Using high-throughput sequencing to assess the impacts of treated and untreated wastewater discharge on prokaryotic communities in an urban river

Yaohui Bai · Weixiao Qi · Jinsong Liang · Jiuhui Qu

Received: 21 June 2013 / Revised: 8 July 2013 / Accepted: 9 July 2013 / Published online: 3 August 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract In many megacities wastewater is an important source of surface water, particularly during drought periods. While changes in surface water chemistry associated with effluent inflow have generally been well-studied, few data have been collected on the effects to prokaryotic communities. The objective of this study was to explore the impacts of treated and untreated wastewater discharges on prokaryotic community in an urban river. High-throughput sequencing was conducted for analyzing the prokaryotic community composition and function in river water, treated wastewater and untreated wastewater. Results revealed that the prokaryotic community compositions in the upstream river reach were dominated by treated wastewater discharge. In the middle- and downstream river reaches, untreated effluent volumes are higher, thus affecting the structure of the prokaryotic community, promoting a rise in *Cyanobacteria* and *Thaumarchaeota*. Function annotation revealed a number of genes associated with xenobiotic metabolism and human diseases were observed in river and wastewater samples, suggesting wastewater discharge to river may pose a risk to human health. Quantitative real-time PCR results revealed that the treated and untreated wastewater discharges also affected the abundance of ammonia oxidation bacteria (AOB) and ammonia oxidation archaea (AOA) in river.

Keywords Effluent-dominated river · Wastewater discharge · Prokaryotic community · High-throughput sequencing · Community composition and function

Electronic supplementary material The online version of this article (doi:10.1007/s00253-013-5116-2) contains supplementary material, which is available to authorized users.

Y. Bai · W. Qi · J. Liang · J. Qu (✉)

State Key Laboratory of Environmental Aquatic Chemistry,
Research Center for Eco-Environmental Sciences, Chinese Academy
of Sciences, Beijing 100085, People's Republic of China
e-mail: jhqu@rcees.ac.cn

Introduction

The rapid development of megacities, especially in Africa and Asia, has led to a deterioration in the quality of many surface waters (Heeb et al. 2012; Suthar et al. 2010). Wastewater discharge is the primary cause of this deterioration, which can be particularly acute during drought periods. Most of the previous studies for the effects of wastewater discharge on surface water quality have focused on conventional and trace chemical agents, such as organic matter, nitrogen, phosphate, and heavy metals (Li et al. 2008; Ort et al. 2009; Pernet-Coudrier et al. 2012; Suthar et al. 2010; Tsuzuki et al. 2010; Zhang et al. 2011), while few data have been collected on the effects to prokaryotic communities. This lack of detailed microbiological information represents a major gap in the overall understanding of the impacts of wastewater on aquatic ecosystems since prokaryotic communities play a fundamental role in the biodegradation and biotransformation of chemical compounds. The response of various taxonomic assemblages to the introduction of chemical stressors may influence the resulting breakdown and transformation of those compounds (Kreves et al. 2013). Collecting detailed information on the prokaryotic community composition and function, and fluctuations occurring in the community is therefore crucial to a better understanding of the effects of wastewater discharge.

The Beijing–Tianjin region of China has one of the highest population growth rates in the world due to the rapid urbanization. The resulting pressure on natural resources has both increased water pollution and decreased available water supplies. Because many regional rivers are wastewater-dominated during drought periods, this region was selected for our study as representative of the effects of wastewater on surface water supplies.

The use of high-throughput sequencing has become an increasingly popular technique for the analysis of prokaryotic communities in natural or artificial environments (Lin et al.

2012; Thomas et al. 2012; Yu and Zhang 2012; Zhang et al. 2010). In this study we integrated physicochemical and molecular approaches (based on high-throughput sequencing) to explore the effect of the wastewater discharges into the river ecosystem, in particular, prokaryotic community composition and function.

Materials and methods

Collection of water samples and physiochemical analysis

The Beijing-Tianjin region of Haihe River System originates from an upstream reservoir (Fig. 1). It is the major river system in the Beijing-Tianjin area that receives wastewater. Of the total upstream inflow, 93 % is treated water from Beijing's urban area and 4 % is untreated domestic wastewater (Heeb et al. 2012). The river system runs for 240 km to Bohai Bay. The largest fraction of treated wastewater entering the river comes from the Beijing wastewater treatment plant through upstream tributaries (between sites 1 and 7) as well as numerous untreated sewage inputs that exist along the entire length of the river. River midstream reach (between sites 8 and 11) is an irrigation area. Over 90 % of the upstream river water is used for crop irrigation in this area (Pernet-Coudrier et al. 2012). River downstream reach (after site 12) mainly received the untreated wastewater from Tianjin. A description of the hydrology of the sampling area was presented in a previous study (Pernet-Coudrier et al. 2012).

Surface water samples were collected from the river (14 samples), untreated wastewater (3 samples, between sites 1 and 7), and treated wastewater (5 samples, between sites 1 and 7) in November 2011. No rainfall or snowfall events occurred within a month prior to sample collection. Temperature, redox potential, pH, and dissolved oxygen (DO) were measured in the field. Collected water samples were filtered in the field through cellulose nitrate membranes of 0.45 μm pore size. The filtrate conserved with 100 μl CHCl_3 for photometric analyses of $\text{NH}_3\text{-N}$, $\text{NO}_3\text{-N}$, and total dissolved phosphorus (TDP) (State Environmental Protection Administration of China 1989). For dissolved organic carbon (DOC) determination, water samples were filtered with cellulose glass membranes and then analyzed with a TOC-V_{CPH} total organic carbon.

DNA extraction

For DNA extraction, 30–400 ml (depending on turbidity) of each water sample were filtered (Millipore, 0.2 μm pore size; 47 mm diameter) using a vacuum pump. Total DNA was extracted from each membrane using a PowerWater DNA isolation kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA). DNA extraction was performed in triplicate for each water sample.

16S rRNA pyrosequencing analysis for community composition

Pyrosequencing for 16S rRNA genes was used to elucidate the prokaryotic community compositions in river, untreated and treated wastewater. Bacterial and archaeal 16S rRNA genes

Fig. 1 Sampling sites from the Beijing-Tianjin region of Haihe River System (based on Qi et al. (2013)). Numbers 1–14 on the map denote the locations of the river sampling sites; W1–W3 denote the locations of sampling sites from three untreated wastewater; TW1–TW5 denote the locations of sampling sites from five domestic wastewater treatment plants. W1–W3 and TW1–TW5 sites are distributed between river sites 1 and 7. The longitudes and latitudes of sampling sites are given in Table 1. WWTP represents wastewater treatment plant

were amplified from triplicate water DNA samples by PCRs using barcoded primers modified with 5'Adaptor-A and 5'Adaptor-B sequences (454 Life Sciences), respectively. PCR primer pairs, thermal program, and barcode primers were described in supplemental Table S1 and a previous study (Bai et al. 2012). The PCR products were then purified with AxyPrep DNA Gel Extraction Kit (Axygen, USA). The DNA concentration of the purified PCR products was measured by TBS-380 Fluorometer (Turner Biosystems, CA, USA). Prior to sequencing, the amplicons from triplicate water samples were pooled in equal amounts based on concentration and subjected to emulsion PCR, and amplicon libraries were generated as recommended by 454 Life Sciences. Sequencing was performed using a Roche Genome Sequencer GS-FLX. Pyrosequencing reads were analyzed using QIIME standard pipeline (Caporaso et al. 2010b). Briefly, sequences reads were initially filtered for removing low quality reads and barcode primers. Then the treated sequences were checked with OTUPIPE (Haas et al. 2011) and the putative chimeric sequences were excluded for further analysis. The remaining set of high quality 16S rRNA sequences were clustered into OTUs using UCLUST (Edgar 2010), with 97 % sequence identity threshold. Representative sequences from each OTU were aligned using PyNAST (Caporaso et al. 2010a) and the Greengenes (DeSantis et al. 2006) database. Taxonomy was assigned using the Ribosomal Database Project (RDP) classifier (minimum confidence of 80 %) (Wang et al. 2007). Alpha diversity statistics including library coverage (Good 1953), nonparametric richness Chao 1 (Chao 1987), and Shannon index, were calculated for each sample. The weighted UniFrac distance metrics (Lozupone et al. 2011) were used to compare community diversity among five samples.

All original 454 sequences are archived at NCBI Sequence Read Archive (SRA) under accession SRX275913.

Metagenomic analysis for community function

Based on the water chemistry and prokaryotic community analyses, we merged extracted DNA from multiple sample sets into composite samples. Three untreated wastewater DNA samples (W1-W3) were combined at same amount, and five treated wastewater DNA samples (TW1-TW5) were combined, resulting in two composite samples, labeled W and TW, respectively. River samples from sites 1-7 (upstream), 8-11 (midstream), and 12-14 (downstream) were combined as R1, R2, and R3, respectively. The merged DNA samples (W, TW, R1, R2 and R3) were used for metagenomic analysis.

Five paired-end libraries (350 bp insert size) were constructed and sequenced using the Illumina Hiseq 2000 (Illumina, Inc., San Diego, CA, USA) sequencing system (90 bp pair-end reads). Sequence reads were initially filtered to remove those that were of low quality or ambiguous. The

filtered reads (~2Gbp for each dataset) were directly uploaded to the MG-RAST (Meta Genome Rapid Annotation using Subsystem Technology, v3.2.2) website (<http://metagenomics.anl.gov/>) for function annotation using MG-RAST pipeline (Meyer et al. 2008). We used SEED subsystem and KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation to assign reads to different functional groups (SEED) or metabolic pathways (KEGG) in MG-RAST with recommended parameters (Gregoracci et al. 2012) (minimum alignment length, 15 amino acids; E-value cutoffs, $1e^{-5}$). Xenobiotics and nitrogen biodegradation or biotransformation genes present in five samples were identified and compared using the KEGG database. The abundance of the gene sequences was quantified by calculating the number in 100,000 metagenomic sequences.

Five metagenomic datasets are archived at NCBI Sequence Read Archive (SRA) under accession SRX275933.

Real-time PCR for community quantification

Real-time PCR was used to obtain the quantitative distribution of bacteria, archaea, ammonia oxidation bacteria (AOB) and ammonia oxidation archaea (AOA) in the river and wastewaters. Bacterial and archaeal 16S rRNA or *amoA* were quantified by an ABI 7300 fast real-time PCR system based on the SYBR Green I method. The primer pairs and the thermal programs of the PCR amplification are described in the Supplemental Table S1. Plasmid standards and calibration curves construction were described in detail previously (Bai et al. 2012).

Results

Physiochemical properties of water samples

Characteristics of the water samples collected from each sampling site are shown in Table 1. pH ranged from 6.5 to 8.3 in all water samples. Dissolved oxygen (DO) in river samples fluctuated from 2.2 to 8.6 mg/l. Conductivity ranged from 1046 to 3420 $\mu\text{S}/\text{cm}$ and dissolved organic carbon (DOC) was measured at between 9.97 and 26.1 mg/l. Both conductivity and DOC increased along the river flow direction. Ammonia ($\text{NH}_3\text{-N}$) and total dissolved phosphate (TDP) were measured in relatively high concentrations (13.4–32.1 mg/l and 1.54–6.73 mg/l, respectively). Elevated concentrations of both these compounds can lead to eutrophication in surface waters.

Although untreated wastewater displayed the highest concentrations of DOC (20.39–46.07 mg/l), $\text{NH}_3\text{-N}$ (24.6–44.6 mg/l) and TDP (2.94–5.37), treated effluents were responsible for the greatest loading to the river because of the sheer volume entering the river system.

Table 1 Water properties of sampling sites

Sample	Location	Temperature (°C)	pH	Conductivity (μS/cm)	DOC (mg/l)	NH ₃ -N (mg/l)	NO ₃ ⁻ -N (mg/l)	TDP (mg/l)
1	40°07'53"N 116°19'58"E	3.0	6.65	1095	12.68	22.4	0.1	2.76
2	40°09'05"N 116°24'15"E	4.2	6.77	1104	12.39	24.9	0.1	2.76
3	40°05'36"N 116°29'09"E	6.7	6.75	1046	12.07	32.1	0.1	2.76
4	40°03'22"N 116°32'21"E	6.7	6.62	1056	9.97	16.8	0.2	1.54
5	39°55'22"N 116°39'23"E	6.0	6.66	1081	10.10	27.0	0.3	1.64
6	39°50'37"N 116°46'03"E	5.7	6.75	1092	12.51	25.8	0.8	2.10
7	39°46'05"N 116°54'16"E	4.3	6.80	1195	14.21	27.0	0.5	2.35
8	39°44'50"N 116°56'08"E	3.2	6.77	1197	13.60	30.9	0.3	6.33
9	39°38'16"N 116°58'10"E	2.0	7.52	1192	14.19	18.7	0.3	4.90
10	39°29'08"N 117°01'37"E	2.2	6.76	1238	13.76	18.5	0.4	2.46
11	39°23'46"N 117°04'33"E	4.6	6.88	1508	14.38	25.3	0.2	2.46
12	39°14'42"N 117°05'22"E	1.6	8.33	2560	13.97	18.4	0.1	2.13
13	39°16'29"N 117°22'44"E	1.6	6.49	3250	26.10	14.8	0.3	2.97
14	39°11'06"N 117°32'01"E	1.3	6.94	3420	24.98	13.4	0.5	6.73
TW1	39°54'18"N 116°30'59"E	12.6	6.71	1037	8.831	5.6	1.1	3.88
TW2	39°57'50"N 116°30'32"E	12.2	6.88	1010	9.604	6.3	0.8	3.12
TW3	40°01'40"N 116°21'38"E	11.3	6.60	1017	10.81	8.5	0.8	0.83
TW4	39°49'35"N 116°26'11"E	13.0	6.97	1339	10.35	7.7	0.8	1.59
TW5	40°0'18"N 116°29'52"E	8.4	6.92	1073	9.88	8.4	1.1	2.22
W1	40°09'03"N 116°24'15"E	7.5	6.90	1058	46.07	44.6	0.1	5.37
W2	40°03'26"N 116°32'24"E	7.0	6.48	1664	20.39	24.6	0.1	3.33
W3	40°07'14"N 116°28'01"E	7.2	6.83	1021	31.23	25.9	0.1	3.94

Prokaryotic community composition

The 454 pyrosequencing analysis of 16S rRNA gene amplicons from river, treated and untreated wastewater samples produced 216,072 reads, leaving 117,892 reads after quality filtering and removal of chimeric sequences. Species richness, coverage, and diversity estimations were calculated for each data set (Table S2). Good's coverage revealed that these libraries represented the majority of bacterial and archaeal 16S rRNA sequences present in each sample, with values ranging from 93.6 % to 99.0 %. Chao 1 values clearly showed that the bacterial estimated richness in treated wastewater samples was higher than in untreated wastewater and river samples; archaeal richness in the river samples was higher than in wastewater samples. Shannon indices among three types of water samples were not significant different, ranging from 6.41 to 8.23.

We compared the bacterial and archaeal community compositions among the three types of water samples using principal coordinate analysis (PCoA) (Fig. 2). The bacterial community composition in the upstream river segment was similar to the treated wastewater composition but different from the untreated wastewater (Fig. 2a). The similarity in bacterial community composition between upstream

river samples and treated wastewater is likely due to the treated wastewater dominance of the upstream river reach. PCoA analysis of the archaeal community (Fig. 2b) revealed composition in the upstream segment of the river was also comparatively similar to treated wastewater.

Of our filtered sequences, we identified a total of 39 bacterial phyla and 5 archaeal phyla across all water samples with RDP classifier. The dominant groups (greater than 1 % abundance) of each sample are displayed as Fig. 3. *Proteobacteria* (20 %–81 %) was the most abundant phylum in most water samples but was lower in wastewater samples. Among *Proteobacteria*, *Betaproteobacteria* dominated in most of water samples and *Deltaproteobacteria* were only present in treated wastewater. The phylum *Bacteroidetes* was also abundant in all samples, accounting for 5 %–28 % of the occurrences. *Cyanobacteria* were only present in the river samples and the abundance of this group increased in middle- and downstream. For archaea (Fig 3(b)), *Methanomicrobia* (phylum *Euryarchaeota*) was the dominant archaeal class in all samples, accounting for 61 %–92 % of the abundance. Comparing to upstream, *Thaumarchaeota* increased in middle- and downstream of the river.

A total of 355 bacterial and 37 archaeal families were identified in our pyrosequencing datasets. The numbers of

The metabolism genes and pathways associated with xenobiotic biodegradation and nitrogen transformation were investigated using KEGG annotation (Table 2). Biodegradation pathways for 12 types of xenobiotic compounds were identified in five metagenomic datasets. In the river samples, the abundance of xenobiotic biodegradation genes tended to increase along with rising river discharge (Upstream, 35 xenobiotic biodegradation genes per 100,000 sequences; Midstream, 111; Downstream, 116). Among various xenobiotic degradation genes, the genes associated with benzoate biodegradation were the most abundant in all samples. This is not surprising since benzoate is a key intermediary in the microbial metabolism of various aromatic compounds. Naphthalene degradation genes were found to be dominant among the polycyclic

Fig. 2 Principal coordinate analysis (PCoA) of community composition using the weighted UniFrac distance matrix method. (a) Bacterial community (b) Archaeal community

aromatic hydrocarbon (PAH) degradation genes. In addition, the degradation genes associated with a common pesticide (atrazine) or other chemical compounds (toluene, nitrotoluene, etc.) were also present in all samples.

Gene sequences associated with nitrogen metabolism were included in this assessment and are shown in Table 2. A complete set of denitrification genes (*nar*, *nir*, *nor*, *nos*) was found in all samples, suggesting that denitrification might be the primary N removal pathway in the river or wastewaters. However, for nitrification genes, we did not locate any ammonia oxidation (*amo*) gene sequences and hydroxylamine oxidation (*hao*) genes in the datasets.

The occurrence of genes associated with human disease was also assessed, and found to account for approximately 8 % of identified function genes (KEGG annotation) in metagenomic datasets. Comparing their abundances among five datasets (Fig. S1), the abundance in the upstream river reach was lower than either midstream or downstream. The

Fig. 3 Relative abundance of dominant phylogenetic groups at the sampling sites. Phylogenetic groups accounting for <1 % of all classified sequences or unidentified sequences are summarized as “other” in the figure. (a) Bacterial community (b) Archaeal community

abundance of these same genes was lower in treated wastewater than in untreated wastewater.

Prokaryotic community quantification

The abundances of prokaryotic 16S rRNA and *amoA* genes are shown in Fig. 5. The density of the bacterial 16S rRNA gene was 10^8 – 10^{11} copies/l of water, which was 92–8,822 times the measured quantity of archaea. These data indicate the bacteria were dominant in river water and both wastewater types, and were primarily responsible for pollutant biodegradation. AOB abundance was significantly higher than AOA (AOB/AOA ratio of 4 to 96) in the treated wastewater samples, suggesting that AOB may be primarily responsible for ammonia oxidation. In river samples (except site 1), AOB/AOA ratios

decreased from 7.24 to 0.04 in conjunction with the river flow. The AOB/AOA ratios in untreated wastewater samples fluctuated from 0.35 to 1.87.

The bacterial *amoA*/16S rRNA ratio in river and treated wastewater samples ranged from 10^{-4} to 10^{-6} , which were higher than those in the untreated wastewater samples (about 10^{-7}). Archaeal *amoA*/16S rRNA ratios in the three types of water samples ranged from 10^{-1} to 10^{-4} and were not significantly different among samples.

Discussion

In many megacities wastewater plays a crucial role in maintaining the water quality and contributing to river

Fig. 4 Community function in five metagenomic datasets based on SEED annotation in the MG-RAST server (Meyer et al. 2008) using a maximum e-value of $1e-5$, a minimum identity of 60 %, and a minimum alignment length of 15 amino acids. (a) Hierarchical classification of function annotation. Data have been normalized to values between 0 and 1 to allow for comparison of differently sized samples (b) Principal coordinate analysis (PCoA) for community function comparison

discharge, particularly during drought periods. The Beijing–Tianjin region of Haihe river system, was selected for study since it is dominated by both treated and untreated effluent. High-throughput sequencing was used to evaluate the effects of wastewater discharge on the river prokaryotic community.

Prokaryotic community composition (Fig. 2) in the upstream river reach was found to be more similar to the treated wastewater, most likely due to the high proportion of treated effluent discharged into the upstream segments. However, in midstream and downstream river, the continuous input of untreated water

Table 2 Relative abundance of xenobiotic biodegradation or nitrogen biotransformation-related gene sequences retrieved from five metagenomic datasets using KEGG annotation. Values were normalized by randomly sampling 100,000 sequences per sample

Category	Compounds metabolism	R1	R2	R3	TW	W
Xenobiotics biodegradation	Benzoate	8.03	28.92	19.38	33.50	15.94
	Atrazine	7.27	21.00	25.92	15.73	10.50
	Naphthalene	6.28	16.77	19.83	13.65	10.77
	Chlorocyclohexane and chlorobenzene	5.64	22.84	21.08	2.70	16.93
	Bisphenol	3.26	9.86	11.76	1.56	2.30
	Drug	2.27	9.51	11.59	0	7.46
	Polycyclic aromatic hydrocarbon	1.16	0.21	0	0	9.90
	Toluene	1.09	1.88	5.89	0.21	0.83
	Nitrotoluene	0.10	0.28	0.46	0.79	0.08
	Aminobenzoate	0.05	0.04	0	0.17	0.23
Nitrogen biotransformation	DDT	0	0	0	0	0.28
	Nitrogen fixation	0.39	0	0	2.62	12.40
	Nitrate reduction	2.33	2.14	0.33	0.45	13.50
	Nitrite reduction	3.74	11.26	8.57	13.07	8.25
	Nitric-oxide reduction	5.32	30.87	17.82	9.57	4.49
	Nitrous-oxide reduction	1.32	0	0	11.90	6.75

led to the changing of river chemical composition. As a consequence, the composition of the prokaryotic community was also altered. Taxonomically, the most significant middle- and downstream alteration was the rapid increase in abundance of *Cyanobacteria* and *Thaumarchaeota* (Fig. 3). The increase of *Cyanobacteria* suggests eutrophication in the middle- or downstream river segments may be more pronounced than upstream since a proliferation of *Cyanobacteria* is often associated with eutrophication in freshwater, estuarine, and marine ecosystems (Riedinger-Whitmore et al. 2005). The field observation also supported this deduction. For the dominant prokaryotic families in the samples, *Comamonadaceae* belonging to the class *Betaproteobacteria*, is a frequently observed microbial family found in rivers (Crump et al. 2009; Kasalický et al. 2013), and was dominant in most of our river samples (Table S3). *Campylobacteraceae* was another dominant family in all untreated wastewater samples, two treated wastewater samples, and two river samples obtained from a wastewater discharge area. This family includes *Arcobacter* and *Campylobacter* which, along with other genera, may pose a significant public health concern (Moreno et al. 2003). *Campylobacter*, for example, includes some of the most common human enteric pathogens (Frost 2001). Furthermore, a number of viruses were also identified in the collected samples. Their proportion in the river and treated wastewaters was lower (0.01–0.02 % of the identified microorganisms), than was found in the untreated wastewaters (0.09 %). These data indicate that wastewater treatment is significantly reducing certain pathogenic organisms, and lack of treatment, or poor treatment, may be negatively affecting the safety of the river water by allowing high pathogen loads, (particularly in the summer), thus posing a potential risk to

human health and contributing to overall surface water impairment (Mitch et al. 2010).

Metagenomics, based on high-throughput sequencing, is being increasingly used as a tool for analyzing microbial community function (Thomas et al. 2012). From our metagenomic results, a number of xenobiotic biodegradation genes were observed in three types of water samples (Table 2). Recently it has been suggested that there is a link between the abundance of xenobiotic degradation genes and xenobiotic biodegradation rates (Fleming et al. 1993; Park and Crowley 2006; Yergeau et al. 2012). It may be beneficial, therefore, to use degradation genes, particularly those in high abundance, as indicators of the possible presence of xenobiotics and their metabolites. In this study we found that the total amounts of xenobiotic biodegradation genes in midstream and downstream segments of the river were significantly higher than in upstream areas (Table 2). This difference may be associated with the presence of higher volumes of untreated wastewater in the mid- and downstream reached. This proposition is supported by the measured DOC concentration data (Table 1). Furthermore, metagenomic results revealed that several refractory xenobiotic (such as atrazine, bisphenols and halohydrocarbon) degradation genes remained in the treated wastewater and river water at relatively high abundance, suggesting that these xenobiotics were not efficiently removed with conventional treatment. All of these xenobiotic compounds had previously been identified through chemical analysis (Heeb et al. 2012). Atrazine is among the most frequently detected herbicides in surface water, groundwater and wastewater, and is typically measured at very low (ng/l to µg/l) levels (Hua et al. 2006). Bisphenols have been used in plastics manufacturing for many years and are ubiquitous in the environment (Ballesteros-Gomez et al. 2007). In addition, in

Fig. 5 Quantitative analysis of bacterial and archaeal 16S rRNA and *amoA* genes in the water samples. Error bars represent standard deviation calculated from three independent experiments. (a) Bacterial and archaeal 16S rRNA (b) Bacterial and archaeal *amoA*

our metagenomic study, a large number of drug related genes were detected in the river and untreated wastewater, which indicate the pharmaceuticals of the river were mainly originated from untreated wastewater. Although most pharmaceuticals demonstrated no significantly acute toxicity to test organisms at environmentally relevant concentrations, such studies (which may be required for registration purposes) often do not adequately evaluate potential effects on non-target organisms, chronic toxicity, and possible additive effects (Huang et al. 2011). Moreover, the presence of pharmaceuticals in the river

may have consequences for the evolution of environmental microbiota through the spread of antibiotic resistance genes by horizontal gene transfer (Dutta and Pan 2002) and dissemination of drug-resistant bacteria (Li et al. 2010; Martinez 2009). So discharging untreated wastewater to the river may be serious harmful to human health.

In this study we did not identify any nitrification genes in the samples from metagenomic analysis due to the nonspecific amplification and/or the formation of a primer dimer (Ye et al. 2012), in particular for ammonia oxidation genes (*amo*). For

a supplement study, we detected bacterial and archaeal *amoA* genes using real-time PCR (Fig. 5). The abundances of bacterial *amoA* were significantly higher than archaeal *amoA* in the treated wastewaters, which indicated that AOB were mainly responsible for ammonia oxidation. As a result, AOB were dominant in the river upstream because of the significant present of treated wastewater entering the upstream segment. Middle- and downstream, however, the continuous inflow of untreated wastewater promotes an increase in the proportion and dominance by AOA.

Based on our data, it appears likely that both the chemical and microbial impacts of wastewaters on surface water may be reflected in the composition and function of microbial communities. Given the rapid development of high-throughput sequencing techniques (lower cost and higher sequencing read length), it could be developed as a formal monitoring tool for the qualitative assessments of water quality.

Acknowledgments This work was supported by the Key Program of National Natural Science Foundation of China (51008003 and 51138006) and the State Key Joint Laboratory of Environment Simulation and Pollution Control (12L03ESPC). The authors kindly thank Meng Qiao, Jingjing Ma for their assistance in chemical analysis.

References

- Bai YH, Shi Q, Wen DH, Li ZX, Jefferson WA, Feng CP, Tang XY (2012) Bacterial communities in the sediments of Dianchi Lake, a partitioned eutrophic waterbody in China. *PLoS One* 7
- Ballesteros-Gomez A, Ruiz FJ, Rubio S, Perez-Bendito D (2007) Determination of bisphenols A and F and their diglycidyl ethers in wastewater and river water by coextractive extraction and liquid chromatography-fluorimetry. *Anal Chim Acta* 603:51–59
- Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R (2010a) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26:266–267
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Tumbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010b) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336
- Chao A (1987) Estimating the population size for capture recapture data with unequal catchability. *Biometrics* 43:783–791
- Crump BC, Peterson BJ, Raymond PA, Amon RMW, Rinehart A, McClelland JW, Holmes RM (2009) Circumpolar synchrony in big river bacterioplankton. *Proc Natl Acad Sci USA* 106:21208–21212
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–5072
- Dutta C, Pan A (2002) Horizontal gene transfer and bacterial diversity. *J Biosci* 27:27–33
- Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461
- Fleming JT, Sanseverino J, Sayler GS (1993) Quantitative relationship between naphthalene catabolic gene-frequency and expression in predicting PAH degradation in soils at town gas manufacturing sites. *Environ Sci Technol* 27:1068–1074
- Frost JA (2001) Current epidemiological issues in human campylobacteriosis. *J Appl Microbiol* 90:85s–95s
- Good IJ (1953) The population frequencies of species and the estimation of population parameters. *Biometrika* 40:237–264
- Gregoracci GB, Nascimento JR, Cabral AS, Paranhos R, Valentin JL, Thompson CC, Thompson FL (2012) Structuring of bacterioplankton diversity in a large tropical bay. *PLoS One* 7
- Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E, Methe B, DeSantis TZ, Petrosino JF, Knight R, Birren BW, Consortium HM (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 21:494–504
- Heeb F, Singer H, Pernet-Coudrier B, Qi WX, Liu HJ, Longree P, Muller B, Berg M (2012) Organic micropollutants in rivers downstream of the megacity Beijing: sources and mass fluxes in a large-scale wastewater irrigation system. *Environ Sci Technol* 46:8680–8688
- Hua WY, Bennett ER, Maio XS, Metcalfe CD, Letcher RJ (2006) Seasonality effects on pharmaceuticals and s-triazine herbicides in wastewater effluent and surface water from the Canadian side of the upper Detroit River. *Environ Toxicol Chem* 25:2356–2365
- Huang QX, Yu YY, Tang CM, Zhang K, Cui JL, Peng XZ (2011) Occurrence and behavior of non-steroidal anti-inflammatory drugs and lipid regulators in wastewater and urban river water of the Pearl River Delta, South China. *J Environ Monit* 13:855–863
- Kasalický V, Jezbera J, Hahn MW, Šimek K (2013) The diversity of the *Limnohabitans* genus, an important group of freshwater bacterioplankton, by characterization of 35 isolated strains. *PLoS One* 8:e58209
- Kreus A, Darginaviciene J, Gilyte B, Grigutyte R, Jurkoniene S, Karionas R, Kucinskiene A, Pakalnis R, Sadauskas K, Vitkus R, Manusadzianis L (2013) Ecotoxicological effects evoked in hydrophytes by leachates of invasive *Acer negundo* and autochthonous *Alnus glutinosa* fallen off leaves during their microbial decomposition. *Environ Pollut* 173:75–84
- Li D, Yang M, Hu JY, Zhang Y, Chang H, Jin F (2008) Determination of penicillin G and its degradation products in a penicillin production wastewater treatment plant and the receiving river. *Water Res* 42:307–317
- Li D, Yu T, Zhang Y, Yang M, Li Z, Liu MM, Qi R (2010) Antibiotic Resistance Characteristics of environmental bacteria from an oxy-tetracycline production wastewater treatment plant and the receiving river. *Appl Environ Microbiol* 76:3444–3451
- Lin XJ, McKinley J, Resch CT, Kaluzny R, Lauber CL, Fredrickson J, Knight R, Konopka A (2012) Spatial and temporal dynamics of the microbial community in the Hanford unconfined aquifer. *ISME J* 6:1665–1676
- Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R (2011) UniFrac: an effective distance metric for microbial community comparison. *ISME J* 5:169–172
- Martinez JL (2009) The role of natural environments in the evolution of resistance traits in pathogenic bacteria. *Proc R Soc B* 276:2521–2530
- 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 Bioinforma* 9
- Mitch AA, Gasner KC, Mitch WA (2010) Fecal coliform accumulation within a river subject to seasonally-disinfected wastewater discharges. *Water Res* 44:4776–4782
- Moreno Y, Botella S, Alonso JL, Ferrus MA, Hernandez M, Hernandez J (2003) Specific detection of *Arcobacter* and *Campylobacter* strains in water and sewage by PCR and fluorescent in situ hybridization. *Appl Environ Microbiol* 69:1181–1186

- Ort C, Hollender J, Schaerer M, Siegrist H (2009) Model-based evaluation of reduction strategies for micropollutants from wastewater treatment plants in complex river networks. *Environ Sci Technol* 43:3214–3220
- Park JW, Crowley DE (2006) Dynamic changes in *nahAc* gene copy numbers during degradation of naphthalene in PAH-contaminated soils. *Appl Microbiol Biotechnol* 72:1322–1329
- Pernet-Coudrier B, Qi WX, Liu HJ, Muller B, Berg M (2012) Sources and pathways of nutrients in the semi-arid region of Beijing-Tianjin, China. *Environ Sci Technol* 46:5294–5301
- Qi W, Liu H, Pernet-Coudrier B, Qu J (2013) Polycyclic aromatic hydrocarbons in wastewater, WWTPs effluents and in the recipient waters of Beijing, China. *Environ Sci Pollut Res*. doi:10.1007/s11356-012-1435-6
- Riedinger-Whitmore MA, Whitmore TJ, Smoak JM, Brenner M, Moore A, Curtis J, Schelske CL (2005) Cyanobacterial proliferation is a recent response to eutrophication in many Florida lakes: A paleolimnological assessment. *Lake Reserv Manage* 21:423–435
- State Environmental Protection Administration of China (1989) Monitoring and analysis method of water and wastewater, 3rd edn. China Environmental Science Press, Beijing
- Suthar S, Sharma J, Chabukdhara M, Nema AK (2010) Water quality assessment of river Hindon at Ghaziabad, India: impact of industrial and urban wastewater. *Environ Monit Assess* 165:103–112
- Thomas T, Gilbert J, Meyer F (2012) Metagenomics - a guide from sampling to data analysis. *Microbial Informatics and Experimentation* 2:3:doi:10.1186/2042-5783-2-3
- Tsuzuki Y, Fujii M, Mochihara Y, Matsuda K, Yoneda M (2010) Natural purification effects in the river in consideration with domestic wastewater pollutant discharge reduction effects. *J Environ Sci-China* 22:892–897
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267
- Ye L, Zhang T, Wang TT, Fang ZW (2012) Microbial structures, functions, and metabolic pathways in wastewater treatment bioreactors revealed using high-throughput sequencing. *Environ Sci Technol* 46:13244–13252
- Yergeau E, Sanschagrin S, Beaumier D, Greer CW (2012) Metagenomic analysis of the bioremediation of diesel-contaminated Canadian high Arctic soils. *PLoS One* 7
- Yu K, Zhang T (2012) Metagenomic and Metatranscriptomic Analysis of microbial community structure and gene expression of activated sludge. *PLoS One* 7
- Zhang HS, Ziv-El M, Rittmann BE, Krajmalnik-Brown R (2010) Effect of dechlorination and sulfate reduction on the microbial community structure in denitrifying membrane-biofilm reactors. *Environ Sci Technol* 44:5159–5164
- Zhang YZ, Song XF, Kondoh A, Xia J, Tang CY (2011) Behavior, mass inventories and modeling evaluation of xenobiotic endocrine-disrupting chemicals along an urban receiving wastewater river in Henan Province, China. *Water Res* 45:292–302