with a threshold similarity of 99% or above. The raw DGGE sequences were clustered with the human pathogenic bacteria 16S rRNA gene database30,38 to identify potential pathogens.

### Illumina Sequencing and Data Analysis

Illumina Sequencing and Data Analysis. The 16S rRNA gene libraries for paired-end Illumina sequencing were constructed with the primers targeting V4 region (U515-F and E786-R) of the 16S rRNA gene using a two-step amplicon approach.39 The libraries were multiplexed with other libraries not used in this study and sequenced using the paired-end approach on a MiSeq Illumina sequencing machine at the BioMicro Center (Massachusetts Institute of Technology [MIT], Cambridge). FASTA/Q clipper and FASTA/Q trimmer components of FASTX-toolkit40 were used for adaptor clipping and base trimming, respectively. FASTX-toolkit is distributed under the Affero GPL (AGPL) version 3. Forward and reverse reads that were long enough were turned into FastQ joined reads using the FLASH tool41 for merging the forward and reverse reads, thus capturing the complete V4 region of the 16S rRNA gene. Quantitative Insights Into Microbial Ecology (QIIME)42 was used for the processing and downstream analyses. The closed-reference operational taxonomic unit (OTU) calling was performed against Greengenes as a reference database.43 OTUs were assigned using QIIME implementation of the UCLUST44 algorithm for clustering, with a threshold of 0.97 (97%) similarity. 243,704 hits were identified when compared with Greengenes database sequences, and this gave rise to 5,085 distinct OTUs (Figure 3).

The unidentified 34% of sequences could be due to chimera, PCR artifacts, sequencing errors or novel sequences not present in Greengenes. Representative OTU sequences were classified taxonomically using the Ribosomal Database Project (RDP)45 classifier against RDP core-set, with a minimum confidence interval of 0.80 (80%). OTU heat maps and sequence counts were generated to summarize the taxonomic communities. In the second approach, the sequences after the initial preprocessing step were clustered against the human pathogenic bacteria 16S rRNA database38 which contains 259 annotated sequences of pathogenic bacteria. The threshold value was set to 0.99 (99%) similarity. For alignment of the OTUs we used the QIIME environment. Multiple sequence alignment was done using PyNast46 and sequences were filtered against RDP core set. The FastTree tool47 was used to generate phylogenetic trees of Gram-negative and Gram-positive pools including identified OTUs, sequence profiles from human pathogenic database, and closely related species from the reference database SILVA (ARB).48 FastTree tool implements neighbor-joining on sequence profiles and uses heuristics to identify candidate joins. Interactive Tree of Life (iTOL),49 visualization server, was used for visualizing the phylogenetic tree and making multivalued bar chart using the data set file constructed that indicates the relative abundance of each identified OTU in all samples.

## RESULTS

### Screening of Microbial Diversity and Pathogens Using PCR-DGGE Analysis of 16S rRNA Genes

Out of six samples studied in duplicate runs, the postdisinfected samples showed the lowest diversity with some of its replicates having maximum of 5–8 bands (SI Figure S1a, lanes labeled PD). The pretreated and prechlorinated samples with higher diversity shared many prominent DNA bands of same melting behavior (SI Figure S1a, lanes labeled PT and PC). Although a significant shift in the population could be observed between the stages and time points, many prominent bands with similar melting behavior were detected between the two time points. However, some bands such as 10B and 9C were only prominent in April 2013 (SI Figure S1a).

---

**Figure 3.** Schematic representation of bioinformatics analysis for Illumina sequencing using Quantitative Insights Into Microbial Ecology (QIIME). A total of six water samples from each stage of the treatment plant at two time-points were sequenced. Raw reads were merged with a 90% overlapping length (up to 200 bp). A total of 365,782 reads were taken for downstream bioinformatic analysis using closed reference OTU calling with two databases.

---

dx.doi.org/10.1021/es502546f | Environ. Sci. Technol. XXXX, XXX, XXX–XXX
Approximately 60 prominent DGGE bands were excised for sequencing, out of which only 28 generated sequences of very high quality, ranging from 400 bp to 500 bp length. The sequences were initially blasted against the raw sequences obtained from Illumina sequencing of the same samples in order to compare the bacteria targeted by these two methods and their universal 16S rRNA primer pairs. Out of 28 sequences, 25 were found to have a 99% and above match with the sequences obtained with Illumina sequencing (SI Table S2). Additionally, 14 of these 25 sequences were classified up to the Genus-level using Greengenes and RDP as reference databases. The DGGE sequence analysis against the human pathogenic database showed that none of the sequences matched with 99% or above similarity. The DGGE profiles of individual grab samples from different WWTP stages collected in the April 2013 event showed same profiling compared to composite samples which were used for final analysis (results not shown).

**Microdiversity Study of Escherichia coli.** The primers 16E1F (forward) and 16E2R + 16E3R (reverse) originally designed for identifying both pathogenic and nonpathogenic E. coli members in water samples were successfully used for DGGE profile analysis in this study by clamping a GC rich region to the forward primer. The DGGE profiling of E. coli 16S rRNA gene products showed maximum of 3−4 bands (SI Figure S1b) indicating a low diversity. Subsequently, the bands were excised and sequenced to determine the identity of the strains, and nine bands generated sequences of very high quality. The sequences were compared against raw sequences generated from Illumina sequencing in the current study and...
were 100% identical to the sequences of pathogenic *E. coli* O157:H7 str EDL933 (SI Table S3). The OTU calling procedure matched sequences of pathogenic *E. coli* O157:H7 str EDL933 also, when compared with the human pathogenic database.

**Screening and Relative Abundance Estimation of Potential Pathogens Using 16S rRNA Next-Generation Sequencing.** Two different approaches were used for analyzing the Illumina sequences (Figure 3). The raw Illumina sequences were compared against the human pathogenic sequences and, in addition, closed-reference OTU calling was conducted against Greengenes general database, 16S rRNA gene sequences matched 21 different pathogenic sequences belonging to various genera and species within the human pathogenic database. The relative abundance of each of these sequences of potential pathogens were calculated for all the three stages of WWTP samples collected in December 2012 and April 2013. Figure 4 displays the phylogenetic affiliations of the pathogens screened while Figure 5 shows their relative abundance in the different water samples. The most dominant sequences of all Gram-negative pathogens screened was *Salmonella enterica*, followed by strains of *E. coli*, *Vibrio cholerae*, and *Yersinia pestis*. Among the Gram-positive human pathogenic bacterial sequences identified, *Mycobacterium tuberculosis*, *Bacillus anthracis*, and *Streptococcus agalactiae* were present at significantly higher relative abundance in water samples when compared to other Gram-positive pathogens. There was a good indication that the relative abundance of sequences affiliated with Gram-negative enteric pathogens, such as *Salmonella enterica* and different strains of *E. coli*, decreased from pretreated to postdisinfected stage. Whereas, the abundance of sequences of Gram-positive pathogens, such as *Mycobacterium tuberculosis* and *Bacillus anthracis*, increased in the postdisinfected samples (Figure 5). However, with some Gram-negative pathogens such as *Yersinia pestis* and *Vibrio cholerae* there was also an increase in relative abundance along the treatment train, predominantly in December 2012.

The closed-reference OTU calling against Greengenes general database showed that 243704 of the sequences clustered to 5085 OTUs (Figure 3), from which 2523 were classified up to the Genus-level (SI Figure S4). The most relatively abundant genera were *Mycobacterium*, *Arcobacter*, and *Clostridium*. The results also showed high abundance of sequences belonging to various phyla such as Alpha-, Beta-, and Gamma-proteobacteria.

**DISCUSSION**

The results from the three 16S rRNA molecular approaches indicated the importance of choosing the right molecular tool for initial screening of human pathogenic bacteria as the approach determines the outcome and the future monitoring strategy. The 16S rRNA gene Illumina NGS with universal bacterial primers (targeting V4 region of the gene) is more effective at screening pathogens than 16S rRNA DGGE (with primers targeting both V3 and V4 regions) combined with the subsequent sequencing of DGGE bands. Around 28 DGGE bands generated 16S rRNA gene sequences of very high quality, however, none of the bands’ sequences were identical to the 16S rRNA gene sequences listed in the human pathogenic database used in this study. The main reason for this discrepancy was that not all bands in the DGGE gel (SI Figure S1a) could be sequenced with the approximate success percentage of only 50%. It is, however, important to note that all of these 28 sequences matched with greater than 99% similarity to the 16S rRNA sequences of nonpathogenic bacteria obtained through Illumina sequencing (SI Table S2). This result gave strong evidence that DGGE universal primers (341F and 907R) and Illumina primers (U 515F and E 786R) targeted the same kind of bacterial DNA in the PCR. The strategy of using DGGE analysis for pathogen detection can be improved dramatically by running a reference lane with DGGE fragments amplified from the pathogenic bacteria of interest along with the PCR products of water samples. This approach will help in excising only the DGGE bands with similar melting profiles to that of the bands of pathogenic bacteria rather than randomly excising dominant bands. It has been demonstrated in detection of fish pathogens, where universal primers were used with PCR-DGGE with a reference lane of known pathogens for a more focused detection.33 Another potential problem with sequencing of excised DGGE bands preamplified with universal primers is that DGGE bands from pathogens in the samples may not necessarily correlate with bands corresponding to pathogenic bacteria.

*E. coli* detection and enumeration in WWTPs is normally performed using culture based methods such as MPN analysis and plating.8 Recently, many researchers have developed molecular methods such as PCR assays targeting toxic genes specific to *E. coli* such as uidA gene, rfbE gene, shiga toxin genes for detection of *E. coli* in water samples.60 Few studies have assessed diversity of different strains of *E. coli* through methodologies such as rep-PCR and pulsed field electrophoresis.51 In the current study, DGGE microdiversity analysis using 16S rRNA gene primers targeting both pathogenic and nonpathogenic *E. coli* in water samples demonstrated reliable and reproducible results. Apart from showing reproducible band patterns, this methodology was also useful in obtaining sequences of approximately 500 bp matching pathogenic *E. coli* O157:H7 str EDL933, which was also picked up by Illumina NGS (SI Table S3). The results proved that this methodology can be efficiently used for studying microdiversity of *E. coli* strain in different kinds of wastewater samples. Furthermore, it is important to note that the sensitivity of the *E. coli* PCR was significantly improved using a nested PCR approach and therefore it can be used to target *E. coli* strains of very low abundance in the samples analyzed. Methodologies such as rep-PCR and pulsed field electrophoresis have already been developed for microdiversity studies of *E. coli* and were reported to be more discriminatory between the closely related species than the 16S rRNA based molecular methods.53 However, methods such as rep-PCR need to have pure cultures for the PCR step. Therefore, a precultivation step and isolation of pathogens is required in these techniques, unlike DGGE where genomic DNA extracted from water samples can be directly used for the PCR. In addition, DGGE band sequencing gives phylogenetic affiliation of the bacteria present in the samples analyzed. It is also important to note that in addition to *E. coli* O157:H7, Illumina sequencing identified an additional *E. coli* pathogen (*E. coli* CFT073) which was not identified by DGGE analysis using *E. coli* specific primers, possibly due to its very low abundance in the water samples.

The main result of this study came via 16S rRNA gene Illumina sequencing of the water samples. NGS methods such as Illumina have several advantages over traditional molecular methods such as 16S rRNA DGGE and clone library analysis. The broader coverage in NGS enables one to identify the microorganisms present in low numbers, which would...
otherwise go missing with traditional molecular tools such as PCR-DGGE that give information only on dominant and preferentially amplified sequences in the samples. There are only a few studies that have compared DGGE with next generation sequencing,31,32 and, to the best of our knowledge, no such analysis has been conducted in studies to assess the quality and safety of treated wastewater.

The raw Illumina sequences were initially compared with the Greengenes database using QIIME clustering algorithm (SI Figure S4). Later, the sequences were compared with the human pathogenic database developed by Cai and Zhang (2013).30 As shown in Sontakke et al. (2009)34 broad range 16S rRNA amplification has proven to be a useful diagnostic tool in clinical microbiology, especially for well-documented pathogens. The method is limited due to unknown pathogenicity of closely related strains. However, in dealing with these false positives, the findings of our approach are instructive for quantification methods such as qPCR, where suitable marker genes can be selected due to the tight linking of the human pathogenic bacteria database deployed here and the Virulence Factor Database. On the other hand, it is likely that 16S rRNA profiles exhibit the lowest false negative rate as compared to shotgun metagenomics and marker gene identification. Finally, the substantially lower cost per sample makes 16S rRNA sequencing more feasible for successive sampling at high temporal resolution than shot-gun metagenomic analysis, thus enabling the monitoring of intruding and recalcitrant strains in WWTPs.

This approach identified 21 different potential pathogens in water samples from different stages of the treatment plant (Figure 4 and SI Figures S2, and S3). Apart from E. coli, intestinal Enterococcus, and Salmonella spp, none of these other pathogens are routinely screened at the WWTP facility. Also, several of the pathogens were not included in the list of common pathogens associated with wastewater reuse in the WHO guidelines.6 The current results clearly indicate that the information obtained through next generation sequencing tools targeting 16S rRNA genes is more inclusive and has the advantage over pathogen targeted assays such as qPCR and microarray, which are more sensitive and specific but also introduce the risk of missing nontargeted pathogens. The results also showed a specific trend in the relative abundance of the pathogenic sequences. The sequences of Gram-positive pathogens such as Mycobacterium tuberculosis and Bacillus anthracis were more abundantly present in the postdisinfected water samples (Figure S). The same trend was observed for other Gram-positive pathogens’ sequences belonging to the genus Clostridium spp, Staphylococcus spp, Streptococcus spp. (Figure S). Conversely, the relative abundance of sequences of Gram-negative pathogens steadily decreased from pretreated to postdisinfected with the exception of Yersinia pestis and Vibrio cholerae. A similar trend of increase in the abundance of Mycobacterium tuberculosis has been previously reported in municipal WWTP facilities.30 The Gram-positive bacteria are reported to have protective cell wall structure and spore forming ability, which allow them to better withstand disinfection processes such as chlorination that was used in this wastewater treatment plant.35,36 This could be one of the primary reasons for high abundance of the sequences of Gram-positive pathogens in the postdisinfected water samples of the current study, which needs further detailed investigation. There are also studies which indicated that pathogens can effectively multiply and increase in number in the different stages of wastewater treatment plants such as activated sludge and post-treated storage tanks.57

It is important to note that the relative abundance estimation in Illumina sequencing cannot be the absolute quantification considering all the DNA extraction, PCR, sequencing biases, as well as unequal read depths and variable gene copy numbers among organisms in the samples. Moreover, bioinformatics tools such as QIIME can be used only to calculate relative abundance of the pathogen of interest in the water samples analyzed. Previous researchers58 have indicated the non-reliability of relative abundance by performing quantitative PCR using primers targeting Mycobacterium spp. Other researchers60 have recommended the use of qPCR to confirm and quantify the relative abundance data obtained from next generation sequencing. Therefore, the current screening results, though alarming for the reuse of postdisinfected water, should be further confirmed using qPCR or quantitative reverse transcription-PCR (qRT-PCR) targeting live cells.

Although 16S rRNA gene NGS provides valuable indication about the presence of pathogens, even the sequences of >99% similarity to the reference sequences do not conclusively prove their presence. Researchers in clinical microbiology have recommended to use this 16S rRNA approach with the caution of using an approximately 500 bp read length and >99% similarity to the reference pathogenic sequences. They have highlighted that some genera need further validation in addition to 16S rRNA sequences for species classification.20 While the presence of pathogens such as Salmonella enterica and Escherichia coli (with the exception to E.coli 0157: H7, where virulent factors are normally targeted)59–61 have been confidently detected using 16S rRNA methods by previous studies, pathogens such as Vibrio cholera, Mycobacterium tuberculosis, Bacillus anthracis, Streptococcus spp, and others need a second line of confirmation targeting their virulence factor genes or other functional genes that can be used as markers.20 This is due to the high similarity of 16S rRNA sequences of species within the genera, which makes the approach less discriminatory. However, a detailed phylogenetic analysis was performed for all the 21 pathogens screened using closely related pathogenic sequences from human pathogenic database and ARB database, which indicate the sequences identified could be potential pathogens (SI Figures S2 and S3). It is also important to note that the length of 16S rRNA fragment analyzed in this study is around ∼250 bp, relatively longer than the fragment length covered in other NGS studies,60 and with a higher degree (~90%) degree of overlap between forward and reverse reads. Therefore, the confidence level here is much higher and could be used as an initial evidence. A detailed literature study on advantages and disadvantages of using 16S rRNA marker gene for the pathogens screened in our current study is provided in the SI (Section II).

This study reconfirmed the proposition made by other researchers that molecular methodologies generate more information on the presence of pathogens than the culture-based methods.15,62 Salmonella spp, which were not identified using routine traditional selective enrichment plating at the wastewater treatment plant, were detected in significand abundance with the Illumina sequencing (Figure S). The E. coli numbers (through MPN analysis) obtained did not reveal the pathogenicity of the bacteria enumerated (SI Table S1), but the PCR-DGGE and Illumina sequencing screened pathogenic strain E. coli. This was also the case with screening of
pathogenic *Enterococcus faecalis* by Illumina sequencing (Figure 5).

Finally, water reclamation and reuse is a growing practice around the world. Generally, recycled water is carefully matched based on its quality to a specific nonpotable restricted end-use. More recently, water scarcity has prompted several regions in the developed world to consider treated municipal wastewater as a source of direct or indirect augmentation of drinking water supplies. Regions such as Singapore and Southern California have implemented such reuse, albeit with further tertiary treatment such as reverse osmosis and advanced oxidation, thereby limiting the risk of human exposure to pathogens. However, much of the developing world still uses treated municipal wastewater directly for landscape irrigation, and in some cases the irrigation of edible crops, leaving the question of exposure risk from residual pathogens in treated effluent an open one.

In conclusion, between the two different methods adopted, 16S rRNA gene Illumina sequencing clearly proved to be a more sensitive methodology compared to 16S rRNA DGGE for initial screening of pathogens. DGGE microdiversity analysis of 16S rRNA fragments of *E. coli* showed reliable and reproducible results, indicating this procedure can be applied in future *E. coli* microdiversity studies of water samples from various sources. The higher abundance of sequences of Gram-positive pathogens such as *Mycobacterium tuberculosis* and *Bacillus anthracis*, as well as Gram-negative pathogens *Yersinia pestis* and *Vibrio cholerae* present in the postdisinfected water samples revealed a significant health risk that needs further evidence such as the presence of toxic (or virulent) genes specific to the pathogens screened, and metagenomics or functional genomics approaches based on next generation sequencing.

**ASSOCIATED CONTENT**

* Supporting Information

NGS datasets of this study were deposited in the SRA server (http://www.ncbi.nlm.nih.gov/sra) and accession numbers are listed in SI Table S4. Additional information as noted in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

*Corresponding Author*
Phone: +971 2 810 9114; fax: +971 2 810 9901; e-mail: fahmad@masdar.ac.ae.

*Notes*  
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This research was funded under research grants 11WAAA1 and 13XAAA2 from the Masdar Institute of Science and Technology.

**REFERENCES**


(22) Boon, N.; Windt, W.; Verstrael, W.; Top, E. M. Evaluation of nested PCR–DGGE (denaturing gradient gel electrophoresis) with


